



ISSN: 2301-5373
E-ISSN: 2654-5101
Volume 13 • Number 1 • August 2024

JELIKU

Jurnal Elektronik Ilmu Komputer Udayana

Informatics Study Program

Faculty of Mathematics and Natural Sciences

Udayana University

Table of Contents

Penerapan Algoritma Slope One dalam Collaborative Filtering Sebagai Sistem Rekomendasi Tempat Kost di Denpasar

Ni Wayan Anti Andari, I Wayan Supriana, I Putu Gede Hendra Suputra,
Luh Gede Astuti 981-992

Sistem Rekomendasi Produk Bodycare Dengan Metode Collaborative Filtering

Ida Ayu Taria Putri Mahadewi, Ida Bagus Gede Dwidasmara, Cokorda Rai Adi Pramatha, I
Gede Santi Astawa 993-1004

Penyisipan Digital Signature Ke Dalam Citra Digital Sebagai Keamanan Hak Cipta Dengan Metode DES dan BPCS

Michael Tanaya, I Gede Arta Wibawa, I Made Widiartha, Luh Gede Astuti..... 1005-1014

Penerapan Metode Content Based Filtering Dan K-Nearest Neighbor Dalam Sistem Rekomendasi Musik

I Made Teja Sarmandana, I Made Widiartha, Luh Arida Ayu Rahning Putri,
I Gede Santi Astawa 1016-1022

Pengolahan Big Data Dengan Sharding Database Dan Kappa Architecture Untuk Data Time-Series

Kompiang Gede Sukadharma, I Putu Gede Hendra Suputra, Ida Ayu Gde Suwiprabayanti
Putra, Luh Arida Ayu Rahning Putri 1023-1034

Aspect-Based Sentiment Analysis of Jkn Mobile Application Reviews Using the Random Forest Method and Information Gain as Feature Selection

Yauw James Fang Dwiputra Harta, I Gede Arta Wibawa, Anak Agung Istri Ngurah Eka
Karyawati, Komang Ari Mogi 1035-1044

Analisis Sentimen Ulasan Aplikasi Solusi Kota Cerdas Menggunakan Algoritma Naïve Bayes dan Support Vector Machine (SVM) dengan Seleksi Fitur Chi-Square

Ni Luh Komang Indira Pramesti, Made Agung Raharja, Ngurah Agus Sanjaya ER,
I Gede Arta Wibawa..... 1045-1058

Deteksi Rasa Buah Jeruk Siam Kintamani Menggunakan SVM dengan Optimasi Algoritma Genetika

Ni Wayan Yulia Damayanti, I Gede Arta Wibawa, I Gede Santi Astawa, Anak Agung Istri Ngurah Eka Karyawati 1059-1068

Implementasi Metode Fuzzy Logic untuk Mendeteksi Asap Dupa di Pasar Tradisional Bali Berbasis IoT

Implementasi Metode Fuzzy Logic untuk Mendeteksi Asap Dupa di Pasar Tradisional Bali Berbasis IoT 1069-1080

Performance impact of Hybrid Cryptography in securing RESTful API messages using ECIES

Albertus Ivan Suryawan, Agus Muliantara, Ngurah Agus Sanjaya ER, I Dewa Made Bayu Atmaja Darmawan 1081-1086

Pengembangan Sistem Pendukung Business Intelligence dan Forecasting untuk UMKM (Studi Kasus: Toko Perlengkapan Ternak)

Gusti Ngurah Deva Wirandana Putra, Cokorda Rai Adi Pramatha, Ida Bagus Gede Dwidasmara, Dr. Made Agung Raharja..... 1087-1098

Penerapan Multi Sensor Berbasis ESP8266 dalam Pengembangan Sistem Pemantauan Siswa di Sekolah

Kadek Yoga Vidya Pradnyaditha, Anak Agung Istri Ngurah Eka Karyawati 1099-1104

Implementasi Business Intelligence Menggunakan Metode OLAP untuk Pengambilan Keputusan Penjualan Properti

Lidya Elisabet Theogracia Silitonga, I Putu Gede Hendra Suputra, Ida Ayu Gde Suwiprabayanti Putra, I Ketut Gede Suhartana 1105-1112

Implementasi Metode Design Thinking dan Prinsip Gestalt Pada Rancang Bangun Dashboard Smart-Farm

Alvin Wiraprathama, I Made Widiartha, I Ketut Gede Suhartana, I Gusti Agung Gede Arya Kadyanan 1127-1138

Penerapan Metode Naive Bayes dalam Sistem Monitoring Akuaponik

Made Rahayu Setyaningrum, I Komang Ari Mogi, Agus Muliantara, Ida Ayu Gde Suwiprabayanti Putra..... 1139-1124

Analisis Serangan Cross Site Scripting (XSS) Pada Website OASE Menggunakan Metode OWASP

Muhammad Arrysatrya Yusuf Putranda, I Komang Ari Mogi, I Gusti Ngurah Anom Cahyadi Putra, I Made Widiartha 1139-1146

Pembangunan Aplikasi Pengenalan Kualitas Rasa Buah Jeruk Siam Kintamani Menggunakan GLCM dan K-NN Berbasis User Centered Design (UCD)

Putu Bagus Dio Pranata, I Gede Arta Wibawa, Made Agung Raharja, I Gusti Ngurah Anom Cahyadi Putra 1147-1158

Pengembangan Sistem Presensi Anti Spoofing dengan Metode Support Vector Machine

Marselinus Putu Harry Setyawan, I Dewa Made Bayu Atmaja Darmawan, I Gede Santi Astawa, I Wayan Santiyasa 1159-1170

Implementasi Metode Human Centered Design Dan Heuristic Evaluations Dalam Improvisasi User Experiences Dan User Interfaces (Studi Kasus : Aplikasi Pendidikan Dan Pelatihan Anggota Koperasi)

I Ketut Santa Wijaya, Cokorda Rai Adi Pramatha, Luh Gede Astuti, I Gede Surya Rahayuda 1171-1180

Faktor Penerimaan dan Penggunaan Siswaskeudes oleh Pengawas Keuangan Desa

Putu Widianari, Made Gede Wirakusuma, Ni Putu Sri Harta Mimba, Putu Agus Ardiana 1181-1190

Real Time Object Detection Menggunakan Mobilenet-SSD pada Sistem Keamanan Ruangan dengan Bot Telegram Sebagai Notifikasi User

Afit Miranto 1191-1196

Klasifikasi Kualitas Air Layak Minum menggunakan Algoritma Random Forest Classifier dan GridsearchCV

Gusti Agung Diah Sri Ari Ningsih, Cokorda Pramatha 1197-1206

This page is intentionally left blank.

**SUSUNAN DEWAN REDAKSI
JURNAL ELEKTRONIK ILMU KOMPUTER
UDAYANA (JELIKU)**

Penanggung Jawab :

Dra. Ni Luh Watiniasih M.Sc., Ph.D.

Redaktur :

Gst. Ayu Vida Mastrika Giri, S.Kom., M.Cs.

Penyunting/Editor :

Agus Muliantara, S.Kom., M.Kom

I Komang Ari Mogi, S.Kom., M.Kom

I Gusti Agung Gede Arya Kadyanan, S.Kom., M.Kom

Dr. Anak Agung Istri Ngurah Eka Karyawati, S.Si., M.Eng

Disain Grafis :

I Gede Yogananda Adi Baskara

I Gusti Agung Ayu Gita Pradnyaswari Mantara

Fotografer :

I Kadek Agus Candra

Widnyana I Komang

Dwiprayoga

Sekretariat :

Ni Ketut Alit Widiastuti, S.Kom.

Anak Agung Raka Darmawan, S.Kom.

I Putu Herryawan, S.Kom.

This page is intentionally left blank.

Penerapan Algoritma Slope One dalam Collaborative Filtering Sebagai Sistem Rekomendasi Tempat Kost di Denpasar

Ni Wayan Anti Andari^{a1}, I Wayan Supriana^{a2}, I Putu Gede Hendra Suputra^{a3}, Luh Gede Astuti^{a4},

^aProgram Studi Informatika, FMIPA, Universitas Udayana
Badung, Indonesia

¹antiandari01@gmail.com

²wayan.supriana@unud.ac.id

³hendra.suputra@unud.ac.id

⁴lg.astuti@unud.ac.id

Abstract

The city of Denpasar serves as the administrative, business, educational, and tourism center in Bali, attracting numerous visitors including students and workers from outside the region. This has led to high demand for temporary accommodation, commonly known as boarding houses or "kost". The appropriate solution to address this issue is by developing a boarding house recommendation system to assist prospective renters in finding a place that suits their needs and preferences. This system is built using two algorithms: collaborative filtering and Slope One, to generate accurate and relevant recommendations. In the evaluation phase, usability testing is conducted using the SUS method. The results of usability testing in terms of Acceptability, Grade Scale, and Adjective Rating obtained an average score of 82.03. Accuracy testing is done using Mean Absolute Error (MAE), which resulted in an accuracy with the highest value being 0.06. A higher accuracy indicates better recommendations. The ultimate goal is to help users, especially students, efficiently find suitable boarding houses in Denpasar through information technology.

Keywords: Boarding houses, Recommendation System, Collaborative Filtering, Usability Testing, Mean Absolute Error (MAE)

1. Pendahuluan

Denpasar sebagai pusat pemerintahan, bisnis, pendidikan, dan pariwisata di Bali menarik banyak orang untuk mengunjungi kota tersebut, termasuk mahasiswa atau pekerja dari luar wilayah. Situasi ini menciptakan kebutuhan tinggi akan akomodasi, dan salah satu pilihan yang diminati adalah hunian kos. Kos atau rumah kos adalah salah satu bentuk akomodasi atau tempat tinggal sementara yang terdiri dari beberapa kamar, di mana setiap kamar dilengkapi dengan fasilitas tertentu. Harganya biasanya ditetapkan oleh pemilik kos, sedangkan jangka waktu sewanya dapat ditentukan oleh penyewa kamar [1]. Saat memilih tempat kos, ada beberapa kriteria yang harus dipertimbangkan seperti harga, fasilitas, lokasi, jarak, aksesibilitas, kondisi lingkungan sekitar, kebersihan, kenyamanan, keamanan, dan lain-lain. Namun, preferensi individu dalam memilih tempat kos seringkali berbeda, dan ketersediaan tempat kos dengan harga dan fasilitas yang bervariasi dapat menjadi alasan mengapa seseorang mengalami kesulitan dalam menemukan tempat kos yang sesuai dengan kebutuhannya [2]. Untuk mengatasi hal ini dibutuhkan sebuah sistem yang dapat membantu untuk mencari dan mendapatkan informasi tentang tempat kos yang ada di Denpasar yaitu berupa sistem rekomendasi tempat kos khususnya di daerah Denpasar.

Sistem rekomendasi merupakan perangkat lunak yang menggunakan alat dan teknik tertentu untuk memberikan saran kepada pengguna mengenai item yang mungkin bermanfaat bagi mereka. Rekomendasi ini berkaitan dengan proses pengambilan keputusan, seperti pemilihan item yang ingin dibeli atau lagu yang ingin didengarkan [3]. Dengan adanya sistem rekomendasi ini diharapkan dapat membantu para pencari kos dalam memilih kos dengan lebih efisien dan efektif. Dengan menggunakan teknologi informasi dan sistem informasi, solusi yang lebih efektif dan efisien dapat dikembangkan

dengan mengimplementasikan sistem rekomendasi yang sesuai dengan preferensi dan kebutuhan pengguna. Sistem ini memanfaatkan metode collaborative filtering untuk menghasilkan rekomendasi yang relevan dan akurat. Metode ini dipadukan dengan algoritma slope one, yang menghitung perbedaan rata-rata rating antara item dalam dataset untuk memberikan rekomendasi. Selain itu, sistem ini juga memanfaatkan data dari pengguna lain yang memiliki preferensi serupa untuk meningkatkan ketepatan rekomendasi.

Penelitian sebelumnya yang dilakukan oleh [4] pada kasus yang serupa telah mengembangkan sebuah sistem rekomendasi tempat kos. Sistem ini memberikan rekomendasi berdasarkan penilaian preferensi yang dipilih oleh pengguna untuk setiap kriteria yang ada. Rekomendasi yang dihasilkan disusun dalam urutan dari alternatif terbaik hingga terburuk sesuai dengan preferensi yang ditentukan. Terkait dengan preferensi dan penilaian, penelitian yang dilakukan oleh [5] menggunakan algoritma collaborative filtering dalam merekomendasikan produk yang memiliki nilai kegunaan tertinggi berdasarkan kesamaan peringkat yang diberikan oleh pengguna terhadap produk yang berbeda.

Sebagai upaya untuk mengatasi kesulitan calon penyewa dalam memilih tempat kost. Penelitian ini bertujuan untuk menyediakan solusi yang efektif dan efisien bagi para pencari kos dalam memilih tempat tinggal sementara. Melalui pengembangan sistem ini, diharapkan dapat memberikan rekomendasi yang akurat dan relevan berdasarkan preferensi, kebutuhan, dan kriteria yang ditetapkan oleh pengguna. Kemudian untuk pengujian menggunakan metode *usability testing* yang sesuai dengan standar ISO, dan menggunakan *mean absolute error* (MAE) untuk mengukur akurasi dari algoritma yang digunakan. Dengan demikian, diharapkan penelitian ini akan memberikan kontribusi yang signifikan dalam mempermudah calon penyewa dalam memilih tempat kost yang sesuai dengan preferensinya melalui pendekatan yang terpadu dan inovatif.

2. Metode Penelitian

Penelitian ini akan memanfaatkan pendekatan metode waterfall sebagai kerangka penelitiannya. Metode waterfall merupakan salah satu pendekatan pengembangan perangkat lunak yang terfokus pada serangkaian tahapan. Pendekatan ini mengikuti urutan langkah yang terstruktur, mencakup analisis kebutuhan, perancangan, implementasi, pengujian, dan pemeliharaan [6]

2.1. Pengumpulan Data

Pengumpulan data tempat kos dengan menggunakan sumber informasi yang tersedia di aplikasi Mamikos dan melalui pencarian data secara daring di internet. Proses pengumpulan data yang komprehensif ini dimaksudkan untuk menyajikan informasi yang tepat dan sesuai bagi pengguna yang mencari tempat kos sesuai dengan preferensi dan kebutuhan mereka. Dalam penelitian ini, sebanyak 150 data tempat kos di wilayah Denpasar telah berhasil dikumpulkan oleh penulis.

2.2. Desain Sistem

Dalam fase perancangan sistem, langkah awal yang dilakukan adalah merancang kebutuhan yang diperlukan oleh sistem sebelum memasuki tahap implementasi dalam bentuk kode program. Proses perancangan ini bertujuan untuk memberikan pemahaman yang komprehensif tentang bagaimana sistem yang direncanakan akan berfungsi dan diintegrasikan secara efektif saat diimplementasikan.

1. Analisis Kebutuhan Sistem

Langkah ini merupakan bagian dari proses pembangunan sistem rekomendasi tempat kos di Denpasar. Analisis kebutuhan dilakukan untuk memahami secara menyeluruh kebutuhan fungsional dan non-fungsional yang harus terpenuhi dalam pengembangan sistem.

a. Kebutuhan Fungsional

Pada tahap ini, yang merupakan bagian dari pengembangan sistem rekomendasi tempat kos di Denpasar, analisis kebutuhan fungsional harus dilakukan untuk menentukan fitur-fitur yang diperlukan oleh sistem.

Tabel 1 Kebutuhan Fungsional

No.	Deskripsi Kebutuhan Sistem	Target Aktor
KF1	Sistem menyediakan proses <i>login</i> dan <i>logout</i> bagi admin sesuai dengan <i>email</i> dan <i>password</i> yang ditentukan.	Admin
KF2	Sistem menyediakan fasilitas untuk melakukan penjelajahan dan pencarian tempat kost berdasarkan kriteria yang telah diinputkan oleh user.	Pengguna

KF3	Sistem menyediakan fasilitas untuk menampilkan hasil rekomendasi dari produk berdasarkan dari rating dan kesamaan item atau kriteria yang diinginkan oleh user.	Pengguna
KF4	Sistem menyediakan fasilitas untuk admin dapat menambahkan, mengedit, dan menghapus data produk pada website melalui halaman admin.	Admin
KF5	Sistem menyediakan fasilitas untuk dapat melihat informasi lebih lengkap dari sebuah tempat kost melalui detail page.	Pengguna

b. Kebutuhan Non-Fungsional

Pada tahap ini, dilibatkan analisis kebutuhan yang berkaitan dengan aspek non-fungsional sistem, termasuk komponen-komponen pendukung yang digunakan untuk mendukung penelitian ini seperti kebutuhan perangkat keras dan perangkat lunak yang diperlukan dalam pengembangan sistem.

Tabel 2 Spesifikasi Perangkat Keras Implementasi Sistem Rekomendasi

No.	Perangkat	Spesifikasi Laptop
1	Processor	MacBook Pro M1
2	Monitor	13,3 inch
3	RAM	8 GB
4	HDD	512 GB

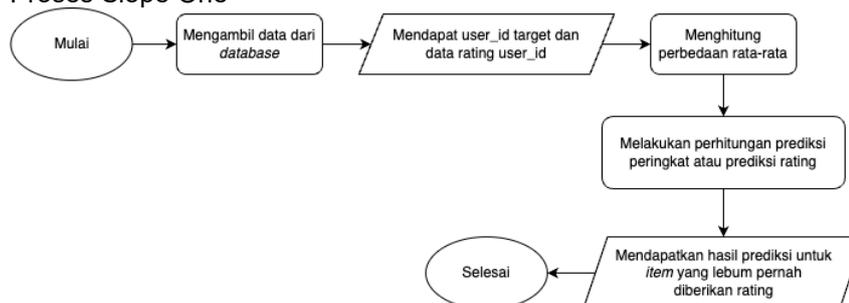
Tabel 3 Daftar Perangkat Keras Implementasi Sistem Rekomendasi

No.	Perangkat	Penjelasan
1	MacOS Sonoma	Sistem operasi
2	XAMPP	Perangkat yang digunakan sebagai server (localhost) local
3	Visual Studio Code	Perangkat yang digunakan sebagai text editor dalam pembuatan web (Javascript, HTML, CSS)
4	HDD	512 GB

2. Flowchart

Flowchart adalah representasi grafis dari langkah-langkah dan urutan prosedur suatu program [7]. Dalam pengembangan sistem rekomendasi, terdapat dua jalur utama yang harus dipertimbangkan. Jalur pertama mencakup proses prediksi rating, sementara jalur kedua berkaitan dengan penyusunan rekomendasi. Oleh karena itu, flowchart menjadi alat yang berguna dalam menggambarkan visualisasi dari proses-proses yang terlibat dalam pengembangan sistem rekomendasi

a. Proses Slope One

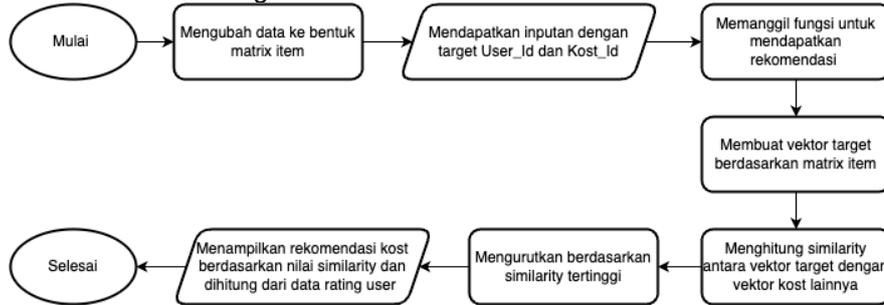


Gambar 1 Flowchart Slope One

Gambar yang disajikan adalah flowchart dari algoritma slope one. Dalam proses prediksi rating, langkah pertama adalah mengekstraksi data dari basis data. Data ini akan mencakup informasi mengenai user_id target serta data rating dari user_id. Selanjutnya, dilakukan perhitungan untuk menemukan perbedaan rata-rata.

Penerapan Algoritma Slope One dalam Collaborative Filtering Sebagai Sistem Rekomendasi Tempat Kost di Denpasar

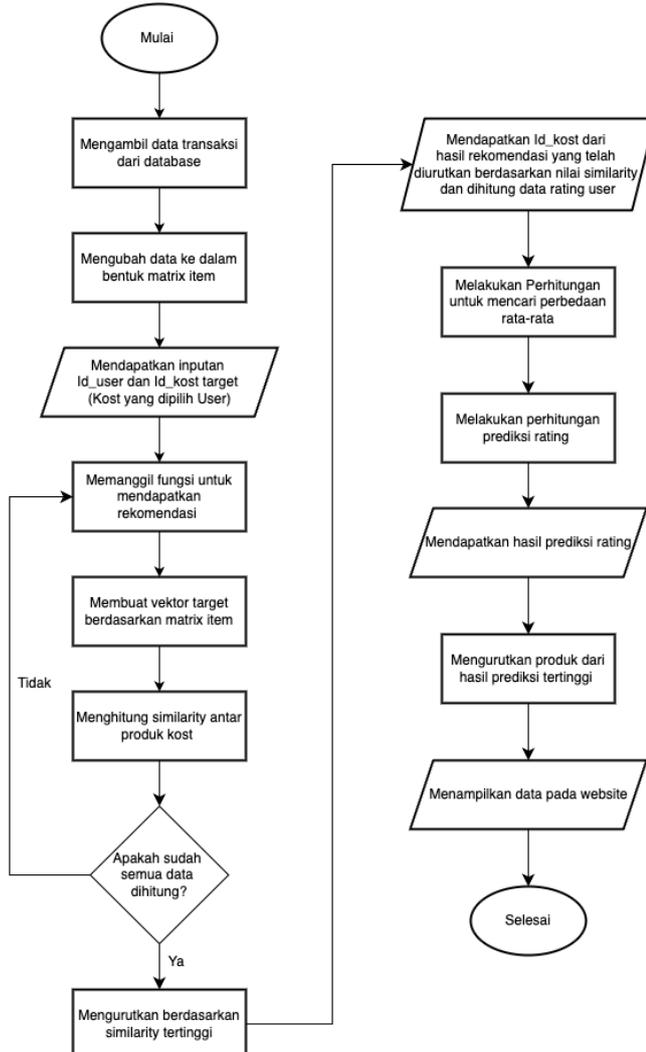
b. Collaborative Filtering



Gambar 2 Flowchart Collaborative Filtering

Metode ini menggunakan collaborative filtering, dimulai dengan mengekstraksi data dari basis data. Langkah selanjutnya adalah inialisasi matriks item, diikuti dengan iterasi melalui data dan pengisian matriks item. Setelah itu, nilai matriks dikembalikan, diikuti dengan pengambilan inputan dan pemanggilan fungsi rekomendasi. Dalam fungsi untuk mendapatkan rekomendasi, terdapat pemanggilan fungsi yang menghitung kesamaan kosinus antara vektor data.

c. Flowchart Sistem



Gambar 3 Flowchart Sistem

Diagram di atas adalah representasi visual dari sistem rekomendasi tempat kos menggunakan algoritma slope one dan collaborative filtering. Proses pertama yang dilakukan yaitu mengolah data menggunakan algoritma collaborative filtering. Selanjutnya hasil dari collaborative filtering akan diproses lagi menggunakan algoritma slope one dengan melakukan perhitungan untuk mencari perbedaan rata-rata antara rating yang telah dihitung. Prediksi rating kemudian dihitung, dan hasil prediksi tersebut digunakan untuk mengurutkan rumah kos berdasarkan

rating tertinggi. Data hasil rekomendasi kemudian ditampilkan pada situs web, dan proses selesai. Dengan demikian, sistem dapat memberikan rekomendasi rumah kos yang sesuai dengan preferensi pengguna dengan efisien.

2.3. Desain Evaluasi Sistem

1. Usability Testing

Pengujian tahap ini dilakukan setelah selesainya pengembangan sistem dan sebelum diperkenalkan kepada pengguna. Tingkat usability dapat dinilai dengan mengukur faktor-faktor seperti kemudahan penggunaan, tingkat kesalahan, kecepatan penggunaan, dan tingkat kepuasan pengguna. Tujuan utama dari pengujian ini adalah memastikan bahwa sistem dapat digunakan dengan mudah dan efektif oleh pengguna [8]. Dalam sistem ini, digunakan *System Usability Scale* (SUS) untuk mengevaluasi dan menilai kegunaan dari website yang telah dibuat. Pada pengujian usability testing terdapat lima ukuran yang dijadikan pedoman untuk mengukur tingkat usability berdasarkan dari standar ISO yaitu *learnability*, *efficiency*, *memorability*, *errors*, dan *satisfaction*.

2. Mean Absolute Error (MAE)

MAE (*Mean Absolute Error*) adalah metode yang digunakan untuk menilai seberapa akurat atau seberapa besar kesalahan dalam hasil prediksi rating suatu sistem terhadap rating aktual yang diberikan oleh pengguna terhadap suatu item. MAE dihitung dengan mengambil selisih absolut antara N pasang rating asli dan prediksi, kemudian diambil rata-ratanya. Semakin mendekati nilai 0, semakin baik hasil prediksinya [9].

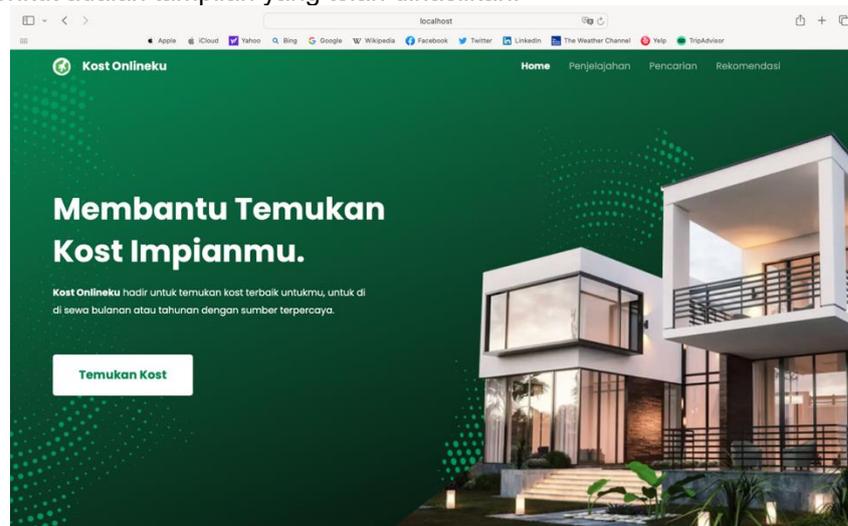
3. Hasil dan Pembahasan

3.1 Proses Pengumpulan Data

Data yang dimanfaatkan dalam penelitian ini adalah informasi mengenai tempat kos yang tersedia di Denpasar. Pengumpulan data untuk tempat kos dilakukan secara manual atau dengan memanfaatkan data primer dari situs web kost, seperti Mamikost, serta melalui internet. Total 150 data tempat kos yang tersebar di empat kecamatan di Denpasar, yaitu Denpasar Utara, Denpasar Timur, Denpasar Selatan, dan Denpasar Barat, digunakan dalam penelitian ini.

3.2 Hasil Desain Sistem

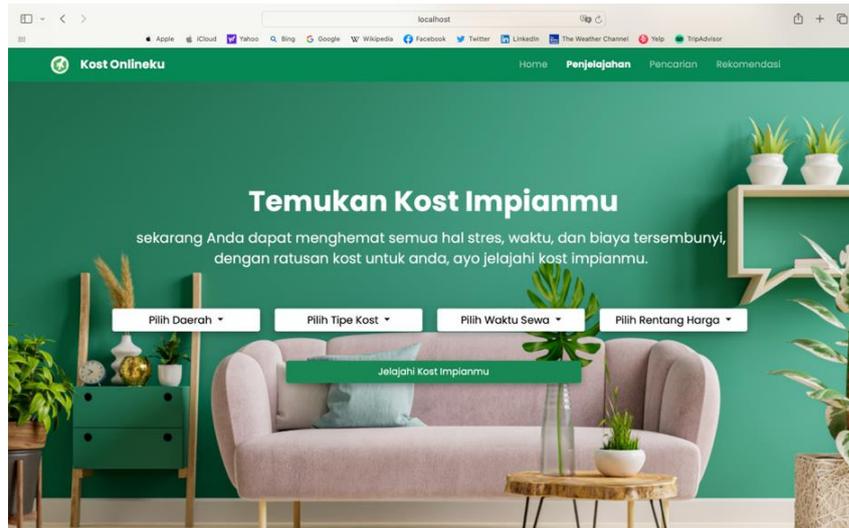
Dalam implementasi desain sistem ini, rancangan desain sistem dibuat menggunakan HTML, CSS, dan Bootstrap. Berikut adalah tampilan yang telah dihasilkan.



Gambar 4 tampilan Dashboard Sistem

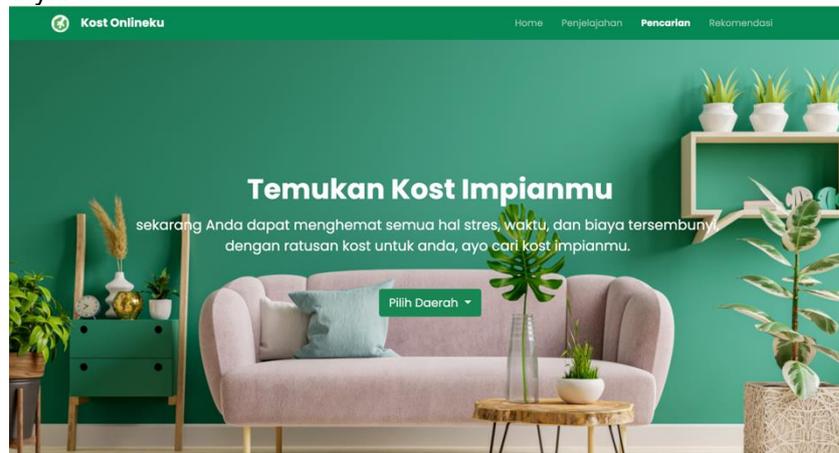
Gambar di atas merupakan tampilan dashboard pada sistem rekomendasi tempat kos. Pada tampilan dashboard ini sistem dapat langsung melihat fitur-fitur yang terdapat pada sistem rekomendasi ini yang terletak di pojok kanan atas yaitu fitur penjelajahan dan fitur pencarian. Pada tampilan dashboard ini guest user juga bisa menekan button “temukan kost” yang akan mengarahkan pada tampilan penjelajahan.

Penerapan Algoritma Slope One dalam Collaborative Filtering Sebagai Sistem Rekomendasi Tempat Kost di Denpasar



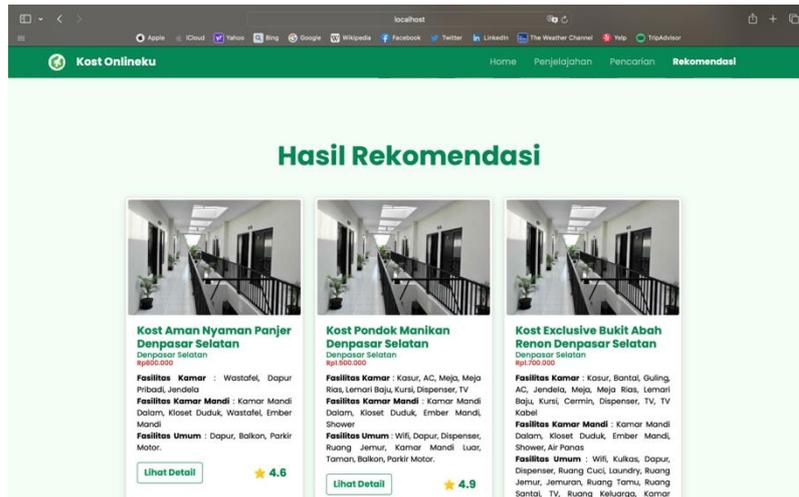
Gambar 5 Tampilan Fitur Penjelajahan Pada Sistem

Pada halaman penjelajahan ini *guest user* dapat melakukan eksplorasi dengan mengisi 4 kriteria yang telah disediakan yaitu memilih daerah yang diinginkan, kemudian memilih tipe kost, pilihan waktu sewa, dan rentang harga yang diinginkan. Setelah itu data akan ditampilkan sesuai dengan kriteria yang telah diinputkan sebelumnya.



Gambar 6 Tampilan Fitur Pencarian Pada Sistem

Pada halaman ini *guest user* dapat melakukan pencarian berdasarkan dari daerah yang diinginkan, terdapat 4 pilihan daerah yaitu Denpasar utara, Denpasar timur, Denpasar selatan, dan Denpasar barat. *Guest user* dapat memilih daerah yang diinginkan kemudian akan ditampilkan seluruh tempat kost dari daerah tersebut.



Gambar 7 Tampilan Hasil Rekomendasi Pada Sistem

Gambar di atas merupakan tampilan dari hasil rekomendasi. Untuk mendapatkan rekomendasi tempat kost dapat dilakukan dengan melakukan pencarian tempat kost pada fitur pencarian. Kemudian, pilih detail tempat kost sehingga guest user dapat melihat rekomendasi lainnya dengan menekan button “Lihat Rekomendasi Lainnya” selanjutnya sistem akan menampilkan rekomendasi dari tempat kost tersebut.

3.3 Implementasi Metode

Pada tahap ini merupakan implementasi dari metode slope one dan algoritma collaborative filtering dengan menggunakan pengkodean program. Dengan mengimplementasikan Node JS dan framework yang digunakan yaitu framework Express JS.

1. Collaborative Filtering

Code di bawah ini berfungsi untuk menghitung kesamaan kosinus antara dua vektor. Fungsi `cosineSimilarity(vectorA, vectorB)` mengukur seberapa serupa dua vektor dalam ruang vektor dengan menggunakan metrik kemiripan kosinus. Nilai kemiripan kosinus yang lebih besar menunjukkan kedua vektor semakin serupa. Proses dimulai dengan menghitung dot produk antara `vectorA` dan `vectorB` menggunakan metode `reduce()`, yang mengalikan setiap elemen pada vektor dan menyimpan hasilnya dalam variabel `dotProduct`. Kemudian, magnitudo (panjang) dari masing-masing vektor dihitung dengan menggunakan Euclidean (akar kuadrat dari jumlah kuadrat semua elemen). Jika salah satu hasil magnitudo adalah 0, ini menandakan bahwa vektor tersebut adalah vektor nol dan tidak memiliki arah. Dalam kasus ini, fungsi mengembalikan 0 sebagai hasil kemiripan kosinus karena tidak mungkin ada kemiripan antara vektor nol dan vektor lainnya. Namun, jika kedua magnitudo tidak nol, kemiripan kosinus dihitung dengan membagi `dotProduct` antara kedua vektor dengan hasil perkalian magnitudenya. Hasil kemiripan kosinus dikembalikan sebagai output dari fungsi.

Tabel 4 Source Code Menghitung Nilai Similarity

```
function cosineSimilarity(vectorA, vectorB) {
  const dotProduct = vectorA.reduce(
    (acc, val, i) => acc + val * vectorB[i],
    0
  );
  const magnitudeA = Math.sqrt(
    vectorA.reduce((acc, val) => acc + val * val, 0)
  );
  const magnitudeB = Math.sqrt(
    vectorB.reduce((acc, val) => acc + val * val, 0)
  );

  if (magnitudeA === 0 || magnitudeB === 0) {
    return 0;
  }

  return dotProduct / (magnitudeA * magnitudeB);
}
```

2. Slope One

Code di bawah ini bertujuan untuk melakukan prediksi rating menggunakan metode slope one. Fungsi `predictRating(userRatings, averages)` digunakan untuk prediksi rating yang mungkin diberikan oleh seorang pengguna untuk properti yang belum dinilainya sebelumnya. Penggunaan prediksi dilakukan dengan mempertimbangkan rating yang diberikan oleh pengguna untuk properti lain serta rata-rata rating yang sudah dihitung sebelumnya. Pertama-tama, fungsi menerima dua parameter, yaitu `userRatings` dan `averages`. Selanjutnya, dua objek kosong, `prediction` untuk menyimpan prediksi rating untuk properti yang belum pernah dinilai oleh pengguna, dan `count` untuk menghitung jumlah properti yang telah dinilai oleh pengguna, dibuat. Proses iterasi dilakukan untuk memeriksa semua properti yang telah dinilai oleh pengguna. Setelah itu, prediksi rating yang telah dihitung dinormalisasi dengan membaginya dengan jumlah properti yang telah dinilai oleh pengguna, untuk mendapatkan prediksi akhir untuk setiap properti yang belum dinilai oleh pengguna. Hasil prediksi akhir disimpan dalam objek result.

Tabel 5 Source Code Prediksi Rating

```
function predictRating(userRatings, averages) {
  let prediction = {};
  let count = {};

  for (let kostIdA in averages) {
    prediction[kostIdA] = 0;
    count[kostIdA] = 0;

    for (let kostIdB in userRatings) {
      if (averages[kostIdA][kostIdB]) {
        prediction[kostIdA] += (userRatings[kostIdB] +
          averages[kostIdA][kostIdB]) *
          userRatings[kostIdB];
        count[kostIdA] += userRatings[kostIdB];
      }
    }
  }

  let result = {};

  for (let kostIdA in prediction) {
    result[kostIdA] = count[kostIdA] === 0 ? 0 :
      prediction[kostIdA] / count[kostIdA];
  }

  return result;
}
```

3.4 Implementasi Evaluasi Sistem

Pada tahap ini merupakan tahapan pengujian dari sistem rekomendasi yang dibuat. Pengujian yang dilakukan pada sistem ini berupa pengujian usability testing dan MAE.

1. Usability Testing

Pengujian ini dilakukan dengan mendistribusikan kuesioner kepada responden melalui Google Form untuk mengumpulkan data. Total 32 responden telah mengisi kuesioner dalam pengujian ini.

a. Pengujian Validitas

Langkah ini penting untuk mengevaluasi dan memeriksa seberapa kuat kesimpulan yang dihasilkan dari pengujian agar mendekati kebenaran. Proses validasi ini melibatkan perhitungan nilai r dari setiap item yang dievaluasi. dari uji validitas ini diketahui bahwa r tabel yang didapatkan dari total 32 responden adalah 0.349. di mana jika nilai r hitung yang didapatkan lebih besar dari r tabel pada masing-masing item maka akan menunjukkan uji validitas pada butir pertanyaan dinyatakan valid.

b. Pengujian Reliabilitas

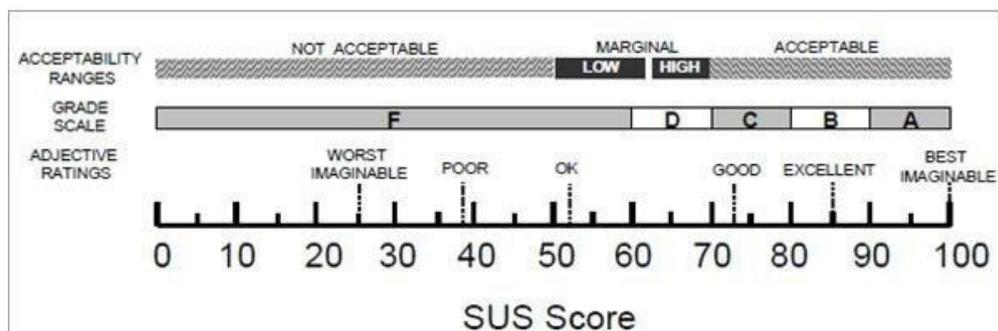
Tabel 6 Hasil Uji Reliabilitas

Cronbach's Alpha	r Tabel	Keterangan
0.801	0.70	Reliabel

Hasil dari pengujian tersebut diperoleh nilai Cronbach's sebesar 0.801 dimana hal ini menunjukkan bahwa Kuesioner sudah reliabel karena nilai dari Cronbach's Alpha (r hitung) lebih besar dari r tabel yaitu $0.801 > 0.70$.

c. Hasil Pengujian Usability Testing

Pada fase ini, data yang dimanfaatkan berasal dari pengisian formulir Google oleh para responden. Pertanyaan yang diajukan dirancang menggunakan skala Likert. Hasil penilaian dari 32 responden memperoleh nilai rata-rata sebesar 82.03. setelah mendapatkan hasil akhir dari responden langkah selanjutnya yaitu menentukan grade hasil penilaian dengan menggunakan metode Acceptability, Grade Scale, dan Adjective Rating.



Gambar 8 Hasil penilaian dengan memanfaatkan metode Acceptability, Grade Scale, dan Adjective Rating

Berdasarkan pada gambar 8 diatas maka hasil penilaian yang dibandingkan akan memiliki hasil sebagai berikut:

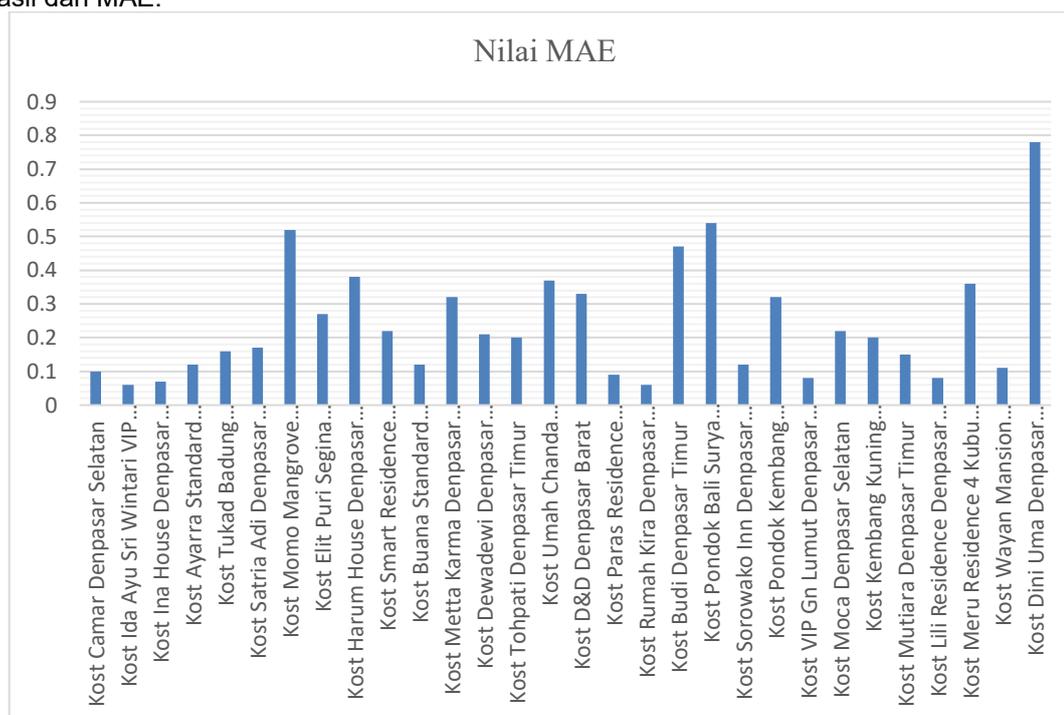
1. Tingkat Acceptability Range pengguna terhadap sistem rekomendasi tempat kost dengan rata-rata 82,03 termasuk ke dalam kategori Acceptable.
2. Tingkat Grade Scale pengguna terhadap sistem rekomendasi tempat kost dengan rata-rata 82,03 termasuk ke dalam kategori B.
3. Tingkat Adjective Rating pengguna terhadap sistem rekomendasi tempat kost dengan rata-rata 82,03 termasuk ke dalam kategori GOOD.

Maka dari itu dapat disimpulkan bahwa sistem rekomendasi tempat kost memiliki tingkat usability yang baik berdasarkan dari skor System Usability Scale (SUS)

2. MAE (Mean Absolute Error)

Pengujian menggunakan Mean Absolute Error (MAE) bertujuan untuk mengevaluasi keakuratan dari Sistem Rekomendasi Tempat Kost yang menggunakan metode slope one dan algoritma collaborative filtering. Dalam proses perhitungan ini, digunakan data nilai dari 30 tempat kost. Rentang nilai MAE berkisar dari 0 hingga 1 secara berkelanjutan. Rekomendasi

dianggap tidak akurat apabila nilai MAE sama dengan atau lebih besar dari 1. Berikut adalah hasil dari MAE:



Gambar 9 Grafik Hasil Nilai MAE

Berdasarkan grafik diatas, tempat kost yang memiliki nilai akurasi tertinggi yaitu dengan nilai 0,06 dan akurasi terendah yaitu 0,78 dimana berdasarkan dari rentang semakin mendekati nilai 0 maka akurasinya semakin tepat atau akurat. Hasil ini menunjukkan bahwa semakin tinggi nilai akurasi, semakin baik juga kualitas rekomendasi yang diberikan. Tingkat akurasi yang tinggi mencerminkan efektivitas sistem dalam memberikan rekomendasi yang bermanfaat dan relevan kepada pengguna, yang pada akhirnya dapat meningkatkan kepuasan dan kepercayaan pengguna terhadap sistem tersebut.

4. Kesimpulan

Dari hasil penelitian yang menyeluruh terhadap sistem rekomendasi tempat kos di Denpasar menggunakan metode slope one dan algoritma collaborative filtering, dapat disimpulkan bahwa:

1. Hasil Pengujian Usability pada sistem rekomendasi tempat kost menggunakan metode SUS (System Usability Scale) memperoleh skor rata-rata sebesar 82,03. Pada tingkat Acceptability Range pengguna terhadap sistem rekomendasi tempat kost termasuk ke dalam kategori Acceptable, pada tingkat Grande Scale pengguna terhadap sistem rekomendasi tempat kost termasuk ke dalam kategori B, dan pada tingkat Adjective Rating pengguna terhadap sistem rekomendasi tempat kost termasuk ke dalam kategori GOOD. Dengan demikian dapat disimpulkan sistem rekomendasi tempat kost memiliki kinerja yang baik.
2. Pengujian akurasi sistem dihitung menggunakan Mean Absolute Error (MAE) menghasilkan akurasi tertinggi mencapai 0,06. Hal ini menunjukkan bahwa sistem rekomendasi menggunakan algoritma slope one dan algoritma collaborative filtering memiliki tingkat keakuratan yang tinggi dalam memberikan rekomendasi kepada pengguna dimana semakin tinggi nilai akurasi yang dihasilkan, semakin baik pula kualitas rekomendasi yang diberikan oleh sistem.

Referensi

- [1] D. Ratnasari, D. B. Qur'ani, and A. Apriani, "Sistem Informasi Pencarian Tempat Kos Berbasis Android," *Jurnal INFORM*, vol. 3, no. 1, Jan. 2018, doi: 10.25139/ojsinf.v3i1.657.
- [2] E. M. Sipayung, C. F. Fiarni, and S. Sutopo, "Sistem Rekomendasi Tempat Kost di Sekitar Kampus ITHB Menggunakan Metode Analytical Hierarchy Process (AHP)," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 7, no. 2, pp. 52–60, Aug. 2021, doi: 10.25077/teknosi.v7i2.2021.52-60.

- [3] K. T. Mukti and I. Mardhiyah, "SISTEM REKOMENDASI PEMBELIAN LISENSI FILM MENGGUNAKAN PENDEKATAN HYBRID FILTERING (STUDI KASUS: FILM ANIMASI JEPANG)," *JURSIKTEKNI (Jurnal Sistem Informasi dan Teknologi Informasi)*, vol. 4, pp. 126–139, 2022.
- [4] E. M. Sipayung, C. F. Fiarni, and S. Sutopo, "Sistem Rekomendasi Tempat Kost di Sekitar Kampus ITHB Menggunakan Metode Analytical Hierarchy Process (AHP)," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 7, no. 2, pp. 52–60, Aug. 2021, doi: 10.25077/teknosi.v7i2.2021.52-60.
- [5] A. Arifin, "Penerapan Sistem Algoritma Collaborative Filtering Untuk Rekomendasi Pemilihan Indeks Berdasarkan Rating," *Teknologipintar.org*, vol. Volume 2 (6), 2022.
- [6] A. A. Wahid, "Analisis Metode Waterfall Untuk Pengembangan Sistem Informasi," *Jurnal Ilmu-ilmu Informatika dan Manajemen STMIK*, 2020.
- [7] Z. Tuasamu *et al.*, "Analisis Sistem Informasi Akuntansi Siklus Pendapatan Menggunakan DFD Dan Flowchart Pada Bisnis Porobico," 2023.
- [8] I. Isnainrajab, S. Hadi Wijoyo, and A. R. Perdanakusuma, "Evaluasi Usability Pada Aplikasi PermataMobile X Dengan Menggunakan Metode Usability Testing Dan System Usability Scale(SUS)," 2020. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [9] Y. Imelda Lubis, D. Josua Napitupulu, and A. Satia Dharma, "Implementasi Metode Hybrid Filtering (Collaborative dan Content-based) untuk Sistem Rekomendasi Pariwisata Implementation of Hybrid Filtering (Collaborative and Content-based) Methods for the Tourism Recommendation System," 2020.

This page is intentionally left blank.

Sistem Rekomendasi Produk *Bodycare* Dengan Metode *Collaborative Filtering*

Ida Ayu Taria Putri Mahadewi^{a1}, Ida Bagus Gede Dwidasmara^{a2}, Cokorda Rai Adi Pramartaha^{a3},
I Gede Santi Astawa^{a4}

^aProgram Studi Informatika, FMIPA, Universitas Udayana
Badung, Indonesia

¹gek.taria@gmail.com

²dwidasmara@unud.ac.id

³cokorda@unud.ac.id

⁴santi.astawa@unud.ac.id

Abstract

Currently, especially among teenage girls, there is a strong trend towards skincare. This includes care for the face, hair, and body. Body care aims not only for appearance but also for skin health and moisture. There are many body care products available on the market, but choosing the right one according to individual needs and skin issues can be challenging due to varying product compositions. Technology is often used to search for information on body care products. However, the information found is often irrelevant or inaccurate. This research aims to assist by developing a body care product recommendation system using Collaborative Filtering method. This recommendation system is built utilizing semantic web technology, and the data will be mapped into an ontology format as the knowledge base. The ontology model will be constructed using the Methontology method, system development using the Prototyping method, and product recommendations offered using the Collaborative Filtering method. The testing phase consists of two stages: usability testing and testing using Mean Absolute Error (MAE). Usability testing using the System Usability Scale (SUS) method from the Acceptability, Grade Scale, and Adjective Rating perspectives obtained an average score of 82.344. The accuracy testing using Mean Absolute Error (MAE) resulted in a value of 0.3556, which implies that the accuracy level obtained is accurate, and the body care product recommendations are good because the closer the value is to 0, the more accurate and better the accuracy towards the given body care product recommendations.

Keywords: Recommendation, Bodycare, Collaborative Filtering, Semantic, Ontology

1. Pendahuluan

Salah satu aspek penting dari gaya hidup manusia saat ini, terutama bagi perempuan, adalah penampilan. Namun, masih banyak perempuan memiliki kepercayaan diri yang rendah untuk tampil di depan publik karena penampilan mereka. Terdapat beberapa hal yang dapat diterapkan guna menaikkan tingkat kepercayaan diri untuk berani tampil di depan umum, diantaranya adalah dengan merawat kulit menggunakan produk perawatan tubuh. Dengan merawat kulit secara teratur, dapat menjaga kesehatan kulit dan mencegah kulit menjadi kering dan kusam. Pada masa saat ini tidak hanya perempuan, namun juga laki-laki mulai memperhatikan penampilan mereka dengan merawat kulit menggunakan produk perawatan tubuh untuk menjaga kebersihan dan kesehatan kulit.

Banyak sekali produk perawatan tubuh yang tersedia di pasaran dengan berbagai jenisnya. Berdasarkan informasi data Badan Pusat Statistik (BPS) pada kuartal I tahun 2020, pertumbuhan perawatan tubuh meningkat hingga 5,59%. Tentunya, dengan banyaknya jenis produk *bodycare* di pasaran menyebabkan banyak pertimbangan oleh konsumen untuk memilih produk *bodycare* yang tepat dan sesuai dengan kebutuhan pengguna. Produk perawatan tubuh yang wajib dimiliki baik oleh perempuan maupun laki-laki antara lain adalah *body wash*, *body scrub*, dan *body lotion*.

Untuk membantu konsumen memilih produk *bodycare* yaitu dengan membangun suatu sistem yang mampu membantu merekomendasikan suatu produk *bodycare* kepada konsumen sesuai dengan kebutuhan pengguna. Ontologi merupakan salah satu pendekatan yang dapat diterapkan, yang diimplementasikan pada *semantic web* yang berkolaborasi dengan sistem rekomendasi. Dengan menggunakan pendekatan ontologi, data pada situs web dapat memiliki koneksi yang jelas, sehingga

dapat memberikan deskripsi data yang lebih baik[1]. *Collaborative filtering* merupakan suatu metode yang kerap diterapkan dalam pembangunan sistem rekomendasi mengenai ontologi[2]. *Collaborative filtering* adalah proses yang digunakan untuk penyaringan item dengan memanfaatkan opini orang lain untuk memberikan rekomendasi. Dalam konteks ini, hubungan antara *collaborative filtering* dan ontologi terletak pada *collaborative filtering* yang berfungsi sebagai pengambil keputusan dengan memprioritaskan perhitungan untuk menentukan rekomendasi produk, sementara ontologi berfungsi sebagai representasi pengetahuan domain produk perawatan tubuh, mengidentifikasi dan mengekstrak informasi dari data, dan membuat pemetaan antara konsep-konsep yang terkait.

Penelitian sebelumnya yang dilakukan oleh Erlangga & Sutrisno (2020) menggunakan metode yang sama, sebuah sistem dapat membantu dalam pemilihan *beauty shop* terbaik dikembangkan dengan menggunakan metode *collaborative filtering* untuk menentukan keputusan dengan mempertimbangkan kriteria yang diberikan oleh pengguna. Studi ini digunakan sebagai acuan untuk menentukan kriteria dan menghitung bobotnya menggunakan *collaborative filtering*. Pratiwi (2022) melakukan penelitian yang terkait dengan proses pencarian dan penjelajahan, dimana ia membangun suatu *framework* untuk sistem penyaringan hibrida dengan menggabungkan aspek-aspek ontologi yang disebut sebagai rekomendasi tempat magang pribadi berbasis ontologi (RMBO). Dalam penelitian tersebut, metode yang digunakan untuk membangun model ontologi adalah *Methontology*.

Penelitian ini akan membangun sistem rekomendasi produk *bodycare* yang cocok dengan kebutuhan pengguna dalam upaya mengatasi kesulitan dari konsumen. Dengan memanfaatkan teknologi web semantik, sistem rekomendasi ini dibangun dengan data yang dihubungkan dengan ontologi sebagai dasar pengetahuan. Pembangunan ontologi akan dilakukan memanfaatkan metode *Methontology*, sementara peningkatan sistem akan memanfaatkan metode *Prototyping*, dan rekomendasi produk akan ditawarkan menggunakan metode *Collaborative Filtering*. Oleh karena itu, diharapkan penelitian ini dapat memberikan bantuan kepada konsumen dalam memilih produk yang cocok dengan keperluan mereka.

2. Metode Penelitian

Design Science Research Methodology (DSRM) merupakan metode yang akan dimanfaatkan pada penelitian ini, yang melibatkan serangkaian tahap dalam pelaksanaan penelitian. Tahapam tersebut meliputi identifikasi permasalahan, solusi objek, desain dan pengembangan, demonstrasi dan evaluasi, serta komunikasi[5].

2.1 Identifikasi Permasalahan

Identifikasi permasalahan adalah langkah pertama dalam penelitian ini, memahami permasalahan yang akan diteliti menjadi fokus utama dalam penelitian ini. Masalah yang ditemukan adalah keragaman produk *bodycare* di pasar yang menyulitkan pengguna dalam menentukan produk yang cocok dengan keperluan individu masing-masing pengguna. Keragaman jenis serta kriteria produk *bodycare* membuat konsumen kebingungan dalam menentukan pilihan yang tepat untuk kebutuhan masing-masing individu. Oleh karena itu, penentuan produk *bodycare* yang sesuai kebutuhan menjadi suatu permasalahan bagi konsumen.

2.2 Solusi Objek

Tahap ini bertujuan untuk menemukan solusi atas permasalahan yang telah diidentifikasi. Solusi tersebut adalah dengan membangun sistem rekomendasi yang mampu memberikan bantuan dalam menentukan produk *bodycare* yang cocok dengan keperluan masing-masing individu. Sistem rekomendasi ini akan menggunakan pendekatan ontologi untuk memproses informasi secara semantik dengan metode *Methontology* dan menggabungkannya dengan pencarian rekomendasi. Serta pengembangan sistem dengan metode *Prototyping* untuk mempercepat proses pengembangan. Metode *collaborative filtering* digunakan untuk pencarian rekomendasi dalam sistem ini.

2.3 Desain dan Pengembangan

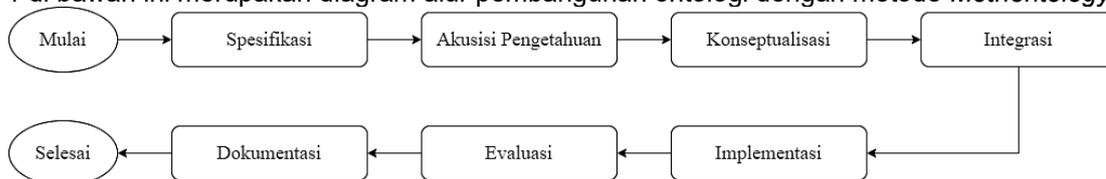
Tahap ini akan dilakukan proses pengumpulan data, pembangunan ontologi, pembangun sistem rekomendasi, dan perancangan sistem.

1. Pengumpulan Data

Penelitian ini menggunakan data produk *bodycare* merek lokal yang tersedia pada *platform* belanja online. Pada penelitian ini, memanfaatkan dua cara yang hendak diterapkan untuk mengumpulkan data, yakni metode yang digunakan untuk pembangunan model ontologi serta metode yang dimanfaatkan guna mengukur serta mengevaluasi sistem. Dalam penelitian ini, sebanyak 105 data produk *bodycare* telah berhasil dikumpulkan oleh penulis.

2. Pembangunan Ontologi

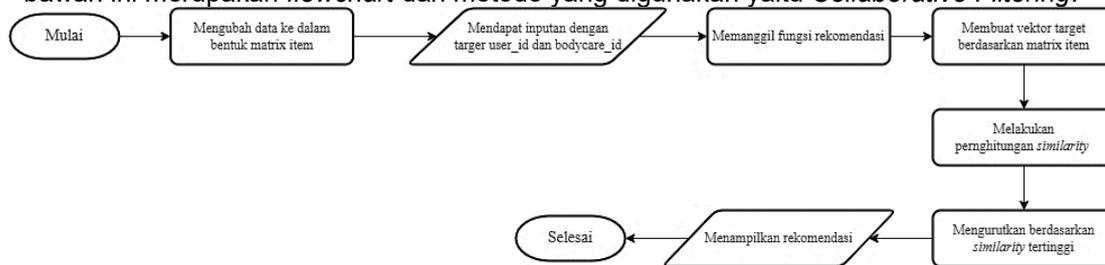
Pada tahap ini, digunakan untuk membangun model ontologi dengan metode *Methontology*. *Methontology* adalah metode terstruktur untuk membangun ontologi dari awal. Pada Gambar 1 di bawah ini merupakan diagram alur pembangunan ontologi dengan metode *Methontology*.



Gambar 1. Diagram Alur *Methontology*

3. Pembangunan Sistem Rekomendasi

Dalam langkah ini, sistem rekomendasi akan dikembangkan dengan memanfaatkan metode *collaborative filtering* dan menerapkan *Adjust Cosine Similarity algorithm*. Pada gambar 2 di bawah ini merupakan *flowchart* dari metode yang digunakan yaitu *Collaborative Filtering*.



Gambar 2. Flowchart Metode *Collaborative Filtering*

4. Perancangan Sistem

Pada langkah perancangan sistem, metode *Prototyping* akan diterapkan. Dengan tujuan untuk menghasilkan model yang dapat dikembangkan menjadi sistem final dengan cepat dan efisien dari segi biaya[6]. Terdapat dua jenis analisis kebutuhan sistem, yakni kebutuhan fungsional dan kebutuhan non-fungsional.

- Kebutuhan Fungsional

Kebutuhan fungsional digunakan untuk menentukan fitur-fitur yang diperlukan oleh sistem. Pada tabel 1 di bawah ini merupakan kebutuhan fungsional sistem.

Tabel 1. Kebutuhan Fungsional Sistem

Code	Deskripsi Kebutuhan Sistem	Target
KF-01	Sistem mempersiapkan proses login serta logout bagi admin sesuai dengan username dan password yang dimiliki.	Admin
KF-02	Sistem menyediakan fasilitas dalam melakukan pencarian serta penjelajahan dari produk berdasarkan kriteria yang telah di inputkan oleh user.	Pengguna
KF-03	Sistem menyediakan fasilitas dalam menampilkan hasil rekomendasi dari produk berdasarkan similarity dari rating yang diberikan oleh user.	Pengguna
KF-04	Sistem menyediakan fasilitas untuk admin dapat menambahkan, mengedit, dan menghapus data produk di dalam website.	Admin
KF-05	Sistem menyediakan fasilitas untuk dapat melihat produk melalui detail page pada setiap produk untuk mendapatkan informasi yang lebih lengkap.	Pengguna

- Kebutuhan Non-Fungsional

Kebutuhan non-fungsional diterapkan sebagai penentu komponen-komponen pendukung yang akan digunakan guna mendukung penelitian ini, seperti *hardware* serta *software* yang

dibutuhkan guna menciptakan sistem ini akan diidentifikasi. Pada tabel 2 dan tabel 3 di bawah ini merupakan kebutuhan non fungsional sistem.

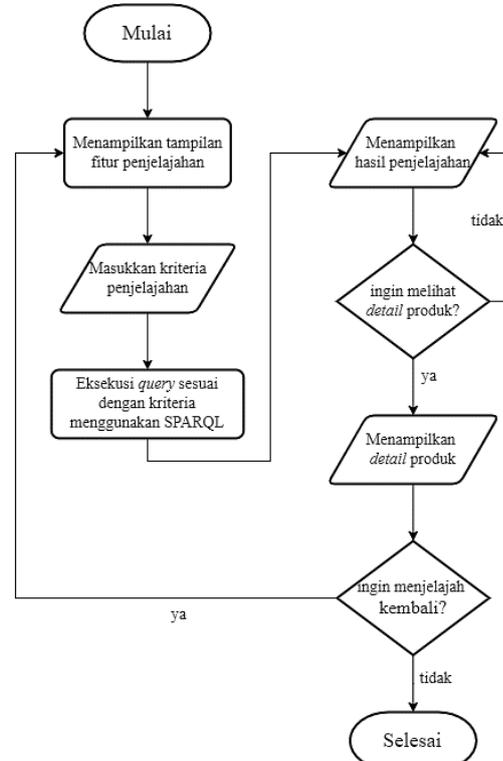
Tabel 2. Kebutuhan Non Fungsional Sistem (*Hardware*)

No.	Perangkat	Spesifikasi Laptop
1	Processor	AMD Ryzen 7 4700U with Radeon Graphics.
2	Monitor	14 inch
3	RAM	8 GB
4	HDD	512 GB

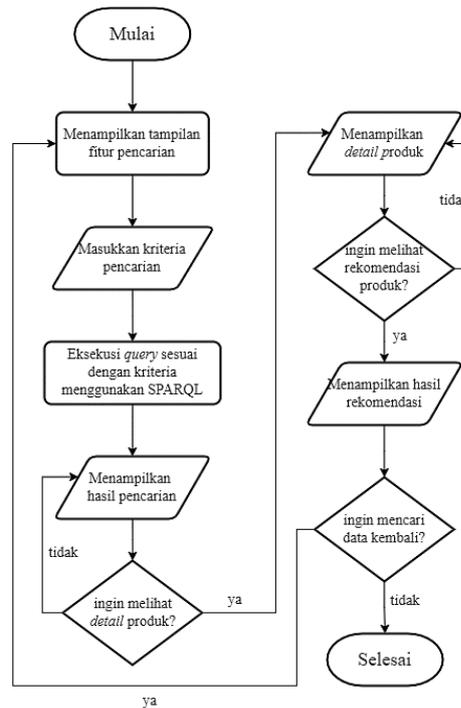
Tabel 3. Kebutuhan Non-Fungsional Sistem (*Software*)

No.	Perangkat	Spesifikasi Laptop
1	Windows 10	Sistem operasi.
2	Protégé 5.5.0	Perangkat yang digunakan untuk membuat, mengedit, dan menyimpan model ontologi.
3	XAMPP	Perangkat yang digunakan untuk server (local host) local.
4	Apache Jena Fuseki	Perangkat yang digunakan untuk mengakses SPARQL dalam mengolah RDF.
5	Visual Studio Code	Perangkat yang digunakan untuk text editor dalam pembuatan program berbasis website.
6	Figma	Perangkat yang digunakan untuk mendesain, membuat wireframe, memastikan pengalaman pengguna yang optimal, UI/UX Design dan mengembangkan prototype aplikasi.
7	HDD	512 GB

Prototype yang akan dirancang berbasis web menggunakan bahasa pemrograman JavaScript. *Prototype* ini akan memiliki tiga fitur yaitu penjelajahan, pencarian, dan rekomendasi. Pada gambar 3 dan gambar 4 di bawah ini merupakan *flowchart* pada setiap fitur yang terdapat pada sistem.



Gambar 3. *Flowchart* Fitur Penjelajahan



Gambar 4. Flowchart Fitur Pencarian dan Rekomendasi

2.4 Demonstrasi dan Evaluasi

Pada tahap demonstrasi bertujuan untuk menguji sistem yang telah dikembangkan berjalan sesuai dengan perencanaan awal atau sesuai dengan keinginan. Lalu, tahap evaluasi memiliki tujuan untuk menilai dan memastikan model ontologi yang dibangun telah sesuai rancangan yang telah disusun. Pada tahap pengujian ini, adapun dua bentuk pengujian. Pengujian pertama akan memanfaatkan *usability testing* untuk mengevaluasi seberapa mudah dan efektif pengguna dapat menggunakan sistem tersebut. Pengujian kedua dilakukan dengan menggunakan *Mean Absolute Error (MAE)* digunakan untuk mengetahui akurasi dari metode rekomendasi yang digunakan[7].

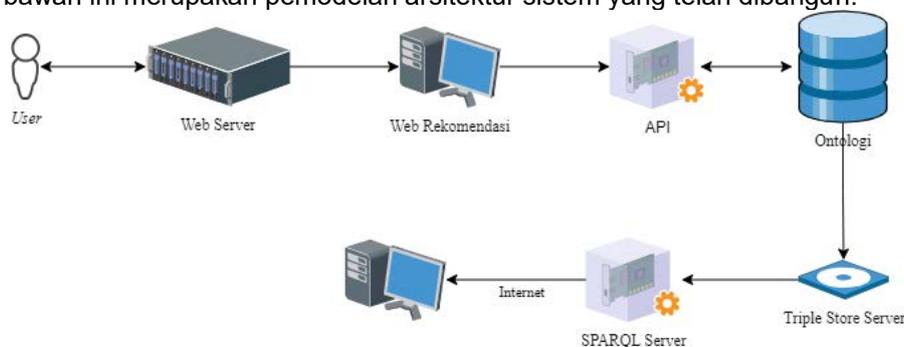
2.5 Komunikasi

Pada tahap komunikasi ini, dilakukan proses dokumentasi seluruh pengetahuan terkait penelitian ini agar hasilnya dapat diarsipkan dalam bentuk tugas akhir yang kemudian dapat dipublikasikan dalam jurnal ilmiah.

3. Hasil dan Pembahasan

3.1 Arsitektur Sistem

Arsitektur sistem (*system architecture*) adalah kerangka kerja yang digunakan untuk merancang, mengintegrasikan, dan mengelola komponen-komponen yang membentuk sebuah sistem. Pada gambar 5 di bawah ini merupakan pemodelan arsitektur sistem yang telah dibangun.

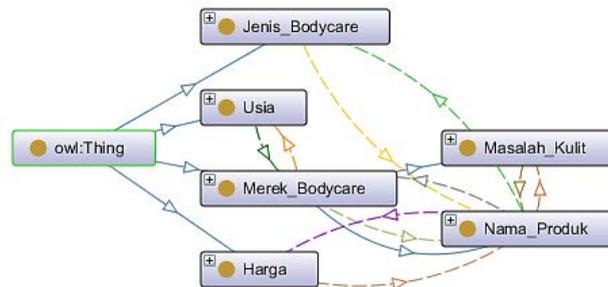


Gambar 5. Pemodelan Arsitektur Sistem

3.2 Pembangunan Model Ontologi

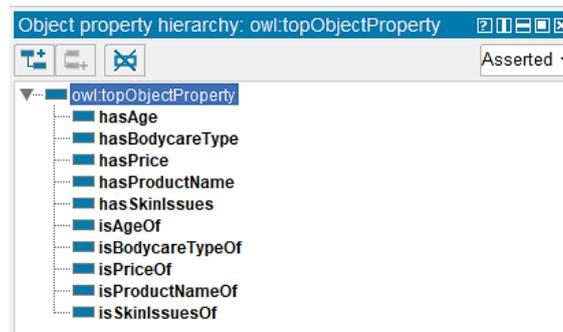
Sistem Rekomendasi Produk *Bodycare* Dengan Metode *Collaborative Filtering*

Model ontologi telah dirancang dengan menggunakan *tools* Protégé. Implementasi perancangan model ontologi dalam domain *bodycare* menghasilkan ontologi dengan total 7 *class*, 11 *object properties* yang dapat dilihat pada gambar 7 di bawah ini, 7 *data properties*, dan 132 *individuals* yang dapat dilihat pada gambar 8 di bawah ini.



Gambar 6. Diagram Ontograf Produk *Bodycare*

Pada gambar 6 di atas merupakan diagram ontograf dari produk *bodycare* yang terdiri dari *class-class* yang telah dibangun dengan menggunakan *tools* Protégé.



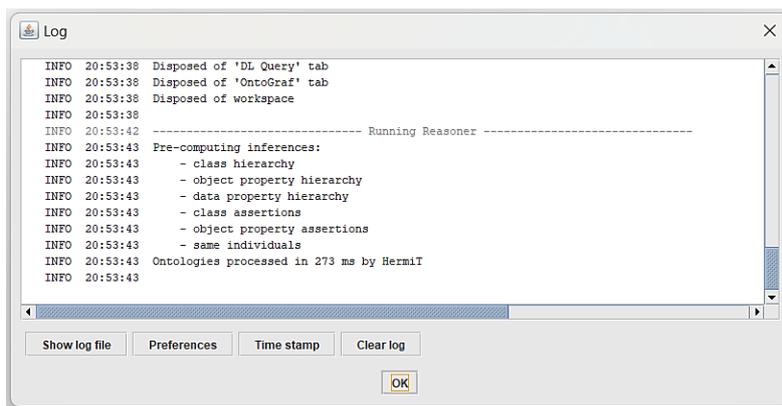
Gambar 7. *Object Properties* Pada Ontologi Produk *Bodycare*

Pada gambar 7 di atas merupakan *object properties* pada ontologi produk *bodycare* yang telah dibangun dengan menggunakan *tools* Protégé.



Gambar 8. *Individuals* Pada Ontologi Produk *Bodycare*

Pada gambar 8 di atas menunjukkan *individuals* pada ontologi produk *bodycare* yang merupakan data produk *bodycare* yang telah didapatkan lalu dibangun dengan menggunakan *tools* Protégé.



Gambar 9. Log Proses Reasoning Ontologi Produk Bodycare

Pada gambar 9 di atas menunjukkan proses *reasoners* telah berhasil dilakukan dan tidak terdapat pesan *error* pada log yang dimana hal tersebut menandakan bahwa ontologi yang dibangun telah konsisten.

Metrics	
Axiom	2133
Logical axiom count	1979
Declaration axioms count	154
Class count	7
Object property count	11
Data property count	7
Individual count	132
Annotation Property count	0

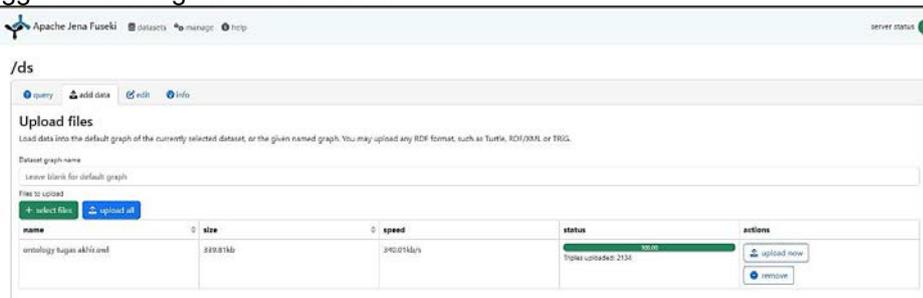
Gambar 10. Informasi Metrics Ontologi Produk Bodycare

Pada gambar 10 di atas merupakan *metric* ontologi yang disusun untuk dapat memberikan gambaran secara matematis komponen yang ada dalam rancangan ontologi yang telah dibangun tersebut.

3.3 Implementasi Sistem

1. Implementasi Ontologi ke Dalam Sistem

Server Apache Jena Fuseki dimanfaatkan untuk mengimplementasikan ontologi produk *bodycare* ke dalam sistem rekomendasi. Pada gambar 11 di bawah ini merupakan proses dari pengunggahan ontologi ke server fuseki.



Gambar 11. Proses Pengunggahan Ontologi ke Server Fuseki

2. Implementasi Fitur

Pada tahap ini merupakan implementasi dari kode program dengan menggunakan metode *collaborative filtering* yang mengimplementasikan Node JS dan menggunakan *framework* Express JS.

a. Implementasi Fitur Penjelajahan

Tabel 4. Source Code Fitur Penjelajahan

```
fetch(
    `http://localhost:3000/web/api/query-by-
merek?merek=${merekPilih}`
)
```

```

        .then((respons) => respons.json())
        .then((bodycareList) => {
            console.log("List", bodycareList[0]);
            bodycareList.forEach((bodycare, index) => {
                const card = document.createElement("div");
                card.className = "card pilihan";

                const          parseharga          =
                parseInt(bodycare.Harga.value);
                const          formattedHarga      =
                parseharga.toLocaleString("id-ID");

```

Pada tabel 4 di atas merupakan *source code* untuk melakukan panggilan ke API dengan menggunakan metode HTTP GET ke endpoint `http://localhost:3000/web/api/query-by-merek?merek=${merekPilih}`, di mana ``${merekPilih}`` adalah parameter merek yang dipilih. Setelah mendapatkan respons dari server, kode tersebut mengonversi respons menjadi objek JSON menggunakan method `.json()`. Setelah itu, dengan menggunakan method `.forEach()`, kode tersebut melakukan iterasi melalui setiap elemen dalam array `bodycareList`. Untuk setiap elemen, kode membuat elemen HTML baru dengan menggunakan `document.createElement("div")` yang memiliki class `card pilihan`. Kemudian, kode mengonversi nilai harga dari string menjadi bilangan bulat menggunakan `parseInt(bodycare.Harga.value)`, dan mengonversi bilangan bulat tersebut menjadi string dengan format mata uang Indonesia menggunakan `toLocaleString("id-ID")`. Pada gambar 12 di bawah ini merupakan tampilan fitur penjelajahan pada sistem.



Gambar 12. Tampilan Fitur Penjelajahan Pada Sistem

b. Implementasi Fitur Pencarian

Tabel 5. *Source Code* Fitur Pencarian

```

fetch(
`http://localhost:3000/web/pencarian?queryFilters=${encodeURIComponent(
    queryFilters
)}` ,
{
    method: "GET",
}
)
.then((response) => response.json())
.then((data) => {
    console.log("Data: ", typeof data);

```

Pada tabel 5 di atas merupakan *source code* untuk melakukan panggilan ke endpoint pencarian API lokal dengan menggunakan metode GET. QueryFilters yang dikirimkan dienkripsi menggunakan `encodeURIComponent` untuk memastikan karakter khusus seperti spasi atau tanda kurung kurawal dienkripsi dengan benar. Setelah menerima respons dari server, respons tersebut diubah menjadi objek JSON. Selanjutnya, jenis data respons dicetak ke konsol untuk keperluan debug. Pada gambar 13 di bawah ini merupakan tampilan fitur pencarian pada sistem.



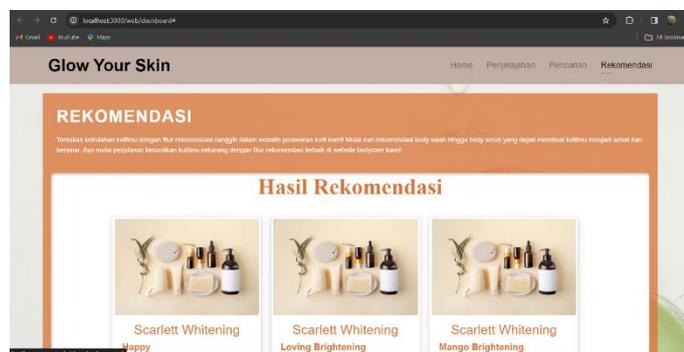
Gambar 13. Tampilan Fitur Pencarian Pada Sistem

c. Implementasi Fitur Rekomendasi

Tabel 6. Source Code Fitur Rekomendasi

```
fetch(`http://localhost:3000/web/dataAlgoritma`)
    .then((response) => response.json())
    .then((dataAlgoritma) => {
        const formattedData =
dataAlgoritma.map((item) => ({
            transaksi_id: item.id_transaksi,
            user_id: item.id_pelanggan,
            produk_id: item.id_produk,
            rating: item.rating,
        }));
```

Pada tabel 6 di atas merupakan *source code* untuk mengirimkan permintaan ke endpoint `http://localhost:3000/web/dataAlgoritma` menggunakan metode `fetch()`. Setelah menerima respons dari server, respons tersebut diubah menjadi objek JSON. Selanjutnya, data algoritma yang diterima diolah dengan melakukan pemetaan setiap item dari array `dataAlgoritma` ke objek baru dengan properti `transaksi_id`, `user_id`, `produk_id`, dan `rating`, yang kemudian disimpan dalam variabel `formattedData`. Pada gambar 14 di bawah ini merupakan tampilan fitur rekomendasi pada sistem.



Gambar 14. Tampilan Fitur Rekomendasi Pada Sistem

3.4 Implementasi Hasil Pengujian Sistem

Dalam langkah ini akan dilakukan tahapan pengujian sistem rekomendasi yang telah dibangun. Pengujian yang dilakukan pada sistem ini berupa pengujian dengan *usability testing* dan MAE.

1. *Usability Testing*

a. Pengujian Validitas

Uji validitas adalah tahap untuk menilai apakah suatu instrumen pengukur, seperti kuesioner, memiliki nilai yang valid atau tidak. Instrumen pengukur yang dimaksud merujuk kepada pernyataan-pernyataan yang ada dalam kuesioner.. Pengujian validitas ini melibatkan perhitungan nilai r tabel dan r hitung dari setiap item yang akan dievaluasi. Dari uji validitas yang dilakukan, dapat diketahui bahwa nilai r tabel yang didapatkan dari 32 responden yaitu 0,349 yang korelasi antara setiap item variabel usability dan total skor A1 hingga E10 menunjukkan hasil yang signifikan. Nilai r tabel untuk setiap item lebih kecil dari r hitung, yang menunjukkan bahwa semua item variabel konstruk adalah valid.

b. Pengujian Reliabilitas

Agar dapat mengevaluasi keseragaman atau kestabilan respons yang diberikan oleh responden mengenai pertanyaan dalam kuesioner, maka dilakukan uji reliabilitas ini. Dalam pengujian ini, diterapkan koefisien Cronbach's Alpha dengan nilai koefisiennya yaitu 0,70 yang apabila melebihi nilai tersebut, maka dapat disimpulkan bahwa instrumen penelitian tersebut dapat dianggap reliabel[8].

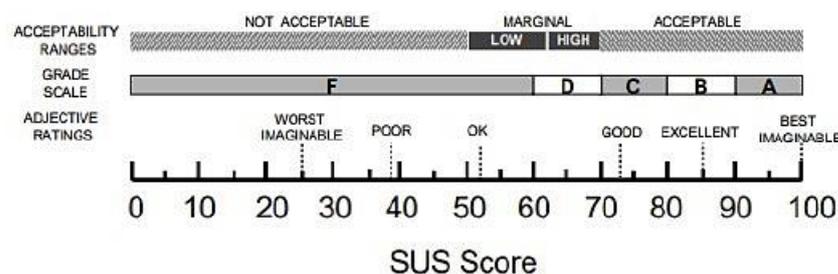
Tabel 7. Hasil Uji Reliabilitas

Cronbach's Alpha	r Tabel	Keterangan
0,809	0,70	Reliabel

Pada tabel 7 di atas menunjukkan hasil dari uji reliabilitas yang telah dilakukan, dari pengujian tersebut diperoleh nilai Cronbach's Alpha sebesar 0,809 yang dimana hal tersebut menunjukkan bahwa kuesioner dinyatakan reliabel karena nilai Cronbach's Alpha lebih besar dari nilai r table yaitu $0,809 > 0,70$.

c. Pengujian Kebergunaan Sistem

Dalam tahapan pengujian ini, data yang digunakan merupakan data yang diperoleh dari pengumpulan data online yang dilakukan melalui Google Form yang telah dilengkapi oleh partisipan. Kuesioner tersebut disusun menggunakan skala *Likert* untuk memastikan bahwa responden dapat memberikan jawaban dengan keyakinan terhadap pernyataan-pernyataan yang tercantum dalam kuesioner[9]. Data kuesioner tersebut akan dihitung dengan menggunakan metode *System Usability Scale* (SUS). Dalam penilaian yang telah dilakukan dari 32 partisipan, didapatkan skor rata-rata yakni sebesar 82,344. Langkah berikutnya adalah menentukan grade penilaian dengan metode *Acceptability*, *Grade Scale*, dan *Adjective Rating*.

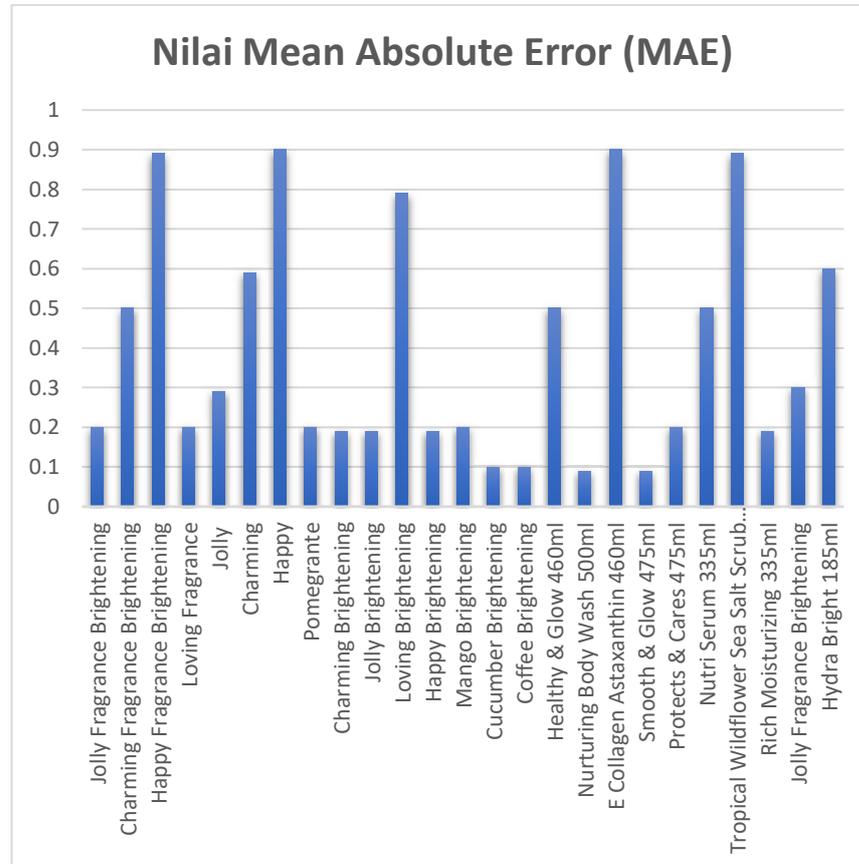
Gambar 15. Penilaian Dengan *Acceptability*, *Grade Scale*, dan *Adjective Rating*

Gambar 15 di atas menunjukkan skala penilaian yang digunakan untuk menetapkan tingkat *Acceptability*, *Grade Scale*, dan *Adjective Rating*. Berdasarkan hasil penilaian, akan dibandingkan dengan skor rata-rata sebesar 82,344 yakni:

- Tingkat *Acceptability Range* pengguna terhadap sistem rekomendasi produk perawatan tubuh yang dibangun masuk dalam kategori *Acceptable*.
- Tingkat *Grade Scale* pengguna terhadap sistem rekomendasi produk perawatan tubuh yang dibangun masuk dalam kategori B.

- Tingkat *Adjective Rating* pengguna terhadap sistem rekomendasi produk perawatan tubuh yang dibangun masuk dalam kategori GOOD.

2. Pengujian Akurasi Sistem



Gambar 16. Grafik Hasil Pengujian *Mean Absolute Error* (MAE)

Pada gambar 16 di atas merupakan pengujian dengan persamaan *Mean Absolute Error* (MAE) diterapkan untuk mengetahui tingkat akurasi dari suatu sistem rekomendasi yang memanfaatkan metode *collaborative filtering*. Dalam melakukan perhitungan ini, digunakan hasil nilai *Mean Absolute Error* (MAE) dari 25 produk bodycare. Berdasarkan informasi grafik di atas, pengujian yang telah dilakukan dengan menggunakan *Mean Absolute Error* (MAE) menghasilkan nilai sebesar 0,3556. Apabila semakin mendekati nilai 0, maka tingkat akurasi semakin tepat atau akurat dan semakin baik terhadap rekomendasi produk yang diberikan [10]. Hasil nilai MAE yang diperoleh yaitu 0,3556 menunjukkan bahwa tingkat akurasi yang dihasilkan sudah akurat dan rekomendasi produk *bodycare* sudah baik.

4. Kesimpulan

Pada pengujian *usability system* yang memanfaatkan metode *System Usability Scale* (SUS) diperoleh skor rata-rata yaitu sebesar 82,344. Kemudian dilihat dari perspektif metode *Acceptability*, *Grade Scale* dan *Adjective Rating*, dihasilkan tingkat *Acceptability Range* pengguna terhadap sistem termasuk ke dalam kategori *Acceptable*, tingkat *Grade Scale* termasuk ke dalam kategori B, dan tingkat *Adjective Rating* termasuk ke dalam kategori GOOD. Maka dapat disimpulkan bahwa sistem rekomendasi ini memiliki kinerja *usability system* yang baik.

Sedangkan, untuk pengujian akurasi sistem dengan menggunakan *Mean Absolute Error* (MAE) diperoleh nilai sebesar 0,3556 yang dapat diartikan bahwa tingkat akurasi yang dihasilkan sudah akurat dan rekomendasi produk *bodycare* sudah baik karena semakin mendekati nilai 0, maka tingkat akurasi semakin tepat atau akurat dan semakin baik terhadap rekomendasi produk *bodycare* yang diberikan.

Referensi

- [1] Pramarta, C. (2020). Pengembangan Ontologi Tujuan Wisata Bali Dengan Pendekatan Kukul Knowledge Framework. *SINTECH (Science and Information Technology) Journal*, 3(2), 77–89. <https://doi.org/10.31598/sintechjournal.v3i2.592>.
- [2] Wulandari, I. A., Pahu, G. Y. K. S. S., & Rahayu, P. (2020). Peran Ontologi dalam Pengembangan Sistem Rekomendasi pada Domain Online Learning. *Jurnal Komtika (Komputasi Dan Informatika)*, 4(1), 1–8. <https://doi.org/10.31603/komtika.v4i1.3535>.
- [3] Erlangga, E., & Sutrisno, H. (n.d.). *Jurnal Manajemen Sistem Informasi dan Teknologi Sistem Rekomendasi Beauty Shop Berbasis Collaborative Filtering*. 10(2), 2745–2765.
- [4] Puspa Setia Pratiwi. (2022). Perancangan Sistem Rekomendasi Berbasis Model Ontologi untuk Rekomendasi Tempat Magang Mahasiswa. *Jurnal Ilmiah Indonesia*, 7(2), 7645–7654.
- [5] I Kadek Anom Sukawirasa Putra, Cokorda Rai Adi Paramartha, I Gusti Ngurah Anom Cahyadi Putra, I Ketut Gede Suhartana, I Made Widiartha, I Komang Ari Mogi, "Pengembangan Sistem Informasi Banten Menggunakan Web Semantik," *JELIKU (Jurnal Elektronik Ilmu Komputer Udayana)*, 11(2), 319. <https://doi.org/10.24843/jlk.2022.v11.i02.p11>.
- [6] Sari, A. M., Yani, D., & Suryani, D. (2021). Implementasi Aplikasi Mobile Peta NKRI (Negara Kesatuan Republik Indonesia) Berbasis Android Menggunakan Metode Prototype. *Journal of Information System Research (JOSH)*, 2(4), 288–292. <https://doi.org/10.47065/josh.v2i4.793>.
- [7] Andik Adi Suryanto, Asfan Muqtadir, 2019. PENERAPAN METODE MEAN ABSOLUTE ERROR (MAE) DALAM ALGORITMA REGRESI LINEAR UNTUK PREDIKSI PRODUKSI PADI. *Jurnal Sains dan Teknologi*, Volume 11, p. 79.
- [8] Aviva Lutfiana, Hasna Hafizhah Salma, Gary Collins Brata Winardy, dan Ali Nina Liche Seniati. (2023). UJI RELIABILITAS DAN VALIDITAS SKALA KARAKTER BAIK (SKB) UNTUK SISWA SMA/SEDERAJAT DI INDONESIA. *Jurnal Penelitian dan Pengukuran Psikologi*, 114.
- [9] Made Agung Raharja, Susy Purnawati, I Putu Gede Adiatmika, and I Nyoman Adiputra, 2021. Usability Analysis of Tembang Sekar Alit Learning (SekARAI) Applications Using The Human Computer Interaction (HCI) Model In Bali Students. *Proceedings of the Second Asia Pacific International Conference on Industrial Engineering and Operations Management Surakarta, Indonesia*, p. 2874.
- [10] Imelda Lubis, Y., Josua Napitupulu, D., & Satia Dharma, A. (2020). Implementasi Metode Hybrid Filtering (Collaborative dan Content-based) untuk Sistem Rekomendasi Pariwisata Implementation of Hybrid Filtering (Collaborative and Content-based) Methods for the Tourism Recommendation System.

Penyisipan Digital Signature Ke Dalam Citra Digital Sebagai Keamanan Hak Cipta Dengan Metode DES dan BPCS

Michael Tanaya.^{a1}, I Gede Arta Wibawa.^{a2}, I Made Widiartha.^{b3}, Luh Gede Astuti.^{b4}

^aInformatika, Universitas Udayana
Kuta Selatan, Badung, Bali, Indonesia
¹michaeltanaya94@gmail.com
²gede.arta@unud.ac.id (Corresponding author)

^bInformatika, Universitas Udayana
Kuta Selatan, Badung, Bali, Indonesia
³madewidiartha@unud.ac.id
⁴lg.astuti@unud.ac.id

Abstrak

Dalam era digital saat ini, distribusi dan reproduksi konten digital menjadi sangat mudah, sehingga keamanan hak cipta citra menjadi isu yang semakin penting. Untuk melindungi integritas dan keaslian citra digital, penyisipan tanda digital telah menjadi pilihan yang efektif sebagai tanda integritas dan keaslian dari dokumen tersebut. Penelitian ini mengusulkan pendekatan inovatif yang menggabungkan metode Data Encryption Standard (DES) dan Bit-Plane Complexity Segmentation (BPCS) untuk menyisipkan tanda digital ke dalam citra digital guna meningkatkan keamanan hak cipta. Pada tahap pertama, tanda digital dienkripsi menggunakan algoritma DES untuk memastikan kerahasiaan data sebelum penyisipan. Tahap berikutnya melibatkan penerapan metode BPCS, di mana tanda digital yang telah dienkripsi disisipkan ke dalam bit-bit rendah signifikan citra digital. Keunggulan metode BPCS terletak pada kemampuannya menyisipkan data tanpa mengorbankan kualitas visual citra secara signifikan, menjaga estetika visual yang penting dalam konteks hak cipta. Dengan menggabungkan kedua metode ini, diharapkan citra digital dapat dilindungi secara lebih efektif dari ancaman pembajakan atau manipulasi yang mengancam hak cipta.

Keywords: *Penyisipan tanda digital, citra digital, keamanan hak cipta, Data Encryption Standard (DES), Bit-Plane Complexity Segmentation (BPCS).*

1. Pendahuluan

Pada Era Revolusi Industri 4.0, penggunaan teknologi digital semakin berkembang di berbagai sektor, mulai dari bisnis hingga pemerintahan. Maka dari itu, keamanan data digital menjadi isu yang semakin penting, karena data digital sangat mudah untuk disalin, dipalsukan dan dimodifikasi oleh pihak yang tidak bertanggung jawab. Oleh karena itu, banyak teknologi pengamanan telah dikembangkan, salah satunya adalah digital signature.

Digital signature adalah tanda digital yang berfungsi sebagai penanda kepemilikan sebuah dokumen digital. Digital Signature memungkinkan seseorang untuk memastikan keaslian dan integritas dokumen digital tersebut. Dalam sistem, digital signature hanya diketahui oleh pemilik dokumen, dikarenakan terdapat sebuah kunci rahasia yang digunakan untuk menghasilkan nilai hash atau tanda digital dari dokumen tersebut. [6] Nilai hash ini yang kemudian disisipkan ke dalam dokumen sebagai suatu tanda digital. Namun, dalam beberapa kasus, digital signature masih rentan terhadap serangan oleh pihak yang

tidak bertanggung jawab. Salah satu serangan yang mungkin terjadi adalah serangan man-in-the-middle (MITM), dimana pihak yang tidak bertanggung jawab akan memodifikasi tanda digital pada dokumen tersebut sehingga dokumen tersebut terlihat sah meskipun sebenarnya telah dimodifikasi. Untuk mengatasi masalah ini, teknologi bernama steganografi dapat digunakan untuk menyisipkan digital signature pada citra digital.[3]

Steganografi adalah suatu teknik yang digunakan untuk menyembunyikan suatu pesan atau informasi rahasia dalam sebuah media yang tampak normal. Citra digital adalah salah satu media yang sering digunakan untuk penyisipan pesan atau informasi rahasia.

Dalam proposal ini, akan membahas mengenai penyisipan digital signature ke dalam citra digital sebagai keamanan hak cipta menggunakan metode Bit-Plane Complexity Segmentation (BPCS). Metode BPCS adalah salah satu teknik steganografi yang menggunakan analisis kompleksitas bit-plane dalam citra digital untuk menyisipkan pesan rahasia.[2] Dalam kasus ini, digital signature akan disisipkan pada bit-plane yang memiliki kompleksitas tertentu sehingga tidak merusak kualitas visual dari citra digital tersebut. [5]

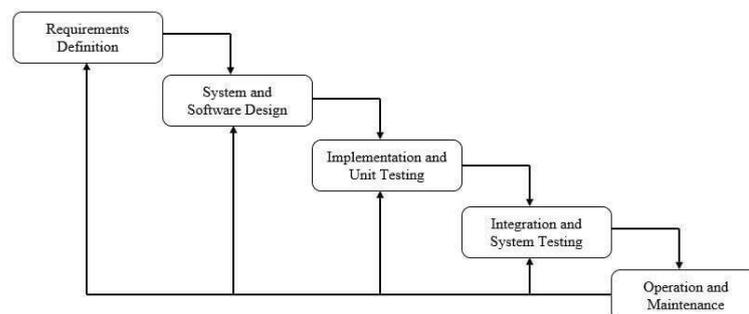
Pada penelitian sebelumnya, metode BPCS ini digunakan untuk menyisipkan sebuah pesan rahasia pada pas foto digital. Dimana pesan yang digunakan pada penelitian sebelumnya itu adalah sebuah teks, sehingga pada penelitian ini akan digunakan digital signature berformat image untuk nantinya disisipkan pada sebuah citra digital.

Penyisipan digital signature pada citra digital dengan menggunakan steganografi dapat meningkatkan keamanan dokumen digital dengan cara menyembunyikan informasi rahasia yang penting.[8] Oleh karena itu, proposal ini diharapkan dapat memberikan kontribusi dan manfaat dalam pengembangan teknologi keamanan informasi khususnya pada penggunaan digital signature. Selain itu, dengan adanya proposal ini, diharapkan dapat memicu minat untuk pengembangan teknologi steganografi pada digital signature di Indonesia.

2. Metodologi Penelitian

2.1 Metode Pengembangan Sistem

Aplikasi steganografi ini menggunakan metode pengembangan sistem Waterfall Mode. Alur pengembangan sistem dengan metode waterfall dapat dilihat pada gambar 1 :



Gambar 1. Waterfall Method

Berikut adalah penjelasan dari gambar 1 untuk setiap proses yang ada di dalam pengembangan sistem menggunakan metode Waterfall :

2.1.1 Requirement Definition

Pada tahapan ini melakukan persiapan dan analisa kebutuhan sistem yang ingin dikembangkan. Persiapan dan analisa sistem seperti, persiapan data,

manfaat dari sistem yang akan dibuat dan batasan-batasan dari sistem yang akan dibuat.

2.1.2 System and Software Design.

Pada tahapan ini melakukan proses perancangan desain sistem dengan menggunakan diagram UML seperti, use case diagram, flowchart dan activity diagram.

2.1.3 Implementation

Pada tahapan ini merupakan tahapan gambaran sistem ke dalam bentuk code. Jadi pada tahapan ini lebih berfokus pada hal teknis dari hasil desain yang telah dibuat pada tahap sebelumnya dan desain tersebut akan diterjemahkan ke dalam bahasa pemrograman.

2.1.4 System Testing

Pada tahapan ini sistem yang telah dibuat akan diuji dari berbagai sisi baik dari aspek desain dan fungsionalitas. Dengan tujuan untuk menentukan apakah terjadi kesalahan pada sistem yang telah dibuat, dimana jika ada kesalahan masih bisa dicegah dan diperbaiki kembali.

2.1.5 Operation and Maintenance

Pada tahapan ini merupakan tahapan pemeliharaan sistem yang mencakup perbaikan kesalahan seperti bug, error, dan penambahan fitur-fitur pada sistem yang telah dibuat.

2.2 Variabel Penelitian

Dalam penelitian ini terdapat dua variabel, yaitu variabel bebas dan variabel terikat. Variabel bebas pada penelitian ini adalah digital signature yang dienkripsi menjadi file txt dan file image dengan format **.jpg*, **.png* dan **jpeg*. Sedangkan variabel terikat pada penelitian ini adalah file steganografi dengan format **png*.

2.3 Data Penelitian

Data penelitian yang digunakan berdasarkan cara memperolehnya yaitu data primer dan data sekunder. Data primer adalah data yang diperoleh peneliti secara langsung yaitu digital signature berbentuk image. Data sekunder adalah data yang diperoleh dari berbagai sumber yang telah ada, seperti sebuah foto hasil dari pengambilan gambar oleh seorang photographer.

2.4 Skenario Pengujian

Skenario pengujian pada penelitian ini bertujuan untuk mengukur kualitas dari citra yang telah dimodifikasi dengan citra asli. Citra yang dimodifikasi akan lebih baik jika citra yang telah dimodifikasi tersebut menyerupai aslinya, agar kualitas citra tetap terjaga. Perhitungan Mean Square Error (MSE) ini diperlukan untuk mengetahui besarnya error yang dihasilkan dari proses penyisipan. Perhitungan ini dilakukan untuk setiap piksel dalam citra. MSE yang dihitung adalah MSE dari hasil citra yang telah di steganografi dan citra asli memiliki hasil yang mendekati sama dengan citra asli.

Pada perhitungan Peak Signal Noise Ratio (PSNR) ini diperlukan untuk membandingkan citra hasil dengan citra asli. Dimana dengan semakin tinggi nilai PSNR maka citra yang dimodifikasi mirip dengan citra aslinya.

2.4.1 Mean Square Error (MSE)

Pengujian nilai error kuadrat rata-rata antara citra asli dengan citra hasil modifikasi

2.4.2 Peak Signal Noise Ratio (PSNR)

Metode pengujian untuk menghitung rasio antara nilai maksimum yang mungkin dari sebuah sinyal dan kekuatan gangguan distorsi yang mempengaruhi kualitas representasinya.

3. Hasil dan Pembahasan

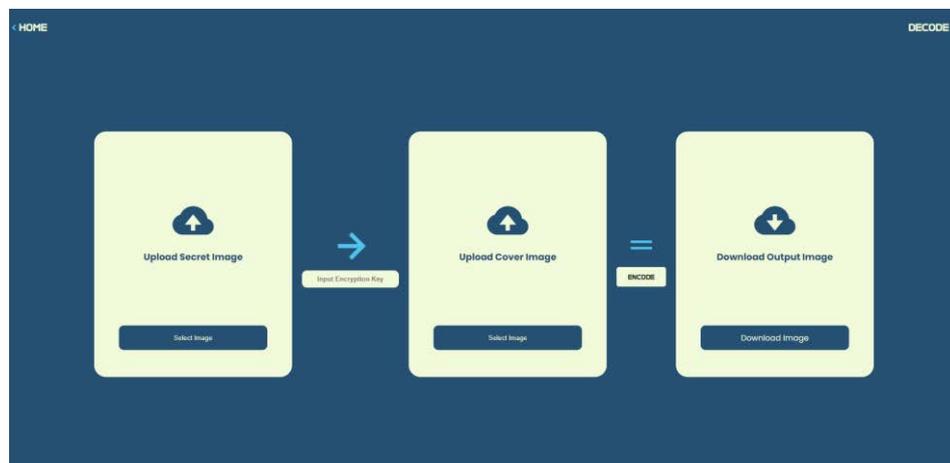
3.1 Implementasi Tampilan

Pada tahap implementasi, website yang telah dibuat akan dijalankan pada server local. Dimana ketika website dibuka, maka akan memunculkan halaman tampilan awal yang berisi sebuah pesan welcome untuk user.



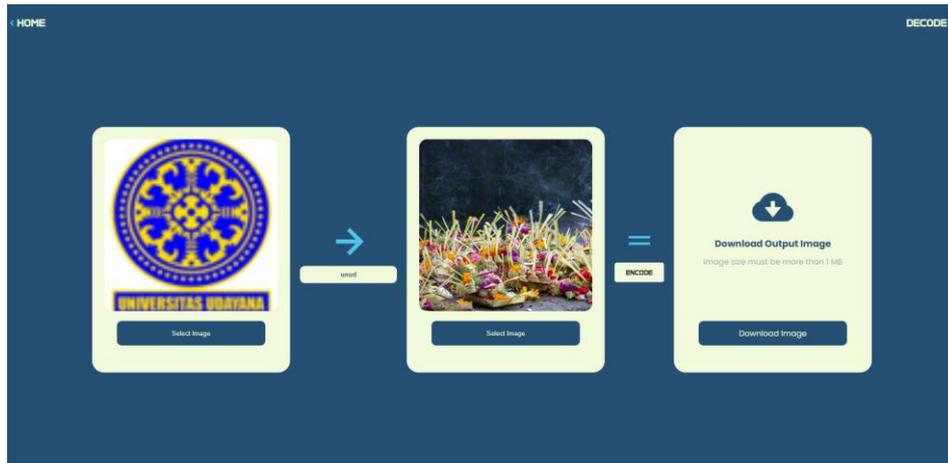
Gambar 2. Tampilan Halaman Home

Bisa dilihat pada gambar 2 diatas. Pada tampilan awal website user akan menemukan sebuah kalimat welcome dan tombol yang akan mengarahkan user ke halaman encode. Setelah user meng-klik tombol tersebut, maka user akan diarahkan ke halaman encode seperti gambar berikut :



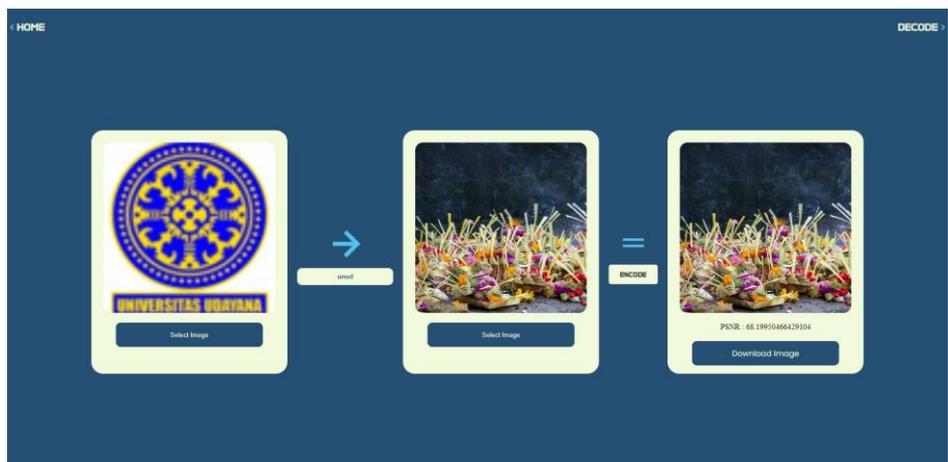
Gambar 3. Tampilan Halaman Encode

Bisa dilihat pada gambar 3 diatas. Ketika user sudah berada di halaman ini maka user sudah bisa melakukan penyisipan digital signature ke dalam citra digital.



Gambar 4. Tampilan Halamn Encode dengan Input Image

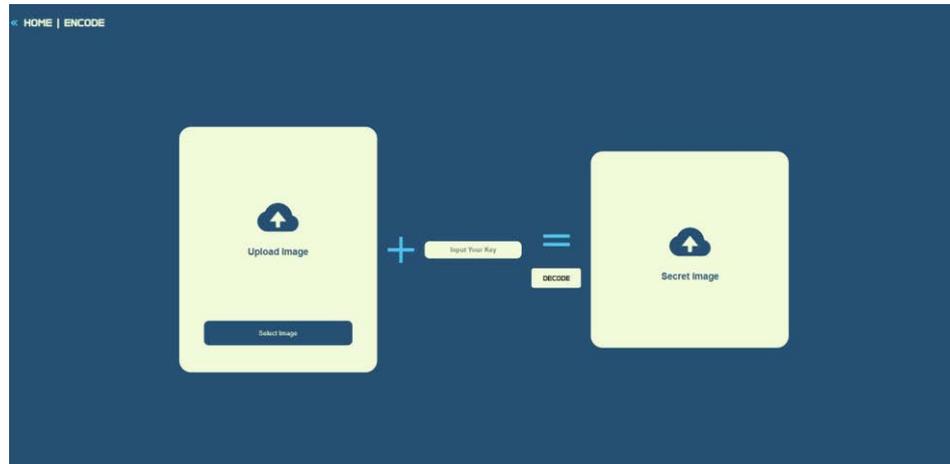
Bisa dilihat pada gambar 4 diatas. Untuk percobaan proses penyisipan digunakan file gambar sebuah desain yang dibuat oleh salah satu mahasiswa unud dan pesan yang disisipkan berupa logo unud sebagai digital signature dari desain mahasiswa tersebut.



Gambar 5. Tampilan Halaman Encode dengan Hasil Image Penyisipan

Bisa dilihat pada gambar 5 diatas. Jika proses enkripsi dan encode berjalan lancar maka akan tampil sebuah gambar. Dimana gambar tersebut adalah stego image yang dikeluarkan sebagai hasil oleh website tersebut, sehingga stego image tersebut dapat di download. Untuk pengujian digunakan pengujian Peak Signal to Noise Ratio (PSNR) yang membandingkan antara nilai sinyal yang diukur dengan besarnya error yang berpengaruh pada sinyal tersebut atau untuk mengetahui perbandingan kualitas citra asli dan citra stegano yang ada dalam satuan decibel (dB).[7] Sebelum menghitung PSNR, nilai error harus dicari terlebih dahulu dengan perhitungan Mean Square Error (MSE). Untuk hasil uji dari citra desain tersebut dapat dilihat ketika proses enkripsi dan encode berhasil dilakukan. Untuk hasilnya dapat dilihat pada gambar diatas. Dari pengujian

diatas didapatkan nilai PSNR yaitu 68.19950466429104, dimana kualitas stegano image yang baik berada pada nilai diatas 40dB. Jadi dapat disimpulkan bahwa hasil penyisipan digital signature pada gambar desain tersebut berkualitas baik.



Gambar 6. Tampilan Halaman Decode

Bisa dilihat pada gambar 6 diatas. Untuk proses dekripsi dan decode file dapat dilakukan dengan klik menu decode terlebih dahulu dan website akan membawa user ke halaman decode. Dimana pada halaman ini user bisa menginputkan stego image dan decrypt key. Untuk tampilannya dapat dilihat pada gambar diatas.



Gambar 7. Tampilan Halaman Decode dengan Input Image

Bisa dilihat pada gambar 7 diatas. Untuk percobaan digunakan file stego image yang merupakan hasil penyisipan yang dilakukan pada proses sebelumnya dan dibutuhkan juga kata kunci yang sebelumnya digunakan.



Gambar 8. Tampilan Halaman Decode dengan Hasil Ekstraksi Image

Bisa dilihat pada gambar 8 diatas. Kemudian user dapat meng-klik tombol decode dan akan dilakukan proses decode dan dekripsi. Jika proses decode dan dekripsi berhasil dilakukan maka digital signature akan muncul.

3.2 Pengujian

Pengujian merupakan salah satu aspek penting dalam pengembangan perangkat lunak. Tujuan dari dilakukannya proses pengujian adalah untuk menjamin bahwa perangkat lunak yang dibangun memiliki kualitas yang handal, dimana dengan adanya pengujian dapat menjamin kualitas serta mengetahui kekurangan dari perangkat lunak yang dibangun.

Hasil gambar stegano yang diciptakan akan berformat *.png sehingga gambar yang dihasilkan tentu memiliki ukuran file yang lebih besar namun tidak mengubah ukuran pixel panjang dan lebar dari gambar asli.

Karena digunakan metode BPCS maka bagian pixel pada gambar yang berubah tentu akan lebih banyak dibandingkan tidak menggunakan metode BPCS, sehingga akan mempengaruhi kualitas dari gambar itu sendiri, maka dilakukan percobaan pada 10 gambar berbeda untuk pesan rahasia yang sama dan kata kunci yang sama, contoh pesan yang digunakan adalah logo unud dan kata kunci "unud".

Untuk hasil uji kualitas atau PSNRnya dapat dilihat pada tabel 1 berikut :

Tabel 1. Hasil Pengujian PSNR

No	Citra Asli	Citra Stegano	PSNR (dB)
1			71.40

2			66.35
3			73.59
4			74.80
5			74.05
6			69.48

7			69.94
8			66.45
9			69.96
10			72.92

Rata-rata dari kesepuluh kali percobaan PSNR :

$$\begin{aligned}
 \text{Rata-rata} &= (\text{Total seluruh PSNR}) / \text{Jumlah percobaan PSNR} \\
 &= (71.40 + 66.35 + 73.59 + 74.80 + 74.05 + 69.48 + 69.94 + \\
 &\quad 66.45 + 69.96 + 72.92) \\
 &= 708.96 / 10 \\
 &= 70.89
 \end{aligned}$$

Dari hasil 10 kali uji coba diatas didapatkan nilai rata-rata PSNR masih berada diatas 40dB [7] yang berarti hasil uji untuk pesan dan kata kunci tersebut dengan gambar-gambar diatas dapat dikatakan kualitas stegano image yang dihasilkan termasuk tinggi.

4. Kesimpulan

4.1 Kesimpulan

- a. Metode steganografi BPCS (Bit-Plane Complexity Segmentation) yang dibangun pada aplikasi ini berhasil menyisipkan file gambar digital signature pada gambar cover image dengan format *.jpg dan dapat menyimpan stegoimage serta berhasil mengekstraksi stegoimage dan melihat file gambar digital signature terenkripsi yang tersimpan.
- b. Dari 10 kali pengujian kemiripan gambar dengan metode PSNR didapatkan bahwa rata-rata PSNR yang dihasilkan adalah 70.89 dB. Penggunaan metode DES dan BPCS termasuk baik karena rata-rata nilai PSNR yang dihasilkan masih diatas standar kategori baik yaitu diatas 40dB.[1]

4.2 Saran

Saran yang dapat diberikan penulis untuk penelitian berikutnya adalah sistem ini dapat dikembangkan dengan menggunakan media penampung selain citra atau gambar seperti media suara atau video juga dapat digunakan dalam mengimplementasikan metode steganografi.

Daftar Pustaka

- [1] Aprilia, I., Ariyanti, D. and Izzuddin, A. (2019) 'Analisa Pengukuran Kualitas Citra Hasil Steganografi', pp. 116–121.
- [2] Fatma, Y., Mukhtar, H. and Taufik, M. (2018) 'Implementasi Steganografi Pada Teks Terenkripsi Dengan Algoritma Rsa Menggunakan Metode Bpcs', *Jurnal Fasilkom*, 7(2), pp. 260–265. Available at: <https://doi.org/10.37859/jf.v7i2.783>.
- [3] Gonzalez, R.C. and Woods, R.E. (2018) *4TH EDITION Digital image processing*.
- [4] Jurnal, J. *et al.* (2023) 'Implementasi Keamanan Aset Informasi Steganografi Menggunakan Metode Least Significant Bit (LSB)', 3(1), pp. 40–46.
- [5] Raharja, B.D. and Harsadi, P. (2018) 'Implementasi Kompresi Citra Digital Dengan Mengatur Kualitas Citra Digital', *Jurnal Ilmiah SINUS*, 16(2), pp. 71–77. Available at: <https://doi.org/10.30646/sinus.v16i2.363>.
- [6] Rizqa, I., Safitri, A.N. and Harkespan, I. (2022) 'Kriptostegano Menggunakan Data Encryption Standard dan Least Significant Bit dalam Pengamanan Pesan Gambar', *Jurnal Masyarakat Informatika*, 13(2), pp. 111–120. Available at: <https://doi.org/10.14710/jmasif.13.2.44547>.
- [7] Sajati, H. (2018) 'The Effect of Peak Signal to Noise Ratio (PSNR) Values on Object Detection Accuracy in Viola Jones Method', *Conference SENATIK STT Adisutjipto Yogyakarta*, 4. Available at: <https://doi.org/10.28989/senatik.v4i0.139>.
- [8] Z, H. (2018) 'Implementasi Metode Bit Plane Complexity Segmentation pada Citra Digital dalam Penyembunyian Pesan Rahasia', *MEANS (Media Informasi Analisa dan Sistem)*, 3(2), pp. 159–165. Available at: <https://doi.org/10.54367/means.v3i2.286>.

Penerapan Metode *Content Based Filtering* Dan *K-Nearest Neighbor* Dalam Sistem Rekomendasi Musik

I Made Teja Sarmandana^{a1}, I Made Widiartha^{a2}, Luh Arida Ayu Rahning Putri^{a3}, I Gede Santi Astawa^{a4}

^aInformatika, Universitas Udayana
Bali, Indonesia
¹tejamade01@gmail.com
²madewidiartha@unud.ac.id
³rahningputri@unud.ac.id
⁴santi.astawa@unud.ac.id

Abstract

Current technological developments are able to change the way the younger generation enjoys music, where music can now be packaged in digital form, which is a new innovation in the music industry in Indonesia. Given the large amount of music data available on the internet, a system that provides services for users to search for their favorite music is really needed. The recommendation system will provide relevant information based on the preferences that the user wants to search for. Content Based Filtering recommends that users utilize the information contained in the data to use as parameters. The K-Nearest Neighbor (K-NN) algorithm is a method of classifying objects based on the closest training data to the object under test. In this study, accuracy testing techniques were used to measure the performance of the classification that has been carried out. The classification process that was created succeeded in obtaining the highest accuracy value at 90,49% with a value of k=9 which shows that the classification and recommendation process can run quite well.

Keywords: *Accuracy, Content Based Filtering, Classification, Recommendation System, K-Nearest Neighbor, Music*

1. Pendahuluan

Masyarakat di penjuru dunia sangat dekat dengan yang namanya musik. Musik mampu menjadi peneman kita dalam melaksanakan aktivitas sehari-hari serta menjadi media yang mampu membawa kita ke perasaan senang, sedih, haru, dan beragam emosi lainnya. Teknologi saat ini berkembang hingga mampu mengubah cara generasi muda dalam menikmati musik dimana saat ini musik sudah mampu dikemas dalam bentuk digital yang menjadi inovasi baru di dalam industri musik di Indonesia [1]. Masyarakat saat ini dapat mendengarkan musik dimanapun kapanpun secara *online* karena telah tersedia banyak media layanan *streaming* musik secara *online*. Indonesia dipandang mampu menjadi pasar yang potensial dengan 220 juta lebih penduduknya. Hal tersebut dibuktikan dengan survey yang menunjukkan sekitar 46,9 juta atau 35,5% dari populasi pengguna internet di Indonesia menikmati musik secara *online* [2]. Ada pula survey yang dilakukan oleh *DailySocial* kepada para responden yang menunjukkan 85 persen responden menyatakan bahwa mereka sering menikmati musik secara *online* dalam waktu kurang lebih 6 bulan terakhir dimana sebagian besar pengguna layanan *streaming* musik tersebut berusia 16 sampai 24 tahun yang termasuk dalam kategori milenial atau "Generasi Z" [3]. Hal tersebut mendorong aplikasi-aplikasi penyedia layanan *streaming* musik seperti *Spotify*, *Joox*, *Youtube Music*, dan masih banyak lagi agar meningkatkan *user experience* untuk menarik *user-user* baru supaya ikut menggunakan aplikasi mereka serta menambah nilai kenyamanan dalam menggunakan aplikasi tersebut.

Mengingat banyaknya data musik yang tersedia di Internet, sistem yang menyediakan layanan bagi *user* untuk mencari musik favoritnya sangat dibutuhkan. Sistem rekomendasi menjadi jawaban untuk menghindari masalah kelebihan informasi tersebut [4]. Sistem rekomendasi merupakan sistem yang memperkirakan beberapa item atau data yang nantinya akan direkomendasikan dari yang teratas. Sistem rekomendasi akan menyajikan informasi yang sesuai dengan preferensi yang dicari oleh pengguna[5]. Sistem rekomendasi dirasa akan berguna menjadi pemecahan masalah dalam penelitian ini yang dimana menggunakan metode *Content-Based Filtering*. *Content-based filtering* merekomendasikan pengguna dengan memanfaatkan informasi yang terdapat pada data untuk

digunakan sebagai parameternya. Metode ini memilih dan memberi peringkat item berdasarkan kesamaan atribut itemnya. Keuntungan dari metode ini adalah bahwa informasi dalam setiap bagian data diketahui dalam representasi, sehingga pengguna bisa mendapatkan informasi tentang informasi yang dianggap relevan baginya [6]. Metode ini dijalankan dengan menggunakan algoritma *K-Nearest Neighbor* (K-NN).

Algoritma K-NN merupakan suatu teknik yang digunakan untuk mengklasifikasikan objek berdasarkan data pembelajaran yang memiliki jarak paling dekat dengan objek yang sedang diuji. Algoritma ini akan mengelompokkan hasil perhitungan dengan data pembelajaran yang memiliki kesamaan tertinggi dalam jarak yang telah ditentukan. Jarak antara data uji dan data latih diukur menggunakan persamaan *Euclidean* [7]. K-NN adalah salah satu algoritma yang mudah dan efektif untuk digunakan pada data besar [8]. Selain itu algoritma K-NN juga merupakan pendekatan yang paling mudah digunakan untuk menghitung jarak. Dalam menguji kinerja penerapan algoritma K-NN untuk sistem rekomendasi, kita dapat mengevaluasi sejauh mana keberhasilan algoritma ini dalam memberikan rekomendasi yang sesuai kepada pengguna [9].

Pada penelitian ini sistem rekomendasi musik dibuat menggunakan metode *Content-Based Filtering* untuk merekomendasikan musik berdasarkan dengan preferensi pengguna dengan menerapkan algoritma *K-Nearest Neighbor* untuk membandingkan dan mencari nilai kedekatan dari musik yang disukai pengguna dengan musik yang ada pada dataset. Dengan dibuatnya sistem rekomendasi ini diharapkan masyarakat di luar sana khususnya kelompok milenial dan “Generasi Z” mendapatkan musik yang sesuai dengan preferensi mereka dengan cepat.

2. Metode Penelitian

Adapun metode penelitian yang dilakukan pada penelitian ini yaitu:

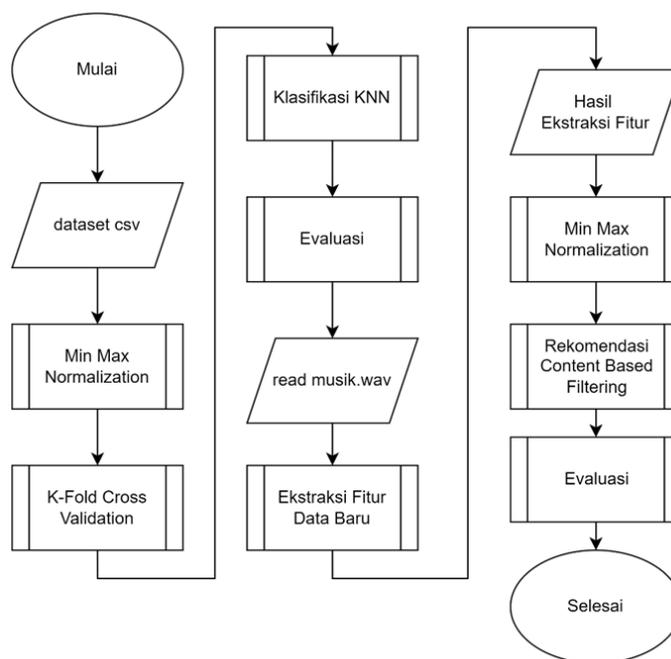
a. Pengumpulan Data

Data yang digunakan adalah data sekunder yang diambil dari *website kaggle*. Terdapat dua dataset yang akan digunakan yaitu dataset yang terdapat 4 folder yang menyimpan file musik yang telah dikelompokkan berdasarkan genrenya yaitu *blues, classical, metal, dan pop*, serta terdapat file CSV yang menyimpan hasil ekstraksi fitur dari seluruh file musik pada dataset ini. Total data musik pada dataset ini adalah sebanyak 400 data musik dengan format *.wav* yang di setiap genrenya terdiri dari 100 data musik dengan frekuensi 22050Hz yang berdurasi 30 detik pada setiap file audionya. Dataset ini akan digunakan untuk membangun model klasifikasi K-NN.

Satu dataset lagi merupakan folder yang menyimpan file musik preferensi user untuk digunakan dalam proses pencarian rekomendasi musik. Total data musik pada dataset ini adalah sebanyak 80 data musik dengan format *.wav* yang dengan genre yang sama seperti dataset sebelumnya namun pada dataset ini, tiap genrenya terdiri dari 20 data musik dengan frekuensi yang sama dan durasi yang sama dengan dataset sebelumnya. Data musik preferensi user ini yang nantinya akan diekstrak fiturnya untuk mendapatkan fitur *Zero Crossing Rate, Harmonic Percussive* dan MFCC yang kemudian akan dijadikan sebagai musik acuan yang akan dicari rekomendasinya.

b. Desain Penelitian

Metode *Content-Based Filtering* adalah sebuah teknik dalam sistem rekomendasi yang dapat menghasilkan rekomendasi dari preferensi *user* berdasarkan fitur yang terkait dengan item lain yang dibandingkan. Pada penelitian ini preferensi *user* adalah musik yang pernah didengarkan oleh user dan telah melewati proses ekstraksi fitur. Proses pencarian rekomendasinya dibantu dengan K-NN untuk menentukan genre dari data musik preferensi user. Langkah-langkah dalam penelitian yaitu diawali dengan proses pembuatan model K-NN sampai mendapatkan parameter k terbaik lalu dilanjutkan dengan proses pencarian rekomendasi musik menggunakan parameter k terbaik tersebut. *Flowchart* desain penelitian dapat dilihat pada Gambar 1.



Gambar 1. Desain Penelitian

c. *Min Max Normalization*

Tahap ini merupakan proses mengubah nilai-nilai fitur dataset ke rentang nilai antara 0 - 1. Tujuan utamanya adalah untuk memastikan bahwa semua fitur memiliki skala yang seragam, sehingga tidak ada fitur yang mendominasi yang lain dalam proses analisis atau pemodelan. Min Max Normalization dapat dilakukan dengan menerapkan rumus yang dapat dilihat pada Persamaan 1 [10].

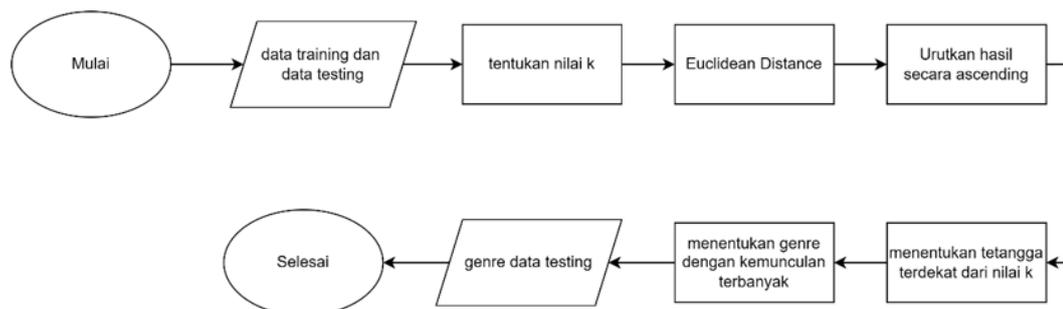
$$normalized(x) = \frac{minRange + (x - minValue)(maxRange - minRange)}{maxValue - minValue} \quad (1)$$

d. *K-Fold Cross Validation*

Proses dari teknik *K-Fold Cross Validation* dilakukan untuk memisahkan data menjadi sama rata tergantung nilai fold yang dimasukkan. Teknik *K-Fold Cross Validation* memiliki tujuan agar setiap data memiliki kesempatan untuk menjadi *data training* dan *data testing*.

e. *Klasifikasi K-NN*

Proses klasifikasi akan dilakukan terhadap dataset yang telah di split menggunakan *K-Fold Cross Validation*. Parameter k untuk menentukan jumlah tetangga terdekat dari *data testing* terhadap *data training* yang telah dihitung jaraknya dengan menggunakan metrik *Euclidean Distance*. K-NN akan menentukan label atau genre dari *data testing* dengan mengambil genre dengan kemunculan terbanyak pada jumlah tetangga terdekat yang telah ditetapkan. Proses klasifikasi memiliki tujuan untuk mendapatkan nilai k yang terbaik yang digunakan sebagai parameter untuk memprediksi genre dari data baru sekaligus menentukan rekomendasinya. Flowchart dari proses klasifikasi menggunakan K-NN dapat dilihat pada Gambar 2.



Gambar 2. Flowchart K-NN

f. Evaluasi

Tahap evaluasi merupakan tahap penting dimana kita bisa menilai kinerja dari algoritma dan metode yang digunakan untuk merekomendasikan musik ini. Tahap evaluasi ini dilakukan untuk menguji parameter k yang berbeda agar mendapatkan nilai k dengan *accuracy* tertinggi menggunakan data yang telah di-split menggunakan *K-Fold Cross Validation*. Pada penelitian ini menggunakan teknik evaluasi *accuracy* yang dapat dilihat pada Persamaan 2 [11].

$$Accuracy = \frac{Total\ Prediksi\ yang\ Benar}{Total\ Sampel} \quad (2)$$

Pada penelitian ini terdapat dua proses evaluasi yaitu evaluasi *accuracy* untuk mendapatkan nilai k terbaik dan evaluasi *accuracy* pada proses klasifikasi menggunakan data preferensi *user*.

1. Evaluasi *Accuracy* untuk Mendapatkan Nilai k Terbaik

Nilai k yang digunakan dapat ditentukan secara bebas oleh peneliti [12]. Dalam tahap pengujian ini, nilai k yang digunakan dimulai dari k=1 sampai dengan akar dari jumlah *data training* karena pembatasan ini diharapkan mampu memperoleh akurasi maksimal dengan waktu komputasi yang lebih singkat karena beberapa penelitian telah membuktikan bahwa nilai k yang terlalu tinggi juga cenderung menghasilkan *accuracy* yang semakin rendah [13].

Nilai k yang akan digunakan adalah bilangan ganjil untuk menghindari dua genre berbeda yang memiliki jumlah kemunculan yang sama pada tetangga terdekat. Oleh karena itu, skenario pengujian *accuracy* akan dilakukan dengan data yang telah dibagi menjadi 5 fold yang mana *data training* pada penelitian ini sebanyak 320 data training, sehingga nilai k yang dipilih adalah bilangan ganjil dalam rentang 1 hingga 17 [13].

Pada penelitian ini menggunakan teknik pengujian *accuracy* yang dapat dilihat pada persamaan (2). *Accuracy* digunakan untuk mengukur seberapa akurat model klasifikasi dalam mengklasifikasikan sampel untuk semua kelas [11].

2. Evaluasi *Accuracy* pada Data Preferensi User

Evaluasi ini dilakukan dengan tujuan untuk menilai seberapa efektif sistem tersebut dalam menentukan genre dari data preferensi *user*. Data preferensi *user* akan dinormalisasi menggunakan *Min Max Normalization* kemudian akan diklasifikasikan dengan K-NN. Nilai yang digunakan adalah nilai k yang telah ditetapkan sebagai nilai k terbaik berdasarkan evaluasi penentuan nilai k terbaik. Hasil dari K-NN tersebut akan dicari akurasi dan akurasi tersebutlah yang akan dijadikan tingkat keberhasilan sistem dalam menentukan genre dari data preferensi *user*.

g. Ekstraksi Fitur

Pada langkah ini dilakukan ekstraksi fitur pada musik-musik yang pernah didengarkan oleh *user* seperti yang dijelaskan pada data penelitian untuk menjalankan proses pencarian rekomendasi musik. Fitur-fitur yang akan diekstrak sesuai dengan fitur-fitur yang terdapat pada dataset musik yang digunakan yaitu *zero crossing rate*, *harmony percussive source separation* serta MFCC. Panjang vektor MFCC yang akan diekstrak adalah sepanjang 13 vektor fitur. Setelah seluruh fitur berhasil diekstrak, maka selanjutnya akan dibuat satu DataFrame yang menampung hasil dari ekstraksi fitur tersebut.

h. Rekomendasi *Content Based Filtering*

Setelah *data file* audio musik sebelumnya telah didapatkan fiturnya dan menjadi data vektor, data tersebut akan dicari kedekatannya dengan kelas genre terdekat diantara genre *Blues*, *Classical*, *Metal* dan *Pop* dengan menentukan nilai k. Nilai k dapat mempengaruhi tingkat akurasi perhitungan sehingga diperlukan perhitungan untuk menemukan nilai optimal k. *Data input* akan dihitung jaraknya dengan *data training* dan akan di sortir mulai dari yang terdekat sampai yang terjauh. Setelah diketahui jarak antara *data input* dan *data training* maka selanjutnya menentukan nilai k yang digunakan untuk menentukan kelas dari *data input*. Penentuan nilai k dilakukan dengan menggunakan beberapa nilai k yang nantinya akan dioptimasi dengan memperhatikan nilai k mana yang memiliki jarak terdekat. Setelah mengetahui kelas label genre maka rekomendasi musik diambil dari data-data yang termasuk di dalam kelas label genre tersebut dengan mengambil 5 data musik dengan jarak terdekat akan dijadikan output dengan menggunakan *Euclidean Distance* yang akan diurutkan mulai dari yang terdekat.

3. Hasil dan Diskusi

Implementasi sistem dilakukan dengan menggunakan bahasa pemrograman *python* dan beberapa *library* yang didukung oleh *python*. *Code* yang dibuat dijalankan di *google colab*.

3.1. Data Penelitian

Data yang digunakan untuk membangun model klasifikasi K-NN adalah file CSV yang menyimpan hasil ekstraksi fitur dari data-data musik pada dataset ini. Total data musik pada dataset ini adalah sebanyak 400 data musik yang terdiri dari 4 genre yaitu blues, classical, metal, dan pop dan di setiap genrenya terdiri dari 100 data musik.

Data musik preferensi user digunakan sebagai data input dalam proses pencarian rekomendasi musik. Total data musik pada dataset ini adalah sebanyak 80 data musik dengan format *.wav* yang dengan genre yang sama seperti dataset sebelumnya namun pada dataset ini, tiap genrenya terdiri dari 20 data musik. Data musik preferensi user ini yang nantinya akan diekstrak fiturnya untuk mendapatkan fitur Zero Crossing Rate, Harmonic Percussive dan MFCC yang kemudian akan dijadikan sebagai musik acuan yang akan dicari rekomendasinya.

Output yang akan dimunculkan adalah 5 data musik terdekat akan dijadikan sebagai rekomendasi musik. Tingkat keberhasilan rekomendasi dapat dilihat dari evaluasi model klasifikasi K-NN menggunakan nilai k terbaiknya dalam memprediksi genre dari data musik preferensi user.

3.2. Evaluasi

Tahap evaluasi ini dilakukan untuk menguji parameter k yang berbeda agar mendapatkan nilai k dengan *accuracy* tertinggi menggunakan data yang telah di-split menggunakan *K-Fold Cross Validation*.

a. Evaluasi *accuracy* untuk mendapatkan nilai k terbaik

Hasil pengujian *accuracy* dari kelima *fold* tersebut kemudian akan dirata-ratakan untuk mendapatkan *accuracy* nilai k secara keseluruhan. *Accuracy* tersebut disimpan ke dalam *DataFrame*. Hasil perhitungan *accuracy* dapat dilihat pada Tabel 1.

Tabel 1. *Accuracy* K-NN dengan beberapa Nilai k

Nilai k	<i>Accuracy</i>
1	0,9009
3	0,9018
5	0,9032
7	0,9043
9	0,9049
11	0,9035
13	0,9022
15	0,9005
17	0,8984

Berdasarkan hasil yang ditampilkan Tabel 1, didapatkan nilai *accuracy* dari yang terendah hingga yang tertinggi dengan rentang *accuracy* 0,8984 atau 89,84% sampai 0,9049 atau 90,49%. *Accuracy* tertinggi pada klasifikasi K-NN diperoleh pada nilai k=9 dengan *accuracy* berada di angka 0,9049 atau 90,49% sedangkan *accuracy* terendahnya diperoleh pada nilai k=17 dengan *accuracy* berada di angka 0,8984 atau 89,84%. Oleh karena itu, pada proses pencarian rekomendasi akan menggunakan nilai k=9 untuk menentukan genre dari data musik preferensi *user*.

b. Evaluasi pada Data Preferensi *User*

Evaluasi ini dilakukan untuk mengetahui keberhasilan sistem dalam menentukan genre dari data preferensi *user*. Data preferensi *user* akan dinormalisasi menggunakan *Min Max Normalization* kemudian akan diklasifikasikan dengan K-NN dan terakhir akan dilakukan perhitungan *accuracy*. Hasil evaluasi pada data preferensi *user* dapat dilihat pada Tabel 2.

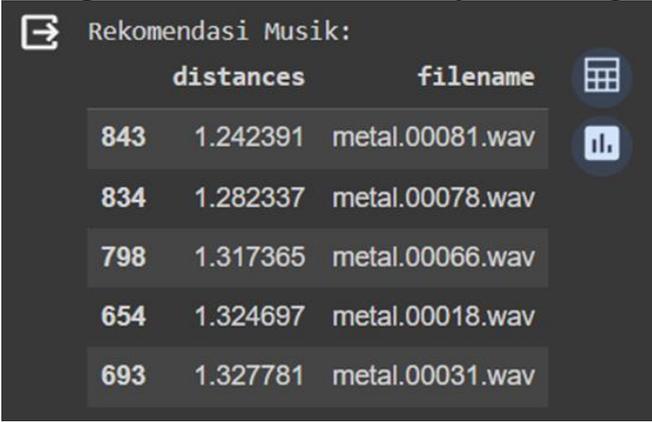
Tabel 2. Hasil *accuracy* K-NN dengan beberapa Nilai k

Nilai k	<i>Accuracy</i>
9	0,9018

Berdasarkan hasil yang ditampilkan Tabel 2, didapatkan tingkat keberhasilan sistem dalam menentukan genre pada data preferensi *user* dengan nilai *accuracy* sebesar 0,9018 atau 90,18%.

3.3. Hasil Rekomendasi Musik

Berdasarkan hasil klasifikasi sebelumnya, percobaan pada data baru telah mendapatkan hasil bahwa data tersebut termasuk ke dalam label *genre* 'metal' maka data yang memiliki label 'metal' akan dipertahankan dan 5 data dengan label 'metal' teratas akan dijadikan sebagai rekomendasi musik.



Rekomendasi Musik:

	distances	filename
843	1.242391	metal.00081.wav
834	1.282337	metal.00078.wav
798	1.317365	metal.00066.wav
654	1.324697	metal.00018.wav
693	1.327781	metal.00031.wav

Gambar 3. Hasil Rekomendasi Musik

4. Kesimpulan

Penelitian ini telah mampu melakukan rekomendasi dengan menerapkan metode *Content-Based Filtering* dan *K-Nearest Neighbor*. Kesimpulan pada penelitian ini yaitu:

1. Hasil evaluasi pada klasifikasi yang dijalankan dengan menggunakan K-NN telah berhasil mendapatkan nilai *accuracy* tertinggi berada pada nilai $k=9$ dengan *accuracy* sebesar 0,9049 atau 90,49% sedangkan *accuracy* terendahnya diperoleh pada nilai $k=17$ dengan *accuracy* berada di angka 0,8984 atau 89,84%.
2. Hasil evaluasi pada data preferensi *user* yang dijalankan menggunakan K-NN dengan menggunakan nilai k terbaik berdasarkan evaluasi sebelumnya yaitu $k=9$ berhasil mendapatkan *accuracy* sebesar 0,9018 atau 90,18%.

References

- [1] D. Noviani, R. Pratiwi, S. Silvianadewi, M. B. Alexandri, And M. A. Hakim, "Pengaruh Streaming Musik Terhadap Industri Musik Di Indonesia," 2020.
- [2] E. P. Utomo And N. Probosini, "Pengaruh Sikap Dan Norma Subjektif Terhadap Intensitas Penggunaan Aplikasi Streaming Pada Generasi 'Z,'" *Jurnal Ilmu Sosial Dan Humaniora*, Vol. 9, No. 2, P. 241, Sep. 2020, Doi: 10.23887/Jish-Undiksha.V9i2.18581.
- [3] A. , R. Abdurokhim & Hidayat, "Pengaruh Sikap Dan Norma Subjektif Pada Intensitas Generasi 'Z' Menggunakan Aplikasi Media Streaming," *Jurnal Ilmiah Indonesia*, Januari, Vol. 2020, No. 1, Pp. 9–22, 2020, [Online]. Available: [Http://Cerdika.Publikasiindonesia.Id/Index.Php/Cerdika/Indexhttp://Cerdika.Publikasiindonesia.Id/-9-](http://Cerdika.Publikasiindonesia.Id/Index.Php/Cerdika/Indexhttp://Cerdika.Publikasiindonesia.Id/-9-)
- [4] H.-C. Chen And A. L. P. Chen, "A Music Recommendation System Based On Music And User Grouping *," 2005.
- [5] N. Ula, C. Setianingsih, And R. A. Nugrahaeni, "Sistem Rekomendasi Lagu Dengan Metode Content-Based Filtering Berbasis Website Web-Based Song Recommendation System Using Content-Based Filtering," 2021.
- [6] A. I. Putra And R. R. Santika, "Edumatic: Jurnal Pendidikan Informatika Implementasi Machine Learning Dalam Penentuan Rekomendasi Musik Dengan Metode Content-Based Filtering," Vol. 4, No. 1, 2020, Doi: 10.29408/Edumatic.V4i1.2162.
- [7] M. M. Baharuddin, H. Azis, And T. Hasanuddin, "Analisis Performa Metode K-Nearest Neighbor Untuk Identifikasi Jenis Kaca," *Ilkom Jurnal Ilmiah*, Vol. 11, No. 3, Pp. 269–274, Dec. 2019, Doi: 10.33096/Ilkom.V11i3.489.269-274.

- [8] H. A. Sheeren, L. W. Santoso, And J. Andjarwirawan, "Sistem Rekomendasi Pembelian Laptop Dengan K-Nearest Neighbor (Knn)," 2019.
- [9] M. Teknologi Informasi, P. Pascasarjana, And U. Teknologi Yogyakarta, "This Work Is Licensed Under A Creative Commons Attribution-Sharealike 4.0 International License Sistem Rekomendasi Pada Tokopedia Menggunakan Algoritma K-Nearest Neighbor," 2022, Doi: 10.31294/Jtk.V4i2.
- [10] D. Azzahra Nasution, H. H. Khotimah, And N. Chamidah, "Perbandingan Normalisasi Data Untuk Klasifikasi Wine Menggunakan Algoritma K-Nn," 2019.
- [11] M. Rangga, A. Nasution, And M. Hayaty, "Perbandingan Akurasi Dan Waktu Proses Algoritma K-Nn Dan Svm Dalam Analisis Sentimen Twitter," *Jurnal Informatika*, Vol. 6, No. 2, Pp. 212–218, 2019, [Online]. Available: [Http://Ejournal.Bsi.Ac.Id/Ejournal/Index.Php/Ji](http://ejournal.bsi.ac.id/ejournal/index.php/ji)
- [12] R. L. Hasanah *Et Al.*, "Klasifikasi Penerima Dana Bantuan Desa Menggunakan Metode Knn (K-Nearest Neighbor)," *Jurnal Techno Nusa Mandiri*, Vol. 16, No. 1, P. 1, 2019, [Online]. Available: [Http://Nusamandiri.Ac.Id/](http://nusamandiri.ac.id/)
- [13] A. B. Hassanat, M. A. Abbadi, G. A. Altarawneh, And A. A. Alhasanat, "Solving The Problem Of The K Parameter In The Knn Classifier Using An Ensemble Learning Approach," 2014. [Online]. Available: [Http://Sites.Google.Com/Site/ljcsis/](http://sites.google.com/site/ljcsis/)

This page is intentionally left blank.

Pengolahan Big Data Dengan *Sharding Database* Dan *Kappa Architecture* Untuk *Data Time-Series*

Kompiang Gede Sukadharma^{a1}, I Putu Gede Hendra Suputra^{a2}, Ida Ayu Gde
Suwiprabayanti Putra^{a3}, Luh Arida Ayu Rahning Putri^{a4}

^aProgram Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas
Udayana Kuta Selatan, Badung, Bali, Indonesia

¹kompiang.sukadharma@gmail.com

²hendra.suputra@unud.ac.id

³iagsuwiprabayantiputra@unud.ac.id

⁴rahningputri@unud.ac.id

Abstract

In the digital era, managing and processing Big Data presents challenges. Historical or time-series data is crucial for decision-making. Previous research has indicated that databases on a server become increasingly irrelevant over time for handling continuous and ever-growing data. Distributed database systems, specifically database sharding, efficiently distribute CPU load and memory usage. Kappa Architecture outperforms Lambda Architecture by 220% in terms of speed, though Lambda has 9% higher reliability rate than Kappa Architecture. This research integrates database sharding with Kappa Architecture using time-series data obtained from the Kaggle platform. In testing results using a load-testing method. Load testing showed 60.46% performance for data retrieval across distributed databases. Remarkably, the system's reliability reached 100%, even when one of the services failed, handling 10000 new data entries. Furthermore, with the same configuration, it was found that the utilization of Kappa Architecture and sharding database resulted in a 70.26% better performance compared to the system solely implementing sharding database.

Keywords: *Kappa Architecture, Time-Series Data, Sharding Database, Big Data Processing, Reliability, Performance*

1. Pendahuluan

Setiap detik, lebih dari 30000 gigabytes data baru dihasilkan, dan laju pertumbuhan produksi data ini diperkirakan akan terus meningkat [1]. Karena jumlah data yang tercipta sangat besar, sistem yang mampu mengolah data dengan volume dan kecepatan seperti ini seringkali dianggap sebagai sistem pengolah *big data*. Pemanfaatan data yang melimpah ini sangat berharga untuk menggali informasi yang berguna untuk mengoptimalkan strategi pengambilan keputusan [2]. Terdapat beberapa jenis data yang tersedia, *data time-series* menjadi salah satu sumber informasi kunci untuk mendukung pengambilan keputusan. Di era saat ini, salah satu tantangan besar dalam penelitian *big data* adalah mengolah *data time-series* yang jumlahnya sangat banyak. Ini karena database relasional konvensional sudah tidak lagi memadai untuk menangani *data time-series*, akibat dari keterbatasan dalam kecepatan pemrosesan saat volume data bertambah, redundansi data, dan kurangnya dukungan terhadap operasi pada data bertipe waktu. Hal ini menyebabkan proses pengolahan data menjadi tidak efisien [3].

Untuk menangani *big data*, yang dicirikan oleh volume, kecepatan, dan keragaman data yang besar, dibutuhkan sistem basis data khusus yang mampu menyimpan dan mengelola data dalam skala besar. Pendekatan penyimpanan data terpusat dapat diadaptasi untuk memenuhi kebutuhan skalabilitas, keandalan, dan pengelolaan data yang efisien. Namun, menghadapi tantangan untuk menciptakan sistem basis data yang efektif biaya dan berkinerja tinggi, solusi yang ditemukan adalah penerapan basis data terdistribusi melalui teknik *sharding database*. *Sharding database* memungkinkan pemecahan dan pendistribusian *database* menjadi segmen-segmen lebih kecil ke berbagai *node database* dalam sebuah kluster, dimana setiap segmen beroperasi secara mandiri, menawarkan solusi skalabilitas dan efisiensi [4].

Dalam pemrosesan *big data*, dua pendekatan utama yang digunakan adalah pemrosesan *batch* dan pemrosesan *stream*. Pemrosesan *batch* dikenal dengan kinerjanya yang solid namun memiliki latensi yang tinggi, yang membuatnya kurang ideal untuk aplikasi data *real-time*. Di sisi lain, pemrosesan *stream* menawarkan latensi yang rendah dan ketersediaan yang tinggi tetapi bisa berdampak pada penurunan tingkat akurasi. Karenanya, untuk mencapai hasil yang optimal dalam berbagai skenario, seringkali diperlukan gabungan dari kedua metode tersebut agar dapat memperoleh proses yang cepat namun tetap detail [5]. Dua arsitektur utama yang digunakan dalam implementasi metode pemrosesan ini adalah *lambda architecture* dan *kappa architecture*. *Lambda architecture* mendukung pemrosesan secara *real-time* dan *batch*, tetapi menimbulkan tantangan dalam hal biaya dan kompleksitas sistem karena memerlukan pemeliharaan dua teknologi yang berbeda [6]. Sebaliknya, *kappa architecture*, yang lebih sesuai untuk menangani data besar yang perlu diproses secara cepat dan *real-time* seperti data dari sensor, robot, dan jaringan, menonjolkan pemrosesan melalui jalur yang sama untuk *real-time* dan *micro batching*. Keuntungan menggunakan *kappa architecture* adalah hanya membutuhkan satu teknologi pemrosesan *real-time* untuk menangani kedua jenis pemrosesan data tersebut [7].

Oleh karena itu, sangat penting untuk melakukan penelitian yang lebih mendalam. Penelitian ini akan berfokus pada pengembangan teknik dan struktur yang mendukung keberlangsungan dan optimalisasi ekstraksi informasi dari *data time-series*. Hasil dari penelitian ini diharapkan dapat menyediakan wawasan terkait efisiensi pemrosesan *big data* dengan data berjenis *time-series* menggunakan teknik *sharding database* dan *kappa architecture*, yang keduanya berpotensi meningkatkan kinerja dan keandalan sistem. Kinerja, dalam konteks ini, merujuk pada kemampuan sistem untuk merespons permintaan.

2. Metodologi Penelitian

Penelitian ini penulis bagi ke dalam lima tahap, yaitu identifikasi masalah, pengumpulan data, perancangan kebutuhan sistem, dan perancangan pengujian dan evaluasi sistem.

2.1. Identifikasi Masalah

Isu yang dibahas dalam penelitian ini muncul dari kebutuhan sebuah perusahaan pengembang perangkat lunak yang ingin menghasilkan laporan berbasis waktu dengan cepat dan biaya yang efisien. Perusahaan ini menghadapi kendala dalam mengakses data dari *database* berukuran besar. Terutama, *data time-series* yang menjadi fokus dalam penelitian ini. Seiring waktu, *data time-series* akan bertambah banyak yang menyebabkan peningkatan waktu akses data serta konsumsi memori dan CPU pada server [4]. Hal ini tidak hanya memperlambat proses pengambilan data tapi juga berdampak pada performa *service* lain yang bergantung pada *database* yang sama. Penyebab utamanya adalah sistem dari operasi *database* pada server yang harus mengiterasi lebih banyak data, karena bertambahnya volume data, mengakibatkan penurunan kinerja server.

Berdasarkan analisis masalah ini, yang didukung oleh referensi penelitian lain, penulis memilih untuk mengeksplorasi penggunaan metode *sharding database* bersamaan dengan penerapan arsitektur kappa. Tujuannya adalah untuk mengolah *big data*, khususnya *data time-series*, melalui stream processing. Pendekatan ini diharapkan dapat meningkatkan kecepatan operasi *database* besar dan secara simultan menurunkan biaya operasional *database* bagi perusahaan.

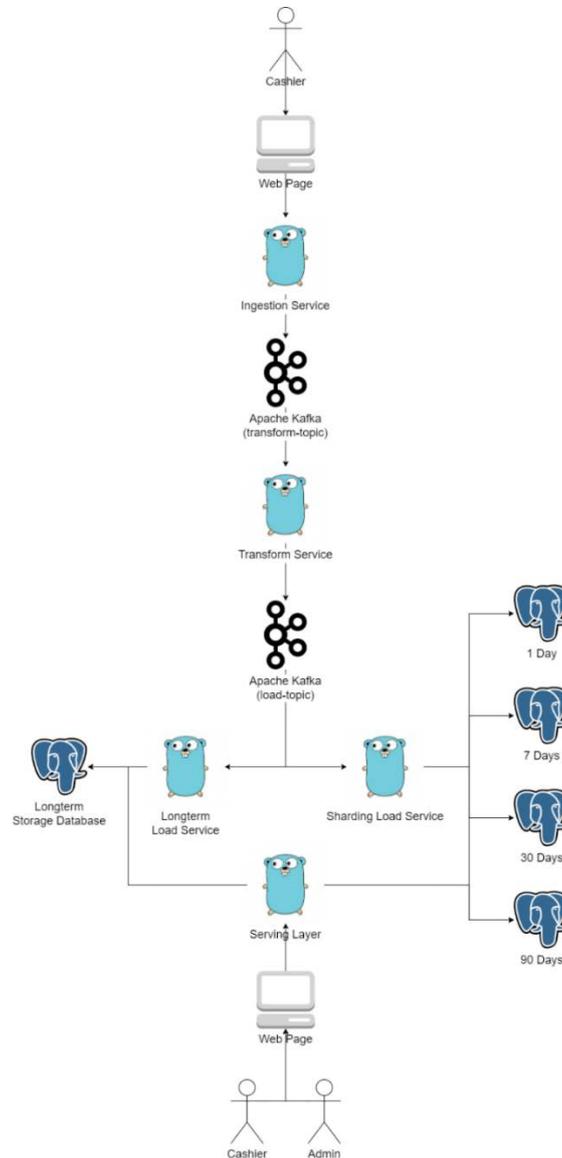
2.2. Pengumpulan Data

Penulis kemudian beralih ke tahap pengumpulan data untuk penelitian ini. Menggunakan *Dataset* dari Kaggle sebagai sumber data, penulis memilih pendekatan ini karena keterbatasan akses ke *data time-series* asli dari perusahaan yang bersifat rahasia dan tidak dapat dibagikan secara publik. *Dataset* yang dipilih mengandung sebuah tabel dengan 1 juta baris data yang berkaitan dengan *time-series*. Untuk lebih mendekati skenario *big data* yang sesungguhnya, data ini akan direplikasi sebanyak 10 kali, menggambarkan aspek volume yang besar, salah satu karakteristik penting dari *big data*. *Dataset* yang penulis peroleh berjudul "ECommerce Data Analysis" yang diunggah oleh M MOHAIMINUL ISLAM pada platform Kaggle. Data yang terdapat pada *dataset* ini adalah data dari *ecommerce* yang terletak pada negara Bangladesh. Oleh karena itu, data *time-series* yang penulis gunakan merupakan data order.

2.3. Perancangan Kebutuhann Sistem

2.3.1. Perancangan High-Level Kappa Architecture

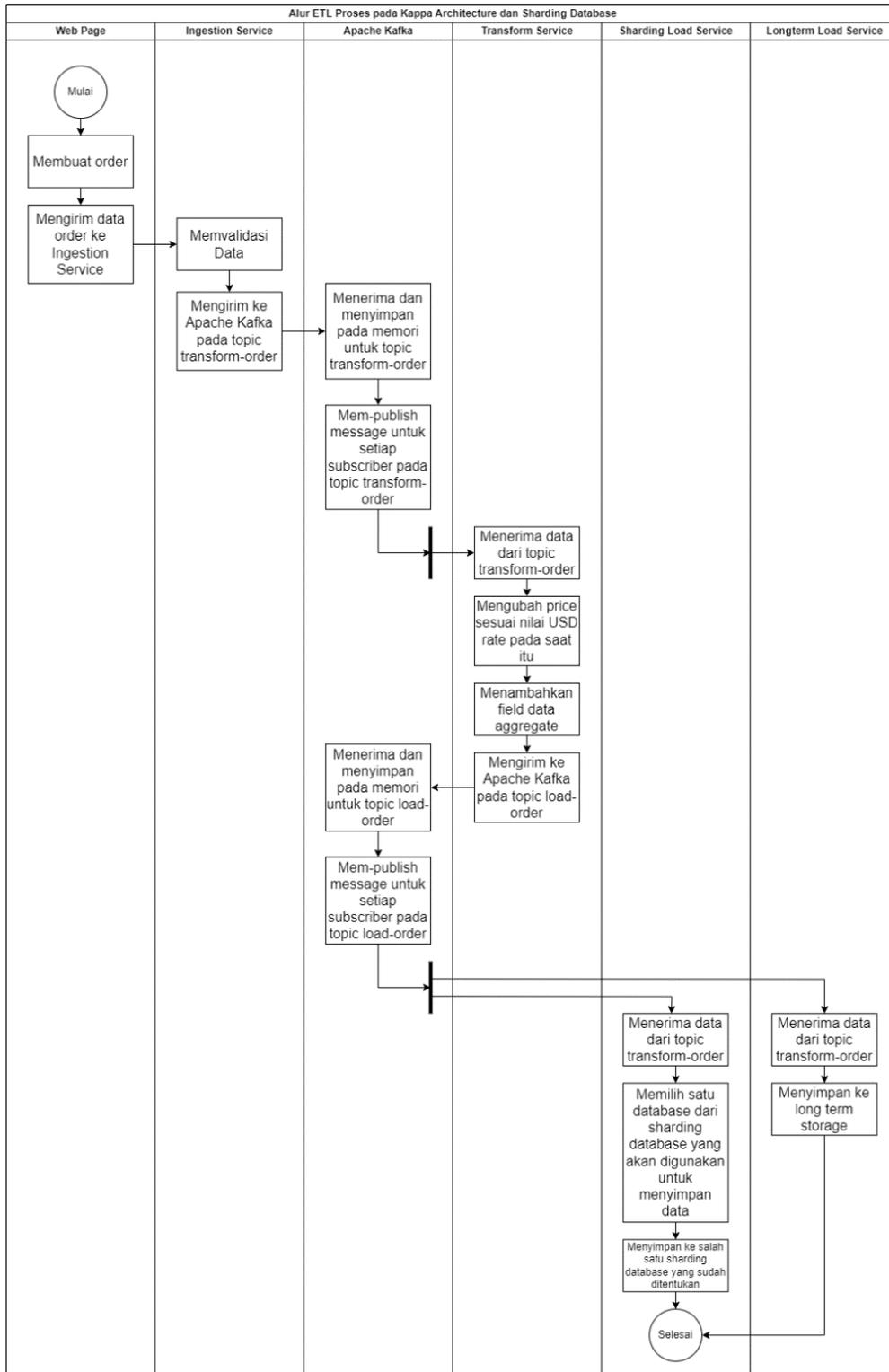
Perancangan *High-Level Arsitektur Kappa*, dapat dilihat pada Gambar 1. *Entrypoint* dari sistem ini terletak pada ingestion service, *service* ini akan mendapatkan data yang dikirim dari web page via API. Kemudian terdapat Apache Kafka yang memiliki peran sebagai *message broker* untuk setiap *publisher* dan *subscriber* yang ada pada sistem ini.



Gambar 1. High-Level Kappa Architecture

Kappa architecture merupakan sebuah pemrosesan *big data* dengan metode *stream processing* pada sebuah jalur dan juga implementasi *pipeline pattern*. Jadi, sistem ini merupakan sekumpulan proses yang akan mengirimkan data dari satu *service* ke *service* lainnya menggunakan *kappa architecture* untuk mengimplementasikan *event-driven microservices*. Pada *kappa architecture* yang penulis teliti juga terdapat Postgres yang menjadi *SQL database* dan menjadi *storage layer* dari arsitektur ini. Kemudian juga terdapat *serving layer* yang diimplementasikan menggunakan bahasa pemrograman Go dan menyediakan API yang dapat digunakan untuk mengambil data dari *database* yang tersedia.

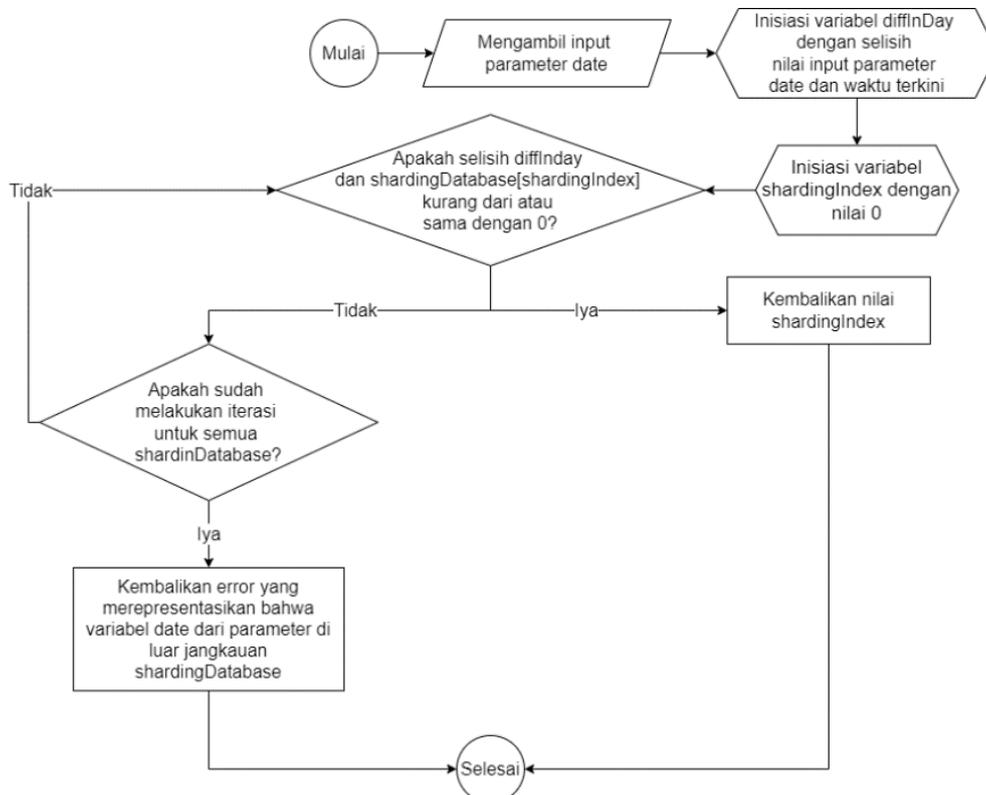
2.3.2. Perancangan Algoritma dan Alur Untuk Melakukan Penulisan ke Storage Layer



Gambar 2. Activity Diagram Alur Pembuatan Data Time-Series

Penulis menjelaskan alur proses pembuatan order dari data *order* yang masuk ke *ingestion service* hingga penyimpanan data pada *sharding load service* dan *longterm load service* dan ditunjukkan pada *activity diagram* pada Gambar 2. *Ingestion service* berperan dalam ekstraksi dan validasi data sebelum dipublikasikan ke Apache Kafka. *Transform service* menambahkan

nilai mata uang, melakukan konversi, dan melakukan agregasi data sebelum mempublikasikannya kembali ke Apache Kafka untuk disimpan pada *sharding load service* dan *longterm load service*. Proses *sharding load service* mencari *database* yang tepat untuk menyimpan data, sedangkan *longterm load service* langsung menyimpan data ke *database* tanpa proses tambahan. Algoritma yang dibuat untuk pemilihan *sharding database* mempertimbangkan rentang data retensi pada setiap *database*. Jika data tidak sesuai dengan rentang semua *sharding database*, data tersebut akan disimpan pada *longterm database*.

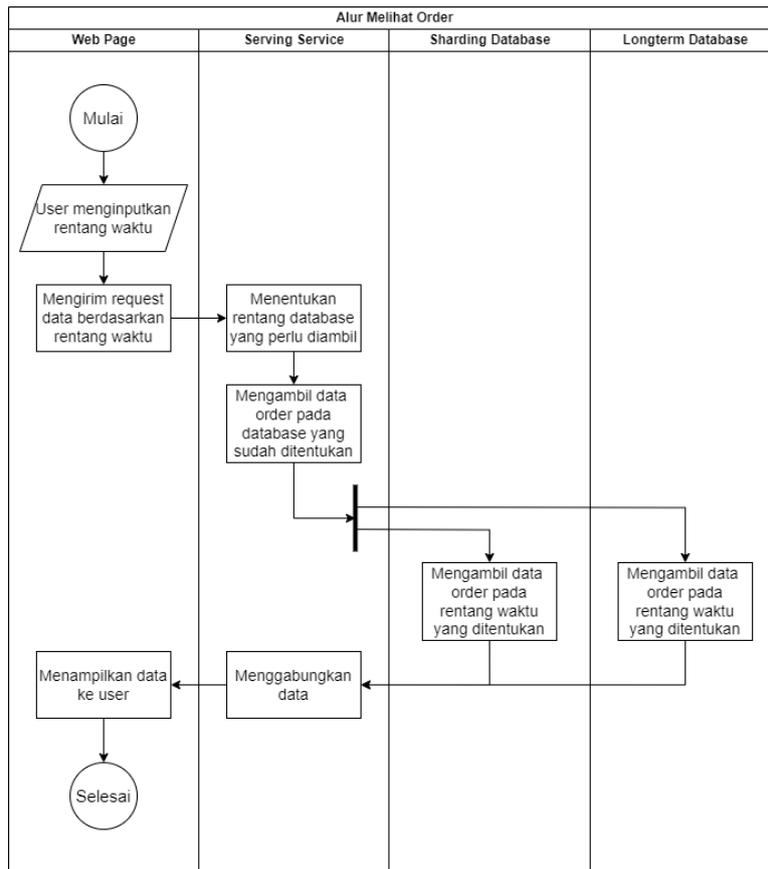


Gambar 3. Alur Algoritma Penentuan Sharding Database Untuk Sebuah Data Time-Series

Karena penelitian ini menggunakan *sharding database*, penulis perlu membuat sebuah algoritma yang akan menentukan *sharding database* indeks yang diperlukan. Algoritma pada Gambar 3 akan meminta sebuah *input parameter* yang merupakan sebuah data yang memiliki tipe data *date*. Setelah itu, nilai dari *parameter* tersebut akan dihitung selisih hari dengan waktu terkini. Kemudian, algoritma akan melakukan perulangan pada konfigurasi dari kumpulan *sharding database* yang merupakan sebuah *array*. Setiap perulangan, algoritma akan menghitung selisih data retensi dari *sharding database* tersebut dengan selisih nilai parameter dengan waktu terkini. Apabila selisih tersebut memiliki nilai 0 atau kurang dari 0, artinya data tersebut berada pada rentang *sharding database* tersebut. Setelah itu, algoritma akan mengembalikan nilai indeks dari konfigurasi *sharding database* yang akan digunakan untuk membuat koneksi ke *sharding database* tersebut. Namun, apabila sudah melakukan perulangan ke semua *sharding database* dan tidak ada selisih yang memiliki nilai 0 atau kurang dari 0, algoritma akan mengembalikan nilai *error* yang merepresentasikan kondisi data tersebut tidak berada dalam rentang dari semua *sharding database*.

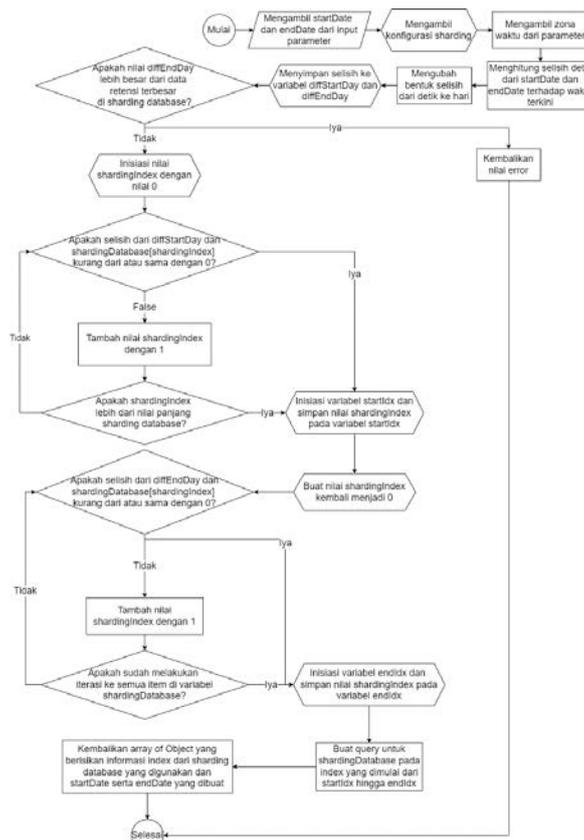
2.3.3. Perancangan Algoritma dan Alur Untuk Melakukan Pembacaan Dari Serving Layer

Alur dalam pembacaan data dapat dilihat pada Gambar 4. Secara umum, pengguna memasukkan rentang waktu pada *web page*, lalu request dikirim ke *servicing service*. *Servicing service* menggunakan sebuah algoritma untuk menentukan rentang *database*, kemudian mengambil data dari setiap rentang *database*, menggabungkannya, dan menampilkan hasilnya pada *web page*.



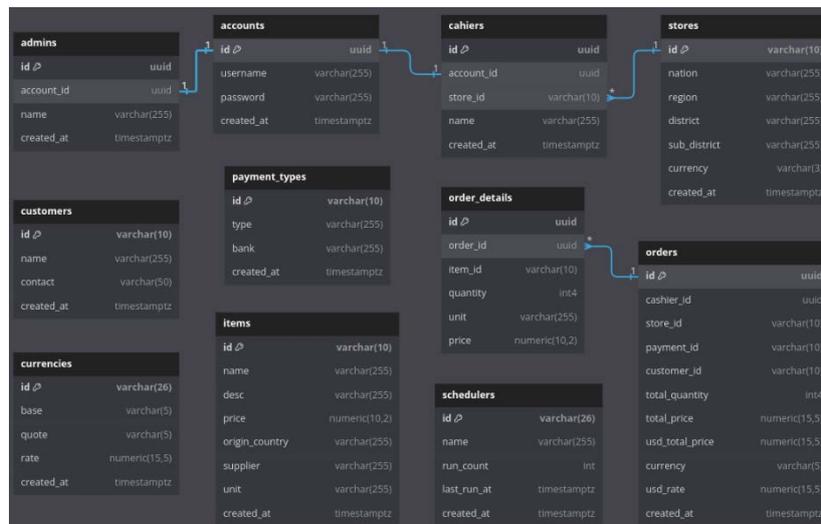
Gambar 4. Activity Diagram Untuk Membaca Data Dari Serving Layer

Perancangan alur dari algoritma untuk pembacaan data dari beberapa *sharding database* dapat dilihat pada Gambar 5. Dengan menggunakan algoritma tersebut, sistem dapat dengan efisien menentukan dan memilih *database* yang tepat untuk mengambil data yang diperlukan, sehingga meminimalkan beban dan meningkatkan kinerja secara keseluruhan. Algoritma tersebut beroperasi dengan langkah-langkah sebagai berikut, dimulai dengan validasi terhadap parameter rentang waktu yang dimasukkan oleh pengguna. Validasi ini penting untuk memastikan bahwa *query* yang akan dilakukan tidak mengakses semua *database* jika tidak diperlukan, sehingga menghindari beban yang tidak perlu. Setelah itu, algoritma memeriksa urutan tanggal awal dan akhir, dan jika tanggal akhir lebih awal dari tanggal awal, sistem akan memberikan pesan *error* untuk memperingatkan pengguna. Langkah selanjutnya adalah menyesuaikan zona waktu dari parameter *input* dengan konfigurasi sistem, sehingga *query* yang dihasilkan akan sesuai dengan zona waktu yang relevan. Algoritma kemudian menghitung perbedaan hari antara tanggal awal dan tanggal akhir, yang menjadi dasar untuk menentukan rentang *database* yang akan digunakan. Proses perulangan dilakukan untuk setiap *sharding database* yang terdaftar dalam konfigurasi sistem. Setiap iterasi perulangan mencari perbedaan antara data retensi dari *sharding database* dan selisih hari dari tanggal awal dan akhir dengan waktu saat ini. Jika perbedaannya kurang dari atau sama dengan nol, itu menandakan bahwa *sharding database* tersebut akan menjadi bagian dari rentang yang diperlukan. Proses ini diulangi untuk menentukan rentang akhir dari *sharding database* yang akan digunakan. Setelah rentang *sharding database* ditentukan, *query* yang sesuai dibuat untuk mengambil data dari rentang waktu yang ditentukan. Terakhir, algoritma mengembalikan *array* yang berisi informasi tentang indeks dari *sharding database* terpilih bersama dengan *query* yang dibuat, sehingga sistem dapat dengan mudah mengambil data yang diperlukan dari *database* yang sesuai dengan rentang waktu yang diminta. Dengan demikian, algoritma ini tidak hanya memastikan efisiensi dalam penggunaan sumber daya, tetapi juga memberikan fleksibilitas dan akurasi dalam pengambilan data yang sesuai dengan kebutuhan pengguna.



Gambar 5. Alur Algoritma Penentuan Rentang Sharding Database

2.3.4. Perancangan Skema Database



Gambar 6. Skema Database

Dalam penelitian yang dilakukan, sistem yang dirancang memanfaatkan tiga jenis *database* yang berbeda, yang masing-masing diadaptasi untuk fungsi tertentu. Pemisahan ini merupakan bagian dari penerapan metode *sharding database* dan arsitektur *microservices*, yang bertujuan untuk memaksimalkan efisiensi dan skalabilitas sistem. Oleh karena itu, perencanaan skema *database* menjadi langkah krusial dalam proses implementasi, di mana harus ditentukan bagaimana

masing-masing *database* akan diorganisir dan diintegrasikan dengan bagian lain dari sistem. Pendekatan ini memungkinkan untuk pemrosesan data yang lebih cepat dan lebih efisien, sekaligus memudahkan pengelolaan data yang berskala besar dan kompleks.

Pada gambar Gambar 6, tabel *admins*, *accounts*, *cashiers*, *stores*, *schedulers*, *payment_types*, *items*, *customers*, dan *currencies* akan berada pada *general database*. Hal ini dilakukan, karena tabel-tabel tersebut memiliki data yang bersifat non-transaksional. Untuk tabel *orders* dan *order_details* akan tersimpan pada tiap-tiap indeks pada *sharding database* dan juga tersimpan pada *longterm database*.

2.4. Perancangan Pengujian dan Evaluasi Sistem

Dalam fase evaluasi dari sistem, peneliti memutuskan untuk mengimplementasikan *load-testing* menggunakan *k6* untuk mengukur performa dan reliabilitas sistem yang dikembangkan. Proses pengujian ini dibagi menjadi dua bagian utama untuk menguji aspek-aspek kritis dari sistem,

1. Pengujian Performa Pengambilan Data *Time-Series*:
Bagian kedua berfokus pada efektivitas sistem dalam mengambil atau meng-query *data time-series* yang besar dari *database*. Ini bertujuan untuk menilai seberapa cepat sistem dapat mengakses dan menyajikan data yang diminta. Pengujian ini akan membandingkan sistem ketika tidak menggunakan *sharding database* dan menggunakan *sharding database*.
2. Pengujian Reliabilitas
Bagian pertama dari pengujian fokus pada kemampuan sistem untuk menangani data yang masuk ketika salah satu *stream processor service* dalam *kappa architecture* mati, pada penelitian ini, penulis mematikan *transform service*. Pengujian ini bertujuan untuk melihat berapa banyak data yang hilang.
3. Pengujian Performa Penulisan Data *Time-Series*:
Bagian ketiga berfokus pada efektivitas sistem dalam menulis data *time-series* yang memiliki kecepatan yang tinggi dan jumlah yang banyak. Pengujian ini bertujuan untuk menilai seberapa cepat sistem dalam menangani penulisan data.

Pengujian reliabilitas dan pengujian performa penulisan data *time-series* akan membandingkan sistem ketika menggunakan *kappa architecture* dan *sharding database* dengan sistem yang tidak menggunakan *kappa architecture*, tetapi menggunakan *sharding database*. Untuk sistem yang tidak menggunakan *kappa architecture* pada pengujian ini, setiap *service* akan saling berkomunikasi secara langsung menggunakan protokol HTTP dan tidak menggunakan Apache Kafka.

Hasil dari pengujian akan menjadi kunci untuk mengevaluasi performa keseluruhan dan keandalan dari sistem yang dibangun. Adapun beberapa metris yang digunakan untuk melakukan pengukuran dalam pengujian ini. Metris-metris yang digunakan dapat dilihat pada Tabel 1.

Tabel 1. Metris Pengujian

Nama Metris	Satuan Ukur	Deskripsi
<code>http_req_waiting</code>	milliseconds	Seberapa lama menunggu respon balik dari awal request. Dapat pula dikatakan dengan "time to first byte"
<code>http_req_failed</code>	persentase	Seberapa banyak failed rate atau rata - rata request yang gagal

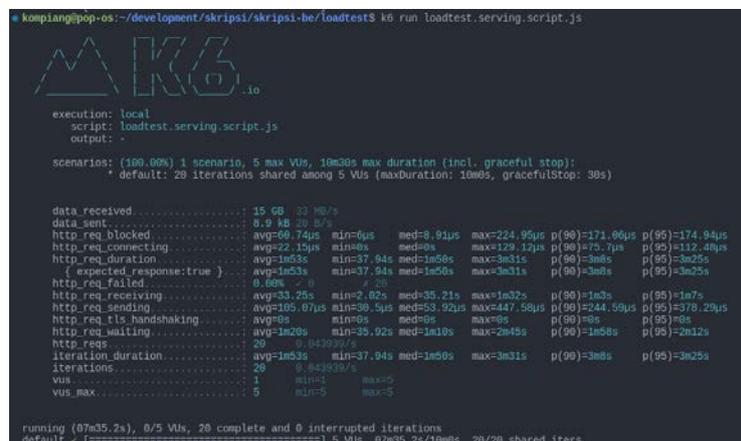
Dalam analisis performa sistem menggunakan *load-testing*, metris `http_req_waiting` menjadi kunci untuk memahami waktu yang dibutuhkan untuk menunggu respons dari server. Untuk memastikan analisis yang akurat dan menghindari distorsi yang mungkin diakibatkan oleh *data outlier*, persentil ke-95 digunakan dalam evaluasi. Pendekatan ini membantu dalam menyaring data yang tidak representatif dengan mengesampingkan 5% latensi teratas, yang dapat memberikan gambaran lebih jelas mengenai performa sistem dalam kondisi normal. Terdapat penelitian sebelumnya yang menggunakan persentil ke-99 untuk mendapatkan gambaran yang lebih dekat dengan kondisi ekstrem dimana latensi tertinggi menandakan potensi masalah dalam sistem [8].

Sementara itu, metris *http_req_failed* mengukur keandalan sistem dengan menghitung jumlah *request* yang gagal, yang ditandai dengan *status code* HTTP kurang dari 300. Metris ini memberikan *insight* tentang seberapa sering pengguna mengalami kegagalan saat berinteraksi dengan sistem, sehingga dapat menjadi indikator langsung dari kualitas *service* yang diberikan. Menggunakan kedua metris ini bersama-sama memberikan gambaran komprehensif tentang performa dan reliabilitas sistem, mulai dari seberapa cepat sistem dapat merespons permintaan hingga seberapa sering permintaan tersebut gagal.

3. Hasil dan Diskusi

3.1. Hasil Pengujian Performa Pengambilan Data Time-Series

Dalam skenario pengujian yang tidak menggunakan metode *sharding* untuk pengambilan data yang dapat dilihat pada Gambar 7, pengujian dilakukan dengan 5 pengguna virtual dan 20 iterasi, memakan waktu total 7 menit dan 35 detik. Dari hasil pengujian, tercatat pada metris *http_req_duration* nilai tertinggi yang dicapai adalah 3 menit 31 detik, dengan nilai untuk persentil ke-95 adalah 3 menit 25 detik. Untuk metris *http_req_waiting*, nilai tertinggi yang tercatat adalah 2 menit 45 detik, sedangkan untuk persentil ke-95, nilai yang dicapai adalah 2 menit 12 detik. Adapun metris *http_req_failed*, menunjukkan nilai 0%, yang berarti tidak ada kegagalan yang terjadi selama proses pengujian. Ini mengindikasikan bahwa meskipun tanpa *sharding*, sistem mampu menangani permintaan tanpa mengalami kegagalan, namun dengan waktu tunggu yang signifikan.



```
kompiang@pop-os: ~/development/skripsi/skripsi-be/loadtest$ k6 run loadtest.serving.script.js

      M K6
     .io

execution: local
script: loadtest.serving.script.js
output: -

scenarios: (100.00%) 1 scenario, 5 max VUs, 10m30s max duration (incl. graceful stop):
  default: 20 iterations shared among 5 VUs (maxDuration: 10m0s, gracefulStop: 30s)

data received .....: 15 GB 30 MB/s
data sent .....: 0.9 kB 20 B/s
http_req_blocked .....: avg=69.74µs min=6µs med=8.91µs max=224.95µs p(90)=171.06µs p(95)=174.94µs
http_req_connecting .....: avg=22.15µs min=0µs med=0µs max=129.12µs p(90)=75.7µs p(95)=112.40µs
http_req_duration .....: avg=1m53s min=37.94s med=1m56s max=3m31s p(90)=3m0s p(95)=3m25s
  ( expected response:true )
http_req_failed .....: 0.00% / 0 / 20
http_req_receiving .....: avg=33.25s min=2.62s med=35.21s max=1m32s p(90)=1m3s p(95)=1m7s
http_req_sending .....: avg=105.07µs min=39.5µs med=53.52µs max=447.58µs p(90)=244.59µs p(95)=378.29µs
http_req_tls_handshaking .....: avg=0µs min=0µs med=0µs max=0µs p(90)=0µs p(95)=0µs
http_req_waiting .....: avg=1m20s min=35.92s med=1m18s max=2m45s p(90)=1m58s p(95)=2m12s
http_reqs .....: 20 0.943039/s
iteration_duration .....: avg=1m53s min=37.94s med=1m56s max=3m31s p(90)=3m0s p(95)=3m25s
iterations .....: 20 0.943039/s
vus .....: 1 min=1 max=5
vus_max .....: 5 min=5 max=5

running (07m35.2s), 0/5 VUs, 20 complete and 0 interrupted iterations
default ✓ (=====) 5 VUs 07m35.2s/10m0s 20/28 shared iters
```

Gambar 7. Hasil Pengujian Performa Sistem Tanpa Menggunakan Metode *Sharding* Database

Hasil pengujian performa sistem menggunakan metode *sharding database* pada dilihat pada Gambar 8. Dengan penggunaan *sharding*, membutuhkan waktu total 3 menit 57 detik untuk 5 *virtual user* dan 20 iterasi. Metris *http_req_duration* menunjukkan nilai tertinggi 1 menit 52 detik dan pada persentil ke-95 adalah 1 menit 33 detik. Untuk *http_req_waiting*, nilai tertinggi adalah 1 menit 14 detik dan pada persentil ke-95 adalah 1 menit 8 detik. Sama seperti skenario pertama, metris *http_req_failed* tetap menunjukkan 0%. Perbandingan antara kedua skenario ini menunjukkan bahwa penggunaan *sharding database* secara signifikan meningkatkan kecepatan pengambilan data, dengan pengurangan waktu tunggu yang signifikan. Peningkatan kecepatan ini diperoleh dari pembagian *data time-series* ke berbagai *database*, memungkinkan proses pencarian dan pengambilan data menjadi lebih efisien karena setiap *database* hanya mengelola jumlah *row* yang lebih sedikit.

```

kompiang@pop-os:~/development/skripsi/skripsi-be/loadtest$ k6 run loadtest.serving.script.js

      M K6
      .io

execution: local
script: loadtest.serving.script.js
output: -

scenarios: (100.00%) 1 scenario, 5 max VUs, 10m30s max duration (incl. graceful stop):
  default: 20 iterations shared among 5 VUs (maxDuration: 10m0s, gracefulStop: 30s)

data_received ..... 15 GB 43 MB/s
data_sent ..... 8.0 KB 30 B/s
http_req_blocked ..... avg=152.36µs min=4.6µs med=10.95µs max=654.49µs p(90)=567.44µs p(95)=572.6µs
http_req_connecting ..... avg=34.69µs min=0µs med=0µs max=160.40µs p(90)=133.04µs p(95)=144.33µs
http_req_duration ..... avg=50.30ms min=25.49ms med=53.31ms max=1m22s p(90)=1m29s p(95)=1m33s
  { expected_response:true } ..... avg=50.90ms min=25.49ms med=53.31ms max=1m25s p(90)=1m29s p(95)=1m33s
http_req_failed ..... 0.00% / 0 / 20
http_req_receiving ..... avg=10.7µs min=573.29ms med=7.62µs max=98.22µs p(90)=24.88µs p(95)=31.75µs
http_req_sending ..... avg=240.80µs min=30µs med=55.25µs max=2.91ms p(90)=302.94µs p(95)=1.48ms
http_req_tls_handshaking ..... avg=0µs min=0µs med=0µs max=0µs p(90)=0µs p(95)=0µs
http_req_waiting ..... avg=40.27ms min=24.92ms med=44.95ms max=1m14s p(90)=50.22s p(95)=1m8s
http_reqs ..... 20 / 9943317 / 0
iteration_duration ..... avg=50.99ms min=25.49ms med=53.32ms max=1m22s p(90)=1m29s p(95)=1m33s
iterations ..... 20 / 0.0013517%
vus ..... 1 min=1 max=5
vus_max ..... 5 min=5 max=5

running (03m57.1s), 0/5 VUs, 20 complete and 0 interrupted iterations
default / [=====] 5 VUs 63m57.1s/10m0s 20/20 shared iters
    
```

Gambar 8. Hasil Pengujian Performa Sistem Menggunakan Metode *Sharding Database*

3.2. Hasil Pengujian Reliabilitas Sistem

```

kompiang@pop-os:~/development/skripsi/skripsi-be$ k6 run loadtest/loadtest.ingestion.script.js

      M K6
      .io

execution: local
script: loadtest/loadtest.ingestion.script.js
output: -

scenarios: (100.00%) 1 scenario, 10000 max VUs, 10m30s max duration (incl. graceful stop):
  default: 10000 iterations shared among 10000 VUs (maxDuration: 10m0s, gracefulStop: 30s)

data_received ..... 2.3 MB 790 B/s
data_sent ..... 5.4 MB 1.9 MB/s
http_req_blocked ..... avg=98.47ms min=766ns med=108.35ms max=412.82ms p(90)=177.40ms p(95)=182.01ms
http_req_connecting ..... avg=96.02ms min=9µs med=108.59ms max=277.20ms p(90)=176.58ms p(95)=186.77ms
http_req_duration ..... avg=1.20s min=33.41ms med=1.31s max=2.64s p(90)=2.15s p(95)=2.26s
http_req_failed ..... 100.00% / 10000 / 0
http_req_receiving ..... avg=30.3µs min=5.79µs med=21.46µs max=19.47ms p(90)=49.47µs p(95)=65.61µs
http_req_sending ..... avg=8.6ms min=5.63µs med=3.5ms max=217.54ms p(90)=19.17ms p(95)=23.49ms
http_req_tls_handshaking ..... avg=9µs min=0µs med=9µs max=6µs p(90)=9µs p(95)=8µs
http_req_waiting ..... avg=1.29s min=52.44ms med=1.31s max=2.63s p(90)=2.14s p(95)=2.25s
http_reqs ..... 10000 / 3400 / 240112%
iteration_duration ..... avg=1.38s min=35.64ms med=1.42s max=2.83s p(90)=2.25s p(95)=2.36s
iterations ..... 10000 / 3700 / 340112%
vus ..... 500 min=500 max=10000
vus_max ..... 10000 min=10000 max=10000

running (09m02.9s), 0/10000 VUs, 10000 complete and 0 interrupted iterations
default / [=====] 10000 VUs 09m02.9s/10m0s 10000/10000 shared iters
    
```

Gambar 9. Hasil Pengujian Reliabilitas Sistem yang Tidak Menggunakan *Kappa Architecture*

Dapat dilihat pada Gambar 9 yang merupakan hasil dari pengujian *reliabilitas* dari sistem yang tidak menggunakan *kappa architecture*. Ketika kita membuat order dengan kondisi salah satu *stream processor service* mati dan tidak menggunakan *kappa architecture*, menghasilkan nilai metris *http_req_failed* dengan nilai 100% atau semua *request* gagal dikirim dan tidak ada data yang berhasil diproses.

```

kompiang@pop-os:~/development/skripsi/skripsi-be/loadtest$ k6 run loadtest.ingestion.script.js

      M K6
      .io

execution: local
script: loadtest.ingestion.script.js
output: -

scenarios: (100.00%) 1 scenario, 10000 max VUs, 10m30s max duration (incl. graceful stop):
  default: 10000 iterations shared among 10000 VUs (maxDuration: 10m0s, gracefulStop: 30s)

data_received ..... 4.2 MB 1.1 MB/s
data_sent ..... 5.4 MB 1.4 MB/s
http_req_blocked ..... avg=124.37ms min=1.11µs med=126.61ms max=565.02ms p(90)=193.59ms p(95)=256.93ms
http_req_connecting ..... avg=118.29ms min=0µs med=124.65ms max=407.70ms p(90)=184.36ms p(95)=215.05ms
http_req_duration ..... avg=1.97s min=61.13ms med=1.97s max=3.58s p(90)=3.11s p(95)=3.22s
  { expected_response:true } ..... avg=1.97s min=61.13ms med=1.97s max=3.55s p(90)=3.11s p(95)=3.23s
http_req_failed ..... 0.00% / 0 / 10000
http_req_receiving ..... avg=71.61µs min=9.28µs med=42.02µs max=6.0ms p(90)=157.95µs p(95)=211.02µs
http_req_sending ..... avg=12.43ms min=7.60µs med=5.54ms max=412.57ms p(90)=28.42ms p(95)=47.77ms
http_req_tls_handshaking ..... avg=0µs min=0µs med=0µs max=0µs p(90)=0µs p(95)=0µs
http_req_waiting ..... avg=1.96s min=61.04ms med=1.96s max=3.53s p(90)=3.1s p(95)=3.21s
http_reqs ..... 10000 / 2500 / 410037%
iteration_duration ..... avg=2.4s min=100.61ms med=2.11s max=3.81s p(90)=3.24s p(95)=3.37s
iterations ..... 10000 / 2500 / 410037%
vus ..... 400 min=400 max=10792
vus_max ..... 10000 min=10000 max=10000

running (09m03.8s), 0/10000 VUs, 10000 complete and 0 interrupted iterations
default / [=====] 10000 VUs 09m03.8s/10m0s 10000/10000 shared iters
    
```

Gambar 10. Hasil Pengujian Reliabilitas Sistem yang Menggunakan *Kappa Architecture*

Dapat dilihat pada Gambar 10 pengiriman tidak mendapatkan *error* apapun. Hal ini dapat dilihat pada metris *http_req_failed* yang menghasilkan nilai 0% yang memberikan informasi, bahwa semua *request* sukses dikirimkan dan tidak ada *error*, meskipun *transform service* dalam kondisi mati. Hal ini sangat penting dalam keandalan sistem, karena kita tidak ingin pengguna mendapatkan masalah karena sistem tidak bisa melakukan pembuatan order. Hal ini juga dimungkinkan karena data disimpan secara sementara dalam *message broker*. Setelah *stream processor system* diaktifkan kembali, semua data yang tersimpan di *message broker* berhasil dipulihkan dan disimpan kembali ke *longterm database* dan *sharding database*, tanpa ada kehilangan data.

3.3. Hasil Pengujian Performa Penulisan Data Time-Series

```
kompiang@pop-os:~/development/skripsi/skripsi-be$ k6 run loadtest/loadtest.ingestion.script.js
┌───────────┴───────────┐
┆  M  K 6              ┆
┆  .io                 ┆
└───────────┬───────────┘

execution: local
script: loadtest/loadtest.ingestion.script.js
output:

scenarios: (100.00%) 1 scenario, 10000 max VUs, 10m30s max duration (incl. graceful stop):
  default: 10000 iterations shared among 10000 VUs (maxDuration: 10ms, gracefulStop: 30s)

data_received ..... 4.2 MB 411 KB/s
data_sent ..... 5.4 MB 534 KB/s
http_req_blocked ..... avg=110.50ms min=832ns med=120.28ms max=820.04ms p(90)=180.53ms p(95)=303.32ms
http_req_connecting ..... avg=107.56ms min=0s med=117.15ms max=320ms p(90)=189.16ms p(95)=330.19ms
http_req_duration ..... avg=6.65s min=0.83ms med=7.26s max=9.82s p(90)=9.2s p(95)=9.35s
  [ expected_response:true ] ..... avg=6.65s min=0.83ms med=7.26s max=9.82s p(90)=9.2s p(95)=9.35s
http_req_failed ..... 0.00% 0 / 10000
http_req_receiving ..... avg=3.01ms min=11.39µs med=3.32µs max=390.81ms p(90)=150.54µs p(95)=229.39µs
http_req_sending ..... avg=33.91ms min=0.89µs med=4.31ms max=523.89ms p(90)=69.40ms p(95)=107.79ms
http_req_tls_handshaking ..... avg=0s min=0s med=0s max=0s p(90)=0s p(95)=0s
http_req_waiting ..... avg=6.61s min=31.63ms med=7.22s max=9.74s p(90)=9.16s p(95)=9.35s
http_reqs ..... 10000 300.200000/s
iteration_duration ..... avg=6.17s min=103.21ms med=7.34s max=10.1s p(90)=9.33s p(95)=9.53s
iterations ..... 10000 300.200000/s
vus ..... 10000 1000.000000 / 10000
vus_max ..... 10000 min=10000 max=10000

running (00m10.1s), 00000/10000 VUs, 10000 complete and 0 interrupted iterations
default [=====] 10000 VUs 00m10.1s/10m30s 10000/10000 shared iters
```

Gambar 11. Hasil Pengujian Performa Penulisan Data Time-Series Dengan Tidak Menggunakan Kappa Architecture Tetapi Menggunakan Sharding Database

Hasil pengujian performa penulisan data *time-series* dengan tidak menggunakan *kappa architecture*, tetapi menggunakan *sharding database* dapat dilihat pada Gambar 11. Pada pengujian ini dilakukan pembuatan 10000 data oleh 10000 *virtual users*. Hasil dari *load testing* pada metris *http_req_waiting* menghasilkan nilai 9,35 detik pada persentil ke-95 dan metris *http_req_duration* memiliki nilai 0%, artinya semua *request* berhasil terkirim.

```
kompiang@pop-os:~/development/skripsi/skripsi-be$ k6 run loadtest/loadtest.ingestion.script.js
┌───────────┴───────────┐
┆  M  K 6              ┆
┆  .io                 ┆
└───────────┬───────────┘

execution: local
script: loadtest/loadtest.ingestion.script.js
output:

scenarios: (100.00%) 1 scenario, 10000 max VUs, 10m30s max duration (incl. graceful stop):
  default: 10000 iterations shared among 10000 VUs (maxDuration: 10ms, gracefulStop: 30s)

data_received ..... 4.2 MB 1.2 MB/s
data_sent ..... 5.4 MB 1.5 MB/s
http_req_blocked ..... avg=154.62ms min=500ns med=186.56ms max=675.55ms p(90)=288.25ms p(95)=298.71ms
http_req_connecting ..... avg=159.25ms min=0s med=184.83ms max=374.19ms p(90)=285.95ms p(95)=295.22ms
http_req_duration ..... avg=1.62s min=14.49ms med=1.62s max=3.36s p(90)=2.73s p(95)=2.8s
  [ expected_response:true ] ..... avg=1.62s min=14.49ms med=1.62s max=3.36s p(90)=2.73s p(95)=2.8s
http_req_failed ..... 0.00% 0 / 10000
http_req_receiving ..... avg=77.28µs min=9.42µs med=49.32µs max=21.35ms p(90)=139.97µs p(95)=263.35µs
http_req_sending ..... avg=15.15ms min=4.85µs med=5.41ms max=483.41ms p(90)=33.09ms p(95)=43.77ms
http_req_tls_handshaking ..... avg=0s min=0s med=0s max=0s p(90)=0s p(95)=0s
http_req_waiting ..... avg=1.61s min=14.46ms med=1.6s max=3.26s p(90)=2.72s p(95)=2.78s
http_reqs ..... 10000 300.000000/s
iteration_duration ..... avg=2.78s min=14.57ms med=1.8s max=3.47s p(90)=2.93s p(95)=3.04s
iterations ..... 10000 300.000000 / 10000
vus ..... 3715 min=3715 max=3715
vus_max ..... 10000 min=10000 max=10000

running (00m03.5s), 00000/10000 VUs, 10000 complete and 0 interrupted iterations
default [=====] 10000 VUs 00m03.5s/10m30s 10000/10000 shared iters
kompiang@pop-os:~/development/skripsi/skripsi-be$
```

Gambar 12. Hasil Pengujian Performa Penulisan *Data Time-Series* Dengan Menggunakan *Kappa Architecture* dan Menggunakan *Sharding Database*

Hasil pengujian pada skenario pengujian performa penulisan data *time-series* menggunakan *kappa architecture* dan *sharding database* dapat dilihat pada Gambar 12. Pada skenario ini, dilakukan pembuatan 10000 data *time-series* oleh 10000 *virtual users* yang menggunakan *kappa*

architecture dan *sharding database* dalam pembuatan order. Hasil dari *load testing* pada skenario 5 pada metris *http_req_waiting* menghasilkan nilai 2,78 detik pada persentil ke-95 dan metris *http_req_duration* memiliki nilai 0%, artinya semua *request* berhasil terkirim.

4. Kesimpulan

Berdasarkan penelitian yang telah dilaksanakan dan juga hasil yang diperoleh dari penelitian, dapat ditarik beberapa kesimpulan sebagai berikut:

- a. Dari hasil pengujian performa pengambilan data dalam skenario 5 *virtual users* dan 20 iterasi dengan rentang waktu 89 hari, terbukti bahwa *sharding database* meningkatkan kecepatan pemrosesan hingga 60,46%.
- b. Dalam penelitian ini, ketidakaktifan *transform service* diuji dengan mengirimkan 10000 data melalui 10000 pengguna virtual. Hasilnya, sistem berhasil mengirimkan 100% data tersebut. Ini menunjukkan kemampuan arsitektur Kappa dalam menjaga kelangsungan proses dan data meskipun menghadapi gangguan pada salah satu *service*-nya.
- c. Pada pengujian performa *kappa architecture* yang dikombinasikan dengan *sharding database*, dikirimkan 10000 data oleh 10000 *virtual users*. Sistem yang mengimplementasikan *kappa architecture* yang dikombinasikan dengan *sharding database* menunjukkan peningkatan kecepatan sebesar 70,26% dibandingkan dengan sistem yang hanya mengimplementasikan *sharding database*, tetapi tidak mengimplementasikan *kappa architecture*.

References

- [1] N. Marz and J. (James O.) Warren, *Big data : principles and best practices of scalable real-time data systems*, vol. 1. 2015.
- [2] F. N. Fote, S. Mahmoudi, A. Roukh, and S. A. Mahmoudi, "Big Data Storage and Analysis for Smart Farming," in *Proceedings of 2020 5th International Conference on Cloud Computing and Artificial Intelligence: Technologies and Applications, CloudTech 2020*, Institute of Electrical and Electronics Engineers Inc., Nov. 2020. doi: 10.1109/CloudTech49835.2020.9365869.
- [3] M. Tahmassebpour, "A New Method for Time-Series Big Data Effective Storage," *IEEE Access*, vol. 5, pp. 10694–10699, 2017, doi: 10.1109/ACCESS.2017.2708080.
- [4] S. Samidi, R. Y. Suladi, and A. B. Lesmana, "Implementation of Database Distributed Sharding Horizontal Partition in MySQL. Case Study of Application of Food Serving On Kemkes," *JURNAL SISFOTEK GLOBAL*, vol. 12, no. 1, p. 50, Mar. 2022, doi: 10.38101/sisfotek.v12i1.477.
- [5] Y. Kumar, "Lambda Architecture-Realtime Data Processing," 2020, doi: 10.13140/RG.2.2.19091.84004.
- [6] J. B. Nkamla Penka, S. Mahmoudi, and O. Debauche, "A new Kappa Architecture for IoT Data Management in Smart Farming," in *Procedia Computer Science*, Elsevier B.V., 2021, pp. 17–24. doi: 10.1016/j.procs.2021.07.006.
- [7] O. Debauche *et al.*, "RAMi: A New Real-Time Internet of Medical Things Architecture for Elderly Patient Monitoring," *Information (Switzerland)*, vol. 13, no. 9, Sep. 2022, doi: 10.3390/info13090423.
- [8] N. A. N. Sobri *et al.*, "INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION: Study of Database Connection Pool in Microservice Architecture," pp. 566–571, 2022, [Online]. Available: www.joiv.org/index.php/joiv

Analisis Sentimen Berbasis Aspek Terhadap Ulasan Aplikasi Mobile JKN Dengan Metode Random Forest Dan Information Gain Sebagai Seleksi Fitur

Yauw James Fang Dwiputra Harta^{a1}, I Gede Arta Wibawa^{a2}, Anak Agung Istri Ngurah Eka Karyawati^{b3}, Komang Ari Mogi^{b4}

^{a1}Informatics Department
^{a2}informatics Department, Udayana University
Bali, Indonesia

¹jamesfangyauw@gmail.com

²gede.arta@unud.ac.id

^{b1}Informatics Department
^{b2}informatics Department, Udayana University
Bali, Indonesia

³eka.karyawati@unud.ac.id

⁴arimogi@unud.ac.id

ABSTRACT

The Mobile JKN application plays a crucial role in providing access to healthcare services in Indonesia. However, understanding user sentiments through reviews is essential for continual improvement. Thus, employing sentiment analysis techniques becomes imperative to gauge user feedback effectively. This study employs the Random Forest algorithm and Information Gain method to conduct aspect-based sentiment analysis (ABSA) on user reviews of the Mobile JKN application. The ABSA model demonstrates impressive classification accuracies: 85% for aspect classification, 93% for application-related sentiment, and 87% for service-related sentiment. Notably, a 40% feature selection rate significantly enhances model accuracy. Experimentation confirms a 1-3% boost in sentiment classification accuracy with feature selection. Beyond technical insights, this research bears broader implications, potentially refining healthcare service quality through the Mobile JKN application. By integrating user sentiments with precision, developers can discern nuanced areas for improvement, fostering a more user-centric approach to application development and service delivery. This study underscores the pivotal role of sentiment analysis in enhancing digital user experiences and guiding iterative improvements in mobile application development.

Keywords : *analysis, sentiment, aspect, random forest, information gain, mobile JKN*

1. Introduction

Perkembangan digital saat ini berlangsung sangat pesat, dapat dilihat dari banyaknya aplikasi-aplikasi yang beredar salah satunya Mobile JKN. Mobile JKN (Jaminan Kesehatan Nasional) adalah aplikasi yang diluncurkan dan dikelola oleh BPJS (Badan Penyelenggaraan Jaminan Sosial) Kesehatan Indonesia mulai tahun 2019. Aplikasi ini dibuat dengan tujuan untuk memudahkan masyarakat khususnya pengguna BPJS Kesehatan untuk mendapatkan asuransi pelayanan kesehatan mereka kapanpun dan dimanapun dengan hanya menggunakan akses internet dalam aplikasi Mobile JKN ini. Di setiap aplikasi yang beredar di Google Play Store termasuk mobile JKN terdapat kolom ulasan terhadap aplikasi tersebut. Kolom ulasan ini merupakan suatu wadah bagi pengguna aplikasi untuk memberikan umpan balik (opini) melalui tulisan kepada pengelola aplikasi tersebut. Harapannya ulasan-ulasan yang diberikan oleh pengguna dapat dijadikan bahan evaluasi dan peningkatan kualitas untuk kemajuan suatu aplikasi yang khususnya terdapat di Google Play Store.

Mengetahui pentingnya untuk mengevaluasi ulasan dari aplikasi mobile JKN ini. Maka dari itu peneliti tertarik untuk menganalisis sentimen pengguna terhadap aplikasi Mobile JKN

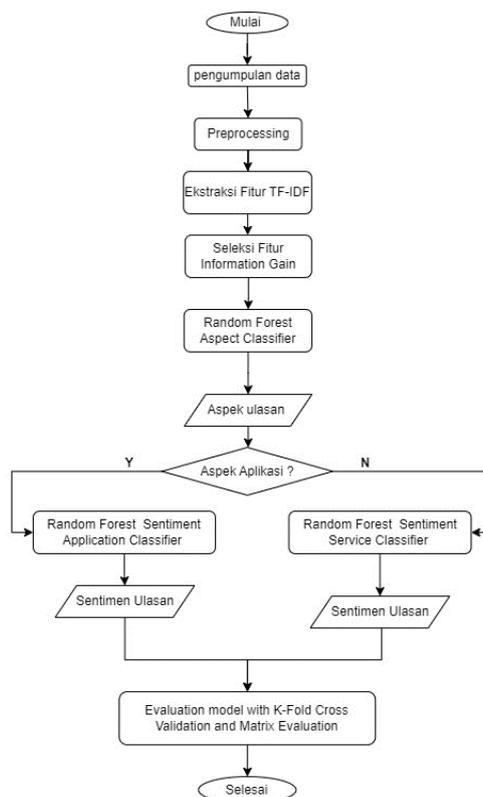
yang dikelola oleh lembaga BPJS Kesehatan pada kolom ulasan Google Play Store dengan menggunakan salah satu metode machine learning. Penelitian ini berfokus dalam menganalisis sentimen berbasis aspek terhadap ulasan oleh pengguna pada aplikasi pelayanan asuransi kesehatan masyarakat berupa Mobile JKN yang tersedia di Google Play Store.

Sebelumnya telah terdapat penelitian terkait analisis sentimen yang dilakukan oleh peneliti lain, seperti yang dilakukan pada tahun 2021 terkait analisis sentimen pada ulasan aplikasi Mobile JKN dengan metode Maximum Entropy dan mendapatkan akurasi sebesar 80% (Rohman & Adinugroho, 2021). Terdapat pula penelitian mengenai analisis sentimen berbasis aspek terhadap ulasan restoran yang dilakukan pada tahun 2021 dengan menggunakan metode Support Vector Machine dan mendapatkan akurasi sebesar 15,119% (Pamungkas & Romadhony, 2021). Selain itu, terdapat penelitian lain pada tahun yang sama terkait analisis sentimen terhadap ulasan film dengan menggunakan metode Modified Balanced Random Forest yang merupakan modifikasi dari metode Random Forest dan mendapatkan akurasi sebesar 79% (Khairunnisa dkk., 2021).

Melihat nilai akurasi dari penelitian sebelumnya yang masih belum terlalu tinggi (<81%) dan belum ada yang menerapkan analisis sentiment berbasis aspek pada ulasan aplikasi mobile JKN. Oleh karena itu peneliti memilih topik penelitian terkait dengan analisis sentimen berbasis aspek terhadap aplikasi Mobile JKN dengan metode klasifikasi yang diangkat penulis untuk digunakan dalam melakukan klasifikasi aspek dan sentimen adalah algoritma Random Forest (RF) dan Information Gain sebagai seleksi fiturnya disertai melakukan implementasi hyperparameter tuning dengan harapan mampu untuk meningkatkan nilai dari akurasi sistem yang dirancang dengan menggunakan algoritma Random Forest (RF).

2. Research Methods

Metode penelitian yang digunakan pada penelitian ini adalah dengan menggunakan metode eksperimen dengan urutan langkah-langkah dalam alur metode penelitian terlihat seperti gambar 2.1.



Gambar 2.1 Alur Metode Penelitian

Gambar 2.1 merupakan flowchart untuk alur metode penelitian ini yang dimulai dari pengumpulan data. Berikut adalah penjelasan lebih lanjut mengenai diagram alur metode penelitian pada gambar 2.1 :

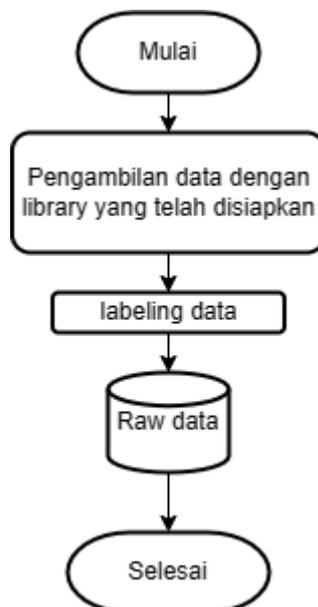
1. Pengumpulan Data

Langkah awal dalam penelitian ini adalah melakukan pengumpulan data ulasan Mobile-JKN pada kolom komentar Google Play Store. Proses pengambilan data ini menggunakan library dari package pandas yaitu google-play-scraper dengan bahasa pemrograman Python[1]. Hasil dari scrapping ini akan disimpan pada file berformat .csv. Bentuk data yang disimpan tersebut masih merupakan bentuk data mentah (raw data) dalam bentuk teks berbahasa Indonesia.

Selanjutnya, dilakukan pelabelan kelas data sentimen akan dikelompokkan menjadi 2 kelas yaitu kelas negatif dan kelas positif yang didasarkan pada fitur score/rating yang terdapat pada data ulasan. Nilai score/rating akan merepresentasikan label data dengan sentimen positif atau negatif. Jika semisal scorenya berupa bintang 4 dan 5 maka akan tergolong sentimen positif. Sedangkan jika semisal scorenya berupa bintang 1 dan 2 akan tergolong sentimen negatif [2].

Adapun untuk pelabelan data aspek, akan dikelompokkan menjadi 2 yaitu aspek Aplikasi dan aspek Pelayanan. Pelabelan aspek nanti akan dilakukan oleh pakar yang memiliki kompetensi dalam bidang bahasa Indonesia. Sehingga ulasan tersebut nantinya akan bisa dikelompokkan kedalam aspek fitur aplikasi atau aspek kebermanfaatan aplikasi.

Proses pengumpulan data dan pelabelan ini dapat dilihat secara detail pada gambar 2.2.



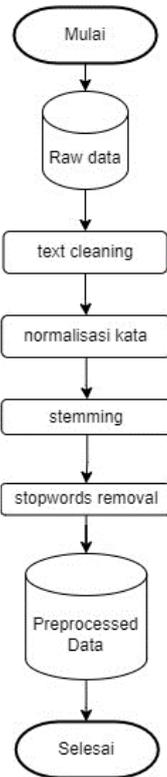
Gambar 2.2. Alur Pengumpulan Data

Pada gambar 2.2, setelah data selesai diambil maka akan dilakukan labeling secara manual. Dan setelah itu akan terbentuk sebuah raw data yang berisi terkait ulasan, sentiment, dan aspeknya.

2. Preprocessing

Sebelum dilakukan ekstraksi fitur, dataset teks yang telah didapatkan sebelumnya akan melalui tahap text preprocessing yang mencakup beberapa langkah seperti yakni: (1) text cleaning, (2) normalisasi kata, (3) stemming, (4) stopwords removal [3]. Teks cleaning melibatkan penghapusan karakter tidak diinginkan dan mengubah semua huruf menjadi huruf kecil. Normalisasi kata mengubah variasi kata menjadi bentuk standar. Stemming menghapus akhiran kata untuk menyisakan bentuk dasarnya. Penghapusan stopwords membuang kata-kata umum yang tidak memberikan informasi penting. Gabungan dari keempat proses ini membantu membersihkan dan mempersiapkan data

sehingga dataset teks menjadi lebih bersih dan siap untuk diolah selanjutnya. Alur proses text preprocessing yang telah dijelaskan dapat dilihat pada Gambar 2.3 dan contoh hasil untuk masing masing tahapan text preprocessing dapat dilihat pada table 2.1



Gambar 2.3 Alur Proses Text Preprocessing

Tabel 2.1 Contoh Penerapan *Text Preprocessing*

No	Nama Proses	Hasil
1	Data Awal	Ngga guna aplikasinya.
2	<i>Punctuation</i>	Ngga guna aplikasinya
3	<i>Casefolding</i>	ngga guna aplikasinya
4	<i>Tokenizing</i>	['ngga', 'guna', 'aplikasinya']
5	<i>Normalizing</i>	['tidak', 'guna', 'aplikasinya']
6	<i>Stemming</i>	['tidak', 'guna', 'aplikasi']
7	<i>Stopword</i>	['tidak', 'guna', 'aplikasi']
8	Data Hasil	tidak guna aplikasi

Pada tabel 2.1 dapat dilihat ketika data awalnya berupa “Ngga guna aplikasinya”, setelah dilakukan *punctuation*, *casefolding*, *tokenizing*, *normalizing*, *stemming*, dan *stopword* maka didapatkan data yang bersih menjadi “tidak guna aplikasi”.

3. TF-IDF

Kemudian dalam penelitian ini, perhitungan nilai bobot tersebut akan dilakukan dengan menggunakan pemanggilan modul yang terdapat pada pustaka *scikit-learn*. Pustaka ini merupakan salah satu pustaka yang dapat digunakan untuk keperluan pembelajaran mesin dan ekstraksi fitur, tidak terkecuali dengan menggunakan TF-IDF. Dalam pustaka ini, *class TfidfVectorizer* akan digunakan untuk mengubah *dataset*

menjadi suatu vektor yang digunakan sebagai representasi dari bobot setiap *term* untuk sebuah dokumen dari seluruh ulasan yang tersedia dalam *dataset* [4].

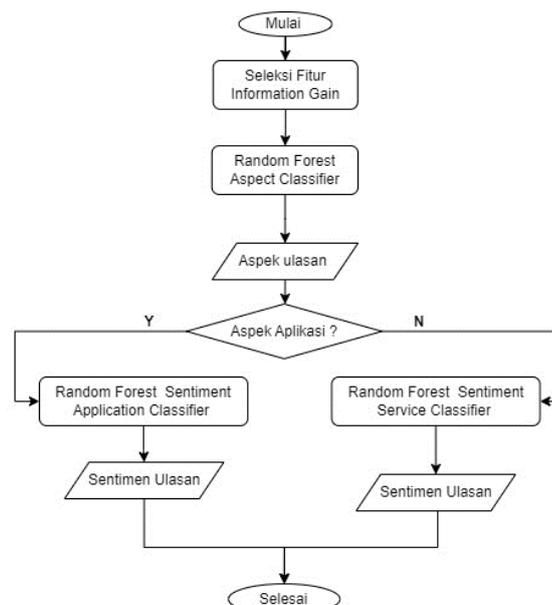
4. Seleksi Fitur dengan Information Gain

Pada penelitian ini metode yang dapat digunakan untuk memilih fitur terbaik adalah metode Information Gain (IG) yang akan dilakukan pada tahap ini dengan menggunakan library *Sklearn* dari python yang berupa metode *Mutual Information*. Dan untuk jumlah fitur dari hasil seleksi nanti akan di-*tuning* mulai dari 20%, 40%, 60%, dan 80%. Proses dasar yang ada dalam Information gain untuk mengukur seberapa baik suatu fitur memisahkan data menjadi kelas-kelas yang berbeda yaitu dengan menghitung nilai information gainnya. Semakin tinggi information gain, semakin baik fitur tersebut dalam membedakan atau mengklasifikasikan data [5].

5. Klasifikasi dengan Random Forest

Pada tahap ini akan dilakukan klasifikasi aspek dengan membangun model klasifikasi menggunakan algoritma Random Forest untuk mengklasifikasikan dataset ke dalam dua kategori aspek yaitu aspek aplikasi mobile-JKN dan aspek pelayanan. Hal ini dilakukan dengan menggunakan fitur-fitur yang relevan dan berkaitan dengan aspek yang dimaksud, seperti kata-kata yang sering atau biasanya digunakan dalam konteks yang berkaitan dengan aspek tersebut.

Setelah berhasil melakukan klasifikasi aspek tersebut, maka selanjutnya tahap kedua dilakukan untuk mengklasifikasikan sentimen berdasarkan aspek yang telah diidentifikasi pada tahap sebelumnya apakah masuk ke kelas positif atau negatif. Pada tahap ini, akan dilakukan pembuatan model klasifikasi random forest dengan memanfaatkan fitur-fitur yang berkaitan dengan sentimen positif atau negatif dalam setiap aspek. Model ini kemudian digunakan untuk mengklasifikasikan sentimen dalam setiap aspek, sehingga diperoleh informasi mengenai sentimen tersebut apakah masuk ke sentimen positif atau negatif pada setiap aspek [6]. Flowchart alur sistem untuk melihat pemilihan model yang digunakan dapat dilihat pada pada gambar 2.4



Gambar 2.4 Flowchart Pemilihan Model

. Pada gambar 2.4 dapat dilihat jika aspek ulasan berupa aspek aplikasi maka akan masuk ke model klasifikasi untuk aspek aplikasi, jika tidak maka akan masuk ke model klasifikasi untuk aspek pelayanan. Kemudian hasil keluaran dari masing masing model klasifikasi yaitu sentiment ulasan apakah positif atau negatif.

6. Evaluasi Model

Pada penelitian ini, penulis akan memanfaatkan dataset yang terdiri dari dua bagian, yakni 80% sebagai data training (data latih) dan 20% sebagai data validasi (data uji). Pada tahap validasi ini, data training akan digunakan dalam proses K-fold Cross Validation dengan mengimplementasikan hyperparameter tuning pada algoritma Random Forest, termasuk *n_estimators*, *max_features*, *max_depth*. Juga diimplementasikan seleksi fitur dengan jumlah fitur terbaik maksimal 80%. Jumlah fold yang digunakan dalam penelitian ini yaitu 5 Fold, dimana dari 5 fold ini nanti akan dicari rata rata dari hasil evaluasinya [7].

Dalam penelitian ini, evaluasi model yang telah didapatkan sebelumnya akan dilakukan dengan menggunakan perhitungan confusion matrix. Untuk menilai performa dari model klasifikasi yang telah dibuat, matrix evaluasi yang digunakan pada penelitian ini yaitu accuracy yang nilainya didapatkan dari persamaan masing matrix evaluasi tersebut. Perhitungan matrix evaluasi tersebut dilakukan dengan membandingkan kelas prediksi yang terdapat pada confusion matrix yang telah dibuat sebelumnya.

3. Result and Discussion

Data yang digunakan diperoleh dari review terhadap aplikasi Mobile JKN di Google Play Store. Data tersebut merupakan data mentah dalam bentuk teks review yang diperoleh melalui teknik scraping dengan menggunakan library google-play-scraper dalam bahasa pemrograman Python. Setelah diambil, kemudian data disimpan dalam format file .csv. Setiap review memiliki atribut konten yang berisi ulasan pengguna, atribut aspek yang mengklasifikasikan aspek dari ulasan, dan atribut skor yang menentukan polaritas ulasan.

Data review ini terdiri dari 1.176 review aplikasi yang terbagi menjadi dua aspek, di mana setiap aspek memiliki dua kategori sentimen dari review tersebut. Distribusi data yang digunakan dapat dilihat pada Tabel 3.1

Tabel 3.1 Distribusi Data Ulasan

No	Aspek	Sentimen	Jumlah
1	Aplikasi	Positif	294
2		Negatif	294
3	Pelayanan	Positif	294
4		Negatif	294
Total			1.176

Pada tabel 3.1 dapat dilihat bahwa total ulasan yang digunakan untuk masing masing aplikasi dan pelayanan kesehatan sejumlah 588 ulasan. Dimana dari 588 ulasan ini akan terbagi lagi menjadi 2 yaitu ulasan untuk sentimen yang positif dan negatif dengan total masing masing 294.

Kemudian sebelum performa model diujikan menggunakan data uji yang sebelumnya dipisahkan pada tahap pembangunan model klasifikasi, model klasifikasi aspek dan model – model klasifikasi sentimen akan dilakukan validasi terlebih dahulu dengan menggunakan metode K-fold cross validation dengan library GridSearchCrossValidation. Dengan metode ini, data latih yang sebelumnya berasal dari data utuh akan dibagi kembali menjadi 5 fold (lipatan). Kemudian, akan dilakukan iterasi terhadap setiap fold agar setiap fold dapat menjadi data uji baru dan fold sisanya dapat menjadi data latih baru. Hasil dari setiap iterasi adalah akurasi model dalam memprediksi data uji baru. Selanjutnya, setiap hasil akurasi akan dijumlahkan dan dibagi dengan jumlah fold yang ditentukan sebelumnya untuk memperoleh rerata akurasi dari model yang telah dibangun dengan K-fold cross validation [7].

Adapun hasil akurasi dari setiap model yang tervalidasi tanpa seleksi fitur Information Gain terlihat pada tabel 3.2 dengan keterangan kolom “k” sebagai nilai akurasi terhadap data validasi setiap iterasinya dan “k_avg” sebagai rata rata dari k1 hingga k5 dan kolom name yang bernilai “aspect” merupakan model klasifikasi aspect, “app” merupakan singkatan “application” untuk model klasifikasi sentimen untuk aspek aplikasi (application), dan “service” merupakan model klasifikasi sentiment untuk aspek pelayanan.

Tabel 3.2 Mesin Klasifikasi Terbaik (Tanpa Seleksi Fitur)

name	n_estimator	Max_feature	Min_samples_leaf	Kfold_avg
aspect	100	Log2	1	0.86
application	300	Log2	1	0.92
service	300	Log2	1	0.88

Pada tabel 3.2 diatas dapat dilihat bahwa model untuk klasifikasi aspek mendapat nilai kfold 86%, model klasifikasi untuk aspek aplikasi mendapat nilai kfold 92% dan untuk klasifikasi ulasan service mendapat nilai kfold 88%. Kemudian untuk hasil akurasi terbaik untuk nilai rata rata kfold dari masing masing model klasifikasi yang menerapkan seleksi fitur dengan persertanse 80% dapat dilihat pada tabel 3.3.

Tabel 3.3 Mesin Klasifikasi Terbaik (Persentase Seleksi Fitur : 80%)

name	n_estimators	max_features	min_samples_leaf	k1	k2	k3	k4	k5	k_avg
aspect	100	log2	1	0.85	0.85	0.82	0.84	0.91	0.85
applicati on	500	log2	1	0.89	0.93	0.91	0.91	0.93	0.91
service	500	log2	1	0.93	0.83	0.91	0.86	0.86	0.88

Pada tabel 3.3 diatas dapat dilihat bahwa model untuk klasifikasi aspek mendapat nilai kfold 85%, model klasifikasi untuk aspek aplikasi mendapat nilai kfold 91% dan untuk klasifikasi ulasan service mendapat nilai kfold 88%. Kemudian untuk hasil akurasi terbaik untuk nilai rata rata kfold dari masing masing model klasifikasi yang menerapkan seleksi fitur dengan persertanse 60% dapat dilihat pada tabel 3.4.

Tabel 3.4 Mesin Klasifikasi Terbaik (Persentase Seleksi Fitur : 60%)

name	n_estimators	max_features	min_samples_leaf	k1	k2	k3	k4	k5	k_avg
aspect	500	log2	1	0.89	0.87	0.84	0.84	0.89	0.87
applicati on	100	log2	1	0.91	0.91	0.91	0.88	0.91	0.91
service	300	log2	1	0.93	0.84	0.91	0.86	0.85	0.88

Pada tabel 3.4 diatas dapat dilihat bahwa model untuk klasifikasi aspek mendapat nilai kfold 87%, model klasifikasi untuk aspek aplikasi mendapat nilai kfold 91% dan untuk klasifikasi ulasan service mendapat nilai kfold 88%. Kemudian untuk hasil akurasi terbaik untuk nilai rata rata kfold dari masing masing model klasifikasi yang menerapkan seleksi fitur dengan persertanse 40% dapat dilihat pada tabel 3.5.

Tabel 3.5 Mesin Klasifikasi Terbaik (Persentase Seleksi Fitur : 40%)

name	n_estimators	max_features	min_samples_leaf	k1	k2	k3	k4	k5	k_avg
aspect	300	log2	1	0.87	0.86	0.83	0.83	0.89	0.86
applicati on	500	log2	1	0.91	0.91	0.93	0.9	0.93	0.92
service	100	log2	1	0.91	0.79	0.94	0.88	0.89	0.88

Pada tabel 3.5 diatas dapat dilihat bahwa model untuk klasifikasi aspek mendapat nilai kfold 86%, model klasifikasi untuk aspek aplikasi mendapat nilai kfold 92% dan untuk klasifikasi ulasan service mendapat nilai kfold 88%. Kemudian terakhir untuk hasil akurasi terbaik untuk nilai

rata rata kfold dari masing masing model klasifikasi yang menerapkan seleksi fitur dengan persertanse 20% dapat dilihat pada tabel 3.6.

Tabel 3.6 Mesin Klasifikasi Terbaik (Persentase Seleksi Fitur : 20%)

name	n_es timat ors	max_f eatur es	min_s ample s_leaf	k1	k2	k3	k4	k5	k_av g
aspect	100	sqrt	1	0.88	0.85	0.81	0.82	0.87	0.85
applicati on	100	log2	1	0.88	0.9	0.89	0.86	0.91	0.89
service	100	nan	1	0.89	0.83	0.91	0.83	0.87	0.87

Pada tabel 3.6 diatas dapat dilihat bahwa model untuk klasifikasi aspek mendapat nilai kfold 85%, model klasifikasi untuk aspek aplikasi mendapat nilai kfold 89% dan untuk klasifikasi ulasan service mendapat nilai kfold 87%.

Dari tabel 3.2 hingga 3.6 terlihat untuk mesin klasifikasi dengan nama "aspect" yang merupakan mesin klasifikasi aspek dari suatu ulasan memiliki rata rata nilai kfold berkisar 0.85 – 0.87. Untuk mesin yang tidak menerapkan seleksi fitur memiliki nilai akurasi 0.85 sedangkan terdapat mesin yang menerapkan seleksi fitur dengan persentase jumlah fitur yang dipertahankan sebesar 60% memiliki nilai akurasi rata rata kfold 0.87. Dari hasil ini dapat dikatakan bahwa penyesuaian jumlah fitur menjadi 60% dari total fitur mula mula mampu memberikan kenaikan nilai rata rata akurasi kfold sebesar 2%.

Untuk mesin klasifikasi dengan nama "application" yang merupakan mesin klasifikasi sentiment dari ulasan yang memiliki aspek aplikasi memiliki rata rata nilai kfold berkisar 0.89 – 0.92. Untuk mesin yang tidak menerapkan seleksi fitur memiliki nilai akurasi 0.92 sedangkan terdapat mesin yang menerapkan seleksi fitur dengan persentase jumlah fitur yang dipertahankan sebesar 40% memiliki nilai akurasi rata rata kfold 0.92 juga. Dari hasil ini dapat dikatakan bahwa penyesuaian jumlah fitur menjadi 40% dari total fitur mula mula belum mampu memberikan kenaikan nilai rata rata akurasi kfold, tetapi jika kita bandingkan nilai akurasi dari iterasi k1 hingga k5 terlihat bahwa model yang menerapkan seleksi fitur 40% memiliki konsistensi yang lebih baik dengan jangkauan 0.3 dibandingkan model yang tidak menerapkan seleksi fitur memiliki jangkauan 0.4. Juga model yang menerapkan seleksi fitur tentunya akan lebih cepat waktu komputasinya dan bentuk modelnya akan lebih sederhana.

Untuk mesin klasifikasi dengan nama "service" yang merupakan mesin klasifikasi sentiment dari ulasan berasppek servis memiliki rata rata nilai kfold berkisar 0.85 – 0.87. Untuk mesin yang tidak menerapkan seleksi fitur memiliki nilai akurasi 0.86 sedangkan terdapat mesin yang menerapkan seleksi fitur dengan persentase jumlah fitur yang dipertahankan sebesar 60% memiliki nilai akurasi rata rata kfold 0.87. Dari hasil ini dapat dikatakan bahwa penyesuaian jumlah fitur menjadi 60% dari total fitur mula mula mampu memberikan kenaikan nilai rata rata akurasi kfold sebesar 1%.

Dari pembahasan diatas dapat dibuktikan bahwa dengan menerapkan seleksi fitur dapat meningkatkan nilai akurasi dari mesin klasifikasi, Kemudian adapun hasil akurasi uji 3 model terbaik berdasarkan nilai k_avg pada tabel tabel diatas dari setiap model klasifikasi terbaik terlihat seperti pada tabel 3.7:

Tabel 3.7 Nilai Akurasi Prediksi Data Uji Model

Nama	Seleksi Fitur	kfold_avg	test_acc	acc_diff
aspect	Tanpa Seleksi Fitur	0.86	0.82	4%
	60%	0.87	0.84	3%
	40%	0.86	0.85	1%
application	Tanpa Seleksi Fitur	0.92	0.91	1%
	60%	0.91	0.89	2%
	40%	0.92	0.93	1%
service	Tanpa Seleksi Fitur	0.86	0.84	2%

	60%	0.87	0.86	1%
	40%	0.86	0.87	1%

Pada tabel 3.7 dapat dilihat nilai – nilai akurasi untuk data uji yang telah diperoleh dari menguji model yang telah divalidasi, berdasarkan nilai nilai tersebut terlihat bahwa tidak terdapat model yang nilai perbedaannya melebihi 10% sehingga dapat dikatakan bahwa model model ini tidak mengalami *overfitting data* saat melakukan pelatihan model [8]. Dan juga dapat dilihat dengan menerapkan seleksi fitur dengan persentase 40% dapat meningkatkan nilai akurasi dari data uji dibandingkan jika tidak menerapkan seleksi fitur.

4. Conclusion and Suggestion

Penelitian ini bertujuan untuk mengevaluasi kinerja model dalam menganalisis sentimen pengguna terhadap aplikasi mobile JKN, dengan fokus pada aspek aplikasi dan pelayanan. Melalui penggunaan algoritma random forest, diperoleh hasil sebagai berikut:

1. Model ABSA yang menggunakan metode Random Forest dan seleksi fitur dengan metode Information Gain mampu mengklasifikasikan aspek, sentimen terkait aspek aplikasi, dan sentimen terkait aspek pelayanan pada teks ulasan bahasa Indonesia dengan tingkat akurasi mencapai 85%, 93%, dan 87% secara berturut-turut. Seleksi fitur dan Hyperparameter yang digunakan untuk masing-masing model klasifikasi adalah seleksi fitur = 40%, $n_estimators = 300$, $max_features="log2"$, dan $min_samples_leaf = 1$ untuk klasifikasi aspek. Seleksi fitur = 40%, $n_estimators = 500$, $max_features="log2"$, dan $min_samples_leaf = 1$ untuk klasifikasi sentimen dari aspek aplikasi. Serta seleksi fitur = 40%, $n_estimators=400$, $max_features="log2"$, $min_samples_leaf = 1$ untuk klasifikasi sentimen dari aspek pelayanan.
2. Pengaruh penerapan seleksi fitur dari algoritma random forest terhadap akurasi klasifikasi sentimen opini pengguna mengenai aplikasi mobile JKN diuji dengan k-fold cross validation dan data uji. Hasil eksperimen menunjukkan bahwa penyesuaian nilai seleksi fitur sebesar 40% memiliki dampak peningkatan nilai akurasi prediksi model sebesar 1-3% dibandingkan jika tidak menggunakan seleksi fitur.

References

- [1] Khairunnisa, S., Adiwijaya, A., & Faraby, S. Al. (2021). Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19). *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(2), 406. <https://doi.org/10.30865/mib.v5i2.2835>
- [2] Kulsumarwati, A., Purnamasari, I., & Darmawan, B. A. (2021). Penerapan SVM dan Information Gain Pada Analisis Sentimen Pelaksanaan Pilkada Saat Pandemi. *Jurnal Teknologi Informatika dan Komputer*, 7(2), 101–109. <https://doi.org/10.37012/jtik.v7i2.641>
- [3] Mauludin Rohman, M., & Adinugroho, I. (2021). *Analisis Sentimen pada Ulasan Aplikasi Mobile JKN Menggunakan Metode Maximum Entropy dan Seleksi Fitur Gini Index Text* (Vol. 5, Nomor 6). <http://j-ptiik.ub.ac.id>
- [4] Mayasari, L., & Indarti, D. (2022). KLASIFIKASI TOPIK TWEET MENGENAI COVID MENGGUNAKAN METODE MULTINOMIAL NAÏVE BAYES DENGAN PEMBOBOTAN TF-IDF. *Jurnal Ilmiah Informatika Komputer*, 27(1), 43–53. <https://doi.org/10.35760/ik.2022.v27i1.6184>
- [5] Pamungkas, T. J., & Romadhony, A. (2021). *Analisis Sentimen Berbasis Aspek Terhadap Ulasan Restoran Berbahasa Indonesia menggunakan Support Vector Machines*.
- [6] Shinta Prima Astuti. (2020). *Analisis Sentimen Berbasis Aspek Pada Aplikasi Tokopedia Menggunakan LDA dan Naive Bayes*.

- [7] Zamzami, A. P. (2021). Analisis Sentimen Terhadap Review Film Menggunakan Metode Modified Balanced Random Forest dan Mutual Information. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5, 415–421.
- [8] Kusuma, Arlan (2023). *Penerapan Metode Klasifikasi Random Forest Terhadap Analisis Sentimen Berbasis Aspek Pada Ulasan Aplikasi Pelayanan Masyarakat*.

Analisis Sentimen Ulasan Aplikasi Solusi Kota Cerdas Menggunakan Algoritma *Naïve Bayes* dan *Support Vector Machine (SVM)* dengan Seleksi Fitur *Chi-Square*

Ni Luh Komang Indira Pramesti^{a1}, Made Agung Raharja^{a2}, Ngurah Agus Sanjaya ER^{a3},
I Gede Arta Wibawa^{a4}

^aFakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana
Badung, Bali, Indonesia

¹indiprames064@student.unud.ac.id

²made.agung@unud.ac.id

³agus_sanjaya@unud.ac.id

⁴gede.arta@unud.ac.id

Abstrak

Masyarakat yang semakin bergantung dengan teknologi dalam kegiatan sehari-hari menyebabkan banyaknya aplikasi yang hadir dalam membantu kegiatan ini. Salah satunya adalah aplikasi SpeedID yang berfungsi sebagai solusi kota cerdas. Fitur yang dimiliki beragam, mulai dari verifikasi identitas *online*, antrean *online*, manajemen usaha kuliner, manajemen usaha UKM, dan masih banyak lagi. Popularitas aplikasi ini berujung pada banyaknya ulasan yang diberikan oleh pengguna, baik itu positif, negatif, maupun netral. Dengan demikian, perlu dilakukan suatu analisis sentimen ulasan guna mengetahui pandangan pengguna terhadap aplikasi. Metode klasifikasi sentimen yang digunakan adalah *Naïve Bayes* (NB) dan *Support Vector Machine* (SVM) dengan seleksi fitur *chi-square*. Hasil evaluasi model menunjukkan bahwa seleksi fitur *chi-square* memiliki pengaruh positif terhadap performa model NB yang ditandai dengan meningkatnya nilai akurasi hingga sebesar 3,12%. Namun, seleksi fitur *chi-square* ini tidak memiliki pengaruh terhadap performa model SVM yang tidak mengalami peningkatan atau penurunan nilai akurasi saat ditambahkan *chi-square*.

Keywords: *SpeedID*, *Naïve Bayes (NB)*, *Support Vector Machine (SVM)*, *Chi-Square*, *Sentiment Analysis*

1. Pendahuluan

Hidup di era digital menyebabkan masyarakat semakin bergantung dengan teknologi dalam kegiatan sehari-hari mereka. Hal ini menjadi suatu peluang bagi perusahaan yang bergerak di bidang teknologi untuk hadir dengan berbagai macam inovasi untuk membantu masyarakat melakukan aktivitas secara daring. Salah satu aplikasi tersebut adalah SpeedID yang dirilis pada tahun 2018 sebagai solusi untuk kota cerdas. Dikutip dari situs resminya, SpeedID memiliki visi untuk menjadi solusi bagi penerapan IT dan identitas digital baru kota cerdas di seluruh dunia. Berbagai fitur yang dimiliki, mulai dari verifikasi identitas *online*, antrian *online*, manajemen usaha kuliner, manajemen usaha UKM, dan masih banyak lagi, membuat semakin banyak orang menggunakan aplikasi ini. Per April 2023, aplikasi ini telah diunduh sebanyak lebih dari 100 ribu kali. Hal ini juga berdampak pada semakin banyaknya ulasan yang diberikan oleh pengguna, baik itu positif, negatif, maupun netral. Dengan demikian, perlu dilakukan suatu analisis sentimen ulasan pengguna terhadap aplikasi ini guna mengetahui seberapa baik performa aplikasi dan apa yang dapat ditingkatkan oleh perusahaan ke depannya. Analisis sentimen adalah metode yang mengelompokkan polaritas dari data tekstual ke dalam kategori positif, netral, dan negatif [1]. Analisis sentimen ini dapat membantu dalam memperoleh wawasan yang berguna dalam pengambilan keputusan terkait pengembangan aplikasi, pemasaran, dan layanan pelanggan. Dengan demikian, perusahaan diharapkan dapat mengidentifikasi kekuatan dan kelemahan aplikasi mereka, serta meningkatkan kepuasan pengguna.

Ada berbagai macam metode yang dapat dilakukan dalam analisis sentimen. Umumnya, metode tersebut dibagi menjadi dua, yaitu pendekatan *supervised learning* dan *lexicon-based*. *Lexicon-based* merupakan pendekatan yang memanfaatkan kamus kata yang telah berisi definisi kata-kata untuk mencari polaritasnya. *Supervised learning* adalah metode yang bergantung kepada data latih yang telah memiliki label atau kelas [2]. Beberapa algoritma yang banyak digunakan dalam analisis sentimen adalah *Naïve Bayes* dan *Support Vector Machine*.

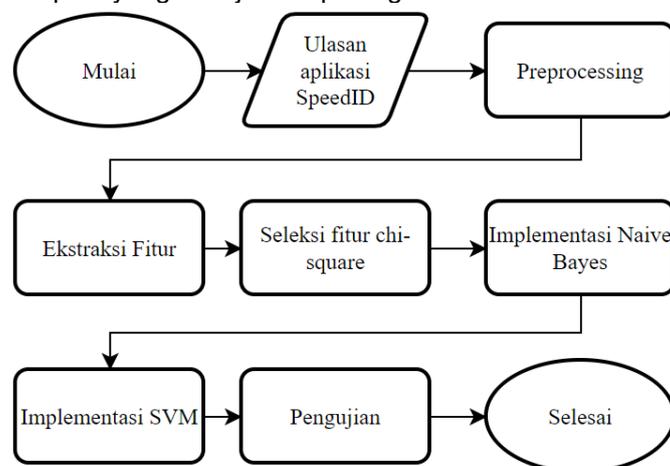
Naïve Bayes merupakan algoritma klasifikasi yang menggolongkan data ke dalam kategori yang tepat berdasarkan probabilitas dari data sebelumnya [3]. Salah satu kelebihan dari metode ini adalah dapat mencapai *accuracy* yang tinggi walaupun data latih yang digunakan sedikit [4]. Selain itu, metode *Naïve Bayes* juga mudah diimplementasikan dan dapat memberikan hasil yang baik pada berbagai kasus. Sedangkan, kekurangannya adalah fitur-fitur bersifat independen dan keterkaitan antar fitur tidak dapat dimodelkan oleh *Naïve Bayes* [5].

Support Vector Machine (SVM), merupakan metode pembelajaran mesin yang memanfaatkan hipotesis dalam bentuk fungsi linier dalam fitur dengan dimensi tinggi yang telah dilatih menggunakan algoritma pembelajaran berdasarkan teori optimisasi [6]. Kelebihan dari metode SVM adalah metode ini dapat menentukan *hyperplane* yang memaksimalkan margin untuk memisahkan kelas yang berbeda [7]. Selain itu, SVM juga mampu memberikan solusi untuk *overfitting* dan *optimal local*, serta memiliki rasio konvergensi yang rendah [8]. SVM juga memiliki kekurangan, yaitu data yang memiliki properti yang sama akan berpengaruh terhadap nilai akurasi secara signifikan [9].

Penelitian ini akan fokus pada analisis sentimen ulasan aplikasi SpeedID pada *Google Play Store* menggunakan metode *Naïve Bayes* dan *Support Vector Machine* (SVM) dengan menerapkan seleksi fitur *chi-square*. Seleksi fitur *chi-square*, merupakan metode yang digunakan untuk menguji ketergantungan dari dua kejadian [10]. Seleksi fitur ini diharapkan dapat mengurangi fitur yang kurang penting dan meningkatkan performa model algoritma dalam memprediksi kelas sentimen. Metode *Naïve Bayes* dan SVM akan dibandingkan untuk mengetahui algoritma mana yang memiliki performa lebih baik dalam klasifikasi sentimen.

2. Metode Penelitian

Tahapan penelitian dimulai dari pengumpulan data yang digunakan (ulasan aplikasi SpeedID), *preprocessing*, ekstraksi fitur, seleksi fitur, implementasi *Naïve Bayes*, implementasi SVM, dan pengujian atau evaluasi seperti yang ditunjukkan pada gambar 1.



Gambar 1. Alur Penelitian

2.1 Pengumpulan Data

Data yang digunakan merupakan ulasan aplikasi solusi kota cerdas, yaitu SpeedID. Data ini diperoleh dari 2 platform, yaitu *Google Play Store* dan *App Store* dengan cara *scraping*. Proses *scraping* dari *Play Store* menghasilkan data berbentuk file .csv yang terdiri dari 557 baris dan 11 kolom, yaitu *thumbsUpCount*, *reviewCreatedVersion*, *at*, *replyContent*, *repliedAt*, *reviewId*, *userName*, *userImage*, *content*, *score*, dan *appVersion*. Dari semua kolom tersebut, data yang digunakan adalah kolom *content* yang mengandung teks ulasan dan kolom *score* yang digunakan untuk menentukan label data.

Pengumpulan data dari platform *App Store* memperoleh data berupa file .csv yang terdiri dari 112 baris dan 7 kolom, yaitu *date*, *review*, *rating*, *isEdited*, *username*, *title*, dan *developerResponse*. Data yang digunakan adalah kolom *review* yang berisi teks ulasan dan *rating* untuk menentukan label sentimen.

2.2 Preprocessing Data

Sebelum masuk ke tahap implementasi algoritma, data mentah harus melalui tahap *preprocessing* terlebih dahulu. Tujuannya adalah untuk meningkatkan kualitas data dan memudahkan proses lebih lanjut. *Preprocessing* data terdiri dari beberapa bagian seperti berikut.

1. *Data cleansing*: Pembersihan atau penghilangan karakter-karakter tertentu, seperti tanda baca, *username*, atau url. Tahap ini juga meliputi *case folding*, yaitu mengubah semua karakter huruf dalam teks menjadi karakter huruf kecilnya.
2. Normalisasi kata: Mengubah kata-kata yang sebelumnya tidak baku menjadi bentuk kata bakunya.
3. *Stemming*: Mengubah kata-kata dalam teks menjadi bentuk dasar atau kata dasar, seperti mengubah kata "makanan", "makanan-makanan", dan "makan" menjadi kata dasar "makan".
4. *Stopwords removal*: Menghilangkan kata-kata yang umum dan tidak memiliki arti khusus, seperti "dan", "atau", dan "saja".

2.3 TF-IDF

Metode ekstraksi fitur yang digunakan adalah TF-IDF. TF-IDF merupakan metode yang menggabungkan antara *Term Frequency* (TF) dan *Inverse Document Frequency* (IDF). Metode ini akan menghasilkan nilai bobot yang tinggi untuk fitur yang banyak muncul pada suatu dokumen, tetapi jarang muncul pada kumpulan dokumen keseluruhan. Rumus untuk menghitung TF-IDF adalah berikut [11].

$$TF * IDF(d, t) = TF(d, t) * \log \frac{N}{df(t)} \quad (1)$$

Keterangan:

$TF * IDF(d, t)$ = bobot TF-IDF

$TF(d, t)$ = Frekuensi kemunculan fitur t pada dokumen d

N = Jumlah semua kumpulan dokumen

$df(t)$ = Jumlah dokumen yang mengandung fitur t

2.4 Chi-Square

Dalam statistika, *chi-square* sering digunakan untuk menguji ketergantungan antara dua kejadian [12]. Metode *chi-square* juga merupakan salah satu metode yang banyak digunakan dalam melakukan seleksi fitur. Metode ini dikatakan mampu menghapus fitur-fitur yang tidak dibutuhkan tanpa mengurangi nilai *accuracy* yang diperoleh [13]. Rumus *chi-square* dijabarkan sebagai berikut [14].

$$X^2(t, c) = \frac{N(AD-CB)^2}{(A+C)(B+D)(A+B)(C+D)} \quad (2)$$

Keterangan:

X^2 = nilai *chi-square*

t = term atau fitur

c = kelas atau label

N = jumlah semua dokumen

A = jumlah kemunculan t dalam kelas c

B = jumlah kemunculan t dalam kelas selain c

C = jumlah kemunculan kata selain t dalam kelas c

D = jumlah kemunculan kata selain t dalam kelas selain c

2.5 Naïve Bayes (NB)

Algoritma *Naïve Bayes* merupakan algoritma klasifikasi yang bisa digunakan untuk memprediksi probabilitas kelas berdasarkan teorema Bayes [10]. Prinsip dasar algoritma *Naïve Bayes* adalah menghitung probabilitas posterior untuk setiap kelas target berdasarkan kemunculan fitur atau atribut dalam dokumen. Probabilitas posterior adalah probabilitas kelas target setelah melihat data. Berikut adalah rumus untuk mencarinya [15].

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)} \quad (3)$$

Keterangan:

$P(A|B)$ = probabilitas dari atribut B masuk ke kelas A

$P(B|A)$ = probabilitas dari munculnya atribut B pada kelas A

$P(A)$ = probabilitas dari data yang termasuk kelas A

$P(B)$ = jumlah semua kata yang ada di dataset

$P(B|A)$, disebut juga sebagai probabilitas kondisional, dapat dihitung dengan menggunakan rumus berikut.

$$P(B|A) = P(B_1|A) \times P(B_2|A) \times \dots \times P(B_d|A) \quad (4)$$

$P(B_1|A)$ merupakan probabilitas dari munculnya fitur B_1 pada kelas A. Probabilitas munculnya fitur tertentu pada suatu kelas A dihitung menggunakan persamaan berikut.

$$P(B_i|A) = \frac{\text{bobot TF-IDF fitur } B_i \text{ di kelas A}}{\text{total bobot TF-IDF di kelas A}} \quad (5)$$

$P(A)$ atau probabilitas prior dapat diperoleh menggunakan persamaan berikut.

$$P(A) = \frac{\text{jumlah dokumen pada kelas A}}{\text{jumlah semua dokumen pada training set}} \quad (6)$$

2.6 Support Vector Machine (SVM)

SVM (*Support Vector Machine*) adalah algoritma *machine learning* yang digunakan untuk melakukan klasifikasi dan regresi pada data yang berdimensi tinggi. SVM bekerja dengan membuat sebuah *hyperplane* yang memisahkan antara dua kelas data yang berbeda secara maksimal [16]. Algoritma ini dapat digunakan untuk klasifikasi data yang memiliki lebih dari 2 kelas dengan menerapkan strategi *one-vs-all*. Strategi ini memberikan satu *classifier* per kelasnya. Untuk masing-masing *classifier*, satu kelas dipisahkan dengan kelas lainnya [17].

SVM dapat bekerja pada data yang tidak linier dengan memanfaatkan penggunaan kernel. Kernel adalah fungsi transformasi yang mengubah data ke dalam dimensi yang lebih tinggi, sehingga memungkinkan SVM untuk membuat *hyperplane* yang dapat memisahkan data yang tidak linier [18]. Beberapa kernel yang banyak digunakan adalah linear kernel, *polynomial* kernel, *sigmoid* kernel, dan *radial basis function* (RBF) kernel.

1. Linear kernel

Kernel ini adalah kernel SVM paling sederhana yang cocok digunakan ketika kelas data dapat dipisah secara linear. Linear kernel adalah kernel yang digunakan secara default pada pemodelan SVM [19].

$$f(x) = w^T x + b \quad (7)$$

2. Polynomial kernel

Polynomial kernel banyak digunakan saat semua data latih telah dinormalisasi.

$$K_{x_i x} = (x_i^T x + 1)^d \quad (8)$$

3. Sigmoid kernel

Sigmoid kernel merupakan pengembangan dari jaringan saraf tiruan yang dinyatakan dalam rumus berikut.

$$K_{x_i x} = \tanh(\gamma x_i^T x + r) \quad (9)$$

4. Radial basis function (RBF)

RBF kernel menggunakan 2 parameter untuk optimalisasi SVM, yaitu *Gamma* dan *Cost*. *Gamma* merupakan nilai dari satu data, sedangkan *Cost* merupakan variabel yang bertujuan untuk optimalisasi informasi data [19].

$$K_{x_i x} = \exp[-\gamma \|x_i - x\|^2] \quad (10)$$

2.7 Confusion Matrix

Confusion matrix adalah alat yang digunakan untuk mengukur kinerja model klasifikasi. Di dalamnya, terdapat informasi mengenai nilai yang diprediksi oleh sistem dan nilai data yang sebenarnya [20].

Tabel 1. Confusion Matrix

		Kelas Prediksi	
		Positif	Negatif
	Positif		
	Negatif		

Kelas Sebenarnya	Positif	TP	FN
	Negatif	FP	TN

Keterangan:

TP (*True Positive*) = data kelas positif yang diprediksi positif

FP (*False Positive*) = data kelas negatif yang diprediksi positif

FN (*False Negative*) = data kelas positif yang diprediksi negatif

TN (*True Negative*) = data kelas negatif yang diprediksi negatif

Dari *confusion matrix* yang dapat dilihat pada tabel 1, dapat diperoleh kinerja algoritma dengan menghitung nilai *accuracy*, *precision*, *recall*, dan *F1-score* dengan rumus berikut [21].

$$accuracy = \frac{TP+TN}{Total} \tag{11}$$

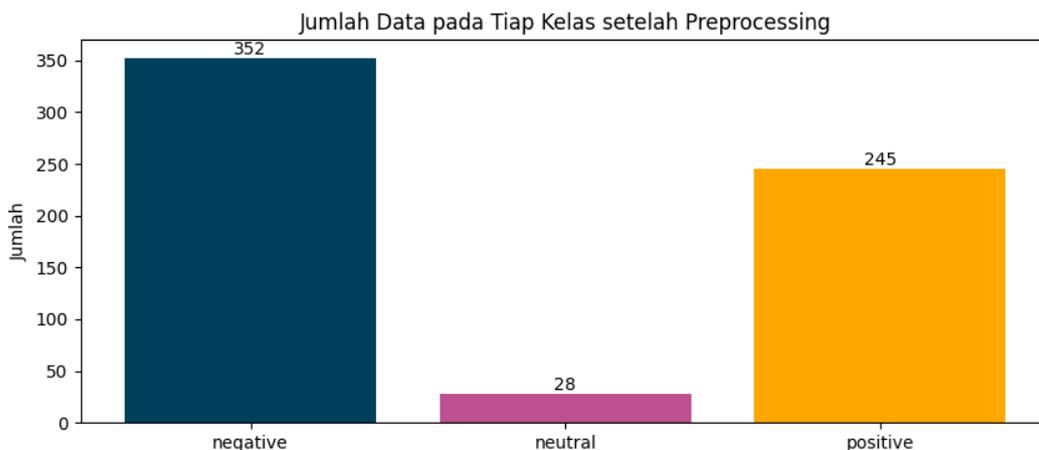
$$precision = \frac{TP}{TP+FP} \tag{12}$$

$$recall = \frac{TP}{TP+FN} \tag{13}$$

$$F1\ score = 2 * \frac{recall*precision}{recall+precision} \tag{14}$$

3. Hasil dan Pembahasan

Total data yang diperoleh dari tahap pengumpulan data adalah 669 data ulasan aplikasi SpeedID yang berasal dari *Play Store* dan *App Store*. Setelah melewati tahap preprocessing, terdapat sebanyak 44 data yang kosong sehingga data yang bisa digunakan adalah 625 data dengan distribusi kelas yang ditunjukkan pada gambar 2.



Gambar 2. Distribusi Kelas setelah *Preprocessing*

Contoh hasil dari *preprocessing* data ditunjukkan pada tabel 2.

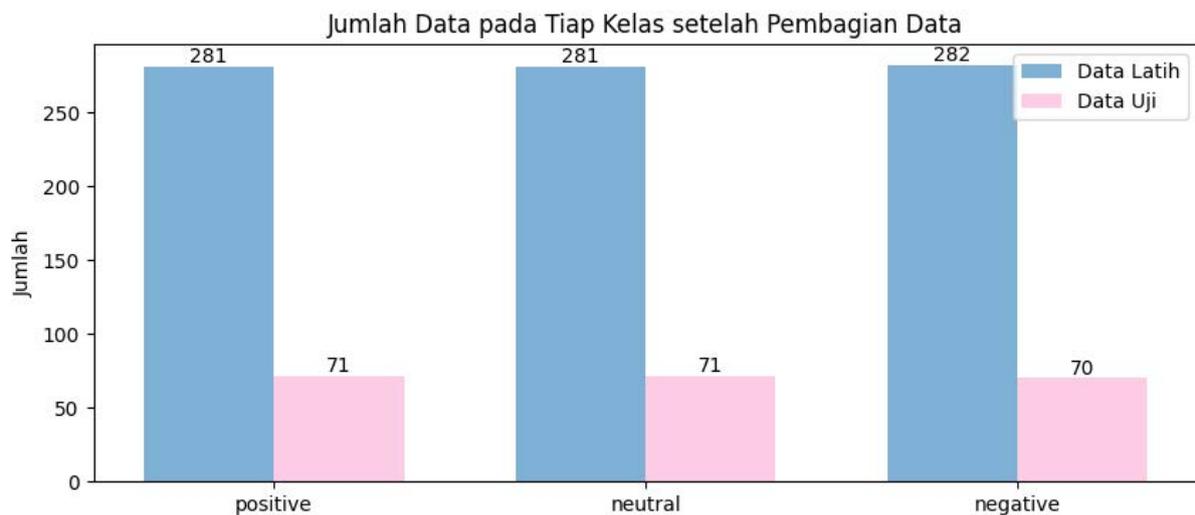
Tabel 2. Hasil *Preprocessing*

Sebelum Preprocessing	Setelah Preprocessing
Gak ada kelebihannya,cuma untuk daftar cari no antrean,,,,,, no antrean juga tidak sesuai dengan yang tertera dilayar □□□□,,,,,, ngantre di bank bpd paling lama,,,keteller nunggu lagi 9 orang perlu waktu 1 jam lebih□□□□□□□□,,ada yang nyalib lewat samping juga□□□□ harus perlu trening ke bank swasta untuk perbandingan kinerja ,,,,,, LOKASI BPD UBUNG	daftar cari antre antre sesuai tera layar antri bank bpd keteller tunggu orang waktu jam nyalib samping trening bank swasta banding kerja lokasi bpd ubung
Kenapa eror terus ya dari kemarin, tolong dong di perbaiki, padahal saya selalu menggunakan speedQ untuk daftar antrean rumah sakit,	eror kemarin speedq daftar antre rumah sakit
Aplikasi masih jauh dari efektif.. cara pendaftaran tidak praktis.. telalu bertele-tele. Boleh ribet tapi aplikasi harus bisa	aplikasi efektif daftar praktis talu tele tele ribet aplikasi simpan data

nyimpen data, supaya tidak menetik berulang ulang.. Ini bukan speedid namanya.. tapi SLOWID..	etik ulang ulang speedid nama slowid
---	--------------------------------------

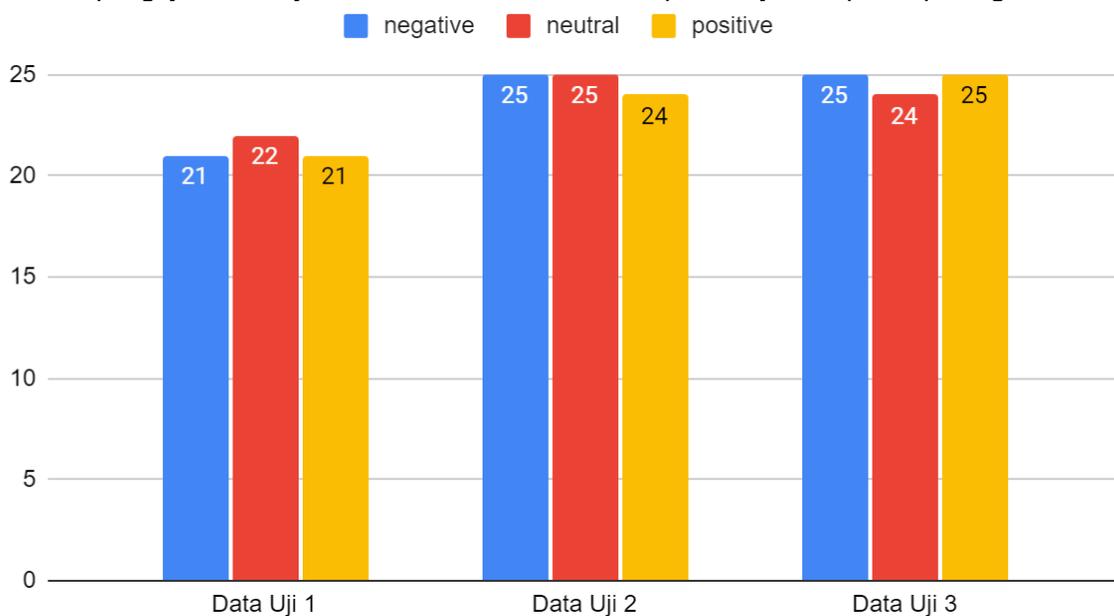
Karena jumlah kelas data yang tidak seimbang, dilakukan *over sampling* untuk menyeimbangkan data sebelum diproses dalam klasifikasi. *Over sampling* diterapkan secara acak pada data latih. *Over sampling* ini mengambil sampel pada kelas positif dan netral secara acak agar seimbang dengan kelas mayoritas, yaitu negatif. Data yang awalnya berjumlah 625, kini menjadi 1056 setelah dilakukan *over sampling* dengan jumlah data pada masing-masing kelas yaitu 352.

Data yang telah melalui tahap *over sampling* kemudian dibagi menjadi data latih sebanyak 80% dan data uji sebanyak 20% dari total data 1056. Data latih yang digunakan dalam pelatihan model klasifikasi terdiri dari 844 dan data uji yang digunakan dalam pengujian performa model terdiri dari 212 data. Adapun distribusi jumlah data pada masing-masing kelas di data latih dan data uji ditunjukkan melalui gambar 3 di bawah.



Gambar 3. Distribusi Kelas Data Latih dan Data Uji

Selain membagi data menjadi data latih dan data uji, data uji juga dibagi menjadi 3 bagian untuk melakukan pengujian sebanyak 3 kali. Distribusi kelas di tiap data uji ditampilkan pada gambar 4.



Gambar 4. Distribusi Kelas Data Uji

Setelah mendapatkan data latih yang seimbang di tiap kelasnya, tahap selanjutnya adalah melatih model klasifikasi NB dan SVM menggunakan data latih. Masing-masing pelatihan model dilakukan dalam 2 skenario, yaitu tanpa menggunakan seleksi fitur dan dengan menggunakan seleksi fitur *chi-square*. Model terbaik pada masing-masing skenario dinilai berdasarkan nilai akurasi tertinggi pada data latih. Model tersebut kemudian dievaluasi menggunakan data uji pada tahap evaluasi.

Pelatihan model tanpa seleksi fitur meliputi *hyperparameter tuning* untuk mengetahui kombinasi nilai parameter yang menghasilkan performa terbaik. Setelah itu, kombinasi *hyperparameter* terbaik tersebut digunakan untuk eksperimen lebih lanjut menggunakan seleksi fitur *chi-square*. Adapun nilai yang diuji dalam pelatihan model dengan *chi-square* adalah jumlah fitur yang dipertahankan seperti yang ditampilkan pada tabel 3.

Tabel 3. Eksperimen *Chi-Square*

Persentase fitur	Jumlah fitur
20%	174
40%	349
60%	524
80%	699

Dalam pelatihan model NB tanpa seleksi fitur, *hyperparameter* yang di-*tuning* adalah *alpha*, yaitu parameter untuk menangani agar probabilitas tidak bernilai 0. Hasil dari pelatihan model NB tanpa *chi-square* ditampilkan pada gambar 5.

Accuracy: 0.8626, alpha: 0.1
Accuracy: 0.8566, alpha: 0.5
Accuracy: 0.8531, alpha: 1
Accuracy: 0.8531, alpha: 1.5
Accuracy: 0.8507, alpha: 2

Gambar 5. Hasil Pelatihan NB

Terlihat bahwa akurasi data latih tertinggi, yaitu 86,26% diperoleh saat nilai *alpha*=0,1 sehingga nilai *hyperparameter* ini yang digunakan untuk melakukan eksperimen lebih lanjut dengan *chi-square*. Adapun hasil dari pelatihan model NB dengan *chi-square* dapat dilihat pada gambar 6 berikut.

Jumlah fitur = 174
Accuracy: 0.8400, alpha: 0.1
Jumlah fitur = 349
Accuracy: 0.8839, alpha: 0.1
Jumlah fitur = 524
Accuracy: 0.8791, alpha: 0.1
Jumlah fitur = 699
Accuracy: 0.8685, alpha: 0.1

Gambar 6. Hasil Pelatihan NB + *Chi Square*

Model NB terbaik diperoleh dengan *hyperparameter alpha*=0,1 dan akurasi data latih sebesar 86,26%. Saat dievaluasi menggunakan data uji sebanyak 3 kali, model NB tersebut menghasilkan performa yang dapat dilihat pada tabel 4.

Tabel 4. Hasil Pengujian Model Terbaik NB

Data	Akurasi (%)	Precision (%)	Recall (%)	F1-Score (%)
Data uji 1	92,19	93	92	91,67
Data uji 2	82,43	84,67	82,33	82,67
Data uji 3	83,78	84,33	84	84

Model NB + Chi Square terbaik memperoleh akurasi data latih sebesar 88,39% dengan mempertahankan jumlah fitur 40%. Adapun performa model terbaik NB dengan *chi-square* selengkapnya dapat dilihat pada tabel 5 berikut.

Tabel 5. Hasil Pengujian Model Terbaik NB + *Chi Square*

Data	Akurasi (%)	Precision (%)	Recall (%)	F1-Score (%)
Data uji 1	95,31	95,33	95	95
Data uji 2	82,43	84,67	82,33	82,67
Data uji 3	85,13	85,67	85,33	85,33

Secara keseluruhan, perbandingan performa kedua model NB dapat dilihat pada tabel 6 berikut.

Tabel 6. Perbandingan Akurasi (%) Model NB

Model	Data Uji 1	Data Uji 2	Data Uji 3
NB	92,19	82,43	83,78
NB + <i>chi square</i>	95,31	82,43	85,13

Berdasarkan tabel 6, seleksi fitur *chi-square* dapat meningkatkan nilai akurasi dari model NB sebanyak 2 kali. Nilai akurasi data uji 1 meningkat sebesar 3,12%, sedangkan akurasi data uji 3 meningkat sebesar 1.35%. Hal ini menunjukkan bahwa seleksi fitur *chi-square* memiliki pengaruh positif terhadap performa model NB.

Dalam pelatihan model SVM tanpa *chi-square*, *hyperparameter* yang diuji dapat dilihat pada tabel 7 berikut.

Tabel 7. Nilai *Hyperparameter* SVM yang Diuji

<i>Hyperparameter</i>	Nilai Parameter	Kernel
C	0,01; 0,1; 1; 10; 100	Linear, RBF, <i>sigmoid</i> , <i>polynomial</i>
Gamma	0,01; 0,1; 1; 10; 100	RBF, <i>sigmoid</i> , <i>polynomial</i>
Degree	2, 3, 4	<i>Polynomial</i>

Pelatihan model linear SVM tanpa seleksi fitur ini menghasilkan nilai akurasi pada data latih yang dapat dilihat pada gambar 7.

```
Accuracy: 0.4797, Parameters: {'C': 0.01, 'kernel': 'linear'}
Accuracy: 0.7559, Parameters: {'C': 0.1, 'kernel': 'linear'}
Accuracy: 0.8803, Parameters: {'C': 1, 'kernel': 'linear'}
Accuracy: 0.8886, Parameters: {'C': 10, 'kernel': 'linear'}
Accuracy: 0.8709, Parameters: {'C': 100, 'kernel': 'linear'}
Best Parameters: {'C': 10, 'kernel': 'linear'}
```

Gambar 7. Hasil Pelatihan Linear SVM

Pelatihan linear SVM dengan *chi-square* selanjutnya dilakukan menggunakan nilai parameter C=10 dengan hasil yang ditampilkan pada gambar 8.

```
Accuracy: 0.8839, Parameters: {'C': 10, 'kernel': 'linear'}
Best Parameters: {'C': 10, 'kernel': 'linear'}
```

```
Accuracy: 0.8898, Parameters: {'C': 10, 'kernel': 'linear'}
Best Parameters: {'C': 10, 'kernel': 'linear'}
```

```
Accuracy: 0.8839, Parameters: {'C': 10, 'kernel': 'linear'}
Best Parameters: {'C': 10, 'kernel': 'linear'}
```

```
Accuracy: 0.8886, Parameters: {'C': 10, 'kernel': 'linear'}
Best Parameters: {'C': 10, 'kernel': 'linear'}
```

Gambar 8. Hasil Pelatihan Linear SVM + *Chi Square*

Untuk pelatihan model SVM dengan kernel *Radial Basis Function* (RBF), *hyperparameter* yang digunakan adalah C dan gamma dengan hasil pelatihan seperti pada tabel 8.

Tabel 8. Nilai Akurasi (%) Hasil Pelatihan RBF SVM

C	Gamma				
	0,01	0,1	1	10	100
0,01	48,09	48,68	47,38	44,30	44,30
0,1	48,09	50,10	71,33	64,33	64,33
1	49,51	80,57	89,81	84,36	84,36
10	81,16	89,45	89,81	84,48	84,36
100	88,62	88,63	89,81	84,48	84,36

Berdasarkan tabel 8, dapat dilihat bahwa nilai akurasi terbaik diperoleh ketika nilai C=10 dan gamma=1. Nilai *hyperparameter* tersebut kemudian digunakan dalam eksperimen dengan *chi-square* dengan hasil seperti pada gambar 9.

Accuracy: 0.8851, Parameters: {'C': 10, 'gamma': 1, 'kernel': 'rbf'}
 Best Parameters: {'C': 10, 'gamma': 1, 'kernel': 'rbf'}

Accuracy: 0.8981, Parameters: {'C': 10, 'gamma': 1, 'kernel': 'rbf'}
 Best Parameters: {'C': 10, 'gamma': 1, 'kernel': 'rbf'}

Accuracy: 0.8981, Parameters: {'C': 10, 'gamma': 1, 'kernel': 'rbf'}
 Best Parameters: {'C': 10, 'gamma': 1, 'kernel': 'rbf'}

Accuracy: 0.8993, Parameters: {'C': 10, 'gamma': 1, 'kernel': 'rbf'}
 Best Parameters: {'C': 10, 'gamma': 1, 'kernel': 'rbf'}

Gambar 9. Hasil Pelatihan RBF SVM + *Chi Square*

Pada pelatihan *sigmoid SVM*, *hyperparameter* yang digunakan adalah C dan gamma juga dengan hasil pelatihan yang ditunjukkan pada tabel 9.

Tabel 9. Nilai Akurasi (%) Hasil Pelatihan *Sigmoid SVM*

C	Gamma				
	0,01	0,1	1	10	100
0,01	47,97	47,97	47,85	45,85	50,12
0,1	47,97	47,97	75	65,16	56,40
1	47,97	75,59	85,54	59,24	53,79
10	75,59	88,03	83,77	56,99	54,15
100	88,03	88,86	84,12	58,30	55,68

Model dengan kombinasi nilai parameter terbaik pada *Sigmoid SVM*, yaitu C=100 dan gamma=0,1. Nilai *hyperparameter* tersebut kemudian digunakan dalam eksperimen dengan *chi-square* dengan hasil pelatihan seperti pada gambar 10.

Accuracy: 0.8815, Parameters: {'C': 100, 'gamma': 0.1, 'kernel': 'sigmoid'}
 Best Parameters: {'C': 100, 'gamma': 0.1, 'kernel': 'sigmoid'}

Accuracy: 0.8898, Parameters: {'C': 100, 'gamma': 0.1, 'kernel': 'sigmoid'}
 Best Parameters: {'C': 100, 'gamma': 0.1, 'kernel': 'sigmoid'}

Accuracy: 0.8839, Parameters: {'C': 100, 'gamma': 0.1, 'kernel': 'sigmoid'}
 Best Parameters: {'C': 100, 'gamma': 0.1, 'kernel': 'sigmoid'}

Accuracy: 0.8886, Parameters: {'C': 100, 'gamma': 0.1, 'kernel': 'sigmoid'}
 Best Parameters: {'C': 100, 'gamma': 0.1, 'kernel': 'sigmoid'}

Gambar 10. Hasil Pelatihan *Sigmoid SVM* + *Chi Square*

Pada pelatihan *polynomial SVM*, *hyperparameter* yang digunakan adalah C, gamma, dan *degree* dengan hasil pelatihan yang ditunjukkan di tabel 10, 11, dan 12.

Tabel 10. Nilai Akurasi (%) Hasil Pelatihan Polynomial SVM (Degree=2)

C	Gamma				
	0,01	0,1	1	10	100
0,01	46,31	46,31	46,31	88,86	89,10
0,1	46,31	46,31	71,57	89,10	89,10
1	46,31	46,31	88,86	89,10	89,10
10	46,31	71,57	89,10	89,10	89,10
100	46,31	88,86	89,10	89,10	89,10

Tabel 11. Nilai Akurasi (%) Hasil Pelatihan Polynomial SVM (Degree=3)

C	Gamma				
	0,01	0,1	1	10	100
0,01	45,36	45,36	45,36	87,08	87,08
0,1	45,36	45,36	69,67	87,08	87,08
1	45,36	45,36	86,97	87,08	87,08
10	45,36	45,36	87,08	87,08	87,08
100	45,36	69,67	87,08	87,08	87,08

Tabel 12. Nilai Akurasi (%) Hasil Pelatihan Polynomial SVM (Degree=4)

C	Gamma				
	0,01	0,1	1	10	100
0,01	45,01	45,01	45,01	86,61	86,61
0,1	45,01	45,01	68,60	86,61	86,61
1	45,01	45,01	86,61	86,61	86,61
10	45,01	45,01	86,61	86,61	86,61
100	45,01	45,01	86,61	86,61	86,61

Berdasarkan tabel 10, 11, dan 12 di atas, model dengan kombinasi nilai parameter terbaik pada *polynomial SVM*, yaitu C=0,01; gamma=100; dan *degree*=2. Nilai *hyperparameter* tersebut kemudian digunakan dalam eksperimen dengan *chi-square* dengan hasil seperti pada gambar 11.

Accuracy: 0.8720, Parameters: {'C': 0.01, 'degree': 2, 'gamma': 100, 'kernel': 'poly'}
 Best Parameters: {'C': 0.01, 'degree': 2, 'gamma': 100, 'kernel': 'poly'}

Accuracy: 0.8637, Parameters: {'C': 0.01, 'degree': 2, 'gamma': 100, 'kernel': 'poly'}
 Best Parameters: {'C': 0.01, 'degree': 2, 'gamma': 100, 'kernel': 'poly'}

Accuracy: 0.8756, Parameters: {'C': 0.01, 'degree': 2, 'gamma': 100, 'kernel': 'poly'}
 Best Parameters: {'C': 0.01, 'degree': 2, 'gamma': 100, 'kernel': 'poly'}

Accuracy: 0.8862, Parameters: {'C': 0.01, 'degree': 2, 'gamma': 100, 'kernel': 'poly'}
 Best Parameters: {'C': 0.01, 'degree': 2, 'gamma': 100, 'kernel': 'poly'}

Gambar 11. Hasil Pelatihan *Polynomial SVM* + *Chi Square*

Untuk menentukan 1 model SVM tanpa *chi-square* terbaik dan 1 model SVM dengan *chi square* terbaik, nilai akurasi data latih model SVM dirangkum pada tabel 13 di bawah.

Tabel 13. Rangkuman Hasil Pelatihan Model SVM

Model	Akurasi Data Latih (%)
-------	------------------------

Linear SVM	88,86
Linear SVM + <i>chi square</i>	88,98
<i>Sigmoid SVM</i>	88,86
<i>Sigmoid SVM + chi square</i>	88,98
RBF SVM	89,81
RBF SVM + <i>chi square</i>	89,93
<i>Polynomial SVM</i>	89,10
<i>Polynomial SVM + chi square</i>	88,62

Tabel 13 menunjukkan bahwa model terbaik dengan nilai akurasi data latih tertinggi adalah model SVM dengan kernel RBF yang memiliki *hyperparameter* C=10 dan gamma=1. Adapun performa model SVM terbaik selengkapny dapat dilihat pada tabel 14 berikut.

Tabel 14. Performa Model Terbaik SVM

Data	Akurasi (%)	Precision (%)	Recall (%)	F1-Score (%)
Data uji 1	95,31	95,33	95	95,33
Data uji 2	83,78	85,67	84	84
Data uji 3	89,19	89,67	89,33	89,33

Model terbaik SVM dengan *chi-square* yang dihasilkan adalah model yang mempertahankan 80% jumlah fitur yang mencapai nilai akurasi data latih sebesar 89,93%. Performa model SVM dengan *chi-square* selengkapny dapat dilihat pada tabel 15 berikut.

Tabel 15. Performa Model Terbaik SVM + *Chi Square*

Data	Akurasi (%)	Precision (%)	Recall (%)	F1-Score (%)
Data uji 1	95,31	95,33	95	95,33
Data uji 2	83,78	85,67	84	84
Data uji 3	89,19	89,67	89,33	89,33

Secara keseluruhan, perbandingan performa kedua model SVM dapat dilihat pada tabel 16.

Tabel 16. Perbandingan Akurasi (%) Model SVM

Model	Data Uji 1	Data Uji 2	Data Uji 3
SVM	95,31	83,78	89,19
SVM + <i>chi square</i>	95,31	83,78	89,19

Berdasarkan tabel 16 di atas, seleksi fitur *chi-square* sama sekali tidak dapat mengubah nilai akurasi dari model SVM. Nilai akurasi data uji 1, 2, dan 3 tidak mengalami peningkatan maupun penurunan saat ditambahkan *chi-square*. Hal ini menunjukkan bahwa seleksi fitur *chi-square* tidak memiliki pengaruh positif maupun negatif terhadap performa model SVM.

4. Kesimpulan

Penelitian ini bertujuan untuk mengetahui performa khususnya nilai akurasi dari model *Naive Bayes* (NB) dan *Support Vector Machine* (SVM) dalam analisis sentimen ulasan aplikasi SpeedID serta mengetahui pengaruh dari seleksi fitur *chi square* terhadap performa model NB dan SVM tersebut. Adapun beberapa kesimpulan yang dapat diambil dari hasil penelitian ini adalah sebagai berikut.

1. Model NB terbaik diperoleh dengan *hyperparameter* $\alpha=0,1$ yang menghasilkan nilai akurasi sebesar 92,18%; 82,43%; dan 83,78%, sedangkan model SVM terbaik diperoleh dengan kernel *Radial Basis Function* (RBF), C=10, dan gamma=1 yang menghasilkan nilai akurasi sebesar 95,31%; 83,78%; dan 89,18 yang semuanya lebih tinggi dibandingkan dengan model NB.
2. Dari 3 kali pengujian yang dilakukan pada tiap model terbaik, seleksi fitur *chi-square* terbukti dapat meningkatkan nilai akurasi model NB hingga 3,12%. Namun, nilai akurasi model SVM

tidak mengalami peningkatan maupun penurunan saat ditambahkan dengan seleksi fitur *chi-square*. Dengan demikian, seleksi fitur *chi-square* disimpulkan memiliki pengaruh positif terhadap performa NB, tetapi tidak memiliki pengaruh terhadap performa SVM.

References

- [1] L. Ardiani, H. Sujaini, dan T. Tursina, "Implementasi Sentiment Analysis Tanggapan Masyarakat Terhadap Pembangunan di Kota Pontianak," *Jurnal Sistem dan Teknologi Informasi (Justin)*, vol. 8, no. 2, hlm. 183–190, Apr 2020, doi: 10.26418/justin.v8i2.36776.
- [2] M. Cindo, D. P. Rini, dan Ermatita, "Literatur Review: Metode Klasifikasi Pada Sentimen Analisis," *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)*, hlm. 66–70, Jan 2019.
- [3] F. S. Pattiiha dan Hendry, "Perbandingan Metode K-NN, Naïve Bayes, Decision Tree untuk Analisis Sentimen Tweet Twitter Terkait Opini Terhadap PT PAL Indonesia," *JURIKOM (Jurnal Riset Komputer)*, vol. 9, no. 2, hlm. 506–514, Apr 2022.
- [4] R. Setiawan dan A. Triayudi, "Klasifikasi Status Gizi Balita Menggunakan Naïve Bayes dan K-Nearest Neighbor Berbasis Web," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 6, no. 2, hlm. 777–785, Apr 2022, doi: 10.30865/mib.v6i2.3566.
- [5] F. Solikhah, M. Febianah, A. L. Kamil, W. A. Arifin, dan S. J. S. Tyas, "Analisis Perbandingan Algoritma Naive Bayes Dan C.45 Dalam Klasifikasi Data Mining Untuk Memprediksi Kelulusan," *Tematik : Jurnal Teknologi Informasi Komunikasi (e-Journal)*, vol. 8, no. 1, hlm. 96–103, Jun 2021.
- [6] I. M. Parapat, M. T. Furqon, dan Sutrisno, "Penerapan Metode Support Vector Machine (SVM) Pada Klasifikasi Penyimpangan Tumbuh Kembang Anak," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 2, no. 10, hlm. 3163–3169, Okt 2018.
- [7] F. Bei dan S. Saepudin, "ANALISIS SENTIMEN APLIKASI TIKET ONLINE DI PLAY STORE MENGGUNAKAN METODE SUPPORT VECTOR MACHINE (SVM)," *SISMATIK (Seminar Nasional Sistem Informasi dan Manajemen Informatika)*, vol. 1, no. 1, hlm. 91–97, Agu 2021.
- [8] V. K. S. Que, A. Iriani, dan H. D. Purnomo, "Analisis Sentimen Transportasi Online Menggunakan Support Vector Machine Berbasis Particle Swarm Optimization," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 9, no. 2, hlm. 162–170, Mei 2020.
- [9] S. I. Nurhafida dan F. Sembiring, "Analisis Sentimen Aplikasi Novel Online Di Google Play Store Menggunakan Algoritma Support Vector Machine (SVM)," *Jurnal Sains Komputer & Informatika (J-SAKTI)*, vol. 6, no. 1, hlm. 317–327, Mar 2022.
- [10] A. Falasari dan M. A. Muslim, "Optimize Naïve Bayes Classifier Using Chi Square and Term Frequency Inverse Document Frequency For Amazon Review Sentiment Analysis," *Journal of Soft Computing Exploration*, vol. 3, no. 1, hlm. 31–36, Mar 2022, doi: 10.52465/josce.v3i1.68.
- [11] C. J. E. Munthe, N. A. Hasibuan, dan H. Hutabarat, "Penerapan Algoritma Text Mining Dan TF-RF Dalam Menentukan Promo Produk Pada Marketplace," *Resolusi : Rekayasa Teknik Informatika dan Informasi*, vol. 2, no. 3, hlm. 110–115, Jan 2022, doi: 10.30865/resolusi.v2i3.309.
- [12] U. I. Larasati, M. A. Muslim, R. Arifudin, dan A. Alamsyah, "Improve the Accuracy of Support Vector Machine Using Chi Square Statistic and Term Frequency Inverse Document Frequency on Movie Review Sentiment Analysis," *Scientific Journal of Informatics*, vol. 6, no. 1, hlm. 138–149, Mei 2019, doi: 10.15294/sji.v6i1.14244.
- [13] L. Luthfiana, J. C. Young, dan A. Rusli, "Implementasi Algoritma Support Vector Machine dan Chi Square untuk Analisis Sentimen User Feedback Aplikasi," *Ultimatics : Jurnal Teknik Informatika*, vol. 12, no. 2, hlm. 125–126, Des 2020, doi: 10.31937/ti.v12i2.1828.
- [14] W. Winata, A. Zaidiah, dan N. Chamidah, "ANALISIS SENTIMEN PADA ULASAN PRODUK MASKER DI MARKETPLACE SHOPEE MENGGUNAKAN SUPPORT VECTOR MACHINE DAN SELEKSI FITUR CHI SQUARE," *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)*, Agu 2022.

- [15] M. Christianto, J. Andjarwirawan, dan A. Tjondrowiguno, "Aplikasi Analisa Sentimen Pada Komentar Berbahasa Indonesia Dalam Objek Video di Website YouTube Menggunakan Metode Naïve Bayes Classifier," *JURNAL INFRA*, vol. 8, no. 1, 2020.
- [16] R. W. Pratiwi, S. F. H, D. Dairoh, D. I. Af'idah, Q. R. A, dan A. G. F, "Analisis Sentimen Pada Review Skincare Female Daily Menggunakan Metode Support Vector Machine (SVM)," *Journal of Informatics, Information System, Software Engineering and Applications (INISTA)*, vol. 4, no. 1, hlm. 40–46, Des 2021, doi: 10.20895/inista.v4i1.387.
- [17] S. Chatterjee, P. George Jose, dan D. Datta, "Text Classification Using SVM Enhanced by Multithreading and CUDA," *International Journal of Modern Education and Computer Science*, vol. 11, no. 1, hlm. 11–23, Jan 2019, doi: 10.5815/ijmecs.2019.01.02.
- [18] A. Z. Praghakusma dan N. Charibaldi, "Komparasi Fungsi Kernel Metode Support Vector Machine untuk Analisis Sentimen Instagram dan Twitter (Studi Kasus : Komisi Pemberantasan Korupsi) ," *Jurnal Sarjana Teknik Informatika*, vol. 9, no. 2, 2021.
- [19] A. Zeputra dan F. Utamingrum, "Perbandingan Akurasi untuk Deteksi Pintu berbasis HOG dengan Klasifikasi SVM menggunakan Kernel Linear, Radial Basis Function dan Polinomial pada Raspberry Pi," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 5, no. 11, hlm. 4746–4757, Nov 2021.
- [20] F. Novitasari dan M. D. Purbolaksono, "Sentiment Analysis Aspect Level on Beauty Product Reviews Using Chi-Square and Naïve Bayes," *JOURNAL OF DATA SCIENCE AND ITS APPLICATIONS*, vol. 4, no. 1, hlm. 18–30, Jan 2021.
- [21] D. Normawati dan S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," *Jurnal Sains Komputer & Informatika (J-SAKTI)* , vol. 5, no. 2, hlm. 697–711, Sep 2021.

This page is intentionally left blank.

Deteksi Rasa Buah Jeruk Siam Kintamani Menggunakan SVM dengan Optimasi Algoritma Genetika

Ni Wayan Yulia Damayanti^{a1}, I Gede Arta Wibawa^{a2}, I Gede Santi Astawa^{b3},
Anak Agung Istri Ngurah Eka Karyawati^{b4}

^aProgram Studi Informatika, Universitas Udayana
Kuta Selatan, badung, Bali, Indonesia
¹yuliadamayanti2016@gmail.com
²Gede.arta@unud.ac.id

^bProgram Studi Informatika, Universitas Udayana
Kuta Selatan, badung, Bali, Indonesia
³santi.astawa@unud.ac.id@gmail.com
⁴eka.karyawati@unud.ac.id@unud.ac.id

Abstract

The Siam Kintamani orange is one of the vital commodities in Indonesian agriculture, particularly in Bangli Regency, Bali. However, the evaluation of orange taste quality often relies on subjective manual identification. In an effort to enhance objectivity and consistency in assessing orange quality, this research proposes the utilization of the Support Vector Machine (SVM) algorithm optimized with genetic algorithms. This study aims to discern the taste quality of Siam Kintamani oranges based on texture features in orange images. Test results indicate that SVM optimized with genetic algorithms achieves superior accuracy compared to SVM without optimization. For instance, SVM without optimization yields an accuracy of 0.78, while post-optimization with genetic algorithms increases accuracy to 0.80. These findings underscore the significant potential of genetic algorithms in enhancing the performance of SVM in detecting the taste quality of Siam Kintamani oranges, thereby facilitating the enhancement of efficiency and consistency in the orange industry.

Abstrak

Buah jeruk siam kintamani adalah salah satu komoditas penting dalam pertanian Indonesia, terutama di Kabupaten Bangli, Bali. Namun, menilai kualitas rasa jeruk masih sering mengandalkan identifikasi manual yang subjektif. Dalam upaya meningkatkan objektivitas dan konsistensi dalam menilai kualitas jeruk, penelitian ini mengusulkan penggunaan algoritma Support Vector Machine (SVM) yang dioptimalkan dengan algoritma genetika. Penelitian ini bertujuan untuk mendeteksi kualitas rasa jeruk siam kintamani berdasarkan ciri tekstur pada citra jeruk. Hasil pengujian menunjukkan bahwa SVM yang dioptimalkan dengan algoritma genetika memiliki akurasi yang lebih baik daripada SVM tanpa optimasi. Sebagai contoh, SVM tanpa optimasi memberikan akurasi sebesar 0.78, sedangkan setelah dioptimalkan dengan algoritma genetika, akurasi meningkat menjadi 0.80. Hasil ini menunjukkan potensi besar algoritma genetika dalam meningkatkan kinerja SVM dalam mendeteksi kualitas rasa jeruk siam kintamani, yang dapat membantu meningkatkan efisiensi dan konsistensi dalam industri jeruk.

Keywords: Jeruk Siam Kintamani, Deteksi, SVM, GLCM, Genetic Algorithm, Parameter

1. Pendahuluan

Buah jeruk siam kintamani merupakan salah satu komoditas penting dalam sektor hortikultura Indonesia. Kabupaten Bangli, khususnya, memiliki peran signifikan dalam produksi jeruk, dengan Kecamatan Kintamani sebagai penyumbang utama [1][2]. Kualitas jeruk sangat penting dalam produksi buah, dengan rasa menjadi salah satu faktor utama. Saat ini berdasarkan hasil survey dan wawancara, identifikasi rasa jeruk kintamani masih dilakukan secara manual oleh petani, pedagang, dan konsumen. Metode manual ini melibatkan penilaian berdasarkan tekstur, warna, ketebalan kulit, dan pemeriksaan bagian bawah jeruk. Misalnya, jeruk kintamani yang masih muda cenderung memiliki warna kulit yang lebih gelap atau hijau tua dan rasa yang lebih asam, sedangkan yang

sudah matang memiliki warna yang lebih cerah atau oranye. Namun, konsumen sering kesulitan menentukan kualitas jeruk tanpa bantuan penjual, yang informasinya dapat bersifat subjektif dan tidak konsisten.

Dalam upaya meningkatkan konsistensi dan obyektivitas dalam penilaian kualitas buah jeruk, penelitian sebelumnya telah menunjukkan potensi machine learning, terutama algoritma Support Vector Machine (SVM). SVM, dengan prinsip Struktural Risk Minimization (SRM), telah terbukti efektif dalam mengklasifikasikan buah-buahan, seperti pisang dan melon, serta memberikan jaminan mutu atas produk pertanian. Namun, dalam mengimplementasikan SVM, menentukan parameter optimal menjadi tantangan. Algoritma Genetika telah diusulkan sebagai metode untuk mengoptimalkan parameter SVM, dan penelitian sebelumnya telah menunjukkan keberhasilan dalam hal ini.

Terdapat penelitian sebelumnya dilakukan dengan menggunakan algoritma genetika. Penelitian Harafani & Maulana (2019) menunjukkan akurasi optimal sebesar 89% setelah dilakukan optimasi menggunakan algoritma genetika. Penelitian Sudriyanto dkk. (2022) menemukan bahwa kombinasi algoritma support vector machine (SVM) dan algoritma genetika menghasilkan nilai root mean square error (RMSE) sebesar 101.208+/- 9.475 dan nilai root mean square error (SE) sebesar 10323.858+/- 1965.237. Algoritma genetika digunakan sebagai metode untuk mencari nilai optimal dan solusi dari berbagai permasalahan.

Berdasarkan konteks yang disebutkan, tujuan penelitian ini adalah untuk memanfaatkan algoritma genetika dalam menyesuaikan parameter SVM untuk mengidentifikasi kualitas rasa buah jeruk siam kintamani berdasarkan karakteristik tekstur pada citra jeruk tersebut. Harapannya, penelitian ini akan memberikan bantuan dalam meningkatkan keobjektifan dan konsistensi dalam menilai kualitas jeruk, serta memberikan solusi untuk tantangan optimasi dalam menerapkan SVM.

2. Metode Penelitian

2.1. Pengumpulan Data

Penelitian ini menggunakan data primer yang diperoleh melalui proses observasi lapangan. Metode pengumpulan data yang digunakan dalam penelitian ini adalah studi literatur dan studi observasi.

a. Studi literatur

Metode ini digunakan untuk mengumpulkan informasi dan memeriksa informasi yang terkait dengan teknik Gray Level Co-occurrence Matrix (GLCM) untuk menganalisis tekstur citra, K-Fold Cross Validation untuk pembagian data citra, algoritma Support Vector Machine (SVM) untuk mengklasifikasikan citra, serta jurnal yang membahas klasifikasi rasa buah jeruk siam kintamani yang dapat dijadikan referensi dalam penelitian ini. Peneliti merujuk pada sumber-sumber literatur berupa buku, jurnal, website, dan melakukan wawancara dengan petani jeruk untuk mendukung penelitian ini.

b. Studi Observasi

Metode ini dilakukan dengan mengumpulkan data secara langsung ke perkebunan jeruk di Kecamatan Kintamani, Kabupaten Bangli, Bali untuk melakukan penelitian dan menerapkannya. Jumlah data yang diambil terdiri dari 345 buah jeruk siam kintamani yang dipilih langsung oleh petani. Dalam penelitian dibagi menjadi 3 class yaitu asam, sedang dan manis. Jumlah data dari masing-masing label adalah 115 data. Dari 345 data tersebut dibagi menjadi 2, yaitu 80% data digunakan sebagai data latih untuk proses pencarian model SVM dengan optimasi dan tanpa optimasi. Dan 20% data digunakan digunakan sebagai data uji untuk menguji model SVM tersebut yang terbaik.

Cara pengambilan data dengan memetik buah jeruk dan diletakkan didalam sebuah kardus yang berada di dalam ruangan, setelah itu dilakukan pengambilan gambar pada buah jeruk menggunakan kamera *smartphone* Samsung A52 dengan jarak pengambilan 20 cm, dan tambahan pencahayaan serta angle gambar yang di terapkan sesuai kebutuhan. Pengambilan gambar dilakukan di dalam ruangan dan di dalam kardus bertujuan untuk memberikan perlakuan yang sama pada setiap data hal ini disebut dengan homogen. Berikut ini beberapa contoh data jeruk

Untuk melakukan pelabelan dan validasi data pada penelitian ini dilakukan dengan mengetes buah jeruk dengan menggunakan alat yang bernama refraktometer. Pada tabel 2.1. berikut ini contoh dataset yang digunakan setelah melakukan pelabelan.

Tabel 2.1. Pelabelan Data

Gambar	Nilai Brix	Label
	3.9%	Asam
	6.3%	Sedang
	12.2%	Manis

2.2. Preprocessing

Preprocessing merupakan langkah awal dalam pengolahan citra yang bertujuan untuk memastikan citra yang akan diproses memiliki kualitas yang optimal. Proses preprocessing ini dapat dilihat pada Gambar 2.1 di bawah ini.



Gambar 2.1 Preprocessing data

Pada tahap preprocessing data, dilakukan proses pengubahan skala gambar pada citra kemudian dilakukan penghapusan background pada citra.

2.3. Ekstraksi Fitur dengan Gray Level Co-occurrence Matrix

Pada langkah ini, citra yang telah melalui tahap preprocessing akan diteruskan ke tahap ekstraksi fitur menggunakan GLCM. Studi ini memanfaatkan nilai-nilai ekstraksi fitur GLCM orde kedua, termasuk kontras, korelasi, energi, entropi, dan homogenitas. Hasil ekstraksi fitur ini akan disimpan dalam format file .csv yang akan digunakan dalam proses klasifikasi berikutnya. Gambar 2.2 berikut ini merupakan tahapan dari proses ekstraksi fitur dengan GLCM.



Gambar 2.2. Ekstraksi Fitur

GLCM adalah teknik yang membantu menganalisis dan mengekstrak fitur gambar. GLCM bekerja dengan membentuk matriks yang mencerminkan hubungan antara dua piksel pada jarak dan arah tertentu pada gambar yang dianalisis [3]. Proses ekstraksi fitur di GLCM dilakukan dengan mempertimbangkan empat arah sudut: 0, 45, 90, dan 135 derajat. Fitur-fitur yang dapat diekstraksi menggunakan GLCM adalah:

1. Contrast
 merupakan sebuah ciri yang dipakai untuk menilai seberapa besar perbedaan intensitas dalam sebuah citra.

$$\text{Contrast} = \sum \sum (i - j)^2 P_{ij} \quad k, j=1 \dots k, i=1 \dots \dots \dots (1)$$
2. Correlation
 adalah ciri yang digunakan untuk menilai hubungan antar piksel dalam citra.

$$\text{Correlation} = \sum \sum (i - m) (j - m) p_{ij} \quad \sigma_i \sigma_j \quad k, j=1 \dots k, i=1 \dots \dots (2)$$
3. Energy
 Energy adalah fitur untuk mengukur keseragaman intensitas pada citra.

$$\text{Energy} = \sum \sum p^2_{ij} \quad k, j=1 \dots k, i=1 \dots \dots \dots (3)$$
4. Entropy
 adalah ciri yang mengukur tingkat ketidakaturan dalam distribusi intensitas citra.

$$\text{Entropy} = - \sum \sum p_{ij} \log_2 p_{ij} \quad k, j=1 \dots k, i=1 \dots \dots \dots (4)$$
5. Homogeneity
 adalah ciri yang mengevaluasi seberapa seragam variasi intensitas dalam citra.

$$\text{Homogeneity} = \sum \sum P_{ij} \frac{1}{1+|i-j|} \quad k, j=1 \dots k, i=1 \dots \dots \dots (5)$$

2.4. Penentuan Parameter Optimal pada SVM menggunakan Algoritma Genetika

2.4.1. Konsep Algoritma Genetika

Algoritma genetika merupakan teknik pencarian yang terinspirasi dari prinsip seleksi alam. Digunakan untuk memecahkan masalah optimasi pencarian. Algoritma genetika mempunyai tahapan yang umum yaitu: [3].

- 1) Pembentukan Populasi Awal

Langkah pertama dalam algoritma genetika adalah membuat populasi awal sebanyak N individu yang dihasilkan secara acak. Setiap individu memiliki kromosom yang mewakili solusi yang diinginkan. Dalam penelitian ini, jumlah kromosom per

individu sesuai dengan jumlah parameter yang akan dioptimasi dengan SVM. Evaluasi.

2) Penentuan Nilai Fitness

Setiap individu dalam suatu populasi dievaluasi dengan menghitung nilai fitness. Dalam penelitian ini akurasi model klasifikasi SVM digunakan untuk mengukur nilai kebugaran. Akurasi setiap individu dievaluasi hingga kriteria keluar terpenuhi. Jika tidak, maka akan terbentuk kelompok-kelompok baru.

3) Seleksi

Salah satu metode seleksi yang bisa Anda gunakan adalah roda roulette. Langkah pertama dalam menggunakan route roulette adalah membuat interval nilai kumulatif peringkat kebugaran individu dibagi dengan total peringkat setiap orang. Kemudian nilai acak dihasilkan. Jika nilai acak berada dalam nilai kumulatif untuk orang tersebut, orang tersebut dipilih.

4) Crossover

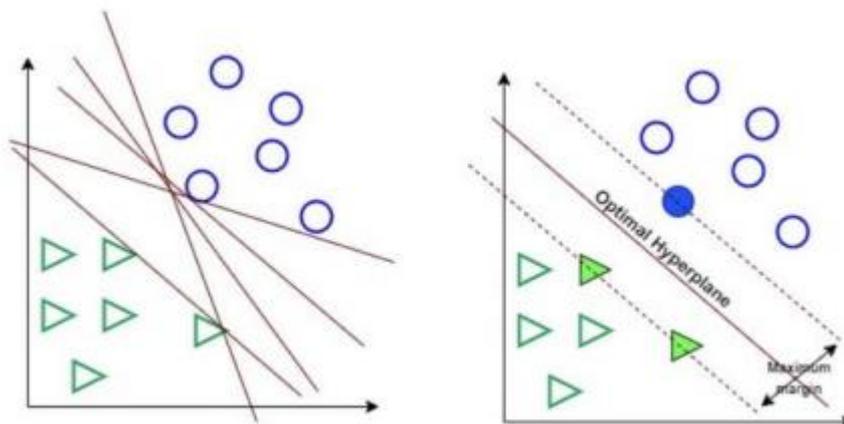
Perkawinan silang Induk yang terpilih pada proses seleksi disilangkan untuk menghasilkan individu baru. Sebelumnya, terjadinya persilangan ditentukan dengan membandingkan bilangan acak yang dihasilkan dengan probabilitas persilangan (pc) yang ditentukan.

5) Mutasi

Mutasi adalah proses perubahan nilai gen suatu individu. Proses mutasi dilakukan bila kondisi tertentu terpenuhi, yaitu ketika bilangan acak yang dihasilkan suatu individu kurang dari probabilitas mutasi yang telah ditentukan (pm) [4].

2.5. Konsep Support Vector Machine (SVM)

SVM merupakan teknik klasifikasi yang membuat garis pemisah untuk membedakan kelas positif (+1) dan negatif (-1). Garis pemisah yang dimaksud ini yakni hyperplane. Gambar 2.3 di bawah menunjukkan penentuan hyperplane pada SVM.



Gambar 2.3. Hyperplane Support Vector Machine

Untuk mencari hyperplane yang optimal, yang perlu dilakukan yakni mengukur atau mencari tepi terbesar dari hyperplane tersebut. Margin adalah jarak antara hyperplane dan support vector untuk setiap kelas [5]. Pada gambar 2.3. di atas, segitiga hijau dan lingkaran biru merupakan vektor pendukung. SVM memungkinkan Anda melakukan klasifikasi dalam kasus linier dan nonlinier. Dalam kasus nonlinier, SVM menggunakan konsep kernel untuk memproyeksikan data ke dalam ruang berdimensi tinggi[6].

Berikut adalah beberapa pilihan fungsi kernel yang digunakan dalam penghitungan mesin vektor dukungan nonlinier (SVM):

a) Kernel Linear

$$(x, y) = x \cdot y \dots \dots \dots (2.11)$$

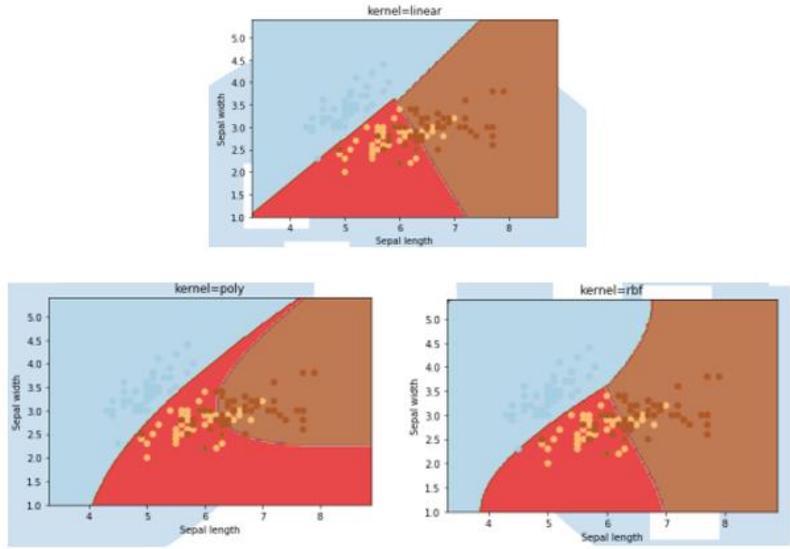
b) Kernel Polynomial

$$(x, y) = (\gamma x \cdot y + r)^d \dots \dots \dots (2.12)$$

c) Kernel Radial Basis Function (RBF)

$$K(x_i, y_j) = \exp(-\gamma ||x_i - y_j ||^2) \dots \dots \dots (2.13)$$

Pemilihan fungsi kernel ini memiliki dampak yang signifikan terhadap akurasi model yang dihasilkan. Gambar 2.4. berikut adalah perbedaan bentuk kernel pada SVM:



Gambar 2.4. Perbedaan Bentuk Kernel pada SVM

3. Hasil dan Pembahasan

Pada pengujian ini dihitung akurasinya melalui proses klasifikasi dengan SVM menggunakan algoritma genetika. Selanjutnya bandingkan hasilnya dengan SVM tanpa optimasi. Nilai k yang digunakan pada validasi k-fold cross adalah 5 dan kernel SVM yang diuji adalah kernel RBF. Kernel RBF memiliki beberapa parameter yaitu C dan Gamma.

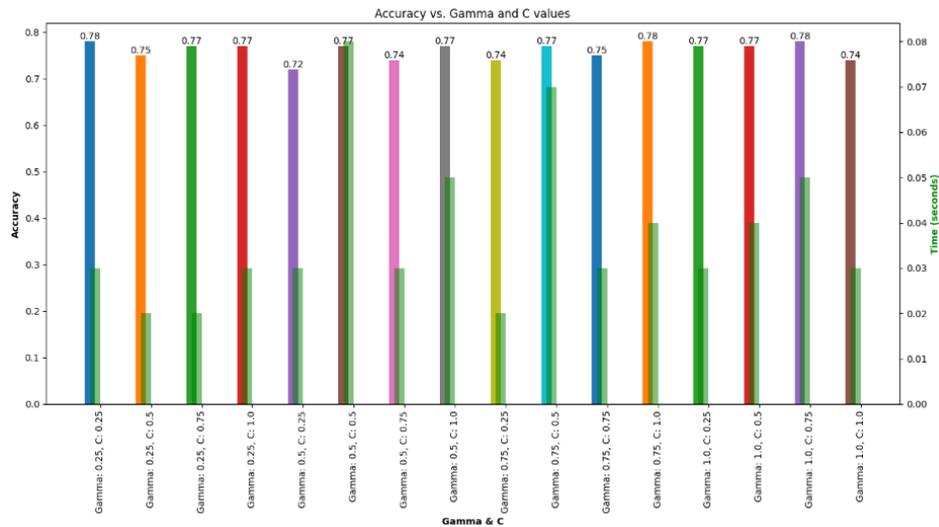
3.4. Pengujian Tanpa Optimasi

Pengujian ini menguji model SVM dengan kernel RBF menggunakan beberapa kombinasi nilai parameter yang ditentukan. Tabel 3.1 menunjukkan model SVM dengan kernel RBF menggunakan 16 kombinasi nilai parameter. Setiap iterasi menghasilkan nilai akurasi yang berbeda-beda untuk setiap kombinasi parameter. Dari tabel 3.1 tersebut, dapat dilihat bahwa nilai akurasi tertinggi adalah 0.78, diperoleh dengan menggunakan kombinasi parameter gamma = 0.25 dan C = 0.25, dengan waktu komputasi sekitar 0.03 detik.

Tabel 3.1. Pengujian Tanpa Optimasi

Gamma	C	Accuracy	Time
0.25	0.25	0.78	0.03 detik
0.25	0.50	0.75	0.02 detik
0.25	0.75	0.77	0.02 detik
0.25	1.00	0.77	0.03 detik
0.50	0.25	0.72	0.03 detik
0.50	0.50	0.77	0.08 detik
0.50	0.75	0.74	0.03 detik
0.50	1.00	0.77	0.05 detik
0.75	0.25	0.74	0.02 detik
0.75	0.50	0.77	0.07 detik
0.75	0.75	0.75	0.03 detik
0.75	1.00	0.78	0.04 detik
1.00	0.25	0.77	0.03 detik
1.00	0.50	0.77	0.04 detik
1.00	0.75	0.78	0.05 detik
1.00	1.00	0.74	0.03 detik

Pada gambar diagram berikut ini menunjukkan bahwa akurasi terbaik 0.78 terletak pada kombinasi nilai gamma = 0.25, C = 0.25, dengan waktu yang dibutuhkan 0.03 detik.



Gambar 3.1. Diagram Akurasi Pengujian Tanpa Optimasi

3.5. Pengujian dengan Optimasi

Pengujian ini menguji model SVM dengan kernel RBF menggunakan beberapa kombinasi nilai parameter tertentu yang dioptimalkan menggunakan algoritma genetika. Tabel 3.2 berikut menampilkan hasil 16 kombinasi parameter untuk model SVM dengan kernel RBF. Setiap iterasi menghasilkan nilai akurasi yang bervariasi untuk setiap kombinasi parameter. Dari tabel 3.2 tersebut, terlihat bahwa nilai akurasi tertinggi adalah 0.80, diperoleh dengan menggunakan kombinasi parameter gamma = 0.0478515625 dan C = 0.365234375, dengan waktu komputasi sekitar 16.09 detik.

Tabel 3.2. Pengujian dengan Optimasi

Pop	Tingkat Mutasi	Probabilitas Crossover	Jml. Gen	Gamma	C	Akurasi	Time
30	0.1	0.6	20	0.0478515625	0.365234375	0.80	16.09 detik
30	0.1	0.7	20	0.017578125	0.5966796875	0.80	16.51 detik
30	0.1	0.8	20	0.00390625	0.423828125	0.72	16.10 detik
30	0.1	0.9	20	0.0595703125	0.48046875	0.80	16.85 detik
30	0.2	0.6	20	0.04296875	0.529296875	0.80	16.77 detik
30	0.2	0.7	20	0.08203125	0.349609375	0.80	16.22 detik
30	0.2	0.8	20	0.015625	0.3408203125	0.77	16.86 detik
30	0.2	0.9	20	0.0166015625	0.2373046875	0.71	17.81 detik
30	0.3	0.6	20	0.01171875	0.5478515625	0.78	16.18 detik
30	0.3	0.7	20	0.0146484375	0.626953125	0.80	17.12 detik
30	0.3	0.8	20	0.0751953125	0.2607421875	0.80	17.45 detik
30	0.3	0.9	20	0.0498046875	0.3017578125	0.80	16.26 detik
30	0.4	0.6	20	0.0146484375	0.3974609375	0.77	17.74 detik

30	0.4	0.7	20	0.06347656 25	0.2265625	0.80	16.35 detik
30	0.4	0.8	20	0.04492187 5	0.18164062 5	0.78	16.31 detik
30	0.4	0.9	20	0.0625	0.29003906 25	0.80	16.60 detik

3.6. Perbandingan Nilai Akurasi SVM Tanpa Optimasi dan Dengan Optimasi Algoritma Genetika

Dapat dilihat perbedaan nilai akurasi SVM dengan SVM +Algoritma Genetika pada gambar 3.2 berikut, dimana nilai akurasi SVM tanpa optimasi memperoleh nilai 0.78 sedangkan nilai akurasi SVM dengan optimasi Algoritma Genetika itu memperoleh akurasi sebesar 0.80. Sehingga bisa dikatakan algoritma genetika mampu mengoptimalkan parameter yang ada pada SVM khususnya pada kernel RBF.



Gambar 3.2. Perbandingan Akurasi SVM dan SVM+GA

4. Kesimpulan

Dari hasil penelitian yang didapatkan, dapat diambil beberapa kesimpulan yakni akurasi model *Support Vector Machine* (SVM) yang dihasilkan dengan optimasi nilai parameter menggunakan algoritma genetika untuk menentukan kualitas rasa buah jeruk siam kintamani adalah untuk kernel RBF memperoleh akurasi sebesar 0.80 dengan nilai populasi = 30, tingkat mutasi = 0.1, probabilitas *crossover* = 0.6, jumlah gen = 20, γ = 0.0478515625 dan C = 0.365234375. sehingga bisa dikatakan bahwa algoritma genetika mampu meningkatkan akurasi pada kernel *Support Vector Machine* (SVM).

Referensi

- [1] BPS Provinsi Bali. (2022, Maret 1). Produksi Buah Jeruk Provinsi Bali Menurut Kabupaten/Kota. Diakses pada 2 April 2023, dari <https://bali.bps.go.id/indicator/55/200/1/produksi-buah-jeruk-provinsi-bali-menurut-kabupaten-kota.html>
- [2] Balitoursclub. (2019, September 19). Jeruk Kintamani Bali. Retrieved April 4, 2023, from <https://www.balitoursclub.net/jeruk-kintamani-bali/>
- [3] Widodo, R., Widodo, A. W., & Supriyanto, A. (2018). Pemanfaatan Ciri Gray Level Co-Occurrence Matrix (GLCM) Citra Buah Jeruk Keprok (*Citrus reticulata* Blanco) untuk Klasifikasi Mutu. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(11), 5769–5776. <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/3420>
- [4] Fadilah, N. I., Rahayudi, B., & Furqon, M. T. (2018). Implementasi Algoritme Support Vector Machine (SVM) Untuk Klasifikasi Penyakit Dengan Gejala Demam. *Jurnal Pengembangan*

Teknologi Informasi Dan Ilmu Komputer, 2(11), 5619–5625.

- [5] Gunawan, V. A., Fitriani, I. I., & Putra, L. S. A. (2020). Sistem Diagnosis Otomatis Identifikasi Penyakit Jantung Coroner Menggunakan Ekstraksi Ciri GLCM dan Klasifikasi SVM. *Informatika Mulawarman : Jurnal Ilmiah Ilmu Komputer*, 15(1), 13. <https://doi.org/10.30872/jim.v15i1.2495>
- [6] Neneng, N., Adi, K., & Isnanto, R. (2016). Support Vector Machine Untuk Klasifikasi Citra Jenis Daging Berdasarkan Tekstur Menggunakan Ekstraksi Ciri Gray Level Co-Occurrence Matrices (GLCM). *Jurnal Sistem Informasi Bisnis*, 6(1), 1. <https://doi.org/10.21456/vol6iss1pp1-10>

This page is intentionally left blank.

Implementasi Metode *Fuzzy Logic* untuk Mendeteksi Asap Dupa di Pasar Tradisional Bali Berbasis IoT

Ira Arituddiniyah^{a1}, Cokorda Pramatha^{a2},
I Dewa Made Bayu Atmaja Darmawan^{a3}, I Gede Surya Rahayuda^{a4}

^aProgram Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana

Kuta Selatan, Badung, Bali, Indonesia

¹iraafd6@email.com

²cokorda@unud.ac.id

²dewabayu@unud.ac.id

²igedesuryarahayuda@unud.ac.id

Abstract

This research employs Fuzzy Logic for incense smoke detection in Bali's traditional markets, supported by IoT. It aims to develop a system for automatic identification, user reminders, and online monitoring. By using Fuzzy Logic, the system assesses smoke concentration, enabling appropriate responses. IoT integration facilitates object connectivity and communication. The outcome is a web platform for smoke monitoring, notifications, and remote device control. This study innovatively merges IoT with local wisdom, offering a valuable contribution.

Keywords: *IoT, Fuzzy Logic, Incense Smoke Detection, Reminder, Monitoring, Local Wisdom, Bali Markets.*

1. Pendahuluan

Indonesia merupakan negara yang memiliki karakteristik yang unik dan membedakannya dari negara lain, dengan keanekaragaman suku, ras, dan budaya menjadi ciri khas masyarakat Indonesia. Setiap daerah di Indonesia memiliki budaya yang berbeda-beda, terutama di Provinsi Bali. Di Bali, masyarakatnya mayoritas beragama Hindu. Umat Hindu di Bali memiliki budaya dalam melakukan upacara keagamaan, salah satunya adalah menggunakan dupa sebagai sarana persembahyangan yang wajib digunakan oleh masyarakat Hindu dalam kehidupan sehari-hari.

Dalam kepercayaan Hindu, terdapat sebuah bentuk pengabdian atau bhakti yang dilakukan dengan memberikan persembahan. Persembahan yang dimaksud adalah upakaraning bebanten (atau biasa disebut sesaji/sesajen), yang diberikan sebelum acara persembahyangan (kramaning sembah) dimulai [1].

Kegiatan mebanten ini biasanya dilakukan setiap hari oleh masyarakat Hindu di Bali. Pada tingkatan yang paling dasar, terdapat suatu kegiatan mebanten yang dapat dilakukan sebagai tanda pengabdian. Kegiatan tersebut melibatkan penyajian canang dan dupa sebagai bentuk persembahan, yang merupakan suatu ritual yang dianggap penting dalam kepercayaan masyarakat setempat, dan biasanya dilakukan setiap hari. Sehingga tidak jarang jika canang dan dupa di temui di berbagai tempat seperti, rumah, jalan raya, tempat usaha atau bahkan pasar tradisional di Bali.

Pasar tradisional di Bali merupakan wadah bagi seluruh lapisan masyarakat, terutama golongan ekonomi menengah ke bawah, untuk melakukan transaksi jual beli. Keberadaannya yang tak terpisahkan dari kehidupan masyarakat Bali, telah memperkuat peran pasar tradisional sebagai tempat pertemuan sosial dan budaya yang memungkinkan interaksi antara masyarakat dan budaya lokal [2].

Padatny aktivitas di pasar tradisional membuat masyarakat terkadang lalai atau lupa untuk mematikan api dupa sehabis sembahyang. Sehingga dapat memicu terjadinya kebakaran. Salah satu tragedi kebakaran pasar tradisiona di Bali yang dipicu karena api dupa yaitu kebakaran yang terjadi di pasar Rubaya di Br. Dinas Beluhu Kangin, Desa Tulamben,

Kecamatan Kubu yang terjadi pada hari Sabtu, 3 Agustus 2019 sekitar pukul 12.40 WITA. Tidak terdapat korban jiwa pada kejadian ini. Namun, kerugian ditaksir mencapai 5 juta rupiah [3].

Tragedi serupa terjadi pasar subagan, Link. Karangasem, Kel,Subagan Karangasem pada hari Jum'at, 29 Maret 2019 sekitar pukul 12.10 WITA. Kebakaran dipicu oleh api dupa yang lupa dimatikan sehabis sembahyang. Tidak terdapat korban jiwa pada tragedi tersebut. Namun kerugian material ditaksir mencapai 3 juta rupiah [4].

Berdasarkan uraian diatas maka, diperlukan adanya solusi yang tepat untuk menyelesaikan permasalahan diatas. Maka dari itu penulis akan membangun sistem "Implementasi Metode Fuzzy Logic Untuk Mendeteksi Asap Dupa di Pasar Tradisional Bali Berbasis IoT" yang diharapkan dapat membantu untuk melakukan pendeteksian asap dupa jika pengguna lupa mematikan dupa sehabis sembahyang. Sehingga permasalahan seperti diatas diharap bisa diminimalisir atau bahkan tidak terulang kembali.

2. Metode Penelitian

2.1. Data dan Metode Pengumpulan Data

Pengumpulan data pada penelitian ini menggunakan data primer. Data tersebut ambil langsung oleh peneliti melalui bantuan alat yang telah dirancang sebelumnya. Terdiri dari sensor MQ-2 yang mendeteksi kandungan asap dupa. Data yang telah diambil akan diolah dan dikirim oleh microcontroller ESP8266 ke database melalui perantara jaringan lokal yang tersedia. Dalam pengembangan sistem peneliti mengimplementasikan metode SDLC (System Development Life Cycle). Metode SDL yang akan digunakan disini adalah metode Waterfall. Metode waterfall dipilih karena sifatnya yang linear dan terstruktur.

Data yang yang diperoleh melalui alat yang telah dirancang tertulis pada tabel 2.1, dimana data ini diperoleh dari jarak yang berbeda diantaranya jarak 100 cm, 70 cm, 50 cm, dan 30 cm dalam waktu masing-masing pengambilan data adalah 5 menit. Data ini digunakan sebagai acuan untuk membuat batas-batas interval pada logika *fuzzy* yang digunakan pada sistem yang dirancang.

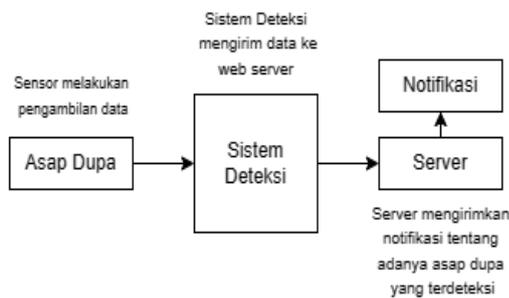
Tabel 2.1. Data ppm

Konsentrasi Asap Dalam 5 Menit		
Jarak 100 cm	Ppm dengan kategori rendah (tidak ada dupa)	< dari 46
	Ppm dengan kategori sedang	46, 50, 51, 55, 52, 49, 49, 49, 56, 56, 56, 56, 53, 49, 48, 45, 50, 50, 50, 50, 50, 55, 55, 52, 52, 56, 56, 51, 50, 45, 46, 50, 53, 53, 53, 55, 55, 51, 51, 52, 52, 53, 53, 53, 50, 50, 48, 46, 46, 46, 49, 49, 48, 48, 52, 52, 53, 55, 55, 50, 50, 50, 48, 49, 49, 48, 48, 48, 47, 47, 50, 50, 53, 53, 53, 50, 50, 52, 50, 55.
	Ppm dengan kategori tinggi	60, 60, 61, 61, 61, 67, 67, 66, 58, 57, 58, 59, 56, 56, 59, 63, 63, 64, 64, 64, 62, 62, 62, 62, 67, 67, 67, 67, 68, 68, 65, 65, 65, 62, 62, 63, 63, 67, 67, 64, 64, 64, 60, 60, 60, 58, 58, 58, 59, 59, 59, 57, 57, 57, 63, 63, 63, 62, 62, 62, 68, 68, 68, 65, 65, 65, 65, 62, 62, 58, 58, 58, 57, 57, 59, 59, 60, 60, 60, 64.
	Ppm dengan kategori sangat tinggi	> 68
	Ppm dengan kategori rendah (tidak ada dupa)	<57

Jarak 70 cm	Ppm dengan kategori sedang	58, 57, 57, 57, 58, 58, 62, 62, 68, 69, 68, 69, 68, 68, 60, 62, 62, 58, 58, 63, 66, 62, 58, 57, 57, 61, 62, 58, 57, 58, 58, 59, 58, 59, 57, 58, 58, 59, 63, 65, 65, 53, 63, 60, 60, 60, 62, 62, 60, 61, 61, 61, 62, 62, 59, 59, 61, 61, 60, 59, 65, 66, 67, 66, 67, 65, 64, 67, 67, 67, 69, 69, 69, 65, 65, 65, 62, 62, 67, 67.
	Ppm dengan kategori tinggi	72, 72, 75, 75, 77, 77, 79, 78, 80, 82, 82, 79, 76, 75, 72, 72, 73, 70, 72, 73, 73, 74, 74, 76, 79, 82, 83, 83, 84, 84, 80, 81, 80, 81, 81, 76, 79, 79, 77, 77, 80, 81, 82, 82, 83, 83, 82, 82, 84, 80, 79, 76, 75, 77, 76, 77, 76, 74, 74, 74, 73, 75, 79, 81, 80, 80, 82, 84, 84, 80, 75, 73, 73, 77, 77, 78, 79, 77, 79, 75.
	Ppm dengan kategori sangat tinggi	>84
Jarak 50 cm	Ppm dengan kategori rendah	<75
	Ppm dengan kategori sedang	76, 76, 75, 75, 77, 79, 80, 82, 85, 85, 82, 86, 87, 90, 90, 90, 88, 89, 89, 92, 92, 93, 93, 91, 91, 92, 92, 90, 90, 89, 89, 85, 84, 88, 89, 92, 91, 92, 88, 87, 88, 88, 89, 92, 92, 91, 87, 85, 85, 87, 88, 85, 85, 86, 88, 87, 85, 83, 82, 85, 89, 92, 92, 91, 89, 87, 86, 85, 82, 83, 84, 82, 83, 83, 81, 80, 79, 79, 77, 79
	Ppm dengan kategori tinggi	95, 97, 97, 102, 103, 105, 105, 99, 99, 97, 103, 101, 100, 100, 98, 97, 104, 105, 107, 108, 111, 114, 115, 115, 116, 116, 114, 110, 111, 109, 107, 108, 109, 105, 106, 102, 105, 106, 109, 111, 113, 114, 114, 114, 113, 112, 115, 116, 116, 113, 111, 110, 111, 111, 109, 108, 112, 115, 115, 116, 116, 113, 115, 115, 110, 110, 109, 107, 107, 107, 107, 107, 105, 104, 101, 100, 98, 97, 98, 98, 97, 99, 99, 98, 99, 98, 100, 101, 101, 102, 102, 101, 102, 102, 105, 105, 105, 102, 105, 105
Ppm dengan kategori sangat tinggi	>116	
Jarak 30 cm	Ppm dengan kategori rendah (tidak ada dupa)	<115
	Ppm dengan kategori sedang	117, 117, 118, 115, 120, 120, 121, 119, 118, 122, 122, 122, 123, 125, 128, 130, 131, 132, 135, 139, 138, 140, 145, 147, 150, 150, 151, 155, 155, 156, 161, 162, 164, 165, 167, 168, 169, 170, 172, 173, 175, 176, 177, 178, 178, 179, 180, 181, 181, 182, 182, 180, 179, 177, 176, 176, 176, 175, 175, 177, 179, 175, 175, 177, 179, 180, 180, 181, 182, 182, 180, 177, 176, 176, 177, 175, 175, 174.

	Ppm dengan kategori tinggi	185, 185, 184, 183, 184, 185, 189, 190, 190, 192, 192, 193, 194, 194, 195, 194, 192, 195, 195, 196, 197, 199, 202, 201, 203, 204, 204, 207, 207, 209, 206, 208, 210, 211, 212, 215, 217, 220, 222, 221, 221, 217, 218, 218, 216, 216, 215, 214, 214, 213, 213, 215, 215, 214, 212, 215, 216, 214, 213, 212, 211, 217, 219, 221, 221, 220, 222, 219, 218, 218, 217, 217, 218, 218, 215, 214, 214, 213.
	Ppm dengan kategori sangat tinggi	>222

2.2. Alur Pendeteksian Asap Dupa



Gambar 2.1. Alur kerja sistem

Data hasil deteksi tersebut akan dikirimkan oleh sistem atau alat pendeteksi yang telah dirancang seperti pada gambar 2.1 menuju ke server website melalui jaringan internet. Selanjutnya data yang telah dikirimkan web sever tersebut akan langsung dianalisis menggunakan sistem cerdas yang dirancang berdasarkan metode *fuzzy logic*, dimana dalam proses analisis tersebut akan memberikan hasil berupa notifikasi yang dikirimkan kepada *user* (pengguna) melalui email berupa notifikasi seabagai informasi tentang adanya asap dupa kepada pengguna dan memberikan informasi melalui tampilan website yang dirancang. Dalam merancang alur kerja sistem yang diuraikan diatas diperlukan adanya beberapa komponen pendukung agar dapat merancang sistem sesuai yang diharapkan. Komponen tersebut meliputi, *Hardware* dan *Software*.

2.3. Fuzzy Logic

Sistem yang dirancang pada penelitian ini menggunakan ppm dari asap dupa sebagai parameter yang digunakan untuk mendeteksi asap dupa. Parameter tersebut dibuat menjadi variabel *fuzzy* agar dapat digunakan pada logika *fuzzy*. Variabel *fuzzy* tersebut memiliki himpunan atau fungsi keanggotaannya masing. Fungsi keanggotaan tersebut di bawah ini :

$$\mu_{\text{Rendah}}(x) = \begin{cases} 1; & x \geq a \text{ dan } x < b \\ 1 - \frac{x-b}{c-b}; & x \geq b \text{ dan } x < c \\ 0; & x \geq c \text{ dan } x < a \end{cases}$$

$$\mu_{\text{Sedang}}(x) = \begin{cases} 1; & x \geq c \text{ dan } x < d \\ 1 - \frac{x-d}{e-d}; & x \geq d \text{ dan } x < e \\ 0; & x < c \text{ dan } x > e \end{cases}$$

$$\mu_{\text{Tinggi}}(x) = \begin{cases} 1; & x \geq e \text{ dan } x \leq f \\ 1 - \frac{x-c}{d-c}; & x \geq c \text{ dan } x < d \\ 0; & x \geq e \text{ dan } x < d \end{cases}$$

$$\mu_{\text{Sangat Tinggi}}(x) = \begin{cases} 1; & x \geq g \\ \frac{x-f}{g-f}; & x \geq f \text{ dan } x < g \\ 0; & x < f \end{cases}$$

Keterangan :

- a = Menyatakan batas bawah interval Rendah
- b = Menyatakan batas atas interval Rendah
- c = Menyatakan batas bawah interval Sedang
- d = Menyatakan batas atas interval Sedang
- e = Menyatakan batas bawah interval Tinggi
- f = Menyatakan batas atas interval Tinggi
- g = Menyatakan batas bawah interval Sangat Tinggi

Batas interval jarak 100 cm diantaranya a=0, b=45, c=46, d=56, e=57, f=68, g=69, Batas interval jarak 70 cm diantaranya a=0, b=56, c=57, d=72, e=73, f=84, g=85, Batas interval jarak 50 cm diantaranya a=0, b=74, c=75, d=93, e=94, f=116, g=117, Batas interval jarak 30 cm diantaranya a=0, b=114, c=115, d=182, e=183, f=222, g=223.

Inferensi yang merupakan tahap pengelolaan data berdasarkan aturan-aturan pada metode *fuzzy logic*. Aturan tersebut berupa himpunan aturan IF-THEN, yang terdiri dari premis yang berisi variabel linguistic. Contohnya, IF konsentrasi asap dupa tinggi, THEN Terdeteksi. Pada tabel 2.2 dapat dilihat aturan inferensi secara sederhana bagaimana pengambilan keputusan pendeteksian asap dupa pada sistem dilakukan.

Tabel 2.2. Aturan Inferensi 1

Asap(ppm)	StatusDeteksi	StatusMeja
Rendah	Tidak	Netral
Sedang	Ya	Terdeteksi
Tinggi	Ya	Terdeteksi
Sangat Tinggi	Berbahaya	Berbahaya

Tahap selanjutnya yaitu penjabaran dari aturan inferensi sebelumnya, yang mana pada tahap ini di terapkan aturan IF-THEN seperti yang tertulis pada uraian di atas. Penjabaran aturan inferensi ini dapat dilihat pada tabel 2.3 di bawah ini.

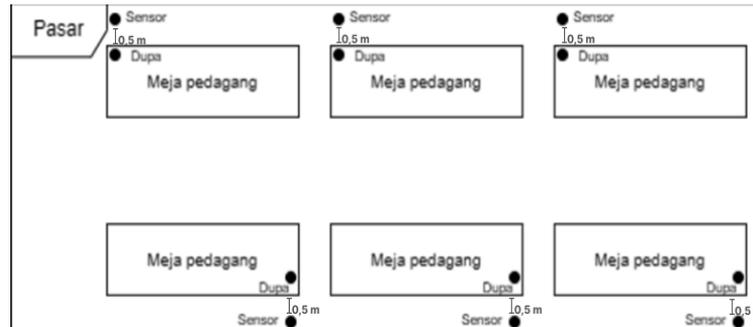
Tabel 2.3. Penjabaran Aturan Inferensi

Aturan 1	IF ppm Asap is Rendah And StatusDeteksi Tidak THEN StatusMeja is Netral
Aturan 2	IF ppm Asap is Sedang And StatusDeteksi Ya THEN StatusMeja is Terdeteksi
Aturan 3	IF ppm Asap is Tinggi And StatusDeteksi Ya THEN StatusMeja is Terdeteksi
Aturan 4	IF ppm Asap is Sangat Tinggi And StatusDeteksi Tidak THEN StatusMeja is Berbahaya

2.4. Perancangan Penempatan Sensor

Perancangan penempatan sensor dapat dilihat pada gambar 2.2. Dimana sensor diletakkan berdekatan dengan posisi dupa dan canang. Dengan jarak antara dupa dengan alat adalah 0,5 meter. Jarak ini disesuaikan dengan waktu pendeteksian atau pemrosesan data yang ada agar proses pengambilan keputusan dapat memberikan hasil yang baik dan lebih efektif. Penempatan dupa diharapkan menghindari adanya angin (kipas angin) karena dapat

mengurangi keefektifan dari sistem. Jumlah alat yang digunakan disesuaikan dengan kebutuhan pasar.

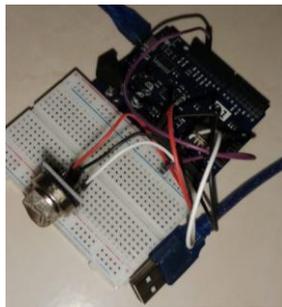


Gambar 2.2. Rancangan Penempatan Sensor

3. Hasil dan Pembahasan

3.1. Perangkat Keras

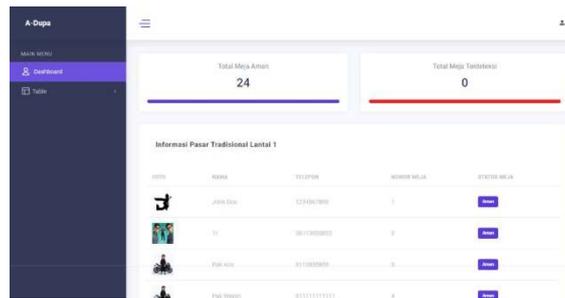
Setiap komponen *hardware* tersebut memiliki fungsi yang berbeda-beda. Sensor MQ-2 berfungsi sebagai pendeteksi kadar asap yang ada di udara. Kabel jumper dan *project board* digunakan sebagai penghubung antar masing-masing komponen *hardware* yang ada. Sedangkan *wemos D1R2 include ESP8266* digunakan untuk memproses data yang diambil oleh sensor serta mengirimkannya ke server database. Dimana sebelum melakukan proses tersebut microcontroller telah diprogram menggunakan bahasa pemrograman agar dapat menjalankan fungsinya sesuai kebutuhan. Gambar perancangan dan implementasi dari perangkat keras tersebut dapat dilihat pada gambar 3.1 berikut.



Gambar 3.1. Alat Deteksi Asap Dupa

3.2. Tampilan Website

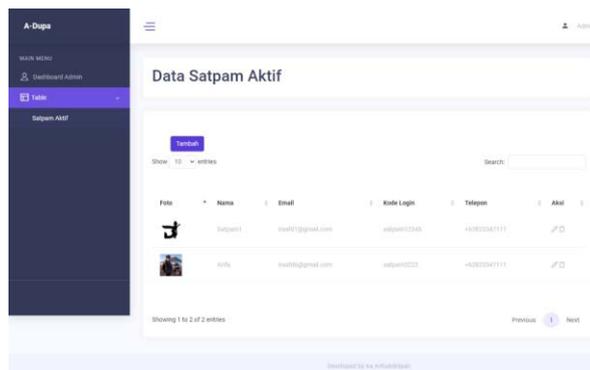
Beberapa tampilan dari halaman website yang telah dirancang dapat dilihat pada beberapa gambar dibawah ini. Gambar 3.2 menampilkan halaman yang tersedia pada website khusus satpam yaitu halaman dashboard. Tampilan dashboard khusus admin dapat dilihat pada gambar 3.3 yang mana pada halaman tersebut menampilkan grafik deteksi asap dupa setiap bulannya. Gambar 3.4 merupakan menu satpam aktif pada website khusus admin yang menampilkan data satpam yang aktif bekerja pada pasar tradisional bali yang mana data pada tabel tersebut juga diperlukan sebagai informasi *login* pada website satpam. Gambar 3.5 menampilkan history meja yang telah terdeteksi asap dupa beserta data pedangan yang menempati masing-masing meja untuk berdagang.



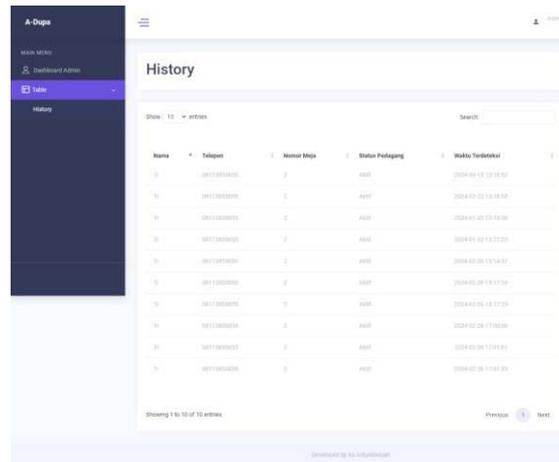
Gambar 3.2. Tampilan Dashboard Khusus Satpam



Gambar 3.3. Tampilan Dashboard Khusus Admin



Gambar 3.4. Tampilan Satpam Aktif



Gambar 3.5. Tampilan History

3.3. Pengujian Sistem

Hasil pengujian pembacaan dan pengiriman nilai sensor dalam bentuk tabel yang menyatakan bahwa kedua proses yang dilakukan telah berhasil dilakukan dan dapat dilihat pada tabel 3.1 di bawah ini.

Tabel 3.1. Pengujian pembacaan dan pengiriman nilai sensor

Pengujian Pada	Cara Kerja	Berhasil	Gagal
Pembacaan nilai sensor	Membaca dan menampilkan data sensor	√	
Pengiriman data	Mengirimkan data hasil deteksi ke web server	√	

Hasil pengujian pengiriman email dalam bentuk tabel yang menunjukkan bahwa pengujian berhasil dilakukan sesuai dengan harapan dimana pengujian dilakukan pada dua kondisi yakni, pada saat StatusMeja=Terdeteksi dan StatusMeja=Aman dapat dilihat pada tabel 3.2 sebagai berikut :

Tabel 3.2. Pengujian pengiriman email

Pengujian Pada	Cara Kerja	Berhasil	Gagal
Pengiriman Email (StatusMeja=Terdeteksi)	Mengirimkan email kepada pengguna (satpam)	√	
Pengiriman Email (StatusMeja=Aman)	Email tidak dikirimkan	√	

Pengiriman email dilakukan setiap 5 menit sekali disesuaikan dengan pendeteksian dan pemrosesan data , selain itu hal ini dilakukan untuk menghindari spam notifikasi pada email.

Proses pengujian logika *fuzzy* dilakukan untuk mengetahui apakah logika *fuzzy* yang telah dibuat dapat berfungsi dengan baik dan memberikan output yang diharapkan dengan benar untuk jarak penempatan sensor dan dupa 100 cm dapat dilihat pada tabel 3.3 di berikut ini.

Tabel 3.3.Pengujian Logika Fuzzy pada Jarak 100 cm

Waktu	Kondisi Dupa	Output		Berhasil	Gagal
		Hasil yang diharapkan	Hasil Sebenarnya		
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Tidak Terdeteksi		√
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Tidak Terdeteksi		√
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Tidak Terdeteksi		√
5 Menit	Menyala	Tereteksi	Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Terdeteksi		√
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	

Akurasi sistem = $16/20 \times 100\% = 80\%$

Proses pengujian logika *fuzzy* dilakukan untuk mengetahui apakah logika *fuzzy* yang telah dibuat dapat berfungsi dengan baik dan memberikan output yang diharapkan dengan benar untuk jarak penempatan sensor dan dupa 70 cm dapat dilihat pada tabel 3.4 di berikut ini.

Tabel 3.4.Pengujian Logika Fuzzy pada Jarak 70 cm

Waktu	Kondisi Dupa	Output		Berhasil	Gagal
		Hasil yang diharapkan	Hasil Sebenarnya		
5 Menit	Menyala	Terdeteksi	Tidak Terdeteksi		√
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	

5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Tidak Terdeteksi		√
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Terdeteksi		√
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	

Akurasi sistem = $17/20 \times 100\% = 85\%$

Proses pengujian logika *fuzzy* dilakukan untuk mengetahui apakah logika *fuzzy* yang telah dibuat dapat berfungsi dengan baik dan memberikan output yang diharapkan dengan benar untuk jarak penempatan sensor dan dupa 50 cm dapat dilihat pada tabel 3.5 di berikut ini.

Tabel 3.5. Pengujian Logika Fuzzy pada Jarak 50 cm

Waktu	Kondisi Dupa	Output		Berhasil	Gagal
		Hasil yang diharapkan	Hasil Sebenarnya		
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi		√
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	

5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	

Akurasi sistem = $19/20 \times 100\% = 95\%$

Proses pengujian logika *fuzzy* dilakukan untuk mengetahui apakah logika *fuzzy* yang telah dibuat dapat berfungsi dengan baik dan memberikan output yang diharapkan dengan benar untuk jarak penempatan sensor dan dupa 30 cm dapat dilihat pada tabel 3.6 di berikut ini.

Tabel 3.6. Pengujian Logika Fuzzy pada Jarak 30 cm

Waktu	Kondisi Dupa	Output		Berhasil	Gagal
		Hasil yang diharapkan	Hasil Sebenarnya		
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Menyala	Terdeteksi	Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Terdeteksi		√
5 Menit	Mati	Tidak Terdeteksi	Terdeteksi		√
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Terdeteksi	√	√
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	
5 Menit	Mati	Tidak Terdeteksi	Tidak Terdeteksi	√	

Akurasi sistem = $17/20 \times 100\% = 85\%$

4. Kesimpulan

Berdasarkan uraian pada penelitian yang telah terlaksana diatas pada sistem deteksi asap dupa berbasis IoT dapat diambil kesimpulan bahwa:

1. Pengimplementasian metode *fuzzy logic* pada sistem ini dilakukan dengan terurut mulai dari tahap fuzzifikasi, inferensi dan defuzzifikasi. Pada tahap fuzzifikasi dilakukan pengonversian

data menjadi variabel linguistik yang memiliki fungsi keanggotaan dimana fungsi keanggotaan tersebut memiliki batas interval yang telah ditentukan berdasarkan data primer pada masing-masing jarak yakni 100 cm, 70 cm, 50 cm, dan 30 cm dengan masing-masing durasi adalah 5 menit untuk mempermudah perhitungan dan pengambilan keputusan dan dilanjutkan dengan tahap inferensi untuk pengambilan keputusan berdasarkan variabel yang ada menggunakan aturan IF-THEN, terakhir menerapkan defuzzifikasi untuk mengubah kembali variabel linguistik sebelumnya menjadi data numerik menggunakan metode centroid.

2. Berdasarkan pengujian yang telah dilakukan pada jarak 100 cm, 70 cm, 50 cm, dan 30 dengan durasi masing-masing pengujian adalah 5 menit diperoleh tingkat akurasi yang berbeda. Masing-masing diuji sebanyak 20 kali, dimana 10 kali dilakukan pada saat dupa menyala dan 10 kali pada saat dupa padam. Pada jarak 100 cm memperoleh tingkat akurasi sebesar 80%, pada jarak 70 cm memperoleh tingkat akurasi sebesar 85%, pada jarak 50 cm memperoleh tingkat akurasi sebesar 90%, dan pada jarak 30 cm memperoleh tingkat akurasi sebesar 85%. Sehingga dapat ditarik kesimpulan bahwa sistem yang dirancang telah berhasil sesuai harapan. Dan sistem yang dirancang paling efektif pada jarak 50 cm karena memberikan tingkat akurasi yang paling baik jika dibandingkan dengan lainnya.

References

- [1] I. G. Ketut Widana, *Etika Sembahyang Umat Hindu*. 2020.
- [2] N. K. Sutrisnawati, A.A.A Ribeka Martha Purwahita, I Ketut Saskara, A.A. Sagung Ayu Srikandi Putri, and Putu Bagus Wisnu Wardhana, "Strategi Pengembangan Pasar Tradisional sebagai Daya Tarik Wisata di Kota Denpasar Bali: Study Kasus Pasar Kumbasari," *J. Kaji. dan Terap. Pariwisata*, vol. 2, no. 1, pp. 37–46, 2021, doi: 10.53356/diparojs.v2i1.45.
- [3] Admindamkar, "Telah Terjadi Kebakaran 1 unit Toko di Pasar Subagan, Lingk. Karangasem, Kel. Subagan, Kec. Karangasem," *Dinas Pemadam Kebakaran dan Penyelamatan Kabupaten Karangasem*, 2019. <http://damkar.karangasemkab.go.id/telah-terjadi-kebakaran-1-unit-toko-di-pasar-subagan-lingk-karangasem/> (accessed Apr. 02, 2023).
- [4] Admindamkar, "telah terjadi kebakaran Pasar Rubaya di Br. Dinas Beluhu Kangin, Desa Tulamben, Kec. Kubu.," *Dinas Pemadam Kebakaran dan Penyelamatan Kabupaten Karangasem*, 2019. [http://damkar.karangasemkab.go.id/telah-terjadi-kebakaran-pasar-rubaya-di-br-dinas-beluhu-kangin-desa-tulamben-kec-kubu/nggallery/thumbnails](http://damkar.karangasemkab.go.id/telah-terjadi-kebakaran-pasar-rubaya-di-br-dinas-beluhu-kangin-desa-tulamben-kec-kubu/) (accessed Apr. 02, 2023).

Performance impact of Hybrid Cryptography in securing RESTful API messages using ECIES

Albertus Ivan Suryawan^{a1}, Agus Muliantara^{a2}, I Dewa Made Bayu Atmaja Darmawan^{a3},
Nurrah Agus Sanjaya ER^{a4}

^aInformatics Department, Udayana University
Bali, Indonesia

¹albertusivan15@gmail.com

²muliantara@unud.ac.id

³dewabayu@unud.ac.id

⁴agus_sanjaya@unud.ac.id

⁴Corresponding Author: Agus Muliantara

Abstract

As technology advances this past decades, many businesses start to integrating technology into their business, which making transaction more convenience. However, this convenience also introducing several security threats against these transmitted data that often involving highly private data. While ideally payment services should have some security standards for its users, research show that misconfigured TLS could actually expose some security threats caused by flaw on certain revisions, which could then be potentially used in dictionary attack. In this research, the author tries to implements a hybrid cryptography implementation involving use of Elliptic Curve Cryptography algorithm and AES in form of Elliptic Curve Integrated Encryption Scheme to secure highly private message over REST APIs and assess its impact in term of performance. In the proposed system, every data that contains a personal data will be secured using end-to-end approach, where each data sent and received will be encrypted using ECIES with AES on the top HTTPS connection. As the result, there is a slight performance degradation at rate of 57 to 230 milliseconds or about 15.57% of the original implementation without any encryption involved inside the system. Although this degradation may seem minimal, it underscores the critical trade-off between performance and security. This increased duration is also still under the currently accepted standard for any transaction request maximum duration which is 8 seconds, and estimated duration for certain request to be completed by the proposed system can be predicted using following formula: $y=0.01156x+1.23$ with RMSE of 3.71.

Keywords: Hybrid Cryptography, Elliptic Curve, AES, REST API, performance analysis, ECIES

1. Introduction

As technology advances this past decades, many businesses start to integrating technology into their business, which making transaction more convenience. However, this convenience also introducing several security threats against these transmitted data that often involving highly private data. The concern come as this business often relies on other third-parties services that support their business operational such as payment gateway, point of sale software, and management software to name some. While ideally service offering such degree of service should have some security standards, but according to some research, misconfigured TLS [1], [2] could actually expose some security threats caused by flaw on certain TLS revisions [3]. That leaked data can be potentially used in dictionary attack which often relies on high volume of valid data.

Cryptography is often used to secure a secret message to ensure that only the authorized receiver has access to it. Some common cryptography algorithm includes Advanced Encryption Standard (AES), Rivest-Shamir-Adleman (RSA), and Elliptic Curve Cryptography. AES is a symmetric cryptography algorithm in which used a single key for both encryption and decryption [4]. While both RSA and Elliptic Curve Cryptography are an asymmetric cryptography which used 2 type of key, namely private key for decryption and public key for encryption, both also has different underlying structure in how the message are processed. Elliptic Curve Cryptography also has advantage of having a smaller size key for roughly the same security power as larger key of RSA or AES [5].

Table 1. Difference between Hoobi's proposed system and this paper's proposed system

	Hoobi's Proposed System	Proposed System
Key Security	DES key is embedded inside the data, and directly encrypted with ECC	AES encryption key derived from ECDH with KDF
Data Encryption Algorithm	DES	AES

Use of hybrid cryptography using Elliptic Curve has been done by Hoobi [6], where Elliptic Curve is being used with DES to secure a message, and resulted a better security with Elliptic Curve and DES compared to only using DES. Another research with similar interest that has been done is [7] where Elliptic Curve Cryptography in form of ECIES with enhanced hash function is used to improve IoT privacy by utilizing blockchain within. The author found that by changing the hash function, an additional 12% of improved data security in term of coefficient correlation between payload and ciphertext, while using 7% less computing time. Research on effectivity of payload encryption also has been done by Varma and UniKrishnan [8], where they measured the effectivity against Man-in-the-Middle (MiTM) attack between application with payload encrypted using AES, and one without any encryption on application level, which both are tested on HTTPS connection. They found that by encrypting sent payload, attacker that gained access to the connection due to MiTM attack, cannot have direct access to the payload.

Based on the related research, this paper aims to measure the performance impact against implementation of hybrid cryptography using AES and ECC with ECIES on existing system, as there are some security advantages can be gained by incorporating a hybrid cryptography into an existing system.

2. Research Methods

Due to the nature of the subject, this research used Waterfall Methodology which consist of 5 stages such as System Requirement Analysis, System Design, Implementation, Testing and Evaluation as shown in Figure 1.

2.1. System Requirement Analysis

System Requirement Analysis is the first stage of System Development Life Cycle (SDLC) where the developer will conduct an analysis to list any requirement needed for the system. Based on observation of technology companies, the system is required to be able to securely transmitted data to another system, while also maintaining stability on high load condition.

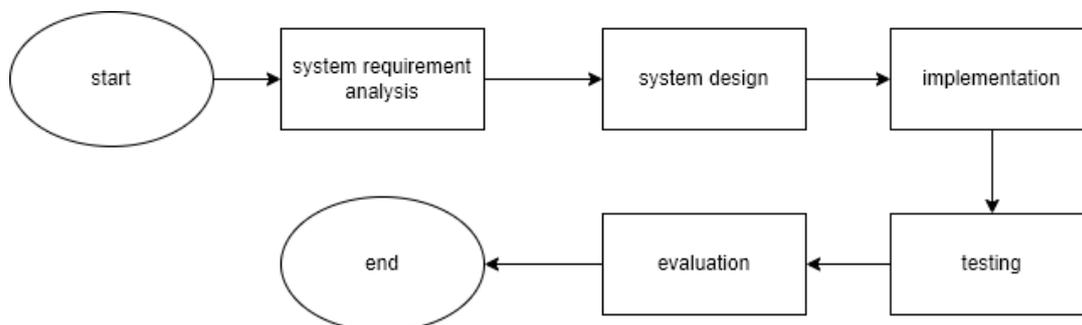


Figure 1. Research Methodology

2.2. System Design

In this research, there are 2 main system to be developed namely Client and Server. Client system is responsible to store user's credentials needed to be able transmitting data securely through the system. This system consists of 2 subsystems, desktop client, and security worker. The desktop client is a GUI-based application used by user to interact with the system. The security worker is program running independently from the main desktop application, and connected by Inter-Process Communication

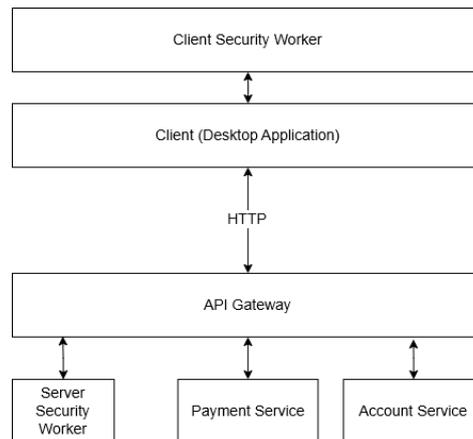


Figure 2. High Level Architecture of Proposed System

(IPC) to the desktop application. It used to handle all cryptographic related process. While on the server side, there is 4 subsystems namely account service, payment service, API gateway, and server security service. The server security worker behaves almost the same as the client counterpart, except it also stores client's public key to be used in server-side encryption. Account service is used to handle all user related process such as user management, and user session managements. While payment service is used to handle all business logic involving payment services, such as deposit, withdrawal, fund transfer, and payment. All request towards server is handled by API gateway which then encrypt or decrypt payload received based on several condition using security service, before then forward that payload to each respective service. Client and Server are connected using HTTPS in RESTful manner, and utilized JSON as its body content type. High Level Architecture of the proposed system is shown Figure 2.

2.3. Implementation

The proposed system is developed using several technologies. In Client side, desktop subsystem is developed using both Rust language for security worker, and Svelte for desktop frontend. While in Server side, Rust language is being used for security worker, and Go language is being used for other services. The proposed system is also using Hybrid Cryptography implementation of ECIES which consist of ECC and AES to encrypt a message. To encrypt a message from either side of the system which shown in Figure 3, each system will fetch the recipient public key, sender secret key, and message to be encrypted. Then, a shared secret would be calculated using both secret key and public key using ECDH, in which used as key material to derive a MAC key and encryption key.

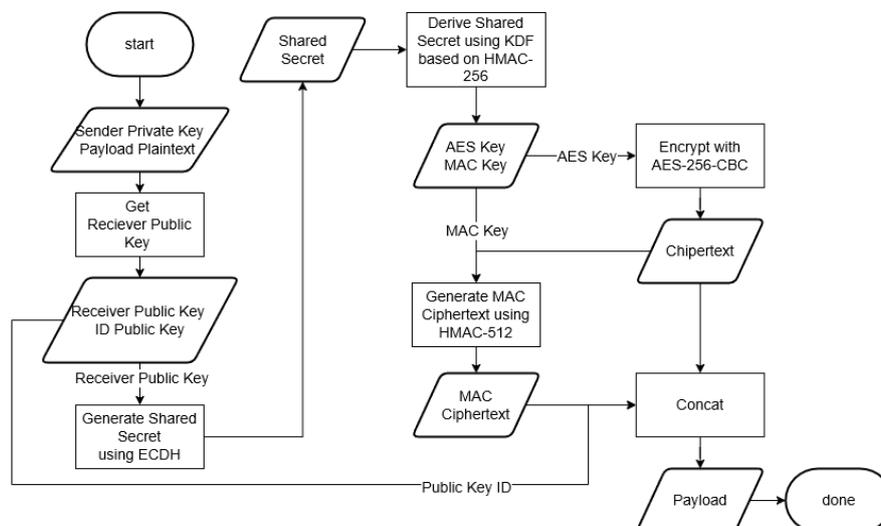


Figure 3. ECIES Encryption Scheme

Message is then encrypted using respective AES configuration using derived encryption key. MAC value of the ciphertext is then calculated using HMAC and MAC key. MAC, Ciphertext and Public Key ID then concatenated together as a single message to be send to the authorized recipient.

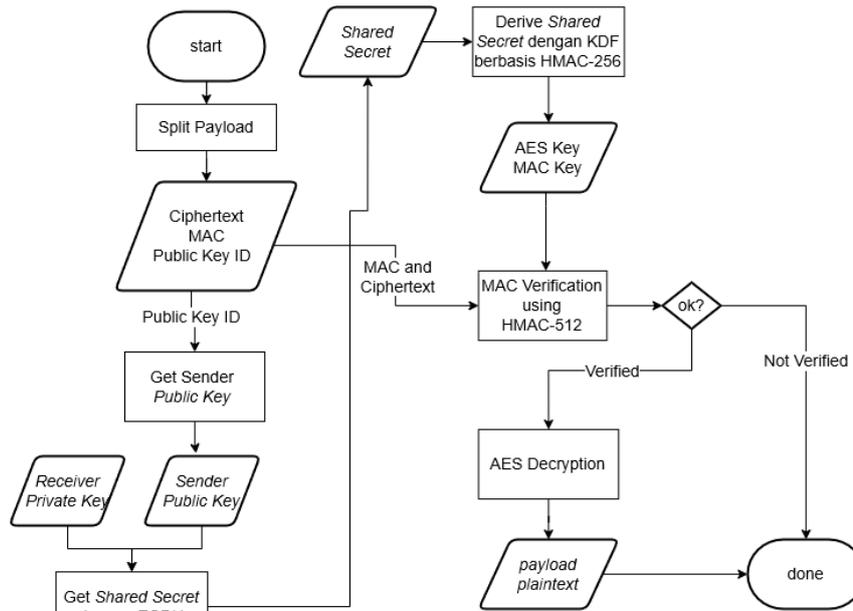


Figure 4. ECIES Decryption Scheme

To decrypt a message from either side of the system which shown in Figure 4, each system will fetch the recipient public key and encrypted message. The concatenated message then split into 3 parts consist of MAC, Encrypted Message, and Public Key ID. A shared secret then calculated from public key and secret key, which then used to derive both MAC key and Encryption Key. Message’s MAC value then validated to ensure encrypted message is still intact. After validation passed, message then decrypted using encryption key.

3. Result and Discussion

In this research, the performance of the proposed application is measured by running a load test between proposed implementation of ECIES in the system and similarly built application without implementation of ECIES. Load test is done using k6 with fixed time of 60 seconds and range of virtual users. Virtual User is used to simulate number of users connecting to the Server side of application concurrently in a time. For this proposed system, the number of virtual users used is 100, 200 and 500 which denoted the time from sending requests to first byte received. This load test measures the time from sending request to first byte received in 3 types of number of virtual users across 2 scenario which is encrypted traffic and unencrypted traffic. The overall result is shown in Table 2.

Table 2. Overall Performance Result on Proposed System

Client Count	Request Duration (ms)					
	Avg	Min	Max	Avg	Min	Max
	Encrypted			Not Encrypted		
100	174.36	136.69	311.99	162.14	124.23	255.69
200	345.75	198.25	725.54	328.21	179.89	651.90
500	663.70	278.90	2003.48	637.87	257.68	1733.10

3.1. 100 Virtual Users

This scenario simulates low traffic of users interacting with the server. As shown in Figure 5, there are some minor bumps of average duration taken for each API call between encrypted and unencrypted traffic. Unencrypted traffic has average duration between 124ms and 255ms on the last iteration, while encrypted traffic has higher average duration between 136ms and 312ms.

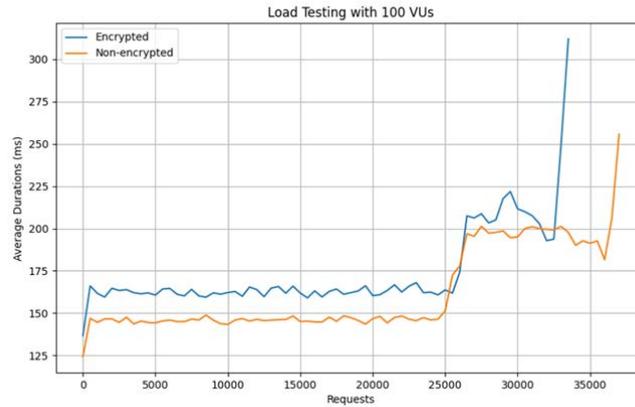


Figure 5. Load testing with 100 VUs

3.2. 200 Virtual Users

This scenario simulates moderate traffic of users interacting with the server. As shown in Figure 6, there are also some minor bumps of average duration taken for each API call between encrypted and unencrypted traffic. Unencrypted traffic has average duration between 180ms and 652ms on the last iteration, while encrypted traffic has higher average duration between 198ms and 725ms.

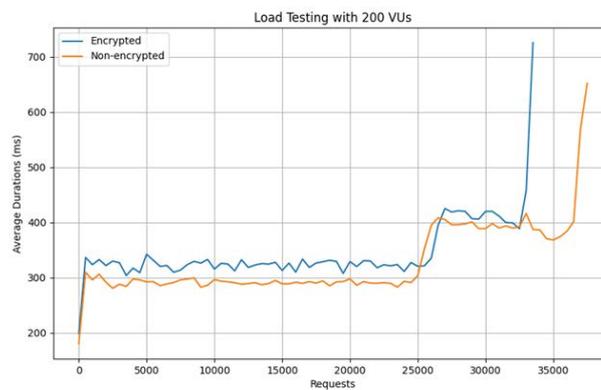


Figure 6. Load testing with 200 VUs

3.3. 300 Virtual Users

This scenario simulates a slight high traffic of users interacting with the server. As shown in Figure 7, there are some minor bumps of average duration taken for each API call between encrypted and unencrypted traffic. Unencrypted traffic has average duration between 258ms and 1733ms on the last iteration, while encrypted traffic has higher average duration between 279ms and 2003ms.

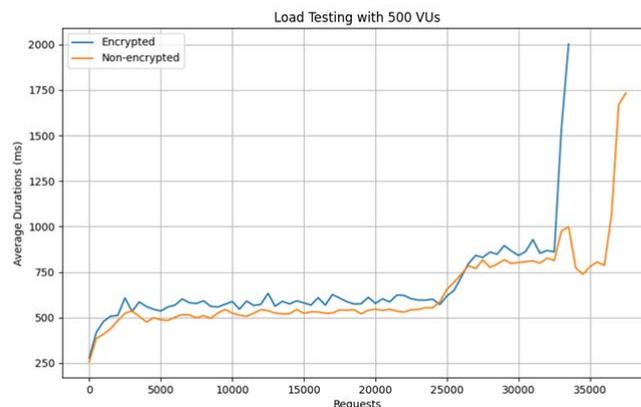


Figure 7. Load testing with 500 VUs

3.4. Performance in term of data size

This section shows the computation time required to process certain number of transactions in the proposed system incorporating encrypted payload. The system is tested against 10.000, 30.000, 60.000, 120.000, and 240.000 requests. As shown in Figure 8, the time it takes for the system to process such data range from 115 seconds (1 minutes and 55 seconds) for 10.000 data to 2.776 seconds (46 minutes and 16 seconds). The computation time required for certain number of transactions in this proposed system can also be predicted using following linear regression formula: $y=0.01156x+1.23$, where y is the computation time required in seconds, and x is the number of data to be processed, with RMSE of 3,70.

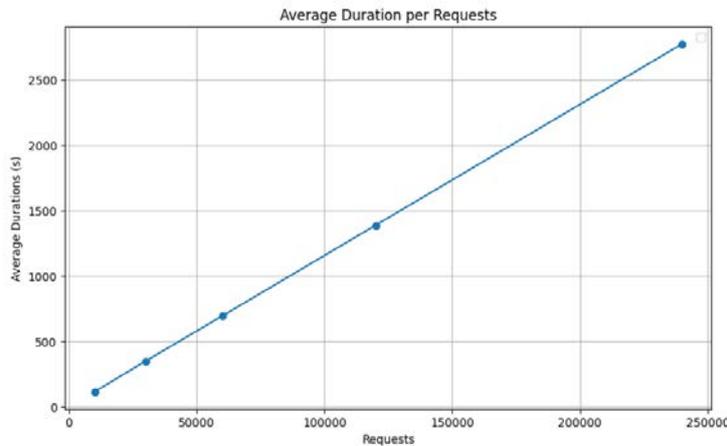


Figure 8. Comparison between number of request and computation time

4. Conclusion

There are many ways to improve security of a system, especially one with wide access to the internet. One of those are by incorporating cryptography into the system, by either encrypting the data, ensuring data integrity with hash, or even both. This research aimed to measure performance impact on incorporating encryption scheme which include both hashing (by using MAC) and encryption, into an existing system. Based on the result shown in the previous section, there are certainly a performance degradation ranging from 57ms until 230ms or roughly 15.57% compared to system without any implementation of ECIES, and can be predicted using linear regression with following equation $y = 0.01156x + 1.23$ with RMSE of 3.71. Such degradation of performance also still below the standard needed for transactional request in the industries. Although this degradation may seem minimal, it underscores the critical trade-off between performance and security.

References

- [1] I. Ali, "Examining cyber security implementation through TLS/SSL on academic institutional repository in Indonesia," *Berkala Ilmu Perpustakaan dan Informasi*, vol. 17, no. 2, pp. 238–249, 2021, doi: 10.22146/bip.v17i1.2082.
- [2] J. K. Huang, Z. X. Zhang, W. J. Li, and Y. Xin, "Assessment of the impacts of TLS vulnerabilities in the HTTPS ecosystem of China," *Procedia Comput Sci*, vol. 147, pp. 512–518, Jan. 2019, doi: 10.1016/J.PROCS.2019.01.238.
- [3] A. Satapathy and J. Livingston, "A Comprehensive Survey on SSL/ TLS and their Vulnerabilities," *Int J Comput Appl*, vol. 153, no. 5, pp. 31–38, Nov. 2016, doi: 10.5120/ijca2016912063.
- [4] M. J. Dworkin *et al.*, "Advanced Encryption Standard (AES)," Nov. 2001, doi: 10.6028/NIST.FIPS.197.
- [5] D. Mahto and D. Kumar Yadav, "RSA and ECC: A Comparative Analysis," 2017. [Online]. Available: <http://www.ripublication.com>
- [6] M. M. Hoobi, "EFFICIENT HYBRID CRYPTOGRAPHY ALGORITHM," *Journal of Southwest Jiaotong University*, vol. 55, no. 3, 2020, doi: 10.35741/issn.0258-2724.55.3.
- [7] Y. P. Khanal *et al.*, "Utilizing Blockchain for IoT Privacy through Enhanced ECIES with Secure Hash Function," *Future Internet*, vol. 14, no. 3, Mar. 2022, doi: 10.3390/fi14030077.
- [8] A. Varma and S. UniKrishnan, "Effect of payload security in MQTT protocol over transport and application layer," *IOP Conf Ser Mater Sci Eng*, vol. 1166, no. 1, p. 012019, Jul. 2021, doi: 10.1088/1757-899x/1166/1/012019.

Pengembangan Sistem Pendukung *Business Intelligence* dan *Forecasting* untuk UMKM (Studi Kasus: Toko Perlengkapan Ternak)

Gusti Ngurah Deva Wirandana Putra^{a1}, Cokorda Pramatha^{ab2},
Ida Bagus Gede Dwidasmara^{a3}, Made Agung Raharja^{a4}

^aProgram Studi Informatika, Universitas Udayana

^bCenter for Interdisciplinary Research on the Humanities and Social Sciences, Universitas Udayana
Bali, Indonesia

¹wirandanaputra@gmail.com

²cokorda@unud.ac.id

³dwidasmara@unud.ac.id

⁴made.agung@unud.ac.id

Abstrak

Penelitian ini bertujuan untuk mengembangkan sistem pendukung *Business Intelligence* pada Toko Perlengkapan Ternak guna menyediakan informasi secara *real-time* yang mendukung pengambilan keputusan yang cepat dan tepat. Metode pengembangan sistem *Business Intelligence Roadmap* digunakan dalam proses pengembangan sistem ini untuk memastikan tahapan-tahapan pengembangan dilakukan dengan terstruktur dan sesuai dengan kebutuhan. Berdasarkan hasil evaluasi, sistem ini mampu memberikan informasi yang terstruktur dan tepat waktu serta memiliki tingkat penerimaan yang tinggi dari pengguna menggunakan evaluasi *Technology Acceptance Model (TAM)* yaitu dengan nilai evaluasi persepsi kegunaan (*perceived usefulness*) mencapai 89% dan evaluasi persepsi kemudahan (*perceived ease of use*) mencapai 86,14%. Hasil dari *ETL Testing* juga menunjukkan bahwa data berhasil diproses sebesar 100% tanpa data loss atau kehilangan data. Selain itu, penelitian juga melakukan *Data Mining* yaitu *Forecasting* menggunakan metode *Support Vector Regression (SVR)* untuk data transaksi penjualan, dengan hasil yang menunjukkan kinerja yang baik dalam prediksi jumlah transaksi di masa depan. Hasil evaluasi menunjukkan bahwa metode *SVR* memiliki *Mean Absolute Error (MAE)* sebesar 21,39, *Mean Squared Error (MSE)* sebesar 642,85, dan *Root Mean Squared Error (RMSE)* sebesar 25,35. Namun, evaluasi jangka panjang diperlukan untuk mengevaluasi kinerja *SVR* dalam periode waktu yang lebih luas dan untuk memahami dampaknya terhadap keberlanjutan bisnis Toko Perlengkapan Ternak.

Keywords: *Business Intelligence, Forecasting, BI Roadmap, Data Warehouse, Dashboard*

1. Pendahuluan

Dalam era digitalisasi yang berkembang pesat, perusahaan dituntut untuk memiliki kemampuan untuk merespons dan beradaptasi dengan cepat terhadap perubahan konstan di lingkungan ekonomi, sosial, dan teknologi [1][2][3]. Data merupakan aset yang sangat berharga bagi semua jenis perusahaan, termasuk UMKM, karena memberikan peluang untuk pemahaman yang lebih baik terhadap masalah dan peningkatan efisiensi operasional [4]. Namun, sebagian besar UMKM belum memanfaatkan potensi data ini secara optimal, dengan hanya mengandalkan prediksi dari pikiran saja dalam mengambil keputusan bisnis [5]. *Business Intelligence (BI)* menjadi kunci untuk mengatasi tantangan ini, dengan mengubah data mentah menjadi informasi yang bermakna untuk meningkatkan fungsionalitas perusahaan dan mendukung keunggulan kompetitif di pasar [6][7].

Implementasi *BI* menjadi semakin penting dalam lingkungan bisnis yang semakin kompetitif, dengan *BI* tidak lagi hanya sebagai teknologi yang 'baik untuk dimiliki', tetapi sebagai kebutuhan yang harus dipenuhi untuk tetap bersaing [8]. Untuk itu, pengembangan sistem pendukung *BI* yang mampu memberikan informasi *real-time* menjadi krusial dalam mendukung pengambilan keputusan yang efisien [9]. Selain itu, kebutuhan akan data *real-time* semakin meningkat, sehingga perusahaan membutuhkan sistem *ETL (Extract, Transform, Load)* yang mampu menyediakan data *real-time* ke dalam *Data Warehouse* [10]. Selain itu, implementasi metode peramalan atau *forecasting* pada *BI* telah

terbukti efektif dalam membuat peramalan terhadap nilai variabel target pada masa mendatang berdasarkan data historis [11][12].

Toko Perlengkapan Ternak yang merupakan UMKM yang menjadi fokus penelitian ini, mengadopsi dan mengembangkan sistem pendukung BI menjadi penting untuk memungkinkan pengambilan keputusan yang lebih baik dan responsif terhadap perubahan pasar [13][14]. Dengan menggunakan metode pengembangan sistem *Business Intelligence Roadmap*, UMKM dapat memaksimalkan manfaat dari solusi BI yang mereka terapkan [15]. Selain itu, pengujian sistem menggunakan *Technology Acceptance Model* (TAM) dapat memvalidasi penerimaan pengguna terhadap sistem berdasarkan faktor *Perceived Usefulness* dan *Perceived Ease of Use*, sehingga memastikan penggunaan yang optimal [16]. Secara keseluruhan, penelitian ini bertujuan untuk mengembangkan sistem pendukung BI yang dapat memberikan informasi *real-time* dan melakukan *forecasting* dengan akurat untuk mendukung keberlangsungan dan pertumbuhan Toko Perlengkapan Ternak.

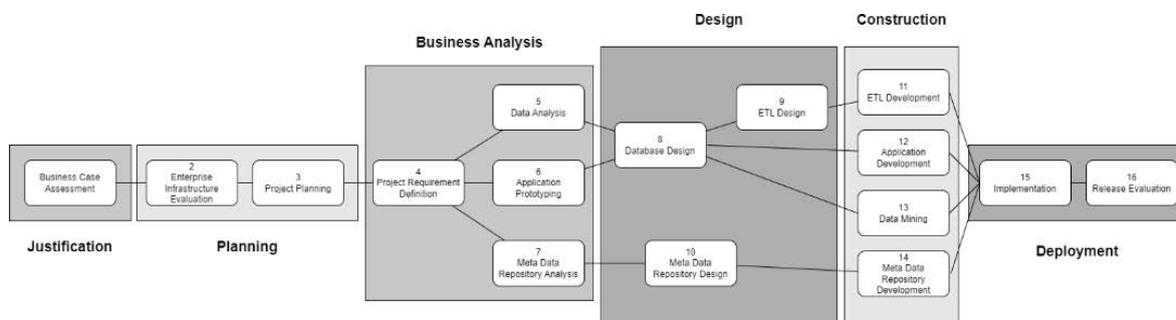
2. Metode Penelitian

2.1. Metode Pengumpulan Data

Proses pengumpulan data didapatkan dengan cara menghubungi pihak bisnis Toko Perlengkapan Ternak yang berlokasi di Kabupaten Badung, Provinsi Bali. Bisnis ini didirikan pada 13 Oktober 2017. Data penjualan yang digunakan adalah data dari tahun 2018 sampai saat ini, sedangkan data pembelian dan keuangan digunakan data dari tahun 2023 sampai saat ini. Data yang dikumpulkan merupakan data yang berupa sebuah *file* CSV. Data yang dikumpulkan mencakup data produk, data pegawai, data pemasok, data kategori produk, data penjualan, data pembelian, dan data keuangan.

2.2. *Business Intelligence Roadmap*

Business Intelligence Roadmap adalah metode yang digunakan oleh perusahaan untuk memperoleh manfaat maksimal dari solusi BI mereka. Metode ini mencakup serangkaian tahap, yang membantu perusahaan mencapai tujuan mereka dalam mengembangkan solusi BI. Dalam penelitian ini, metode yang digunakan yaitu metode dari Moss & Atre pada tahun 2003 [15].



Gambar 1. *Business Intelligence Roadmap*

Dalam tahap pertama yaitu *Justification*, dilakukan *Business Case Assessment* untuk menentukan kebutuhan bisnis, evaluasi sistem, menentukan objektif dari pembangunan BI, mengajukan solusi BI, dan analisis risiko. Tahap kedua yaitu *Planning*, dilakukan evaluasi infrastruktur perusahaan dan *project planning*. Tahap ketiga yaitu *Business Analysis*, dilakukan *project requirement definition*, *data analysis*, *application prototyping*, dan perancangan lainnya. Selanjutnya tahap keempat yaitu *Design*. Disini dilakukan pembuatan *database design*, dan *ETL design*. Tahap kelima yaitu *Construction*, dilakukan pengembangan atau *development* ETL, lalu pengembangan aplikasi, dan *data mining*. Lalu tahap terakhir atau tahap keenam yaitu *Deployment*, dimana disini dilakukan implementasi dari tahap sebelumnya dan setelah selesai diimplementasikan, maka akan *dibuatkan release evaluation* atau evaluasi terhadap hasil dari implementasi.

3. Hasil dan Pembahasan

3.1. Fase *Justification*

Fase *Justification* merupakan tahap pertama dari *Business Intelligence Roadmap*. Kegiatan yang dilakukan pada tahap ini yaitu *Business Case Assessment*. Dalam *Business Case Assessment*, dilakukan identifikasi permasalahan dan analisis kebutuhan bisnis di Toko Perlengkapan Ternak,

dilanjutkan dengan identifikasi dan analisis sistem yang sedang berjalan. Solusi sistem usulan baru selanjutnya diajukan untuk mengatasi permasalahan yang dihadapi pada Toko Perlengkapan Ternak.

Pada tahap pembelian produk oleh *Customer*, pelayanan masih menggunakan sistem konvensional dimana Bagian Penjualan masih melakukan pemberian bukti pembayaran menggunakan nota belanja dengan menulis secara manual dan mencatat penjualan toko pada *File Excel Laporan Penjualan*. Bagian Pembelian dan Gudang mencatat pembelian toko pada *File Excel Laporan Pembelian* dan mengalami masalah mengatur ketersediaan produk yang kurang efisien dikarenakan harus melakukan pemantauan dan *update* secara terus menerus data produk yang ada. Untuk Bagian Keuangan melakukan pencatatan laporan keuangan pada *File Excel Laporan Keuangan*.

Lalu dari setiap laporan yaitu laporan pembelian, laporan penjualan, dan laporan keuangan dilakukan pengecekan oleh Manajer agar nanti bisa dilaporkan kepada *Owner*. Dari laporan tersebut *Owner* dapat mengecek kembali setiap laporannya dan menentukan keputusan untuk bisnis kedepannya. Laporan yang didapat *Owner* juga masih belum terlalu efektif dapat membantu dalam pengambilan keputusan bisnis.

Setelah melakukan analisis sistem berjalan pada Toko Perlengkapan Ternak, maka ditemukan adanya beberapa masalah yang dihadapi yaitu:

1. Keterbatasan dalam Pengolahan dan Analisis Data
2. Kendala dengan Sistem Pencatatan Manual
3. Kesulitan dalam Akses Informasi
4. Tantangan dalam Pengecekan Perkembangan Bisnis
5. Perkiraan Perkembangan Bisnis di Masa Depan

Toko Perlengkapan Ternak menghadapi tantangan dalam meningkatkan efisiensi transaksi, ketersediaan produk, dan pengolahan data. Dengan mengimplementasikan sistem otomatisasi yang meminimalkan kesalahan manusia, termasuk laporan penjualan, pembelian, dan keuangan yang otomatis, serta *dashboard* analisis yang mudah digunakan, toko ini berharap untuk meningkatkan akurasi data, keandalan informasi, dan kepuasan pelanggan. Dengan demikian, diharapkan bahwa penerapan teknologi ini dapat memberikan keunggulan kompetitif, memastikan pertumbuhan berkelanjutan, dan meningkatkan kualitas layanan pelanggan.

Setelah menganalisa kebutuhan *user* akan informasi dan permasalahan yang dihadapi, maka sistem yang diusulkan adalah membangun sebuah Sistem *Business Intelligence* Toko Perlengkapan Ternak. Sistem ini dirancang untuk menyediakan pengalaman berbelanja yang efisien dan lancar bagi pelanggan dan berfungsi sebagai pusat pengelolaan data serta analisis untuk mendukung pengambilan keputusan yang lebih baik.

Sistem ini juga akan memungkinkan pemilik toko dan manajer untuk memantau kinerja toko secara *real-time*, mengidentifikasi tren penjualan, dan mengelola persediaan dengan lebih efisien. Dengan integrasi sistem yang menyeluruh, informasi yang terkumpul akan disajikan dalam *dashboard* yang intuitif dan interaktif, memungkinkan pengguna untuk dengan mudah mengakses dan memahami data untuk mendukung pengambilan keputusan yang strategis. Sistem yang diusulkan akan memberikan beberapa keuntungan diantaranya sebagai berikut:

1. Efisiensi Transaksi
2. Akurasi Pencatatan
3. Pengelolaan Persediaan yang Lebih Baik
4. Analisis Tren Penjualan
5. Laporan Keuangan yang Rinci
6. Pelacakan Performa Toko dan Produk
7. *Dashboard* Interaktif
8. *Forecasting* Penjualan dan Pembelian

3.2. Fase Planning

Fase *Planning* merupakan tahap kedua dari *Business Intelligence Roadmap*. Kegiatan yang dilakukan pada tahap ini meliputi *Enterprise Infrastructure Evaluation* dan *Project Planning*.

a. Enterprise Infrastructure Evaluation

Untuk membangun sistem yang diusulkan diperlukan berbagai infrastruktur untuk menunjang keberhasilan implementasi. Infrastruktur yang dibutuhkan oleh bisnis terdiri dari 2 komponen yaitu Infrastruktur Teknis dan Infrastruktur Non-Teknis. Infrastruktur teknis mencakup perangkat

keras dan perangkat lunak yang digunakan saat ini, sementara infrastruktur non-teknis mencakup peraturan dan keamanan akses data sistem yang didalamnya terdapat *roles* dan *responsibility* serta *requirement* dari data yang diperlukan.

b. Project Planning

Perencanaan proyek merupakan hal penting agar proyek yang direncanakan dapat berjalan sesuai dengan rencana dan hasil yang didapatkan sesuai target yang diinginkan. Proyek yang baik adalah proyek yang dapat selesai tepat pada waktu perencanaan dan sesuai hasil yang diharapkan. Kegiatan yang dilakukan yaitu menentukan *Project Goals and Objectives*, *Project Scope*, *Project Risk*, dan *Project Constraint*.

3.3. Fase Business Analysis

Fase *Business Analysis* merupakan tahap ketiga dari *Business Intelligence Roadmap*. Kegiatan yang dilakukan pada tahap ini meliputi *Project Requirement Definition*, *Data Analysis*, *Application Prototyping*.

a. Project Requirements Definition

Pada tahap ini aktivitas yang dilakukan adalah mendefinisikan kebutuhan tambahan infrastruktur teknis dan mendefinisikan kebutuhan tambahan infrastruktur non-teknis. Kebutuhan teknis untuk Toko Perlengkapan Ternak sudah terpenuhi dalam hal perangkat keras dan perangkat lunak yang digunakan dalam sistem ini. Namun, untuk mendukung implementasi sistem *Business Intelligence*, toko membutuhkan sebuah server yang akan bertanggung jawab atas penyimpanan dan pengelolaan data. Selain itu, diperlukan juga *Thermal Printer* untuk mencetak struk belanja bagi pelanggan dengan cepat dan efisien. Kebutuhan non-teknis juga sudah terpenuhi dari *Roles* dan *Responsibility* yang dibuat sesuai dengan *Use Case Diagram* dan data-data yang diperlukan sudah memenuhi dalam pengembangan sistem BI.

b. Data Analysis

Tahap analisis data dalam perancangan sistem *Business Intelligence* untuk Toko Perlengkapan Ternak berfokus pada penggunaan data internal saja, yang awalnya dalam format CSV. Data ini disesuaikan dengan model ERD untuk memudahkan pengolahan dalam *database* OLTP melalui proses ETL. Tahap ETL mencakup pembersihan data dan perubahan format tanggal, serta pembuatan ID untuk setiap tabel sebagai kunci primer dan asing.

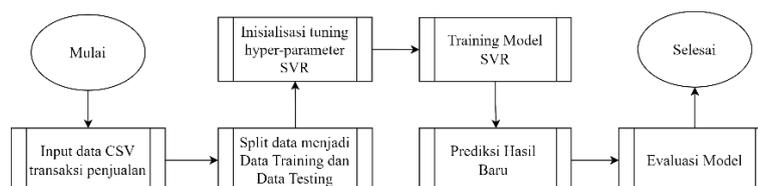
c. Application Prototyping

Prototyping merupakan representasi awal dari *Sistem Business Intelligence* Toko awal tentang cara sistem berinteraksi dengan pengguna yang mempermudah pembuatan program. Figma dipilih sebagai alat untuk merancang desain *prototype* karena fleksibel dan telah digunakan oleh banyak pengguna untuk merancang antarmuka pengguna aplikasi. *Prototype* ini juga dapat dijadikan panduan dalam penggunaan sistem oleh pengguna. Silakan kunjungi URL *prototype* Figma untuk melihat desainnya. Bisa diakses pada <https://s.id/FigmaTPTApps>.

3.4. Fase Design

Setelah melakukan analisis sistem, tindak selanjutnya adalah merancang sistem yang sesuai dengan hasil analisis pada Toko Perlengkapan Ternak. Perancangan sistem termasuk dalam Fase Design, yang merupakan tahap keempat dari *Business Intelligence Roadmap*.

a. Perancangan Forecasting

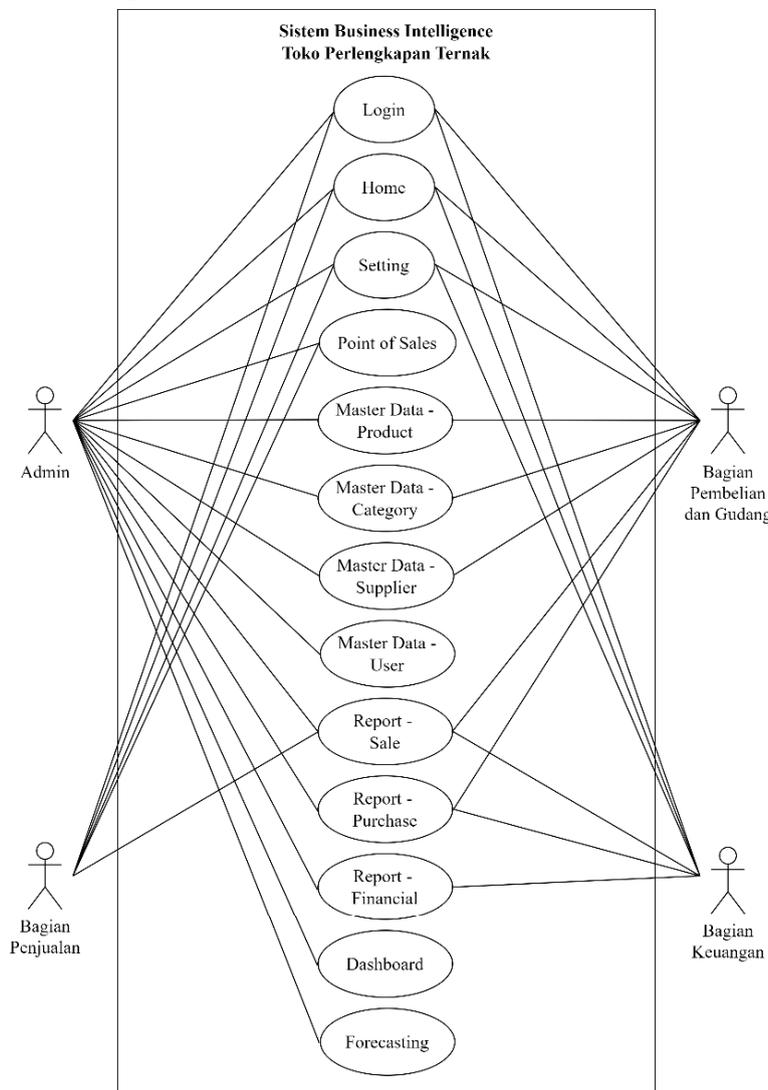


Gambar 2. Flowchart Pembangunan Model Forecasting

Perancangan Data Mining untuk memprediksi data transaksi penjualan masa depan menggunakan *Support Vector Regression (SVR)* dengan *kernel Radial Basis Function (RBF)*, dibuat melalui *Google Collab*. Sesuai Gambar 2, Langkah awal melibatkan *input* data CSV dari *Data Warehouse* melalui *OLAP Query*, diikuti dengan *export* ke *Google Collab*. Data tersebut langsung digunakan tanpa *preprocessing*, kemudian dibagi menjadi data *training* dan *testing*. *Hyper-parameter SVR* di-*tune* untuk mendapatkan hasil terbaik, lalu model dilatih dan dievaluasi. Prediksi dilakukan untuk 3 bulan mendatang, diikuti dengan evaluasi. Integrasi dilakukan dengan *backend Flask* untuk memperbarui prediksi secara dinamis dan menyimpan data baru ke *database*. Proses integrasi ini memudahkan jika model prediksi lain diperlukan di masa mendatang.

b. Use Case Diagram

Use Case Diagram merupakan salah satu jenis diagram yang digunakan untuk menggambarkan interaksi antara aktor-aktor eksternal dan fungsionalitas sistem. Diagram ini membantu untuk memvisualisasikan bagaimana sistem berinteraksi dengan pengguna atau sistem lain dalam berbagai skenario. *Use Case Diagram* bisa dilihat pada Gambar 3 dibawah.



Gambar 3. Use Case Diagram

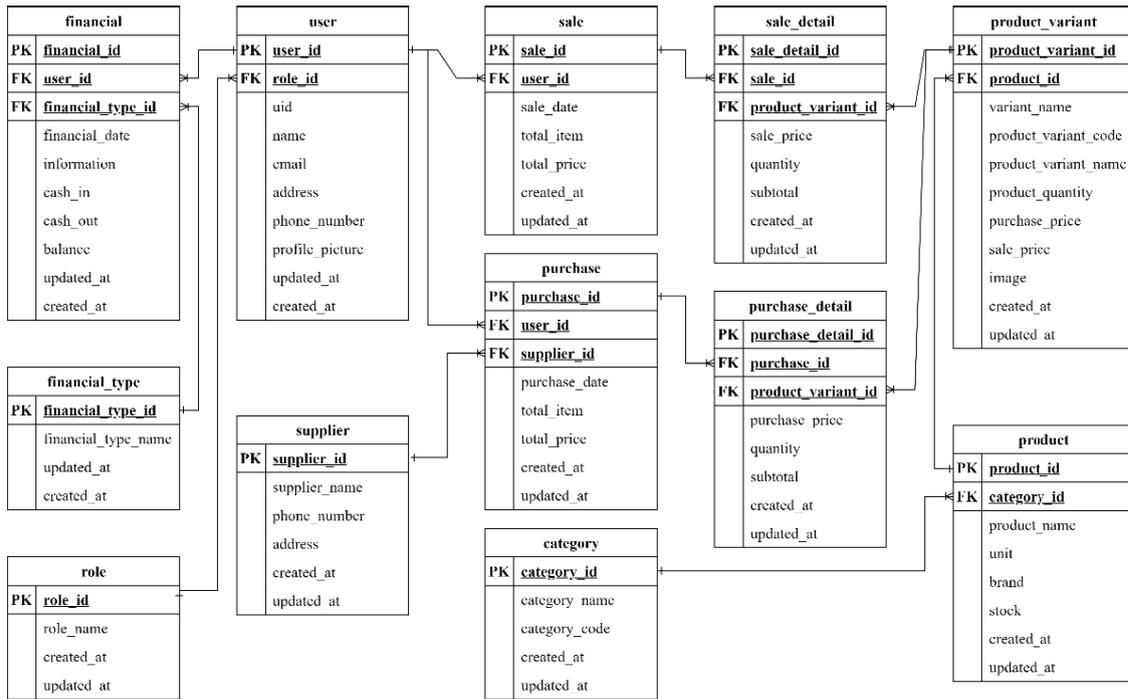
Ada empat aktor dalam sistem. Pertama, Admin memiliki akses penuh ke semua fitur, termasuk modul *view*, *create*, *edit*, dan *delete*, serta menu-menu seperti *Home*, *Setting*, *Point of Sales*, dan lainnya. Kedua, Bagian Penjualan hanya memiliki akses terkait penjualan, seperti *Point of Sales* dan *Report - Sale*. Ketiga, Bagian Pembelian dan Gudang memiliki akses terkait pembelian dan

gudang, seperti *Master Data - Product*, *Report - Sale*, dan lainnya. Keempat, Bagian Keuangan

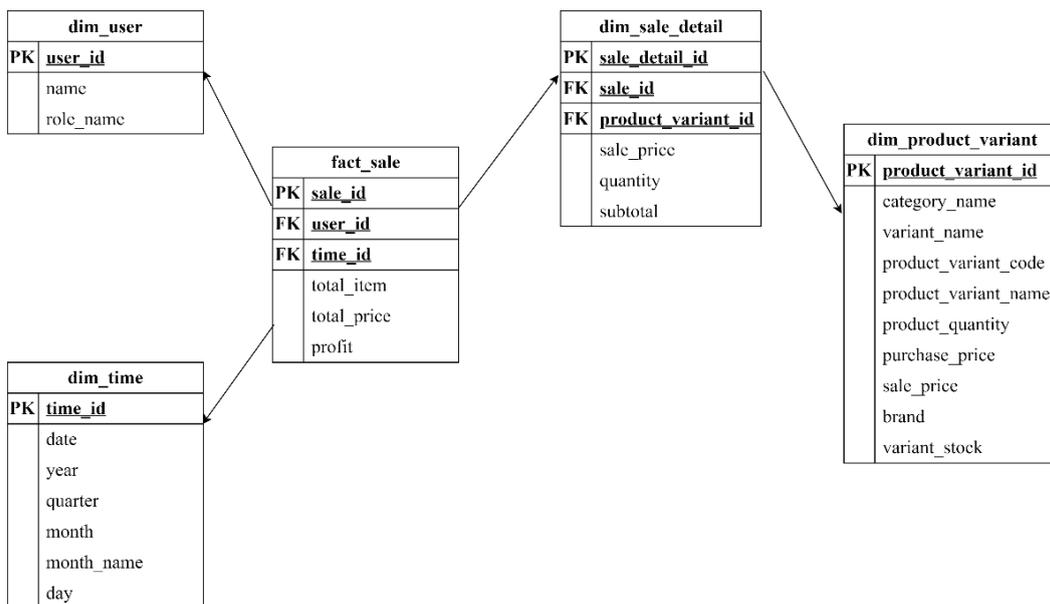
hanya memiliki akses ke fitur terkait keuangan, seperti *Report - Financial*.

c. Database Design

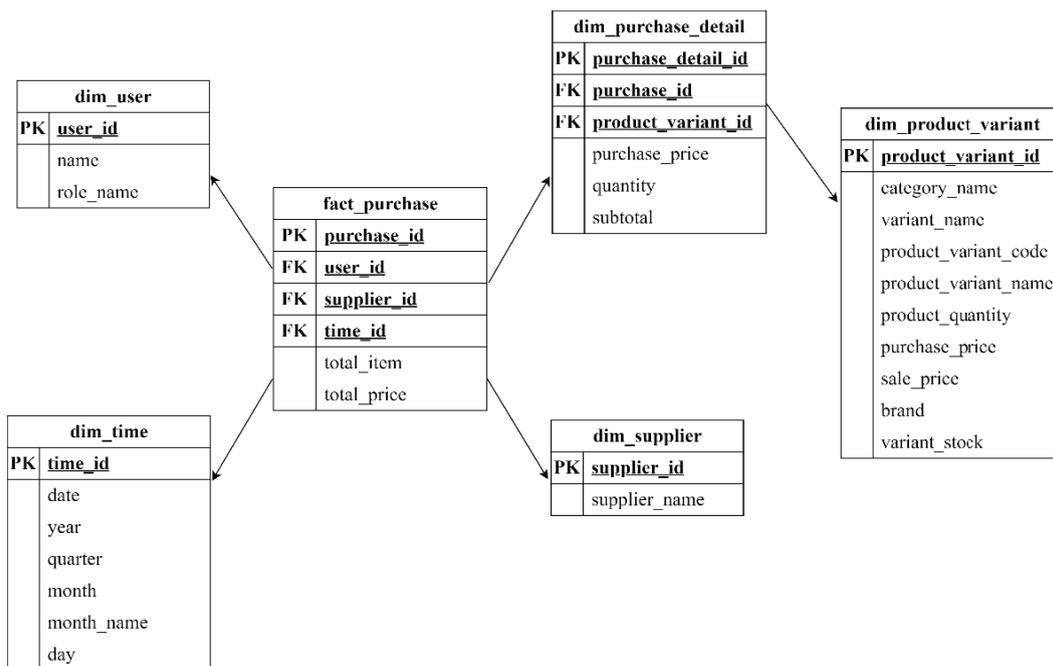
Tahap ini merupakan tahap dilakukannya perancangan *database* yang akan digunakan dalam Sistem *Business Intelligence* Toko Perlengkapan Ternak. *Database* yang diperlukan yaitu *Database OLTP (Online Transaction Processing)* sebagai sumber data transaksi operasional dan *Data Warehouse* yang digunakan sebagai penyimpanan dan analisis data untuk *OLAP (Online Analytical Processing)*.



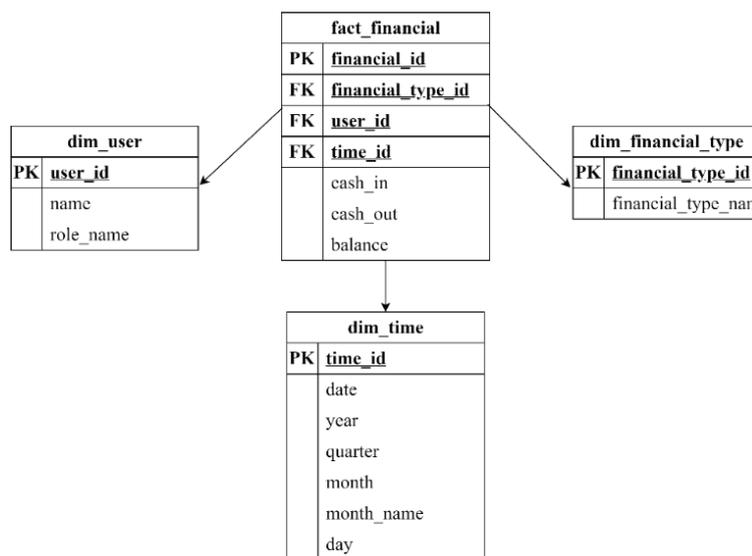
Gambar 4. Entity Relationship Diagram (ERD) Database OLTP



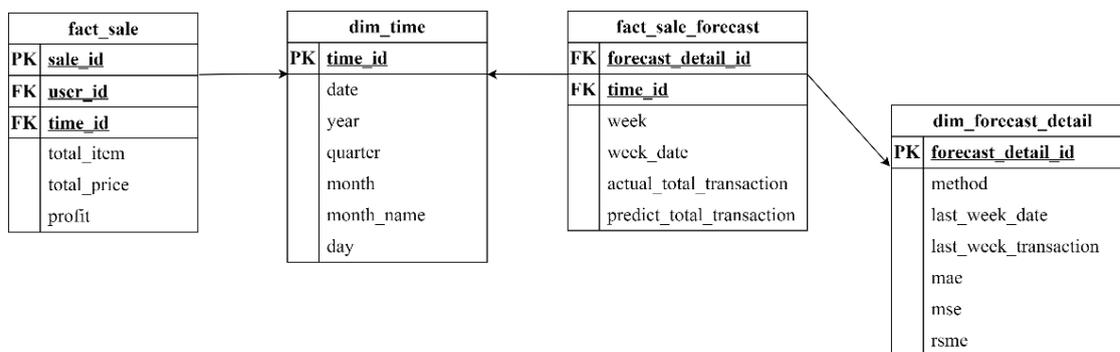
Gambar 5. Snowflake Schema Table Fact_Sale



Gambar 6. Snowflake Schema Table Fact_Purchase



Gambar 7. Star Schema Table Fact_Financial

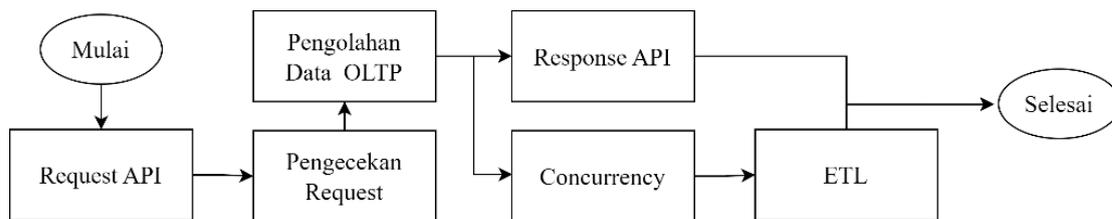


Gambar 8. Constellation Schema Table Fact_Sale_Forecast

Entity Relationship Diagram (ERD) dari *Database OLTP* bisa dilihat pada Gambar 4 diatas. Sedangkan *Data Warehouse* pada penelitian ini dibangun berdasarkan 3 model skema yaitu *Star Schema*, *Snowflake Schema*, dan *Constellation Schema* yang telah dirancang sesuai dengan kebutuhan bisnis dan kebutuhan pengguna. Ketiganya menggambarkan struktur data yang diperlukan untuk operasional dan analisis toko perlengkapan ternak. ketiga konsep *data modeling* ini memastikan bahwa informasi yang relevan dapat diakses dan dianalisis secara efisien sesuai kebutuhan bisnis. Berikut merupakan skema yang telah dibuat.

Pada penelitian ini dibuat 4 skema yaitu *Snowflake Schema Table Fact_Sale* bisa dilihat pada Gambar 5, *Snowflake Schema Table Fact_Purchase* bisa dilihat pada Gambar 6, *Star Schema Table Fact_Financial* bisa dilihat pada Gambar 7, dan *Constellation Schema Table Fact_Sale_Forecast* bisa dilihat pada Gambar 8.

d. ETL Design



Gambar 9. Flowchart ETL Real-time

Untuk proses ETL dilakukan ketika *Frontend* sistem *Business Intelligence* melakukan *request* ke *Backend* seperti menambahkan data transaksi baru. Lalu pada *Backend* tersebut akan dilakukan pengecekan parameter *request* apakah sudah sesuai. Jika tidak sesuai maka proses tidak dilanjutkan.

3.5. Fase Construction

Setelah menyelesaikan tahap perancangan sistem, tahap selanjutnya adalah implementasi sistem yang dilakukan sesuai dengan rencana sistem yang telah dibuat sebelumnya. Implementasi sistem termasuk dalam *Fase Construction*, tahap kelima dari *Business Intelligence Roadmap*.

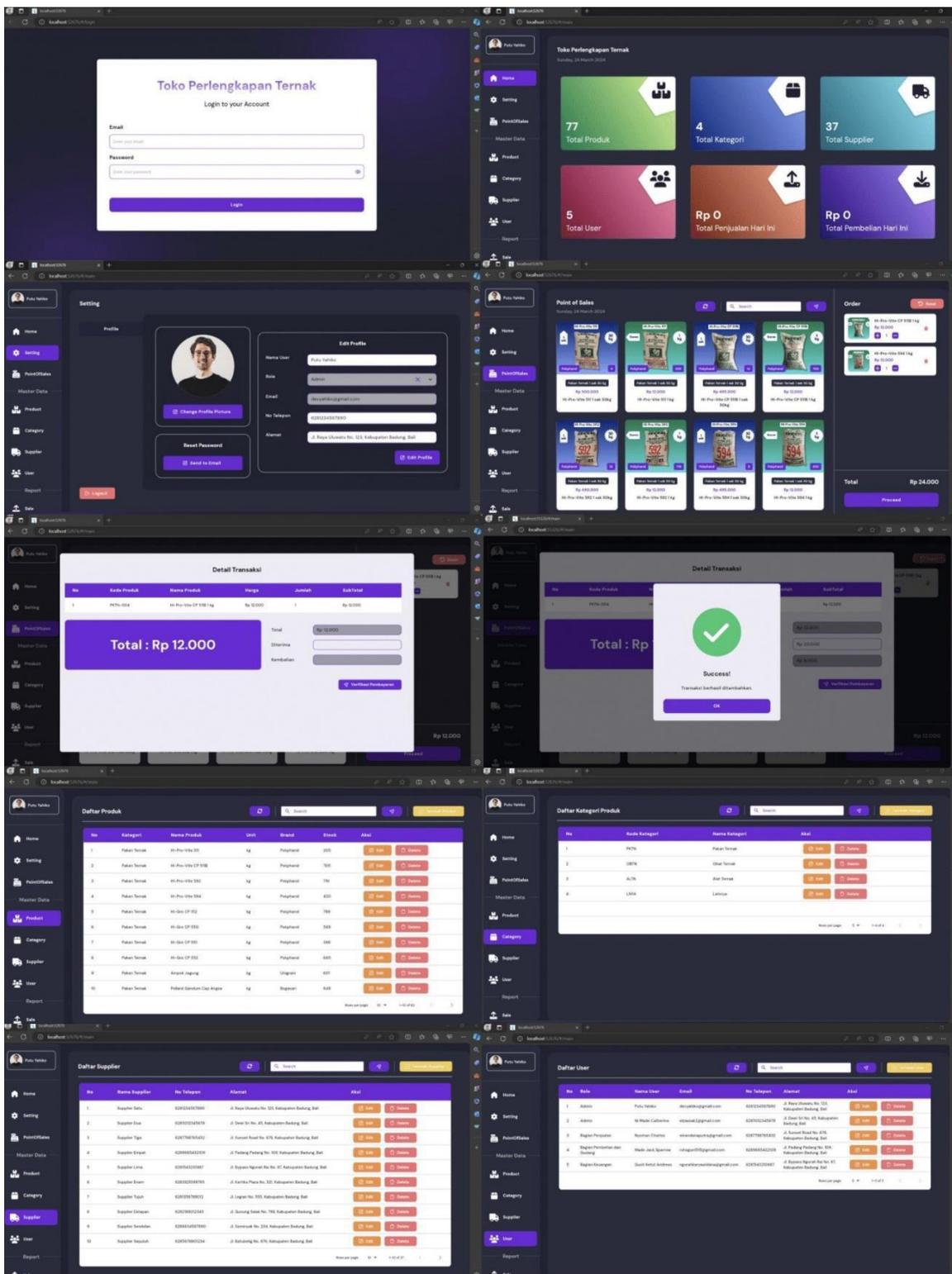
a. ETL Development

Tahap ini menggambarkan proses pengembangan *Extract, Transform, dan Load* (ETL) data yang dilakukan dalam pembuatan Sistem *Business Intelligence* Toko Perlengkapan Ternak. Dalam tahap Pengembangan ETL atau *ETL Development*, dilakukan pembuatan ETL data historis dan ETL *real-time* sesuai dengan rancangan yang telah dibuat pada ETL Design. Pada ETL data historis menggunakan tools *Pentaho Data Integration 9.4.4.0* sedangkan ETL *real-time* diimplementasikan pada *Backend* dengan konsep *concurrency*.

b. Application Development

Pada tahap ini, dilakukan pengembangan sistem. Implementasi dilakukan dengan membangun sistem sesuai dengan perancangan yang telah dibuat sebelumnya. Pada tahap ini, dilakukan pengembangan Sistem *Business Intelligence* Toko Perlengkapan Ternak. Implementasi dilakukan dengan membangun sistem sesuai dengan perancangan yang telah dibuat sebelumnya. Proses pengerjaan atau implementasi akan menggunakan *Visual Studio Code* untuk melakukan proses pemrograman *Frontend* dan *Backend* dan *MySQL phpMyAdmin* untuk *Database Client*.

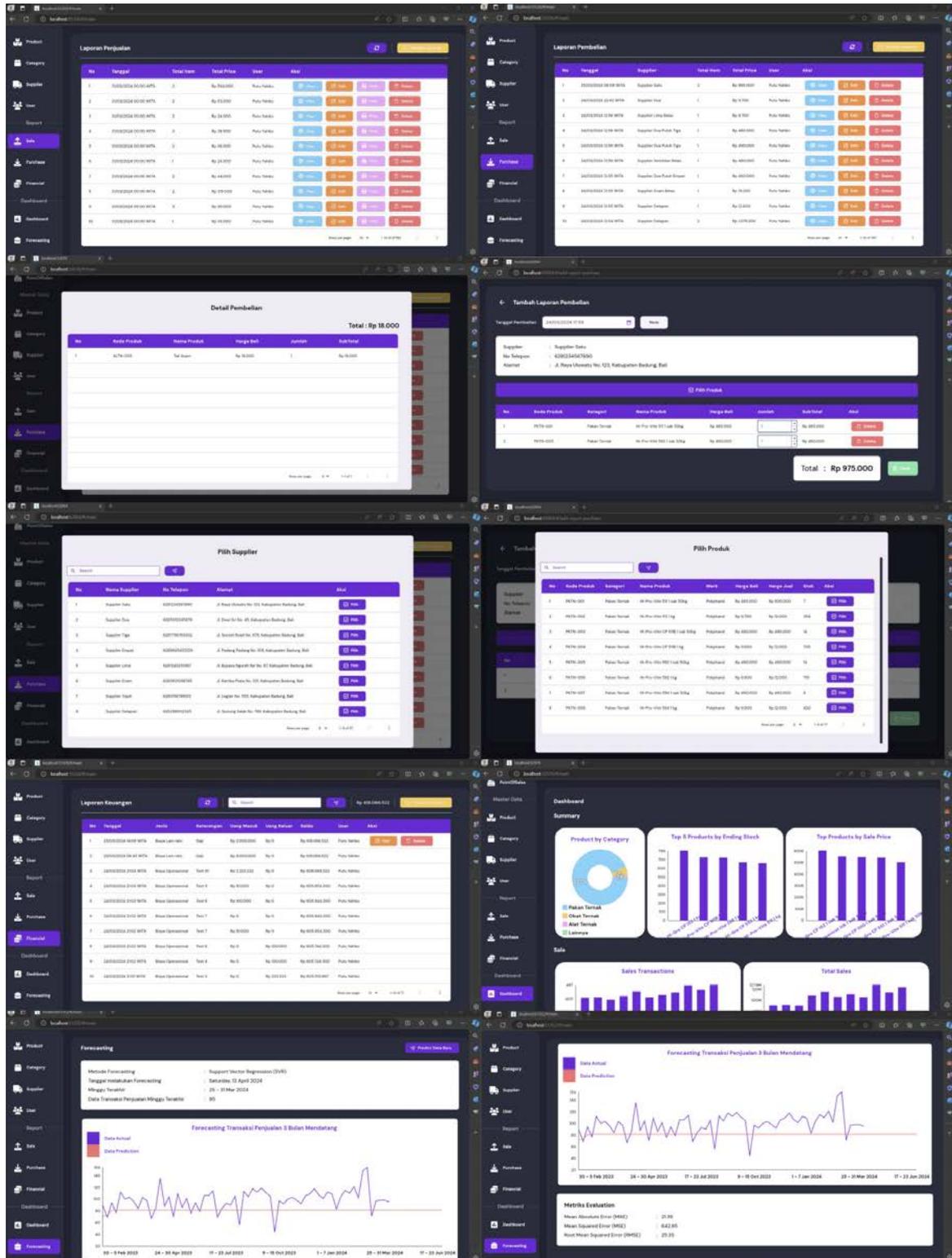
Untuk *Version Control System* yang akan digunakan yakni *Git* dengan *remote repository* yaitu *GitHub*. Setiap integrasi ke dalam *branch main* harus melalui *Pull Request* agar dapat di-review dan juga akan menjadi sebuah histori yang dapat ditinjau kembali nantinya. Jadi nanti pengembangan sistem dapat terkontrol dengan baik. Berikut merupakan hasil dari Sistem *Business Intelligence* Toko Perlengkapan Ternak.



Gambar 10. Hasil 1 Sistem *Business Intelligence* Toko Perlengkapan Ternak

Pada fitur *Login*, pengguna dapat masuk menggunakan email dan *password*. Setelah berhasil *login*, fitur yang dapat diakses akan sesuai dengan peran pengguna seperti yang tercantum dalam *Use Case Diagram*. Halaman *Home* menyajikan rangkuman data utama dan laporan seperti Total Produk, Total Kategori, Total *Supplier*, Total Pengguna, Total Penjualan Hari ini, dan Total Pembelian Hari ini, yang berasal dari *Database OLTP*. Fitur *Setting* memungkinkan pengguna untuk mengedit profil, termasuk foto profil, biodata, dan *reset password*, serta melakukan *logout* dari sistem.

Putra, Pramarta, Dwidasmara, dan Raharja
 Pengembangan Sistem Pendukung *Business Intelligence* dan *Forecasting* untuk UMKM
 (Studi Kasus: Toko Perlengkapan Ternak)



Gambar 11. Hasil 2 Sistem *Business Intelligence* Toko Perlengkapan Ternak

Pada fitur *Point of Sales*, terdapat daftar produk yang dapat dibeli oleh pelanggan, dengan informasi seperti nama, stok, kategori, harga, dan merk produk. Pengguna dapat mencari produk, menambahkannya ke dalam pesanan, mengatur jumlah pembelian, dan melakukan proses pembayaran dengan menginput jumlah uang yang diterima serta melihat kembalian. Fitur *Master Data* memungkinkan pengguna untuk melihat, menambah, mengedit, dan menghapus data produk, kategori, pemasok, dan pengguna. Fitur *Report* menyajikan laporan penjualan, pembelian,

dan keuangan, dengan opsi melihat, menambah, mengedit, mencetak, dan menghapus laporan. *Dashboard* menampilkan informasi grafis yang mendukung pengambilan keputusan, seperti ringkasan data, penjualan, pembelian, dan keuangan. Sedangkan fitur *Forecasting* memberikan informasi hasil peramalan dengan grafik dan metrik evaluasi, serta opsi untuk melakukan peramalan ulang dengan data baru.

3.6. Fase Deployment

Fase terakhir dari *Business Intelligence Roadmap* yaitu Fase Deployment dimana melakukan pengujian dan evaluasi sistem untuk memastikan bahwa sistem telah berfungsi dengan baik.

a. ETL Testing

sale_id	user_id	time_id	total_item	total_price	profit	purchase_id	user_id	supplier_id	time_id	total_item	total_price	financial_type_id	user_id	time_id	cash_in	cash_out	balance
27348	1	20240324	1	3000	31000	402	1	15	20240324	1	9700	4	1	20240324	0	0	605044300
27367	1	20240324	1	50000	33000	401	1	23	20240324	1	485000	4	1	20240324	100000	0	605844300
27366	1	20240324	2	33000	16000	400	1	23	20240324	1	490000	4	1	20240324	0	100000	605744300
27365	1	20240324	2	53000	26000	399	1	19	20240324	1	480000	4	1	20240324	0	100000	605744300
27364	1	20240324	2	325000	98000	398	1	24	20240324	1	460000	5	1	20240324	222222	0	606068522
27363	1	20240324	2	34000	23000	397	1	16	20240324	1	19200	5	1	20240324	500000	0	606344300
27362	1	20240324	2	25000	3400	396	1	8	20240324	1	12800	5	1	20240324	0	1000000	613844300
27361	1	20240324	2	24000	2300	395	1	8	20240324	2	1079200	4	1	20240324	0	333333	605519967
27360	1	20240324	3	1044000	6000	394	1	2	20240324	1	9800	5	1	20240324	0	8000000	605844300
27359	1	20240324	2	515000	8000	393	1	1	20240324	1	9700	4	1	20240324	10000	0	605854300

Gambar 11. Hasil Pengujian ETL

Pada proses ETL tanpa terstruktur dengan 10 data penjualan, 10 data pembelian, dan 10 data keuangan, tidak terjadi *data loss*. Seluruh data yang dipindahkan dari *Backend* ke *Data Warehouse* berhasil tanpa kehilangan data. Hasil analisis menunjukkan bahwa ETL *Real-time* berhasil 100% tanpa adanya data yang hilang, sehingga semua data penjualan, pembelian, dan keuangan terisi dengan baik di *Data Warehouse*.

b. Data Mining Testing

Metriks Evaluation	
Mean Absolute Error (MAE)	: 21.39
Mean Squared Error (MSE)	: 642.85
Root Mean Squared Error (RMSE)	: 25.35

Gambar 12. Hasil Pengujian Data Mining

Dengan menggunakan metrik evaluasi seperti Mean Absolute Error (MAE), Mean Squared Error (MSE), dan Root Mean Squared Error (RMSE), dilakukan evaluasi terhadap *forecasting* menggunakan metode *Support Vector Regression* (SVR) pada data transaksi penjualan terakhir hingga 31 Maret 2024, dengan prediksi 3 bulan ke depan. Hasil menunjukkan MAE sebesar 21,39, MSE sebesar 642,85, dan RMSE sebesar 25,35. Meskipun prediksi awal menunjukkan kinerja yang baik, evaluasi lebih lanjut diperlukan terutama jika digunakan untuk jangka waktu yang lebih panjang, karena kinerja SVR sangat dipengaruhi oleh distribusi data pada *dataset*.

c. Technology Acceptance Model (TAM)

Tabel 1. Hasil Pengujian Technology Acceptance Model (TAM)

Statistics	Perceived Usefulness	Perceived Ease of Use
Mean	6,23	6,03

Hasil akhir pengujian TAM melalui kuesioner online kepada pihak Toko Perlengkapan Ternak menunjukkan tingkat kegunaan Sistem *Business Intelligence* dengan rata-rata nilai 6.23, serta tingkat kemudahan penggunaan dengan rata-rata nilai 6.03. Tidak ada responden yang memberikan nilai di bawah 4, menandakan bahwa sistem yang ditawarkan dinilai berguna dan mudah digunakan oleh pengguna.

4. Kesimpulan

Berdasarkan penelitian yang telah dilaksanakan dan juga hasil yang diperoleh dari penelitian, maka dapat didapatkan kesimpulan sebagai berikut:

- a. Pengembangan Sistem *Business Intelligence* pada Toko Perlengkapan Ternak dengan metode pengembangan sistem *Business Intelligence Roadmap* dapat menghasilkan sistem yang lebih terarah dan terstruktur dalam implementasinya. Penerapan ETL *Real-time* pada sistem dapat menghasilkan informasi secara *real-time* dan dari hasil pengujian ETL Testing didapatkan bahwa data 100% berhasil tanpa *data loss*. *Dashboard* yang tersedia pada sistem dapat digunakan sebagai pendukung dalam pengambilan keputusan. Hasil dari *Technology Acceptance Model* (TAM) juga menunjukkan angka 89% untuk evaluasi persepsi kemudahan (*perceived usefulness*) dan 86,14% untuk evaluasi persepsi kemudahan (*perceived ease of use*) yang menyatakan bahwa sistem yang ditawarkan berguna dan mudah untuk digunakan bagi user.
- b. Berdasarkan nilai *Mean Absolute Error* (MAE) sebesar 21,39, *Mean Squared Error* (MSE) sebesar 642,85, dan *Root Mean Squared Error* (RMSE) sebesar 25,35 untuk *Forecasting* menggunakan metode *Support Vector Regression* (SVR) pada data transaksi penjualan terakhir yaitu 31 Maret 2024 dan prediksi 3 bulan atau 12 minggu ke depan, dapat disimpulkan bahwa kinerja *Forecasting* menghasilkan hasil yang baik ketika prediksi awal, namun apabila digunakan pada jangka waktu panjang SVR perlu di evaluasi lebih jauh. Karena hasilnya sangat dipengaruhi oleh persebaran data pada *dataset*.

Referensi

- [1] Lindvall, J., "Business Intelligence Implementation from a User Perspective-A case study of an international company implementing BI in the procurement function Malin Lisa Jansson" 2021.
- [2] De Araujo, M., Priadana, S., Paramarta, V., & Sunarsi, D., "Digital leadership in business organizations: an overview" *International Journal of Educational Administration*, vol. 2, no. 1, 2021.
- [3] Tzedec, B. C., Neacșu, M., & Bucur, M., "Analysis of Knowledge and Implementation of Business Intelligence Solutions in Romanian SMEs" 2022.
- [4] Ranjbarfarid, M., & Hatami, Z., "Critical success factors of BI project implementation: An implementation methodology perspective" *Interdisciplinary Journal of Information, Knowledge, and Management*, vol. 15, pp. 175–202, 2020
- [5] Muchlis, T. I., Susanti, N., & Gumilar, I., "The Factors Affecting The Performing Of Indonesian Accounting Standards For Non-Publicly Accountable Entities (Sak Etap) In Smes Assisted By Chamber Of Commerce And Industry (Kadin) Of West Java Province, Indonesia" *Turkish Journal of Computer and Mathematics Education*, vol. 12, no. 8, pp. 1451–1455, 2021.
- [6] Eryadi, R. A., & Hidayanto, A. N., "Critical success factors for business intelligence implementation in an enterprise resource planning system environment using dematel: A case study at a cement manufacture company in Indonesia" *Journal of Information Technology Management*, 2020.
- [7] Zafary, F., "Implementation of business intelligence considering the role of information systems integration and enterprise resource planning" 2020.
- [8] Skyrius, R., "Business Intelligence: A Comprehensive Approach to Information Needs, Technologies and Culture" Springer, 2021.
- [9] Choi, L. K., Panjaitan, A. S., & Apriliasari, D., "The Effectiveness of Business Intelligence Management Implementation in Industry 4.0" *Startuppreneur Business Digital (SABDA Journal)*, vol. 1, no. 2, pp. 115–125, 2022.
- [10] Li, X., & Mao, Y., "Real-Time Data ETL Framework for Big Real-Time Data Analysis" 2015.
- [11] Putra, A. L., & Kurniawati, A., "Analisis Prediksi Harga Saham PT. Astra International Tbk Menggunakan Metode Autoregressive Integrated Moving Average (ARIMA) dan Support Vector Regression (SVR)" *Jurnal Ilmiah Komputasi*, vol. 20, no. 3, 2021.
- [12] P. W. A. Wibawa dan C. Pramatha, "Systematic Literature Review: Machine Learning Methods in Emotion Classification in Textual Data," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 12, no. 3, hal. 425–433, 2023, doi: 10.32736/sisfokom.v12i3.1787.
- [13] Lopes, J., Guimarães, T., & Santos, M. F., "Adaptive business intelligence: A new architectural approach" *Procedia Computer Science*, vol. 177, pp. 540–545, 2020.
- [14] Pramatha, C., Mahendra, I. M. Y., Rajeg, G. P. W., & Arka, I. W. (2023). The Development of Semantic Dictionary Prototype for the Balinese Language. *International Journal of Cyber and IT Service Management*, 3(2), 96-106.
- [15] Moss, L. T., & Atre, S., "Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications" Addison-Wesley, 2003.
- [16] C. Pramatha, J. G. Davis, dan K. K. Y. Kuan, "A Semantically-Enriched Digital Portal for the Digital Preservation of Cultural Heritage with Community Participation," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11196 LNCS, no. October, hal. 560–571, 2018, doi: 10.1007/978-3-030-01762-0_49.

Penerapan Multi Sensor Berbasis ESP8266 dalam Pengembangan Sistem Pemantauan Siswa di Sekolah

K.Y.V. Pradnyaditha^{a1}, AAIN Eka Karyawati^{a2}

Program Studi Teknik Informatika, Jurusan Informatika,
Fakultas Matematika Dan Ilmu Pengetahuan Alam, Universitas Udayana
Jalan Menuri Gg III, Indonesia
¹vidyayoga7@gmail.com
²eka.karyawati@unud.ac.id

Abstract

Penelitian ini mengembangkan sistem pemantauan siswa dengan sensor multi berbasis ESP8266 untuk memantau kehadiran dan lokasi siswa di sekolah. Data dari sensor RFID dikumpulkan, diproses, dan ditampilkan pada halaman website yang telah dibuat. Setelah dilakukan pengujian, hasil menunjukkan bahwa sistem data logger server (direct) menunjukkan kinerja yang optimal pada malam hari, dengan rata-rata waktu respon mencapai 55 detik dan kecepatan transmisi tertinggi mencapai 0,6 kartu/detik. Pada pagi hari, rata-rata waktu respon adalah 116,9 detik dengan kecepatan transmisi tertinggi 0,2 kartu per detik, sedangkan pada siang hari, rata-rata waktu respon mencapai 96,8 detik dengan kecepatan transmisi tertinggi sebesar 0,23 kartu/detik. Hasil yang didapaat dari uji t menunjukkan bahwa tidak ada perbedaan yang signifikan dalam kecepatan data antara konfigurasi data-logger client dan server. Pengujian fitur web juga menunjukkan bahwa keseluruhan sistem berjalan sesuai dengan test case yang telah ditetapkan, dengan fungsionalitas situs web absensi yang telah teruji dan dapat digunakan dengan baik.

Keywords: Pemantauan Siswa, Multi Sensor, ESP8266, Kehadiran, Lokasi, Keamanan, IoT.

1. Pendahuluan

Pendidikan memiliki peran yang krusial dalam kemajuan suatu negara. Di era digital seperti sekarang ini, penggunaan teknologi dalam bidang pendidikan semakin dibutuhkan untuk meningkatkan efektivitas dan efisiensi dalam proses pembelajaran. Salah satu permasalahan yang sering terjadi di dunia pendidikan adalah memantau kehadiran siswa secara efektif. Kehadiran siswa menjadi salah satu parameter penting dalam mengevaluasi kinerja siswa dan mengukur efektivitas pengajaran. Namun, memantau kehadiran siswa secara manual dapat memakan waktu dan tenaga yang besar bagi staf sekolah.

Salah satu contoh kasus nyata terjadi di SD Negeri 3 Batur, Batur Utara, Kecamatan Kintamani, Kabupaten Bangli, Bali, Indonesia. Sekolah tersebut masih melakukan absensi secara manual, yang menyebabkan masalah saat orang tua siswa menanyakan lokasi terakhir anak mereka. Dengan menerapkan sistem absensi elektronik, orang tua siswa dan guru dapat memonitor lokasi siswa setiap kali mereka memasuki dan meninggalkan kelas. Hal ini dapat dilakukan karena setiap siswa harus melakukan absensi ketika masuk dan pulang sekolah. Selama observasi di SD Negeri 3 Batur, penelitian mengalami kendala terkait cuaca selama musim hujan, yang menyebabkan gangguan pada koneksi Wifi. Meskipun demikian, dengan perencanaan yang cermat dan mungkin penerapan alternatif seperti koneksi internet cadangan atau penyempurnaan infrastruktur, implementasi sistem absensi elektronik masih dapat dijalankan secara efektif untuk meningkatkan efisiensi dan keamanan di lingkungan sekolah.

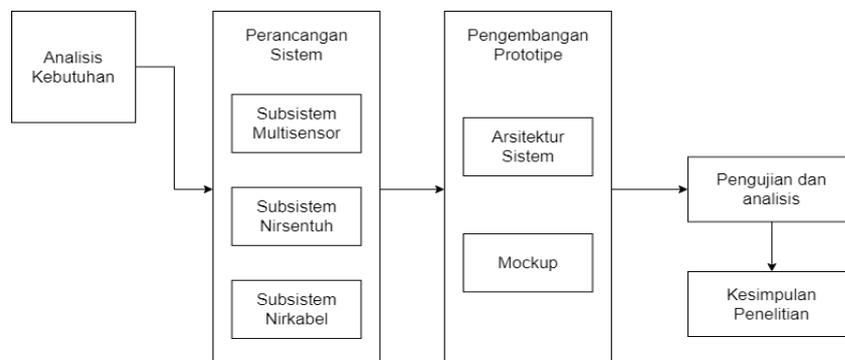
Dalam rangka meningkatkan efisiensi dan akurasi pemantauan kehadiran siswa, dikembangkanlah Sistem Monitoring Siswa dengan Multi-Sensor berbasis ESP8266. Sistem ini dapat memantau kehadiran siswa secara otomatis menggunakan teknologi RFID dan memanfaatkan mikrokontroler ESP8266 sebagai pusat pengolahan data dan menggunakan

data logger untuk merekam data sensor secara terus menerus dalam waktu tertentu. Pada penelitian (Sedghy, 2019) menggunakan teknologi RFID di lingkungan sekolah yang memiliki tujuan untuk mencatat kehadiran atau absensi siswa [1]. Kehadiran siswa dianggap sebagai nilai penting sehingga dapat mempengaruhi nilai akademik dan kinerja di sekolah. Penelitian (Siregar et al., 2024) menggunakan metode data logger untuk merekam data yang telah diukur ke dalam Secure-Digital Card yang mana data tersebut dapat diolah nantinya [2]. Dimana tujuan dari penelitian ini untuk membandingkan nilai akurasi dari radiasi matahari yang diperoleh dari alat konvensional dan alat yang telah dirancang.

Penelitian ini bertujuan untuk merancang sebuah sistem yang cocok untuk memantau keaktifan siswa dengan menggunakan teknologi RFID di beberapa lokasi. Sistem tersebut diimplementasikan dalam bentuk prototipe yang menggunakan teknologi multisensor, termasuk teknologi RFID, mikrokontroler, dan teknologi Wifi untuk melakukan pencatatan otomatis. Penelitian ini juga membandingkan kecepatan transmisi ke database dengan menggunakan Wifi atau disebut data-logger sebagai client (indirect) dan menggunakan komunikasi langsung dua perangkat antara server (komputer) dan alat atau disebut data-logger sebagai server (direct).

2. Metode Penelitian

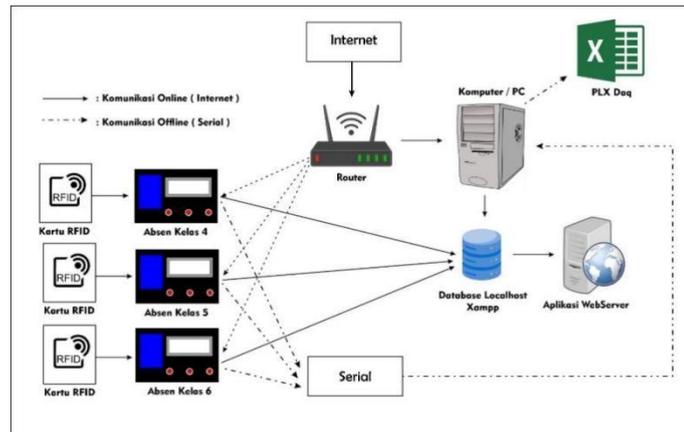
Penelitian ini melibatkan perancangan dan pengujian prototipe sistem dengan melakukan percobaan. Pada penelitian ini, perancangan sistem monitoring menggunakan metode waterfall yang telah dimodifikasi [3]. Hal ini dapat dilihat pada Gambar 1 yang menunjukkan tahapan-tahapan yang telah dimodifikasi dalam metode waterfall tersebut. Langkah-langkah yang akan dilakukan sebagai berikut:



Gambar 1. Konsep multisensor

Pada Gambar 1, merupakan proses kerja multisensor yang melibatkan beberapa sensor yang berfungsi untuk mengumpulkan data dari lingkungan sekitarnya. Data yang dikumpulkan oleh sensor-sensor tersebut kemudian diteruskan ke sistem akuisisi data, di mana informasi yang diterima lalu dianalisis secara menyeluruh [4]. Proses analisis ini melibatkan beberapa tahapan, termasuk korelasi data, fusi data, dan inversi bersama.

Pertama, dalam tahap korelasi data, berbagai dataset yang dikumpulkan dari sensor-sensor yang berbeda diselaraskan dan disinkronisasi untuk memastikan kesesuaian dan konsistensi dalam informasi yang diperoleh. Langkah ini membantu memperjelas hubungan antara data yang diperoleh dari sensor-sensor yang berbeda [5]. Kemudian, dalam tahap fusi data, informasi dari berbagai sensor digabungkan atau disintesis untuk memberikan gambaran yang lebih lengkap dan akurat tentang lingkungan atau objek yang diamati. Proses fusi data ini memungkinkan untuk mendapatkan pemahaman yang lebih dalam dan komprehensif tentang situasi yang diamati. Terakhir, tahap inversi bersama melibatkan analisis lebih lanjut untuk mendapatkan pemahaman yang lebih mendalam tentang fenomena atau objek yang diamati.

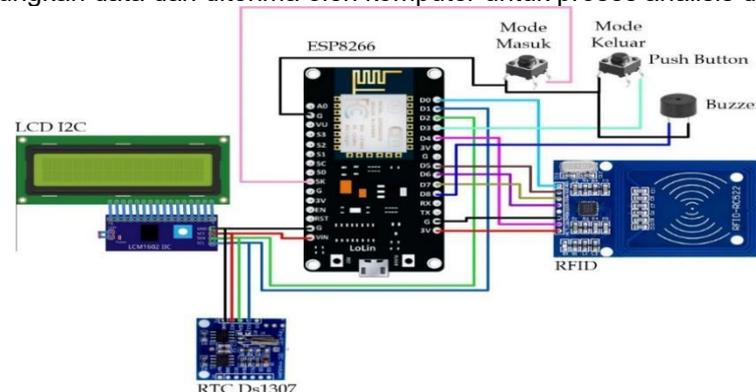


Gambar 2. Arsitektur Sistem Absensi

Pada pengembangan prototipe menggunakan arsitektur sistem yang digambarkan pada Gambar 2, dimana sistem ini terdiri dari beberapa komponen utama yang bekerja bersama untuk mencatat kehadiran siswa dan mengirimkan data ke database serta aplikasi web server, serta output data serial ke data logger server yang kemudian dikoneksikan secara offline ke PC. Dengan arsitektur ini, sistem dapat menghasilkan data kehadiran siswa secara akurat dan efisien, serta memungkinkan pengguna untuk memantau kehadiran siswa secara real-time melalui aplikasi web, atau menganalisis data secara lebih mendalam dengan mengakses data logger server melalui PC.

3. Hasil dan Pembahasan

Pada implementasi multisensor pada “Sistem Monitoring Siswa Berbasis ESP8266” berfungsi untuk mendapatkan output nilai yaitu data dari 3 sensor yang telah dipasang dalam masing-masing ruang kelas yang mana dalam kasus ini kelas 4, 5, dan 6. Pada desain sistem terdiri dari tiga bagian utama yaitu RFID, data-logger, dan database. Sistem ini memiliki fungsi untuk merekam data dari RFID, kemudian di transmisikan ke data-logger lalu data diakuisisi ke komputer melalui PLX-DAQ. Pengembangan prototipe, menghasilkan sistem yang terdiri dari tiga modul mikrokontroler ESP8266 yang masing masing di pasang di ruang kelas. Pada Gambar 3 ditunjukkan rangkaian sistem monitoring siswa dengan multisensor berbasis ESP8266. Komunikasi data pada sistem ini dimulai dari RFID sebagai sensor nirsentuh, lalu ESP8266 untuk komunikasi nirkabel. Kemudian data-logger mengakuisisi data dan menggabungkan data dan diterima oleh komputer untuk proses analisis data.

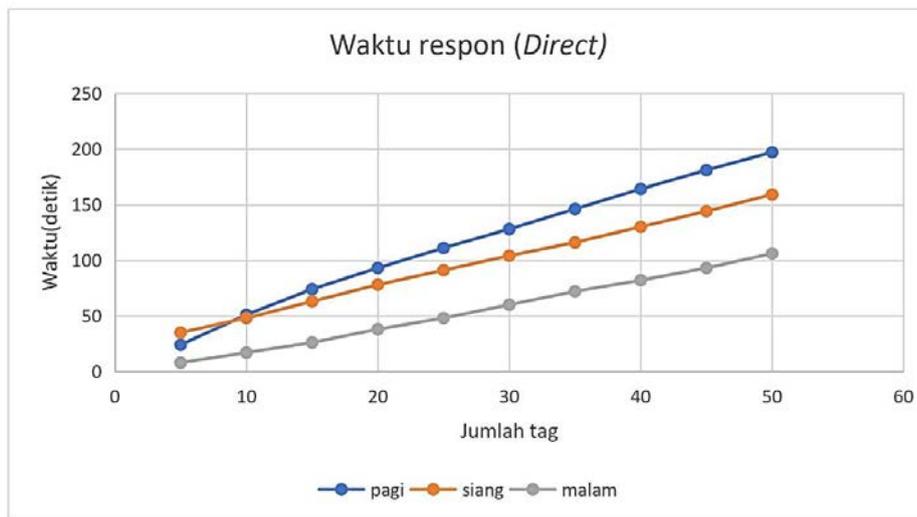


Gambar 3. Rangkaian Sistem

Saat membaca data dari RFID, deteksi RFID ditandai dengan buzzer. RFID dipasang di tiga ruangan dan dihubungkan dengan dengan modul ESP8266. Informasi yang diperoleh berupa kode unik dari tag RFID, lokasinya, waktu dan informasinya disimpan dalam database. Tag RFID dikirim melalui komunikasi nirkabel, dan sistem juga menandakan siswa keluar dan masuk dengan menekan tombol switch masuk dan keluar.

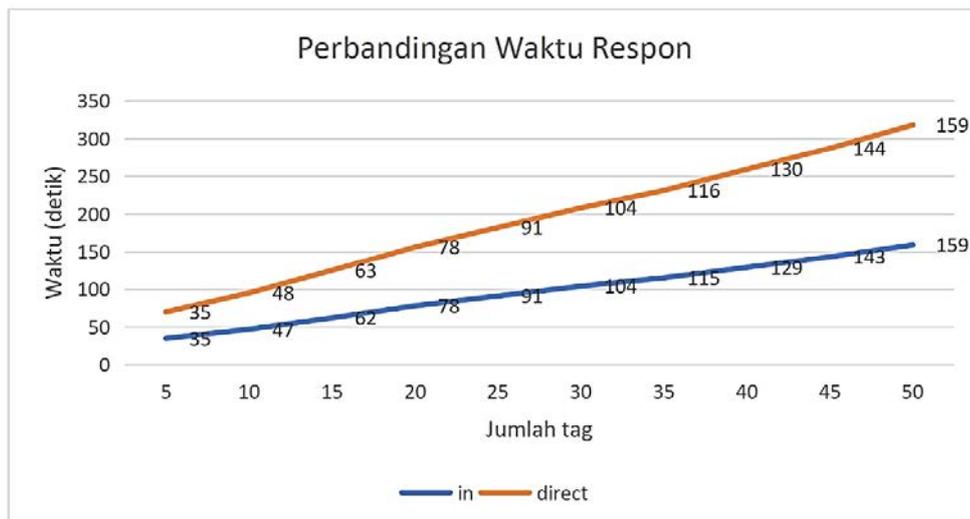
Pengujian dilaksanakan dengan mencatat interval waktu dari pembacaan, pengiriman, hingga penerimaan data. Proses pengujian ini diulang sebanyak 10 kali dengan peningkatan setiap kelipatan 5 pada setiap RFID-tag. Hasil pengujian menunjukkan bahwa

waktu respons langsung pada waktu malam mengalami peningkatan, menandakan perbaikan dalam kinerja sistem. Ini dapat dijelaskan oleh dua faktor utama. Pertama, gangguan elektromagnetik yang umumnya lebih rendah pada malam hari, ketika aktivitas listrik dan kendaraan berkurang, menghasilkan kualitas sinyal yang lebih stabil antara perangkat dan data-logger. Kedua, suhu yang lebih rendah di malam hari meningkatkan efisiensi operasi komponen elektronik, memungkinkan pengiriman data yang lebih cepat. Oleh karena itu, peningkatan waktu respon langsung pada malam hari dapat diatribusikan pada penurunan gangguan elektromagnetik dan kondisi suhu yang mendukung kinerja perangkat keras elektronik.



Gambar 4. Grafik Waktu Respon (Direct)

Pada Gambar 4 menunjukkan data waktu respon terkait kenaikan waktu pada konfigurasi langsung (Direct) terlihat konstan.



Gambar 5. Grafik Perbandingan Waktu Respon

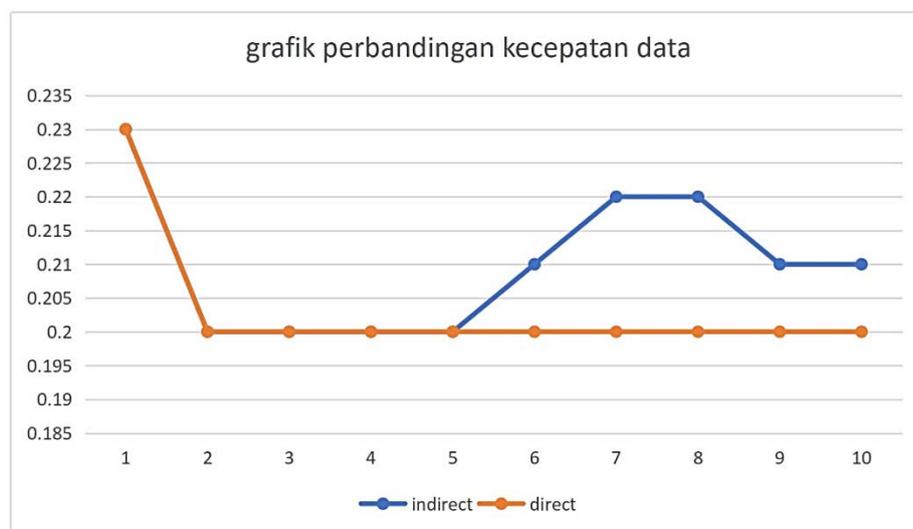
Pada Gambar 5 menunjukkan data perbandingan waktu respon, terdapat dua garis linier biru yang menandakan data konfigurasi tidak langsung (indirect) dan orange menandakan data konfigurasi langsung (direct). Berdasarkan data hasil tersebut, terlihat konfigurasi tidak langsung lebih unggul dengan persenan selisih sebesar 0,5% yang mana hal ini mungkin saja bisa terjadi dikarenakan pada konfigurasi indirect memiliki wifi yang stabil dan pada konfigurasi direct memerlukan software tambahan yaitu PLX-DAQ. Penelitian juga dilakukan dilakukan ketika malam hari dan menghasilkan selisih yang sama yaitu 0,5% antar dua konfigurasi tersebut.

Group Statistics				
Konfigurasi	N	Mean	Std. Deviation	Std. Error Mean
Nilai 1	10	116.5000	56.97807	18.01805
2	10	116.9000	56.89454	17.99163

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Nilai	Equal variances assumed	.000	.988	-.016	18	.988	-.40000	25.46270	-53.89514	53.09514
	Equal variances not assumed			-.016	18.000	.988	-.40000	25.46270	-53.89515	53.09515

Gambar 6. Hasil Uji T waktu Respon

Pada Gambar 6 menunjukkan data hasil uji t untuk waktu respon, tidak ada perbedaan yang signifikan antara nilai konfigurasi langsung dan tidak langsung, baik ketika diasumsikan bahwa variansa sama maupun tidak sama. Ini dapat disimpulkan dari nilai p yang tinggi (lebih besar dari 0.05), yang menunjukkan tidak adanya bukti yang cukup untuk menolak hipotesis nol bahwa tidak ada perbedaan yang signifikan antara kedua kelompok tersebut.



Gambar 7. Grafik Perbandingan Kecepatan Data

Gambar 7 menampilkan perbandingan kecepatan data antara kedua konfigurasi. Garis oranye menunjukkan konfigurasi langsung (direct), sedangkan garis biru mewakili konfigurasi tidak langsung (indirect). Dari grafik tersebut, terlihat bahwa kecepatan data pada konfigurasi tidak langsung lebih tinggi daripada konfigurasi langsung. Ini menunjukkan bahwa komunikasi pada konfigurasi client lebih cepat daripada konfigurasi server.

Group Statistics				
Konfigurasi	N	Mean	Std. Deviation	Std. Error Mean
Nilai 1	10	.21	.011	.003
2	10	.21	.013	.004

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Nilai	Equal variances assumed	.262	.615	.768	18	.452	.004	.005	-.007	.015
	Equal variances not assumed			.768	17.433	.453	.004	.005	-.007	.015

Gambar 8. Hasil Uji T Kecepatan Data

Pada Gambar 8 menunjukkan data hasil uji t untuk kecepatan data yang tidak terdapat perbedaan yang signifikan antara kecepatan data dari kedua kelompok, baik saat varian diasumsikan sama maupun tidak sama. Hal ini diperkuat oleh nilai p yang tinggi (lebih

besar dari 0.05), menandakan bahwa tidak ada bukti yang cukup untuk menolak hipotesis nol bahwa tidak ada perbedaan yang signifikan antara kedua kelompok tersebut.

Setelah merancang dan melakukan pengujian, hasilnya dicatat dalam form pengujian. Kemudian, melakukan analisis dan menarik kesimpulan berdasarkan penilaian keselarasan dengan standar yang diharapkan. Apabila hasil sesuai akan diberi kesimpulan bahwa pernyataan valid dan begitu juga sebaliknya. Pengujian dilakukan terhadap website absensi yang mencakup 16 pengujian yang berfokus pada berbagai menu dan fitur yang disediakan oleh website tersebut. Setiap pengujian dirancang untuk menguji fungsionalitas dan responsivitas dari menu-menu utama, termasuk penambahan data baru, proses absensi siswa, serta manajemen data siswa. Dengan demikian, pengujian mencakup aspek-aspek krusial dalam operasional dan penggunaan sehari-hari dari sistem monitoring kehadiran siswa ini. Proses pengujian dilakukan dengan cermat dan terstruktur, menggunakan serangkaian test case yang telah dipersiapkan sebelumnya. Setiap langkah dalam pengujian direkam dan dievaluasi untuk memastikan bahwa fungsi-fungsi yang diuji berjalan sesuai dengan yang diharapkan dan tidak menghasilkan kesalahan atau kegagalan dalam prosesnya.

4. Kesimpulan

Berdasarkan hasil penelitian mengenai implementasi Sistem Monitoring Siswa dengan Multi-Sensor berbasis ESP8266, dapat disimpulkan bahwa penggunaan teknologi RFID dan sensor lainnya secara efektif meningkatkan efisiensi dalam memantau kehadiran siswa di sekolah. Hasil pengujian dengan sistem langsung, kinerja terbaik terjadi pada malam hari dengan waktu respons rata-rata sebesar 55 detik dan kecepatan transmisi tertinggi mencapai 0,6 kartu per detik. Di pagi hari, waktu respons rata-ratanya adalah 116,9 detik dengan kecepatan transmisi tertinggi 0,2 kartu per detik, dan pada siang hari, waktu respons rata-ratanya mencapai 96,8 detik dengan kecepatan transmisi tertinggi sebesar 0,23 kartu per detik. Berdasarkan hasil uji t, tidak ada perbedaan signifikan dalam kecepatan data antara konfigurasi data-logger client dan server. Nilai p yang tinggi menunjukkan bahwa tidak ada bukti cukup untuk menolak hipotesis nol bahwa tidak ada perbedaan signifikan antara kedua konfigurasi tersebut. Oleh karena itu, dari segi kecepatan transmisi dan waktu respons, kedua konfigurasi data-logger dapat dianggap memiliki kinerja yang serupa. Selanjutnya, pengujian fitur web menunjukkan bahwa keseluruhan sistem berjalan sesuai dengan test case yang telah ditetapkan. Fungsionalitas situs web absensi telah diuji dan dapat digunakan dengan baik. Dengan demikian, hasil penelitian ini memberikan pemahaman yang lebih baik tentang kinerja sistem dan menunjukkan bahwa sistem ini dapat diandalkan dalam memantau kehadiran siswa dengan efisien dan efektif.

Referensi

- [1] B. M. Sedghy, "Evolution of Radio Frequency Identification (RFID) in Agricultural Cold Chain Monitoring: A Literature Review," *J. Agric. Sci*, vol. 11, no. 3, p. 43, 2019.
- [2] N. S. A. A. A. F. B. A. -S. H. A. H. D. Rjeib, "Attendance and information system using RFID and web-based application for academic sector," *Int. J. Adv. Comput. Sci. Appl*, vol. 9, no. 1, p. 266–274, 2018.
- [3] N. H. M. F. A. W. Siregar, "Rancang Bangun Alat Ukur Radiasi Matahari Berbasis Arduino Uno Dan Data Logger," *J. Tek. Fis*, 2024.
- [4] Y. W. Q. N. Z. Wang, "Multi-Sensor Fusion in Automated Driving: A Survey," *IEEE Access*, vol. 8, p. 2847–2868, 2020.
- [5] Y. A. F. A. A. N. Z. Zakiah, "Sistem Multi-Sensor Nirkabel Berbasis RFID Untuk Pemantauan Keaktifan Siswa," *J. Rekayasa Elektr*, vol. 15, no. 3, 2020.

Implementasi Business Intelligence Menggunakan Metode OLAP untuk Mengambil Keputusan Penjualan Properti

Lidya Elisabet Theogracia Silitonga^{a1}, I Putu Gede Hendra Suputra^{a2},
Ida Ayu Gde Suwiprabayanti Putra, I Ketut Gede Suhartana^{a4}

Informatics Department, Udayana University
Jimbaran, Indonesia

¹lidaelisabet83gmail.com

²hendra.suputra@gmail.com

³iagsuwiprabayantiputra@unud.ac.id

⁴ikg.suhartana@unud.ac.id

Abstract

Based on data from the Central Statistics Agency, population growth in Indonesia reaches 2 million – 3 million people every year. In this case, with increasing population growth, the need for housing also increases. The high demand for housing provides opportunities for property companies, one of which is PT Anugerah Indah Prima in the city of Medan. Apart from providing great opportunities, business in the property sector also has intense competition so that in order to continue to progress and develop, PT Anugerah Indah Prima must continue to be innovate in developing its business. In making a decision to develop its business, the company still faces obstacles, namely processing property sales data manually using Microsoft Excel. This data processing can not meet the needs of managers and marketing parties in making decision because the reports produced are still general information and difficult to understand.

Therefore, this research focuses on designing a dashboard that can provide insightful information in company decision making. This research utilizes business intelligence, ETL(extract, transform, and loading), OLAP(online analytical processing), and forecasting.

Keywords: Information Dashboard, Business Intelligence, Property Business, OLAP, ETL

1. Pendahuluan

Berdasarkan data Badan Pusat Statistik, pertumbuhan penduduk di Indonesia mencapai 2 juta – 3 juta jiwa per tahun. Dalam hal ini, adanya peningkatan pertumbuhan penduduk, maka kebutuhan akan rumah juga meningkat. Tingginya permintaan akan perumahan memberikan peluang bagi perusahaan *real estate*, termasuk PT Anugerah Indah Prima di Kota Medan. Bisnis di bidang *real estate* tidak hanya memiliki peluang besar, tetapi juga memiliki persaingan yang ketat sehingga upaya yang dimiliki, salah satu adalah kemampuan dalam menggunakan teknologi.

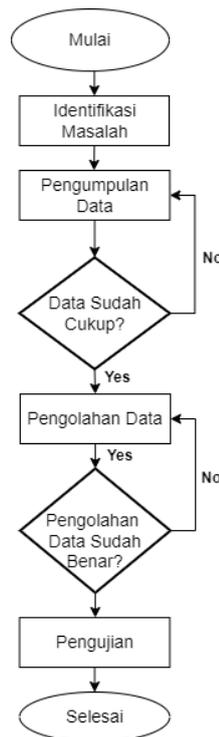
Untuk terus maju dan berkembang, PT Anugerah Indah Prima harus melakukan inovasi dalam mengembangkan bisnisnya. Dalam membuat suatu keputusan untuk mengembangkan bisnisnya, perusahaan tersebut masih menghadapi kendala, yaitu melakukan pengolahan data penjualan properti secara manual menggunakan Microsoft Excel. Pengolahan data tersebut belum dapat memenuhi kebutuhan manajer dan pihak marketing dalam pengambilan keputusan karena laporan yang dihasilkan masih berupa informasi umum dan sulit dipahami. Untuk mengatasi hal tersebut, pada penelitian ini akan memanfaatkan inovasi teknologi informasi, yakni *business intelligence* dalam mengolah data yang diharapkan dapat membantu PT Anugerah Indah Prima

dalam mengembangkan bisnis dan meningkatkan kualitasnya. Pada umumnya, *business intelligence* digunakan untuk pengambilan keputusan dari data yang berjumlah besar [1]. *Business intelligence* digunakan untuk membantu dalam menganalisis sebuah masalah dan dapat menentukan performa masing-masing usaha [2]. *Business intelligence* juga digunakan untuk mengukur kinerja perguruan tinggi dalam mengambil keputusan bagi pimpinan perguruan tinggi [3]. *Business Intelligence* digunakan untuk membuat keputusan penjualan properti di PT Jatiwangi Grahatama Properti [4]. *Business Intelligence* digunakan untuk menampilkan keuntungan pada data superstore [5].

Penelitian ini berfokus pada perancangan *dashboard* informasi untuk memvisualisasi data penjualan dan data penting lainnya menggunakan *business intelligence*. Hal tersebut bertujuan untuk meningkatkan daya saing dan kualitas pengambilan keputusan karena data yang tersimpan belum dapat digunakan secara maksimal untuk menambah wawasan perusahaan. Pada penelitian ini dibuat *dashboard* yang berisi visualisasi data-data yang dimiliki perusahaan, seperti data penjualan perumahan, keuntungan yang didapat dari penjualan tersebut, dan data-data yang berasal dari luar perusahaan, seperti data pertumbuhan penduduk Kota Medan dan pertumbuhan harga properti menggunakan Microsoft Power BI. Microsoft Power BI adalah salah satu bagian dari aplikasi *business intelligence* yang digunakan untuk meningkatkan wawasan terhadap instansi atau organisasi.

2. Metode Penelitian

Pada bagian ini membahas langkah-langkah yang akan dilakukan pada penelitian ini. Adapun tahapan penelitian ini, yaitu identifikasi masalah, pengumpulan data, pengolahan data, dan pengujian. Visual dari desain metode penelitian terdapat pada Gambar 1.



Gambar 1. Desain Metodologi Penelitian

2.1 Identifikasi Masalah

PT Anugerah Indah Prima merupakan perusahaan yang bergerak di bidang *real estate* di Kota Medan. Dalam membuat suatu keputusan untuk mengembangkan bisnisnya, perusahaan tersebut masih menghadapi kendala, yaitu melakukan pengolahan data penjualan properti secara manual menggunakan Microsoft Excel. Pengolahan data tersebut belum dapat memenuhi kebutuhan manajer dan pihak marketing dalam pengambilan keputusan karena laporan yang dihasilkan masih berupa informasi umum dan sulit dipahami. Perancangan *dashboard* informasi

dibuat untuk menyelesaikan permasalahan tersebut yang mana *dashboard* dapat memberikan *insight* dari data-data yang dimiliki yang mana data-data yang digunakan dalam metode tersebut merupakan data yang bersifat nyata, relevan dengan kebutuhan pasar, mudah dipahami dan diolah sehingga dapat memberikan wawasan mengenai kondisi yang sebenarnya terjadi di pasar.

2.2 Metode Pengumpulan Data

Dalam penelitian ini menggunakan dua sumber data, yaitu data primer dan data sekunder. Data primer merupakan data yang dikumpulkan secara langsung di lapangan. Pada penelitian ini, data yang akan dikumpulkan adalah data penjualan dua perumahan, yaitu Perumahan Mencirim Indah dan Perumahan Griya Harapan Indah. Data sekunder merupakan data yang tidak didapat dari tempat objek penelitian dan bersumber dari website, buku, jurnal, dan sebagainya yang mana data tersebut digunakan sebagai pelengkap dan memperkuat penelitian. Data sekunder tersebut, seperti data jumlah penduduk Kota Medan dan laju pertumbuhan penduduk di setiap wilayah Kota Medan, data penduduk Kota Medan berdasarkan umur, data penduduk Kota Medan Umur 15+ yang sudah bekerja, data UMP masyarakat Kota Medan, dan data indeks harga properti Kota Medan. Adapun teknik pengumpulan data primer dan sekunder, yaitu wawancara, kuesioner, studi literatur, dan observasi. Tabel 1 merupakan salah satu data primer yang digunakan dalam penelitian ini.

Tabel 1. Data Penjualan Rumah

Nama Perumahan	Lebar Jalan	Tipe Ruma	Luas Bangunan
Griya Mencirim Indah	6 m	36/91	36 m ²
Griya Mencirim Indah	6 m	36/91	36 m ²
Griya Mencirim Indah	6 m	36/91	36 m ²
Griya Mencirim Indah	6 m	36/91	36 m ²
Griya Mencirim Indah	6 m	36/91	36 m ²

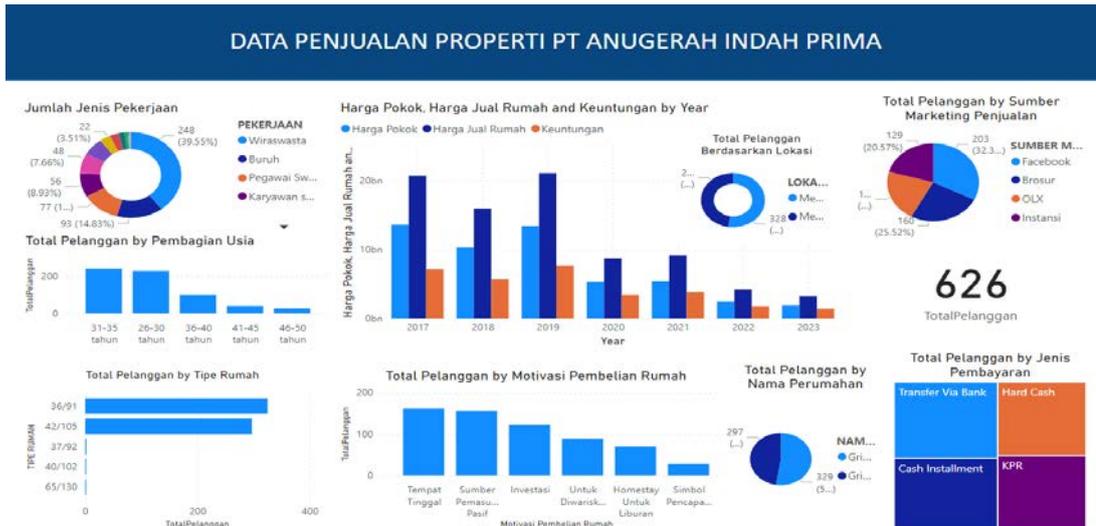
2.3 Pengolahan Data

Pada tahap ini dilakukan pengolahan data, seperti *extract, transform, loading*(ETL), pembuatan *data warehouse*, pembuatan *dashboard*, implementasi OLAP(*online analytical processing*), dan pengujian evaluasi sistem *dashboard*.

3. Hasil dan Pembahasan

3.1. Wawasan Data Penjualan Perumahan

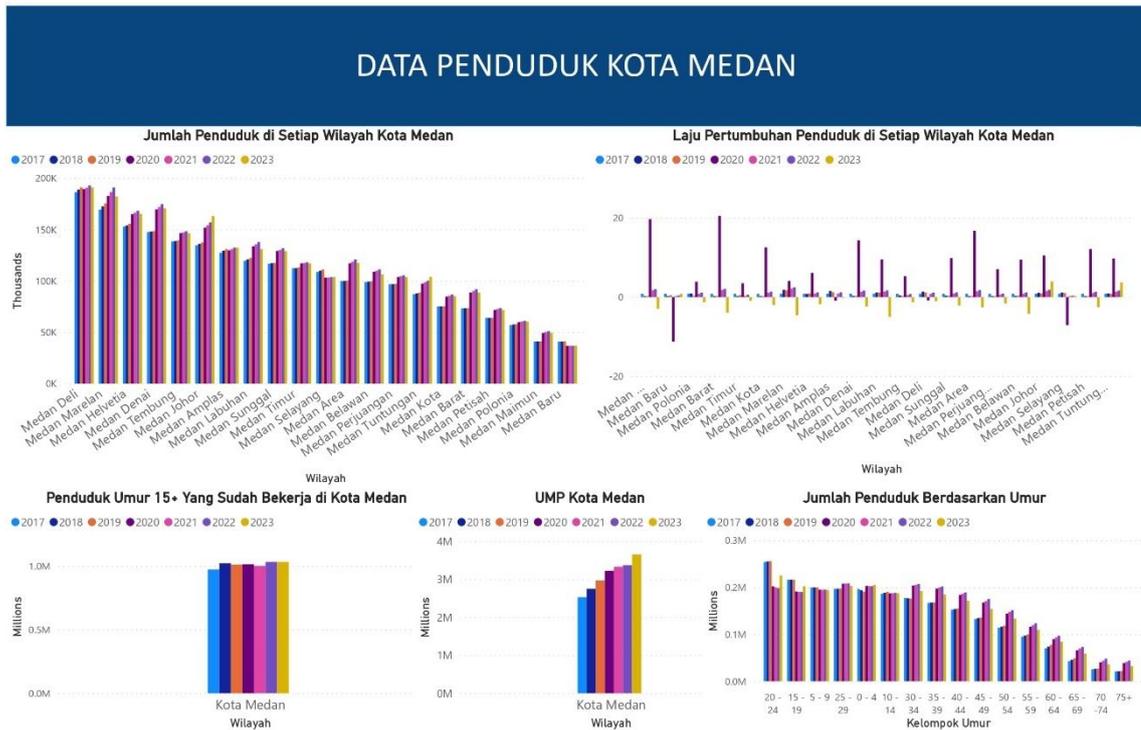
Gambar 2 merupakan dashboard visualisasi data penjualan perumahan PT Anugerah Indah Prima. Dari *dashboard business intelligence* tersebut, didapatkan bahwa dari jenis pekerjaan yang ada, jenis pekerjaan yang paling banyak dari pelanggan properti PT Anugerah Indah Prima adalah Wiraswasta. Adapun total pelanggan terbanyak berdasarkan usia berada di rentang usia 31-35 tahun. Untuk perumahan yang terjual ada 2, yaitu perumahan Griya Mencirim Indah dan perumahan Griya Harapan Indah yang berlokasi di Medan Sunggal dan Medan Tuntungan II. Selain itu, motivasi pembelian rumah pelanggan paling banyak adalah untuk tempat tinggal. Selain itu, keuntungan yang didapatkan PT Anugerah Indah Prima dari dua perumahan tersebut mengalami fluktuatif. Untuk jenis pembayaran yang digunakan oleh pembeli perumahan ada empat, yaitu transfer via Bank, Hard Cash, Cash Installment, KPR. Untuk jenis marketing penjualan terdapat empat jenis dan pembeli paling banyak mendapatkan informasi terkait penjualan perumahan melalui Facebook. Selain itu, terdapat 5 tipe rumah dan tipe rumah yang paling banyak disukai oleh masyarakat, yaitu tipe 36/91.



Gambar 2. Dashboard Data Penjualan Properti PT Anugerah Indah Prima

3.2. Wawasan Data Penduduk Kota Medan

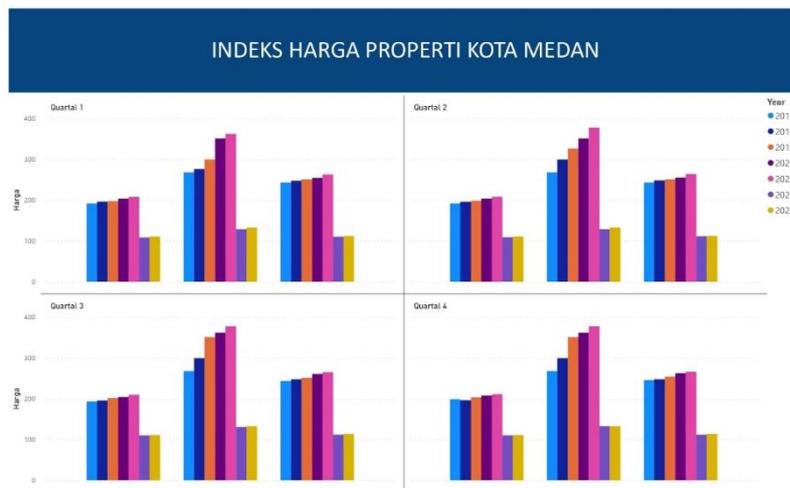
Gambar 3 adalah dashboard visualisasi data penduduk kota Medan. Pada dashboard terdapat 21 wilayah di kota Medan, terdapat empat wilayah dengan total penduduk paling banyak pada tahun 2017-2023, yaitu Medan Deli dengan jumlah penduduk adalah 1.514.775, diurutkan kedua adalah Medan Marelان dengan jumlah penduduk adalah 1.427.182, diurutkan ketiga adalah Medan Helvetia dengan jumlah penduduk adalah 1.278.556, dan diurutkan keempat adalah Medan Denai dengan jumlah penduduk adalah 1.277.283. Selain itu, empat wilayah dengan total penduduk paling sedikit pada tahun 2017-2023, yaitu Medan Baru dengan jumlah penduduk adalah 310.297, diurutkan kedua adalah Medan Maimun dengan jumlah penduduk adalah 363.860, diurutkan ketiga adalah Medan Polonia dengan jumlah penduduk adalah 470.306, dan diurutkan terakhir adalah Medan Petisah dengan jumlah penduduk adalah 545.050. Dari dashboard visualisasi tersebut dapat dilihat terdapat beberapa wilayah yang mengalami penurunan laju pertumbuhan, adapun empat wilayah dengan laju pertumbuhan penduduk yang mengalami penurunan signifikan, yaitu Medan Labuhan, Medan Marelان, Medan Belawan, dan Medan Barat. Selain itu, dapat dilihat juga bahwa bahwa jumlah penduduk umur 15+ yang sudah bekerja di Kota Medan mengalami fluktuatif yang mana salah satu penyebab hal tersebut adalah Pandemi Covid-19 yang terjadi di Indonesia. Pada dashboard tersebut dapat dilihat bahwa UMP Kota Medan meningkat setiap tahunnya yang mana pada tahun 2017, UMP Kota Medan adalah Rp 2.528.815. Pada tahun 2018, UMP Kota Medan adalah 2.749.074. Pada tahun 2019, UMP Kota Medan adalah 2.969.825.



Gambar 3. Dashboard Data Penduduk Kota Medan

3.3. Wawasan Data Indeks Harga Properti Kota Medan

Gambar 4 merupakan dashboard visualisasi Indeks Harga Properti Kota Medan. Dari dashboard visualisasi Indeks Harga Properti Kota Medan dapat dilihat bahwa Indeks Harga Properti Kota Medan mengalami peningkatan pada tahun 2023 di setiap kuartal dan disetiap jenis baik itu besar, kecil, dan menengah. Pada Quartal 1, jenis besar, Indeks Harga Properti pada tahun 2023 sebesar 110.55 yang mana pada tahun 2022, Indeks Harga Properti sebesar 108.22. Untuk jenis kecil, Indeks Harga Properti pada tahun 2023 adalah 132.70 yang mana pada tahun 2022, Indeks Harga Properti sebesar 128.55. Dan untuk jenis menengah, Indeks Harga Properti pada tahun 2023 adalah 111.71 yang mana pada tahun 2022, Indeks Harga Properti sebesar 109.95. Adanya peningkatan Indeks Harga Properti menunjukkan bahwa minat pembeli dalam membeli rumah juga mengalami peningkatan



Gambar 4. Dashboard Indeks Harga Properti Kota Medan

3.4. Pengujian Data dan Evaluasi Dashboard

Berikut ini pembahasan tentang pengujian data untuk melakukan prediksi (*forecasting*) harga jual rumah dengan *support vector regression*(SVR) dan pengujian *dashboard* menggunakan *usability testing*.

3.4.1. Pengujian Data

Pengujian data ini bertujuan untuk melakukan *forecasting* harga jual rumah di tahun berikutnya dan untuk mengetahui apakah data yang digunakan dalam melakukan prediksi harga jual rumah sudah cukup atau tidak. *Forecasting* harga jual rumah dilakukan menggunakan algoritma *Support Vector Regression* yang mana proses *forecasting* dilakukan dengan menggunakan data sebelumnya(data *actual*). Untuk melakukan *forecasting* harga jual rumah pada tahun 2019 diperlukan data tahun 2017 dan 2018, untuk melakukan prediksi harga jual rumah pada tahun 2020 diperlukan data tahun 2018,2019, dan 2020, dan seterusnya. Pada tabel 2 menunjukkan bahwa prediksi harga jual rumah cenderung rendah daripada harga jual rumah sebelumnya. Pada tahun 2019, hasil prediksi harga jual rumah adalah 126.345.600, sedangkan harga jual rumah sebenarnya adalah 130.000.000 yang mana MAE pada tahun tersebut 3.425.170. Hal tersebut menunjukkan tingkat kesalahan prediksi relatif rendah. Namun, pada tahun 2022, prediksi harga rumah tidak akurat yang mana dapat dilihat bahwa hasil prediksi harga jual rumah 137.162.604 dan harga jual rumah sebenarnya adalah 150.000.00, dan MAE adalah 10.333.460. Adanya ketidakuratan prediksi karena kurangnya data penjualan harga rumah dan diperlukan data pendukung lainnya untuk melakukan prediksi harga jual rumah yang mana pada penelitian ini untuk melakukan prediksi harga jual rumah hanya menggunakan data tanggal akad kredit, data harga jual rumah setiap tahunnya, dan data inflasi. Tabel 2 merupakan tabel hasil prediksi harga jual rumah, MAE(*Mean Absolute Error*), dan harga jual rumah sebenarnya di PT Anugerah Indah Prima.

Tabel 2. Hasil Prediksi Harga Jual Rumah dan MAE

Tahun	Harga Jual Rumah Sebenarnya	Hasil Prediksi Harga Jual Rumah	MAE(<i>Mean Absolute Error</i>)
2019	Rp 130.000.000	Rp 126.345.600	3.425.170
2020	Rp 140.000.000	Rp 132.183.999	5.568.353
2021	Rp 150.000.000	Rp 132.430.999	6.236.397
2022	Rp 150.000.000	Rp 137.162.604	10.333.460
2023	Rp 150.500.000	Rp 133.393.905	8.825.572

3.4.2. Pengujian Dashboard

Pengujian *dashboard* dilakukan menggunakan kuesioner *usability testing* untuk mengetahui apakah *dashboard* informasi sudah sesuai dengan kebutuhan perusahaan dan apakah perusahaan dapat dengan mudah menggunakan desain *dashboard* informasi yang telah dibuat. Hasil kuesioner dari 5 responden perusahaan dengan 5 variabel *usability* terdapat pada Tabel 3.

Tabel 3. Nilai Rata-Rata Variabel *Usability*

No	Variabel	Rata-Rata	Kategori
1	Learnability	4.50	Sangat Baik
2	Efficiently	4.75	Sangat Baik
3	Memorability	4.40	Sangat Baik
4	Error	2.20	Kurang
5	Satisfaction	4.10	Baik
	Mean	3.99	Baik

Dari hasil keseluruhan nilai rata-rata variabel yang terdapat pada Tabel 3 menunjukkan bahwa Usability dari aplikasi Dashboard Business Intelligence tergolong baik dengan *mean* adalah 3.99.

4. Kesimpulan

Berdasarkan hasil penelitian tersebut, maka simpulan dalam penelitian ini adalah:

1. Dari implementasi *business intelligence* tersebut dihasilkan *dashboard* informasi yang dapat memberikan wawasan untuk membantu perusahaan dalam membuat keputusan.
2. Dari *dashboard* penduduk Kota Medan dapat diketahui bahwa Medan Baru merupakan salah satu lokasi dengan jumlah penduduk paling sedikit di Wilayah Kota Medan sehingga pihak PT Anugerah Indah Prima dapat mengembangkan bisnis mereka di daerah tersebut.
3. Dari dashboard tersebut dapat dilihat bahwa Indeks Harga Properti di Kota Medan mengalami peningkatan setiap tahunnya yang mana dalam hal ini menunjukkan bahwa minat pembeli dalam membeli rumah juga mengalami peningkatan.
4. Berdasarkan penelitian di atas diketahui bahwa MAE dalam prediksi harga jual rumah mengalami peningkatan setiap tahunnya yang mana adanya ketidakuratan prediksi karena kurangnya data penjualan harga rumah dan diperlukan data pendukung lainnya untuk melakukan prediksi harga jual rumah.
5. Dari hasil pengujian dashboard menggunakan usability testing didapat bahwa nilai rata-rata variabel menunjukkan bahwa Usability dari aplikasi Dashboard Business Intelligence tergolong baik dengan *mean* adalah 3.99.

Referensi

- [1] He Hanifah, A. L. (2020). *PEMBANGUNAN BUSINESS INTELLIGENCE PADA TOSERBA KOPERASI KARYAWAN SEMEN PADANG (KKSP) BERBASIS DASHBOARD SYSTEM* (Doctoral dissertation, Universitas Andalas)
- [2] Purnama, A., Wijaya, I. N. Y. A., & Sudiatmika, I. B. K. (2022). PENERAPAN BUSINESS INTELLIGENCE UNTUK MENGANALISIS DATA PENJUALAN MENGGUNAKAN POWER BI. *Jurnal Informatika Teknologi dan Sains*, 4(2), 118-123
- [3] Poerbo, H., & Susilowati, M. (2020). Pemanfaatan Business Intelligence di Perguruan Tinggi. *Kurawal-Jurnal Teknologi, Informasi dan Industri*, 3(1), 40-57
- [4] SAPTOMO, A. B. (2022). Implementasi Business Intelligence Dalam Membuat Keputusan Penjualan Properti Dengan Metode Online Analytical Processing (Olap)(Studi Kasus: Pt. Jatiwangi Grahatama Properti).
- [5] Achmad, M., & Susilawati, D. (2020). Penerapan Business Intelligence Untuk Menampilkan Keuntungan Pada Data Superstore Dengan Menggunakan Metode Olap. *ALGOR*, 2(1), 48-56

This page is intentionally left blank.

Penerapan Metode Naive Bayes dalam Sistem Monitoring Akuaponik

Made Rahayu Setyaningrum^{a1}, I Komang Ari Mogi^{a2}, Agus Muliantara^{a3},
Ida Ayu Gde Suwiprabayanti Putra^{a4}

^aProgram Studi Informatika Fakultas Matematika Dan Ilmu Pengetahuan Alam, Universitas Udayana

¹made.rahayu21@gmail.com

²arimogi@unud.ac.id

³muliantara@unud.ac.id

⁴iagsuwiprabayantiputra@unud.ac.id

Abstrak

Akuaponik merupakan gabungan antara budidaya akuakultur dan hidroponik yang menciptakan sistem pertanian yang memanfaatkan ikan dan tanaman secara simbiosis mutualisme. Pentingnya menjaga keseimbangan pH dan ketinggian air dalam akuaponik untuk pertumbuhan yang optimal menjadi fokus utama. Untuk mengatasi tantangan tersebut, sebuah sistem website dirancang untuk memantau parameter seperti suhu, pH, kekeruhan, dan ketinggian air, serta memberikan informasi tentang kondisi lingkungan. Data yang digunakan dikumpulkan melalui node-red dan disimpan dalam PHPMyAdmin. Proses pengolahan data menggunakan algoritma Naive Bayes yang mengklasifikasikan kondisi menjadi normal atau tidak normal, dengan hasil pengujian k-fold cross validation dan confusion matrix menunjukkan akurasi sebesar 92%. Selain itu, pengujian kinerja sistem website dengan black box testing menunjukkan bahwa 7 dari 8 fitur berjalan dengan baik, menunjukkan potensi pengembangan lebih lanjut untuk optimalitas penggunaannya.

Keywords: *Akuaponik, node-red, naive bayes, k-fold cross validation, confusion matrix*

1. Pendahuluan

Peningkatan jumlah penduduk di Indonesia dan kemajuan teknologi berdampak pada menyusutnya lahan pertanian. Data BPS Indonesia 2023 menunjukkan peningkatan jumlah penduduk Indonesia menjadi 275 juta jiwa pada 2022, naik 1,13% dari tahun sebelumnya. Perkembangan ini dapat mengancam lahan pertanian, seperti terlihat dari data BPS 2022 tentang penutupan lahan hutan dan non-hutan. Namun, berkat perkembangan zaman, telah muncul teknik pertanian modern, salah satunya adalah teknik akuaponik.

Akuaponik menggabungkan akuakultur dan hidroponik yang menciptakan simbiosis antara ikan dan tanaman. Dalam sistem ini, kotoran ikan menjadi nutrisi bagi tanaman, sementara tanaman memberikan oksigen untuk ikan setelah proses penyaringan. Beberapa parameter penting dalam monitoring sebuah akuaponik yang sekaligus menjadi acuan dalam penelitian ini, yaitu suhu, pH, kekeruhan air, dan ketinggian air. Untuk menjaga keseimbangan parameter-parameter tersebut dalam pertanian akuaponik dapat penggunaan metode Naive Bayes yang dapat memprediksi peluang berdasarkan data input dengan probabilitas dan statistik yang cocok untuk data kategorial dan kontinu dengan kecepatan komputasi tinggi.

Beberapa penelitian terkait sistem monitoring pada akuaponik yang telah dilakukan di antaranya, "Sistem Monitoring Tanaman Hidroponik Berbasis Internet of Things" [1], membahas tentang sistem hidroponik yang memanfaatkan komunikasi data melalui Restful Architecture API dan menguji data sensor yang akan dikontrol. Adapun hasil dari penelitian ini menunjukkan bahwa waktu optimal untuk mengirimkan satu interval data dari mikrokontroler ke sistem adalah 30 detik. Penelitian selanjutnya berjudul "Implementasi IoT Cerdas Berbasis Inference Fuzzy Tsukamoto Pada Pemantauan Kadar pH dan Ketinggian Air Dalam Akuaponik" [2], membahas informasi tentang penerapan IoT dalam memantau kadar pH dan ketinggian air dalam sistem akuaponik dengan memanfaatkan logika fuzzy untuk mengolah data. Adapun hasil yang didapatkan adalah sensor pH meter memiliki toleransi kesalahan sekitar 6,3% dari pengukuran manual, sedangkan sensor ultrasonik memiliki toleransi kesalahan sekitar 7,8%.

Penelitian dengan judul “Smart Monitoring dan Kontrol Berbasis Arduino Pada Sistem Akuaponik” [3], membahas terkait monitoring dan kontrol dalam budidaya akuaponik yang terhubung dengan Internet of Things (IoT), di mana data mengenai kondisi sistem yang terkait dengan parameter fisik dikumpulkan melalui sensor-sensor yang dipasang dan disampaikan ke server untuk memberikan informasi kondisi secara real-time kepada pengguna. Adapun hasil yang didapatkan adalah akurasi penggunaan algoritma naive bayes sebesar 70%. Penelitian dengan judul “Rancang Bangun Akuaponik Berbasis Internet of Things” [4], membahas sistem akuaponik yang menggunakan Mikrokontroler Node MCU Esp8266 sebagai kontrol utama, Raspberry Pi 3 B sebagai tampilan visual untuk notifikasi, dan dilengkapi dengan sensor pH 4502C, sensor suhu DS18B20, Motor Servo, serta sensor SRF05. Adapun hasil yang didapatkan adalah hasil rancang bangun akuaponik dengan IoT mendekati hasil yang sebenarnya. Penelitian yang berjudul “Efektivitas Sistem Akuaponik dengan Jenis Tanaman yang Berbeda Terhadap Kualitas Air Media Budidaya Ikan” [5], membahas pengaruh dari sistem akuaponik dengan variasi jenis tanaman terhadap kualitas air dalam media budidaya ikan lele sangkuriang. Adapun hasil yang didapatkan adalah nilai suhu, oksigen terlarut, dan pH berturut-turut adalah 25,1-27,0°C, 6,3-7,5 mg/L, 7,0-7,8. Perlakuan A dengan menggunakan kangkung memiliki nilai amonia terendah, yakni sekitar 0,001 mg/L.

Berdasarkan hal tersebut, penulis bermaksud melakukan penelitian yang dapat membantu dalam melakukan klasifikasi keadaan lingkungan berdasarkan kondisi pH, suhu, kekeruhan, dan ketinggian air dalam sistem akuaponik. Namun, pada penelitian ini, penulis lebih berfokus pada penggunaan metode naive bayes untuk klasifikasi kondisi lingkungan dan data yang digunakan adalah data dummy yang diambil dengan menggunakan node-red.

2. Metode Penelitian

2.1. Metode Pengembangan Sistem

Metode yang akan digunakan untuk mengembangkan sistem adalah SDLC (Software Development Life Cycle). Metode ini terdiri dari beberapa tahapan:

a. Tahap analisis kebutuhan sistem

Pada tahap ini, sistem dianalisis dengan melakukan identifikasi masalah, menganalisis kebutuhan sistem baik yang bersifat fungsional maupun non-fungsional, mempertimbangkan efisiensi, serta memperhatikan aspek-aspek lain yang terkait dengan perancangan sistem.

b. Tahap desain

Tahap ini mencakup penyajian konsep awal, perancangan konseptual, perancangan basis data dan sistem, serta perancangan detail input/output Sistem Informasi. Pada tahap ini, fungsi sistem dipaparkan secara detail, termasuk tata letak antarmuka pengguna, aturan bisnis, diagram alur kerja, dan dokumentasi terkait.

c. Tahap konstruksi

Tahap ini melibatkan penulisan kode program menggunakan bahasa pemrograman seperti HTML dan PHP untuk aplikasi berbasis web, pembuatan dan penghubungan database, pembuatan formulir sistem, dan penyusunan petunjuk pengguna.

d. Tahap implementasi sistem

Tahap ini melibatkan pelaksanaan program dan pengujian sistem untuk memastikan hasil yang sesuai dengan kebutuhan yang telah ditetapkan.

e. Tahap uji coba

Pada tahap ini, sistem diuji dengan menggunakan data aktual untuk memperoleh hasil tes yang akurat. Uji coba dilakukan untuk mengevaluasi konektivitas dan fungsionalitas sistem, serta menilai kelayakan implementasi program.

2.2. Metode Pengumpulan data

Metode pengumpulan data pelatihan akan menggunakan data dari *Kaggle*, sementara data uji akan dihasilkan dari node-red dan disimpan ke dalam basis data sistem. Data dari *Kaggle* yang akan digunakan merupakan dataset yang disediakan oleh Bobby Siswanto pada tahun 2023 dengan judul ‘A Simple Dataset of Aquaponic Fish Pond IoT’ dan dapat diakses melalui tautan berikut: <https://www.kaggle.com/datasets/bobsis/small-aquaculture-fishpond>.

2.3. Naive Bayes

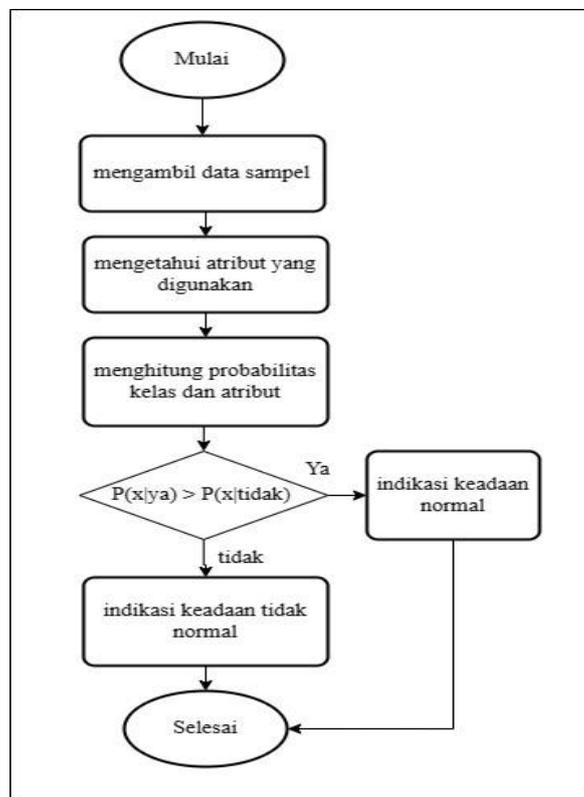
Naive Bayes merupakan salah satu algoritma klasifikasi yang memanfaatkan probabilitas dan statistik untuk meramalkan kemungkinan berdasarkan input yang tersedia. Keuntungan dari Naive Bayes adalah kemampuannya untuk menghitung probabilitas posterior bahkan dengan jumlah atribut yang besar [6]. Prediksi Bayes berasal dari teorema Bayes, yang rumusnya dapat dinyatakan secara umum seperti pada persamaan 1 dengan keterangan seperti yang terlihat pada tabel 1 berikut [7].

$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E)} \dots\dots\dots (1)$$

Tabel 1. Keterangan Formula Teorema Naive Bayes

Parameter	Keterangan
$P(H E)$	Probabilitas bersyarat akhir dari suatu hipotesis H terjadi dapat dihitung berdasarkan bukti E yang diberikan.
$P(E H)$	Kemungkinan terjadinya suatu bukti E dapat mempengaruhi probabilitas dari hipotesis H.
$P(H)$	Probabilitas awal dari hipotesis H terjadi tanpa memperhatikan bukti apa pun.
$P(E)$	Probabilitas awal dari bukti E terjadi tanpa memperhatikan hipotesis atau bukti lainnya..

Flowchart Naive Bayes.



Gambar 1. Flowchart Naive Bayes

Berdasarkan gambar 1 terkait flowchart naive bayes, adapun langkah pertama yang dilakukan dalam penggunaan algoritma naive bayes adalah pengambilan data sampel. Setelah melakukan pengambilan sampel, langkah berikutnya adalah mengetahui atribut-atribut yang digunakan pada data tersebut. Selanjutnya adalah menghitung probabilitas kelas dan atribut, di mana:

$P(\text{normal}|\text{indikasi}) > P(\text{tidak normal}|\text{indikasi}) \rightarrow \text{normal}.$

$P(\text{normal}|\text{indikasi}) < P(\text{tidak normal}|\text{indikasi}) \rightarrow \text{tidak normal}.$

2.4. Confusion Matrix

Confusion Matrix merupakan tabel yang memberikan gambaran detail mengenai kinerja suatu model atau algoritma. Setiap baris dalam tabel mewakili kelas sebenarnya dari data, dan setiap kolom mewakili kelas yang diprediksi oleh model (atau sebaliknya) [8]. Adapun tabel data yang dituangkan ke dalam tabel confusion matrix dapat dilihat melalui tabel 2.

Tabel 2. Confusion Matrix

	<i>Predicted Negative</i>	<i>Predicted Positive</i>
<i>Actual Negative</i>	<i>True Negative</i>	<i>False Positive</i>
<i>Actual Positive</i>	<i>False Negative</i>	<i>True Positive</i>

Melalui empat data seperti yang terlihat pada tabel 2, kita dapat memperoleh informasi tambahan yang sangat bermanfaat untuk mengevaluasi kinerja sebuah model, seperti berikut:

1. Akurasi (*Accuracy*): Menunjukkan seberapa sering model berhasil mengklasifikasikan dengan benar secara keseluruhan. Formula akurasi dapat diungkapkan menggunakan persamaan 2.

$$\frac{TP + TN}{Total} \dots\dots\dots (2)$$

2. Presisi (*Precision*): Menunjukkan seberapa sering prediksi positif model adalah benar ketika model memprediksi suatu instance sebagai positif. Formula presisi dapat diungkapkan menggunakan persamaan 3.

$$\frac{TP}{FP + TP} \dots\dots\dots (3)$$

3. Recall (*Sensitivitas / True Positive Rate*): Menunjukkan frekuensi di mana model memprediksi positif saat kelasnya memang positif. Persamaan untuk menghitung recall dapat dijelaskan dengan rumus 4..

$$\frac{TP}{FN + TP} \dots\dots\dots (4)$$

4. Skor F1 (*F1-Score*): hasil rata-rata harmonik dari Presisi dan Recall. Rumus untuk menghitung skor F1 dapat dijelaskan melalui persamaan 5..

$$2 * \frac{precision * recall}{precision + recall} \dots\dots\dots (5)$$

2.5. K-fold Cross Validation

K-fold cross validation adalah metode yang dipakai untuk menaksir kesalahan prediksi dalam mengevaluasi performa suatu model. Data asal akan dibagi menjadi k bagian dengan ukuran hampir serupa. Model akan dilatih dan diuji sebanyak k kali selama proses klasifikasi. Pada setiap iterasi, satu bagian data akan digunakan sebagai data uji (testing data), sementara yang lainnya sebagai data latih (training data). [9].

Adapun langkah-langkahnya adalah sebagai berikut:

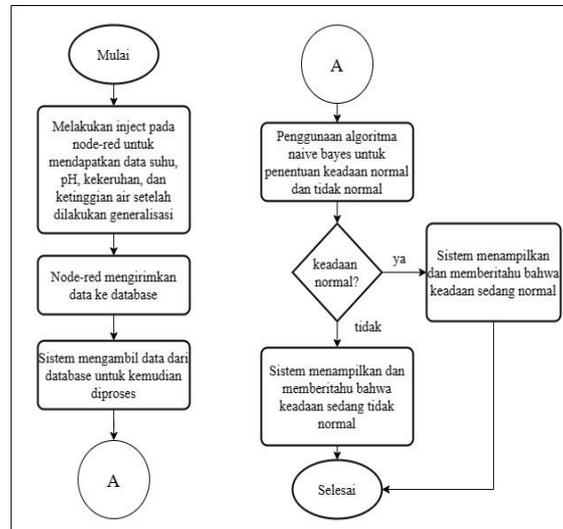
1. Data awal akan terbagi menjadi k bagian yang sama.
2. Pada setiap iterasi, salah satu bagian menjadi data uji dan yang lainnya menjadi data latih. Akurasi dihitung dengan membandingkan hasil prediksi dengan data sebenarnya.
3. Proses tersebut diulang sampai mencapai iterasi ke-k. Setelah itu, rata-rata dari semua akurasi dihitung. Rata-rata tersebut menjadi akurasi final dari k-fold cross validation.

2.6. Blackbox Testing

Pengujian blackbox adalah proses pengujian yang fokus pada fungsi-fungsi perangkat lunak tanpa memperhatikan struktur internalnya. Tujuannya adalah untuk menemukan kesalahan fungsi, antarmuka, struktur data, performa, inisialisasi, dan terminasi. Salah satu teknik yang digunakan adalah partisi ekuivalen, di mana setiap jenis input diuji dan dikelompokkan berdasarkan fungsinya serta validitas hasilnya. Pengujian black box memastikan setiap proses dalam sistem berfungsi sesuai kebutuhan yang diinginkan. Pengujian dilakukan dengan menginterpretasikan rangkaian kondisi masukan dan menguji fungsi khusus dari sistem. Ini dilakukan sebagai bagian dari implementasi program untuk mengidentifikasi dan memperbaiki kesalahan agar sistem dapat dianggap layak digunakan [10].

2.7. Desain Sistem

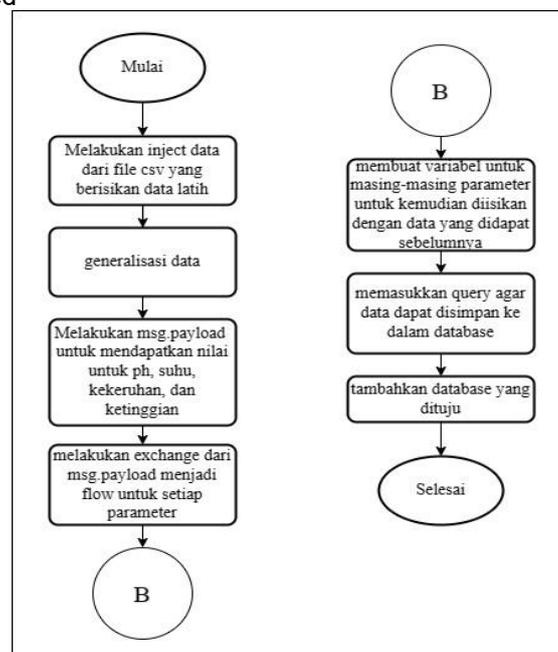
2.7.1. Flowchart Sistem



Gambar 2. Flowchart Sistem

Gambar 2 menunjukkan flowchart inti sistem. Sistem dimulai dengan pengumpulan data dari Node-RED, di mana data suhu, pH, kekeruhan, dan ketinggian air diinjeksikan. Setelah data diperoleh, itu disimpan dalam database. Kemudian, sistem mengambil data dari database untuk diproses. Berdasarkan data tersebut, kondisi lingkungan diklasifikasikan menggunakan metode naive Bayes berdasarkan keempat parameter. Jika kondisi normal, sistem akan memberikan laporan bahwa keadaan normal. Namun, jika tidak normal, sistem akan memberitahu bahwa keadaan tidak normal.

2.7.2. Flowchart node-red



Gambar 3. Flowchart Node-red

Gambar 3 menunjukkan flowchart Node-RED. Langkah pertama adalah melakukan injeksi data timestamp dengan interval waktu 1 menit. Setelah itu, *msg.payload* digunakan untuk mendapatkan

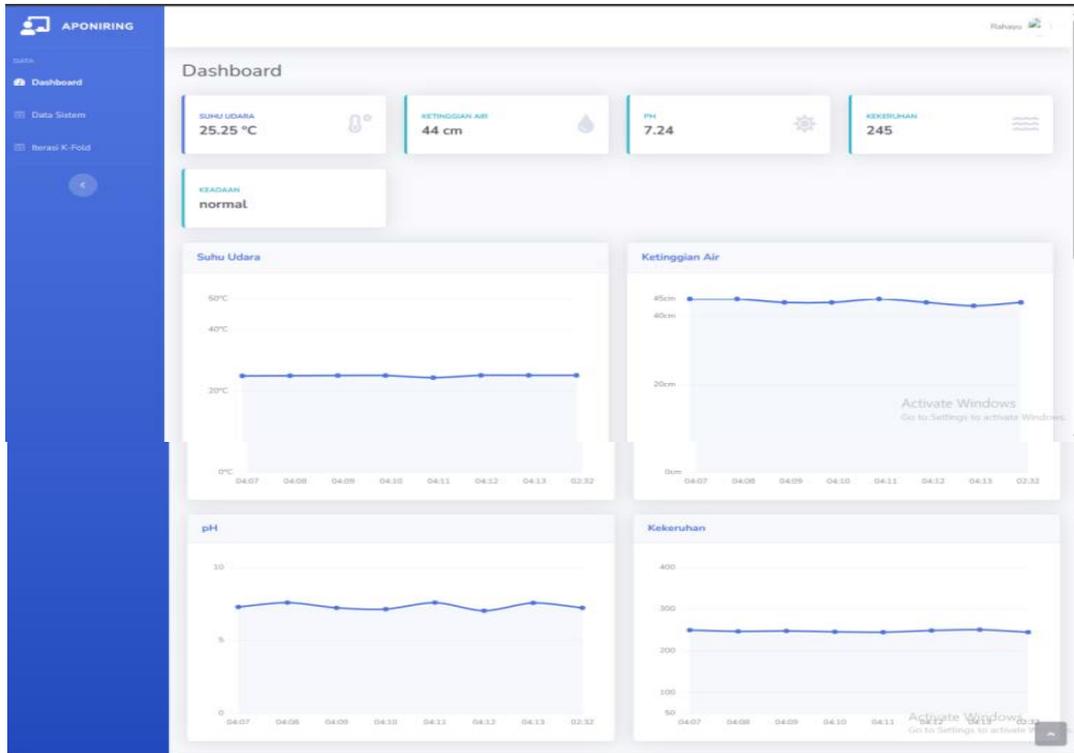
nilai pH, suhu, kekeruhan, dan ketinggian. Data tersebut kemudian diubah dari *msg.payload* menjadi flow untuk setiap parameter agar dapat disimpan dalam variabel. Selanjutnya, variabel untuk masing-masing parameter diisi dengan data yang diperoleh dari payload. Langkah berikutnya adalah membuat query untuk menyimpan data ke dalam database, dengan menambahkan database yang dituju agar data dapat disimpan di sana.

3. Hasil dan Pembahasan

3.1. Hasil Rancangan Website

Setelah melalui proses perancangan terhadap sistem, adapun hasil perancangan website dapat dilihat seperti pada gambar 4 sampai gambar 6 berikut.

a. Halaman Dashboard



Gambar 4. Halaman Dashboard

Berdasarkan gambar 4 di atas, pada halaman dashboard, user dapat melihat informasi suhu, ketinggian air, pH, kekeruhan, dan keadaan kondisi lingkungan baik secara grafik maupun informasi kondisi terkini.

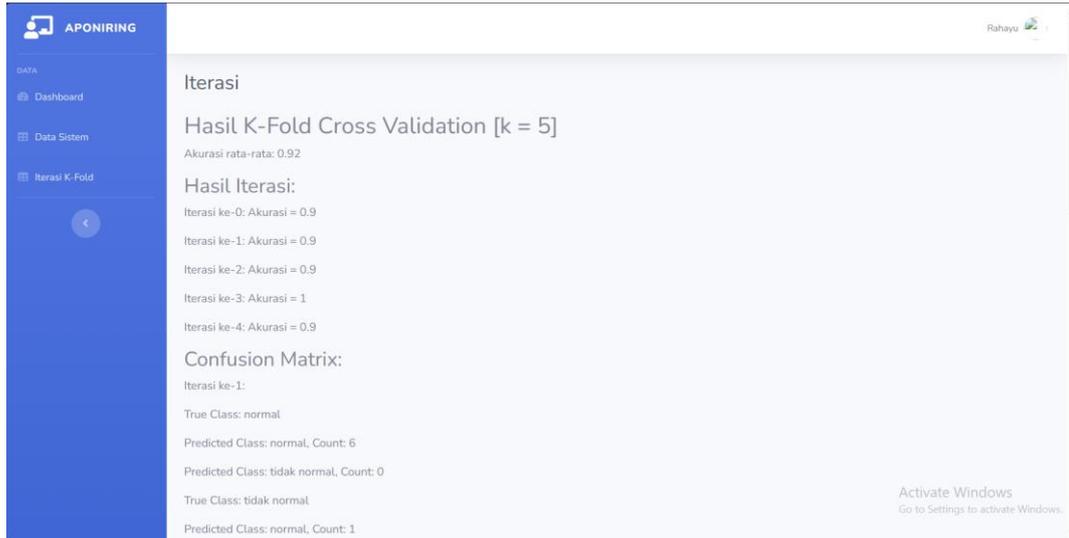
b. Halaman Data Sistem

Waktu	Suhu °C	Ketinggian Air (cm)	pH	Kekeruhan	Status
2024-05-08 00:50:39	25.25	45	7.24	251	
2024-05-08 00:49:05	25.31	45	7.15	251	
2024-05-08 00:48:45	25.31	44	7.6	248	
2024-05-08 00:47:37	25.25	45	7.3	246	
2024-05-08 00:46:37	25.25	45	7.11	249	
2024-05-08 00:45:04	25.25	44	7.09	251	

Gambar 5. Halaman Data Sistem

Berdasarkan gambar 5 di atas, pada halaman data sistem, user dapat melihat informasi suhu, ketinggian air, pH, kekeruhan dalam bentuk tabel. Selain itu, user dapat melakukan pencarian serta penghapusan data melalui halaman data sistem ini.

c. Halaman Iterasi K-Fold

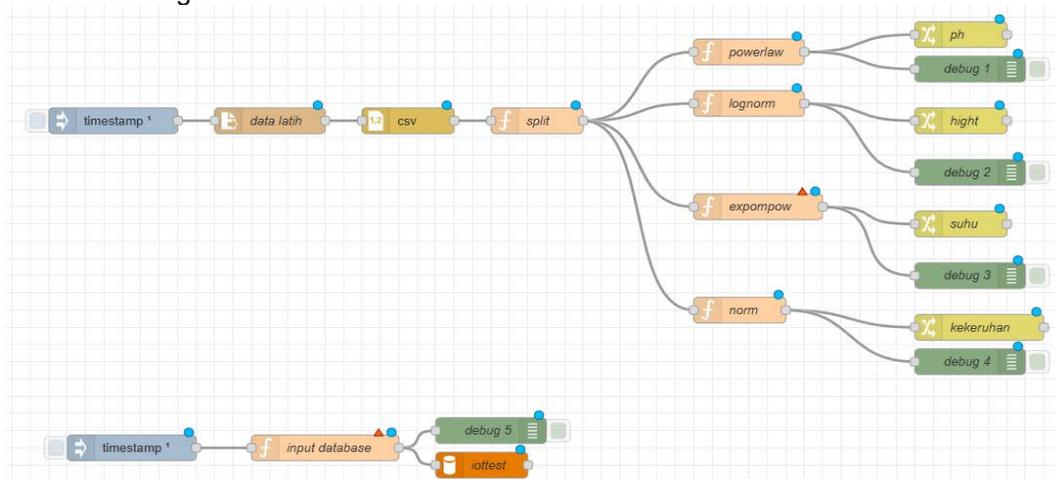


Gambar 6. Halaman Iterasi K-fold

Berdasarkan gambar 6 di atas, pada halaman iterasi k-fold, user dapat melihat informasi akurasi hasil pengujian naive bayes dengan *k-fold cross validation* dan *confusion matrix*.

3.2. Pengimplementasian Node-red

Pembuatan flow sistem pada node-red untuk mendapatkan data baru dari hasil generate data latihan dapat dilihat melalui gambar 7 berikut.



Gambar 7. Flow Node pada Node-red

Berdasarkan gambar 7 di atas, flow diawali dengan melakukan inject timestamp untuk melakukan running flow sekali tiap satu menit. Kemudian, menambahkan node file in yang digunakan untuk membaca file data latihan yang digunakan. Setelah file berhasil dibaca, akan dilakukan pengambilan string pada data dengan menggunakan node csv. Tahap selanjutnya adalah melakukan pemisahan data berdasarkan kolom yang dipisahkan oleh titik koma (;) dengan menggunakan node function. Setelah data berhasil dipisahkan, akan dilakukan generate data berdasarkan distribusi data pada masing-masing parameter agar mendapatkan data baru yang dilakukan dengan menggunakan node function. Setelah mendapatkan data baru, dilakukan penggantian properti value dengan menggunakan node change agar dapat digunakan pada node selanjutnya. Setelah melakukan penggantian properti value, data disimpan ke dalam database dengan menggunakan node function dan node mysql sebagai node yang mengonfigurasi database tujuan.

3.3. Hasil Pengimplementasian Naive Bayes

Adapun data dalam pengimplementasian naive bayes ini adalah data yang didapatkan dari hasil generate data di node-red sebanyak 50 data yang disimpan pada database. Beberapa data yang sudah didapatkan dapat dilihat melalui gambar 8 berikut.

time	pH	kekeruhan	ketinggian	suhu
2024-05-07 23:51:31	7.24	246	44	25.19
2024-05-07 23:52:36	7.6	246	44	25.13
2024-05-07 23:53:08	6.98	248	44	25.06
2024-05-07 23:54:20	7.32	250	45	25.06
2024-05-07 23:55:12	7.6	246	45	24.31

Gambar 8. Data yang didapatkan dari Node-red

Gambar 8 di atas menunjukkan beberapa data hasil generate data dari node-red yang akan diproses dengan menggunakan metode naive bayes untuk mengetahui indikasi kondisi lingkungan saat itu. Setelah melalui pemrosesan dengan menggunakan metode naive bayes, adapun hasil yang didapatkan seperti pada tabel 3 berikut.

Tabel 3. Hasil Implementasi Naive Bayes

No	pH	Kekeruhan	Ketinggian	Suhu	Indikasi
1	7.24	246	44	25.19	Normal
2	7.6	246	44	25.13	Normal
3	6.98	248	44	25.06	Normal
4	7.32	250	45	25.06	Normal
5	7.6	246	45	24.31	Tidak Normal

Berdasarkan tabel 3 di atas, didapatkan bahwa hasil implementasi metode naive bayes pada masing-masing data yang sudah didapatkan sebelumnya pada gambar 8 secara berturut-turut adalah normal, normal, normal, normal, dan tidak normal.

3.4. Pengujian dengan Confusion Matrix dan K-fold Cross Validation

Proses pengujian menggunakan k=5, yang berarti akan ada 5 iterasi pengujian dan hasil akhir akan diambil rata-rata dari 5 pengujian tersebut. Dataset terdiri dari 50 entri, dimana 10 digunakan untuk pengujian dan 40 untuk pelatihan. Pembagian antara data latih dan data uji menggunakan 5-fold cross validation, artinya dataset akan dibagi menjadi 5 bagian dari total 50 data. Setiap iterasi, satu bagian (data uji) akan digunakan untuk pengujian, sementara bagian lainnya digunakan untuk data latih. Adapun visualisasi pembagian data testing dan data training dapat dilihat melalui tabel 4 dan tabel 5 berikut.

Tabel 4. Pembagian Data Testing

ID Data Testing	Jumlah Data
K-1	10
K-2	10
K-3	10
K-4	10
K-5	10

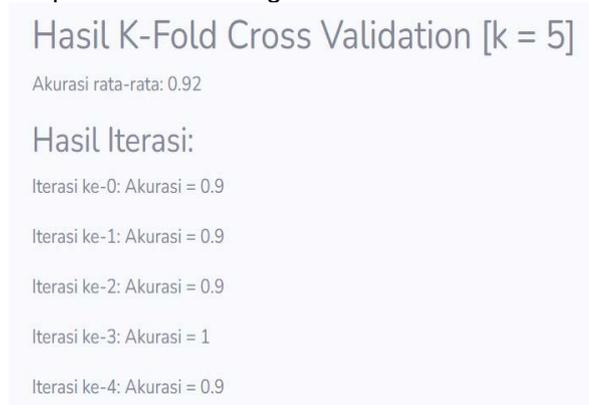
Berdasarkan tabel 4, pada tiap iterasi yang dilakukan, akan digunakan 10 data dari 50 data yang ada sebagai data testing.

Tabel 5. Pembagian Data Training

ID Data Training	Jumlah Data
K-1	40
K-2	40
K-3	40
K-4	40
K-5	40

Berdasarkan tabel 5, pada tiap iterasi yang dilakukan, akan digunakan 40 data dari 50 data yang ada sebagai data training.

Adapun hasil pengujian ini dapat dilihat melalui gambar 9 berikut.



Gambar 9. Hasil Pengujian Naive Bayes

Berdasarkan gambar 9 di atas, didapatkan akurasi sebesar 92%. Di mana pada iterasi ke-0 didapatkan akurasi sebesar 0,9. Iterasi ke-1 didapatkan akurasi sebesar 0,9. Iterasi ke-2 didapatkan akurasi sebesar 0,9. Iterasi ke-3 didapatkan akurasi sebesar 1. Iterasi ke-4 didapatkan akurasi sebesar 0,9.

3.5. Pengujian Blackbox

Dalam melakukan pengujian performa terhadap sistem yang telah dirancang dapat dilakukan dengan melakukan blackbox testing. Adapun hasil pengujian seperti pada tabel 6 berikut.

Tabel 6. Hasil Pengujian *Blackbox*

No	Pengujian	Hasil yang diharapkan	Keterangan
1	Penginputan data dari node-red ke <i>database</i>	Data berhasil diinputkan dari node-red ke <i>database</i>	Sesuai
2	Akses halaman <i>dashboard</i>	Berhasil mengakses halaman <i>dashboard</i>	Sesuai
3	Akses halaman data sistem	Berhasil mengakses halaman data sistem	Sesuai
4	Akses halaman iterasi k-fold	Berhasil mengakses halaman iterasi k-fold	Sesuai
5	Halaman <i>dashboard</i> menampilkan data sesuai <i>database</i>	Berhasil menampilkan data sesuai <i>database</i> pada halaman <i>dashboard</i>	Sesuai
6	Halaman <i>dashboard</i> menampilkan gambar profil pengguna	Berhasil menampilkan gambar profil pengguna	Tidak sesuai
7	Melakukan pencarian pada halaman data sistem	Data berhasil ditampilkan sesuai kata kunci yang diinputkan	Sesuai
8	Melakukan penghapusan data pada halaman data sistem	Data berhasil dihapus baik pada <i>website</i> maupun <i>database</i>	Sesuai

Berdasarkan hasil pengujian pada tabel 6 di atas, didapatkan bahwa 7 dari 8 fitur dapat berjalan dengan baik namun masih terdapat hal yang perlu disempurnakan.

4. Kesimpulan

Kesimpulan dari penelitian ini adalah sebagai berikut:

1. Penelitian ini terdiri dari tiga tahap: perancangan, pengujian, dan pencarian akurasi. Tahap perancangan menggunakan berbagai alat seperti Node-RED, Visual Studio Code, XAMPP, dan database PHPMyAdmin. Bahasa pemrograman yang digunakan adalah PHP. Pengujian dan pencarian akurasi dilakukan melalui dua pengujian: *k-fold cross validation* dan *confusion matrix* untuk pengujian naive Bayes, serta *blackbox testing* untuk pengujian sistem website.

2. Berdasarkan hasil pengujian naive Bayes, ditemukan bahwa akurasi rata-rata mencapai 92%, menunjukkan bahwa algoritma tersebut dapat berjalan dengan baik.
3. Hasil pengujian blackbox menunjukkan bahwa 7 dari 8 fitur berjalan lancar. Namun, beberapa fitur masih memerlukan perbaikan dan pengembangan di masa mendatang.

Referensi

- [1] Rivaldo, K.L., Mogi, I.K.A., Suputra, P.G.H., Sanjaya, N.A., Darmawan, B.A., dan Dwidasmara, B.G., "Sistem Monitoring Tanaman Hidroponik Berbasis Internet of Things", *Jurnal Elektronik Ilmu Komputer Udayana*, Vol. 11, No. 1, p. 101-110, 2022.
- [2] Kuswinta, A.J., Wedashwara, W., dan Arimbawa, A., "Implementasi IoT Cerdas Berbasis Inference Fuzzy Tsukamoto Pada Pemantauan Kadar pH dan Ketinggian Air Dalam Akuaponik", *J-COSINE*, Vol. 3, No. 1, p. 65-74, 2019.
- [3] Jagat, L., Sundari, E.M., dan Apriani W., "Smart Monitoring dan Kontrol Berbasis Arduino Pada Sistem Akuaponik", *JASISFO*, Vol. 3, No. 1, p. 260-267, 2022.
- [4] Widiantera, K., Linawati, dan Wiharta, D.M., "Rancang Bangun Akuaponik Berbasis Internet of Things" *Jurnal SPEKTRUM*, Vol. 8, No. 1, p. 243-253, 2021.
- [5] Zidni, I., Rizal, A., Andriani, Y., dan Ramadan, R., "Efektivitas Sistem Akuaponik dengan Jenis Tanaman yang Berbeda Terhadap Kualitas Air Media Budidaya Ikan", *Jurnal Perikanan dan Kelautan*, Vol. 9, No. 1, p. 81-94, 2019.
- [6] Firmansyah, A., Syangy, D., dan Prasetyo, B.H., "Implementasi Algoritma Naive Bayes pada Sistem Monitoring dan Klasifikasi Kualitas Air Akuarium Ikan Mas Koki", *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*. Vol. 6, No. 9, p. 4128-4135, 2022.
- [7] Azizah, N., Goejantoro, R., dan Sifriyani, "Metode Naive Bayes dengan Pendekatan Dstribusi Gauss untuk Klasifikasi Peminatan Peserta Didik" pada *Prosiding Seminar Nasional Matematika, Statistika, dan Aplikasinya*, Samarinda, 2019, vol. I, pp. 8-14.
- [8] Kurniawan, I.B., Candiasa, I.M., dan Aryanto, K.Y.E., "Sistem Pendukung Keputusan Pemilihan Mahasiswa Berprestasi di Universitas Dhyana Pura Menggunakan Metode AHP, Electre, dan Topsis", *Jurnal Ilmu Komputer Indonesia*, Vol. 4, No.1, p. 22-33, 2019.
- [9] Mardiana, L., Kusnandar, D. dan Satyahadewi, N., "Analisis Diskriminan dengan K-Fold Cross Validation untuk Klasifikasi Kualitas Air di Kota Pontianak", *Buletin Ilmiah Mat. Stat. dan Terapannya*, Vol. 11, No. 1, p. 97-102, 2022.
- [10] Wijaya, Y.D., dan Astuti, M.W., "Pengujian Blackbox Sistem Informasi Penilaian Kinerja Karyawan PT Inka (Persero) Berbasis Equivalence Partitions", *Jurnal Digital Teknologi Informasi*, Vol. 4, No. 1, p. 22-26, 2021.

This page is intentionally left blank.

Implementasi Metode Design Thinking dan Prinsip Gestalt Pada Rancang Bangun Dashboard Smart-Farm

Alvin Wiraprathama^{a1}, I Made Widiartha^{a2}, I Ketut Gede Suhartana^{a3},
I Gusti Agung Gede Arya Kadyanan^{a4}

^aProgram Studi Informatika Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana

¹ wiraprathamaalvin@gmail.com

² madewidiartha@unud.ac.id

³ ikg.suhartana@unud.ac.id

⁴ gungde@unud.ac.id

Abstract

Bali's agricultural sector experienced a notable downturn in growth during 2022 compared to the preceding year, which had seen substantial expansion. One of the key factors behind this decline was the limited adoption of digital farming practices in Bali. This underutilization of digital technologies in agriculture, including tools for crop monitoring, soil analysis, weather prediction, and market connectivity, likely contributed to decreased productivity and efficiency in farming activities. Addressing this issue could be pivotal for Bali to rejuvenate its agriculture sector and foster growth moving forward. To tackle this challenge, a smart farming dashboard has been developed specifically to address the needs of farmers in Bali. This dashboard boasts a user-friendly interface design and incorporates features tailored to assist farmers in Bali. The dashboard was crafted through the application of design thinking methodology, integrating Gestalt principles to optimize the user interface. The results of the dashboard's usability testing, as measured by the System Usability Scale (SUS), indicate a strong performance with a score of 79.75, reflecting its excellent usability.

Keywords: Agriculture, Smart Farming, Design Thinking, Gestalt Principle, Dashboard

1. Pendahuluan

Bali merupakan salah satu provinsi di Indonesia yang terkenal turut aktif dalam perkembangan sektor pertanian. Namun berdasarkan Laporan Perekonomian Provinsi Bali yang diterbitkan oleh Bank Indonesia pada Februari 2023, perkembangan sektor pertanian di Bali pada tahun 2022 mengalami penurunan yang cukup signifikan dibandingkan dengan peningkatan yang dialami pada tahun 2021. Salah satu variabel yang menyebabkan menurunnya perkembangan sektor pertanian di Bali adalah rendahnya penerapan digitalisasi pertanian. Berdasarkan survei Bank Indonesia, hanya sekitar 26,7% petani di Bali yang sudah memiliki teknologi pertanian dan menggunakannya dalam bertani. Hal ini disebabkan oleh terbatasnya fasilitas teknologi pertanian yang ada. [1]

Digitalisasi pertanian merupakan perubahan cara yang dilakukan khususnya pada aspek pertanian yang dimulai dari pengolahan lahan pertanian hingga pemasaran komoditas pertanian itu sendiri menggunakan teknologi terkini atau lebih dikenal dengan nama *Smart Farm*. Teknologi *Smart Farm* sendiri terdiri dari berbagai jenis mulai dari sensor IoT, analitik data dari produk pertanian, hingga automasi sistem pertanian yang dapat meningkatkan produktivitas pertanian. Namun untuk menggunakan teknologi ini, para petani membutuhkan sebuah interface yang mudah untuk digunakan sehingga para petani dapat mengontrol teknologi tersebut untuk mendapatkan data real-time dan hasil analisis tentang data pertanian mereka. [2]

Design Thinking merupakan sebuah framework yang dapat digunakan pada pembuatan perangkat lunak berfokus pada pembuatan desain produk yang sesuai dengan kebutuhan dan keinginan dari pengguna karena memiliki empat elemen yaitu : *Human-Centered* yang berarti berdasarkan pengguna. *Highly Creative* yang berarti dapat menggunakan ide dan kreativitas yang sebebas – beasnya. *Hands On* yang berarti melakukan percobaan secara langsung tanpa tidak hanya

menggunakan studi dan teori. Dan yang terakhir iteratif yang berarti proses yang dapat dilakukan berulang kali untuk improvisasi pada setiap Langkah. *Design Thinking* dapat menjadi metode yang efektif dalam mengembangkan interface smart farm yang sesuai dengan kebutuhan dari para petani yang ada di Bali namun mudah untuk digunakan. [3]

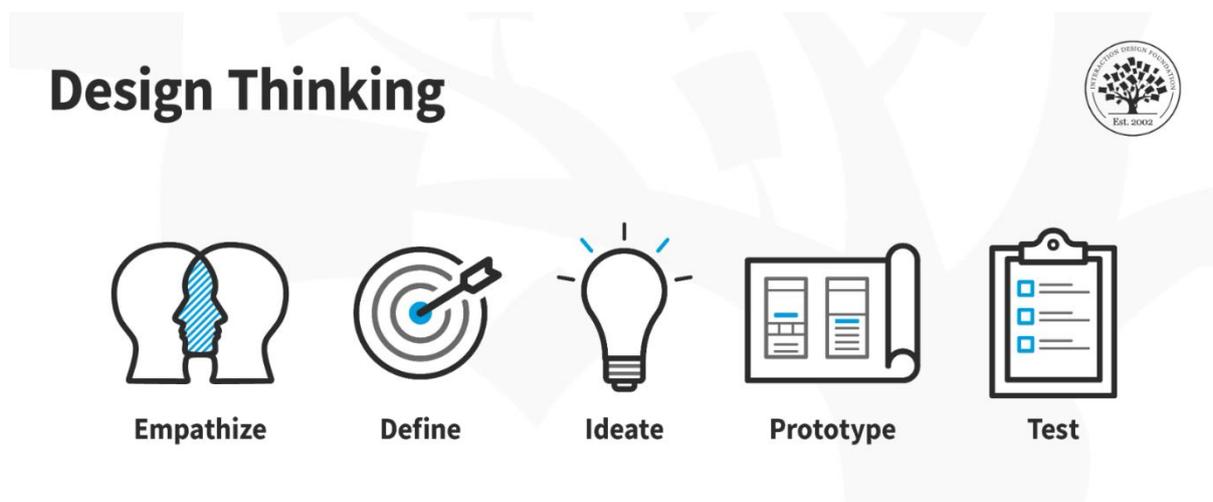
Prinsip Gestalt merupakan sebuah kumpulan prinsip bagaimana manusia melihat dan mengolah informasi secara visual. Prinsip Gestalt ini dapat diaplikasikan kedalam pembuatan desain interface aplikasi sehingga dapat menciptakan sebuah desain *interface* aplikasi yang intuitif, mudah digunakan dan nyaman untuk dilihat, dan *user-friendly*. Prinsip Gestalt berperan penting dalam membangun sebuah *UI Design*. [4]

Dengan implementasi dari metode Design Thinking dan prinsip Gestalt, diharapkan dapat tercipta sebuah *interface dashboard* dari sistem smart farm yang mudah untuk digunakan dan efektif dalam meningkatkan digitalisasi pertanian yang dapat membantu para petani di Bali untuk meningkatkan produktivitas di sektor pertanian. Oleh karena itu, berdasarkan permasalahan yang dikemukakan, penulis melakukan sebuah penelitian dengan judul "Implementasi Metode Design Thinking dan Prinsip Gestalt Pada Rancang Bangun Dashboard Smart-Farm".

2. Metode Penelitian

Dalam penelitian ini, penulis menggunakan metode Design Thinking sebagai pendekatan metode penelitian dan Prinsip Gestalt sebagai salah satu prinsip untuk merancang desain antarmuka.

2.1. Metode Design Thinking



Gambar 1. Metode Design Thinking [5]

Gambar 2 diatas memperlihatkan tiap langkah yang ada pada metode *design thinking*. Metode *design thinking* merupakan sebuah pendekatan untuk mendapatkan solusi dari sebuah masalah yang ada. Metode design thinking merupakan metode baru dalam proses perancangan sebuah desain aplikasi. *Design Thinking* bersifat *human-centered* atau *user-centered* yang dimana metode perancangan ini akan menyelesaikan masalah yang berfokus pada pengguna atau *user* dengan cara melakukan pembiasaan ulang suatu masalah dengan sudut pandang dan cara – cara yang berpusat kepada pengguna atau manusia sehingga menciptakan lebih banyak ide sebagai solusi untuk permasalahan yang dihadapi dan menjadikan proses perancangan desain sesuai dengan kebutuhan pengguna sehingga membantu pengguna dalam mencapai tujuan. Metode Design Thinking memiliki 5 tahapan dalam prosesnya yaitu (1) *Empathize*, (2) *Define*, (3) *Ideate*, (4) *Prototype* dan (5) *Test*. Setiap tahapan pada metode ini dibuat berdasarkan kebutuhan dari pengguna dan juga dilakukan berulang sesuai dengan kebutuhan [6]. Berikut adalah tahapan – tahapan dalam metode *Design Thinking* :

a. *Empathize*

Tahapan ini bertujuan untuk memahami lebih dalam permasalahan apa yang dihadapi oleh pengguna, apa yang diinginkan, dan apa yang dibutuhkan oleh pengguna untuk pengembangan sistem kedepannya. Pada tahapan ini, penulis melakukan wawancara dan observasi pada objek dan tempat penelitian yang telah ditentukan. Objek dan tempat penelitian yang akan digunakan pada penelitian ini adalah para petani atau para pekerja yang pernah ataupun sedang aktif

bekerja di sektor pertanian yang berdomisili di provinsi Bali yang dijadikan narasumber untuk mendapatkan analisis kebutuhan dari sistem.

b. *Define*

Pada tahapan *Define*, terdapat proses pendefinisian masalah ke dalam bentuk problem statement yang konkrit dan jelas berdasarkan *insight* permasalahan yang didapatkan sebelumnya. Pendefinisian masalah ke dalam bentuk problem statement yang konkrit dan jelas berdasarkan *insight* permasalahan yang didapatkan sebelumnya. Pada tahapan ini, penulis membuat *user persona*, *pain points*, dan juga *how-might-we*.

c. *Ideate*

Ideate merupakan tahapan ketiga dalam metode *Design Thinking*. Pada tahapan *Ideate*, akan terjadi *brainstorming* ide dan solusi terhadap *problem statement How-Might-We* yang telah didapatkan dari tahapan *Define*. Langkah – Langkah yang terdapat pada tahapan ini termasuk pembuatan *solution idea*, *prioritization matrix*, hingga pembuatan *task flow*.

d. *Prototype*

Prototype merupakan tahapan keempat dalam metode *Design Thinking*. Pada tahapan *Prototype*, akan mulai dilakukan perancangan prototipe pada sistem yang akan dibuat menggunakan acuan dari hasil yang telah didapatkan dari tahapan – tahapan sebelumnya. Pada tahapan ini akan dilakukan pembuatan *Design System*, *Wireframe*, dan *High Fidelity Prototype*. Prototipe merupakan rancangan awal yang dibuat untuk mewakili sistem yang sebenarnya sebelum dilakukan iterasi pengembangan sistem yang sebenarnya sehingga rancangan aplikasi pada tahapan ini bertujuan sebagai representasi dari model sistem yang sebenarnya untuk menemukan kesalahan yang ada tanpa harus menghabiskan waktu yang lama dan mengerjakan hal yang terlalu kompleks pada saat perancangan aplikasi

e. *Test*

Test atau *Testing* merupakan tahapan paling akhir yang ada pada metode *Design Thinking*. Pada tahapan *testing* dilakukan pengujian terhadap prototipe yang telah dibuat. Tujuan dari tahapan *testing* adalah memperoleh respon dan feedback dari pengguna. Pada tahapan ini, metode *testing* yang digunakan adalah *black box testing*, *system usability scale (SUS)*. *Black Box Testing* merupakan sebuah metode pengujian fungsionalitas pada sistem yang dibuat. Pada *black box testing*, pengujian ditekankan pada pola masukan dan keluaran yang sesuai dengan skenario sistem yang diinginkan. *System Usability Scale* merupakan sebuah standar kuisioner yang digunakan untuk mengukur tingkat kegunaan atau *usability* dari sebuah sistem perangkat lunak

2.2. Prinsip Gestalt

Prinsip Gestalt merupakan sebuah prinsip yang menjelaskan bagaimana proses manusia mempersepsikan informasi visual lalu mengorganisasikan komponen – komponen dari informasi tersebut yang memiliki hubungan, pola atau kemiripan menjadi kesatuan. Prinsip Gestalt banyak diaplikasikan dalam proses mendesain dan beberapa jenis seni rupa lainnya karena dapat memberikan penjelasan bagaimana sebuah proses kognitif individu dalam menafsirkan informasi yang diperoleh oleh indera visual dapat terbentuk. [7]

3. Hasil dan Pembahasan

Pada penelitian ini, hasil dari proses tiap tahapan *design thinking* diimplementasikan menjadi sebuah sistem lalu diuji dan dievaluasi menggunakan instrument pengujian antarmuka sistem.

3.1. Empathize

Pada tahapan pertama ini, telah dilakukan wawancara dan observasi terhadap narasumber langsung yaitu para petani yang aktif bekerja pada sektor pertanian. Narasumber berpendapat bahwa akses teknologi pertanian yang ada saat ini masih terbatas serta sulit digunakan. Narasumber juga berpendapat bahwa kemudahan dalam penggunaan teknologi pertanian serta tampilan yang ramah pengguna dengan kondisi mereka saat ini sangat penting untuk membantu proses bertani. Hasil dari wawancara dan observasi terhadap pengguna tersebut kemudian dituangkan kedalam bentuk *empathy map*. Dari *empathy map* tersebut, akan dipilih beberapa masalah utama dari pengguna yang akan difokuskan untuk diselesaikan. Pada tabel 1 berikut merupakan hasil *empathy map* yang didapat dari observasi dan wawancara bersama para petani :

Tabel 1. Empathy Map

No	Aspek Empathy Map	Keterangan
1	Says	Saya hanya pernah menggunakan teknologi pertanian seperti tractor dan pemotong rumput
		Saya tidak pernah melakukan pencatatan atau pendataan terkait hasil produksi pertanian
		Saya tidak pernah menggunakan data dan informasi cuaca dalam melakukan penanaman komoditas pertanian saya
		Saya masih kurang paham terhadap penggunaan teknologi pertanian yang ada
2	Thinks	Saya ingin melakukan pendataan dan pencatatan terkait hasil produksi pertanian
		Saya ingin mendapatkan pelatihan atau panduan tentang penggunaan teknologi pertanian yang ada
		Saya ingin mendapatkan informasi terkait yang dapat membantu proses bertani saya
3	Feels	Saya khawatir tidak paham dalam menggunakan teknologi pertanian yang ada
		Saya khawatir melakukan pengambilan keputusan yang salah
		Saya khawatir hasil panen memburuk karna keadaan lingkungan yang tidak bisa ditebak
		Bekerja penuh waktu sebagai petani
4	Does	Menggunakan teknologi pertanian sederhana seperti tractor dan alat pemotong rumput
		Mengambil keputusan berdasarkan intuisi dan keadaan yang dilihat sekarang

3.2. Define

Pada tahapan *define*, hasil wawancara, observasi dan empathy map dari tahapan sebelumnya lebih didefinisikan kembali kedalam bentuk yang lebih rinci sehingga inti masalah dapat lebih fokus untuk diselesaikan.

a. User Persona

Pada bagian ini, penulis membuat sebuah *user persona* yang dijadikan acuan nanti berdasarkan hasil dari wawancara dan observasi. Pada gambar 2 berikut merupakan hasil dari *user persona* yang telah dibuat :



Gambar 2. User Persona

b. Pain Points

Berdasarkan hasil – hasil dari tahapan sebelumnya, didapatkan *pain points* yang berisikan kesulitan yang dialami oleh pengguna saat ini. Berikut adalah hasil *pain points* yang didapatkan dalam penelitian ini :

1. Para petani di Bali masih memiliki keterbatasan dalam pengelolaan informasi serta data pertanian yang mereka punya
2. Masih rendahnya tingkat ketrampilan para petani dalam menggunakan teknologi yang dapat membantu mereka dalam bertani
3. Keterbatasan akses informasi dan teknologi yang mereka butuhkan dalam bertani
4. Ketergantungan pada tengkulak dalam memproses hasil produksi pertanian
5. Kurangnya pelatihan serta dukungan dari instansi terkait praktik penggunaan teknologi pertanian

c. How Might We

Metode ini mengubah sebuah pernyataan menjadi pertanyaan sehingga inti masalah yang sebelumnya telah didapatkan pada proses sebelumnya diubah menjadi sebuah bentuk pertanyaan (How). Lalu setiap pertanyaan yang dibentuk dari permasalahan tersebut dijawab dengan kemungkinan solusi atau jawaban yang mengacu pada pertanyaan tersebut. Pada tabel 2 berikut merupakan hasil dari proses *How Might We* :

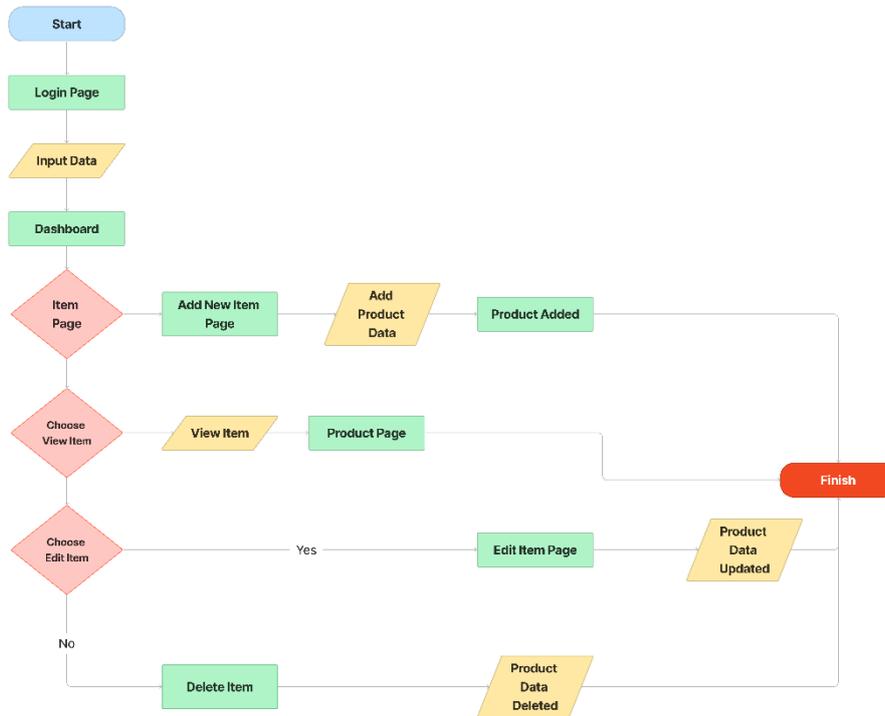
Tabel 2. How Might We

No	<i>Pain Points</i>	<i>How?</i>	<i>Might We</i>
1	Para petani di Bali masih memiliki keterbatasan dalam pengelolaan informasi serta data pertanian yang mereka punya	Bagaimana cara petani dapat mengelola informasi serta data pertanian yang mereka punya?	Merancang sebuah aplikasi yang mudah digunakan oleh petani untuk mendata informasi sehari - hari terkait informasi pertanian mereka
2	Masih rendahnya tingkat ketrampilan para petani dalam menggunakan teknologi yang dapat membantu mereka dalam bertani	Bagaimana cara membantu petani dalam meningkatkan ketrampilan mereka dalam menggunakan teknologi pertanian?	Memberikan pelatihan langsung dengan materi yang edukatif dan mudah dipahami serta akses teknologi pertanian yang mudah
3	Keterbatasan akses informasi dan teknologi yang mereka butuhkan dalam bertani	Bagaimana cara membantu petani dalam mendapatkan akses informasi dan teknologi yang mereka butuhkan dalam bertani?	Membangun sebuah pusat informasi petani yang mudah diakses dan digunakan
4	Ketergantungan pada tengkulak dalam memproses hasil produksi pertanian	Bagaimana cara mengatasi ketergantungan petani terhadap tengkulak?	Meningkatkan akses petani terhadap fasilitas pengolahan dan penanganan hasil produksi pertanian dengan membangun atau meningkatkan infrastruktur
5	Kurangnya pelatihan serta dukungan dari instansi terkait praktik penggunaan teknologi pertanian	Bagaimana cara petani mendapatkan dukungan serta pelatihan dalam praktik pertanian?	Mengembangkan aplikasi yang menyediakan akses ke kursus online, tutorial praktis, dan forum diskusi untuk petani tentang praktik penggunaan teknologi pertanian.

3.3. Ideate

Pada tahapan ideate, dilakukan pemilihan terhadap solusi – solusi terbaik yang paling sesuai dengan penelitian. Solusi yang telah dipilih tersebut kemudian dikembangkan menjadi fitur – fitur dalam rancang bangun dashboard smart-farming nantinya sehingga dapat membantu dan mempermudah pengguna dalam menyelesaikan masalah mereka. Selanjutnya ide – ide tersebut dimasukkan kedalam sebuah matriks yang akan mengurutkan prioritas fitur yang memiliki urgensi dan dampak yang lebih krusial. Sehingga berdasarkan urutan prioritas tersebut maka fitur – fitur yang dibangun pada aplikasi adalah sebagai berikut :

- a. Fitur *Login*
- b. Fitur *Register*

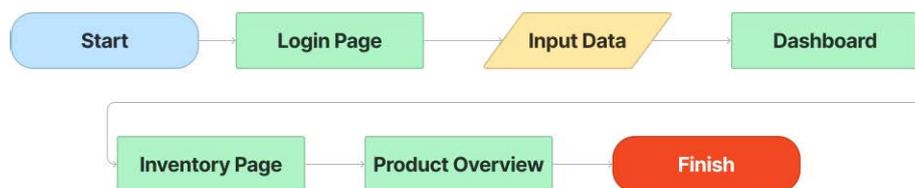


- c. Fitur *Dashboard*
- d. Fitur *Add Item*
- e. Fitur *View Item*
- f. Fitur *Edit Item*
- g. Fitur *Delete Item*
- h. Fitur *Settings*
- i. Fitur *Inventory*
- j. Fitur *Weather*

Setiap fitur dibuatkan task flownya masing – masing yang berisikan representasi sistematis dari serangkaian langkah atau aktivitas yang dilakukan oleh pengguna untuk mencapai tujuan tertentu dalam suatu aplikasi. Berikut beberapa taskflow dari tiap fitur yang ada :

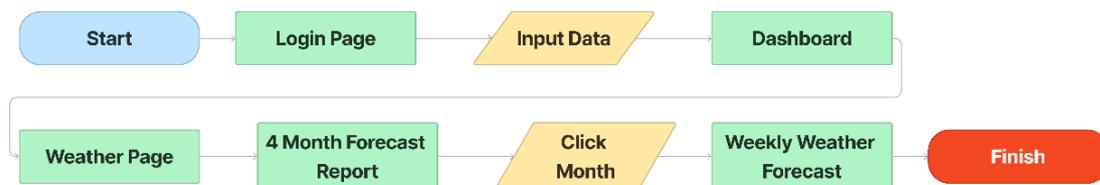
Gambar 3. Fitur CRUD

Pada gambar 3 diatas merupakan taskflow dari fitur CRUD yang memungkinkan user untuk memanajemen hasil produk pertanian.



Gambar 4. Fitur Inventory

Pada gambar 4 diatas merupakan *taskflow* dari fitur *inventory* yang memungkinkan user untuk melihat ringkasan dari isi penyimpanan produk pertanian mereka.



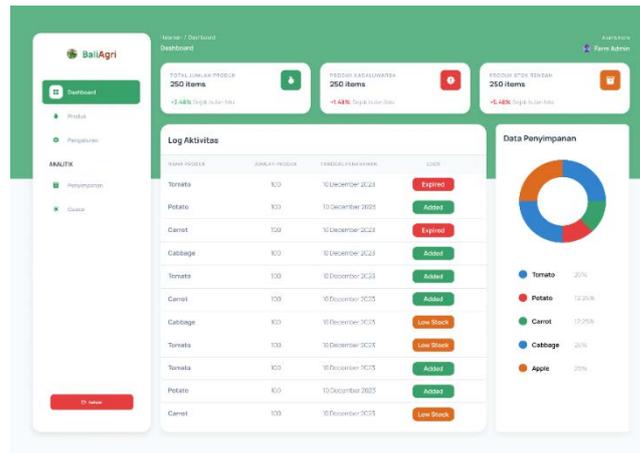
Gambar 5. Fitur Weather

Pada gambar 5 diatas merupakan *taskflow* dari fitur *weather* yang memungkinkan user untuk melihat ramalan cuaca selama 4 bulan kedepan

3.4. Prototype

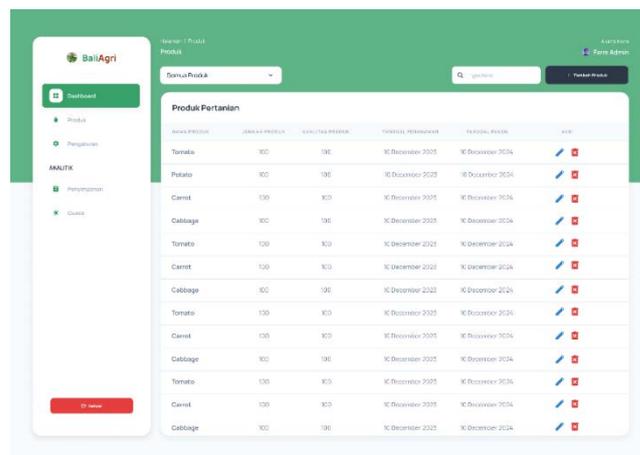
Pada tahapan ini, hasil dari proses *ideate* sebelumnya kemudian didesain menjadi sebuah tampilan antarmuka dari *dashboard* yang akan dikembangkan namun belum nyata. Dalam melakukan pengembangan prototype ini, penulis menggunakan aplikasi Figma.

- a. Low Fidelity Prototype
Low Fidelity Prototype merupakan representasi awal dari sebuah antarmuka pengguna. Pada tahapan ini, bentuk awal dari rancangan antarmuka dashboard dibuat sebagai dasar acuan untuk membuat desain antarmuka pengguna.
- b. Design System
Design system adalah kumpulan aturan, panduan, dan komponen desain yang digunakan secara konsisten dalam pengembangan antarmuka pengguna.
- c. High Fidelity Prototype



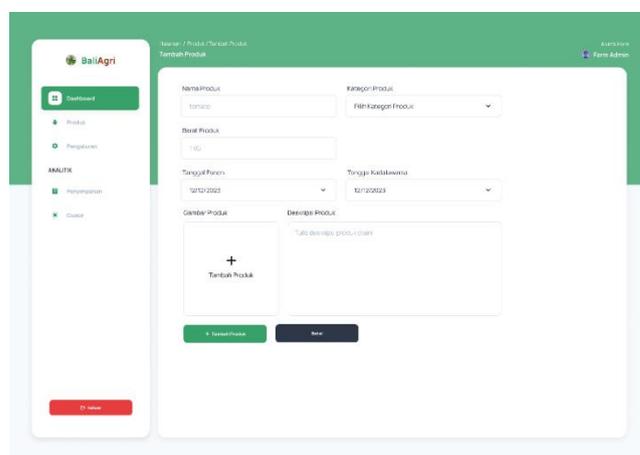
Gambar 6. Halaman Dashboard

Gambar 6 diatas merupakan desain *high fidelity prototype* dari halaman utama dashboard BaliAgri



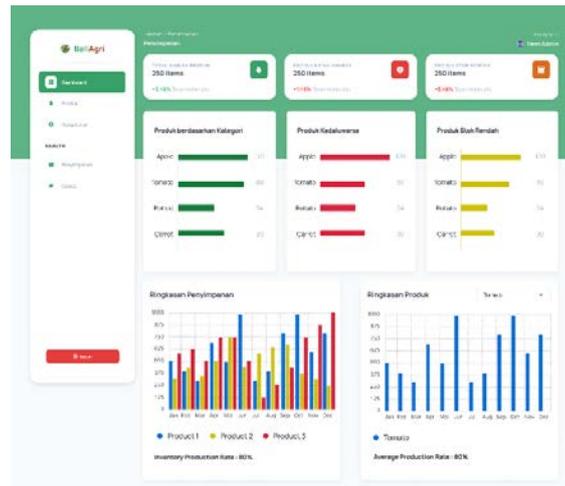
Gambar 7. Halaman Product List

Gambar 7 diatas merupakan desain *high fidelity prototype* dari halaman produk dashboard BaliAgri



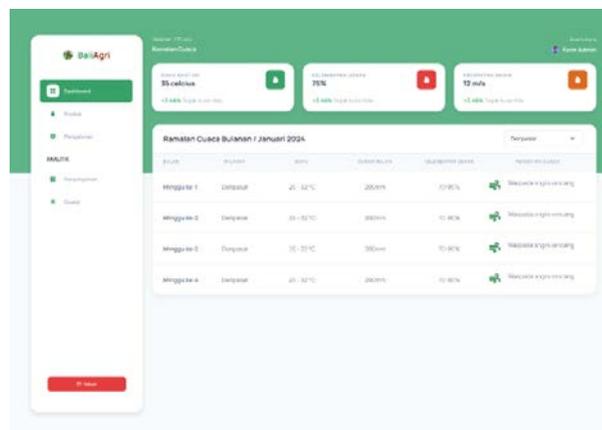
Gambar 8. Halaman Add Product

Gambar 6 diatas merupakan desain *high fidelity prototype* dari halaman tambah produk dashboard BaliAgri



Gambar 9. Halaman Inventory

Gambar 6 diatas merupakan desain *high fidelity prototype* dari halaman penyimpanan dashboard BaliAgri



Gambar 10. Halaman Weather

Gambar 10 diatas merupakan desain *high fidelity prototype* dari halaman cuaca dashboard BaliAgri

3.5. Pengujian

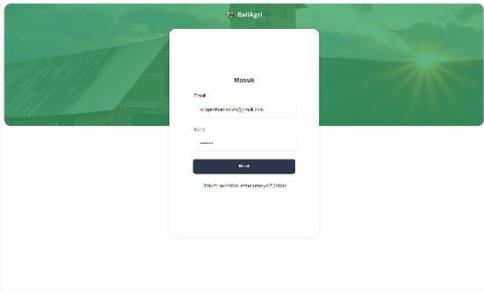
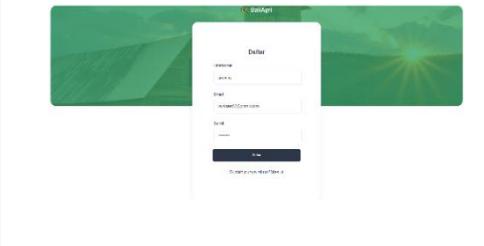
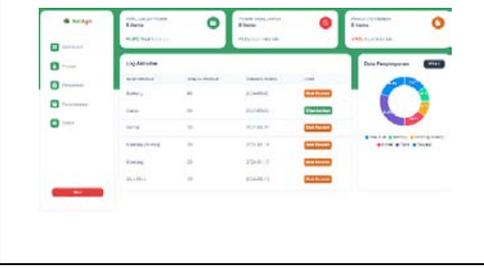
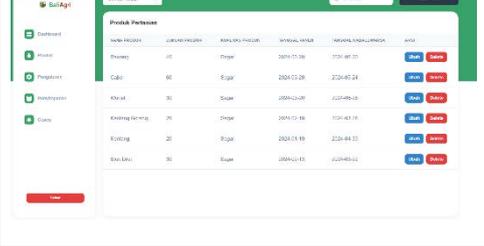
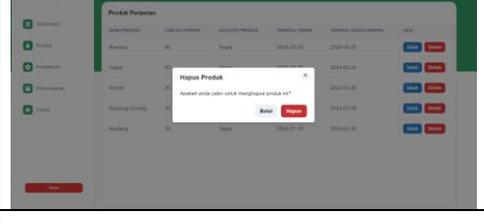
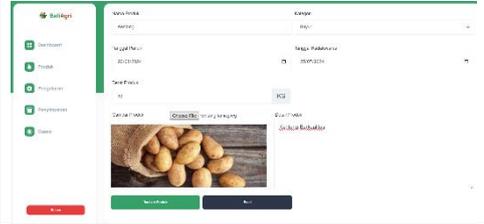
Pada tahapan ini, akan dilakukan pengujian terhadap hasil dari rancangan *dashboard* yang telah diimplementasikan. Pengujian yang dilakukan menggunakan 2 jenis pengujian yaitu *Black Box Testing* dan SUS. Hasil dari tahap pengujian akan dibahas lebih lanjut pada bagian selanjutnya

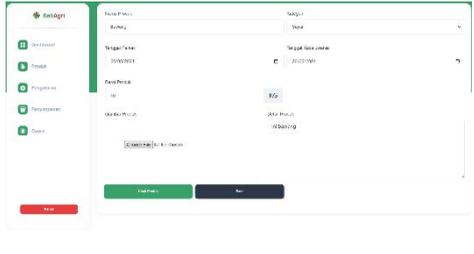
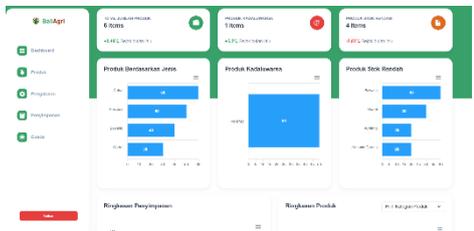
3.6. Hasil Evaluasi Sistem

Setelah implementasi sistem selesai dilakukan, sistem akan diuji menggunakan 2 buah jenis pengujian yaitu *blackbox testing* dan *SUS (System Usability Scale)*. Pada penelitian ini, peneliti memiliki 3 kriteria utama dalam melakukan pengujian, yakni rentan usia 18 – 60 tahun, sedang bekerja atau belajar di sektor pertanian, dan berdomisili di Bali. Berdasarkan kriteria tersebut, penulis melakukan pengujian dengan peserta yang berjumlah 30 orang.

Blackbox Testing atau pengujian *blackbox* merupakan pengujian yang difokuskan kepada pengujian fungsionalitas keseluruhan fitur – fitur dari system. Pada tabel 3 berikut merupakan daftar hasil pengujian menggunakan *blackbox testing* pada tiap halaman *dashboard* BaliAgri

Tabel 3. Pengujian Blackbox

No	Komponen yang diuji	Skenario dan Hasil Uji		
		Tampilan Aplikasi	Hasil yang diharapkan	Kesimpulan
1	Halaman Login – Fitur Login		Menampilkan halaman login yang berisikan input untuk melakukan login	VALID
2	Halaman Register – Fitur Register		Menampilkan halaman register yang berisikan input untuk melakukan register	VALID
3	Halaman Dashboard – Fitur Read Product		Menampilkan halaman dashboard yang berisi data produk dan grafik	VALID
4	Halaman Product List – Fitur Read Product, Search		Menampilkan halaman product list yang berisi data produk dan pencarian	VALID
5	Halaman Product List – Fitur Delete Product		Menampilkan modal konfirmasi untuk menghapus produk	VALID
6	Halaman Add Product - Fitur Create Product		Menampilkan halaman add product yang berisikan input data	VALID

7	Halaman Edit Product - Fitur Update Product		Menampilkan halaman edit product yang berisi data dan input data untuk update	VALID
8	Halaman Inventory – Fitur Read Product		Menampilkan halaman inventory yang berisi informasi grafis produk	VALID
9	Halaman Weather – Fitur Read Data		Menampilkan halaman inventory yang berisi informasi grafis produk	VALID
10	Halaman Profile		Menampilkan halaman profil yang berisi data profil	VALID

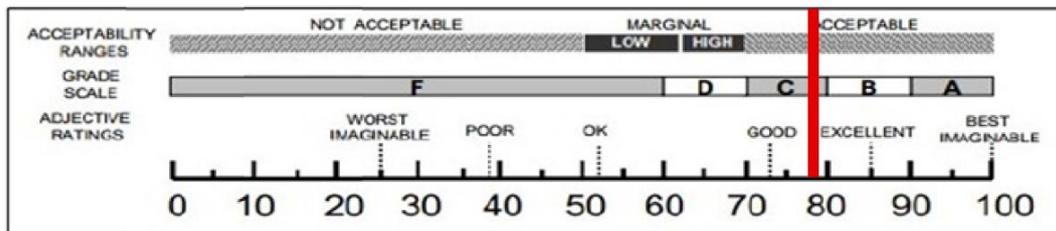
Pengujian sistem yang dilakukan dengan menggunakan metode SUS atau System Usability Scale mewajibkan peserta pengujian untuk melakukan pengisian kuisioner yang telah diberikan oleh penulis. Penilaian ini dilakukan berdasarkan skala Likert yaitu dengan rentang nilai 1 hingga 5. Pengujian ini berfokus pada kenyamanan pengguna bertujuan untuk mengukur sejauh mana sistem dapat memberikan pengalaman pengguna yang intuitif dan memuaskan.[8] Berikut ini merupakan hasil perhitungan skor pengujian SUS yang diperoleh dari 30 orang responden :

Responden	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	Jumlah	Skor SUS
Responden 1	3	4	4	3	4	3	4	3	4	3	35	87,5
Responden 2	2	3	3	2	4	3	3	3	3	2	28	70
Responden 3	4	3	3	3	4	2	4	3	4	1	31	77,5
Responden 4	3	4	4	3	3	2	4	4	4	4	31	77,5
Responden 5	4	3	3	3	4	3	3	3	3	3	35	87,5
Responden 6	3	4	4	4	3	4	4	4	3	4	32	80
Responden 7	3	4	4	4	4	4	3	4	4	4	38	95
Responden 8	3	1	4	3	4	4	4	3	3	2	31	77,5
Responden 9	3	3	3	3	3	3	4	3	3	3	31	77,5
Responden 10	3	4	4	4	3	3	4	4	3	4	36	90

Responden 11	3	2	3	2	3	3	3	2	3	1	25	62,5
Responden 12	3	2	3	2	3	3	3	2	3	1	25	62,5
Responden 13	3	3	2	4	3	3	3	4	4	3	32	80
Responden 14	4	3	3	3	3	1	3	3	3	2	28	70
Responden 15	4	3	4	4	3	2	3	4	4	4	35	87,5
Responden 16	4	4	4	4	4	4	4	4	4	0	36	90
Responden 17	3	3	4	3	2	4	4	3	3	3	32	80
Responden 18	4	4	4	4	4	4	4	4	4	4	40	100
Responden 19	4	4	4	4	4	4	4	4	4	4	40	100
Responden 20	3	4	4	3	4	4	4	4	4	3	37	92,5
Responden 21	3	2	3	2	3	3	3	2	3	1	25	62,5
Responden 22	4	3	4	4	4	4	4	4	4	4	39	97,5
Responden 23	3	3	2	4	3	3	3	4	4	3	32	80
Responden 24	4	4	4	3	4	3	4	4	4	4	38	95
Responden 25	3	1	4	3	4	2	3	3	3	3	29	97,5
Responden 26	3	3	3	2	3	3	3	3	3	3	29	72,5
Responden 27	3	3	3	3	3	3	4	3	3	3	31	77,5
Responden 28	3	4	4	4	3	3	4	4	3	4	36	90
Responden 29	3	2	3	2	3	3	3	2	3	1	25	62,5
Responden 30	4	4	4	4	4	4	4	4	4	4	40	100
Rata – Rata Skor SUS											79,75	

Tabel 4. Hasil Perhitungan Skor SUS

Berdasarkan hasil perhitungan skor SUS pada tabel 4, yang merupakan metode yang umum digunakan untuk mengukur kegunaan sistem dengan memberikan nilai numerik, dashboard BaliAgri mendapatkan skor yang cukup baik yakni 79,75.



Gambar 10. Hasil SUS Skor Pengujian Sistem

Berdasarkan skala SUS pada gambar 10, skor rata – rata tersebut menunjukkan bahwa pengguna memiliki penilaian yang positif terhadap kebergunaan dari dashboard Bali Agri.

4. Kesimpulan

Perancangan *dashboard Smart Farming* menggunakan metode *Design Thinking* dengan mengimplementasikan prinsip Gestalt pada desain antarmukanya berjalan melalui 6 tahapan yaitu tahapan *empathize*, *define*, *ideate*, *prototype*, *testing* dan implementasi. Pada tahapan pengujian *blackbox* sistem dashboard, yang melibatkan 15 orang responden untuk melakukan pengujian komponen sistem dengan 10 skenario pengujian, diperoleh persentase sebesar 100% untuk tingkat ketepatan dan kebenaran sistem. Hasil pengujian ini membuktikan bahwa seluruh sistem berjalan sesuai dengan alur scenario yang ada sehingga dashboard dapat digunakan dengan baik. Selanjutnya pada pengujian *usability dashboard* menggunakan *System Usability Scale (SUS)* dengan memberikan 10 buah pertanyaan dengan 30 orang responden mendapatkan nilai yang cukup baik yaitu 79,75 yang menunjukkan bahwa dashboard memiliki kebergunaan yang baik. Berdasarkan *Adjective Rating*,

dashboard ini diklasifikasikan sebagai “good” dan dalam *grade scale* masuk kedalam skala B yakni kategori baik. Dalam rentang penerimaan, dashboard ini dinilai “acceptable” yang artinya mayoritas pengguna menerima dan merasa puas dalam menggunakan *dashboard* ini. Dengan demikian, berdasarkan hasil kedua pengujian tersebut, *dashboard* BaliAgri memiliki kualitas yang baik dalam memenuhi kebutuhan penggunaanya

Daftar Pustaka

- [1] Indonesia, B., 2023. Laporan Perekonomian Provinsi Bali
- [2] Johan, D., Maarif, M.S. and Zulfainarni, N., 2022. Persepsi Petani Terhadap Digitalisasi Pertanian untuk Mendukung Kemandirian Petani. *Jurnal Aplikasi Bisnis dan Manajemen (JABM)*
- [3] Ilham H, Wijayanto B, Rahayu SP. 2021. Analysis and Design of User Interface/User Experience With the Design Thinking Method in the Academic Information System of Jenderal Soedirman University. *Jurnal Teknik Informatika (Jutif)*.
- [4] Hovorushchenko T, Pavlova O, Kobel K. 2019. Method of Evaluating the User Interface of Software Systems for Compliance with the Gestalt Principles. *IEEE 14th International Conference on Computer Sciences and Information Technologies (CSIT) 2019 Sep 17 (Vol. 2, pp. 138-141)*. IEEE
- [5] Dam, R. F. and Teo, Y. S. 2024. What is Design Thinking and Why Is It So Popular?.
- [6] Shirvanadi, E.C., 2021. Perancangan Ulang UI/UX Situs E-Learning Amikom Center Dengan Metode Design Thinking (Studi Kasus: Amikom Center).
- [7] Khoiruddin, A.Z.K.A.A., 2017. Implementasi Gestalt Principles pada Rancang Bangun Aplikasi Berbasis Android ClearRoute. Institut Teknologi Sepuluh Nopember
- [8] Lewis, J.R., 2018. The system usability scale: past, present, and future. *International Journal of Human-Computer Interaction*

This page is intentionally left blank.

Analisis Serangan Cross Site Scripting (XSS) Pada Website OASE Menggunakan Metode OWASP

Muhammad Arrysatrya Yusuf Putranda^{a1}, I Komang Ari Mogi^{a2}, I Gusti Ngurah Anom Cahyadi Putra^{a3},
I Made Widiartha^{a4}

^aProgram Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana
Jalan Raya Kampus Unud, Jimbaran, Bali, 80361, Indonesia

¹arrysatrya097@student.unud.ac.id

²arimogi@unud.ac.id

³anom.cp@unud.ac.id

⁴madewidiartha@unud.ac.id

Abstract

Berkembangnya internet membuat majunya era teknologi digitalisasi, dimana hampir seluruh sektor kini dapat diakses secara digital, termasuk pendidikan. Salah satunya adalah Universitas Udayana yang memiliki LMS bernama *Online Academic Service for Elearning* (OASE) yang digunakan dalam proses pembelajaran di lingkungan Universitas Udayana. Namun perkembangan ini diikuti oleh potensi ancaman, dengan terwujudnya digitalisasi yang berarti hal tersebut dapat diakses oleh siapa saja, termasuk orang yang merusak suatu sistem. Salah satu jenis serangan yang banyak ditemukan adalah *Cross Site Scripting* (XSS). Untuk memastikan keamanan LMS OASE milik Universitas Udayana, perlu dilakukan Analisis Kerentanan terutama pada serangan XSS yang dilakukan dengan uji penetrasi menggunakan metode OWASP. Dari hasil pengujian ditemukan bahwa meskipun OASE memiliki beberapa potensi celah kerentanan, namun hanya satu fitur saja yang dikonfirmasi memiliki kerentanan, sedangkan fitur lainnya berhasil diproteksi dengan adanya fungsi *filtering* serta kontrol pada eksekusi script dari pengguna.

Kata Kunci: *Cross Site Scripting, OASE, Uji Penetrasi, OWASP, CVSS 3.1, Kerentanan*

1. Pendahuluan

Sektor-sektor tertentu mulai digitalisasi sebagai akibat dari pesatnya kemajuan teknologi, salah satunya adalah pendidikan [1]. Era digital adalah era di mana orang dapat memahami teknologi dan terhubung satu sama lain. Salah satu universitas Indonesia di Bali, Universitas Udayana, menggunakan *website* LMS *Online Academic Service for Elearning* (OASE) untuk mendukung proses pembelajaran. Sangat penting untuk menjaga keamanan data dalam era teknologi saat ini, karena banyak data yang disimpan dalam jaringan dan harus dapat diakses setiap saat.

Berdasarkan Lanskap Keamanan Siber Tahun 2022 yang dirilis oleh Badan Siber dan Sandi Negara [2]. Indonesia menempati peringkat 1 negara tujuan anomali dengan total 539.933.976. Lebih lanjut lagi berdasarkan notifikasi insiden keamanan siber yang diterima Badan Siber dan Sandi Negara, terdapat 1.433 notifikasi selama tahun 2022 dimana 336 notifikasi insiden siber diantaranya berasal dari sektor pendidikan. Terdapat 933 insiden berupa *web defacement*, 138 insiden *data breach*, serta 122 kerentanan *Cross Site Scripting*. Besarnya gap diantara top 3 indikasi insiden dari notifikasi yang dikirimkan tersebut menunjukkan bahwa masih tingginya indikasi serangan siber yang mengarah pada aplikasi *website* [2].

Menurut penelitian yang dilakukan oleh Ghozali *et.al* pada 2019, peneliti melakukan penelitian untuk mendeteksi kerentanan serta menilai risiko keamanan pada aplikasi *website* menggunakan metode *Open Web Application Security Project* (OWASP), dari hasil penelitian didapatkan bahwa *website* tersebut terdapat tiga kerentanan, yakni *SQL Injection*, *Cross Site Scripting* (XSS), dan *Broken Authentication*. Penelitian tersebut juga menghasilkan dua faktor, yakni *impact* dan *likelihood* dimana kedua faktor tersebut digunakan untuk membuat skala prioritas mengenai kerentanan mana yang harus diperbaiki [3].

Penelitian lain yang dilakukan oleh Hakim *et.al* pada 2020 mengenai serangan *Cross Site Scripting* (XSS) berdasarkan *Base Metric* dari CVSS v.2, dimana penelitian ini mengukur tingkat kerentanan serangan XSS berdasarkan *Base Metric* CVSS v.2. Adapun hasil penelitian ini didapatkan bahwa meskipun kerentanan yang ditemukan berupa kerentanan XSS, namun dampak yang dihasilkan berbeda pada tiap *endpoint*, dimana hal ini dipengaruhi oleh beberapa faktor seperti kode pemrograman pada fitur yang diuji, *payload* pengujian yang digunakan, serta kemampuan dari penyerang [4].

Berdasarkan uraian diatas serta hasil dari beberapa penelitian terdahulu yang membahas mengenai serangan XSS, dapat dipahami bahwa kerentanan berupa XSS sangat umum pada suatu *website*. Penulis melakukan penelitian ini untuk melakukan Analisis serangan *Cross Site Scripting* (XSS) Pada *Website* OASE yang merupakan salah satu *website* penunjang proses pembelajaran pada Universitas Udayana. Adapun metode yang digunakan adalah metode *Open Web Application Security Project* (OWASP) sebagai acuan, hal ini didasari dari beberapa penelitian sebelumnya dimana metode OWASP digunakan untuk mencari kerentanan pada *website*, dan hasilnya metode ini mampu mendeteksi kerentanan yang ada, serta memberikan rekomendasi dari hasil analisis keamanan untuk meningkatkan keamanan pada sistem.

2. Metode Penelitian

Metode penelitian yang digunakan adalah metode kualitatif dimana metode ini menggunakan metode kualitatif untuk menentukan kerentanan yang ditemukan serta analisis dampak kerentanan tersebut. Lebih lanjut lagi penelitian ini menggunakan metode OWASP dalam melakukan pengujian, yang terdiri dari *Information Gathering, Penetration Testing, Report and Evaluation*.

2.1 Kajian Pustaka

a. Vulnerability

Vulnerability atau kerentanan merupakan kelemahan atau celah yang terdapat pada *website* maupun sistem yang dapat dimanfaatkan untuk merusak sistem maupun mencuri informasi. Kerentanan dapat terjadi pada *software*, sistem operasi, modul, sampai pengguna yang menggunakannya [5]. *Vulnerability* dapat dideteksi dan ditemukan melalui pemindaian kerentanan serta uji penetrasi.

b. Cross Site Scripting (XSS)

Cross Site Scripting atau XSS merupakan salah satu dari 10 jenis kerentanan yang masuk ke dalam OWASP Top 10 Vulnerability 2017. Serangan ini dilakukan dengan cara menginjeksi *script* berbahaya ke dalam suatu *website* yang memungkinkan penyerang mencuri data ataupun merusak sistem yang ada [6]. Serangan XSS selain berdampak kepada pengguna, juga berdampak kepada sistem, terdapat 3 jenis serangan XSS yakni *Stored* dan *Reflected* XSS yang memanfaatkan kerentanan pada sisi server, serta *DOM-Based* XSS yang memanfaatkan kerentanan pada sisi client.

c. Open Web Application Security Project (OWASP)

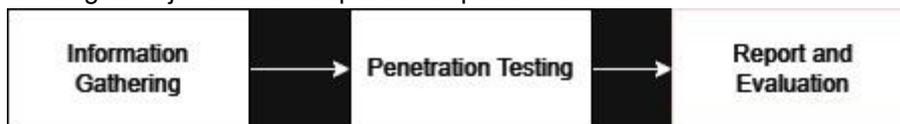
Open Web Application Security Project (OWASP) merupakan sebuah organisasi yang berfokus pada peningkatan keamanan *software* aplikasi *website*. Dimana tujuan utama dari organisasi ini adalah mengedukasi, mengembangkan, dan menyediakan sumber daya bagi industri serta komunitas teknologi informasi dalam upaya peningkatan keamanan *software* aplikasi *website* [7]. OWASP juga memiliki daftar top 10 kerentanan pada *website* yang paling sering ditemukan, berikut ini merupakan top 10 kerentanan pada *website* yang dirilis pada tahun 2017 seperti yang dapat dilihat pada **Tabel 1**.

Tabel 1. OWASP Top 10 Application Security Risk 2017

A1:2017-Injection
A2:2017-Broken Authentication
A3:2017-Sensitive Data Exposure
A4:2017-XML External Entities (XXE)
A5:2017-Broken Access Control
A6:2017-Security Misconfiguration
A7:2017-Cross-Site Scripting (XSS)
A8:2017-Insecure Deserialization
A9:2017-Using Components with Known Vulnerabilities
A10:2017-Insufficient Logging & Monitoring

2.2 Metodologi Uji Penetrasi

Penelitian ini menggunakan metode OWASP dalam melakukan uji penetrasi, dimana pengujian ini dimulai dari *Information Gathering*, *Penetration Testing*, dan *Report and Evaluation*. Adapun *flowchart* dari langkah Uji Penetrasi dapat dilihat pada **Gambar 1**.



Gambar 1. Flowchart Uji Penetrasi Metode OWASP

Seperti yang ditampilkan pada **Gambar 1**, langkah Uji Penetrasi dimulai dengan melakukan *Information Gathering* berupa *endpoint* yang berpotensi memiliki celah keamanan, dalam penelitian ini seperti fitur search, chat, profile, calendar, dan course. Dimana setelah mendapatkan *endpoint* yang akan diuji, dilanjutkan dengan melakukan *Penetration Testing*.

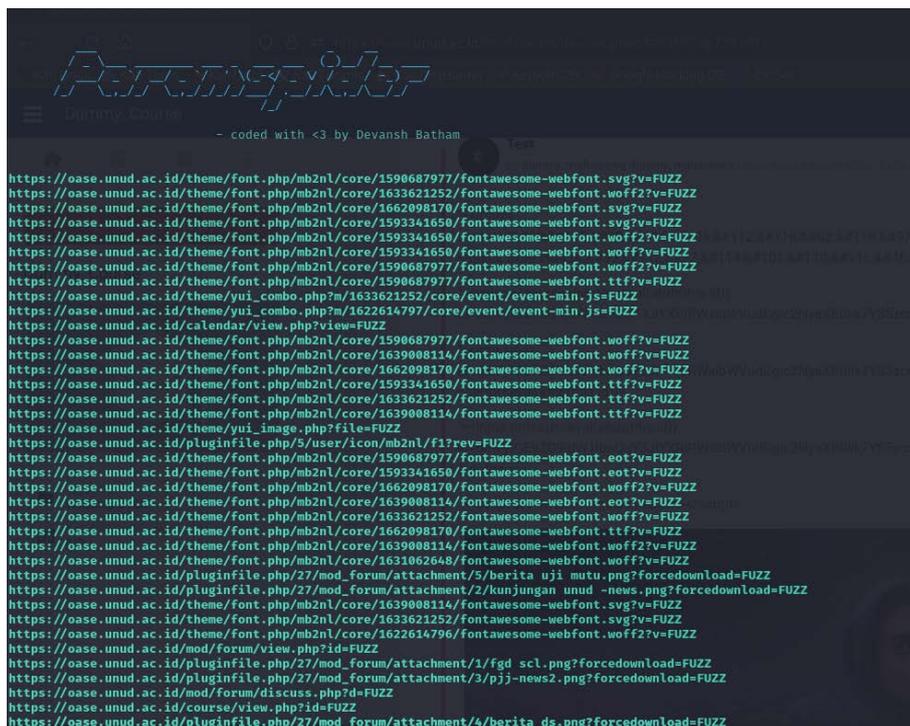
Langkah selanjutnya adalah melakukan *Penetration Testing* pada *endpoint* yang sudah didapatkan pada langkah *Information Gathering*, adapun metode ujinya adalah dengan menggunakan *payload* standar pengujian XSS pada *endpoint* tersebut, jika terdapat kerentanan XSS maka akan dilanjutkan untuk mencari sejauh apa potensi kerusakan dari kerentanan yang ditemukan.

Terakhir adalah membuat *Report and Evaluation*. Jika ditemukan kerentanan dari hasil *Penetration Testing*, maka dilakukan pengujian lebih lanjut untuk mencari tingkat *severity* dari kerentanan yang ditemukan tersebut, selanjutnya hasil pengujian diarsipkan dalam *Proof of Concept* yang memuat informasi mengenai kerentanan yang ditemukan seperti *endpoint*, *payload*, *documentation*, *step to recreate*, serta *recommendation*. Dimana *proof of concept* ini dapat digunakan sebagai acuan bagi pemilik sistem untuk memperbaiki kerentanan yang ditemukan.

3. Hasil dan Pembahasan

3.1 Information Gathering

Pada tahapan ini dilakukan pengumpulan informasi mengenai sistem OASE, dimana ditemukan beberapa *endpoint* yang berpotensi memiliki celah kerentanan terhadap serangan XSS, diantaranya adalah fitur Chat, Search Course, Profile, Forum, Assignment, Calendar, dan Quiz. Pengumpulan informasi juga dibantu menggunakan framework *Paramspider* untuk mencari parameter yang berpotensi memiliki celah kerentanan seperti yang dapat dilihat pada **Gambar 2**.

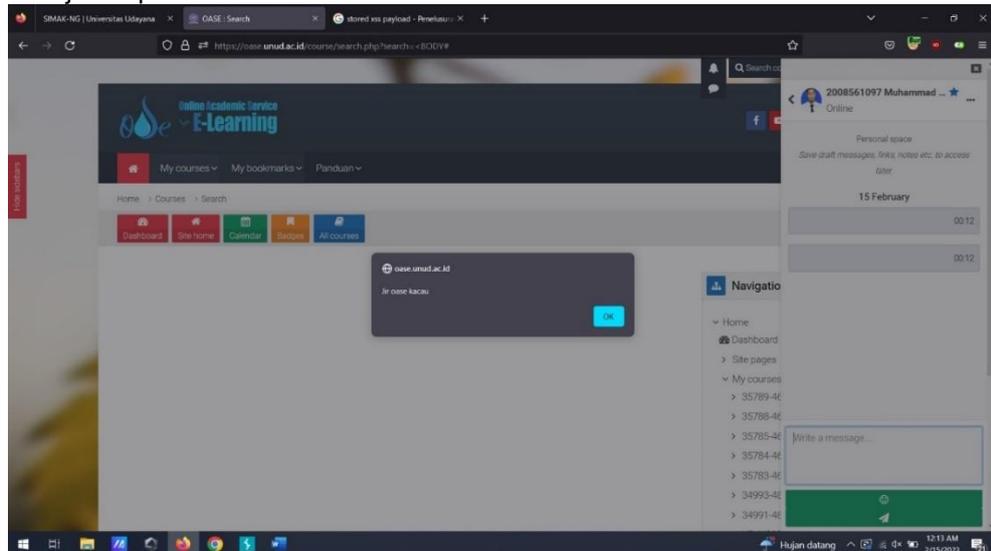


Gambar 2. Proses Pencarian Endpoint dengan Paramspider

3.2 Penetration Testing

a. Fitur Chat

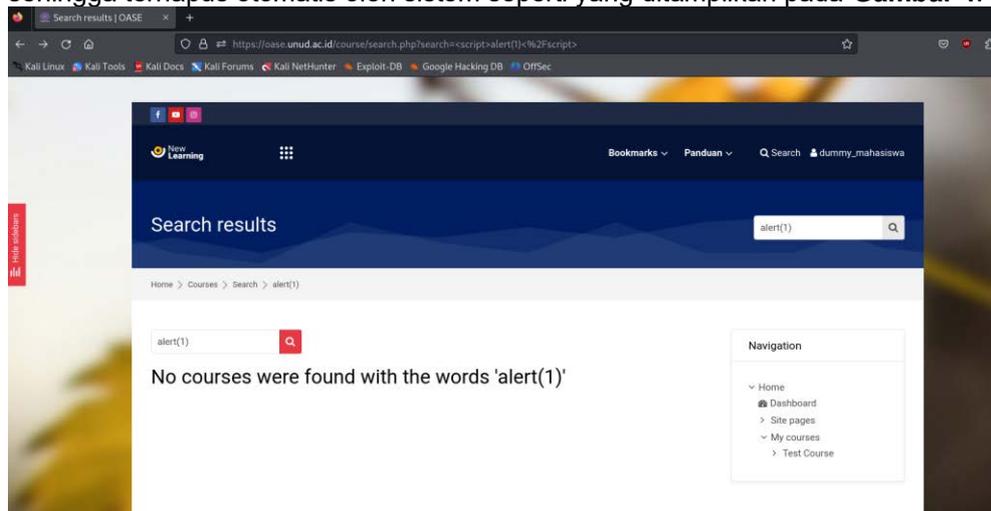
Terdapat kerentanan yang ditemukan pada fitur chat, dimana kerentanan ini memungkinkan pengguna untuk mengirimkan *script* berisi *payload* XSS kepada pengguna lain kemudian *payload* tersebut akan dieksekusi oleh sistem, seperti yang ditunjukkan pada **Gambar 3**.



Gambar 3. XSS Fitur Chat

b. Fitur Search

Terdapat celah kerentanan pada fitur search, dimana fungsi ini dapat menerima masukan user dari field maupun URL. Setelah dilakukan pengujian dengan *payload* standar pengujian XSS, ditemukan bahwa meskipun URL menampilkan *script* secara keseluruhan, namun fungsi *search* melakukan *filtering* kata `<script>` dan tanda `<` sehingga terhapus otomatis oleh sistem seperti yang ditampilkan pada **Gambar 4**.

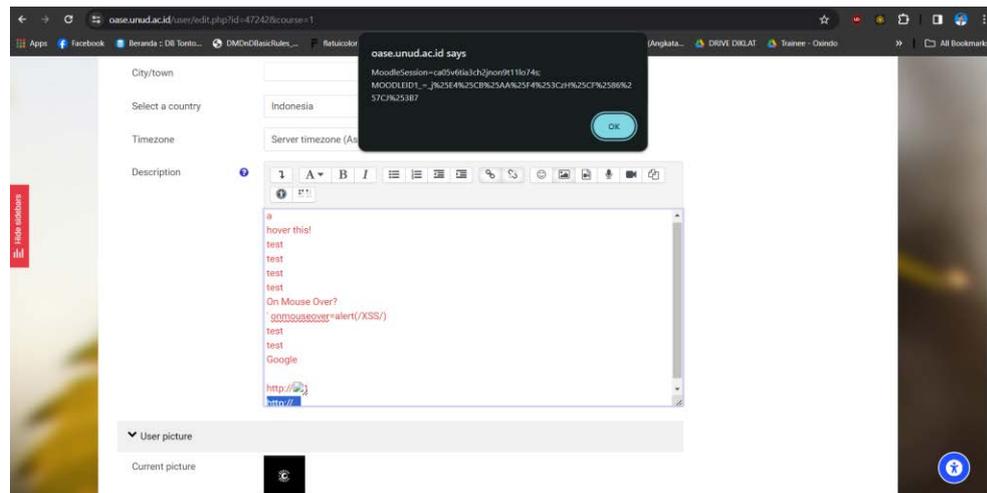


Gambar 4. Function Filtering Pada Fitur Search

Pengujian kemudian dilakukan dengan *payload* yang tidak menggunakan karakter yang difilter, hasilnya tidak ditemukan celah kerentanan karena sistem tidak mengeksekusi *script* yang terdapat pada *payload* tersebut.

c. Fitur Profile

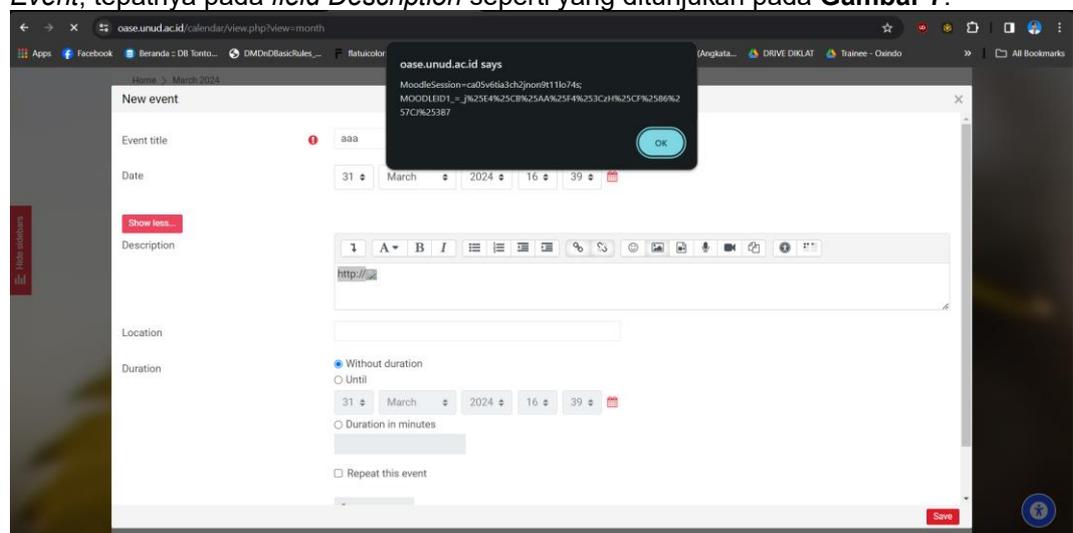
Seperti yang ditunjukkan pada **Gambar 5**, fitur profile memiliki *Field Text* yang dapat digunakan untuk memasukkan *payload* serangan XSS. Pengujian dilakukan dengan menggunakan *payload* secara langsung dan dengan *hyperlink*. *Payload* yang digunakan adalah *payload* yang memasukkan gambar, dimana jika terdapat *error* maka *browser* akan menampilkan *document cookies* dari pengguna. Dimana pada fitur ini dapat dilihat bahwa terdapat kerentanan serangan XSS.



Gambar 5. Payload Profile

d. Fitur Calendar

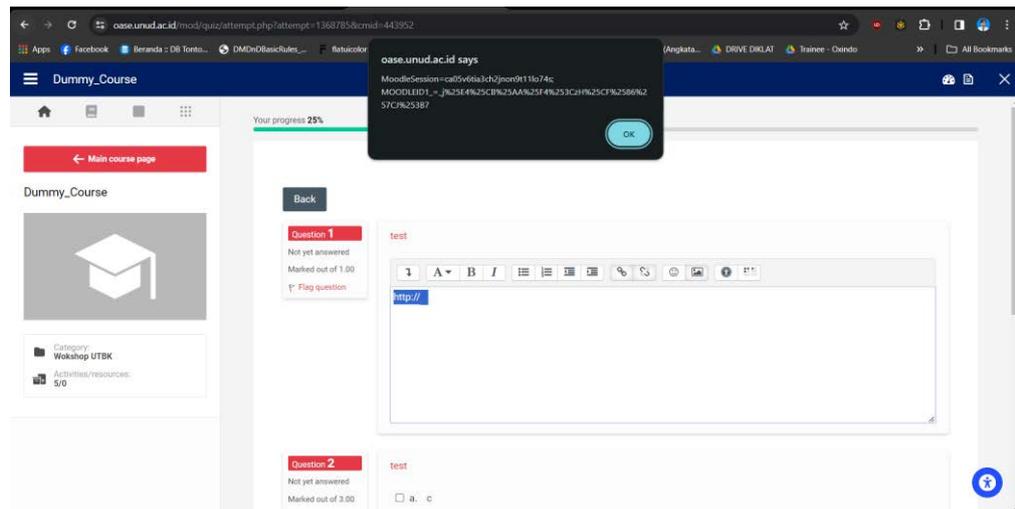
Pada fitur calendar, potensi celah keamanan ditemukan pada fungsi *Add Event*, yakni field *Event Name* dan *Description*. Dilakukan pengujian yang sama seperti pada profile, yakni menggunakan *payload* standar pengujian XSS dan *payload* yang disisipkan menggunakan *hyperlink*. Hasilnya ditemukan kerentanan serangan XSS pada fitur *Add Event*, tepatnya pada *field Description* seperti yang ditunjukkan pada Gambar 7.



Gambar 6. Penetration Testing Pada Calendar

e. Fitur Forum

Potensi celah keamanan pada fitur forum ditemukan pada bagian *Subject* dan *Message forum*. Dimana pada kedua bagian ini terdapat *field* untuk mengisi *text* dan *script*. Setelah dilakukan pengujian menggunakan *payload* standar pengujian XSS dan *payload* yang disisipkan menggunakan *hyperlink* yang akan memunculkan *pop up* ketika terjadi error dalam load gambar, dan hasilnya ditemukan kerentanan *pada field Message* di fitur Forum seperti yang ditunjukkan pada Gambar 7.



Gambar 10. Hasil Penetration Testing Long Answer Quiz



Gambar 11. Hasil Penetration Testing Short Answer Quiz

3.3 Report and Evaluation

Terdapat kerentanan serangan *Cross Site Scripting* atau XSS pada hampir semua fitur yang diuji. Namun meskipun demikian, secara garis besar hanya terdapat dua fitur atau *endpoint* yang memiliki kerentanan yang ditemukan. Pertama adalah *field* pesan pada fitur Chat, pengujian lebih lanjut tidak dapat dilakukan karena fitur tersebut sudah ditiadakan oleh pemilik sistem selama proses penelitian. Dari pengujian awal, didapatkan dampak terburuk dari kerentanan ini adalah penyerang dapat mengambil *cookies* dari pengguna, sehingga akun OASE pengguna dapat diambil alih oleh penyerang.

Kedua adalah *field* message, tepatnya pada fitur *hyperlink*. Meskipun berada pada menu dan halaman yang berbeda, namun seluruh kerentanan XSS yang ditemukan pada dasarnya bersumber dari fitur *hyperlink* yang ada pada *field* message. Pada pengujian lebih lanjut, dampak yang ditimbulkan dari kerentanan ini tidak berdampak besar pada sistem, satu satunya dampak yang dihasilkan adalah *pop up* error tidak dapat ditutup oleh pengguna kecuali pengguna melakukan *refresh* halaman dimana hal ini terbilang mengganggu ketersediaan sistem. Lebih lanjut lagi, dampak yang dihasilkan hanya berlaku kepada sistem milik penyerang dan tidak berdampak kepada pengguna lain, dimana dampaknya akan hilang ketika halaman direfresh.

4. Kesimpulan

Berdasarkan hasil penelitian ini, disimpulkan terdapat beberapa potensi celah kerentanan serangan XSS pada *Website* OASE, dimana terdapat dua fitur yang kerentanannya dikonfirmasi, yaitu fitur Chat yang langsung diperbaiki oleh pengembang sistem untuk menghindari kerusakan lebih lanjut serta fitur *hyperlink* pada *field* message.

Adapun dampak yang dihasilkan tidak terlalu berdampak kepada sistem. Dimana kerentanan pada fitur *hyperlink* ini hanya berdampak kepada sisi penyerang serta tidak memiliki dampak kepada pengguna lain, serta hanya berdampak kepada ketersediaan sistem dimana hal ini akan kembali normal ketika halaman *direfresh*.

References

- [1] I. Riadi, R. Umar, and T. Lestari, "Analisis Kerentanan Serangan Cross Site Scripting (XSS) pada Aplikasi Smart Payment Menggunakan Framework OWASP," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 5, no. 3, pp. 146–152, 2020, doi: 10.14421/jiska.2020.53-02.
- [2] BSSN, "Rilis Launching Lanskap Keamanan Siber Tahun 2022," *Bssn*, vol. 1, no. 9, p. 20, 2023.
- [3] B. Ghozali, K. Kusriani, and S. Sudarmawan, "Mendeteksi Kerentanan Keamanan Aplikasi Website Menggunakan Metode Owasp (Open Web Application Security Project) Untuk Penilaian Risk Rating," *Creat. Inf. Technol. J.*, vol. 4, no. 4, p. 264, 2019, doi: 10.24076/citec.2017v4i4.119.
- [4] A. S. Hakim, T. A. Cahyanto, and H. Azizah, "Serangan cross-site scripting (XSS) berdasarkan base metric CVSS V.2," *J. Smart Teknol.*, vol. 2, no. 1, 2020.
- [5] S. Shekhar and A. Agrawal, "A survey of website security vulnerabilities," *J. Inf. Secur. Appl.*, vol. 53, no. 102536, 2020.
- [6] V. Nithya, S. Lakshmana Pandian, and C. Malarvizhi, "A survey on detection and prevention of cross-site scripting attack," *Int. J. Secur. its Appl.*, vol. 9, no. 3, pp. 139–152, 2015, doi: 10.14257/ijssia.2015.9.3.14.
- [7] OWASP, "OWASP Top Ten Project." Accessed: Mar. 20, 2023. [Online]. Available: <https://owasp.org/www-project-top-ten/>

Pembangunan Aplikasi Pengenalan Kualitas Rasa Buah Jeruk Siam Kintamani Menggunakan GLCM dan K-NN Berbasis *User Centered Design* (UCD)

Putu Bagus Dio Pranata.^{a1}, I Gede Arta Wibawa.^{a2}, Made Agung Raharja.^{b3}, I Gusti Ngurah Anom Cahyadi Putra.^{b4}

^aInformatika, Universitas Udayana
Kuta Selatan, Badung, Bali, Indonesia
¹bagusdio01@gmail.com
²gede.arta@unud.ac.id (Corresponding author)

^bInformatika, Universitas Udayana
Kuta Selatan, Badung, Bali, Indonesia
³made.agung@unud.ac.id
⁴anom.cp @unud.ac.id

Abstrak

Buah Jeruk Siam Kintamani merupakan salah satu komoditas unggulan yang memiliki variasi kualitas rasa yang signifikan. Oleh karena itu, diperlukan suatu alat yang dapat membantu pengguna dan calon konsumen, dalam mengenali kualitas rasa buah dengan cepat dan akurat. Metode GLCM digunakan untuk mengekstraksi fitur tekstur dari citra buah Jeruk Siam Kintamani, sementara algoritma K-NN akan memproses fitur-fitur tersebut untuk mengklasifikasikan kualitas rasa buah. Pendekatan UCD diterapkan dalam pengembangan aplikasi untuk memastikan bahwa antarmuka pengguna dirancang dengan mempertimbangkan kebutuhan dan preferensi pengguna akhir. Pengujian usability testing dan uji akurasi algoritma Gray Level Co – Occurrence Matrix (GLCM) dan K – Nearest Neighbors (KNN) dilakukan, dengan memanfaatkan hasil pengujian usability, serta akurasi algoritma pengenalan kualitas rasa. Hasil dari penelitian ini memiliki hasil tingkat usability sangat baik, dalam masing – masing aspek yang di uji, yaitu aspek learnability sebesar 80%, efficiency 83%, hasil nilai skor system usability scale (SUS) sebesar 86,3%, dan error defective rate sebesar 14% kemudian dikatakan valid serta reliabel, serta sudah baik dan dapat diterima oleh pengguna aplikasi ini.

Keywords: *Android, Jeruk, Citrus, User Experience, User Centered Design, Usability Testing, GLCM, KNN*

1. Pendahuluan

Jeruk (*Citrus*) merupakan jenis tanaman buah tahunan yang berasal dari Asia dan diyakini berasal dari Cina. Buah jeruk dapat dikonsumsi secara langsung, dimana buah jeruk ini merupakan sumber dari *Vitamin C*, berguna untuk kesehatan tubuh manusia, dengan kandungan vitamin *C* beragam, dan buah jeruk memiliki beberapa jenis yang banyak dikembangkan, yaitu seperti jeruk limau, kemudian jeruk manis, dan jeruk siam [1].

Hasil dari budidaya buah jeruk tentu saja memiliki kualitas rasa yang berbeda berdasarkan hasil survey yang dilakukan secara manual. Memilih buah berdasarkan rasa

Pembangunan Aplikasi Pengenalan Kualitas Rasa Buah Jeruk Siam Kintamani Menggunakan GLCM Dan K-NN Berbasis User Centered Design (UCD)

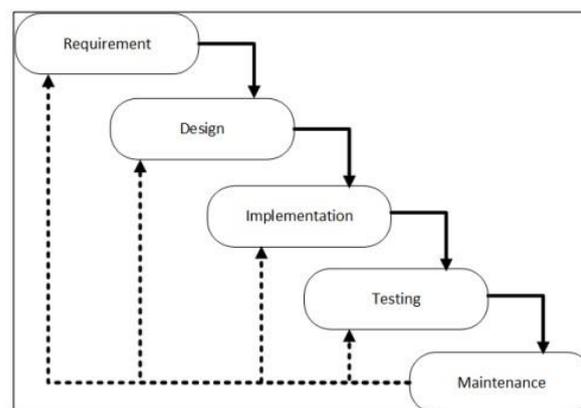
(manis) dari buah jeruk tentu saja menjadi suatu hal yang penting bagi petani serta penjual, dan sering kali pembeli atau konsumen menggunakan informasi yang diberikan oleh penjual sebagai informasi secara 2 deskriptif, serta informasi yang diberikan kepada konsumen memiliki konsistensi rendah, atau secara subjektif. Dengan perkembangan sistem informasi dan teknologi saat ini, muncul banyak aplikasi yang dapat membantu konsumen atau *end user* dalam berbagai hal, di kehidupan sehari – hari. Pada perancangan sebuah aplikasi, memerlukan beberapa pendekatan, salah satunya adalah pendekatan kepada pengguna (*user*).[3]

Pada penelitian dilakukan oleh Iqbal, melakukan penerapan metode *User Centered Design* pada perancangan aplikasi darurat, berbasis android, sesuai dengan kebutuhan pengguna seperti fitur kondisi terkini, kemudian berita, petunjuk arah, dsb [7]. Dalam penelitian ini, melakukan perancangan sebuah aplikasi dengan beberapa pendekatan, salah satunya adalah pendekatan kepada pengguna (*user*), baik dalam tujuan, hingga masukan, dapat menghasilkan aplikasi lebih dipahami oleh penggunanya, kemudian aplikasi akan melakukan klasifikasi pengenalan rasa buah jeruk menggunakan *GLCM* dan *KNN* berbasis mobile dilakukan menggunakan metode pengembangan perangkat lunak seperti *SDLC* (*System Development Life Cycle*), dengan pendekatan *UCD* (*User Centered Design*).

2. Metodologi Penelitian

2.1 Metode Pengembangan Sistem

Aplikasi pengenalan kualitas rasa buah jeruk ini menggunakan metode pengembangan sistem Waterfall Mode. Alur pengembangan sistem dengan metode waterfall dapat dilihat pada Gambar 1 berikut.



Gambar 1. Metode *Waterfall*

Berikut adalah penjelasan dari setiap proses yang ada di dalam pengembangan sistem menggunakan metode Waterfall :

2.1.1 Requirement Definition

Pada tahapan ini melakukan persiapan dan analisa kebutuhan sistem yang ingin dikembangkan. Persiapan dan analisa sistem seperti, persiapan data, manfaat dari sistem yang akan dibuat dan batasan-batasan dari sistem yang

2.1.2 System and Software Design.

Pada tahapan ini melakukan proses perancangan desain sistem dengan menggunakan diagram UML seperti, use case diagram, flowchart dan activity diagram.

2.1.3 Implementation

Pada tahapan ini merupakan tahapan gambaran sistem ke dalam bentuk code. Jadi pada tahapan ini lebih berfokus pada hal teknis dari hasil desain yang telah dibuat pada tahap sebelumnya dan desain tersebut akan diterjemahkan ke dalam bahasa pemrograman.

2.1.4 System Testing

Pada tahapan ini sistem yang telah dibuat akan diuji dari berbagai sisi baik dari aspek desain dan fungsionalitas. Dengan tujuan untuk menentukan apakah terjadi kesalahan pada sistem yang telah dibuat, dimana jika ada kesalahan masih bisa dicegah dan diperbaiki kembali.

2.1.5 Operation and Maintenance

Pada tahapan ini merupakan tahapan pemeliharaan sistem yang mencakup perbaikan kesalahan seperti bug, error, dan penambahan fitur-fitur pada sistem yang telah dibuat.

2.2 Data Penelitian

Data penelitian yang digunakan berdasarkan cara memperolehnya yaitu data primer. Data primer adalah data yang diperoleh peneliti secara langsung yaitu data citra dari jeruk siam kintamani dengan format gambar serta data dari input atau masukan pengguna aplikasi, serta data hasil kuesioner calon pengguna.

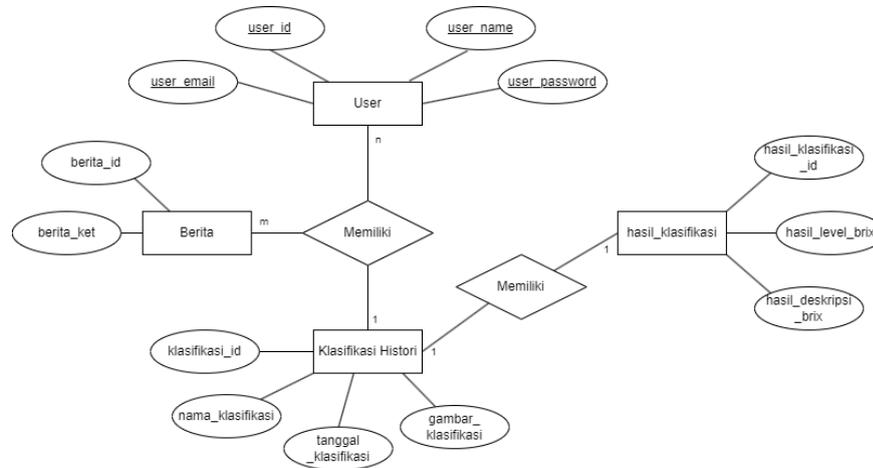
2.3 Desain Perancangan Sistem

Merupakan tahapan dalam perancangan sistem, dengan memberikan gambaran mengenai alur jalan dari sebuah aplikasi, ditunjukkan untuk pengguna. Dengan penerapan metode *User Centered Design*, diharapkan dapat mencapai sebuah *Product design solution*, dirancang sesuai dengan kebutuhan pengguna. Pada tahap ini perancangan diawali dengan rancangan *ERD (Entity Relationship Diagram)*, kemudian *use case diagram*, digunakan untuk menggambarkan interaksi yang dapat dilakukan pengguna dalam alur sebuah aplikasi.

2.3.1 *Entity Relationship Diagram*

Dalam perancangan *ERD* ini, memiliki beberapa field utama, seperti *user*, kemudian informasi jeruk. Dimana dalam field *user* memiliki beberapa data, kemudian dalam informasi jeruk, memiliki field lain seperti data jeruk, kemudian berita, kemudian history hasil klasifikasi, dimana klasifikasi rasa dilakukan didalam data jeruk, serta pada aplikasi ini, memiliki desain *ERD* seperti pada Gambar 2 berikut.

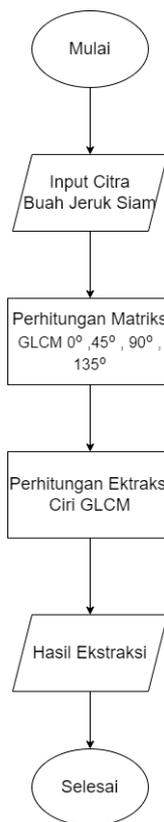
Pembangunan Aplikasi Pengenalan Kualitas Rasa Buah Jeruk Siam Kintamani Menggunakan
GLCM Dan K-NN Berbasis *User Centered Design* (UCD)



Gambar 2. ERD (*Entity Relationship Diagram*)

2.3.2 Alur Proses GLCM

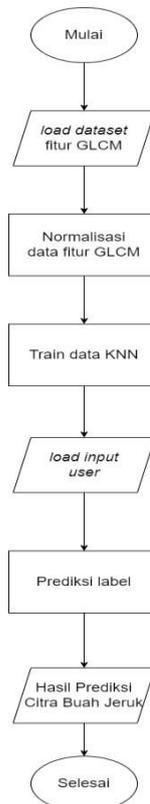
Pada tahapan ini, dilakukan perancangan perhitungan ciri GLCM, berdasarkan citra *grayscale*. Pertama akan dilakukan perhitungan matriks awal pada arah tertentu, seperti 0° , 45° , 90° , 135° , kemudian pembentukan matriks GLCM secara simetris dilakukan dengan melakukan penjumlahan antara matriks awal dengan *transpose* [2]. Setelah mendapatkan hasil matriks yang simetris, selanjutnya dilakukan normalisasi elemen matriks bernilai 1, dengan membagi tiap elemen atau nilai matriks dengan total keseluruhan pixel. Kemudian dilakukan perhitungan ciri (ekstraksi GLCM "*Contrast*", "*Energy*", "*Homogeneity*", "*Correlation*", "*Entropy*", seperti dijelaskan pada Gambar 3 berikut



Gambar 3. *Flowchart* alur proses GLCM

2.3.3 Alur Proses K-NN

Tahapan ini melakukan perancangan alur proses dari K – Nearest Neighbor, dengan mencari jarak terdekat untuk menentukan kelas atau label [6]. Dimana akan dimulai dengan melakukan load dataset fitur dari data gambar (citra) hasil dari ekstraksi fitur GLCM, kemudian akan dilakukan normalisasi data tersebut menggunakan fungsi normalisasi `scaler.fit_transform`, dan hasil dari normalisasi data tersebut yang akan digunakan untuk latih data (train) dimana selanjutnya melakukan load input user dan lakukan normalisasi seperti data sebelumnya. Setelah masing – masing data sudah dilakukan normalisasi, langkah selanjutnya dilakukan prediksi label kelas dari input user, menggunakan `KNN predict`, sehingga akan mendapatkan hasil dari prediksi rasa buah jeruk berupa rasa dan label, dan dijelaskan pada Gambar 4.

Gambar 4. *Flowchart* alur proses K-NN

2.4 Skenario Pengujian

Merupakan tahapan untuk pengujian sistem, dimana bertujuan untuk mengukur seberapa akurat aplikasi, baik dalam pengukuran kegunaan pengguna (*user*) ataupun pada algoritma yang digunakan, pada aplikasi ini yaitu akurasi penggunaan algoritma *gray level co-occurrence matrix* (GLCM) dan *K-nearest neighbor* (KNN). Berikut beberapa pengujian yang digunakan :

2.4.1 *System Usability Scale* (SUS)

System Usability Scale (SUS) merupakan *Usability testing* untuk mengevaluasi seberapa mudah pengguna menggunakan aplikasi atau produk untuk mencapai tujuan [3]. Ini membantu menilai efektivitas, efisiensi, dan kepuasan pengguna. Kuisisioner SUS, atau *System Usability Scale*, digunakan untuk penilaian usability secara terstruktur dan dapat diukur secara kuantitatif, dengan nilai rata - rata diatas 80% sudah memasuki grade "B" [5].

2.4.2 Pengujian Model

Accuracy, Precision, Recall, dan *F-1 Score* adalah metrik evaluasi yang umum digunakan dalam pemodelan klasifikasi untuk mengukur kinerja suatu model. Berikut merupakan pengujian model yang digunakan pada penelitian ini :

1. *Accuracy*

Akurasi mengukur seberapa sering model mengklasifikasikan data dengan benar secara keseluruhan. Ini dihitung dengan membagi jumlah prediksi yang benar dengan total jumlah data yang diamati. Akurasi

memberikan gambaran umum tentang kinerja model, tetapi dapat menjadi bias jika data tidak seimbang.

2. *Precision*

Presisi mengukur proporsi positif yang benar dari semua hasil yang diklasifikasikan sebagai positif oleh model. Dalam kata lain, ini adalah tingkat ketepatan dari prediksi positif. Presisi dihitung dengan membagi jumlah prediksi positif yang benar dengan jumlah total prediksi positif.

3. *Recall*

Recall mengukur proporsi positif yang benar dari semua kasus yang sebenarnya positif dalam data. Ini memberi gambaran tentang kemampuan model untuk menemukan semua instance dari kelas yang positif. Recall dihitung dengan membagi jumlah prediksi positif yang benar dengan jumlah total instance yang sebenarnya positif.

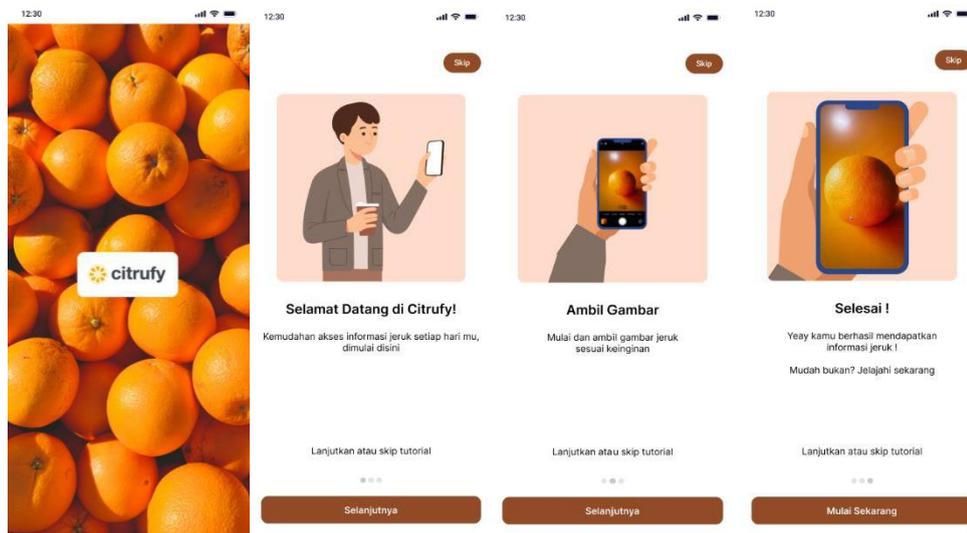
4. *F-1 Score*

F-1 Score adalah rata-rata harmonik dari presisi dan recall. Ini memberikan keseimbangan antara kedua metrik ini. F-1 Score berguna ketika ada ketidakseimbangan kelas dalam data.

3. Hasil dan Pembahasan

3.1 Implementasi Antar Muka

Pada halaman *Splash Screen* adalah halaman yang muncul saat awal aplikasi *Cirtufy* ini dibuka, kemudian akan dilanjutkan secara otomatis menuju halaman *On Boarding*, dimana akan terdapat beberapa konten seperti tombol untuk interaksi dengan aplikasi, kemudian *typography* mengenai informasi mengenai aplikasi.

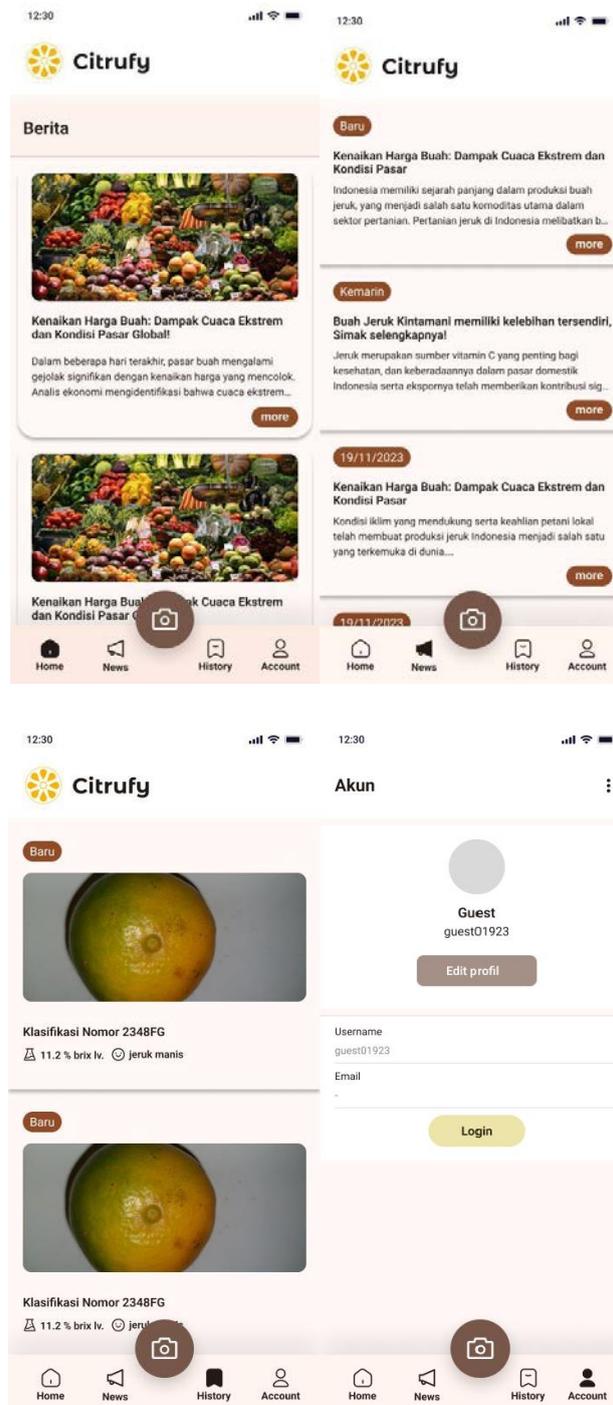


Gambar 5. Halaman *SplashScreen*

Halaman ini adalah halaman yang muncul setelah kamu membuka aplikasi *Cirtufy* dan belum masuk ke akun atau mendaftar. Di sini, kamu bisa melihat beberapa fitur seperti Halaman Awal, Berita, Riwayat, dan Profil Pengguna. Kamu bisa pindah ke

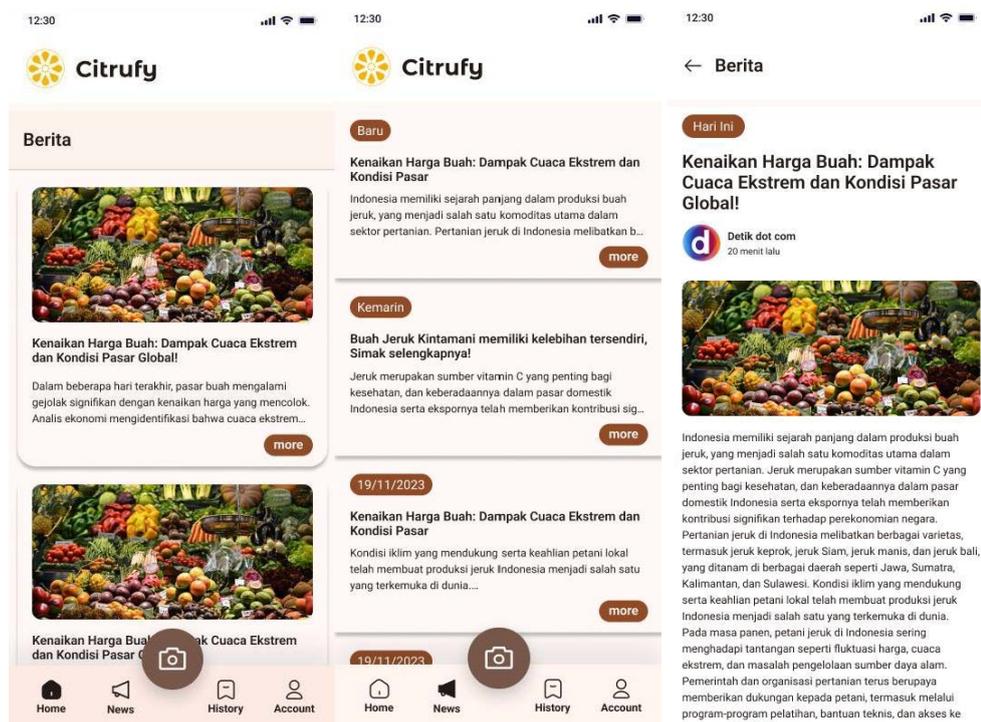
Pembangunan Aplikasi Pengenalan Kualitas Rasa Buah Jeruk Siam Kintamani Menggunakan GLCM dan K-NN Berbasis *User Centered Design* (UCD)

Pembangunan Aplikasi Pengenalan Kualitas Rasa Buah Jeruk Siam Kintamani Menggunakan GLCM Dan K-NN Berbasis *User Centered Design* (UCD) halaman-halaman tersebut dengan menekan tombol-tombol di bawah layar yang memiliki gambar dan judulnya, misalnya tombol rumah untuk Halaman Awal, tombol pengeras suara untuk Berita, tombol bookmark/draft untuk Riwayat, dan tombol user untuk Profil Pengguna. Ada juga tombol khusus untuk Halaman Klasifikasi di bagian atas layar dengan gambar kamera, seperti Gambar 5 berikut.



Gambar 6. Halaman *NewsScreen*, *HistoryScreen*, *Account*

Gambar di bawah ini adalah Halaman Home dalam aplikasi Citruffy, yang menampilkan informasi tentang Berita terkini dari portal berita News API. Informasi yang ditampilkan akan mirip dengan tampilan pada Halaman News. Informasi ini disajikan menggunakan komponen interaktif seperti Card, yang mencakup judul, deskripsi berita, dan tombol "more" untuk melihat detail berita. Selanjutnya, ini adalah tampilan High-Fidelity dari Halaman News, di mana pengguna dapat beralih ke halaman berita menggunakan tombol dengan ikon pengeras suara. Halaman ini memiliki komponen interaktif yang sama dengan Halaman Home, termasuk Card untuk memisahkan informasi berita dan tombol "more" untuk detail berita. Pada Halaman Detail Berita, komponen yang mirip dengan Card digunakan, dengan tambahan informasi tentang penerbit berita.

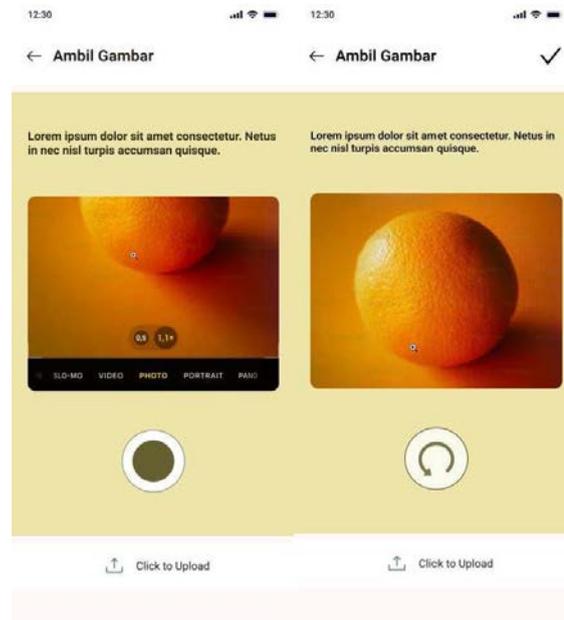


Gambar 7. Halaman NewsScreen

Berikut adalah tampilan High-Fidelity dari Halaman Riwayat, di mana pengguna dapat beralih ke halaman riwayat menggunakan tombol dengan ikon bookmark/draft. Selanjutnya, pada Halaman Detail Berita, memiliki komponen mirip dengan Card pada Halaman Riwayat dan juga News. Terdapat perbedaan informasi tentang level kadar gula (brix level) dan deskripsi mengenai rasa buah tersebut. Di dalam Halaman Klasifikasi, pengguna dapat melihat informasi tentang cara melakukan klasifikasi. Ada petunjuk untuk mengambil gambar dengan objek yang sesuai dengan lingkaran pada preview kamera. Kemudian, ada tampilan preview dari kamera yang akan digunakan untuk mengambil gambar.

Di bagian bawah layar, terdapat tombol untuk menangkap gambar, dan tombol unggah gambar jika ingin menggunakan gambar yang sudah ada di perangkat. Jika pengguna memilih untuk menangkap gambar, setelah menekan tombol, tombol tersebut akan berubah menjadi tombol dengan ikon "memutar" atau ulang, untuk mengulangi proses pengambilan gambar yang diinginkan. Setelah gambar atau foto

Pembangunan Aplikasi Pengenalan Kualitas Rasa Buah Jeruk Siam Kintamani Menggunakan
GLCM Dan K-NN Berbasis User Centered Design (UCD)
yang diinginkan sudah siap, akan muncul tombol dengan ikon centang di pojok kanan
atas. Jika tombol tersebut ditekan, proses klasifikasi akan dilanjutkan, dan komponen
loading akan muncul untuk menunjukkan proses yang sedang berlangsung.



Gambar 7. Halaman Klasifikasi Jeruk

3.2 Pengujian

Pengujian merupakan salah satu aspek penting dalam pengembangan perangkat lunak. Tujuan dari dilakukannya proses pengujian adalah untuk menjamin bahwa perangkat lunak yang dibangun memiliki kualitas yang handal, dimana dengan adanya pengujian dapat menjamin kualitas serta mengetahui kekurangan dari perangkat lunak yang dibangun [4].

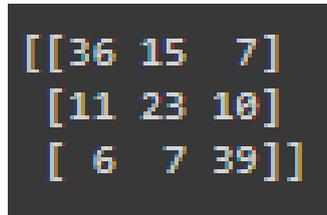
3.1.1 *System Usability Scale*

Total hasil skor masing – masing partisipan akan dijumlah setelah didapatkan, dimana jumlah total skor didapatkan dari hasil normalisasi, pada pernyataan yang memiliki nomor ganjil, digunakan rumus (Skor SUS – 1) sedangkan pada pernyataan yang memiliki nomor genap digunakan rumus (5 - Skor SUS), kemudian masing – masing total skor akan dikalikan 2.5. Setelah total skor pada masing – masing partisipan sudah didapat, total skor di jumlah, kemudian di bagi dengan total partisipan yaitu sejumlah 25 pada penelitian ini. Hasil yang didapat pada pengujian menggunakan kuesioner *System Usability Scale*, pada aplikasi Citruffy ini adalah 86,3%, nilai rata – rata diatas 80% sudah memasuki *grade* “B” pada penilaian *usability testing*, dan menandakan sistem yang dibangun sudah layak untuk digunakan [5].

3.2.1 Pengujian Model

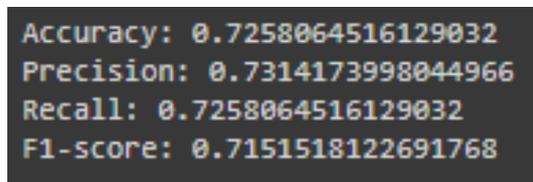
Kemudian dilakukan evaluasi dengan 4 komponen seperti *accuracy*, *precision*, *recall*, *f-1 score*, pada gambar 4.31 memiliki hasil dengan nilai masing – masing seperti *accuracy* sebesar 72%, kemudian *precision* sebesar 73%, lalu *recall* sebesar 72%, dan terakhir yaitu *f1-score* sebesar 71%, dan dapat

dikatakan algoritma dalam penelitian ini memiliki hasil yang cukup baik dalam melakukan klasifikasi pengenalan rasa buah jeruk. Hasil *confusion matrix* dan pengujian *accuracy*, *precision*, *recall*, *f-1 score*. Pada gambar dibawah, matriks representasi 3 x 3, dengan label 0,1,2, yaitu asam, sedang, manis. Kemudian kolom merupakan *predicted* dan baris adalah *true(actual)*, dengan demikian, dimulai dari kanan, yaitu label 0 berhasil diprediksi dengan nilai 36, dari total 58 data *true* (36+15+7), begitu pula seterusnya pada kolom selanjutnya, untuk label 1, dan label 2.



[[36	15	7]
[11	23	10]
[6	7	39]]

Gambar 8. *Confusion Matrix*



```
Accuracy: 0.7258064516129032
Precision: 0.7314173998044966
Recall: 0.7258064516129032
F1-score: 0.7151518122691768
```

Gambar 9. Hasil Pengujian Model

4. Kesimpulan

4.1 Kesimpulan

- a. Penerapan metode user centered design pada aplikasi Citruffy untuk pengenalan rasa buah jeruk berbasis mobile berhasil dikembangkan dan memenuhi 9 kebutuhan pengguna. Tingkat usability aplikasi ini sangat baik, dengan nilai learnability mencapai 80%, efficiency sebesar 83%, dan memoriability serta efficiency diukur melalui System Usability Scale (SUS) mencapai 86,3%, dengan kategori excellent. Tingkat error defective rate sebesar 14%, masih dalam batas wajar. Hasil kuesioner SUS menunjukkan skor 86,3%, dan validitas serta reliabilitas pengujian telah terbukti valid dan reliabel. Dengan demikian, aplikasi ini dapat diterima dengan baik oleh pengguna.
- b. Hasil pengujian dari penggunaan algoritma *Gray Level Co Occurrence Matrix* dan *K – Nearest Neighbors* memiliki nilai sebesar 72% untuk *accuracy*, kemudian 73% untuk *precision*, lalu 72% untuk *recall*, serta 72% untuk *f-1 score*, dan dapat dikatakan kedua algoritma ini sudah berjalan dengan baik pada aplikasi pengenalan rasa buah jeruk.

4.2 Saran

Saran yang dapat diberikan penulis untuk penelitian berikutnya adalah Penerapan fitur cari berita yang dapat memudahkan pengguna untuk mencari berita atau *News* yang di inginkan, serta penambahan fitur ubah *password* untuk membantu serta memudahkan pengguna dalam aplikasi Citruffy. Peningkatan hasil dari pengujian algoritma GLCM serta KNN, dalam usaha meningkatkan kinerja aplikasi serta hasil yang diberikan dapat membantu pengguna dalam melakukan pengenalan rasa buah jeruk.

Daftar Pustaka

- [1] Abdullah, D. (2017). *MERANCANG APLIKASI PERPUSTAKAAN MENGGUNAKAN SDLC*.
- [2] Widodo, R., Widodo, A. W., & Supriyanto, A. (2018). Pemanfaatan Ciri Gray Level CoOccurrence Matrix (GLCM) Citra Buah Jeruk Keprok (*Citrus reticulata* Blanco) untuk Klasifikasi Mutu (Vol. 2, Issue 11). <http://j-ptiik.ub.ac.id>
- [3] Albani. L., & Lombardi. G. (2011). *User Centred Design for EASYRaEACH*.
- [4] Nielsen, J. (2012, January 3). Usability 101: Introduction to Usability. Nielsen Norman Group
- [5] Rizaldy Raka. (2021). PENGEMBANGAN FRONT-END SISTEM INFORMASI PENDATAAN SEKOLAH DESA PENDAR FOUNDATION YOGYAKARTA.
- [6] Fadhlul Barkah, M. (2020). *KLASIFIKASI RASA BUAH JERUK PONTIANAK BERDASARKAN WARNA KULIT BUAH JERUK MENGGUNAKAN METODE K-NEAREST NEIGHBOR* (Vol. 08, Issue 01).
- [7] Iqbal, M., Marthasari, G. I., & Nuryasin, I. (2020). Penerapan Metode UCD (User Centered Design) pada Perancangan Aplikasi Darurat Berbasis Android. *REPOSITOR*, 2(8), 1041–1054.

Pengembangan Sistem Presensi Anti Spoofing dengan Metode Support Vector Machine

Marselinus Putu Harry Setyawan^{a1}, I Dewa Made Bayu Atmaja Darmawan^{a2}, I Gede Santi Astawa^{a3},
I Wayan Santiyasa^{a4}

^aProgram Studi Informatika, Universitas Udayana
Badung, Bali, Indonesia

¹marselinusphs@gmail.com

²dewabayu@unud.ac.id

³santi.astawa@unud.ac.id

⁴santiyasa@unud.ac.id

Abstrak

Penelitian ini bertujuan untuk mengembangkan sistem deteksi spoofing menggunakan metode Support Vector Machine (SVM) dengan ekstraksi fitur tekstur Local Binary Patterns (LBP) dan Gray-Level Co-occurrence Matrix (GLCM). Sistem ini ditujukan untuk mengatasi tantangan kehadiran mahasiswa dengan mengintegrasikan teknologi pengenalan wajah dalam manajemen presensi. Spoofing, yang merupakan usaha pemalsuan wajah, menjadi tantangan dalam sistem keamanan berbasis wajah. Oleh karena itu, sistem ini memfokuskan pada deteksi spoofing dengan membandingkan pola tekstur antara wajah asli dan palsu. Model mampu mengidentifikasi upaya pemalsuan wajah (spoofing) dengan tingkat akurasi sebesar 94% setelah melakukan tuning parameter C dan gamma. Selanjutnya, sistem presensi anti spoofing diuji melalui black box testing dan memberikan hasil yang sesuai dengan harapan. Sistem ini mampu memulai kelas, mencatat kehadiran mahasiswa, serta menghasilkan laporan kehadiran yang valid. Seluruh fungsi sistem telah diuji secara menyeluruh dan memperoleh tingkat akurasi 95% dalam pendeteksian spoofing. Dengan demikian, penelitian ini menghadirkan kontribusi penting dalam pengembangan sistem presensi yang lebih akurat dan aman melalui deteksi spoofing.

Kata kunci: spoofing, attendance system, local binary patterns, gray level co-occurrence matrix, support vector machine, black box testing

1. Pendahuluan

Untuk memverifikasi catatan kehadiran mahasiswa, kampus harus memiliki sistem yang tepat untuk menyetujui dan memelihara catatan kehadiran secara konsisten. Sistem pencatatan kehadiran mahasiswa umumnya terbagi menjadi Manual Attendance System (MAS) dan Automated Attendance System (AAS). Sistem manual melibatkan dosen memanggil mahasiswa satu per satu atau meminta tanda tangan mahasiswa, yang bisa menjadi sulit dan memakan waktu, terutama di kelas besar. Sebaliknya, sistem absensi otomatis menggunakan pengenalan wajah dapat mengurangi beban administrasi staf kampus [1].

Sistem pengenalan wajah memiliki kerentanan karena tidak membedakan antara gambar wajah nyata dan wajah palsu [2]. Tantangan utama adalah mendeteksi spoofing citra wajah, di mana seseorang mencoba mengakses sistem dengan memalsukan wajah korban untuk mendapatkan akses ilegal. Jenis spoofing yang paling sering digunakan yaitu spoofing cetak dan spoofing presentasi. Spoofing cetak adalah pelaku menggunakan foto korban dengan gambar dicetak, sedangkan spoofing presentasi adalah cara yang lebih canggih yang biasanya menggunakan foto wajah korban dari perangkat elektronik [3]. Deteksi spoofing dapat dilakukan dengan pendekatan fitur tekstur yang membandingkan pola tekstur antara wajah asli dan palsu, memanfaatkan perbedaan pantulan dan permukaan [4].

Penelitian deteksi spoofing wajah telah menggunakan analisis tekstur untuk meningkatkan keamanan sistem pengenalan wajah. Penelitian oleh Kusuma dan Kartika yang menggunakan LBP dan Adaptive

Histogram Equalization dengan KNN [4], serta oleh Alexander *et al* yang menggunakan ekstraksi fitur Local Binary Patterns dalam pengenalan wajah [5]. Penelitian tersebut menunjukkan bahwa kombinasi LBP dan GLCM memberikan akurasi yang lebih tinggi dibandingkan dengan penggunaan algoritma tunggal.

Penelitian ini bertujuan membangun sistem presensi yang dapat mengidentifikasi spoofing wajah menggunakan metode LBP dan GLCM untuk ekstraksi fitur tekstur dan SVM untuk klasifikasi. Metode HOG digunakan untuk deteksi wajah, yang efektif berdasarkan penelitian sebelumnya. Diharapkan sistem ini dapat membantu dalam proses otentikasi untuk mencegah spoofing, sehingga mengurangi akses ilegal.

2. Metode Penelitian

Penelitian ini dimulai dengan mengumpulkan data citra spoofing untuk digunakan dalam model dan melakukan preprocessing pada citra. Kemudian dilakukan analisis kebutuhan dan desain dari sistem yang dibuat berupa UML. Setelah itu membuat model anti-spoofing dan mengimplementasikan desain sistem yang sudah dibuat menjadi sebuah sistem presensi. Terakhir dilakukan pengujian terhadap sistem dan menganalisis keluaran yang dihasilkan dari sistem yang sudah dibuat.

2.1. Pengumpulan Data

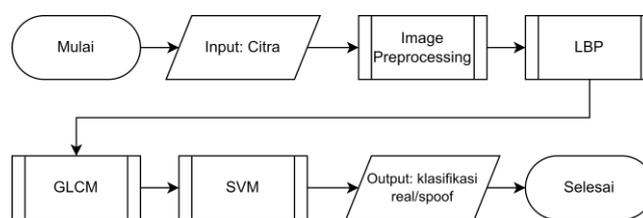
CASIA-FASD adalah dataset anti-spoofing wajah yang terdiri dari 50 subjek. Masing-masing subjek memiliki sebuah video wajah asli, spoofing cetak, dan spoofing presentasi menggunakan smartphone. Video direkam menggunakan berbagai jenis kamera: kualitas rendah, kualitas sedang, dan kualitas tinggi. Seluruh video tersimpan dalam format mp4. Video-video tersebut kemudian diubah menjadi serangkaian gambar menggunakan teknik framing. Proses framing ini mengambil setiap frame dari video dan menyimpannya sebagai file gambar dengan format jpg. Contoh gambar pada dataset dapat dilihat pada Gambar 1.



Gambar 1. Contoh Gambar Dataset CASIA-FASD

2.2. Face Spoofing Detection

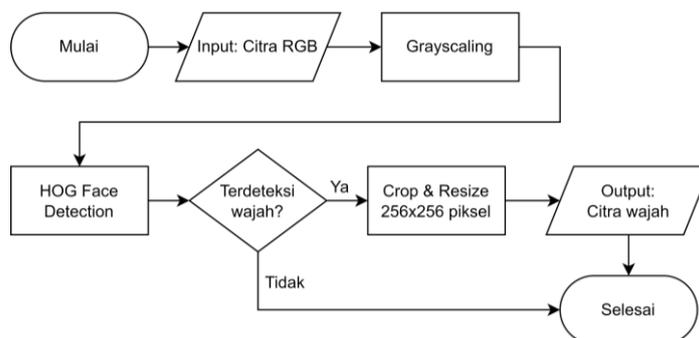
Face spoofing detection adalah tahapan untuk mengidentifikasi apakah wajah yang terdeteksi merupakan wajah asli atau spoof. Masukan pada tahap ini adalah citra wajah. Pada tahap ini, pertama data citra akan mendeteksi wajah menggunakan HOG. Kemudian crop area wajah yang terdeteksi menggunakan HOG. Kemudian dilakukan resizing. Resizing dilakukan bertujuan agar seluruh citra memiliki ukuran yang sama, yakni 256x256 piksel. Selanjutnya, data citra akan dilakukan ekstraksi fitur tekstur LBP dan GLCM. Kemudian untuk pengklasifikasiannya menggunakan metode SVM. Adapun parameter yang dilakukan tuning pada SVM yaitu C dan gamma. Keluaran dari tahap face spoofing detection adalah hasil klasifikasi spoofing. Flowchart tahap image preprocessing ditunjukkan pada Gambar 2.



Gambar 2. Flowchart Face Spoofing Detection

2.2.1.1. Image Preprocessing

Image preprocessing adalah serangkaian teknik yang digunakan untuk mempersiapkan dan memperbaiki kualitas gambar sebelum digunakan dalam analisis atau pemrosesan lebih lanjut. Tujuannya adalah agar gambar lebih mudah dianalisis dan memberikan hasil yang lebih baik. Masukan pada tahap ini adalah data citra berwarna. Pada tahap ini, citra dilakukan grayscale, yaitu mengubah piksel citra berwarna menjadi citra yang memiliki tingkat warna abu-abu. Kemudian dari citra grayscale tersebut akan dideteksi gambar wajah menggunakan metode Histogram of Oriented Gradients (HOG). Dari wajah yang terdeteksi dilakukan crop pada area wajah. Flowchart tahap image preprocessing ditunjukkan pada Gambar 3.



Gambar 3. Flowchart Image Preprocessing

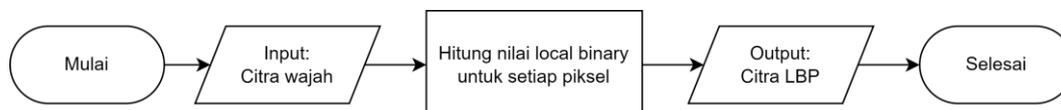
Contoh hasil output pada tahap Image Preprocessing ditunjukkan pada Gambar 4.



Gambar 4. Contoh Hasil Output Image Preprocessing

2.2.1.2. Local Binary Patterns

Local Binary Patterns (LBP) adalah proses ekstraksi fitur pada citra digital yang digunakan dalam pengolahan citra dan pengenalan wajah. Masukan pada tahap ini adalah data citra yang telah dipreprocessing sebelumnya. Proses yang dilakukan pada algoritma LBP adalah dengan mengambil sebuah piksel di citra grayscale, kemudian membandingkan nilai intensitas piksel tersebut dengan nilai intensitas piksel tetangganya. Jika nilai intensitas piksel tetangganya lebih besar dari piksel pusatnya, maka dihasilkan angka 1 pada bit biner. Sebaliknya, jika nilai intensitas piksel tetangganya lebih kecil atau sama dengan piksel pusatnya, maka dihasilkan angka 0 pada bit biner. Dari sini, sebuah pola biner terbentuk pada setiap piksel pada citra. Kemudian bilangan biner tersebut diubah ke bentuk desimal, dan nilai desimal tersebut menjadi nilai baru pada piksel pusat. Keluaran dari tahap ini adalah citra hasil LBP. Flowchart tahapan LBP ditunjukkan pada Gambar 5.



Gambar 5. Flowchart Local Binary Patterns

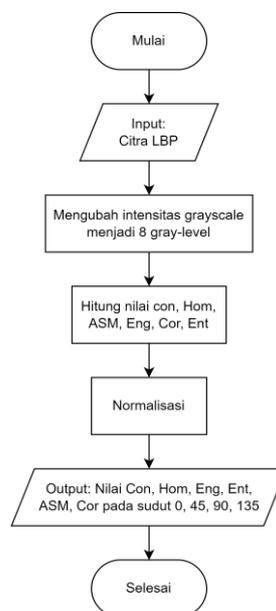
Contoh hasil output pada tahap Local Binary Patterns ditunjukkan pada Gambar 6.



Gambar 6. Contoh Image Hasil Output Local Binary Patterns

2.2.1.3. Gray-Level Co-occurrence Matrix

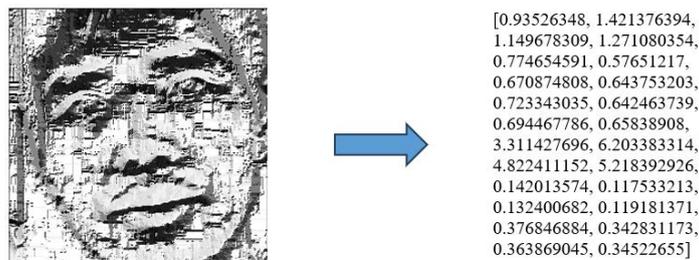
Gray-Level Co-occurrence Matrix (GLCM) adalah sebuah metode untuk mengukur tekstur citra dengan menghitung kejadian kemunculan pasangan piksel dengan intensitas tertentu dalam suatu jendela atau area citra. Flowchart tahapan GLCM ditunjukkan pada Gambar 7.



Gambar 7. Flowchart Gray-Level Co-occurrence Matrix

Masukan dari tahap ini adalah citra hasil LBP. Dalam proses pembentukan matriks GLCM, setiap piksel pada citra dianalisis dengan melihat tetangganya. Setiap pasangan tetangga akan dicatat jarak antar piksel dan arah yang dihitung. Selanjutnya, matriks GLCM dibangun dengan menghitung kemunculan relatif dari pasangan piksel dengan level keabuan tertentu. Matriks GLCM kemudian digunakan untuk mengekstraksi fitur-fitur citra, seperti kontras, korelasi, energi, dan homogenitas. Keluaran dari tahap ini adalah nilai kontras, korelasi, energi, dan homogenitas pada sudut 0, 45, 90, 135 pada citra.

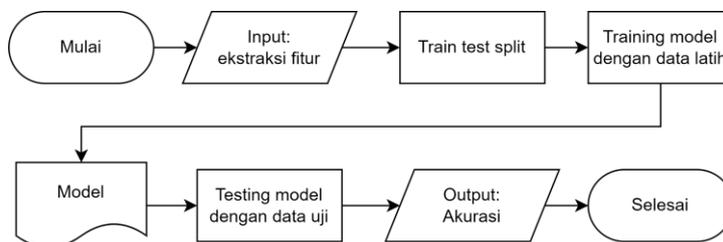
Contoh hasil output pada tahap GLCM ditunjukkan pada Gambar 8.



Gambar 8. Contoh Input dan Output pada Proses Co-occurrence Matrix

2.2.1.4. Support Vector Machine

Support Vector Machine (SVM) adalah algoritma machine learning yang digunakan untuk klasifikasi dan regresi. SVM mencari garis/hyperplane terbaik yang dapat memisahkan data ke dalam kelas-kelas yang berbeda dengan margin maksimal. Masukan pada tahap ini menggunakan data ekstraksi fitur tekstur GLCM. Prosesnya dimulai dengan normalisasi data yang bertujuan untuk mengubah nilai-nilai data menjadi rentang yang serupa. Kemudian, data dibagi menjadi data latih dan data uji dengan perbandingan 80:20. Selanjutnya, data latih digunakan untuk melatih model SVM dan menghasilkan model SVM yang digunakan untuk mengklasifikasikan data uji. Terakhir, dilakukan evaluasi terhadap model SVM dengan menghitung akurasi klasifikasi yang dihasilkan. Keluaran dari tahap ini adalah akurasi dari model. Flowchart tahapan SVM ditunjukkan pada Gambar 9.



Gambar 9. Flowchart Support Vector Machine

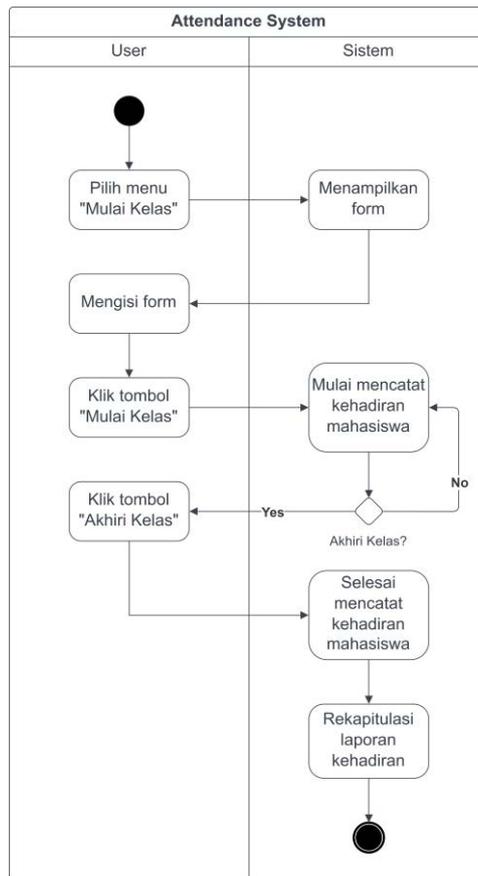
2.3. Sistem Presensi

Sebelum mulai mengembangkan sistem, perlu dilakukan proses analisis kebutuhan pengguna, perancangan sistem, serta pemilihan teknologi yang tepat untuk mengembangkan sistem yang diinginkan. Analisis dapat diterjemahkan dalam bentuk diagram UML untuk memvisualisasikan desain sistem, desain database untuk memodelkan struktur data yang dibutuhkan.

2.3.1.1. Activity Diagram

Activity diagram digunakan untuk merepresentasikan aktivitas atau perilaku suatu sistem, proses bisnis, atau fungsi yang akan dijalankan dalam sebuah sistem. Pada penelitian ini, alur sistem presensi dimulai ketika user memilih menu 'mulai kelas', yang merupakan titik awal dari seluruh proses. Sistem kemudian merespon dengan menampilkan form memulai kelas yang memerlukan input kode mata kuliah. Setelah user mengisi dan mengirimkan form tersebut, sistem melakukan proses training face recognition terhadap mahasiswa yang telah melakukan enroll pada mata kuliah tersebut. Proses ini memastikan bahwa sistem dapat mengenali wajah-wajah mahasiswa dengan akurat.

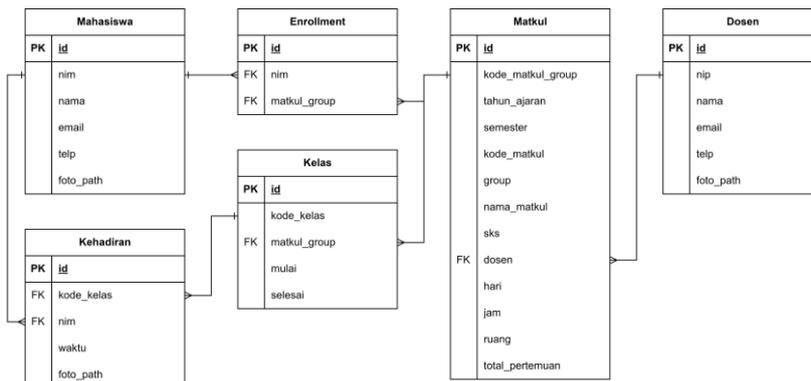
Selanjutnya, sistem secara otomatis mulai merekam kehadiran mahasiswa menggunakan teknologi pengenalan wajah. Proses perekaman ini berlangsung secara terus-menerus hingga dosen memutuskan untuk mengakhiri kelas melalui sistem. Pada tahap akhir, setelah kelas diakhiri, dosen memiliki kemampuan untuk melihat hasil rekapitulasi kehadiran mahasiswa untuk perkuliahan tersebut. Diagram use case yang menggambarkan seluruh proses ini secara lebih rinci dapat dilihat pada Gambar 10. Diagram tersebut memberikan gambaran visual yang membantu dalam memahami bagaimana setiap langkah dalam sistem presensi saling berkaitan dan berinteraksi satu sama lain, sehingga menciptakan sebuah sistem yang efisien dan efektif dalam mencatat kehadiran mahasiswa.



Gambar 10. Activity Diagram

2.3.1.2. Desain Database

Desain database digunakan untuk merancang struktur database untuk memenuhi kebutuhan aplikasi atau sistem informasi. Pada sistem presensi dalam penelitian ini, terdapat enam buah tabel yang masing-masing memiliki peran spesifik untuk memastikan data tersimpan dan dikelola dengan baik. Tabel-tabel tersebut meliputi tabel mahasiswa, dosen, matkul, enrollment, kelas, dan kehadiran. Setiap tabel memiliki atribut-atribut yang saling berhubungan satu sama lain untuk mendukung fungsionalitas sistem. Misalnya, tabel mahasiswa menyimpan data pribadi mahasiswa, tabel dosen menyimpan informasi mengenai dosen, dan tabel matkul menyimpan detail mata kuliah yang diajarkan. Selain itu, tabel enrollment berfungsi untuk mencatat pendaftaran mahasiswa ke dalam mata kuliah tertentu, tabel kelas mencatat informasi mengenai kelas yang diadakan, dan tabel kehadiran mencatat kehadiran mahasiswa dalam setiap sesi perkuliahan. Desain database sistem presensi ini digambarkan secara visual pada Gambar 11, yang memberikan representasi grafis tentang bagaimana tabel-tabel tersebut diatur dan dihubungkan satu sama lain untuk membentuk sebuah sistem yang terpadu dan efisien.



Gambar 11. Desain Database

2.4. Pengujian Sistem

Pengujian sistem adalah metode untuk memeriksa apakah produk perangkat lunak yang sebenarnya sesuai dengan persyaratan yang diharapkan dan memastikan bahwa produk perangkat lunak bebas dari cacat. Pada tahapan pengujian dibagi menjadi dua, yaitu functional testing dan non functional testing.

2.4.1. Functional Testing

Pada tahap ini dilakukan pengujian functional untuk memverifikasi bahwa setiap fitur aplikasi berfungsi sesuai dengan persyaratan yang telah ditentukan. Pada penelitian ini, functional testing menggunakan metode Black Box Testing.

2.4.2. Non-Functional Testing

Pada tahap ini dilakukan pengujian non-functional untuk memverifikasi pada kebutuhan non-fungsional seperti performa dan kegunaan sebuah aplikasi. Pada penelitian ini, non functional testing menggunakan akurasi deteksi spoofing sebagai penilaiannya. Pengujian dilakukan dengan beberapa skenario uji, di antaranya membandingkan ekstraksi fitur LBP, GLCM, dan LBP-GLCM untuk mengetahui pengaruh penggunaan ekstraksi fitur terhadap akurasi deteksi spoofing dan pengenalan wajah. Untuk menghitung akurasi dapat dilakukan dengan cara membagi jumlah prediksi yang benar dengan jumlah percobaan yang dilakukan. Persamaan menghitung akurasi dapat dilihat pada rumus berikut.

$$Akurasi = \frac{TP+TN}{TP+FN+FP+TN} \times 100\% \quad (1)$$

Keterangan:

- TP = True Positive (sistem benar mengidentifikasi wajah spoofing)
- TN = True Negative (sistem benar mengidentifikasi wajah asli)
- FP = False Positive (sistem keliru mengidentifikasi wajah asli sebagai spoofing)
- FN = False Negative (sistem keliru mengidentifikasi wajah spoofing sebagai wajah asli)

Penelitian ini menggunakan beberapa skenario uji untuk menguji keberhasilan deteksi spoofing. Pengujian dilakukan pada beberapa skenario uji, seperti penggunaan atribut wajah kacamata, rotasi citra wajah, dan pencahayaan indoor/outdoor. Hal ini bertujuan untuk mengetahui sejauh mana kemampuan algoritma dalam mengenali wajah yang memiliki kondisi yang berbeda-beda, sehingga dapat diketahui kehandalan dan kelemahan algoritma dalam berbagai situasi. Hasil pengujian pada skenario uji tersebut kemudian dievaluasi dan dibandingkan dengan hasil pengujian pada skenario uji lainnya untuk mendapatkan pemahaman yang lebih lengkap mengenai kinerja algoritma deteksi spoofing.

3. Hasil dan Pembahasan

3.1. Pengujian Spoofing Detection

Dalam pengujian model deteksi spoofing, dilakukan beberapa skenario pengujian untuk mengevaluasi performa dan karakteristik model. Skenario pengujian meliputi pengujian terhadap nilai C dan gamma, pengaruh penggunaan atribut wajah kaca mata, pengaruh angle wajah, dan pengaruh pencahayaan indoor/outdoor. Pengujian dilakukan dengan tujuan untuk menentukan kombinasi parameter terbaik, menguji kemampuan model dalam mendeteksi spoofing dengan variasi atribut wajah kaca mata dan angle wajah

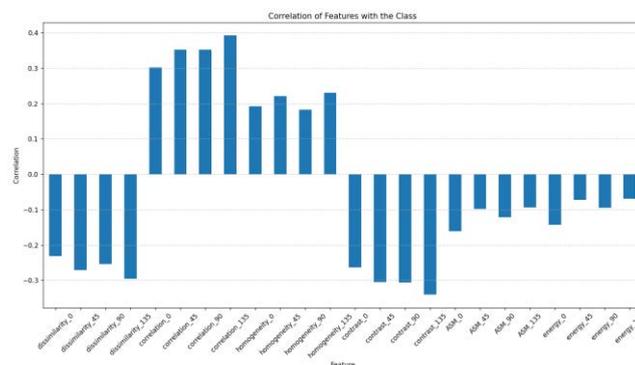
3.1.1. Skenario Pengujian terhadap Nilai C dan Gamma

Dalam skenario pengujian terhadap nilai C dan gamma, dilakukan eksperimen dengan menggunakan beberapa nilai C dan gamma yang berbeda. Rentang nilai C yang diuji meliputi 0.1, 1, 10, 100, 1000, dan 10000, sedangkan untuk gamma diuji dengan nilai 1, 0.1, 0.01, 0.001, dan 0.0001. Hasil pengujian ditunjukkan pada Tabel 1.

Tabel 1. Hasil Akurasi dengan Kombinasi Nilai C dan Gamma

		C					
		0.1	1	10	100	1000	10000
Gamma	1	64%	69%	64%	63%	63%	61%
	0.1	67%	68%	78%	84%	76%	72%
	0.01	63%	66%	66%	84%	94%	86%
	0.001	61%	61%	64%	67%	69%	70%
	0.0001	61%	61%	61%	61%	67%	63%

Setelah melalui pengujian, ditemukan hasil terbaik pada nilai C sebesar 1000 dan gamma sebesar 0.01. Hasil pengujian ini menunjukkan bahwa model SVM dengan konfigurasi parameter tersebut memberikan akurasi sebesar 94% dalam mendeteksi spoofing. Hal ini menunjukkan bahwa kombinasi nilai C dan gamma tersebut mampu memberikan performa terbaik dalam memisahkan kelas-kelas data pada model SVM, sehingga menghasilkan tingkat akurasi yang tinggi dalam deteksi spoofing.



Gambar 12. Korelasi Fitur-Fitur terhadap Kelas Label

Gambar 12 menunjukkan korelasi antara 24 fitur dan 1 kolom kelas. Setiap titik pada grafik mewakili satu fitur, dengan koordinat x menunjukkan nama fitur dan koordinat y menunjukkan nilai koefisien korelasi Pearson antara fitur tersebut dengan kelas. Nilai korelasi dapat berkisar antara -1 hingga 1.

Berdasarkan gambar, terdapat beberapa fitur dengan nilai korelasi tinggi (absolut) terhadap kelas. Fitur-fitur ini kemungkinan besar memiliki pengaruh signifikan terhadap kelas. Contohnya, fitur "dissimilarity_45" memiliki nilai korelasi 0.35, menunjukkan hubungan positif yang cukup kuat dengan kelas. Artinya, semakin tinggi nilai "dissimilarity_45", semakin tinggi kemungkinan data tersebut termasuk dalam kelas tertentu.

Sebaliknya, fitur "homogeneity_90" memiliki nilai korelasi -0.28, menunjukkan hubungan negatif yang cukup kuat dengan kelas. Artinya, semakin tinggi nilai "homogeneity_90", semakin rendah kemungkinan data tersebut termasuk dalam kelas tertentu.

3.1.2. Skenario Pengujian terhadap Pengaruh Penggunaan Atribut Wajah

Pada skenario ini, dilakukan dua kondisi uji coba yang berbeda. Pertama, tanpa atribut kaca mata dan topi, di mana pengujian dilakukan pada wajah asli tanpa perubahan atau penambahan apapun. Kedua, dengan atribut tambahan berupa kaca mata dan topi, di mana pengujian dilakukan untuk menguji kemampuan model dalam mendeteksi spoofing pada wajah yang telah dimodifikasi dengan atribut tersebut.

Tabel 2. Hasil Pengujian Model Terhadap Pengaruh Penggunaan Atribut Wajah

Skenario	Akurasi
Tanpa Atribut	90%
Dengan Atribut	90%

Pada Tabel 2 menunjukkan bahwa model deteksi spoofing menggunakan metode SVM dengan ekstraksi fitur LBP dan GLCM memiliki tingkat keberhasilan yang sangat baik. Model ini mencapai akurasi sebesar 90% dalam mendeteksi spoofing pada wajah tanpa atribut seperti kacamata dan topi, serta akurasi sebesar 85% dalam mendeteksi spoofing pada wajah dengan atribut tersebut. Hasil ini mengindikasikan bahwa metode tersebut memiliki potensi untuk menjadi solusi yang efektif dalam mendeteksi spoofing pada berbagai kondisi wajah, baik dengan maupun tanpa atribut.

3.1.3. Skenario Pengujian terhadap Pengaruh Angle Wajah

Dalam skenario pengujian terhadap pengaruh angle wajah, dilakukan dua kondisi uji coba yang berbeda. Pertama, kondisi angle wajah lurus di mana pengujian dilakukan menggunakan citra atau data dengan angle wajah tegak lurus terhadap kamera. Kedua, kondisi angle wajah miring, di mana variasi angle wajah dilakukan hingga lebih dari 45 derajat.

Tabel 3. Hasil Pengujian Model Terhadap Pengaruh Angle Wajah

Skenario	Akurasi
Lurus	90%
Miring	90%

Hasil pengujian menunjukkan bahwa model deteksi spoofing menggunakan metode SVM dengan ekstraksi fitur LBP dan GLCM sangat handal dalam mendeteksi spoofing pada berbagai angle wajah. Model ini mencapai akurasi 100% dalam mendeteksi spoofing pada wajah dengan angle lurus dan 90% pada wajah dengan angle miring. Tingkat keberhasilan yang tinggi ini menandakan bahwa metode tersebut efektif dalam berbagai kondisi angle wajah.

3.1.4. Skenario Pengujian terhadap Pengaruh Pencahayaan Indoor/Outdoor

Dalam skenario pengujian pengaruh pencahayaan indoor dan outdoor terhadap model deteksi spoofing, dua kondisi uji coba dilakukan. Pertama, pada kondisi pencahayaan indoor di mana citra atau data diambil di dalam ruangan dengan pencahayaan yang terkendali. Pada kondisi ini, tingkat noise cenderung rendah dengan Signal-to-Noise Ratio (SNR) yang tinggi, umumnya di atas 20 dB. Kedua, pada kondisi pencahayaan outdoor di mana citra atau data diambil di luar ruangan dengan variasi pencahayaan yang lebih bervariasi dan terpengaruh oleh faktor-faktor eksternal seperti sinar matahari, bayangan, dan gangguan lainnya. Pada kondisi outdoor, tingkat noise cenderung tinggi dengan SNR yang rendah, dapat di bawah 10 dB.

Tabel 4. Hasil Pengujian Model Terhadap Pengaruh Pencahayaan Indoor/Outdoor

Skenario	Akurasi
Indoor	90%
Outdoor	75%

Dalam skenario pengujian terhadap pengaruh pencahayaan indoor/outdoor, hasilnya menunjukkan bahwa model memberikan akurasi sebesar 90% dalam mendeteksi spoofing pada dalam ruangan, namun hanya 75% dalam mendeteksi spoofing pada luar ruangan. Hal ini menunjukkan bahwa model kurang bekerja dengan baik dalam kondisi luar ruangan, mungkin karena perubahan pencahayaan yang lebih kompleks dan variabel di luar ruangan.

3.2. Pengujian Sistem Presensi

Pengujian sistem presensi dilakukan dengan pendekatan Black box Testing, di mana sistem presensi dievaluasi berdasarkan input dan output yang dihasilkan, tanpa memperhatikan detail implementasi internalnya. Pada tahap ini, berbagai skenario pengujian dilakukan dengan memberikan berbagai input ke sistem presensi dan mengamati keluaran yang dihasilkan. Tujuan dari pengujian ini adalah untuk memverifikasi bahwa sistem presensi berfungsi sesuai dengan yang diharapkan, mencakup fungsionalitas seperti memulai kelas, pengelolaan data mahasiswa, dan pelaporan.

Tabel 5. Skenario Pengujian Sistem Presensi

No	Skenario Pengujian	Hasil Diharapkan	Hasil Uji
1	Mulai kelas	Menambahkan kelas ke database, menampilkan halaman kelas	Sesuai harapan
2	List kelas yang sedang berlangsung	Menampilkan list kelas sedang berlangsung	Sesuai harapan
3	Masuk ke kelas	Menuju ke halaman kelas.	Sesuai harapan
4	Presensi menggunakan spoofing	Menampilkan bounding box berwarna merah di sekitar wajah	Sesuai harapan
5	Presensi menggunakan citra wajah namun tidak dikenali sebagai peserta mata kuliah	Menampilkan bounding box berwarna biru di sekitar wajah, dan keterangan "Unknown"	Sesuai harapan
6	Presensi menggunakan citra wajah peserta mata kuliah	Menampilkan bounding box berwarna biru di sekitar wajah, dan keterangan NIM peserta. Input data kehadiran ke database.	Sesuai harapan

Hasil positif dari 6 skenario pengujian ini menegaskan bahwa sistem presensi yang telah dikembangkan mampu beroperasi sesuai dengan tujuan awal dan mampu mengatasi berbagai situasi yang beragam. Keberhasilan ini membuktikan bahwa sistem mampu mendukung kehadiran mahasiswa secara efektif, mengelola data mahasiswa dan dosen dengan akurat, serta menghasilkan laporan yang relevan dan sesuai dengan kebutuhan. Dengan demikian, sistem presensi ini memenuhi standar kualitas dan performa yang diharapkan, serta siap digunakan untuk membantu dalam manajemen presensi mahasiswa secara efisien.

Tabel 6. Hasil Pengujian Model Terhadap Pengaruh Pencahayaan Indoor/Outdoor

Skenario	Akurasi
Lurus	90%
Miring	90%

Hasil pengujian ini menunjukkan bahwa model memberikan akurasi sebesar 95% dalam mendeteksi spoofing ketika diimplementasikan dalam sistem presensi, yang menunjukkan tingkat keberhasilan yang sangat baik dalam kondisi tersebut. Hal ini menandakan bahwa metode SVM dengan ekstraksi fitur LBP dan GLCM baik dalam mendeteksi spoofing pada wajah ketika diimplementasikan dalam sistem presensi.

4. Kesimpulan

Berdasarkan rangkaian penelitian dan pengembangan model deteksi spoofing menggunakan metode Support Vector Machine (SVM) dengan ekstraksi fitur tekstur Local Binary Patterns (LBP) dan Gray-Level Co-occurrence Matrix (GLCM), serta pengembangan sistem presensi, dapat diambil beberapa simpulan penting sebagai berikut:

- Hasil penelitian menunjukkan bahwa deteksi spoofing menggunakan ekstraksi fitur tekstur Local Binary Patterns (LBP) dan Gray-Level Co-occurrence Matrix (GLCM) dengan metode Support Vector Machine (SVM) memberikan hasil akurasi 94%. Dalam penelitian ini, dilakukan tuning parameter C dan gamma. Hasil terbaik diperoleh dengan nilai C sebesar 1000 dan gamma sebesar 0,01. Dengan nilai parameter ini, model SVM berhasil mencapai tingkat akurasi sebesar 94%. Hasil ini menunjukkan bahwa SVM dengan kombinasi fitur LBP dan GLCM dapat efektif dalam mendeteksi spoofing.
- Peneliti melakukan serangkaian percobaan di berbagai kondisi, termasuk indoor-outdoor, tanpa atribut-dengan atribut kaca dan topi, dan angle wajah lurus dan miring. Hasil

percobaan menunjukkan tingkat keberhasilan yang signifikan, dengan akurasi mencapai 90% untuk kondisi tanpa atribut, 85% untuk kondisi dengan atribut, 90% untuk kondisi wajah lurus, dan 90% untuk kondisi wajah miring. Namun, terdapat penurunan akurasi pada kondisi outdoor, dimana akurasi mencapai 75%, yang lebih rendah dibandingkan dengan kondisi indoor. Penurunan ini mungkin disebabkan oleh faktor lingkungan seperti pencahayaan yang lebih bervariasi, gangguan dari sinar matahari, bayangan, dan elemen lainnya yang dapat mempengaruhi kualitas citra wajah.

- c. Setelah melakukan non-functional testing dan functional testing pada sistem presensi anti-spoofing, dapat disimpulkan bahwa hasilnya memenuhi harapan. Sistem ini mampu memulai kelas, mencatat kehadiran mahasiswa, serta menghasilkan laporan kehadiran yang valid. Seluruh fungsi sistem telah diuji secara menyeluruh dan memperoleh tingkat akurasi 95% dalam pendeteksian spoofing.

Referensi

- [1] M. Surve, et al., "Automatic attendance system using face recognition technique," in *International Journal of Recent Technology and Engineering*, vol. 9, no. 1, pp. 2134-2138, Jan. 2020.
- [2] F. Jiang, P. Liu, and X. Zhou, "Multilevel fusing paired visible light and near-infrared spectral images for face anti-spoofing," *Pattern Recognition Letters*, vol. 128, pp. 30-37, 2019.
- [3] N. Daniel and A. Anitha, "Texture and quality analysis for face spoofing detection," *Computers & Electrical Engineering*, vol. 94, p. 107293, 2021.
- [4] I. B. Kusuma and A. Kartika, "Image spoofing detection using local binary pattern," *J. Masy. Inform. Indones.*, vol. 2, no. 1, pp. 49-54, 2017.
- [5] A. D. Alexander, R. Salkiawati, H. Lubis, F. Rahman, H. Herlawati, and R. T. Handayanto, "Local Binary Pattern Histogram for Face Recognition in Student Attendance System," in *2020 3rd International Conference on Computer and Informatics Engineering (IC2IE)*, vol. 152, no. 156, p. 012014, Sept. 2020, IEEE.

This page is intentionally left blank.

Implementasi Metode Human Centered Design Dan Heuristic Evaluations Dalam Improvisasi User Experiences Dan User Interfaces (Studi Kasus : Aplikasi Pendidikan Dan Pelatihan Anggota Koperasi)

I Ketut Santa Wijaya^{a1}, Cokorda Pramatha^{ab}, Luh Gede Astuti^{a3}, I Gede Surya Rahayuda^{a4}

^aProgram Studi Informatika, Universitas Udayana

^bCenter for Interdisciplinary Research on the Humanities and Social Sciences, Universitas Udayana
Jimbaran, Bali, Indonesia

¹iketutsantawijaya028@student.unud.ac.id

²cokorda@unud.ac.id

³lg.astuti@unud.ac.id

⁴igedesuryarahayuda@unud.ac.id

Abstract

Dalam era globalisasi saat ini, peran teknologi informasi menjadi sangat signifikan, terutama bagi Usaha Mikro, Kecil, dan Menengah (UMKM) seperti koperasi. Laporan "Digital 2022 Indonesia" oleh Kemp (2022) menunjukkan bahwa 73,7% populasi Indonesia sudah melek teknologi, namun hanya 13% koperasi yang mengadopsi sistem digital. Untuk mengatasi hambatan dalam peningkatan partisipasi anggota koperasi, penelitian ini mengembangkan aplikasi pendidikan menggunakan metodologi Human-Centered Design (HCD) dan evaluasi Heuristic. Fokus pada lima kriteria usability yakni learnability, efficiency, memorability, errors, dan satisfaction, penelitian ini bertujuan menciptakan antarmuka pengguna yang intuitif dan pengalaman pengguna yang memuaskan. Hasil evaluasi menunjukkan peningkatan yang signifikan dengan desain baru aplikasi. Learnability meningkat 16%, efisiensi tugas naik 12,09% dengan waktu penyelesaian tugas yang lebih cepat, dan tingkat kesalahan menurun 41,67%. Memoriabilitas pengguna meningkat 16,55%, menandakan desain yang lebih mudah diingat. Peningkatan paling mencolok terlihat pada skor kepuasan pengguna yang melonjak dari 54 menjadi 80,5, menandakan peningkatan sebesar 49,07%. Inisiatif ini menunjukkan bagaimana HCD dan evaluasi Heuristic dapat bersinergi untuk mengimprovisasi User Experience dan User Interface dalam aplikasi pendidikan koperasi, menciptakan pengalaman pengguna yang lebih menyenangkan dan efektif, serta meningkatkan kesejahteraan anggota melalui pendidikan teknologi yang inklusif dan efektif.

Katakunci: Web, Pengujian Kegunaan, Evaluasi Heuristik, Antarmuka Pengguna, Pengalaman Pengguna, Desain yang Berpusat pada Manusia.

1. Pendahuluan

Dalam era globalisasi saat ini, kita tentunya menghadapi berbagai tantangan yang signifikan terkait dengan penggunaan dan perkembangan teknologi informasi. Karena itu pentingnya teknologi informasi dalam kehidupan sehari-hari yang terbukti esensial dalam kehidupan manusia, sebagaimana ditunjukkan oleh laporan Digital Indonesia tahun 2022[3], yang menyatakan bahwa sekitar 73.7% populasi Indonesia telah melek teknologi dan internet.

Meskipun demikian, realitanya menunjukkan bahwa hanya sebagian kecil yang benar-benar memanfaatkan teknologi ini untuk meningkatkan efisiensi pekerjaan mereka. Salah satu bidang yang terkena dampak langsung dari tantangan ini adalah sektor Usaha Mikro, Kecil, dan Menengah (UMKM), khususnya dalam bentuk koperasi. Koperasi, yang didefinisikan sebagai badan usaha kolektif dengan tujuan meningkatkan kesejahteraan ekonomi anggotanya [8], saat ini menghadapi stagnasi dalam

perkembangannya. Data dari Kementerian Koperasi dan UMKM [8] menunjukkan bahwa dari 127.846 koperasi yang aktif di Indonesia, hanya sekitar 13% yang telah mengadopsi sistem digital.

Faktor kunci dalam kemajuan koperasi bukan hanya terletak pada adopsi teknologi, tetapi juga pada pengetahuan dan partisipasi anggotanya dalam mengembangkan koperasi. Terdapat hambatan dalam peningkatan partisipasi anggota, yang sering kali dikaitkan dengan kurangnya pemahaman mereka tentang mekanisme dan bisnis koperasi. Hal ini menyoroti pentingnya pendidikan sebagai alat utama dalam meningkatkan kesadaran dan motivasi anggota koperasi. Berangkat dari pemahaman ini, penulis tertarik untuk mengembangkan teknologi yang dapat meningkatkan pengetahuan anggota koperasi.

Aplikasi pendidikan yang dirancang tidak hanya fokus pada inovasi fungsional, tetapi juga pada aspek kegunaan (*usability*) produk. Dalam hal ini, penulis memilih untuk menerapkan metode *Human-Centered Design* (HCD) untuk pengembangan desain sistem dan *Heuristic Evaluation* untuk Langkah *usability* desain sistem. Kedua metode ini terbukti efektif dalam memastikan *usability*, seperti yang terlihat dalam penelitian sebelumnya yang menggunakan HCD untuk analisis desain sistem dan *usability* [] dan *Heuristic Evaluation* untuk menilai kepuasan pengguna [5].

Penelitian sebelumnya menunjukkan bahwa *HCD* dapat diterapkan dalam analisis ukuran *usability* suatu sistem, dengan fokus pada lima kriteria menurut Nielsen yaitu *learnability*, *efficiency*, *memorability*, *errors*, dan *satisfaction*. Selain itu, ada penelitian yang menggunakan Design Thinking, yang berbeda dengan HCD karena lebih menekankan pada pengembangan solusi bagi masalah pengguna, tidak terbatas hanya pada antarmuka pengguna. Selain itu dimana sebelumnya sudah terdapat sebuah penelitian mengenai penggunaan *Heuristic Evaluation* guna menilai tingkat kepuasan pengguna berdasarkan lima kriteria pada Heuristic yang telah dilakukan [1]. Dimana peneliti tersebut menguji sebuah aplikasi kesehatan berbasis mobile yang menghasilkan, dengan digunakannya metode *Heuristic Evaluation* yang dimana penulis dapat menyimpulkan bahwa aplikasi tersebut telah memiliki tingkat kebergunaan yang cukup baik.

Dengan pertimbangan ini, penulis mengambil inisiatif untuk melakukan penelitian tentang "Implementasi HCD dan Heuristic Evaluation dalam Improvisasi User Experiences dan User Interface (Studi Kasus: Aplikasi Pendidikan dan Pelatihan Anggota Koperasi)". Tujuan utama adalah untuk menghasilkan desain antarmuka yang tidak hanya diterima dengan baik oleh pengguna, tetapi juga memberikan pengalaman pengguna yang nyaman melalui model Human-Centered Design.

2. Metode Penelitian

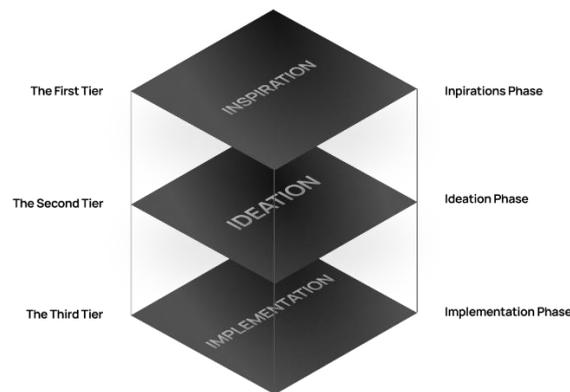
2.1. Pengumpulan Data

Proses untuk pengumpulan data untuk evaluasi dan peningkatan aplikasi pendidikan dan pelatihan koperasi akan dikumpulkan melalui metode evaluasi heuristik dan pengujian Usability. Metode heuristik, memungkinkan identifikasi masalah desain awal secara efisien dan mendapatkan data perbaikan dari responden ahli. Setelah implementasi perbaikan, pengujian Usability dengan pengguna aktual dilakukan untuk menggali data kuantitatif dan kualitatif, memperkaya pemahaman tentang efektivitas aplikasi dan diuji dengan para responden akhir.

2.2. Human-Centered Design

Pada tahap ini, penulis merancang seluruh desain hasil evaluasi heuristik berdasarkan pendekatan *Human-Centered Design*, dalam mengembangkan aplikasi ini, tahapan pertama ada pada Pencarian inspirasi berupa apa jenis desain yang kita kembangkan lalu tahap kedua ada idetasi dimana ini merangkum hasil dari pencarian inspirasi tersebut 1 kerangka utuh yang kita anggap sebagai pedoman dalam merancang fitur yang diperlukan user yang tentunya dapat diimplementasikan, terakhir

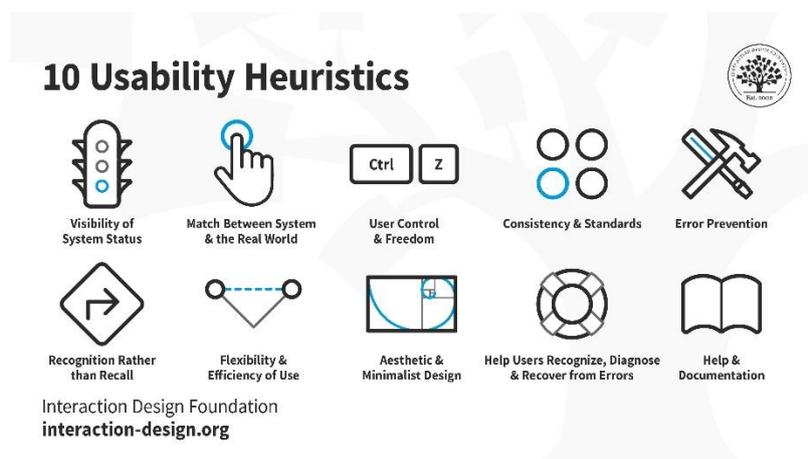
ada tahap implementasi dimana seluruh konsep desain yang telah dipilih akan di implementasikan ke dalam bentuk produk akhir yang akan digunakan oleh user [2].



Gambar 1. HCD Diagram

2.3. Heuristic Evaluations

Evaluasi heuristik adalah metode investigasi yang digunakan dalam desain interaksi dan pengalaman pengguna untuk mengevaluasi dan mengidentifikasi masalah atau cacat antarmuka pengguna [4]. Metode ini dilakukan oleh individu atau sekelompok evaluator yang merupakan ahli dalam bidang pengalaman pengguna yang mengidentifikasi pelanggaran terhadap prinsip-prinsip desain interaksi yang diberikan. Teknik evaluasi heuristik ini dapat dilakukan di awal fase desain sebelum mengujinya dengan pengguna. Prinsip-prinsip desain interaksi yang dijadikan acuan dalam evaluasi heuristik biasanya merupakan hasil penelitian atau pengalaman para ahli UX yang telah terbukti efektif dalam meningkatkan kualitas UX. berikut adalah 10 prinsip desain interaksi yang biasa digunakan dalam evaluasi heuristik yang adalah sebagai berikut [4]:



Gambar 2. Komponen Heuristics Evaluations

2.4. Usability Testing

Proses ini adalah salah satu pengukuran yang populer digunakan di dalam dunia UX/UI Design, proses ini melibatkan pengguna akhir untuk melakukan pengujian aplikasi sebelum di kembangkan lebih lanjut sehingga dapat mengurangi biaya dan dapat melihat masalah yang tidak terlihat di fase pengembangan sehingga ini memudahkan para desainer untuk mengembangkan aplikasi yang lebih baik [7][2]. Ada 5 Komponen dalam usability yang penulis gunakan dalam mengukur hasil desain setelah revisi heuristic desain yaitu *Learnability* (Komponen ini mengacu pada kemudahan bagi pengguna baru untuk mulai menggunakan produk dan dengan cepat menjadi efisien dalam menggunakannya.), *Errors* (komponen yang menilai seberapa banyak kesalahan yang dilakukan user), *Satisfaction* (Komponen yang mengukur sejauh mana kepuasan user dalam menggunakan produk), *Memorability* (Komponen yang mengukur seberapa inget user terhadap produk yang digunakan dalam

jangka waktu tertentu), *Efficiency* (Komponen yang mengukur seberapa butuh usahanya user dalam mengakses suatu fitur atau produk) [1].

3. Hasil dan Pembahasan

3.1 Pengumpulan Evaluator awal

Pada penelitian tahap awal, peneliti atau penulis melakukan penilaian menyeluruh terhadap iterasi desain awal aplikasi pendidikan dan pelatihan anggota koperasi. Tujuan utama dari penilaian ini adalah untuk menunjukkan dengan tepat area dan karakteristik yang memerlukan peningkatan atau modifikasi. Hasil dan perspektif yang diperoleh dari penilaian ini akan menjadi landasan penting dalam pembuatan prototipe desain baru. Prototipe ini dimaksudkan untuk menampilkan desain yang lebih ramah pengguna dan efisien, dengan mempertimbangkan kebutuhan dan preferensi pengguna secara lebih akurat. Dengan mengadopsi pendekatan ini, peneliti bertujuan untuk mengembangkan solusi desain yang tidak hanya mengatasi kekurangan versi sebelumnya, namun juga meningkatkan nilai dan pengalaman pengguna secara keseluruhan. Maka dari itu dibutuhkan 3-5 kelompok ahli dalam melakukan pengujian heuristic ini diantaranya sebagai berikut,

Table 1. Responden Awal/Ahli

No	Nama	Jenis Kelamin	Pekerjaan
1	Angga Rizky	Laki-Laki	Senior UX/UI Designer
2	Bagus Aditya	Laki-Laki	Freelancer UX Designer
3	Muhammad Detya	Laki-Laki	UI Designer

3.2 Hasil pengujian desain awal *Heuristik Evaluation*

Table 2. Hasil Heuristik Evaluations

Kode Heuristik	Jumlah Masalah	Persentase	Rata rata nilai <i>Severity ratings</i>
U001	2	22,22%	4
U002	3	33,22%	3,3
U003	1	11,11%	4
U004	0	0%	0
U005	0	0%	0
U006	0	0%	0
U007	1	11,11%	3
U008	2	22,22%	3,5
U009	0	0%	0
U010	0	0%	0

Berdasarkan hasil pada table 2 diatas didapat pada prinsip desain dengan kode U002 mendapatkan nilai persentase tinggi berupa 33,22%, sedangkan nilai rata rata severity ratings yang tertinggi terdapat pada U001 dan U003 dengan nilai rata rata severity ratings 4. Namun prinsip desain lainnya juga mendapatkan rata rata nilai 3 hingga 3,5 sehingga ini mendekati nilai 4 yaitu catastrophic.

3.3 Implementasi Desain aplikasi

a. Rekomendasi Desain Aplikasi dari para ahli

Rekomendasi yang ditawarkan telah dirumuskan melalui analisis komprehensif terhadap data yang dikumpulkan selama tahap pengujian, dengan tujuan untuk langsung menyoal permasalahan yang dihadapi pengguna. Saran-saran ini mencakup peningkatan respons sistem, keseragaman desain visual, dan penyederhanaan pedoman operasional. Integrasi dari penyempurnaan yang terinformasi ini diharapkan dapat meningkatkan kemudahan penggunaan aplikasi, memfasilitasi navigasi yang lebih lancar dan pada akhirnya meningkatkan pengalaman pendidikan dan keterlibatan pengguna dengan

platform Pendidikan dan Pelatihan anggota koperasi. Berikut adalah table rekomendasi perbaikan berdasarkan severity ratings yang bernilai tinggi.

Table 3. Hasil Rekomendasi desain aplikasi

No	Deskripsi Masalah	Saran perbaikan dari evaluator	Severity Ratings
1	Tidak ada feedback visual atau audio saat pengguna menyelesaikan video pelatihan atau materi	Buat notifikasi yang menyatakan pengguna telah menyelesaikan video, agar pengguna tahu bahwa mereka telah menonton materi keseluruhan video	4
2	Aplikasi terlalu sederhana dan kurang fitur yang merepresentasikan e-learning	Bisa menambahkan fitur fitur seperti kelas, kuis hingga materi pembelajaran	4
3	Label pada tombol tidak konsisten dengan aksi yang diharapkan pengguna.	Sesuaikan label tombol dengan aksi yang umum dipahami oleh pengguna.	4
4	Ikon yang digunakan tidak intuitif dan tidak mencerminkan fungsi sebenarnya.	Pilih ikon yang universal dan mudah dikenali untuk mewakili fungsi yang relevan.	4
5	Tidak ditemukan fungsi shortcut ke kelas maupun ke fitur lain	Bisa menambahkan fitur shortcut agar bisa lebih efisien dalam menavigasikan fungsi 1 ke yang lainnya.	4
6	Desain terlalu sederhana dan tidak mencerminkan aplikasi dimana disini adalah e-learning dimana lebih sering menggunakan warna dominasi hijau maupun biru	Membuat desain lebih menarik dan penggunaan warna juga diperhatikan	4

Pada tabel diatas memberikan ringkasan komprehensif hasil yang diperoleh dari uji evaluasi heuristik yang dilakukan oleh para evaluator atau responden ahli. Masalah yang teridentifikasi telah diklasifikasikan dan dikelompokkan berdasarkan kebutuhan perbaikan antarmuka, menyoroti area spesifik dari desain antarmuka pengguna yang memerlukan perhatian strategis. Ketiga individu yang terlibat dalam proses evaluasi tidak hanya mengidentifikasi masalah-masalah ini tetapi juga memainkan peran penting dalam mempengaruhi pengembangan saran desain di masa depan.

3.3 Perancangan Desain Aplikasi

a. User Persona

User Persona disini adalah orang atau user yang menggunakan segala fitur yang ada pada aplikasi[], dimana disini peneliti mengembangkan aplikasi Pendidikan dan pelatihan anggota koperasi. Berikut penjabarannya,

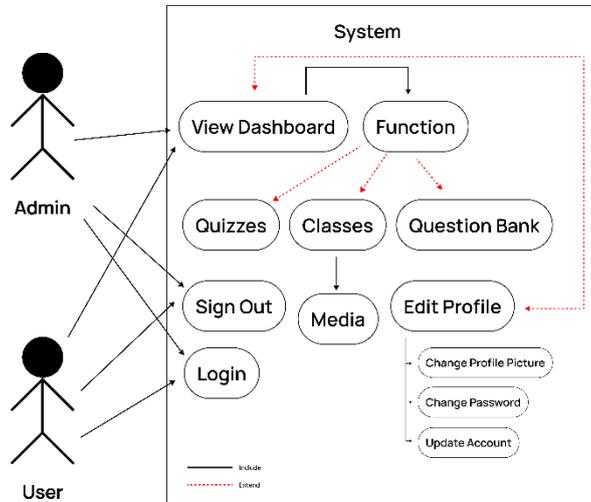
Table 3. User Persona

User	Keterangan
Admin	User/pengguna admin dapat menambahkan atau mendaftarkan pengguna anggota, melihat, masuk dan membuat kelas, kuis, materi pembelajaran hingga mengatur sertifikasi, mengedit password, menghapus password, dan menambahkan password.
Anggota	User Anggota dapat masuk ke kelas yang sudah dibuat, mengerjakan kuis, materi pembelajaran, dan mendapatkan sertifikasi, masuk ke dashboard dengan user dan password yang sudah disediakan, mengedit password, menghapus password dan menambahkan password.

b. Use Case Diagram

Wijaya, Pramarta, Astuti, Rahayuda
 Implementasi Metode Human Centered Design Dan Heuristic Evaluations Dalam Improvisasi User
 Experiences Dan User Interfaces (Studi Kasus : Aplikasi Pendidikan Dan Pelatihan Anggota
 Koperasi)

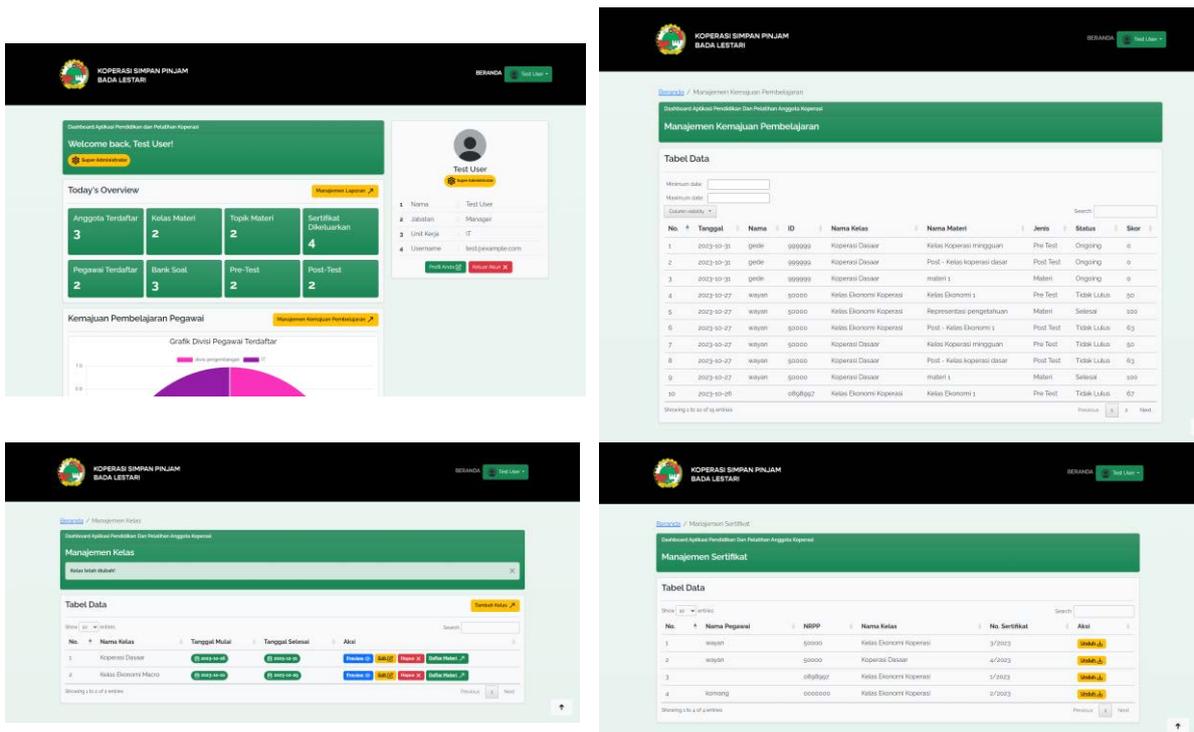
Dalam pengertiannya adalah sebuah diagram vital untuk menggambarkan segala interaksi sistem dan penggunanya [7]. Dalam penelitian ini adalah para pengguna sistem misalnya admin dan user yang terlibat dengan interaksi sistem, tentunya hal ini sangat penting mengingat ini adalah dokumentasi fungsional sistem dan uraian penggunaan yang direncanakan untuk sistem yang sedang mengalami desain ulang antarmuka pengguna.

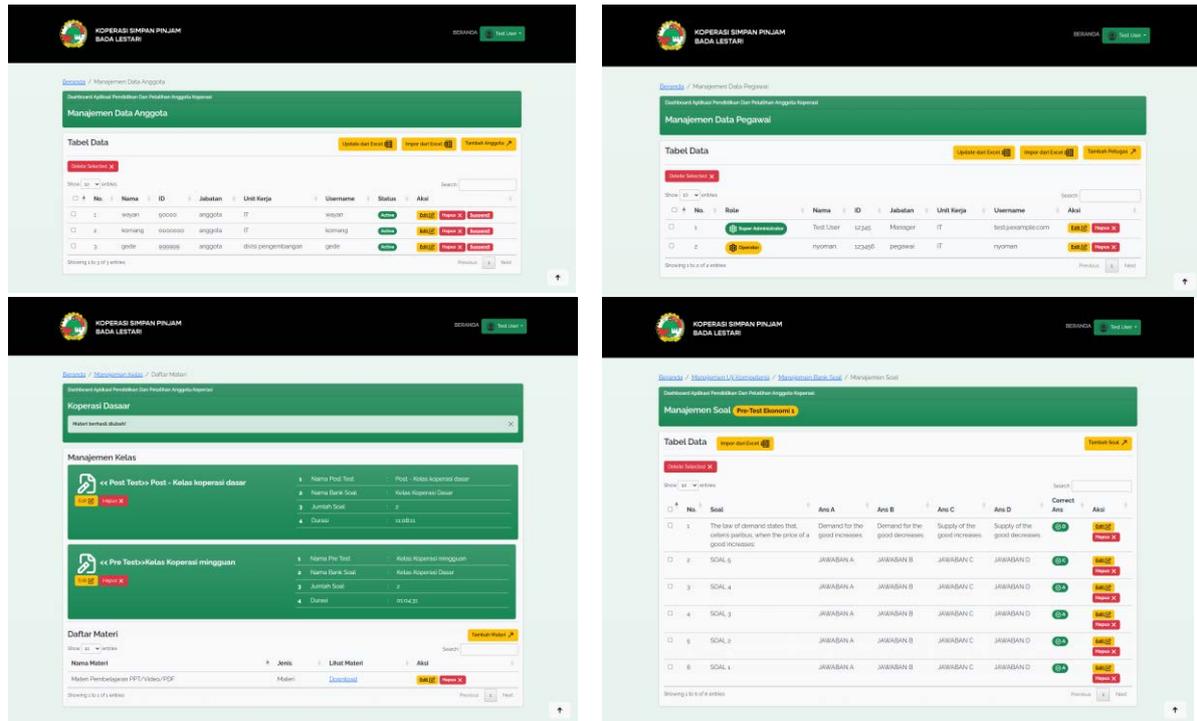


Gambar 3. Use Case Diagram

c. Perancangan High Fidelity Design

Pada tahap ini, penulis melakukan pengembangan desain menjadi high fidelity design atau design yang siap untuk di uji secara usability atau siap untuk diimplementasikan, berikut adalah hasil desainnya.





Gambar 4. Hasil Desain Aplikasi Pendidikan Dan Pelatihan Anggota Koperasi

3.4 Usability Testing

a. Identifikasi User

Tahap identifikasi pengguna merupakan prosedur penting yang bertujuan untuk mengumpulkan informasi yang diperlukan tentang responden yang akan berpartisipasi dalam evaluasi kebergunaan pada aplikasi pendidikan dan pelatihan anggota koperasi, dan berikut adalah anggota koperasi yang menjadi responden dalam uji kebergunaan ini adalah sebagai berikut,

Table 4. Responden Usability Testing

Responden	Nama	Jenis Kelamin	Pekerjaan	Usia
1	I Gede Angga Wijaya	Laki-Laki	Wiraswasta	29
2	I Made Widi Permana	Laki-Laki	PNS	32
3	Ni Wayan Ningsih	Perempuan	Wiraswasta	33
4	Ni Made Aryaning	Perempuan	PNS	35
5	Wahyu Arya Permana	Laki-Laki	Pedagang	25

b. Hasil Akhir Pengukuran Semua Aspek Usability

Table 5. Hasil perhitungan Usability

No	Aspek Usability	Desain Lama	Desain Baru	Peningkatan (%)
1	Learnability	76%	92%	16%
2	Efficiency	16.06 Detik	14.12 detik	12.09%
3	Error	48%	28%	41.67%
4	Memorability	58.33%	68%	16.55%.
5	Satisfaction	54	80.5	26.5%

Pada Tabel diatas memperlihatkan peningkatan yang mengesankan dalam semua aspek kebergunaan aplikasi dengan peralihan dari desain lama ke desain baru, dimana aspek kemudahan belajar (learnability) meningkat 16%, aspek efisiensi tugas meningkat 12.09% dengan penurunan waktu

yang diperlukan untuk menyelesaikan tugas, dan aspek tingkat kesalahan (error) menurun 41.67%, mengindikasikan desain yang lebih intuitif dan mudah digunakan. Aspek Memoriabilitas disini pengguna meningkat 16.55%, menunjukkan desain yang lebih baik dan lebih mudah diingat oleh pengguna. Signifikansi perbaikan ini paling mencolok dalam skor aspek kepuasan pengguna (satisfaction), yang melonjak dari 54 menjadi 80.5, menandai peningkatan kepuasan pengguna sebesar 49.07%. Perubahan ini tentunya menggarisbawahi kesuksesan desain baru dalam menciptakan pengalaman pengguna yang lebih menyenangkan dan efektif.

4. Kesimpulan

Berdasarkan penelitian yang telah disampaikan dan dilaksanakan dan juga hasil yang telah diperoleh selama hasil penelitian, maka peneliti bisa menyimpulkan berbagai data yang sudah disampaikan sebagai berikut:

Evaluasi existing yang dilakukan melalui metode *heuristic evaluation* telah berhasil mengidentifikasi 9 permasalahan desain antarmuka yang signifikan, berkat Analisa yang dilakukan oleh 3 para *evaluator expert*. Permasalahan-permasalahan ini berkisar dari isu-isu kecil hingga masalah serius yang berpotensi menghambat penggunaan aplikasi secara efektif. Untuk mengkategorikan dan menentukan tingkat urgensi setiap masalah, *severity ratings* diterapkan dengan teliti. Sistem penilaian ini memungkinkan tim desain untuk memahami dan memprioritaskan isu berdasarkan dampaknya, dari yang bersifat catastrophic, yang menunjukkan kebutuhan perbaikan segera karena berpotensi mengganggu operasi penting, hingga major dan minor issues, yang meskipun kurang kritikal, tetap memerlukan perhatian untuk memperhalus pengalaman pengguna. Dengan memanfaatkan hasil evaluasi ini, tim pengembang dapat merancang rencana tindak lanjut yang terstruktur untuk mengatasi masalah desain antarmuka secara bertahap, dengan memulai dari permasalahan yang paling kritikal untuk diatasi.

Dapat disimpulkan bahwa revisi desain pada aplikasi Pendidikan dan Pelatihan Anggota Koperasi telah menghasilkan peningkatan signifikan dalam aspek-aspek kunci usability. Secara khusus, aspek pembelajaran (Learnability) mengalami peningkatan sebesar 16%, menunjukkan bahwa pengguna baru kini dapat memahami aplikasi dengan lebih cepat. Efisiensi penggunaan aplikasi juga meningkat, dibuktikan dengan penurunan waktu yang dibutuhkan (Efficiency) dari 16.06 detik menjadi 14.12 detik, mencerminkan peningkatan sebesar 12.09%. Dari segi kesalahan (Error), terjadi penurunan yang dramatis dari 48% menjadi 28%, menunjukkan bahwa aplikasi menjadi lebih intuitif dan mengurangi kemungkinan pengguna melakukan kesalahan sebesar 41.67%. Memoriabilitas (Memorability), atau kemudahan pengguna mengingat cara menggunakan aplikasi setelah tidak menggunakannya untuk sementara waktu, meningkat sebesar 16.55%. Dan terakhir, tingkat kepuasan (Satisfaction) pengguna meningkat signifikan dari skor 54 menjadi 80.5, yang menandakan peningkatan sebesar 49.07%. Kesimpulannya akhirnya, revisi desain telah berhasil membuat aplikasi tidak hanya lebih efisien dan intuitif, tetapi juga lebih memuaskan untuk digunakan oleh anggota koperasi.

References

- [1] Kemp, S. (2022). Digital 2022: Indonesia. DataReportal. <https://datareportal.com/reports/digital-2022-indonesia>
- [2] Sokhom, W., & Mekruksavanich, S. (2021). A cooperative education management system using technology acceptance model. 2021 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunication Engineering, 297-300. <https://doi.org/10.1109/ECTIDAMTNCN51128.2021.9425690>
- [3] Sharma, R., & Singh, J. (2021). Human centred software development approaches. 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), 602-605. <https://doi.org/10.1109/ICAC3N53548.2021.9725715>

- [4] Putra, F., & Tedyyana, A. (2021). Pendekatan human centered design pada perancangan user experience aplikasi pemesanan menu cafe. *SISTEMASI*, 10(2), 336. <https://doi.org/10.32520/stmsi.v10i2.1229>
- [5] Ardiningrum, P., Fabroyir, H., & Akbar, R. J. (2021). Desain dan evaluasi antarmuka pengguna responsif aplikasi MyITS Connect modul job marketplace menggunakan metodologi user-centered design. *Jurnal Teknik ITS*, 10(2). <https://doi.org/10.12962/j23373539.v10i2.62903>
- [6] Hasna, R. N., Tolle, H., & Wijoyo, S. H. (2020). Evaluasi dan perancangan user experience menggunakan metode human centered design dan heuristic evaluation pada aplikasi Dunia Games. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 4(4), 1333–1343. <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/7226>
- [7] Mathews, C., Ye, K., Grozdanovski, J., Marinelli, M., Zhong, K., Khalajzadeh, H., Obie, H., & Grundy, J. (2021). AH-CID: A tool to automatically detect human-centric issues in app reviews. *Proceedings of the International Conference on Software Engineering and Knowledge Engineering (SEKE)*, 386-397. <https://doi.org/10.5220/0010576503860397>
- [8] Sokhom, W., & Mekruksavanich, S. (2021). A cooperative education management system using technology acceptance model. *2021 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunication Engineering*, 297-300. <https://doi.org/10.1109/ECTIDAMTNCON51128.2021.9425690>

This page is intentionally left blank.

Faktor Penerimaan dan Penggunaan Siswaskeudes oleh Pengawas Keuangan Desa

Putu Eka Pratiwi Widiantari^{a1}, Made Gede Wirakusuma^{a2}, Ni Putu Sri Harta Mimba^{a3}, Putu Agus Ardiana^{a4}

^aFakultas Ekonomi dan Bisnis, Universitas Udayana
Jl. P.B. Sudirman, Dangin Puri Klod, Kec. Denpasar Tim., Kota Denpasar, Bali 80112, Indonesia

¹ekapratiwi.widiantari85@gmail.com

Abstract

This research investigates the factors influencing the acceptance and use of the Village Financial Monitoring System (Siswaskeudes) application by village financial supervisors. Utilizing an interpretive approach and hermeneutic phenomenology, the study aims to understand supervisors' experiences and perceptions regarding the usefulness and ease of use of Siswaskeudes. Data were collected through in-depth interviews with five informants from the Denpasar City Inspectorate. The results suggest that application usage is significantly influenced by perceived usefulness, ease of use, attitudes towards the behavior, subjective norms, and perceived behavioral control. Most informants found Siswaskeudes facilitated easier data access and enhance work efficiency, despite identifying initial use challenges and the need for further development. Compliance with legal regulations, reflected in subjective norms, is a primary driver of the application's use. This research recommends ongoing development, training, regular monitoring, team support, and security enhancements to boost Siswaskeudes' acceptance and usage. Future research should explore user satisfaction using quantitative approaches with the Technology Acceptance Model (TAM) and consider additional variables such as security and cost savings.

Keywords: *Siswaskeudes, Theory of Reasoned Action, Theory of Planned Behavior, Technology Acceptance Model, Interpretative Phenomenological Analysis*

1. Pendahuluan

Perkembangan teknologi informasi yang pesat memberikan peluang besar bagi pemerintah untuk meningkatkan efisiensi dan transparansi dalam pengelolaan keuangan. *E-government* menjadi topik yang populer, terutama ketika dikaitkan dengan otonomi daerah. Pada dasarnya, *e-government* merupakan pemanfaatan teknologi informasi untuk meningkatkan hubungan antara pemerintah dan para pemangku kepentingan, termasuk masyarakat, bisnis, dan pemerintah lainnya. Tantangan terbesar ada pada kemampuan dan kesiapan manajemen serta Sumber Daya Manusia (SDM) dari para pelaku. *E-government* memberikan kontribusi dalam memudahkan pemerintahan dan para pemangku kepentingan, salah satunya dalam melakukan pengawasan khusus desa. Inspektorat Daerah dibantu dengan aplikasi Siswaskeudes (Sistem Pengawasan Keuangan Desa) yang terintegrasi dengan aplikasi Siskeudes (Sistem Keuangan Desa) yang digunakan pada masing-masing desa. Siswaskeudes dirancang untuk membantu pengawasan pengelolaan keuangan desa secara lebih efektif dan efisien. Aplikasi ini bertujuan untuk mengurangi penyelewengan dana desa dan meningkatkan akuntabilitas serta transparansi dalam pengelolaan keuangan desa.

Meskipun aplikasi Siswaskeudes memiliki banyak manfaat potensial, keberhasilan implementasinya sangat tergantung pada faktor-faktor yang mempengaruhi penerimaan dan penggunaan oleh para pengawas, yang dalam hal ini adalah aparat pemerintah daerah, Inspektorat kabupaten/kota. Pemahaman yang mendalam mengenai persepsi dan tantangan yang dihadapi oleh para pengawas ini penting untuk memastikan keberhasilan implementasi dan operasionalisasi Siswaskeudes.

Fuad et al. meneliti penerapan model *Unified Theory of Acceptance and Use of Technology* (UTAUT) pada Sistem Keuangan Desa (Siskeudes) dengan etika kerja Islami sebagai variabel moderasi [1]. Penelitian ini menemukan bahwa tiga dari empat variabel dalam model UTAUT, yaitu *performance expectancy*, *social influences*, dan *facilitating conditions*, memiliki pengaruh signifikan. Namun, variabel moderasi etika kerja Islami hanya memoderasi ekspektasi bisnis dan itu pun tidak signifikan. Indriani et al. menggunakan model keberhasilan DeLone dan McLean dan menemukan bahwa kualitas informasi dan kualitas sistem memiliki pengaruh positif terhadap efektivitas organisasi dan kepuasan pengguna [2]. Kualitas pelayanan juga berpengaruh positif terhadap efektivitas organisasi tetapi tidak mempengaruhi kepuasan kerja. Selain itu, efektivitas organisasi terbukti berpengaruh positif terhadap tata kelola. Rosnidah et al. menggabungkan model UTAUT dan *technology readiness* (TR) untuk mengukur akseptabilitas dan kesiapan teknologi pada sistem keuangan desa (Siskeudes). Mereka menemukan bahwa orang-orang di sekitar pengguna, seperti teman atau atasan, sangat mempengaruhi penggunaan sistem keuangan desa oleh perangkat desa [3].

Penelitian awal ini bertujuan untuk mengidentifikasi faktor-faktor yang mempengaruhi penerimaan dan penggunaan Siswaskeudes oleh pengawas keuangan desa, memahami persepsi mereka terhadap manfaat dan kemudahan penggunaan aplikasi, serta mengidentifikasi hambatan-hambatan yang dihadapi dalam implementasinya.

Sistem Pengawasan Keuangan Desa (Siswaskeudes)

Aplikasi Sistem Pengawasan Keuangan Desa (Siswaskeudes) digunakan oleh Aparat Pengawasan Intern Pemerintah (APIP) daerah sebagai alat untuk memeriksa pengelolaan keuangan desa dengan pendekatan berbasis risiko dan menggunakan Teknik Audit Berbantuan Komputer (TABK). Pengembangan aplikasi Siswaskeudes melibatkan Badan Pengawasan Keuangan dan Pembangunan (BPKP) dan Kementerian Dalam Negeri, mengacu pada Peraturan Menteri Dalam Negeri Nomor 73 Tahun 2020 tentang Pengawasan Pengelolaan Keuangan Desa. Walaupun ada delapan jenis pengawasan yang dilakukan oleh Inspektorat Kabupaten/Kota sesuai dengan peraturan tersebut, saat ini Siswaskeudes hanya mendukung satu jenis pengawasan, yaitu pemeriksaan kinerja pengelolaan keuangan dan aset desa.

Setiap tahun nilai APBDesa semakin besar sehingga pemerintah pusat dan daerah semakin mendorong optimalisasi peran APIP daerah kabupaten/kota untuk melakukan pengawasan atas pengelolaan keuangan desa. Permendagri Nomor 73 tahun 2020 menegaskan bahwa pengawasan keuangan desa di wilayah kabupaten/kota dilakukan oleh APIP daerah kabupaten/kota dan camat. Tujuan aplikasi ini adalah membantu inspektorat kabupaten/kota dalam melaksanakan pemeriksaan kinerja pengelolaan keuangan dan aset desa, mencakup mulai dari perencanaan audit berbasis risiko, pengujian bukti audit, dokumentasi aktivitas audit, hingga penyusunan laporan audit.

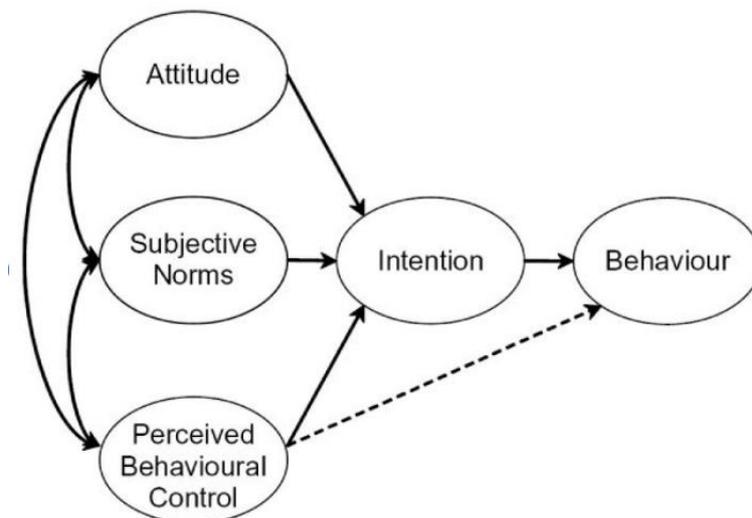
Theory of Reasoned Action (TRA) dan Theory of Planned Behavior (TPB)

Teori Perilaku Terencana (TPB) menjelaskan faktor-faktor yang memengaruhi niat berperilaku. TPB merupakan pengembangan dari Teori Tindakan Beralasan (TRA), yang menyatakan bahwa niat untuk melakukan suatu tindakan dipengaruhi oleh norma subjektif dan sikap terhadap perilaku [4]. Pada tahun 1991, Ajzen menambahkan faktor kontrol perilaku yang dipersepsikan individu atau *perceived behavioral control*, sehingga TRA berkembang menjadi TPB [5].

Teori Tindakan Beralasan (TRA/*Theory of Reasoned Action*) adalah model yang telah terbukti efektif dalam memprediksi dan menjelaskan perilaku di berbagai domain [6]. TRA sangat umum dan dirancang untuk menjelaskan hampir semua perilaku manusia, sehingga cocok digunakan untuk mempelajari penentu perilaku penggunaan sistem informasi. Davis et al. memperkenalkan adaptasi dari TRA, yaitu Model Penerimaan Teknologi (TAM/*Technology Acceptance Model*), yang secara khusus dirancang untuk menjelaskan perilaku penggunaan sistem informasi [7]. TAM menggunakan TRA sebagai dasar teoretis untuk menentukan hubungan sebab akibat antara dua keyakinan utama: kegunaan yang dirasakan dan kemudahan penggunaan yang dirasakan, serta sikap pengguna, niat, dan perilaku yang diadopsi. TAM didesain lebih khusus daripada TRA, TAM dirancang khusus untuk perilaku penggunaan sistem informasi. TAM menjelaskan hubungan sebab akibat antara keyakinan (manfaat sistem informasi

dan kemudahan penggunaannya), perilaku, tujuan atau keperluan, dan penggunaan aktual dari pengguna suatu sistem informasi. Kegunaan yang dirasakan (*perceived usefulness*) dan kemudahan penggunaan yang dirasakan (*perceived ease of use*) adalah variabel utama TAM. Niat untuk menggunakan teknologi sangat dipengaruhi oleh kegunaan yang dirasakan dan kemudahan penggunaan yang dirasakan [8].

Fokus utama TPB adalah niat individu untuk melakukan perilaku tertentu. Niat ini dianggap sebagai motivasi yang memengaruhi perilaku individu dan merupakan prediktor seberapa keras seseorang akan berusaha dan seberapa besar usaha yang akan dikeluarkan untuk melakukan suatu tindakan. Semakin kuat niat seseorang untuk melakukan suatu perilaku, semakin tinggi kemungkinan perilaku tersebut terjadi. Niat berperilaku hanya muncul jika individu merasa memiliki kendali atas perilaku tersebut [5]. Persepsi kendali perilaku atau *perceived behavioral control* adalah tingkat kendali yang dirasakan individu terhadap suatu perilaku, yang menentukan apakah perilaku tersebut akan dilakukan atau tidak. Oleh karena itu, faktor ini menjelaskan tingkat kontrol yang dimiliki individu terhadap suatu tindakan. Gambar 1. menggambarkan faktor-faktor yang dapat memengaruhi niat dan perilaku seseorang [5].



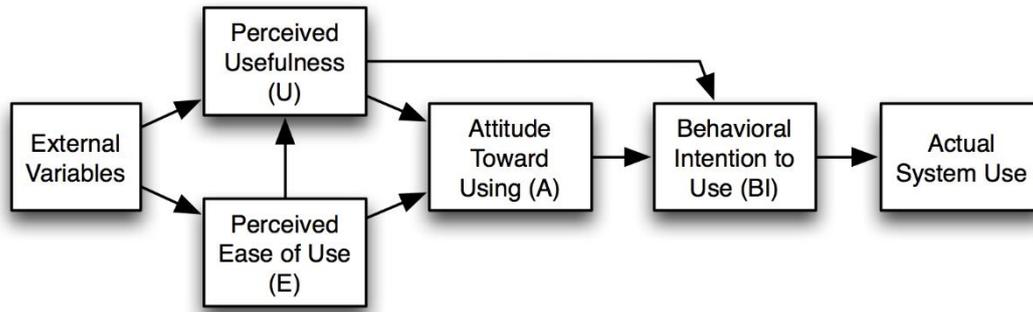
Sumber : Ajzen, 1991 [5]

Gambar 1. Theory of Planned Behavior

Technology Acceptance Model (TAM)

Model penerimaan teknologi (TAM/*Technology Acceptance Model*) adalah adaptasi dari teori tindakan beralasan (TRA). Davis memodifikasi hubungan antara kepercayaan, sikap, niat, dan perilaku pengguna yang berasal dari komponen-komponen TRA untuk secara khusus menjelaskan perilaku pengguna komputer. TAM adalah salah satu model yang dirancang untuk menganalisis dan memahami faktor-faktor yang memengaruhi penerimaan teknologi komputer, sehingga dapat menjelaskan perilaku pengguna akhir teknologi informasi yang beragam [9].

TAM menjelaskan bahwa untuk menggunakan sistem, niat seseorang berperilaku memainkan peran penting, yang dipengaruhi oleh tiga faktor utama yaitu *perceived usefulness* (kegunaan yang dirasakan), *perceived ease of use* (kemudahan penggunaan), dan *attitudes towards usage* (sikap terhadap penggunaan). Ketiga faktor ini digunakan untuk menjelaskan dan memprediksi adopsi sistem teknologi yang ditunjukkan oleh Gambar 2.



Gambar 2. *Technology Acceptance Model*

2. Metode Penelitian

Paradigma Penelitian

Penelitian ini menggunakan paradigma interpretif. Pendekatan interpretif diadopsi dari orientasi praktis. Secara umum, pendekatan interpretif adalah sebuah sistem sosial yang memaknai perilaku dengan cara mengobservasi secara langsung [10]. Pendekatan ini berusaha mencari penjelasan tentang peristiwa-peristiwa sosial atau budaya berdasarkan perspektif dan pengalaman orang yang diteliti. Interpretif memandang fakta sebagai sesuatu yang unik dengan konteks dan makna khusus yang esensial untuk memahami makna sosial. Fakta dalam pendekatan interpretif dianggap fleksibel dan terkait dengan sistem makna.

Penelitian ini bertujuan untuk menggali makna pengalaman pengawas (inspektorat kabupaten/kota) dalam menggunakan Siswaskeudes. Penelitian ini menggunakan pendekatan kualitatif. Penelitian kualitatif adalah penelitian yang menempatkan manusia sebagai subjek utama. Menurut Creswell, penelitian kualitatif adalah proses eksplorasi dan pemahaman terhadap makna perilaku individu dan kelompok, serta menggambarkan masalah sosial atau kemanusiaan [11]. Tujuan dari penelitian kualitatif adalah untuk memperoleh pemahaman mendalam tentang bagaimana orang-orang merasakan proses hidup mereka, memberikan makna, dan menjelaskan bagaimana pengalaman tersebut diinterpretasikan.

Pendekatan Penelitian

Pendekatan penelitian kualitatif yang digunakan dalam penelitian ini adalah pendekatan fenomenologi. Studi fenomenologi menggambarkan makna dari pengalaman hidup beberapa individu terkait dengan suatu konsep atau fenomena [12]. Pokropski, menyatakan bahwa fenomenologi bertujuan untuk menemukan objek dari kesadaran subjek [13]. Dalam penelitian ini, fenomenologi dipilih sebagai metode penelitian, karena dianggap paling cocok untuk mencapai tujuan penelitian, yaitu memaknai pengalaman subjek terhadap suatu fenomena, penelitian ini khususnya memaknai pengalaman pengawas (inspektorat kabupaten/kota) dalam menggunakan Siswaskeudes.

Penelitian ini menggunakan pendekatan fenomenologi hermeneutik. Pendekatan ini menekankan pada pengalaman hidup (fenomenologi) dan penafsiran "teks" kehidupan (hermeneutika). Fenomenologi hermeneutik tidak hanya melibatkan deskripsi tetapi juga proses interpretatif di mana peneliti menghasilkan interpretasi tentang makna pengalaman yang dialami. Proses penyelidikan dan pembuatan makna menjadi fokus dalam fenomenologi hermeneutik, di mana peneliti berusaha memahami bagaimana informan menjelajahi dan memaknai pengalaman mereka sendiri. Oleh karena itu, pendekatan fenomenologi interpretatif tepat digunakan ketika tujuan penelitian adalah untuk memahami sifat pengalaman yang dialami (fenomenologi) dan cara individu memahaminya (interpretasi).

Teknik analisis data dalam penelitian ini menggunakan metode *Interpretative Phenomenological Analysis* (IPA). IPA adalah suatu kerangka metodologi dalam pendekatan

kualitatif yang bertujuan untuk menggali secara detail pengalaman hidup individu. Metode IPA bertujuan untuk mendalami interpretasi subjek terhadap aspek-aspek kehidupan pribadi dan sosial mereka. IPA memperlihatkan sifat hermeneutika ganda, di mana informan (sebagai subjek penelitian) menginterpretasikan pengalaman hidup mereka sendiri, dan kemudian peneliti juga menginterpretasikan pengalaman subjek tersebut [14]. Proses interpretatif melibatkan kedua belah pihak, baik peneliti maupun informan/subjek [15]. Proses pengumpulan data dalam studi IPA memfasilitasi interaksi antara peneliti dan partisipan, di mana pertanyaan awal dapat disesuaikan berdasarkan tanggapan partisipan, dan peneliti memiliki fleksibilitas untuk mengeksplorasi hal-hal menarik lainnya yang muncul selama dialog berlangsung [16].

Informan Penelitian

Dalam penelitian kualitatif, sampel penelitian atau informan biasanya disebut sebagai partisipan. Pemilihan partisipan untuk penelitian ini dilakukan melalui seleksi komprehensif, dimana peneliti secara purposif memilih partisipan yang relevan dengan topik dan tujuan penelitian. Kualifikasi partisipan dalam penelitian ini adalah individu berusia 26-45 tahun yang berperan sebagai pengawas sistem keuangan desa. Rentang usia tersebut dibagi lagi ke dalam dua kelompok usia:

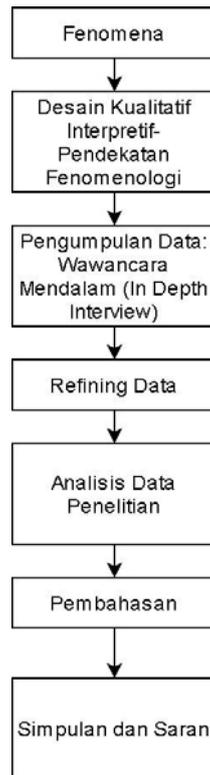
- a. Usia 26-35 tahun: individu dalam kelompok usia ini cenderung menerima dan mengadopsi teknologi baru dengan cepat, lebih inovatif dan berani mencoba hal-hal baru, kemungkinan besar menjadi pengguna awal yang menguji dan memberikan masukan untuk perbaikan teknologi.
- b. Usia 36-45 tahun: individu dalam kelompok usia ini memiliki lebih banyak pengalaman kerja dan kematangan dalam mengambil keputusan berdasarkan manfaat praktis dari teknologi, mereka sering berada pada posisi manajerial sehingga kemungkinan mereka lebih menghargai efisiensi dan peningkatan produktivitas yang ditawarkan oleh teknologi dan menggunakan teknologi untuk tujuan strategis dan pengambilan keputusan yang lebih baik dalam pengelolaan keuangan desa.

Metode Pengumpulan Data

Metode pengumpulan data yang dilakukan dalam penelitian ini adalah wawancara. Metode ini merupakan metode yang paling populer karena memungkinkan peneliti dan peserta terlibat dalam dialog secara *real time* [17]. Prosedur pengumpulan data untuk studi IPA memungkinkan peneliti dan peserta untuk terlibat dalam dialog di mana pertanyaan awal dimodifikasi berdasarkan tanggapan peserta, dan peneliti dapat menanyakan hal-hal menarik lainnya yang muncul kemudian [16].

Model Penelitian

Penelitian ini mengkaji lebih dalam tentang faktor penerimaan dan penggunaan aplikasi Siswaskeudes oleh pengawas keuangan desa dengan menggunakan *theory of planned behavior* dan *technology acceptance model*. Konsep penelitian ini disajikan melalui Gambar 3.



Gambar 3. Model Penelitian

3. Hasil dan Diskusi

3.1. Situs Penelitian

Situs penelitian ini pada inspektorat kota Denpasar. Inspektorat kota Denpasar sudah menggunakan Siswaskeudes dalam kegiatan pengawasannya, dimana sistem keuangan desa pada desa yang berada di wilayah kota Denpasar telah terintegrasi secara *online* pada Sistem Pengawasan Keuangan Desa (Siswaskeudes).

3.2. Deskripsi Informan

Sumber data dalam penelitian ini adalah data primer yang diperoleh dari jawaban informan melalui wawancara secara mendalam (*in depth interview*). Wawancara mendalam adalah metode pengumpulan data yang melibatkan percakapan intensif dan terbuka antara peneliti dan informan. Metode ini bertujuan untuk mendapatkan pemahaman yang mendalam mengenai pengalaman, pandangan, dan perasaan informan tentang topik yang sedang diteliti [18]. Melalui interaksi langsung ini, peneliti dapat mengklarifikasi jawaban, menanyakan lebih lanjut tentang poin-poin yang menarik sehingga memungkinkan peneliti untuk mendapatkan wawasan yang lebih mendalam dan holistik. Tabel 1 menunjukkan data informan yang digunakan dalam penelitian ini.

Tabel 1. Data Informan

No	Partisipan	Jenis Kelamin	Umur	Lama Bekerja di Inspektorat Kota Denpasar
1	Informan 1	Perempuan	42 Tahun	12 Tahun
2	Informan 2	Laki-laki	45 tahun	16 Tahun

No	Partisipan	Jenis Kelamin	Umur	Lama Bekerja di Inspektorat Kota Denpasar
3	Informan 3	Perempuan	35 Tahun	10 Tahun
4	Informan 4	Laki-laki	28 Tahun	5 Tahun
5	Informan 5	Perempuan	25 Tahun	4 tahun

Sumber: Hasil Wawancara (2024)

3.3. Pembahasan

Berdasarkan wawancara dengan 5 (lima) informan yang berperan sebagai pengguna Siswaskeudes, ditemukan bahwa penerapan model penerimaan teknologi atau cara mereka menggunakan teknologi yang ada telah sesuai, namun masih terdapat beberapa kendala yang membutuhkan penanganan lebih lanjut

3.3.1. Kegunaan yang Dirasakan

Sebagian besar informan mengakui bahwa aplikasi Siswaskeudes mempermudah mereka dalam mengakses data secara *online* dan menyelesaikan pekerjaan lebih cepat dan efisien. Aplikasi ini membantu menghemat biaya operasional, terutama dengan mengurangi kebutuhan untuk perjalanan fisik dalam memperoleh data. Penggunaan teknologi ini juga mengurangi konsumsi energi dan waktu yang dibutuhkan untuk tugas-tugas administratif. Informan menekankan bahwa aplikasi ini meningkatkan kecepatan dan ketepatan dalam pekerjaan mereka, yang sangat penting dalam pengelolaan pengawasan keuangan desa. Integrasi Siswaskeudes dengan Siskeudes menjadi salah satu nilai tambah utama, yang mempermudah berbagai aspek dalam pengelolaan keuangan desa dan pengawasan. Meskipun aplikasi ini dinilai memiliki kualitas yang cukup baik, ada catatan dari informan 3:

"Aplikasi Siswaskeudes memberikan banyak manfaat yang bisa meningkatkan kinerja pengawas, namun pengembangan lebih lanjut diperlukan untuk meningkatkan fungsionalitas dan efektivitasnya."

3.3.2. Kemudahan Penggunaan

Dari hasil wawancara dengan lima informan mengenai faktor kemudahan penggunaan aplikasi Siswaskeudes, ditemukan beberapa poin penting yaitu, mayoritas informan menganggap aplikasi ini mudah digunakan, yang menunjukkan bahwa aplikasi ini *user-friendly* bagi kebanyakan pengguna. Kemampuan untuk mengakses data dengan cepat dan menggunakan aplikasi tanpa bantuan merupakan poin penting yang meningkatkan persepsi positif. Terdapat satu informan (Informan 1) yang beropini:

"Saya tidak merasakan kemudahan dalam menggunakan aplikasi Siswaskeudes, aplikasi ini hanya membantu dalam tugas spesifik, yaitu sebatas menentukan desa untuk disampling saja."

3.3.3. Sikap terhadap Perilaku

Berdasarkan wawancara dengan informan terkait faktor sikap terhadap perilaku pada aplikasi Siswaskeudes, dapat disimpulkan bahwa mayoritas responden memiliki sikap positif terhadap aplikasi tersebut. Sebagian besar informan menyatakan bahwa aplikasi Siswaskeudes terintegrasi dengan baik dengan aplikasi Siskeudes dan dapat disinkronkan dengan baik. Mereka juga mengakui bahwa kualitas layanan aplikasi tersebut cukup baik. Informan juga menyatakan bahwa aplikasi Siswaskeudes dapat berfungsi dengan baik dari berbagai fitur yang ditawarkan. Meskipun demikian, ada tanggapan dari Informan 3:

"Aplikasi Siswaskeudes perlu dipantau secara rutin oleh *vendor* untuk memastikan kerjanya tetap optimal, dan aplikasi masih memerlukan kerjasama tim yang baik. Ini menyoroti pentingnya kolaborasi antara pengguna aplikasi dalam menggunakan dan memanfaatkannya secara efektif."

3.3.4. Norma Subjektif

Semua informan sepakat bahwa penggunaan aplikasi Siswaskeudes untuk mengawasi Siskeudes merupakan kewajiban yang ditetapkan oleh pemerintah Republik Indonesia sesuai dengan peraturan yang berlaku. Hal ini menunjukkan bahwa faktor norma subjektif yang berasal dari kepatuhan terhadap regulasi hukum memainkan peran penting dalam penggunaan aplikasi ini. Meskipun ada persepsi yang sama mengenai kewajiban penggunaan aplikasi, beberapa informan menyatakan,

Informan 1: “Masih diperlukan pendampingan dalam menggunakan Siswaskeudes dari narasumber/vendor.”

Informan 3: “Diperlukan kerjasama tim yang baik saat menggunakan aplikasi Siswaskeudes ini.

Dengan demikian, hasil penelitian kualitatif menunjukkan bahwa faktor norma subjektif yang berasal dari kewajiban hukum yang ditetapkan oleh pemerintah menjadi pendorong utama dalam penggunaan aplikasi Siswaskeudes

3.3.5. Kontrol Perilaku yang Dirasakan

Berdasarkan wawancara dengan informan terkait faktor kontrol perilaku yang dirasakan pada aplikasi Siswaskeudes, terdapat beragam tanggapan yang mencerminkan persepsi dan pengalaman pengguna terhadap keamanan dan kemudahan penggunaan aplikasi. Salah satu informan menyatakan bahwa aplikasi Siswaskeudes terasa tidak aman karena merasa belum fasih dalam menggunakan aplikasi tersebut. Hal ini menunjukkan bahwa kesulitan dalam memahami atau mengoperasikan aplikasi dapat memengaruhi persepsi pengguna terhadap tingkat keamanan aplikasi. Sebaliknya, beberapa informan menyatakan bahwa mereka merasa aman dan yakin dapat menggunakan aplikasi Siswaskeudes dengan baik. Hal ini menunjukkan bahwa keyakinan dalam kemampuan menggunakan aplikasi dapat meningkatkan persepsi terhadap keamanan aplikasi.

Informan 3: “Meskipun saya merasa aplikasi Siswaskeudes aman, namun saya masih merasa perlu melakukan pemeriksaan manual sebagai *back up* jika terjadi *system error*.”

Informan 4: “Saya merasa aplikasi Siswaskeudes aman karena terdapat fitur pengamanan seperti *user ID* dan *password*.”

4. Kesimpulan

Studi ini menggali persepsi dan pengalaman para pengawas keuangan desa terkait penerimaan dan penggunaan aplikasi Siswaskeudes. Dalam konteks penerapan aplikasi ini, faktor-faktor seperti kegunaan yang dirasakan, kemudahan penggunaan, sikap terhadap perilaku, norma subjektif, dan kontrol perilaku yang dirasakan memainkan peran penting dalam penggunaan aplikasi Siswaskeudes. Hasil penelitian menunjukkan bahwa mayoritas pengguna merasa bahwa Siswaskeudes memiliki kegunaan yang signifikan dalam memudahkan akses data dan meningkatkan efisiensi pekerjaan. Integrasi dengan aplikasi Siskeudes juga dianggap sebagai nilai tambah yang besar. Meskipun demikian, beberapa informan menyoroti perlunya pengembangan lebih lanjut untuk meningkatkan fungsionalitas aplikasi.

Kemudahan penggunaan juga menjadi faktor penting dalam penerimaan aplikasi ini. Mayoritas informan menganggap aplikasi ini mudah digunakan, meskipun ada beberapa yang mengalami kesulitan pada tahap awal penggunaan. Sikap terhadap perilaku penggunaan aplikasi cenderung positif, dengan mayoritas informan menyatakan kepuasan mereka terhadap kinerja dan kualitas layanan aplikasi. Namun, masih ditemukan tanggapan yang menyoroti perlunya pemantauan rutin dan kerjasama tim yang baik untuk memastikan optimalitas penggunaan.

Norma subjektif yang mendorong kewajiban penggunaan aplikasi sesuai dengan regulasi hukum menjadi faktor penting dalam penerimaan aplikasi ini. Meskipun demikian, masih terdapat kebutuhan akan pendampingan dan kerjasama tim untuk mendukung penggunaan yang efektif.

Kontrol perilaku yang dirasakan juga memengaruhi persepsi pengguna terhadap keamanan aplikasi. Beberapa informan merasa tidak aman dalam menggunakan aplikasi ini karena kesulitan dalam penggunaan, sementara yang lain merasa aman karena adanya fitur pengamanan seperti *user ID* dan *password*.

Saran

Berdasarkan hasil penelitian ini, beberapa saran dapat diajukan untuk meningkatkan penerimaan dan penggunaan aplikasi Siswaskeudes:

- a. Pengembangan Lanjutan: Perlu dilakukan pengembangan lebih lanjut terhadap aplikasi ini untuk meningkatkan fungsionalitasnya sesuai dengan kebutuhan pengguna.
- b. Pelatihan dan Pendampingan: Diperlukan pelatihan dan pendampingan lebih lanjut bagi pengguna, terutama pada tahap awal penggunaan, untuk memastikan pemahaman yang mendalam dan penggunaan yang efektif.
- c. Pemantauan dan Dukungan Tim: Perlunya pemantauan rutin oleh *vendor* dan dukungan tim yang baik untuk mendukung penggunaan aplikasi secara optimal serta menangani masalah yang muncul dengan cepat.
- d. Penguatan Keamanan: Perlu ditingkatkan lagi aspek keamanan dalam aplikasi ini untuk meningkatkan rasa aman pengguna dalam menggunakan aplikasi.

Penelitian selanjutnya dapat mengeksplorasi kepuasan pengguna dalam menggunakan aplikasi Siswaskeudes dengan menggunakan pendekatan kuantitatif. Penelitian dapat menggunakan teori TAM (*technology acceptance model*) yang diperluas dengan tambahan variabel seperti keamanan, kualitas layanan, *time saving*, *cost reduction* dan *energy saving*.

Daftar Pustaka

- [1] K. Fuad, Winarsih, L. M. Ifada, H. Setyawan, and R. T. Handayani, "An Analysis in the Application of the Unified Theory of Acceptance and Use of Technology (UTAUT) Model on Village Fund System (SISKEUDES) with Islamic Work Ethics as a Moderating Effect," in *Complex, Intelligent and Software Intensive Systems*, L. Barolli, K. Yim, and T. Enokido, Eds., Cham: Springer International Publishing, 2021, pp. 347–356. doi: 10.1007/978-3-030-79725-6_34.
- [2] M. Indriani, N. Nadirsyah, R. Fuadi, and N. Evayani, "Can Siskeudes support village governance in Aceh Province of Indonesia?," *Cogent Business & Management*, vol. 7, no. 1, p. 1743617, Jan. 2020, doi: 10.1080/23311975.2020.1743617.
- [3] I. Rosnidah, R. J. Johari, A. M. Musyaffi, R. Marota, and A. Muna, "E-Government Finance System Readiness for Village Government Employees: Lessons from Indonesia," *Quality - Access to Success*, vol. 23, no. 189, pp. 238–244, 2022, doi: 10.47750/QAS/23.189.27.
- [4] M. Fishbein, "A Theory of Reasoned Action: Some Applications and Implications.," 1979.
- [5] I. Ajzen, "The theory of planned behavior," *Organ Behav Hum Decis Process*, vol. 50, no. 2, pp. 179–211, 1991, doi: 10.1016/0749-5978(91)90020-T.
- [6] M. Purnama, N. P. S. H. Mimba, M. M. R. Sari, and D. Ariyanto, "Success Analysis of Flip: A Free Interbank Funds Transfer Mobile Application Using the UTAUT2 Model," *Jurnal Ilmiah Akuntansi*, vol. 8, no. 2, 2023.
- [7] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw, "User Acceptance of Computer Technology: A Comparison of Two Theoretical Models," *Manage Sci*, vol. 35, no. 8, pp. 982–1003, 1989, doi: 10.1016/j.system.2015.06.005.
- [8] C. Pramatha, I. M. Y. Mahendra, G. P. W. Rajeg, and I. W. Arka, "The Development of Semantic Dictionary Prototype for the Balinese Language," *International Journal of Cyber and IT Service Management*, vol. 3, no. 2, pp. 96–106, 2023.
- [9] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Q*, vol. 13, no. 3, pp. 319–339, 1989, doi: 10.2307/249008.
- [10] W. L. Newman, *Social research methods: Qualitative and quantitative approaches*. Allyn and Bacon, 1991.
- [11] J. W. Creswell and C. N. Poth, *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications, 2016.
- [12] J. W. Creswell, R. Shope, V. L. Plano Clark, and D. O. Green, "How interpretive qualitative research extends mixed methods research," *Research in the Schools*, vol. 13, no. 1, pp. 1–11, 2006.
- [13] M. Pokropski, "Phenomenology and functional analysis. A functionalist reading of Husserlian phenomenology," *Phenomenol Cogn Sci*, vol. 19, no. 5, pp. 869–889, 2020.

- [14] L. Tindall, "JA Smith, P. Flower and M. Larkin (2009), *Interpretative Phenomenological Analysis: Theory, Method and Research*. London: Sage." Taylor & Francis, 2009.
- [15] N. H. U. Dewi and P. Diptyana, "Makna Pengendalian Internal dalam Perspektif Lembaga Amil Zakat, Infaq, Shodaqoh: Pendekatan Paradigma Interpretif," *Jurnal Riset Akuntansi & Perpajakan (JRAP)*, vol. 9, no. 1, pp. 41–51, 2022.
- [16] A. Alase, "The Interpretative Phenomenological Analysis (IPA): A Guide to a Good Qualitative Research Approach.," *International journal of education and literacy studies*, vol. 5, no. 2, pp. 9–19, 2017.
- [17] I. Pietkiewicz and J. A. Smith, "A Practical Guide to Using Interpretative Phenomenological Analysis in Qualitative Research Psychology," *Psychological journal*, vol. 20, no. 1, pp. 7–14, 2014.
- [18] J. W. Creswell and J. D. Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Sage publications, 2017.

Real Time Object Detection Menggunakan Mobilenet-SSD pada Sistem Keamanan Ruangan dengan Bot Telegram Sebagai Notifikasi User

Afit Miranto

Program Studi Teknik Elektro, Fakultas Teknologi Industri, Institut Teknologi Sumatera
Lampung Selatan, Indonesia
afit.miranto@el.itera.ac.id

Abstract

CCTV memainkan peran krusial dalam menjaga keamanan suatu lingkungan. Namun, penggunaan CCTV konvensional menghasilkan penggunaan storage yang cukup banyak karena tanpa pengenalan objek CCTV akan selalu merekam kejadian selama sistem menyala. Untuk meningkatkan efisiensi, dibutuhkan sebuah sistem yang dapat mendeteksi objek yang presisi dan hanya memberikan peringatan ketika ada objek yang dikenali. Penelitian ini berfokus pada perancangan sistem otomatis dengan menggunakan Raspberry Pi 4 sebagai pusat kendali untuk mengelola kamera keamanan dalam pendeteksian objek dan juga penyimpanan data. Sistem ini juga memanfaatkan database penghuni rumah untuk membandingkan dan memberikan notifikasi jika mendeteksi kehadiran yang tidak sesuai. Dengan merekam hanya saat terdeteksi pergerakan manusia, sistem mengurangi beban penyimpanan. Solusi ini memberikan pemantauan ruangan yang efisien dan notifikasi melalui aplikasi Telegram, mengatasi kendala penggunaan storage yang berlebihan. Metode yang digunakan dalam pendeteksian objek adalah dengan menggunakan mobilenet-SSD v2. Hasil yang diperoleh adalah sistem mampu mengenali objek manusia dan akan mengirimkan pesan notifikasi dan berupa capture gambar jika objek yang terdeteksi tidak dikenali.

Keywords: CCTV, Deteksi Objek, Mobilenet-SSD, Pengolahan Citra, Sistem Keamanan,

1. Pendahuluan

Peran CCTV dalam keamanan suatu lingkungan sangat penting, sehingga saat ini dalam perumahan maupun kantor sebagian besar memiliki CCTV yang dapat memantau keberadaan suatu lingkungan tersebut, baik dalam kondisi ada maupun tidak ada kejahatan yang terjadi dalam lingkungan tersebut [1]. Pada umumnya CCTV hanya merekam kejadian atau kondisi lingkungan sepanjang waktu tanpa harus mengenali kondisi yang terjadi, baik objek manusia maupun lainnya. Akibatnya penggunaan *storage*/penyimpanan data akan sangat boros walaupun tanpa ada objek yang terdeteksi.

Deteksi objek merupakan hal sangat penting dalam hal keamanan dengan menggunakan kamera/citra digital [2]. Tingkat ketelitian dalam mendeteksi objek menentukan kualitas dari sistem yang dirancang. Oleh karena itu dibutuhkan sebuah sistem otomatis yang bisa melakukan manajemen penggunaan *storage* dari kamera keamanan tersebut dengan sensor kamera sebagai sumber inputnya dan Raspberry Pi 4 sebagai pusat kendali sistem. Sistem tersebut diotomasi menggunakan kamera sebagai sensor yang mendeteksi objek. Dengan memasukkan *database* berdasarkan orang yang terdata didalam rumah tersebut, jika kamera mendeteksi kondisi ruangan yang penghuninya tidak sesuai atau tidak ada dalam *database*, maka sistem akan memberikan notifikasi dan alarm. Sebagai bentuk manajemen *storage*/penyimpanan, sistem hanya akan memberikan perintah *record/capture* hanya ketika ada pergerakan seseorang didalam ruangan tersebut, sehingga data penyimpanan tidak akan cepat penuh karena selalu menyimpan video/gambar.

Real Time Object Detection Menggunakan Mobilenet-SSD pada Sistem Keamanan Ruangn dengan Bot Telegram Sebagai Notifikasi User

Sistem ini dibangun memberikan solusi atas keamanan dalam suatu ruangan sehingga dapat memantau keadaan suatu ruangan dengan memberikan notifikasi peringatan kepada pengguna melalui aplikasi telegram. Selain itu juga penggunaan storage tidak terlalu membebani sistem karena sistem yang dirancang tidak selalu merekam atau meng-*capture* kondisi dalam suatu ruangan.

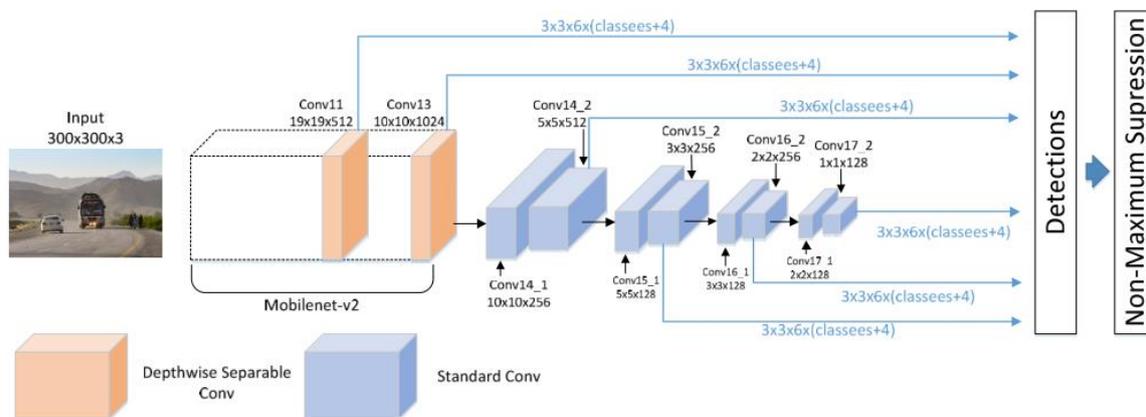
2. Metode Penelitian

Keberadaan kamera pengawas sangat berperan penting untuk mengawasi aktivitas pada suatu area/lingkungan tertentu. Kamera keamanan umumnya bersifat pasif, artinya hanya melihat dan merekam aktifitas sehingga tidak selalu diketahui apakah kondisi lingkungan tersebut aman atau tidak [1]. Sehingga diperlukan sebuah sistem yang tidak hanya memantau namun dapat memberikan notifikasi terhadap kondisi ruang jika orang yang terdeteksi tidak dikenali dan hanya akan merekam kejadian hanya saat ada seseorang didalam area yang terlihat pada kamera sehingga akan menghemat penggunaan storage pada sistem [3].

2.1. MobileNet-SSD

MobileNet-SSD adalah sebuah arsitektur jaringan saraf tiruan (*deep neural network*) yang dikembangkan untuk deteksi objek pada gambar secara real-time, khususnya untuk aplikasi yang berjalan pada perangkat mobile dengan sumber daya terbatas. Arsitektur ini menggabungkan antara MobileNet, sebuah arsitektur jaringan saraf tiruan yang ringan dan efisien secara komputasi, dengan *Single Shot MultiBox Detector* (SSD), yang merupakan sebuah pendekatan untuk deteksi objek yang cepat dan akurat. MobileNet digunakan sebagai bagian dari arsitektur untuk mengekstraksi fitur-fitur dari gambar, sementara SSD digunakan untuk mendeteksi objek dengan menghasilkan beberapa kotak pembatas (*bounding boxes*) dan mengevaluasi kelas serta probabilitas keberadaan objek di dalamnya. Gabungan dari keduanya menghasilkan arsitektur yang efisien secara komputasi dan mampu mendeteksi objek dengan kecepatan tinggi. Sehingga Metode MobileNet-SSD dengan arsitektur yang dipakai *convolutional neural network* (CNN) dapat mengatasi kebutuhan akan *computing resource* yang berlebih [2], [4], [5], [6], [7], [8]. Untuk dapat menjalankan sistem dengan menggunakan arsitektur ini, maka diperlukan *single board computer* yaitu raspberry pi dan kamera sebagai sensor deteksi objek, untuk mengolah data yang diterima kamera tersebut [9]. MobileNet-SSD merupakan arsitektur untuk pendeteksian object yang cukup ringan [2].

MobileNet-SSD V2 merupakan versi kedua dari ekstraktor fitur dan sangat efektif digunakan untuk deteksi objek. Model deteksi ini 35% lebih akurat dibandingkan MobileNetV1 [10] [11][2].



Gambar 1. Arsitektur mobilenet-SSDv2 [2]

2.2. Face Recognition

Pengenalan wajah merupakan teknologi yang dapat dipakai untuk mengidentifikasi atau memverifikasi identitas seseorang dari gambar wajah mereka. Ini merupakan salah satu aplikasi dari pemrosesan

citra dan pengenalan pola yang telah berkembang sangat pesat dalam beberapa tahun ini, terutama berkat kemajuan dalam *deep learning* dan jaringan saraf tiruan.

Teknologi pengenalan wajah biasanya melibatkan beberapa langkah:

Deteksi Wajah: Langkah pertama adalah mendeteksi wajah dalam gambar atau video. Ini dapat dilakukan dengan menggunakan algoritma deteksi wajah seperti *Viola-Jones* atau menggunakan pendekatan yang lebih canggih seperti *deep learning*.

Ekstraksi Fitur: Setelah wajah terdeteksi, langkah berikutnya adalah mengekstraksi fitur-fitur kunci dari wajah tersebut, seperti bentuk mata, hidung, dan mulut. Ini bisa dilakukan dengan berbagai metode.

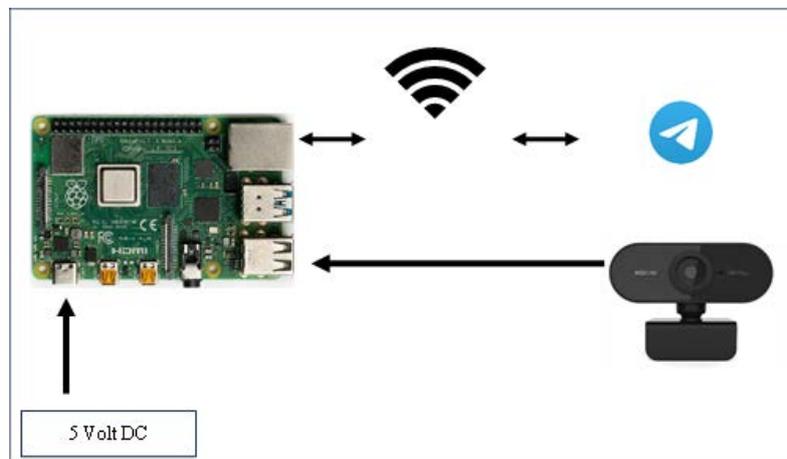
Pembuatan Model: Setelah fitur-fitur kunci diekstraksi, model pembelajaran mesin atau *deep learning* dilatih untuk mengidentifikasi individu berdasarkan fitur-fitur ini. Ini bisa menjadi model klasifikasi yang mempelajari perbedaan antara berbagai wajah dan memutuskan identitas seseorang berdasarkan fitur-fitur yang ada.

Pengenalan dan Verifikasi: Dalam pengenalan wajah, sistem membandingkan fitur-fitur yang diekstraksi dari wajah yang diproses dengan data wajah yang ada dalam basis data. Dalam verifikasi wajah, sistem memverifikasi apakah individu tersebut adalah orang yang dia klaim dengan membandingkan wajah yang diidentifikasi dengan wajah yang telah diotorisasi sebelumnya [12][13].

3. Hasil dan pembahasan

3.1. Hasil Perancangan Prototipe Sistem

Prototipe sistem yang dibangun menggunakan Raspberry Pi 4 sebagai pusat kendali dan webcam sebagai sensor yang mendeteksi objek. Gambar 5 Berikut merupakan rangkaian sistem yang dibangun.



Gambar 2. Blok diagram sistem deteksi objek yang dibangun

Berdasarkan blok diagram sistem pada Gambar 2 dapat dijelaskan bahwa kamera akan menangkap gambar kondisi pada suatu ruangan yang ada kemudian hasil deteksi akan diolah oleh Raspberry pi 4. Dengan menggunakan metode MobileNet-SSD sistem akan mengenali bahwa yang dideteksi tersebut adalah objek seorang manusia. MobileNet-SSD (*single shot detector*) merupakan model deteksi objek yang dapat memberikan inferensi secara *real time* dengan batasan komputasi yang bisa digunakan tidak perlu spesifikasi yang tinggi. Setelah dilakukan proses training terkait objek yang akan dideteksi, model deteksi dengan menggunakan MobileNet-SSD dapat digunakan untuk perangkat komputasi kecil seperti Raspberry Pi 4.

Real Time Object Detection Menggunakan Mobilenet-SSD pada Sistem Keamanan Ruang dengan Bot Telegram Sebagai Notifikasi User



Gambar 3. Prototype sistem deteksi objek yang dibangun

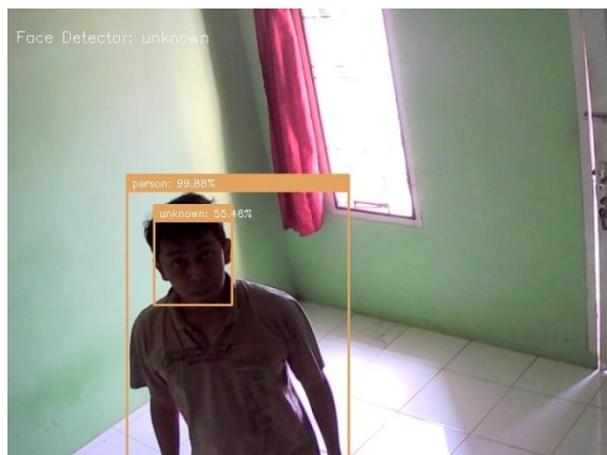
Hasil implementasi dari sistem yang dibangun dari blok diagram pada Gambar 2. Hasil implementasi yang dilakukan berupa prototipe yang dapat dilihat pada Gambar 3. Sistem yang terdiri dari komputer Raspberry pi 4 dan webcam yang diberikan supply daya dengan tegangan 5 Volt 3 Ampere ini mampu untuk menjalankan sistem yang dibangun dengan cukup baik.

3.2. Hasil Pengujian Deteksi Objek

Berikut merupakan beberapa hasil pengujian sistem yang telah dibangun. Pada Gambar 4 status ruangan belum ada objek yang terdeteksi. sehingga sistem tidak mengirimkan atau memberikan notifikasi peringatan.



Gambar 4. Status ruangan saat tidak terdeteksi objek



Gambar 5. Hasil Deteksi Objek yang dikirimkan melalui aplikasi Telegram

Berdasarkan pada Gambar 5 dapat dilihat bahwa sistem yang dibangun sudah mampu untuk mendeteksi objek dan dapat mengirimkan notifikasi ke pengguna melalui aplikasi telegram dan *capture* gambar objek yang dideteksi dengan sangat baik.

Selain mendeteksi objek manusia hasil deteksi wajah yang telah dilakukan juga cukup baik berdasarkan database yang telah dimasukkan sebelumnya rata-rata hasil pengenalan wajah diatas 70%. Tingkat akurasi objek yang diterima dipengaruhi oleh beberapa hal yaitu intensitas cahaya yang diterima oleh kamera harus baik, karena jika pencahayaan minim, maka akan sangat mengurangi akurasi dari pendeteksian objek, serta wajah yang terlihat kamera harus terlihat jelas jika tidak maka akan mengurangi persentase pembacaan wajah.

Tabel 1. Tabel hasil pengujian deteksi objek

Pengujian ke-	Persentase hasil deteksi objek manusia (%)	Persentase pengenalan wajah (%)	Ukuran file (KB)
1	99.85	77.21	91.6
2	99.84	61.88	88.2
3	99.88	75.87	91.0
4	99.63	64.33	78.4
5	99.64	76.58	104
Rata-rata	99.76	71.17	90.64

Berdasarkan pada Tabel 1 dapat dilihat bahwa persentase keberhasilan objek dalam mendeteksi objek manusia sangat baik yaitu diatas 90% dan ukuran file yang dikirimkan sangat kecil yaitu rata-rata dibawah 90Kb ini sangat kecil jika dibandingkan sistem yang harus merekam setiap waktu walau tanpa ada sesuatu yang perlu direkam.

4. Kesimpulan

Kegiatan penelitian telah terlaksana dengan baik dan hasil yang diperoleh adalah sistem mampu mendeteksi objek dengan sangat baik dengan tingkat akurasi tinggi yaitu diatas 90% untuk mendeteksi objek manusia dan juga mampu mengirimkan notifikasi alarm serta bukti hasil *capture* objek yang terdeteksi melalui aplikasi telegram.

Adapun kekurangan yang perlu diperhatikan adalah semakin banyak algoritma yang digunakan dan dimasukkan kedalam sistem maka akan semakin membebani perangkat Raspberry Pi 4 yang digunakan sehingga akan terasa sedikit lag jika dilihat pada layar monitor. Untuk pendeteksian wajah masih kurang optimal karena wajah harus terlihat jelas dikamera, jika tidak maka akan mengurangi keberhasilan. Sehingga perlu dilakukan penelitian lebih lanjut untuk ekstraksi wajah agar semakin optimal dalam pendeteksian objek dan wajah secara bersamaan.

Daftar Pustaka

- [1] A. Miranto, S. R. S. R. Sulistiyanti, and F. X. X. Arinto Setyawan, "Adaptive background subtraction for monitoring system," in *2019 International Conference on Information and Communications Technology, ICOIACT 2019*, 2019, pp. 153–156. doi: 10.1109/ICOIACT46704.2019.8938501.
- [2] Y. C. Chiu, C. Y. Tsai, M. Da Ruan, G. Y. Shen, and T. T. Lee, "Mobilenet-SSDv2: An Improved Object Detection Model for Embedded Systems," *2020 International Conference on System Science and Engineering, ICSSE 2020*, pp. 0–4, 2020, doi: 10.1109/ICSSE50014.2020.9219319.
- [3] S. Sruthy, S. Yamuna, and S. N. George, "An IoT based Active Building Surveillance System using Raspberry Pi and NodeMCU," *ArXiv*, pp. 1–9, 2020.
- [4] A. Younis, L. Shixin, J. N. Shelembi, and Z. Hai, "Real-time object detection using pre-trained deep learning models mobilenet- SSD," *ACM International Conference Proceeding Series*, no. March, pp. 44–48, 2020, doi: 10.1145/3379247.3379264.
- [5] D. Biswas, H. Su, C. Wang, A. Stevanovic, and W. Wang, "An automatic traffic density estimation using Single Shot Detection (SSD) and MobileNet-SSD," *Physics and Chemistry of the Earth*, vol. 110, no. December, pp. 176–184, 2019, doi: 10.1016/j.pce.2018.12.001.
- [6] I. G. Fernandez and C. Wada, "Shoe Detection Using SSD-MobileNet Architecture," *LifeTech 2020 - 2020 IEEE 2nd Global Conference on Life Sciences and Technologies*, no. LifeTech, pp. 171–172, 2020, doi: 10.1109/LifeTech48969.2020.1570618965.
- [7] W. Sun, S. Chen, L. Shi, Y. Li, and Z. Lin, "Vehicle Following in Intelligent Multi-Vehicle Systems Based on SSD-MobileNet," *Proceedings - 2019 Chinese Automation Congress, CAC 2019*, pp. 5004–5009, 2019, doi: 10.1109/CAC48633.2019.8996181.
- [8] A. Heredia and G. Barros-Gavilanes, "Video processing inside embedded devices using SSD-MobileNet to count mobility actors," *2019 IEEE Colombian Conference on Applications in Computational Intelligence, CoICACI 2019 - Proceedings*, pp. 1–6, 2019, doi: 10.1109/CoICACI.2019.8781798.
- [9] R. A. Nadafa, S. M. Hatturea, V. M. Bonala, and S. P. Naikb, "Home Security against Human Intrusion using Raspberry Pi," *Procedia Comput Sci*, vol. 167, no. Iccids 2019, pp. 1811–1820, 2020, doi: 10.1016/j.procs.2020.03.200.
- [10] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," Jan. 2018, [Online]. Available: <http://arxiv.org/abs/1801.04381>
- [11] A. G. Howard *et al.*, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," Apr. 2017, [Online]. Available: <http://arxiv.org/abs/1704.04861>
- [12] A. Raghunandan, Mohana, P. Raghav, and H. V. R. Aradhya, "Object Detection Algorithms for Video Surveillance Applications," in *2018 International Conference on Communication and Signal Processing (ICCSP)*, 2018, pp. 563–568. doi: 10.1109/ICCSP.2018.8524461.
- [13] G. Chandan, A. Jain, H. Jain, and Mohana, "Real Time Object Detection and Tracking Using Deep Learning and OpenCV," *Proceedings of the International Conference on Inventive Research in Computing Applications, ICIRCA 2018*, no. Icirca, pp. 1305–1308, 2018, doi: 10.1109/ICIRCA.2018.8597266.

Klasifikasi Kualitas Air Layak Minum menggunakan Algoritma *Random Forest Classifier* dan *GridsearchCV*

Gusti Agung Diah Sri Ari Ningsih^{a1}, Cokorda Pramatha^{a2}

^aProgram Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Udayana, Bali
Jln. Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, 08261, Bali, Indonesia
¹gstadiahs@gmail.com
²cokorda@unud.ac.id

Abstract

Drinkable water is water that is healthy for humans to drink and does not pose significant health risks. To determine whether water has a quality that meets health standards can be determined through the substances or minerals contained in it. Conventional methods require quite a long time to evaluate and classify water quality as suitable for consumption or not. One approach that can be used to overcome this problem is to utilize machine learning. This research uses a random forest to carry out classification. Using random forest by default cannot produce optimal performance because the parameters used are not necessarily the best. Therefore, this research also uses GridsearchCV to find optimal hyperparameter values in the Random Forest Classifier. After hyperparameter tuning, an optimal model was obtained with each parameter $n_estimators$ 100, max_depth 9, $max_features$ 4, and $min_samples_split$ 2. The performance of Random Forest after hyperparameter tuning increased accuracy, which was initially 76% increase to 84%, precision which was initially 76.19% increase to 81.70%, recall which was initially 74.89% increase to 85.53%, and f1-score which was initially 75.53%, increase to 83.57%.

Keywords: Classification, Drinking Water Quality, Random Forest, Optimal Hyperparameters, Hyperparameter Tuning, GridSearchCV

1 Pendahuluan

Pesatnya pembangunan dan pertumbuhan penduduk menyebabkan terjadinya penurunan kualitas air. Hal ini dikarenakan semakin meningkatnya aktivitas manusia dan kompleksitas industri akan mengakibatkan tingkat pencemaran semakin meningkat yang berdampak pada keseimbangan perairan, kebutuhan sanitasi yang memadai dan akses terhadap air minum sehat untuk konsumsi manusia [3]. Berdasarkan data *World Health Organization* (WHO) sebanyak 663 juta penduduk kesulitan dalam memperoleh air bersih [4]. *World Health Organization* (WHO) menjelaskan bahwa setiap tahunnya sebanyak 829 ribu orang meninggal karena air minum, sanitasi, dan kebersihan tangan yang tidak aman [5]. Setiap tahun sekitar dua juta bayi kehilangan nyawa dikarenakan kurangnya akses air minum yang bersih [6]. Perlu dilakukan suatu prediksi untuk mengetahui air yang akan dikonsumsi layak atau tidak.

Keberadaan zat-zat yang terdapat didalam air dapat menjadi indikator untuk menentukan kualitas air. Umumnya penilaian kualitas air menggunakan metode perhitungan manual seperti menghitung *Water Quality Index* (WQI) [2]. *Water Quality Index* (WQI) merupakan suatu metode untuk menilai kualitas air dengan mempertimbangkan beberapa karakteristik yang merepresentasikan kondisi air [7]. Terdapat sembilan faktor yang digunakan dalam perhitungan WQI secara tradisional meliputi PH, *Hardness*, *Solids*, *Chloramines*, *Sulfate*, *Conductivity*, *Organic_Carbon*, *Trihalomethanes*, dan *Turbidity*. *Water Quality Index* (WQI) dibuat berdasarkan pada standar yang direkomendasikan WHO [8]. Metode konvensional bergantung pada pemahaman terhadap parameter-parameter yang telah ditetapkan untuk mengevaluasi dan mengategorikan kualitas air sebagai air yang aman atau tidak untuk dikonsumsi. Metode konvensional seringkali memerlukan waktu lama dalam perhitungannya, sehingga membutuhkan sistem yang dapat bekerja secara otomatis.

Pemanfaatan algoritma *machine learning* dapat diterapkan untuk mengatasi tantangan ini. *Machine learning* memiliki dua aplikasi utama yaitu klasifikasi dan prediksi. Klasifikasi adalah

proses pemberian label atau kategori tertentu yang sudah ditetapkan sebelumnya pada objek dalam data yang akan diklasifikasikan [9]. Dalam kasus ini klasifikasi digunakan untuk menentukan apakah air layak untuk dikonsumsi atau tidak. Penelitian sebelumnya mengenai klasifikasi kualitas air minum telah dilakukan oleh [4] dengan membandingkan berbagai algoritma klasifikasi seperti *Logistic Regression*, *Support Vector Machine*, *K-NN*, *Decision Trees*, *Naïve Bayes*, *Random Forest Classifier*, *XGBosst Classifier*, dan *Artificial Neural Network* dalam menilai apakah air dapat dikonsumsi atau tidak dapat dikonsumsi, hasil menunjukkan bahwa model klasifikasi *Random Forest Classifier* mencapai akurasi tertinggi sebesar 72,81%. Penggunaan *random forest* secara default yaitu tanpa dilakukan optimasi tidak dapat mencapai kinerja optimal dikarenakan belum tentu parameter yang digunakan adalah parameter yang paling optimal pada model [10].

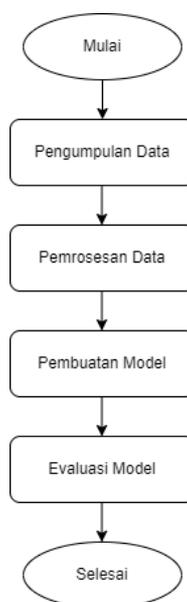
Mengatur *hyperparameter* atau *hyperparameter tuning* merupakan salah satu metode untuk meningkatkan performa model klasifikasi [11]. Hal ini dikarenakan kinerja *random forest* sangat dipengaruhi oleh pengaturan *hyperparameter* yang dipilih. Optimasi akan memberikan hasil yang lebih akurat [12]. Berdasarkan penelitian yang dilakukan oleh [13] memperoleh hasil akurasi setelah dioptimasi 94,8%, akurasi mengalami peningkatan sebesar 21,5% dari 73,31%. Selain itu penelitian oleh [10] yang memanfaatkan algoritma *random forest* untuk mengklasifikasikan malware dan melakukan penyetelan *hyperparameter* menggunakan *GridsearchCV* berhasil meningkatkan kinerja model dengan mencapai tingkat akurasi sebesar 99,23%. Mengacu pada hasil penelitian sebelumnya maka dilakukan penelitian dengan judul *Klasifikasi Kualitas Air Layak Minum dengan Algoritma Random Forest Classifier dan GridsearchCV*

2 Metode Penelitian

Metode penelitian merupakan serangkaian langkah yang dilakukan untuk menyelesaikan masalah selama proses penelitian dengan mematuhi aturan dan pedoman yang telah ditetapkan oleh peneliti untuk mencapai tujuan yang diinginkan serta memperoleh hasil yang optimal [14]. Penelitian ini terdiri dari empat tahapan utama meliputi

1. Pengumpulan data
2. Pemrosesan data
3. Pembuatan model
4. Evaluasi model

Tahapan penelitian ini ditunjukkan pada Gambar 1



Gambar 1. Alur Penelitian

Pada Gambar 1 menunjukkan alur penelitian yang menggambarkan tahapan-tahapan yang dilakukan pada penelitian ini. Tahapan penelitian pada penelitian ini terdiri dari tahap pengumpulan data, pemrosesan data, pembuatan model dan evaluasi model. Penjelasan rinci mengenai alur tersebut disajikan pada bagian berikut

2.1 Pengumpulan Data

Data pada penelitian ini menggunakan data sekunder yang diambil dari *Water Quality* Dataset yang diperoleh melalui website kaggle.com dengan jumlah data 3276 dan 10 atribut. Dataset ini memiliki 2 kelas output yaitu nilai 1 yang artinya air dapat dikonsumsi dan nilai 0 artinya air tidak dapat dikonsumsi. Berikut merupakan tabel atribut dan domain yang digunakan

Tabel 1. Atribut dan Domain Dataset

Atribut	Deskripsi	Domain
Ph	Indikator yang dipakai untuk menilai derajat keasaman atau kebasaaan larutan.	0 – 14
Hardness	Konsentrasi garam-garam mineral tertentu, terutama kalsium dan magnesium yang larut dalam air.	47,432 – 323,124
Solids	Jumlah garam anorganik dan kandungan bahan organik yang terlarut dalam air.	320,94 – 61227,19
Chloramines	Kandungan kimia yang dihasilkan dari interaksi klorin dalam air dan amonia atau senyawa nitrogen lainnya.	0,352 – 13,127
Sulfate	Bahan alami yang terdapat dalam mineral, tanah, dan batuan.	129 – 481.03
Conductivity	Kadar ion positif (kation) dan ion negatif (anion) dalam air.	181,4 – 753,3
Organic Carbon	Materi yang berasal dari senyawa yang terurai dalam air.	2,2 – 28,3
Trihalomethanes	Senyawa yang terbentuk dalam air minum sebagai hasil dari penambahan klorin ke zat organik.	0,738 – 124
Turbidity	Ukuran kekeruhan atau kekeruhan relatif air yang disebabkan oleh partikel-paditersuspensi yang terdapat dalam air.	1,45 – 6,739
Potability	Indikator apakah air tersebut layak untuk diminum atau tidak.	0 – 1

Pada tabel 1 menjelaskan 10 atribut yang ada didalam dataset beserta deskripsi dan rentang nilainya. Atribut-atribut tersebut antara lain pH, *hardness*, *solids*, *chloramines*, *sulfate*, *conductivity*, *organic carbon*, *trihalomethanes*, *turbidity*, dan *potability*. Deskripsi menjelaskan setiap atribut, sementara kolom domain menunjukkan rentang nilai yang ada pada setiap atribut

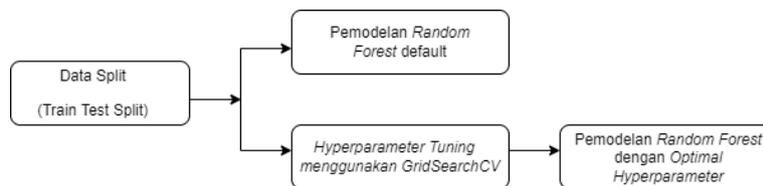
2.2 Pemrosesan Data

Tahapan pemrosesan data meliputi penanganan data yang memiliki nilai null atau kosong dengan cara menghapus nilai null tersebut. Pemrosesan data selanjutnya adalah *feature scaling* untuk mengubah rentang atau skala nilai dari fitur-fitur dalam dataset ke rentang nilai baru dengan menerapkan suatu metode atau pendekatan tertentu sehingga memiliki karakteristik yang serupa antar fitur. Metode untuk melakukan *feature scaling* pada penelitian ini adalah standarisasi. Standarisasi mengubah nilai-nilai fitur sehingga memiliki rata-rata nol dan deviasi standar satu. Selain itu tahapan pemrosesan data yang dilakukan yaitu *resampling* untuk menyeimbangkan kelas data sehingga menghasilkan dataset baru yang memiliki jumlah sampel yang seimbang antara kelas mayoritas dan kelas minoritas. Proses *resampling* pada penelitian ini diawali dengan pengidentifikasian kelas minoritas yang umumnya memiliki jumlah sampel yang lebih sedikit. Kemudian, dengan menggunakan teknik *resampling*, sampel-sampel dari kelas minoritas dipilih secara acak. Penggunaan *resampling* dengan penggantian atau *replacement* memungkinkan

sampel yang sama untuk dipilih beberapa kali, sehingga meningkatkan jumlah sampel kelas minoritas.

2.3 Pembuatan Model

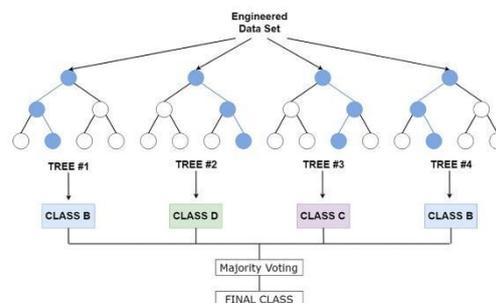
Pada penelitian ini model dibangun dengan dua kondisi. Model pertama yaitu *Random Forest* sebelum *hyperparameter tuning* atau menggunakan nilai default. Model kedua yaitu *Random Forest* dengan menggunakan *GridSearchCV*. Model kedua yakni *Random Forest* dengan optimal parameter yang ditemukan melalui *hyperparameter tuning* menggunakan *GridSearchCV*. Pada model kedua, terlebih dahulu dilakukan *hyperparameter tuning* menggunakan *GridSearchCV* terlebih dahulu, selanjutnya *Random Forest* dibangun dengan hyperparameter optimal yang sudah ditemukan pada proses *hyperparameter tuning*



Gambar 2. Tahapan pembuatan model

Pada Gambar 2 menunjukkan tahapan pembuatan dua model yaitu model pertama yang menggunakan parameter default dan model kedua yang menggunakan parameter optimal setelah proses *hyperparameter tuning*. Tahapan ini dilakukan untuk membandingkan performa model sebelum dan sesudah hyperparameter tuning, sehingga dapat melihat peningkatan yang dihasilkan. Berikut penjelasan untuk setiap tahap pembuatan model yang ditunjukkan pada gambar 2

- a. Split data (Train test split)
 Dalam tahap pembagian data, dataset dibagi menjadi data pelatihan dan data pengujian. Pembagian ini dilakukan dengan rasio 80% untuk data pelatihan dan 20% untuk data pengujian (80:20).
- b. Pemodelan *Random Forest* default
Random Forest (RF) adalah algoritma klasifikasi *ensemble* yang menggunakan banyak pohon keputusan (*Decision Trees*). Setiap pohon keputusan memberikan prediksi, dan prediksi akhir untuk suatu contoh pengujian dihasilkan dari gabungan prediksi semua pohon individu [15]. Proses klasifikasi dalam *random forest* dimulai dengan membagi data, lalu secara acak mengambil sampel untuk dianalisis dalam setiap pohon keputusan. Setelah pohon-pohon terbentuk, setiap kelas dari sampel data dipertimbangkan, kemudian dilakukan pemungutan suara untuk menentukan kelas yang paling sering muncul sebagai prediksi akhir [16].



Gambar 3. Diagram alur *Random Forest* [16]

Pada Gambar 3 menunjukkan diagram alur *Random Forest* yang terdiri dari beberapa *decision tree* yang dibuat dari subset data yang berbeda. Prediksi akhir ditentukan melalui *majority voting*. *Majority voting* adalah metode dalam *Random Forest* di mana keputusan akhir untuk klasifikasi didasarkan pada suara terbanyak dari semua *decision*

tree [15]. Jika mayoritas *decision tree* memprediksi kelas tertentu, maka kelas tersebut akan menjadi hasil prediksi akhir

Berdasarkan *scikit-learn* nilai parameter default *random forest* yang dimaksud adalah jumlah *n_estimators* 100, *max_features* menggunakan *sqrt*, *max_depth* menggunakan *None* dan jumlah *min_samples_split* 2. *n_estimators* yaitu jumlah pohon keputusan. *max_feature* yaitu jumlah fitur acak yang diambil ketika mencari simpul pada *tree*. *min_samples_split* yaitu parameter yang menentukan jumlah minimum sampel yang diperlukan di sebuah node agar node tersebut dapat dibagi selama proses pembangunan pohon.

- c. Pemodelan *Random Forest* menggunakan *Optimal Hyperparameter*
Optimal hyperparameter ditentukan melalui proses *hyperparameter tuning*. *Hyperparameter tuning* merupakan pendekatan yang dapat digunakan dengan mencoba berbagai kombinasi parameter untuk mengevaluasi kinerja masing-masing modelnya [14]. *Grid search* adalah sebuah proses memilih kombinasi nilai dari *hyperparameter* yang digunakan dengan mencoba berbagai kombinasi *hyperparameter* dan mengevaluasi masing-masing kombinasi. *Grid search* digunakan untuk menemukan kombinasi terbaik dari *hyperparameter* untuk suatu model[17].

Grid search cross-validation adalah proses pengujian beberapa kombinasi nilai *hyperparameter* secara simultan dengan menggunakan proses validasi silang [18]. *K-fold cross-validation* adalah salah satu metode *cross-validation* yang memungkinkan data pelatihan dan pengujian diulang hingga k iterasi menggunakan 1/k bagian dari kumpulan data sebagai data pengujian [19]. Algoritma *Random Forest* memiliki banyak *hyperparameter* yang dapat dioptimasi. Penelitian ini berfokus pada beberapa *hyperparameter* dari *random forest* seperti *n_estimators*, *max_features*, *max_depth*, dan *min_samples_split*. Tabel 2 menampilkan daftar parameter yang diuji menggunakan *GridSearchCV* [20] [21] [22]

Tabel 2. Daftar Parameter Model

Parameter	Nilai
<i>n_estimators</i>	50, 100, 300
<i>max_features</i>	2, 3, 4
<i>max_depth</i>	3, 5, 7, 9
<i>min_samples_split</i>	2, 5, 8

Pada Tabel 2 menunjukkan rentang parameter yang diuji. Parameter-parameter ini akan digunakan dalam proses pengujian untuk menentukan konfigurasi terbaik pada model. *Hyperparameter tuning* untuk model *Random Forest* dilakukan dengan menggunakan *Grid Search Cross-Validation* yang menggunakan k-fold sebanyak 5. Setelah menemukan *n_estimators*, *max_features*, *max_depth*, dan *min_samples_split* yang optimal untuk model *Random Forest*, langkah berikutnya adalah menggunakan *hyperparameter* tersebut untuk membangun model klasifikasi *Random Forest*.

2.4 Evaluasi Model

Confusion matrix digunakan sebagai metode evaluasi untuk mengukur kinerja proses klasifikasi dari model dengan membandingkan prediksi yang dihasilkan oleh model dengan data yang sebenarnya dari data yang diuji.

Tabel 3. Confusion Matrix

	Nilai Prediksi	
Nilai Aktual	Positif(1)	Negatif(0)
Positif (1)	<i>True Positive</i> (TP)	<i>False Negative</i> (FN)
Negatif(0)	<i>False Positive</i> (FP)	<i>True Negative</i> (TN)

Pada tabel 3 menampilkan *confusion matrix* yang menunjukkan bagaimana prediksi model sesuai dengan nilai aktual. *True Positive* (TP) yaitu banyak data yang secara benar terklasifikasi sebagai positif. *False Positif* (FP) yaitu banyak data yang sebenarnya negatif tetapi terklasifikasi sebagai positif. *False Negative* (FN) yaitu banyak data yang sebenarnya positif tetapi terklasifikasi sebagai negatif. *True Negative* (TN) yaitu banyak data yang sebenarnya negatif yang secara benar terklasifikasi sebagai negatif. Dalam mengevaluasi kinerja model, performa yang digunakan diantaranya *Accuracy*, *Precision*, *Recal* dan *F1-Score*.

a. *Accuracy*

Akurasi adalah persentase data yang diklasifikasikan dengan benar setelah diuji. Nilai *Accuracy* dapat diperoleh dengan menggunakan persamaan

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \times 100\% \quad (1)$$

b. *Precision*

Precision adalah persentase dari data positif yang diprediksi dengan benar dari total data yang diprediksi sebagai positif. Nilai *Precision* dapat diperoleh dengan menggunakan persamaan

$$Precision = \frac{TP}{TP+FP} \times 100\% \quad (2)$$

c. *Recall*

Recall merupakan persentase dari data positif yang diprediksi dengan benar dari total data positif yang sebenarnya. Nilai *Recall* dapat diperoleh dengan menggunakan persamaan

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (3)$$

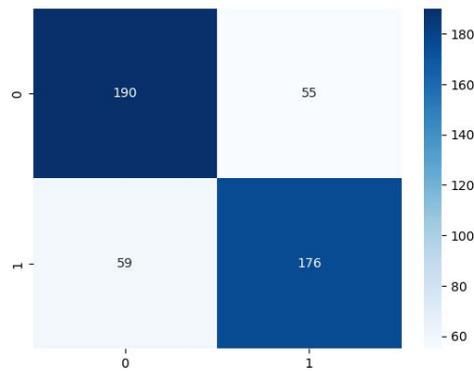
d. *F1-Score*

F1-score adalah metrik yang mengevaluasi hasil dari perhitungan *precision* dan *recall* dengan menghitung rata-rata harmonik dari keduanya. Nilai *F1-Score* dapat diperoleh dengan menggunakan persamaan

$$F1-Score = \frac{2 \times Precision \times Recall}{Precision+ Recall} \times 100\% \quad (4)$$

3 Hasil dan Diskusi

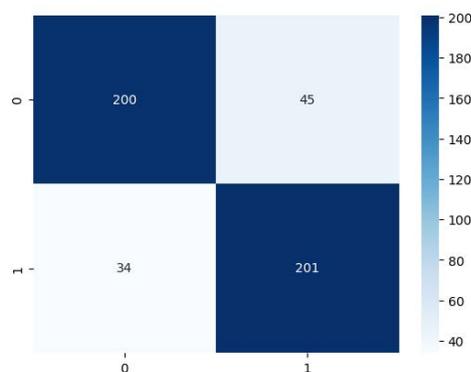
Pada model pertama, *random forest* digunakan dengan konfigurasi default sehingga tidak memerlukan *hyperparameter tuning*. Gambar 4 menggambarkan *confusion matrix* dari proses klasifikasi menggunakan *Random Forest* tanpa melalui proses *hyperparameter tuning*.



Gambar 4. *Confusion Matrix* tanpa *Hyperparameter Tuning*

Pada gambar 4 menunjukkan hasil *confusion matrix* dari model tanpa *hyperparameter tuning*. Data yang benar terklasifikasi sebagai positif berjumlah 190, sedangkan data negatif yang salah terklasifikasi sebagai positif berjumlah 59. Sebaliknya, data positif yang salah terklasifikasi sebagai negatif berjumlah 55, dan data yang benar terklasifikasi sebagai negative berjumlah 176.

Pada model kedua, *random forest* mengalami *hyperparameter tuning* dengan menggunakan *GridSearchCV*. *Hyperparameter tuning* menggunakan *GridSearch Cross-Validation* dengan jumlah *k-fold* bernilai 5. Dalam model *Random Forest* yang telah melalui proses *hyperparameter tuning* ditemukan bahwa model yang paling optimal menggunakan 100 pada *n_estimators*, 9 pada *max_depth*, 4 pada *max_features*, dan 2 pada *min_samples_split*. Gambar 5 merupakan *confusion matrix* dari model klasifikasi *Random Forest* yang telah mengalami proses *hyperparameter tuning* menggunakan *GridsearchCV*.



Gambar 5. *Confusion Matrix* dengan *Hyperparameter Tunning GridSearchCV*

Pada gambar 5 menunjukkan hasil *confusion matrix* dari model tanpa *hyperparameter tuning*. Data yang benar terklasifikasi sebagai positif berjumlah 200, sedangkan data negatif yang salah terklasifikasi sebagai positif berjumlah 34. Sebaliknya, data positif yang salah terklasifikasi sebagai negatif berjumlah 45, dan data yang benar terklasifikasi sebagai negatif berjumlah 201.

Evaluasi model meliputi akurasi, presisi, *recall*, dan *f1-score* dihitung berdasarkan data yang terdapat dalam *confusion matrix*. Perbandingan antara *random forest* tanpa *hyperparameter tuning* atau dengan *hyperparameter default* dan *random forest* dengan menggunakan *GridSearchCV* untuk *hyperparameter tuning* ditunjukkan pada Tabel 4

Tabel 4. Hasil Perbandingan Performa *Random Forest*

Metric	Default	Hyperparameter tuning
Accuracy	76,00%	84,00%
Precision	76,19%	81,70%
Recall	74,89%	85,53%
F1-Score	75,53%	83,57%

Pada Tabel 4 menunjukkan hasil perbandingan performa *Random Forest* sebelum dan sesudah *hyperparameter tuning*. Hasilnya menunjukkan bahwa performa model meningkat setelah *hyperparameter tuning*. Akurasi yang sebelumnya 76% meningkat menjadi 84%, presisi yang sebelumnya 76,19% meningkat menjadi 81,70%, *recall* yang sebelumnya 74,89% meningkat menjadi 85,53%, dan nilai *f1-score* yang sebelumnya 75,53% meningkat menjadi 83,57%.

4 Kesimpulan

Berdasarkan hasil penelitian dapat disimpulkan bahwa melalui proses *hyperparameter tuning* menggunakan *GridsearchCV*, kinerja model meningkat dibandingkan dengan model yang tidak melalui proses *hyperparameter tuning*. Sebelum dilakukan *hyperparameter tuning*, performa algoritma *Random Forest* telah menghasilkan *accuracy* sebesar 76%, *precision* sebesar 76,19%, *recall* sebesar 74,89%, dan *f1-score* sebesar 75,53%. Namun, setelah dilakukan *hyperparameter tuning* dengan parameter *n_estimators* yaitu 100, *max_depth* yaitu 9, *max_features* yaitu 4, dan *min_samples_split* yaitu 2 ditemukan model yang optimal. Hasilnya menunjukkan peningkatan dalam performa model, di mana akurasi meningkat menjadi 84,00%, *precision* meningkat menjadi 81,70%, *recall* meningkat menjadi 85,53%, dan *f1-score* meningkat menjadi 83,57%. Hal ini menunjukkan bahwa pengaturan parameter yang tepat dapat secara substansial meningkatkan kemampuan prediktif algoritma *Random Forest* dalam mengklasifikasikan kualitas air layak minum.

Referensi

- [1] Y. B. Riyanto, E. Setiawan, M. Hannats, and H. Ichsan, "Sistem Klasifikasi Kelayakan Air Minum pada Depot Air Minum Isi Ulang (DAMIU) berdasarkan pH, Total Dissolved Solids, dan Keekeruhan menggunakan Metode Naïve Bayes," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 7, no. 7, pp. 3525-3536, 2023.
- [2] G. L. Pritalia, "Analisis Komparatif Algoritma Machine Learning pada Klasifikasi Kualitas Air Layak Minum," *KONSTELASI: Konvergensi Teknologi dan Sistem Informasi*, vol. 2, no. 1, pp. 43-55, 2022.
- [3] A. Kustanto, "Water quality in Indonesia: The role of socioeconomic indicators," *Jurnal Ekonomi Pembangunan*, vol. 18, no. 1, pp.47-62, 2020, doi: 10.29259/jep.v18i1.11509.
- [4] P. A. Riyantoko, T. M. Fahrudin, K. M. Hindrayani, S. Data, and J. Timur, "Analisis Sederhana Pada Kualitas Air Minum Berdasarkan Akurasi Model Klasifikasi Dengan Menggunakan Lucifer Machine Learning," *Seminar Nasional Sains Data*, vol. 1, no. 01, pp. 12–18, 2021.
- [5] A. Sudin, M. Salmin, M. Fhadli, and AM Mamonto, "Klasifikasi Kelayakan Air Minum Bagi Tubuh Manusia Menggunakan Metode Support Vektor Machine Dengan Backward Elimination," *Jurnal Jaringan dan Teknologi Informasi*, vol. 3, no. 01, pp. 87–95, 2023
- [6] A. M. Graboski, J. Martinazzo, S. C. Ballen, J. Steffens, and C. Steffens, "Nanosensors for water quality control," in *Nanotechnology in the Beverage Industry: Fundamentals and Applications*, 2020. doi: 10.1016/B978-0-12-819941-1.00004-3.
- [7] S. S. Maulina Putri, M. Arhami, and H. Hendrawaty, "Penerapan Metode SVM pada Klasifikasi Kualitas Air," *Journal of Artificial Intelligence and Software Engineering (J-AISE)*, vol. 3, no. 2, pp. 94-101, 2023, doi: 10.30811/jaise.v3i2.4630.

- [8] A. S. Handayani, S. Soim, T. E. Agusdi, Rumiasih, and A. Nurdin, "Klasifikasi Kualitas Udara Dengan Metode Support Vector Machine," *JIRE (Jurnal Informatika & Rekayasa Elektronika)*, vol. 3, no. 2, pp. 187-199, 2020.
- [9] R. J. Alfirdausy and S. Bahri, "Implementasi Algoritma K-Nearest Neighbor untuk Klasifikasi Diagnosis Penyakit Alzheimer," *Techno.Com*, vol. 22, no. 3, pp. 635-642, 2023, doi: 10.33633/tc.v22i3.8393.
- [10] I. Muhamad Malik Matin, "Hyperparameter Tuning Menggunakan GridsearchCV pada Random Forest untuk Deteksi Malware," *MULTINETICS*, vol. 9, no. 1, pp. 43-50, 2023, doi: 10.32722/multinetics.v9i1.5578.
- [11] C. G. Siji George and B. Sumathi, "Grid search tuning of hyperparameters in random forest classifier for customer feedback sentiment prediction," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 9, 2020, doi: 10.14569/IJACSA.2020.0110920.
- [12] I.P.G.H. Suputra and C.R.A. Pramatha, "Rekomendasi Rute Perjalanan Wisata Berbasis Web Menggunakan Algoritma Genetika," *Jurnal Ilmu Komputer*, vol. 13, no. 1, 2020, doi: 10.24843/jik.2020.v13.i01.p03.
- [13] A. Toha, P. Purwono, and W. Gata, "Model Prediksi Kualitas Udara dengan Support Vector Machines dengan Optimasi Hyperparameter GridSearch CV," *Buletin Ilmiah Sarjana Teknik Elektro*, vol. 4, no. 1, pp. 12-21, 2022, doi: 10.12928/biste.v4i1.6079.
- [14] M. A. A. Jihad, A. Adiwijaya, and W. Astuti, "Analisis Sentimen Terhadap Ulasan Film Menggunakan Algoritma Random Forest," *e-Proceeding of Engineering*, vol. 8, no. 5, pp. 1-13, 2021.
- [15] M. Azhar and H. F. Pardede, "Klasifikasi Dialek Pengujar Bahasa Inggris Menggunakan Random Forest," *Jurnal Media Informatika Budidarma*, vol. 5, no. 2, pp.439-446, 2021, doi: 10.30865/mib.v5i2.2754.
- [16] J. D. Muthohhar and A. Prihanto, "Analisis Perbandingan Algoritma Klasifikasi untuk Penyakit Jantung," *Journal of Informatics and Computer Science (JINACS)*, vol.4, no.3, pp. 298-304.2023, doi: 10.26740/jinacs.v4n03.
- [17] D. N. Triwibowo, P. I. A. Ashari, A. S. Sandi, and Y. F. Rahman, "Enkripsi Pesan Menggunakan Algoritma Linear Congruential Generator (LCG) dan Konversi Kode Morse," *Buletin Ilmiah Sarjana Teknik Elektro*, vol. 3, no. 3, pp. 194-201, doi: 10.12928/biste.v3i3.5546.
- [18] Fatmawati and N. A. K. Rifai, "Klasifikasi Penyakit Diabetes Retinopati Menggunakan Support Vector Machine dengan Algoritma Grid Search Cross-validation," *Jurnal Riset Statistika*, vol. 3, no. 1, pp.79-86, 2023, doi: 10.29313/jrs.v3i1.1945.
- [19] K. R. Singh, K. P. Neethu, K. Madhurekaa, A. Harita, and P. Mohan, "Parallel SVM model for forest fire prediction," *Soft Computing Letters*, vol. 3, 2021, doi: 10.1016/j.socl.2021.100014.
- [20] G. S. K. Ranjan, A. Kumar Verma, and S. Radhika, "K-Nearest Neighbors and Grid Search CV Based Real Time Fault Monitoring System for Industries," in *2019 IEEE 5th International Conference for Convergence in Technology, I2CT 2019*, 2019. doi: 10.1109/I2CT45611.2019.9033691.
- [21] S. P. Tamba and Edric, "Prediksi Penyakit Gagal Jantung dengan Menggunakan Random Forest," *Jurnal Sistem Informasi dan Ilmu Komputer Prima(JUSIKOM PRIMA)*, vol. 5, no. 2, pp. 176-181, 2022, doi: 10.34012/jurnalsisteminformasidanilmukomputer.v5i2.2445.
- [22] Pirjatullah, D. Kartini, D. T. Nugrahadi, Muliadi, and A. Farmadi, "Hyperparameter Tuning using GridsearchCV on the Comparison of the Activation Function of the ELM Method to the Classification of Pneumonia in Toddlers," in *Proceedings - 2021 4th International Conference on Computer and Informatics Engineering: IT-Based Digital Industrial Innovation for the Welfare of Society, IC2IE 2021*, 2021. doi: 10.1109/IC2IE53219.2021.9649207.

This page is intentionally left blank.



ISSN



E-ISSN
