



Volume 3, Nomor 1, November 2024

JNATIA

Jurnal Nasional Teknologi Informasi dan Aplikasinya

Program Studi Informatika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Udayana

Susunan Dewan Redaksi
Jurnal Nasional Teknologi Informasi dan Aplikasinya (JNATIA)
Volume 3, Nomor 1, November 2024

Penanggung Jawab

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

Ketua Redaktur

I Gede Surya Rahayuda, M.Kom.

Editor

Drs. I Wayan Santiyasa, M.Si.
Dr. Dra. Luh Gede Astuti, M.Kom.
Ida Bagus Made Mahendra, S.Kom., M.Kom.
I Gede Arta Wibawa, S.T., M.Kom.
I Dewa Made Bayu Atmaja Darmawan, S.Kom., M.Cs.
I Made Widhi Wirawan, S.Si., M.Si., M.Cs.

Desain 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.

Reviewer

Dr. Ir. I Ketut Gede Suhartana, S.Kom., M.Kom., IPM., ASEAN.Eng.
I Gede Arta Wibawa, S.T., M.Kom.
I Made Widiartha, S.Si., M.Kom.
Ida Bagus Made Mahendra, S.Kom., M.Kom.
Ida Bagus Gede Dwidasmara, S.Kom., M.Cs.
Gst. Ayu Vida Mastrika Giri, S.Kom., M.Cs.
I Gusti Agung Gede Arya Kadyanan, S.Kom., M.Kom.
Dr. Ngurah Agus Sanjaya ER, S.Kom., M.Kom.
I Dewa Made Bayu Atmaja Darmawan, S.Kom., M.Cs.

Luh Arida Ayu Rahning Putri, S.Kom., M.Cs.
Agus Muliantara, S.Kom., M.Kom.
Dra. Luh Gede Astuti, M.Kom.
Cokorda Rai Adi Pramartha, S.T., M.M., Ph.D.
I Gusti Ngurah Anom Cahyadi Putra, S.T., M.Cs.
Dr. Anak Agung Istri Ngurah Eka Karyawati, S.Si., M.Eng.
I Gede Santi Astawa, S.T., M.Cs.
Dr. Made Agung Raharja, S.Si., M.Cs.
I Komang Ari Mogi, S.Kom., M.Kom.
Ida Ayu Gde Suwiprabayanti Putra, S.Kom., M.T.
I Putu Gede Hendra Suputra, S.Kom., M.Kom.
Dr. Drs. I Wayan Santiyasa, M.Si.
I Gede Surya Rahayuda, M.Kom.
I Wayan Supriana, S.Si., M.Cs.

Daftar Isi

Rancang Model Ontologi: Representasi Pengetahuan Kuliner Tradisional Kabupaten Bangli, Bali Celia Maureen Chandra, Agus Muliantara	1-10
Analisis Dimensi Gambar Terhadap Klasifikasi Batik Indonesia dengan CNN Ida Bagus Gde Ardita Mahaprawira, Agus Muliantara	11-16
Enkripsi Resep Dokter untuk Meminimalisir Penyalahgunaan Obat Menggunakan Algoritma AES Mode CBC Komang Wahyu Agastya, AAIN Eka Karyawati	17-26
Analisis Penggunaan Logika Fuzzy Mamdani dan Sugeno untuk Memprediksi Shade Foundation Aprinia Salsabila Roiqoh, Hanin Fatma Soraya, Dela Ayu Putri Mayona, Nabila Anggita Luna, Anggraini Puspita Sari	27-36
Segmentasi Pengguna Spotify Berdasarkan Preferensi Musik dengan Algoritma K-Means Clustering Kadek Bisma Dharmasena, Cokorda Pramatha	37-42
Pengujian Prototipe Sistem Jasa Pengembangan Aplikasi Menggunakan Metode System Usability Scale I Kadek Agus Wijaya Kusuma, Gst. Ayu Vida Matrika Giri	43-52
Klasifikasi Tingkat Keparahan Kecelakaan Lalu Lintas Menggunakan Random Forest Classifier I Gusti Ngurah Bagus Lanang Purbhawa, I Gede Arta Wibawa	53-62
Analisis Sentimen Ulasan Traveloka Menggunakan Metode Naïve Bayes Classifier dan Information Gain Kadek Yuni Suratri, I Gede Santi Astawa	63-70
Perancangan Desain Antarmuka Aplikasi Soul Notes dengan Metode Design Thinking berbasis Mobile Ni Komang Purnami, I Gede Surya Rahayuda	71-80
Analisis Prediktif Bitcoin dengan Metode SVM serta Pembobotan TIF-IDF Berbasis Data Narrative Danendra Darmawansyah, I Gusti Agung Gede Arya Kadyanan	81-90
Identifikasi Mekar dan Kuncupnya Bunga Sedap Malam Menggunakan Convolutional Neural Network Kadek Bakti Pramanayoga St, I Gusti Agung Gede Arya Kadyanan	91-98
Penerapan Model Ontologi dalam Perkembangan Game Digital I Komang Maheza Yudistia, I Gusti Ngurah Anom Cahyadi Putra	99-108
Klasifikasi Mood pada Musik Pop dan Jazz dengan Menggunakan Mel Frequency Cepstral Coefficients dan K-Nearest Neighbor I Gusti Bagus Putrawan, I Ketut Gede Suhartana	109-116

Penerapan Enkripsi dan Dekripsi Dokumen Data UMKM Menggunakan Algoritma ChaCh0-Poly1305 I Made Chandra Widjaya, I Komang Ari Mogi	117-126
Klasifikasi Genre Musik Menggunakan Metode Support Vector Machine dengan Multi-Kernel I Gusti Agung Istri Agrivina Shyta Devi, I Made Widiartha	127-132
Analisa Rancangan Desain Antarmuka Aplikasi LibrarySense Menggunakan System Usability Scale Gagas Pradipta Jatmiko, I Putu Gede Hendra Suputra	133-142
Pengaruh Penanganan Ketidakseimbangan Kelas pada Prediksi Cacat Perangkat Lunak dengan Teknik Oversampling I Gusti Agung Ramananda Wira Dharma, I Wayan Santiyasa	143-152
Analisis Performa Algoritma K-Nearest Neighbor dalam Klasifikasi Penyakit Tumor Otak Komang Gede Bagus Devit Aditiya, I Wayan Santiyasa	153-160
Implementasi Algoritma Yolo untuk Deteksi Kebusukan pada Sayur Kembang Kol Alexander Ibrahim, I Wayan Supriana	161-168
Deteksi Hate Speech pada Unggahan Media Sosial dengan Naive Bayes Menggunakan Seleksi Fitur Chi-square Putu Steven Belva Chan, Ida Ayu Gde Suwiprabayanti Putra	169-176
Rancang Model Ontologi dalam Representasi Digital Loloh Cencem Penglipuran Putu Chandra Mayoni, Ida Bagus Gede Dwidasmara	177-188
Simulasi IoT Pemantauan Tanaman Lidah Buaya Berbasis Algoritma Fuzzy Bayu Yudistira Ramadhan, Luh Arida Ayu Rahning Putri	189-194
Analisa Perancangan Sistem Rekomendasi Makanan Untuk Defisit Kalori "Calcraft" Melalui Evaluasi UI/UX Putu Ananda Darma Wiguna, Luh Gede Astuti	195-204
Analisis dan Klasifikasi Genre Musik Menggunakan Algoritma STFT dan Random Forest Merry Royanti Manalu, Made Agung Raharja	205-214
Pengamanan Gambar dengan Metode Cipher Block Chaining Made Yayang Eka Pranand, I Gusti Ngurah Anom Cahyadi Putra	215-222
Pengembangan Model Ontologi pada Domain Oleh-Oleh Khas Bali Putu Ardi Sudarmika, Ngurah Agus Sanjaya ER	223-232

Rancang Model Ontologi: Representasi Pengetahuan Kuliner Tradisional Kabupaten Bangli, Bali

Celia Maureen Chandra^{a1}, Agus Muliantara^{a2}

^aProgram Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹chandra.2208561080@student.unud.ac.id
²muliantara@unud.ac.id

Abstract

The Province of Bali is a preferred tourist destination in Indonesia, including one of the developing tourist areas, Bangli Regency. The tourism development in Bangli is closely linked to traditional cuisine as a key attraction. However, in preserving the traditional cuisine, it is evident that many residents lack in-depth knowledge about the variety of Bangli's traditional dishes. This research aims to develop an ontology model representing the traditional culinary knowledge of Bangli Regency using Methontology method. The implementation, conducted using Protégé software, resulted in 6 classes, 8 object properties, 5 data properties, and 55 individuals. The ontology was evaluated using the Ontology Quality Analysis (OntoQA) method with schema metrics techniques, including RR, IR, and AR measurements. The evaluation revealed that the ontology has a high diversity of information and specific information representation, though the amount of information provided is still limited. With these promising results, the constructed ontology is expected to effectively represent the knowledge of traditional Bangli cuisine and serve as a foundation for developing an information system that facilitates access to this knowledge, thereby supporting its preservation.

Keywords: Ontology, Traditional Bangli Cuisine, Methontology, Protégé, OntoQA

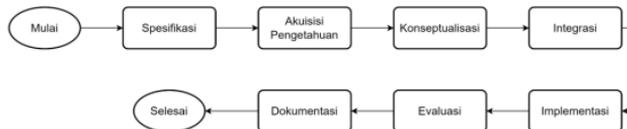
1. Pendahuluan

Provinsi Bali merupakan destinasi wisata Indonesia yang menjadi pilihan para wisatawan, baik sebelum dan setelah pandemi COVID-19. Dapat dilihat pada data banyaknya wisatawan mancanegara tahunan ke Bali menurut pintu udara oleh Badan Pusat Statistik Provinsi Bali, yaitu sebelum pandemi sebanyak 6.027.159 orang pada tahun 2018 dan 6.239.543 orang pada tahun 2019, kemudian setelah pandemi kembali meningkat hingga 2.154.045 orang di tahun 2022 dan 5.248.113 orang di tahun 2023 [1]. Salah satu daerah wisata berkembang di Bali yang mulai ramai dikunjungi oleh wisatawan adalah Kabupaten Bangli, dengan adanya wisata seperti Danau dan Gunung Batur serta Desa Penglipuran. Perkembangan pariwisata Kabupaten Bangli tidak dapat dipisahkan dari peran kuliner tradisional sebagai daya tarik wisatawan. Kuliner tradisional memiliki fungsi sebagai konsumsi sehari-hari, obat-obatan, sarana pelaksanaan adat dan tradisi, serta sebagai warisan budaya yang harus dijaga dan dilestarikan. Dalam pelestarian kuliner tradisional Bangli, sangat diperlukan peran masyarakat setempat yang paham secara mendalam atas pengetahuan ragam kuliner daerahnya karena pada umumnya kuliner tradisional hanya menggunakan bahan-bahan nabati dan hewani yang berasal dari perkebunan, peternakan, dan kolam air tawar pada wilayah tersebut serta diolah dengan metode yang diwariskan secara turun temurun tanpa berpedoman pada resep dan hanya melalui kebiasaan setempat [2]. Namun, dalam penelitian oleh Agustini di Kecamatan Bangli, Kabupaten Bangli, 50% responden memiliki pengetahuan yang sangat minim atas jenis, bahan utama, dan teknik pengolahan kuliner tradisional yang ada pada wilayahnya [3]. Putra, Fredlina, dan Ardyanti telah berupaya dalam melestarikan kuliner tradisional Bali dengan mengembangkan aplikasi ensiklopedia kuliner tradisional Bali yang berbasis mobile geographical information system sehingga mempermudah pengguna, baik wisatawan maupun masyarakat lokal, dalam mencari lokasi penjual kuliner tradisional Bali [4]. Dalam melestarikan kuliner tradisional Bali, terkhusus pada mengenalkan

pengetahuan terkait ragam kuliner tradisional di Kabupaten Bangli, dibutuhkan suatu solusi yang mampu merepresentasikan dalam bentuk digital. Salah satu teknologi yang dapat dimanfaatkan adalah ontologi, yaitu metode untuk merepresentasikan domain pengetahuan tertentu secara eksplisit dengan memberikan makna, properti, dan relasi pada konsep sehingga membentuk kumpulan pengetahuan dalam domain tersebut dan menciptakan basis pengetahuan (knowledge base) [5]. Perancangan model ontologi dilakukan dengan menggunakan metode Methontology yang mengusulkan pengekspresian ide sebagai sekumpulan Intermediate Representations (IR) dan menghasilkan ontologi menggunakan translators [6]. Model ontologi dibangun dengan memanfaatkan perangkat lunak Protégé, yaitu perangkat lunak open source berbasis Java yang mendukung format penyimpanan seperti OWL, RDF, dan XML [7]. Untuk mendapatkan model ontologi yang berkualitas baik, akan dilakukan pengujian menggunakan metode Ontology Quality Analysis (OntoQA) dengan salah satu teknikanya, yaitu schema metrics, yang digunakan dalam proses evaluasi rancangan ontologi berupa potensi knowledge yang dapat dihasilkan [8]. Hasil pemodelan ontologi dapat diterapkan dalam web semantik untuk mendefinisikan data menjadi pengetahuan, keterkaitan antar data, serta representasi graph. Penerapan tersebut melibatkan metode yang dapat dipahami oleh mesin dan memungkinkan otomatisasi integrasi [7]. Melalui penelitian ini, diharapkan model ontologi dapat merepresentasikan pengetahuan kuliner tradisional Bangli dengan baik dan menjadi dasar pengembangan sistem informasi yang mempermudah akses pengetahuan kuliner tradisional Bangli guna mendukung pelestariannya.

2. Metode Penelitian

Metode yang digunakan dalam penelitian ini adalah Methontology. Metode Methontology mampu melakukan pengembangan ontologi dengan menggunakan Intermediate Representations (IR) dan translators. Dengan menyediakan kerangka kerja bagi siklus hidup ontologi dalam cakupan pengetahuan mulai dari identifikasi proses pengembangan hingga perawatan. Metode Methontology memberikan keunggulan pada penelitian dalam mengintegrasikan dan memanfaatkan kembali model ontologi terdahulu untuk pengembangan lebih lanjut serta mampu memberikan panduan terstruktur mengenai tahap menjalankan aktivitas dalam proses pengembangan ontologi, termasuk teknik dan hasil yang diharapkan dari setiap langkah [6]. Adapun tahapan metode Methontology yang harus dilakukan terdiri atas spesifikasi, akuisisi pengetahuan, konseptualisasi, integrasi, implementasi, evaluasi, dokumentasi, dan selesai [9].



Gambar 1. Metode Methontology

2.1 Spesifikasi

Tahap spesifikasi bertujuan dalam menghasilkan dokumen spesifikasi ontologi yang berbentuk formal, semi-formal, atau informal, ditulis dalam bahasa alami (natural language). Metode ini menggunakan serangkaian representasi tingkat menengah atau pertanyaan kompetensi. Berikut merupakan deskripsi ontologi Kuliner Tradisional Bangli.

- Domain : Kuliner Tradisional Bangli
- Tujuan : Membangun model ontologi sebagai representasi pengetahuan kuliner tradisional Kabupaten Bangli, Bali
- Dirancang oleh : Celia Maureen Chandra
- Level Formalitas : Formal
- Ruang Lingkup : Kuliner Tradisional Bangli
- Sumber Pengetahuan : Studi Literatur

2.2 Akuisisi Pengetahuan

Tahap berikutnya merupakan tahap independen dalam pengembangan ontologi. Sebagian besar tahap akuisisi telah dilakukan bersamaan dengan tahap spesifikasi, dan perannya semakin berkurang seiring kemajuan dalam proses pengembangan ontologi. Penggunaan data untuk membangun model ontologi dalam penelitian ini mencakup data terkait kuliner tradisional Bangli. Data tersebut diperoleh dari berbagai sumber studi literatur, yaitu skripsi, jurnal dan internet.

2.3 Konseptualisasi

Pada tahap konseptualisasi, dibuat model konseptual yang mewakili pengetahuan domain, menggambarkan masalah, dan solusinya dengan istilah-istilah yang telah diidentifikasi pada tahap spesifikasi. Tugas utamanya adalah menyusun Glosarium of Terms (GT) yang mencakup konsep, contoh, kata kerja, dan properti terkait. GT bertujuan untuk mengidentifikasi dan mengumpulkan semua pengetahuan domain yang relevan serta memberikan definisi untuk setiap elemen tersebut.

2.4 Integrasi

Tahap selanjutnya merupakan tahap integrasi yang melibatkan pemanfaatan ontologi yang telah ada dan menggabungkannya ke dalam ontologi lain sehingga berkaitan dengan domain kuliner tradisional Bangli. Hal ini memungkinkan pembangunan ontologi tanpa perlu memulai dari awal dan mendapatkan hasil yang sesuai.

2.5 Implementasi

Kemudian dilanjutkan ke tahap implementasi, yaitu proses penerapan rancangan ontologi dari tahap spesifikasi hingga integrasi. Pada tahap ini, rancangan ontologi yang telah dibuat diimplementasikan menggunakan perangkat lunak Protégé 5.5.0.

2.6 Evaluasi

Pada tahap evaluasi, dilakukan penilaian teknis terhadap ontologi, lingkungan perangkat lunak, dan dokumentasi terkait, mengacu pada kerangka referensi di setiap tahap siklus kehidupan mereka. Evaluasi terdiri dari dua komponen utama, yaitu verifikasi dan validasi. Verifikasi bertujuan untuk memastikan akurasi ontologi, lingkungan perangkat lunak, dan dokumentasi terkait dengan kerangka referensi di setiap tahap siklus kehidupan. Sementara itu, validasi memastikan bahwa ontologi, lingkungan perangkat lunak, dan dokumentasi tersebut selaras dengan sistem yang akan diwakili. Evaluasi dilakukan dengan metode Ontology Quality Analysis (OntoQA). Dengan salah satu teknikya, yaitu schema metrics, pengukuran terdiri atas Relationship Richness (RR), Inheritance Richness (IR), dan Attribute Richness (AR) [8]. Relationship Richness menunjukkan keberagaman relasi yang dimiliki dan dihitung dengan membagi jumlah relasi non-inheritance (P) dengan total relasi yang ada (jumlah relasi inheritance (H) ditambah P).

$$RR = \frac{|P|}{|H|+|P|} \quad (1)$$

Nilai RR yang mendekati 0 menandakan sebagian besar relasi adalah inheritance, dan sebaliknya. Dominasi relasi inheritance menunjukkan ontologi memiliki informasi yang lebih sedikit. Acuan dalam penilaian dapat menggunakan tabel klasifikasi seperti pada Tabel 1 [8].

Tabel 1. Skala Klasifikasi Nilai Relationship Richness

Rentang Skala	Nilai Skala
0,00 - 0,20	Minimum
0,21 - 0,40	Kurang

Rentang Skala	Nilai Skala
0,41 - 0,60	Cukup
0,61 - 0,80	Kaya
0,81 - 1,00	Maksimum

Inheritance Richness (IR) mengukur distribusi informasi dengan mendapatkan rata-rata jumlah subclass ($H(c)$) per class (C).

$$IR = \frac{|H(c)|}{|C|} \quad (2)$$

Nilai IR yang kecil menunjukkan bahwa ontologi memiliki representasi informasi yang mendalam atau spesifik terhadap suatu domain dan memiliki level inheritance yang banyak. Sifat ontologi tersebut disebut juga dengan karakter deep. Sebaliknya, ontologi bersifat shallow memiliki nilai yang besar dan representasi informasi kurang mendalam namun mencakup area yang lebih luas. Dalam penilaiannya dapat digunakan acuan yang telah dirangkum oleh Saf dalam jurnal berjudul "Pemodelan Ontologi untuk Sistem Informasi Proyek Akhir (SIMPOA) di Politeknik Caltex Riau" sebagai berikut [8].

Tabel 2. Skala Klasifikasi Nilai Inheritance Richness

Rentang Skala	Nilai Skala
0,00 - 1,34	Spesifik
1,35 - 2,68	Cukup Spesifik
2,69 - 4,02	Cukup Umum
4,03 - 5,36	Umum

Attribute Richness (AR) mengukur jumlah informasi yang ada pada ontologi dengan mendapatkan jumlah rata-rata attribute atau data property (att) per class (C). Semakin banyak attribute menunjukkan semakin banyak informasi yang disediakan oleh ontologi

$$AR = \frac{|att|}{|C|} \quad (3)$$

2.7 Dokumentasi

Tahap terakhir dalam pembangunan ontologi melibatkan proses dokumentasi. Dokumentasi dapat berupa kode ontologi, teks bahasa alami yang dilampirkan pada definisi formal, maupun makalah yang diterbitkan dalam konferensi dan jurnal. Tahap dokumentasi bertujuan dalam menguraikan pertanyaan-pertanyaan penting terkait ontologi yang telah dikembangkan.

3. Hasil dan Diskusi

3.1. Perancangan Ontologi

Perancangan ontologi Kuliner Tradisional Bangli dilakukan dengan membuat tabel yang berisikan class dan subclass ontologi serta tabel perancangan data property, domain, dan range. Hasil perancangan class dan subclass dapat dilihat pada Tabel 3 dan hasil perancangan data property, domain, dan range dapat dilihat pada Tabel 4.

Tabel 3. Rancangan Class dan SubClassOf Ontologi Kuliner Tradisional Bangli

Class	SubClassOf
Kuliner_Tradisional_Bangli	Thing
Nama_Kuliner	Kuliner_Tradisional_Bangli
Jenis_Kuliner	Kuliner_Tradisional_Bangli
Metode_Pengolahan	Kuliner_Tradisional_Bangli
Bahan_Utama	Kuliner_Tradisional_Bangli
Nabati	Bahan_Utama
Hewani	Bahan_Utama

Tabel 4. Rancangan Data Property, Domain, dan Range Ontologi Kuliner Tradisional Bangli

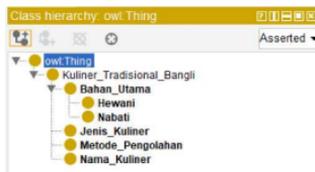
Data Property	Domain	Range
nama_kuliner	Nama_Kuliner	xsd:string
jenis_kuliner	Nama_Kuliner	xsd:string
metode_pengolahan	Nama_Kuliner	xsd:string
bahan_utama_nabati	Nama_Kuliner	xsd:string
bahan_utama_hewani	Nama_Kuliner	xsd:string

3.2. Implementasi

Berdasarkan hasil perancangan ontologi, implementasi dilakukan dengan membuat model ontologi sebagai representasi pengetahuan kuliner tradisional Bangli. Proses implementasi menggunakan perangkat lunak Protégé versi 5.5.0.

3.2.1. Implementasi Class

Class merupakan definisi dari himpunan abstrak yang mengelompokkan objek. Di dalam class, mungkin terdapat instance atau class lain yang dikenal sebagai subclass. Pada hirarki class ontologi Kuliner Tradisional Bangli terdapat 6 class, terdiri 4 class, yaitu "Nama_Kuliner", "Jenis_Kuliner", "Bahan_Utama", dan "Metode Pengolahan". Pada class "Bahan_Utama" terdapat 2 subclass, yaitu "Nabati" dan "Hewani". Berikut merupakan hasil implementasi class dari ontologi Kuliner Tradisional Bangli.



Gambar 2. Hirarki Class Ontologi Kuliner Tradisional Bangli

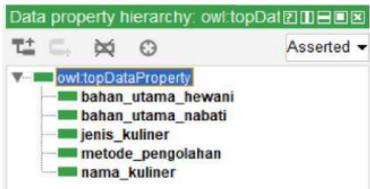
3.2.2. Implementasi Property pada Class

Dalam class, property dibagi menjadi dua jenis, yaitu object property dan data property. Object property mendefinisikan hubungan antar class atau relasi yang terdapat dalam ontologi sehingga berupa kata kerja penghubung antar class dan instance. Sementara data property menghubungkan individu dengan nilai tipe data seperti string, integer, float dan lainnya. Berikut merupakan hasil implementasi property dari ontologi Kuliner Tradisional Bangli.



Gambar 3. Object Property Ontologi Kuliner Tradisional Bangli

Pada ontologi Kuliner Tradisional Bangli terdapat 8 object property yang digunakan untuk mendefinisikan hubungan atau relasi antar individual pada class. Object property "berbahan_utama_hewani" menghubungkan class "Nama_Kuliner" dengan subclass "Hewani" dan inversnya adalah object property "bahan_hewani_dari", sementara object property "berbahan_utama_nabati" menghubungkan class "Nama_Kuliner" dengan subclass "Nabati" dan inversnya "bahan_nabati_dari". Adapun object property "berjenis" yang menghubungkan class "Nama_Kuliner" dengan "Jenis_Kuliner" dan inversnya "jenis_dari", serta object property "diolah_dengan_metode" menghubungkan "Nama_Kuliner" dengan "Metode_Pengolahan" dan inversnya "metode_dari".

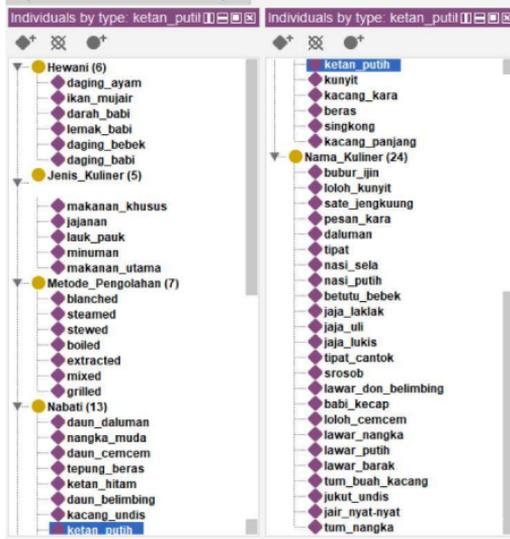


Gambar 4. Data Property Ontologi Kuliner Tradisional Bangli

Data property yang digunakan pada ontologi Kuliner Tradisional Bangli adalah sebanyak 5 data property, yaitu "nama_kuliner" dengan tipe data string yang memberikan value berupa nama kuliner dari setiap individual dari class "Nama_Kuliner", "metode_pengolahan" dengan tipe data string untuk memberikan value berupa metode pengolahan kuliner, "jenis_kuliner" dengan tipe data string untuk memberikan value berupa jenis dari kuliner, "bahan_utama_nabati" dan "bahan_utama_hewani" dengan tipe data string untuk memberikan value berupa bahan utama kuliner.

3.2.3. Implementasi Individuals

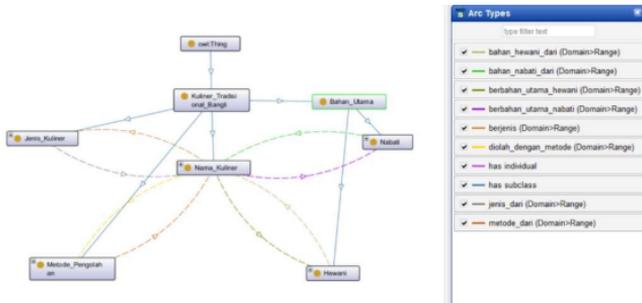
Individuals merupakan atribut dari setiap class atau instance. Dari hasil implementasi individuals pada ontologi Kuliner Tradisional Bangli, dihasilkan beberapa individuals pada setiap class dengan total 55 individuals. Dalam class "Nama_Kuliner" terdapat 24 individuals, pada class "Jenis_Kuliner" terdapat 5 individuals, class "Bahan_Utama" dibagi menjadi subclass "Nabati" dengan 13 individuals dan subclass "Hewani" dengan 6 individuals, dan class "Metode_Pengolahan" terdiri atas 7 individuals.



Gambar 5. Individuals Ontologi Kuliner Tradisional Bangli

3.3. Hasil Ontograf

Ontograf merupakan representasi visual dari hubungan semantik antara class, object property, dan individuals dalam bentuk graf. Tanda panah lurus berwarna biru pada ontograf menunjukkan hubungan antara class dan subclass, sementara tanda panah putus-putus menunjukkan hubungan atau relasi antar class yang dihubungkan dengan object properties. Hasil ontograf dari ontologi Kuliner Tradisional Bangli adalah sebagai berikut.



Gambar 6. Ontograf Ontologi Kuliner Tradisional Bangli

3.4. Evaluasi Model Ontologi

Evaluasi model ontologi melibatkan pengujian menggunakan metode Ontology Quality Analysis (OntoQA) dengan teknik schema metrics. Penilaian yang dilakukan mencakup perhitungan nilai Relationship Richness (RR), Inheritance Richness (IR), dan Attribute Richness (AR). Nilai-nilai yang digunakan dalam penilaian tersebut adalah ontology metrics yang dapat dilihat pada Tabel 5.

Tabel 5. Ontology Metrics Ontologi Kuliner Tradisional Bangli

No	Ontology Element	Count
1	Axiom	342
2	Logical Axiom	267
3	Class	6
4	Data Property	5
5	SubClassOf Axioms	2
6	Object Property	8
7	Object Property Domain Axioms	8
8	Object Property Range Axioms	8
9	Data Property Domain axioms	5

3.4.1. Evaluasi Relationship Richness (RR)

Pada evaluasi Relationship Richness (RR), nilai RR didapatkan dari perhitungan jumlah relasi non-inheritance dibagi dengan total relasi relasi inheritance dan non-inheritance.

$$RR = \frac{|R|}{|I| + |R|} = 0,80 \tag{4}$$

Berdasarkan Tabel 1, nilai RR sebesar 0,80 menunjukkan bahwa ontologi Kuliner Tradisional Bangli telah kaya akan informasi, namun masih dapat ditingkatkan lagi.

3.4.2. Evaluasi Inheritance Richness (IR)

Evaluasi Inheritance Richness (IR) dilakukan dengan mencari rata-rata jumlah subclass per class yang memiliki subclass. Ontologi Kuliner Tradisional Bangli terdiri atas 1 class yang memiliki subclass, dengan total 2 subclass.

$$IR = \frac{|2|}{|1|} = 2,00 \quad (5)$$

Nilai IR = 2,00 yang didapatkan dengan acuan Tabel 2 menunjukkan bahwa ontologi yang dibangun termasuk dalam kategori cukup spesifik, dimana representasi informasi ontologi cukup mendalam (deep).

3.4.3. Evaluasi Attribute Richness (AR)

Pada evaluasi Attribute Richness (AR), perhitungan nilai AR adalah dengan mendapatkan jumlah rata-rata attribute atau data property per class. Dalam ontologi yang dibangun terdapat 5 attribute dan 6 class.

$$AR = \frac{|5|}{|6|} = 0,83 \quad (6)$$

Dengan nilai AR sebesar 0,83 atau rata-rata 1 attribute per class. Nilai tersebut menunjukkan bahwa informasi yang disediakan dalam ontologi yang dibangun masih sedikit informasi. Adapun hasil evaluasi secara keseluruhan dapat dilihat pada Tabel 6.

Tabel 6. Hasil Evaluasi Ontologi Kuliner Tradisional Bangli dengan OntoQA

No	Nama Evaluasi	Hasil	Kesimpulan
1	Relationship Richness (RR)	0,80	Ontologi didominasi relasi non-inheritance, dimana keberagaman informasi cukup tinggi
2	Inheritance Richness (IR)	2,00	Ontologi bersifat deep atau representasi informasi yang cukup mendalam dengan kategori spesifik
3	Attribute Richness (AR)	0,83	Ontologi memiliki rata-rata 1 attribute untuk setiap class, dimana informasi yang disediakan masih sedikit

4. Kesimpulan

Berdasarkan hasil dan pembahasan yang telah dijabarkan, pembangunan model ontologi sebagai representasi pengetahuan kuliner tradisional Bangli telah melalui seluruh tahapan perancangan dan implementasi. Model ontologi dibangun menggunakan perangkat lunak Protégé 5.5.0 dengan metode penelitian Methontology. Perancangan ontologi menghasilkan 6 class, 8 object property, 5 data property, dan total 55 individuals dan telah dievaluasi menggunakan metode Ontology Quality Analysis (OntoQA) dengan teknik schema metrics yang terdiri atas pengukuran RR, IR, dan AR. Dari hasil evaluasi, diketahui bahwa ontologi Kuliner Tradisional Bangli memiliki keberagaman informasi yang cukup tinggi dan representasi informasi yang spesifik. Adapun hal yang harus ditingkatkan adalah jumlah informasi yang disediakan, dimana hasil evaluasi memperlihatkan nilai rata-rata 1 attribute untuk setiap class. Dengan hasil yang cukup baik, pembangunan model ontologi Kuliner Tradisional Bangli diharapkan dapat menjadi dasar pengembangan sistem informasi yang mempermudah akses pengetahuan kuliner tradisional Bangli guna mendukung pelestariannya.

Daftar Pustaka

- [1] Badan Pusat Statistik Provinsi Bali, "Tabel Dinamis", BPS Bali, [Online]. Tersedia: <https://bali.bps.go.id/site/pilihdata.html> [Diakses: 3 Mei 2024].

- [2] A. A. K. A. Pujawan, "Eksplorasi Makanan Tradisional Bali Di Kabupaten Bangli," *Jurnal Gastronomi Indonesia*, vol. 5, no. 1, pp. 1-12, September 2017.
- [3] Agustini. "Eksplorasi Kuliner Tradisional Di Kecamatan Bangli, Kabupaten Bangli", Tugas Akhir. Nusa Dua: STP Nusa Dua Bali, 2016.
- [4] A. A. G. B. Putra, K. Q. Fredlina, I. P. Satwika, dan A. A. A. P. Ardyanti, "Model Aplikasi Ensiklopedia Kuliner Tradisional Bali Berbasis Mobile Geographical Information System," *Jutisi: Jurnal Ilmiah Teknik Informatika dan Sistem Informasi*, vol. 9, no. 1, pp. 95-106, April 2020.
- [5] Y. F. Badron, F. Agus, dan H. R. Hatta, "Studi Tentang Pemodelan Ontologi Web Semantik Dan Prospek Penerapan Pada Bibliografi Artikel Jurnal Ilmiah," *Prosiding Seminar Ilmu Komputer dan Teknologi Informasi*, vol. 2, no. 1, pp. 164-169, 2017.
- [6] K. W. Triyoga, D. E. Cahyani, dan S. W. Sihwi, "Pembangunan Ontology Berbasis Metode Methontology Untuk Domain Tuberculosis," in *Seminar Nasional Inovasi Teknologi*, pp. 47-54, Feb 23, 2019.
- [7] Y. S. Baskoro, H. S., dan H. Jayadianti, "Representasi Pengetahuan Dalam Semantik Ontologi Pada Domain Kriminalitas Kepolisian Sektor Depok Timur Daerah Istimewa Yogyakarta," *TELEMATIKA*, vol. 15, no. 01, pp. 13-29, 2018.
- [8] M. R. A. Saf, "Pemodelan Ontologi untuk Sistem Informasi Proyek Akhir (SIMPOA) di Politeknik Caltex Riau," *Jurnal Sains, Teknologi dan Industri*, vol. 13, no. 1, pp. 46-52, 2015.
- [9] C. Pramatha, "Pengembangan Ontologi Tujuan Wisata Bali Dengan Pendekatan Kulkul Knowledge Framework," *SINTECH (Science and Information Technology) Journal*, vol. 13, no. 2, pp 77-89, 2020.

Analisis Dimensi Gambar Terhadap Klasifikasi Batik Indonesia dengan CNN

¹Ida Bagus Gde Ardita Mahaprawira, ²Agus Muliantara

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

Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia

1mahaprawira.2208561127@student.ac.id

2muliantara@gmail.com

Abstract

Batik motif classification has gained significant attention due to its cultural significance and practical applications in various fields. This study explores the impact of image dimensions on the classification of batik motifs using Convolutional Neural Networks (CNN). The research investigates how variations in image dimensions affect the accuracy and robustness of CNN-based classification models. Through experimentation with different image resolutions and aspect ratios, the study aims to identify optimal settings for achieving high classification performance. Additionally, it examines the computational efficiency of CNN models under varying image dimensions. The findings contribute to enhancing the understanding of image preprocessing techniques and model optimization strategies for batik motif classification tasks.

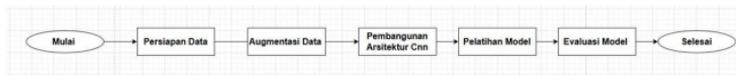
Keywords: *Batik motif classification, Convolutional Neural Networks (CNN), image dimensions, classification accuracy, computational efficiency.*

1. Pendahuluan

Batik, warisan budaya Indonesia yang kaya akan motif dan makna, kini terancam pudar di era digital. Klasifikasi motif batik secara manual membutuhkan waktu, keahlian, dan tenaga yang tidak sedikit. Kemajuan teknologi, khususnya di bidang kecerdasan buatan, membuka peluang baru untuk melestarikan batik melalui otomasi. Masing-masing motif memiliki makna dan filosofi tersendiri, mencerminkan kekayaan budaya Indonesia [1]. Convolutional Neural Network (CNN) telah menunjukkan performa luar biasa dalam klasifikasi gambar. Namun, performa CNN dalam klasifikasi motif batik dapat dipengaruhi oleh berbagai faktor, salah satunya adalah dimensi gambar. Dimensi gambar mengacu pada lebar dan tinggi gambar. Dimensi gambar yang tepat dapat membantu CNN menangkap fitur penting dari motif batik dengan lebih baik, sehingga meningkatkan akurasi klasifikasi. Penelitian ini bertujuan untuk menganalisis pengaruh dimensi gambar terhadap akurasi klasifikasi motif batik menggunakan CNN. Dimensi gambar yang tepat dapat mengoptimalkan performa CNN dalam menangkap ciri khas motif batik dan meningkatkan akurasi klasifikasi. Hasil penelitian ini dapat bermanfaat bagi berbagai pihak, seperti pelestari budaya, kolektor batik, dan pedagang batik [2].

2. Metode Penelitian

2.1 Dataset



Gambar 1. Desain Penelitian

2.2 Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah gambar motif batik yang dikumpulkan dari berbagai sumber, seperti internet, museum batik, dan koleksi pribadi. Data gambar harus direpresentasikan secara memadai dari segi jumlah dan variasi motif batik. Berikut Beberapa contoh jenis gambar batik.



Gambar 2. Batik Betawi



Gambar 3. Batik Bali



Gambar 4. Batik Insang



Gambar 5. Batik Cendrawasih Megamendung



Gambar 6. Batik Parang



Gambar 7. Batik



Gambar 8. Batik Dayak



Gambar 9. Batik Ikat celup

2.3 Preprocessing Data

Sebelum digunakan untuk pelatihan dan evaluasi model CNN, data gambar perlu dipreprocessing untuk meningkatkan kualitas dan konsistensi data. Gambar diubah ukurannya menjadi dimensi yang seragam, sesuai dengan dimensi gambar yang akan digunakan untuk pelatihan dan evaluasi CNN.

2.4 Augmentasi Data

Teknik augmentasi data dapat digunakan untuk meningkatkan variasi data dan mencegah

overfitting. Teknik augmentasi data yang umum digunakan untuk gambar meliputi rotasi, flipping, cropping, dan penambahan noise. Pemilihan teknik augmentasi data yang tepat dapat meningkatkan generalisasi model.

2.5 Arsitektur CNN

Arsitektur CNN yang digunakan dalam penelitian ini akan disesuaikan dengan kebutuhan klasifikasi motif batik. Jumlah layer, ukuran filter, dan fungsi aktivasi pada setiap layer akan dioptimalkan untuk mencapai performa klasifikasi motif batik yang terbaik. Pemilihan arsitektur CNN yang tepat dapat meningkatkan performa klasifikasi.

2.6 Pembagian data

Data gambar dibagi menjadi dua set, yaitu set pelatihan dan set validasi. Set pelatihan digunakan untuk melatih model CNN, sedangkan set validasi digunakan untuk mengevaluasi performa model selama proses pelatihan.

2.7 Evaluasi model

Performa model CNN dievaluasi dengan menggunakan metrik akurasi klasifikasi pada set validasi. Proses evaluasi model akan diulang secara iteratif hingga model mencapai performa klasifikasi motif batik yang optimal.

3. Hasil dan Diskusi

3.1 Evaluasi Kinerja Model

Penelitian ini mengevaluasi pengaruh dimensi gambar terhadap akurasi klasifikasi motif batik dengan CNN. Model CNN dilatih dengan menggunakan data gambar yang telah diproses dan dievaluasi pada set validasi. Hasil evaluasi menunjukkan bahwa dimensi gambar memiliki pengaruh yang signifikan terhadap akurasi klasifikasi motif batik. Dimensi gambar 128x128 menghasilkan akurasi klasifikasi tertinggi, yaitu 45,76 %. Berikut table laporan klasifikasi

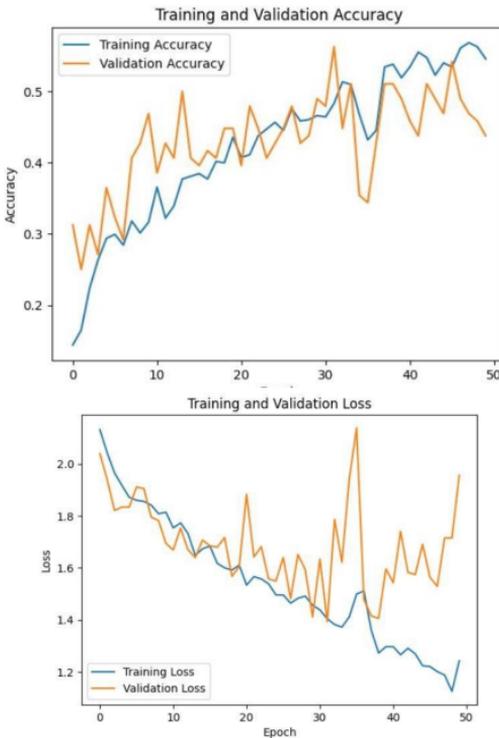
Table 1. Classification Report

	precision	recall	f1-score	support
Batik Parang	0.12	0.05	0.07	20
Batik Ikat Celup	0.22	0.10	0.14	20
Batik Bali	0.05	0.25	0.08	4
Batik Cendrawasih	0.00	0.00	0.00	6
Batik Dayak	0.15	0.35	0.21	20
Batik Insang	0.18	0.10	0.13	20
Batik Megamendung	0.42	0.15	0.31	20
Batik Betawi	0.00	0.00	0.00	8
accuracy			0.15	118
macro avg	0.14	0.14	0.12	118
weighted avg	0.19	0.25	0.15	118

Penelitian ini menunjukkan bahwa dimensi gambar 128x128 menghasilkan akurasi klasifikasi motif batik tertinggi dengan CNN. Selain dimensi gambar, faktor lain seperti jumlah data, arsitektur CNN, dan teknik augmentasi data juga dapat memengaruhi performa klasifikasi. Hasil penelitian ini dapat digunakan untuk mengoptimalkan model CNN untuk klasifikasi motif batik dan meningkatkan akurasinya.

3.2 Diskusi

Hasil penelitian ini dapat digunakan untuk mengoptimalkan model CNN untuk klasifikasi motif batik dan meningkatkan akurasinya. Hal ini dapat membantu dalam pelestarian budaya batik dengan mengidentifikasi dan mengkatalogkan motif batik dengan benar.



Gambar 10. Grafik training and Validation Accuracy

4. Kesimpulan

Proses klasifikasi motif batik secara manual membutuhkan waktu, keahlian, dan tenaga yang tidak sedikit. CNN telah membuktikan performanya yang luar biasa dalam klasifikasi gambar. Namun, performa CNN dalam klasifikasi motif batik dapat dipengaruhi oleh berbagai faktor, salah satunya adalah dimensi gambar. Dimensi gambar mengacu pada lebar dan tinggi gambar. Penelitian ini memberikan kontribusi penting dalam memahami faktor-faktor yang memengaruhi klasifikasi motif batik dengan CNN dan membuka jalan bagi pengembangan sistem otomatis yang lebih efisien dalam melestarikan warisan budaya batik Indonesia.

Daftar pustaka

- [1] Ernawati, "batik design training sebagai upaya pembekalan softskill di bidang desain grafis siswa-siswi di smk negeri no 5 bengkulu," jurnal rekursif, vol. 3, no. 2303-0755, 2015.
- [2] E. D. A. Fathin Ulfah Karimah, "Rancang Bangun Aplikasi Pencarian Citra Batik Besurek Berbasis Tekstur Dengan Metode Gray Level Cooccurrence Matrix Dan Euclidean Distance," Teknologi Informasi, vol. 11, no. 1414-9999, 2015.
- [3] A. H. Rangkuti, "Klasifikasi Motif Batik Berbasis Kemiripan Ciri Dengan Wavelet Transform Dan Fuzzy Neural Network," ComTech, vol. 5, no. 361-372, 2015.
- [4] C. J. Daurat Sinaga, "Klasifikasi Citra Batik Sumatera Menggunakan NaiveBayes Berbasis Fitur Ekstraksi Glcm," Semnas Ristek, no. 2527-5941, 2024.
- [5] C. ., D. D. I. G. D. P. S. Angginy Akhirunnisa Siregar, "Klasifikasi Batik Parang Menggunakan Convolutional Neural Network (CNN)," Jurnal Penelitian Mahasiswa, vol. 3, no. 2963-5306, 2024.

Halaman ini sengaja dibiarkan kosong

Enkripsi Resep Dokter untuk Meminimalisir Penyalahgunaan Obat Menggunakan Algoritma AES Mode CBC

Komang Wahyu Agastya^{a1}, AAIN Eka Karyawati^{a2}

^aProgram Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹wahyuagastyakomang@gmail.com
²eka.karyawati@unud.ac.id

Abstract

Patient safety has always been a top priority, and medication error is a significant concern in this domain. Prescriptions are a crucial factor that can elevate the risk of medication errors. A study by Prof. Dr. Nurul Idrus involving 1,200 drug users revealed that a substantial number of respondents misused prescriptions and experimented with obtained medications to substitute for those they couldn't easily access. To address this issue and enhance patient safety, we propose utilizing the AES encryption algorithm to safeguard the confidentiality of prescriptions. AES encryption offers robust protection against unauthorized access and data breaches. Furthermore, employing the CBC (Cipher Block Chaining) mode provides an additional layer of security. In CBC mode, each block of the message is encrypted not only with the encryption key but also with the ciphertext of the previous block, resulting in a unique encrypted message even for identical plaintexts. This combination of AES encryption and CBC mode effectively safeguards prescription data, minimizing the potential for prescription misuse and medication errors, ultimately contributing to improved patient safety.

Keywords: Prescription misuse, AES, CBC

1. Pendahuluan

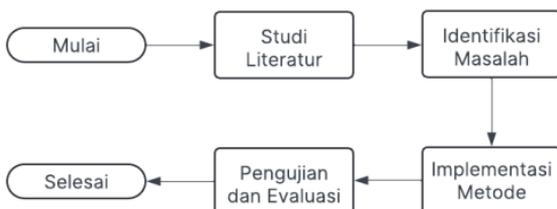
Patient safety telah menjadi perhatian bagi setiap tenaga kesehatan di seluruh dunia sejak dahulu. Kesalahan pengobatan (medication error), yang seharusnya dapat dicegah, masih di bawah kontrol atau tanggung jawab tenaga kesehatan, adalah salah satu hal yang terkait erat dengan patient safety [1]. Pada fase komunikasi non-verbal antara dokter dan apoteker tentang pengobatan pasien, dapat terjadi medication error [2]. Salah satu faktor yang dapat meningkatkan kemungkinan medication error adalah resep. Sebuah laporan dari Institut Kedokteran Amerika menunjukkan bahwa antara 44,000 dan 98,000 orang telah meninggal karena medication error, atau sekitar 7000 orang meninggal setiap tahun sebagai akibat dari medication error yang sering terjadi. [3]. Penyalahgunaan resep dokter marak terjadi dengan dalih mencari alternatif dari obat-obatan yang dilarang keras peredarannya. Dari 1.200 pengguna narkoba yang dipilih untuk penelitian oleh Prof. Dr. Nurul Idrus, sebagian mengatakan bahwa mereka bereksperimen menggunakan obat yang diresepkan dokter sebagai pengganti obat yang sulit mereka dapatkan. [4]. Resep adalah alat komunikasi profesional antara dokter (penulis resep), APA (penyedia/pembuat obat), dan pasien. [5] oleh karena itu, hendaknya resep bersifat rahasia agar tidak dapat dilihat oleh orang yang tidak berhak. Dengan melakukan enkripsi terhadap resep, maka kerahasiaan resep dapat ditingkatkan serta penyalahgunaan resep dapat dikurangi. Enkripsi AES-CBC adalah algoritma enkripsi simetris yang kuat dan aman yang dapat membantu meningkatkan keamanan data resep dan mengurangi penyalahgunaan obat. AES bekerja dengan memecah pesan menjadi blok-blok dan mengenkripsinya menggunakan kunci rahasia. Mode CBC menambahkan lapisan keamanan ekstra dengan menggabungkan setiap blok pesan dengan blok acak (Initialization Vector) sebelum dienkripsi, sehingga menghasilkan pesan terenkripsi yang unik setiap kali meski pesan aslinya sama. Proses algoritma enkripsi AES terdiri

dari empat kategori transformasi byte yaitu SubBytes, ShiftRows, MixColumns, dan AddRoundKey. [6] [7]. Input plaintext akan mempertahankan transformasi byte AddRoundKey pada langkah pertama enkripsi. Pada langkah-langkah berikutnya, state akan mengalami transformasi SubBytes, ShiftRows, MixColumns, dan AddRoundKey berulang sebanyak N_r . State tidak akan mengalami transformasi MixColumns lagi pada putaran terakhir. [8].

2. Metode Penelitian

2.1. Diagram Alir Penelitian

Penelitian ini dimulai dengan melakukan studi literatur tentang penelitian terkait dengan enkripsi AES mode CBC, selanjutnya memahami permasalahan yang ingin diselesaikan, lalu mengimplementasikan algoritma kedalam sistem, dan terakhir dilakukan pengujian dan evaluasi.



Gambar 1. Diagram Alir Penelitian

2.2 Studi Literatur

Langkah pertama dalam penelitian ini adalah melakukan tinjauan pustaka untuk mempelajari kriptografi dan Algoritma AES serta mode CBC. Hal ini dilakukan dengan cara membaca dan memahami berbagai sumber seperti jurnal ilmiah, makalah, dan referensi lainnya. Tujuannya adalah untuk mendapatkan informasi yang mendalam dan akurat yang diperlukan untuk penelitian ini.

2.3 Identifikasi Masalah

Pada tahap ini, dilakukan identifikasi masalah untuk memahami permasalahan utama yang ingin diselesaikan. Permasalahan dalam penelitian ini adalah penyalahgunaan resep dokter sehingga maraknya penyalahgunaan obat-obatan yang dilarang keras peredarannya. Maka dari itu, pengenkripsian file resep dokter menjadi salah satu solusi yang efektif dalam menjaga kerahasiaan isi resep pasien.

2.4 Implementasi Metode

Tahap ini adalah pengimplementasian algoritma AES mode CBC ke dalam sistem. Sistem nantinya akan berjalan pada web, digunakan bahasa HTML, CSS, dan javascript untuk membangun sistem. Algoritma AES mode CBC akan diterapkan menggunakan bahasa pemrograman javascript. Proses enkripsi dan dekripsi akan menggunakan satu kata sandi yang sama berjumlah delapan karakter atau lebih untuk memastikan keamanan dan kesamaan data.

2.5 Pengujian dan Evaluasi

Pada tahap ini dilakukan pengujian jalannya algoritma AES mode CBC yang telah diterapkan, apakah algoritma telah berjalan baik atau ada kendala. Selanjutnya dilakukan evaluasi apabila masih terdapat kekurangan dalam sistem.

2.6 Kriptografi

Kriptografi adalah bidang yang mempelajari cara menyandikan pesan dalam bentuk yang tidak dapat dipahami lagi untuk menjaga kerahasiaan pesan [9]. Kriptografi terdiri dari proses enkripsi dan proses dekripsi. Proses enkripsi adalah mengonversikan informasi yang dapat dibaca (plaintext) menjadi informasi yang tidak terbaca (ciphertext), sedangkan proses dekripsi adalah membalikkan dari informasi yang tidak terbaca (ciphertext) menjadi informasi yang dapat dibaca (plaintext).

2.7 Advanced Encryption Standard

AES, atau Advanced Encryption Standard, adalah algoritma simetri dan cipher blok yang dikenal sebagai Rijndael, yang ditemukan oleh Dr. Vincent Rijmen dan Dr. Joan Daemen [10]. Setiap cipher mengenkripsi dan mendekripsi data dalam blok 128-bit menggunakan kunci kriptografi masing-masing 128, 192, dan 256 bit. Kunci 128-bit, 192-bit, dan 256-bit menjalani 10, 12, dan 14 putaran enkripsi, secara berurutan. Satu putaran terdiri dari beberapa langkah pemrosesan termasuk substitusi, transposisi, dan pencampuran input teks biasa untuk mengubahnya menjadi output ciphertext akhir. Semakin banyak putaran, semakin sulit untuk memecahkan enkripsi, dan semakin aman informasi asli. Baik dalam enkripsi maupun dekripsi, algoritma ini menggunakan satu kunci yang sama, dan input serta outputnya berupa blok sejumlah bit tertentu. Algoritma AES terdiri dari empat proses transformasi byte yaitu SubBytes, ShiftRows, MixColumns, dan AddRoundKey. [6] [7].

2.8 Cipher Block Chaining

CBC (singkatan dari cipher-block chaining) adalah mode block cipher AES yang lebih unggul dari mode ECB dalam hal menyembunyikan pola pada plaintext. Mode CBC mencapai hal ini dengan melakukan XOR antara blok plaintext pertama (B_1) dengan initialization vector sebelum dienkripsi. CBC juga menggunakan chaining blok di mana setiap blok plaintext selanjutnya di-XOR dengan ciphertext dari blok sebelumnya. Jika dituliskan dalam notasi matematis, rumusnya akan menjadi:

$$C_i = E_K(B_i \oplus C_{i-1}) \quad (1)$$

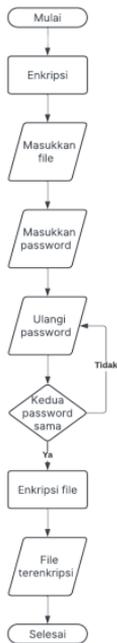
di mana E_K menunjukkan algoritma enkripsi blok menggunakan kunci K , dan C_{i-1} adalah cipher yang sesuai dengan B_{i-1} . Untuk dekripsinya secara matematis akan menjadi:

$$B_i = D_K(C_i) \oplus (C_{i-1}) \quad (2)$$

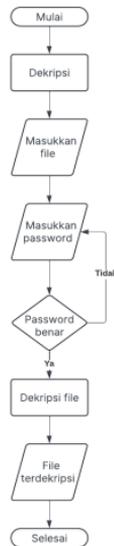
di mana D_K menunjukkan algoritma enkripsi blok menggunakan kunci K .

2.9 Desain Sistem

Sistem akan terdiri dari dua proses yaitu proses enkripsi dan proses dekripsi. Masing-masing proses dapat dilihat pada Gambar 2 dan Gambar 3. Proses Enkripsi dapat dilihat pada Gambar 2, dimulai dari memasukkan file resep yang hendak dienkripsi, selanjutnya dilanjutkan dengan membuat password enkripsi dan mengulangi password yang sama. Sistem akan mengecek apakah kedua password telah sama dan memenuhi persyaratan. Jika kedua password telah sama dan memenuhi persyaratan, akan muncul tombol untuk mengenkripsi file resep. Jika kedua password tidak sama/belum memenuhi persyaratan, maka dilakukan pengulangan untuk memasukkan password. Selanjutnya sistem akan mengenkripsi file resep menggunakan algoritma AES dengan mode CBC, lalu file resep yang telah dienkripsi dapat diunduh. Proses dekripsi dapat dilihat pada Gambar 3, dimulai dari memasukkan file resep yang telah dienkripsi sebelumnya, lalu dilanjutkan dengan memasukkan password yang telah digunakan untuk melakukan enkripsi sebelumnya. Sistem akan mengecek apakah password yang dimasukkan sama dengan password yang digunakan untuk melakukan enkripsi, jika salah maka dapat dilakukan penginputan password lagi, dan jika benar maka file resep terenkripsi akan didekripsi oleh sistem, lalu file terdekripsi dapat diunduh.



Gambar 2. Diagram Alir Enkripsi



Gambar 3. Diagram Alir Dekripsi

3. Hasil dan Diskusi

3.1. Implementasi Sistem

Sistem ini akan terbagi menjadi dua bagian yaitu proses enkripsi dan dekripsi. Masing-masing proses akan berjalan di web tanpa memerlukan koneksi internet agar dapat memastikan bahwa tidak ada data yang bocor dan tersebar ke internet.

3.1.1. Tampilan Proses Enkripsi

Saat hendak melakukan proses enkripsi, dibutuhkan file resep berupa file berformat .txt, .doc, .jpg, .png ataupun .pdf. Selanjutnya dapat memasukkan password dan perulangan password. Jika sudah, dapat dilakukan proses enkripsi. Proses enkripsi akan menghasilkan file berformat sama dengan format asli yang ditambahkan format.enc. Gambar 4 merupakan tampilan proses enkripsi resep dokter.



Gambar 4. Tampilan Proses Enkripsi

3.1.2. Tampilan Proses Dekripsi

Saat hendak melakukan proses dekripsi, dibutuhkan file resep terenkripsi berupa file berformat .txt, .doc, .jpg, .png ataupun .pdf yang telah ditambahkan format .enc sebelumnya. Selanjutnya dapat memasukkan password yang telah dibuat sebelumnya. Jika sudah, dapat dilakukan proses dekripsi. Proses dekripsi akan menghasilkan file berformat sama dengan format asli yang ditambahkan format.dec. Gambar 5 merupakan tampilan untuk proses dekripsi resep dokter.

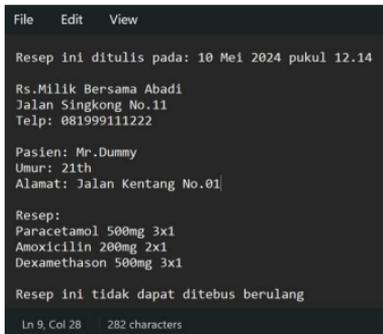


Gambar 5. Tampilan Proses Dekripsi

3.2. Pengujian Sistem

3.2.1. Pengujian Enkripsi AES mode CBC

Pengujian enkripsi algoritma AES dengan mode CBC pada penelitian ini menggunakan data resep dummy yang telah peneliti buat sebelumnya dengan harapan untuk menguji apakah algoritma AES mode CBC berhasil diterapkan atau tidak.



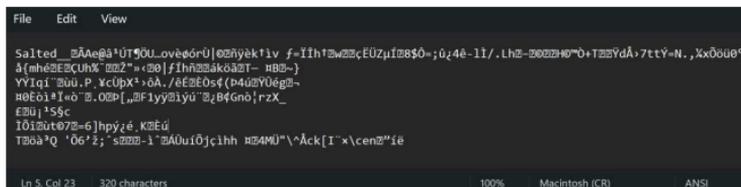
Gambar 6. Data Dummy Plaintext Berformat .txt

Gambar 6 merupakan salah satu contoh data dummy berupa plaintext, data resep tersebut ditulis dalam bahasa Indonesia yang dapat dibaca dan dipahami oleh seseorang. Data dummy adalah informasi yang tidak mengandung data berguna apapun, tetapi berfungsi untuk memesan ruang di mana data nyata secara nominal hadir. Data dummy dapat digunakan sebagai placeholder untuk tujuan pengujian dan operasional. Untuk pengujian, data dummy juga dapat digunakan sebagai stub atau pad untuk menghindari masalah pengujian perangkat lunak dengan memastikan bahwa semua variabel dan bidang data terisi. Data dummy resep dokter untuk pengujian algoritma AES adalah data buatan yang menyerupai resep dokter asli, namun tidak mengandung informasi pasien yang sebenarnya. Data ini digunakan untuk menguji kinerja algoritma AES dalam mengenkripsi dan mendekripsi data medis sensitif.



Gambar 7. Proses Enkripsi

Gambar 7 menunjukkan proses untuk melakukan enkripsi file resep. File dummy bernama "Resep Obat Mr.Dummy.txt" berukuran 297 Bytes telah dimasukkan ke sistem. Pengguna telah memasukkan password yang valid dan file resep telah terenkripsi dilihat dari pemberitahuan "File Encrypted" yang berwarna hijau.



Gambar 8. File Terenkripsi berupa Chipertext

Gambar 8 menunjukkan hasil enkripsi berupa chipertext yang sudah tidak dapat terbaca lagi. Chipertext hasil dari enkripsi resep dokter menggunakan algoritma AES adalah string teks yang tidak dapat dipahami yang mewakili resep dokter asli. Chipertext ini dihasilkan dengan menggabungkan resep dokter asli dengan kunci enkripsi AES. Algoritma AES menggunakan serangkaian substitusi dan permutasi yang kompleks untuk mengubah data asli menjadi chipertext.

3.2.2. Pengujian Dekripsi AES mode CBC

Pengujian dekripsi algoritma AES dengan mode CBC pada penelitian ini menggunakan data resep dummy yang telah peneliti buat dan dekripsi sebelumnya dengan harapan untuk menguji apakah algoritma AES mode CBC berhasil diterapkan atau tidak. Gambar 9 menunjukkan proses untuk melakukan dekripsi. File dummy bernama "Resep Obat Mr.Dummy.txt.enc" berukuran 320 Bytes telah dimasukkan ke sistem. Pengguna telah memasukkan password enkripsi yang valid dan file resep telah berhasil didekripsi, dapat dilihat dari pemberitahuan "File decrypted" yang berwarna hijau. Gambar 10 menunjukkan hasil dekripsi resep dokter berupa plaintext yang dapat terbaca dengan berbahasa Indonesia sesuai dengan plaintext asli sebelum dienkripsi. Ini menunjukkan bahwa proses enkripsi serta dekripsi telah berhasil dilakukan tanpa mengubah informasi dalam file resep asli. File hasil dekripsi yang diperoleh berukuran 297 Bytes sama dengan file asli namun file hasil dekripsi berukuran 320 Bytes menunjukkan ada peningkatan ukuran dalam file hasil dekripsi.



Gambar 9. Proses Dekripsi

```

File Edit View

Resep ini ditulis pada: 10 Mei 2024 pukul 12.14

Rs.Milik Bersama Abadi
Jalan Singkong No.11
Telp: 081999111222

Pasien: Mr.Dummy
Umur: 21th
Alamat: Jalan Kentang No.01

Resep:
Paracetamol 500mg 3x1
Amoxicilin 200mg 2x1
Dexamethason 500mg 3x1

Resep ini tidak dapat ditebus berulang

Ln 1, Col 1 282 characters 100% Windows (CRLF) UTF-8
    
```

Gambar 10. Plaintext Hasil Dekripsi

3.3. Pengujian Akurasi Sistem

Menguji tingkat keberhasilan sistem akan dilakukan pengujian terhadap 20 file resep dokter berbeda.

3.3.1. Pengujian Akurasi Enkripsi

Tabel 1. Hasil Pengujian Enkripsi

No	Nama File	Ukuran Asli	Nama File Terenkripsi	Ukuran File Terenkripsi	Keterangan
1	Resep Obat Dummy.txt	297 Bytes	Resep Obat Dummy.txt.enc	320 Bytes	Berhasil
2	presc1.txt	146 Bytes	presc1.txt.enc	176 Bytes	Berhasil
3	presc2.txt	147 Bytes	presc2.txt.enc	176 Bytes	Berhasil
4	presc3.txt	144 Bytes	presc3.txt.enc	176 Bytes	Berhasil
5	presc4.txt	282 Bytes	presc4.txt.enc	304 Bytes	Berhasil
6	presc5.txt	353 Bytes	presc5.txt.enc	384 Bytes	Berhasil
7	presc6.txt	342 Bytes	presc6.txt.enc	368 Bytes	Berhasil
8	presc7.txt	880 Bytes	presc7.txt.enc	912 Bytes	Berhasil
9	presc8.txt	491 Bytes	presc8.txt.enc	512 Bytes	Berhasil
10	presc9.txt	246 Bytes	presc9.txt.enc	272 Bytes	Berhasil
11	presc10.txt	517 Bytes	presc10.txt.enc	544 Bytes	Berhasil
12	presc11.txt	140 Bytes	presc11.txt.enc	160 Bytes	Berhasil
13	presc12.txt	349 Bytes	presc12.txt.enc	368 Bytes	Berhasil
14	presc13.txt	703 Bytes	presc13.txt.enc	720 Bytes	Berhasil
15	presc14.txt	401 Bytes	presc14.txt.enc	432 Bytes	Berhasil
16	presc15.txt	331 Bytes	presc15.txt.enc	352 Bytes	Berhasil
17	presc16.txt	355 Bytes	presc16.txt.enc	384 Bytes	Berhasil

No	Nama File	Ukuran Asli	Nama File Terenkripsi	Ukuran File Terenkripsi	Keterangan
18	presc17.txt	510 Bytes	presc17.txt.enc	528 Bytes	Berhasil
19	presc18.txt	245 Bytes	presc18.txt.enc	272 Bytes	Berhasil
20	presc19.txt	385 Bytes	presc19.txt.enc	416 Bytes	Berhasil

Telah dilakukan pengujian enkripsi dengan dua puluh file resep berbeda. Pada tabel 1 dapat dilihat bahwa ukuran file asli mengalami perubahan setelah dilakukan enkripsi. Algoritma enkripsi AES CBC (Cipher Block Chaining) membutuhkan blok data input yang memiliki panjang yang kelipatan dari ukuran blok AES (biasanya 128 bit). Jika panjang data input tidak kelipatan dari ukuran blok, maka data input harus dipad dengan bit tambahan agar panjangnya menjadi kelipatan dari ukuran blok. Bit padding ini akan meningkatkan ukuran file setelah enkripsi. Selain itu format file juga akan berubah setelah dienkripsi, sistem akan menambahkan format .enc pada file terenkripsi. Pengujian enkripsi terhadap kedua puluh file tersebut berhasil dilakukan.

Tabel 2. Hasil Pengujian Dekripsi

No	Nama File	Ukuran	Nama File Terdekripsi	Ukuran File Terdekripsi	Keterangan
1	Resep Obat Dummy.txt	320 Bytes	Resep Obat Dummy.txt.enc	297 Bytes	Berhasil
2	presc1.txt.enc	176 Bytes	presc1.txt.enc.dec	146 Bytes	Berhasil
3	presc2.txt.enc	176 Bytes	presc2.txt.enc.dec	147 Bytes	Berhasil
4	presc3.txt.enc	176 Bytes	presc3.txt.enc.dec	144 Bytes	Berhasil
5	presc4.txt.enc	304 Bytes	presc4.txt.enc.dec	282 Bytes	Berhasil
6	presc5.txt.enc	384 Bytes	presc5.txt.enc.dec	353 Bytes	Berhasil
7	presc6.txt.enc	368 Bytes	presc6.txt.enc.dec	342 Bytes	Berhasil
8	presc7.txt.enc	912 Bytes	presc7.txt.enc.dec	880 Bytes	Berhasil
9	presc8.txt.enc	512 Bytes	presc8.txt.enc.dec	491 Bytes	Berhasil
10	presc9.txt.enc	272 Bytes	presc9.txt.enc.dec	246 Bytes	Berhasil
11	presc10.txt.enc	544 Bytes	presc10.txt.enc.dec	517 Bytes	Berhasil
12	presc11.txt.enc	160 Bytes	presc11.txt.enc.dec	140 Bytes	Berhasil
13	presc12.txt.enc	368 Bytes	presc12.txt.enc.dec	349 Bytes	Berhasil
14	presc13.txt.enc	720 Bytes	presc13.txt.enc.dec	703 Bytes	Berhasil
15	presc14.txt.enc	432 Bytes	presc14.txt.enc.dec	401 Bytes	Berhasil
16	presc15.txt.enc	352 Bytes	presc15.txt.enc.dec	331 Bytes	Berhasil
17	presc16.txt.enc	384 Bytes	presc16.txt.enc.dec	355 Bytes	Berhasil
18	presc17.txt.enc	528 Bytes	presc17.txt.enc.dec	510 Bytes	Berhasil
19	presc18.txt.enc	272 Bytes	presc18.txt.enc.dec	245 Bytes	Berhasil
20	presc19.txt.enc	416 Bytes	presc19.txt.enc.dec	385 Bytes	Berhasil

Selanjutnya dilakukan pengujian dekripsi dengan dua puluh file resep berbeda yang telah dienkripsi sebelumnya. Pada tabel 2 dapat dilihat bahwa ukuran file mengalami perubahan ke ukuran file asli sebelum di enkripsi. Saat data dienkripsi dengan AES CBC, padding yang

ditambahkan pada langkah enkripsi dihapus pada langkah dekripsi. Hal ini mengembalikan panjang data ke ukuran semula. Mode CBC AES menambahkan blok ciphertext dari blok sebelumnya ke blok plaintext saat ini sebelum enkripsi. Saat data didekripsi, operasi ini dibalik, sehingga blok ciphertext pertama didekripsi menggunakan kunci enkripsi saja, dan blok ciphertext selanjutnya didekripsi dengan menggunakan blok ciphertext sebelumnya dan kunci enkripsi. Hal ini mengembalikan data ke format plaintext asli, dengan panjang yang sama dengan plaintext sebelum enkripsi. Sistem menambahkan format .dec pada file hasil dekripsi. Pada kedua puluh file yang dilakukan pengujian, semua file berhasil didekripsi.

4. Kesimpulan

Dapat disimpulkan bahwa enkripsi dan dekripsi menggunakan AES-CBC untuk mengamankan resep dokter berhasil dilakukan. Hasil enkripsi menunjukkan bahwa resep dokter yang telah dienkripsi tidak dapat dibaca oleh pihak yang tidak berwenang. Hal ini dibuktikan dengan tidak terbacanya pesan yang telah dienkripsi. Proses dekripsi juga berhasil dilakukan, dimana resep dokter yang telah didekripsi kembali menjadi pesan yang sama dengan resep asli. Hal ini menunjukkan bahwa proses enkripsi dan dekripsi tidak merusak data resep dokter. Secara keseluruhan, penelitian ini menunjukkan bahwa AES-CBC merupakan metode yang efektif untuk mengamankan resep dokter dan meminimalisir penyalahgunaan obat.

Daftar Pustaka

- [1] M. R. COHEN, "Medication Errors," American Journal of Pharmaceutical Education, 2007.
- [2] M. J. Rantucci, Komunikasi Apoteker-Pasien Panduan Konseling Pasien, Jakarta: EGC, 2009.
- [3] I. Lisni, N. E. Gumilang and E. Kusumahati, "Potensi Medication error Pada Resep di Salah Satu Apotek di Kota Kadipaten," Jurnal Sains dan Kesehatan, vol. 3, no. 4, pp. 558-568, 2021.
- [4] Admin, "Universitas Gadjah Mada," Universitas Gadjah Mada, 7 January 2014. [Online]. Available: <https://ugm.ac.id/id/berita/8575-pemakai-narkoba-gunakan-obat-terlarang-dari-resep-dokter/>. [Accessed 9 May 2024].
- [5] C. S. Lestari, Seni Menulis Resep, Jakarta: PT Perca, 2002.
- [6] T. M. Kumar and P. Karthigaikumar, "FPGA implementation of an optimized key expansion module of AES algorithm for secure transmission of personal ECG signals," Design Automation for Embedded Systems, vol. 22, no. 1-2, pp. 13-24, 2018.
- [7] H. M. Mohammad and A. A. Abdullah, "Enhancement process of AES: a lightweight cryptography algorithm-AES for constrained devices," TELKOMNIKA (Telecommunication Computing Electronics and Control), vol. 20, no. 3, pp. 551-560, 2022.
- [8] F. P. Utama, G. Wijaya, R. Faurina and A. Vatesria, "Implementasi Algoritma Aes 256 Cbc, Base 64, Dan Sha 256 Dalam Pengamanan Dan Validasi Data Ujian Online," Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK), vol. 10, no. 5, pp. 945-954, 2014.
- [9] M. Azhari, D. I. Mulyana, F. J. Perwitosari and F. Ali, "Implementasi Pengamanan Data pada Dokumen Menggunakan Algoritma Kriptografi Advanced Encryption Standard(AES)," Jurnal Pendidikan Sains dan Komputer, vol. 2, no. 1, 2022.
- [10] Henry, A. H. Kridalaksana and Z. Arifin, "Kriptografi Aes Mode Cbc Pada Citra Digital," Prosiding Seminar Ilmu Komputer dan Teknologi Informasi, vol. 1, no. 1, 2016.

Analisis Penggunaan Logika Fuzzy Mamdani dan Sugeno untuk Memprediksi Shade Foundation

Aprinia Salsabila Roiqoh^{a1}, Hanin Fatma Soraya^{a2}, Dela Ayu Putri Mayona^{b3}, Nabila Anggita Luna^{b4}, Anggraini Puspita Sari^{c5}

^aFakultas Ilmu Komputer, UPN "Veteran"
Jawa Timur, Surabaya
122081010166@student.upnjatim.ac.id
222081010069@student.upnjatim.ac.id
³22081010008@student.upnjatim.ac.id
⁴22081010057@student.upnjatim.ac.id
⁵anggraini.puspita.if@upnjatim.ac.id

Abstract

The magnificence industry has developed quickly in later decades, expanding request for items that meet an assortment of shopper needs. One imperative angle in choosing magnificence items is finding an establishment color that suits your skin color. This investigates points to analyze the utilize of fluffy rationale, particularly the Mamdani and Sugeno strategy, in foreseeing the correct establishment shade based on varieties in skin color and suggestion. Definition of membership functions, definition of fuzzy rules, fuzzy interference and defuzzification are used as research methods. The data used was obtained from a random experiment by entering skin tone and undertone values into the program. Research results show that the shade of the foundation is greatly influenced by the color and undertone of the skin. Although there is a significant difference between the Mamdan and Sugeno method values, the final predicted base colors are not significantly different. This research strengthens the position of fuzzy logic as an effective method in improving the quality of products and services in the beauty industry, as well as solving the complex problem of determining the appropriate foundation shade for various skin types. It is hoped that the results of this research will make it easier for consumers and beauty professionals to choose the right foundation shade according to individual needs and preferences.

Keywords: Fuzzy logic, Mamdani method, Sugeno method, foundation shade, skin color, undertone

1. Pendahuluan

Industri kecantikan telah mengalami pertumbuhan yang signifikan selama beberapa dekade terakhir, meningkatkan permintaan akan produk yang memenuhi berbagai kebutuhan konsumen. Kuncinya dalam industri ini adalah memilih warna dasar yang sesuai dengan warna kulit seseorang. Memilih warna dasar yang tepat akan meningkatkan kepercayaan diri Anda dan memastikan tampilan natural. Namun, mengingat variasi warna kulit yang kompleks dan beragam, menemukan warna dasar yang sempurna seringkali sulit. Logika fuzzy, salah satu cabang kecerdasan buatan, mengacu pada pengetahuan yang memungkinkan komputer meniru kecerdasan manusia dan melakukan hal-hal yang dilakukan manusia. Logika fuzzy, juga disebut logika fuzzy, mengacu pada metode di mana ruang masukan dapat dipetakan ke ruang keluaran berdasarkan konsep himpunan fuzzy. Logika fuzzy juga mudah dipahami, sangat fleksibel, dan memberikan toleransi terhadap data yang tidak akurat. Oleh karena itu fuzzy sangat cocok digunakan dalam menentukan besaran. [1]. Logika fuzzy memiliki kemampuan untuk mengembangkan sistem fuzzy, sistem kecerdasan dalam lingkungan yang tidak pasti. Beberapa tahapan dalam proses pembentukan sistem fuzzy adalah analisis input dan output, penentuan variabel input dan output, penentuan fungsi keanggotaan setiap himpunan fuzzy, dan penetapan aturan berdasarkan aturan empiris atau pakar. Bidang Pengetahuan dan Implementasi Sistem Fuzzy. Secara keseluruhan logika fuzzy menggunakan konsep matematika yang sangat sederhana, mudah dipahami, dan memberikan toleransi terhadap data yang tidak akurat atau

tidak jelas. Sistem fuzzy dapat langsung membangun dan menerapkan pengalaman pakar tanpa melalui proses pelatihan, mentransfer ilmu dari pakar tersebut ke sistem komputer sehingga dapat melakukan upaya pemodelan untuk menjadi suatu sistem. keputusan benar-benar dapat diandalkan.[2] Pemilihan shade foundation telah menjadi topik penting dalam industri kecantikan, di mana teknologi seperti logika fuzzy dan kecerdasan buatan (AI) telah mulai digunakan untuk memperbaiki proses ini. Penelitian oleh Farnoosh et al. (2016) mengkaji penggunaan spektral imaging dan logika fuzzy dalam analisis warna kulit. Mereka menemukan bahwa logika fuzzy memungkinkan pencocokan yang lebih baik antara warna kulit dan foundation dibandingkan teknik manual, dengan mempertimbangkan berbagai variabel seperti undertone dan intensitas cahaya pada kulit. Pengelompokan Warna Fuzzy untuk Diagnosis Melanoma pada Gambar Dermoskopi (mdpi.com). Penelitian terkait penggunaan logika fuzzy dalam deteksi warna kulit telah banyak dilakukan. N. Nguyen, misalnya, mengusulkan penggunaan logika fuzzy dalam deteksi warna kulit berbasis pembelajaran mendalam. Studi ini menunjukkan bahwa kombinasi antara logika fuzzy dan pembelajaran mesin mampu memberikan hasil yang lebih akurat dalam identifikasi warna kulit yang kompleks. [3] Selain itu, penelitian Nguyen lainnya mengintegrasikan jaringan saraf konvolusional (Convolutional Neural Networks/CNN) dengan logika fuzzy untuk mendeteksi warna kulit, menegaskan bahwa pendekatan hibrida ini memiliki potensi besar dalam memproses data gambar dengan presisi yang tinggi. [4] Tujuan dari penelitian ini adalah menganalisis penggunaan logika fuzzy dalam menentukan rona dasar warna kulit yang berbeda. Dalam penelitian ini, kami akan menggunakan teknologi logika fuzzy untuk mengembangkan model prediktif yang dapat memprediksi warna alas bedak yang tepat berdasarkan perubahan warna kulit seperti warna kulit, rona bawah, dan kegelapan. Model ini bertujuan untuk memudahkan konsumen dan profesional kecantikan dalam memilih warna foundation yang sesuai dengan kebutuhan dan kesukaannya. Dengan pendekatan penelitian yang komprehensif dan inovatif, penelitian ini bertujuan untuk memperkuat posisi logika fuzzy sebagai metode efektif untuk meningkatkan kualitas produk dan layanan serta memecahkan masalah kompleks di berbagai bidang aplikasi.

2. Metode Penelitian

Metode Penelitian yang kami gunakan pada penelitian kali ini adalah metode kuantitatif. hipotesis diuji menggunakan angka. Metode ini ditekankan pada hasil pengukuran yang objektif menggunakan analisis statistik. Metode ini dapat digunakan oleh rumpun sosial maupun sains. [5]

2.1 Penggunaan Metode Mamdani

Pada penelitian Analisa Penggunaan Logika Fuzzy Mamdani dan Sugeno untuk Memprediksi Shade Foundation ini nantinya akan memerlukan beberapa langkah langkah yang akan dilakukan. bertujuan untuk memprediksi shade foundation yang sesuai berdasarkan variasi warna kulit. Model ini akan mempertimbangkan berbagai faktor seperti warna kulit dan undertone. berikut ini adalah langkah langkah metode fuzzy mamdani :

- a. Menentukan Fungsi Keanggotaan
Menentukan himpunan-himpunan fuzzy dengan fungsi keanggotaan yang sesuai. Fungsi keanggotaan adalah pengaturan rentang nilai masukan pada seluruh himpunan fuzzy yang ditetapkan.
- b. Mendefinisikan Aturan Fuzzy
Langkah berikutnya adalah membuat aturan fuzzy yang didalamnya terdiri dari himpunan fuzzy masukan dan keluaran. Aturan yang digunakan adalah "IF" (Masukan "warna kulit") AND (Masukan "undertone") THEN (Keluaran "Warna Foundation")
- c. Interferensi Fuzzy
Selanjutnya aturan aturan fuzzy digunakan untuk menghasilkan nilai keanggotaan pada variabel keluaran "warna foundation"dari kombinasi nilai masukan.
- d. Defuzzifikasi
Terakhir, mengubah nilai nilai hasil interferensi fuzzy menjadi nilai crisp. Metode fuzzy yang digunakan Center of Gravity (COG) untuk membulatkan nilai keluaran fuzzy menjadi nilai crisp. [6]

2.2 Penggunaan Metode Sugeno

Pada metode Sugeno, nilai rata-rata digunakan untuk menghitung nilai crisp output, sehingga menghasilkan persamaan linear konstanta sebagai output dari fungsi keanggotaan Sugeno [7]. Karakteristik inferensi fuzzy, yaitu konsekuensi tidak merupakan himpunan fuzzy, Namun, persamaan ini adalah sebuah persamaan linear dengan variabel-variabel yang sesuai dengan variabel inputnya.[8] Untuk mendapatkan hasil, maka terdapat 3 langkah tahapan sebagai berikut:

- a. Pembentukan Himpunan Fuzzy
Dalam proses pembentukan himpunan, semua variabel yang terkait harus ditentukan dengan jelas. Untuk setiap variabel input, fungsi fuzzifikasi harus disesuaikan. Pada metode Sugeno, semua variabel dibagi menjadi satu atau lebih himpunan, termasuk variabel output.
- b. Aplikasi Fungsi Implikasi
Menyusun basis aturan berupa implikasi-implikasi fuzzy yang menyatakan relasi antara variabel input dengan variabel output. Aturan-aturan ini didefinisikan menggunakan IF...THEN.
- c. Defuzzifikasi
Proses defuzzifikasi menerima himpunan fuzzy yang diperoleh dari komposisi aturan-aturan fuzzy sebagai input, dan menghasilkan output berupa bilangan real yang pasti. Dengan kata lain, diberikan suatu himpunan fuzzy dalam rentang tertentu, maka dapat ditentukan nilai pasti tertentu sebagai output. [9]

2.2 Pengumpulan Data

Data yang akan digunakan sebagai bahan pengembangan sistem ini adalah warna kulit dan undertone pengguna. Data diperoleh dari hasil google form yang telah disebar secara acak. Skenario dari percobaan ini adalah memasukkan nilai warna kulit dan undertone Hasil data yang diperoleh dari percobaan tersebut dijadikan sebagai dasar analisa logika fuzzy mamdani dan sugeno. Alasan menggunakan Google Form dapat mengumpulkan data dengan cepat dan efisien, serta memanfaatkan berbagai kelebihan lainnya seperti tidak dipungut biaya, data otomatis terkumpul, dan analisis data yang mudah dilakukan. Pertanyaan yang digunakan untuk mengumpulkan data yang dibutuhkan sebagai berikut

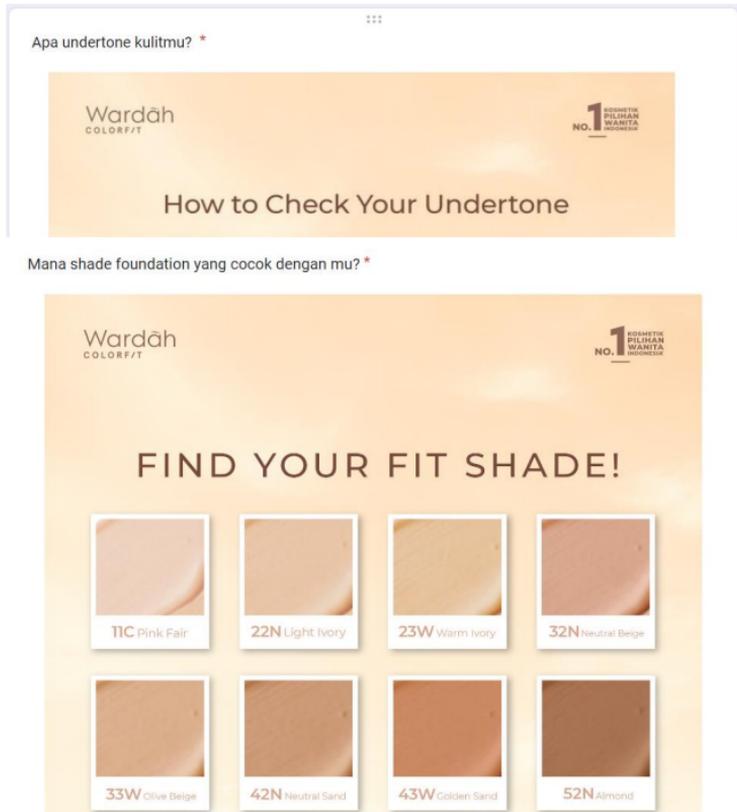
Apa Warna kulitmu? *

Putih Gading

Kuning Langsat

Sawo Matang

Hitam Kecoklatan



2.3 Rancangan Sistem

Variabel fuzzy yang digunakan untuk memprediksi shade foundation dengan metode Mamdani dan Sugeno adalah undertone, warna kulit, dan shade foundation dari data yang diperoleh. Pada penerapan logika fuzzy, langkah pertama yang diupayakan adalah mengatur domain pada himpunan fuzzy di setiap variabel dan fungsi keanggotaannya, dengan mengikuti aturan sebagai berikut:

a. Undertone

Variabel diatur pada skala 1-60 untuk memberikan fleksibilitas yang cukup dalam pengklasifikasian. Ini juga membantu untuk membuat sistem lebih intuitif dan mudah dimengerti oleh pengguna yang mungkin terbiasa dengan skala 1-60.

Tabel 1. Fungsi Keanggotaan Variabel Undertone

Fungsi Keanggotaan	Rentang Nilai
Cool (C)	1-20
Neutral (N)	20-40
Warm (W)	40-60

b. Warna Kulit

Variabel diatur pada skala 1-60 untuk memberikan fleksibilitas yang cukup dalam pengklasifikasian. Ini juga membantu untuk membuat sistem lebih intuitif dan mudah dimengerti oleh pengguna yang mungkin terbiasa dengan skala 1-60.

Tabel 2. Fungsi Keanggotaan Variabel Warna Kulit

Fungsi Keanggotaan	Rentang Nilai
Putih Gading (P)	1-15
Kuning Langsung (K)	16-30
Sawo Matang (S)	31-45
Coklat Kehitaman (C)	46-60

c. Shade Foundation

Variabel diatur pada skala 1-60 mewakili tingkat kecerahan warna foundation dari paling terang ke paling gelap. Ini juga membantu untuk membuat sistem lebih intuitif dan mudah dimengerti oleh pengguna yang mungkin terbiasa dengan skala 1-60

Tabel 3. Fungsi Keanggotaan Variabel Shade Foundation

Fungsi Keanggotaan	Rentang Nilai
Pink Fair	1-10
Light Ivory	11-23
Warm Ivory	24-35
Neutral Beige	36-47
Golden Sand	48-59
Almond	60-70

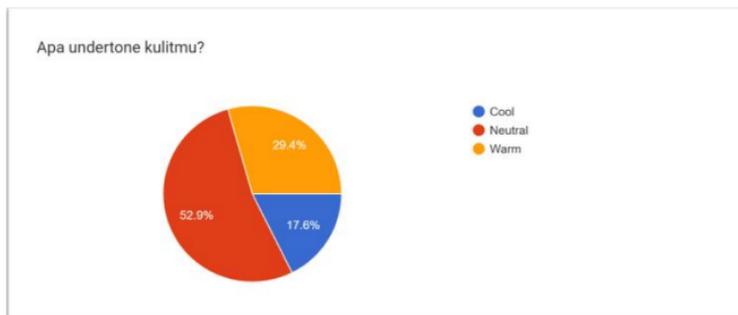
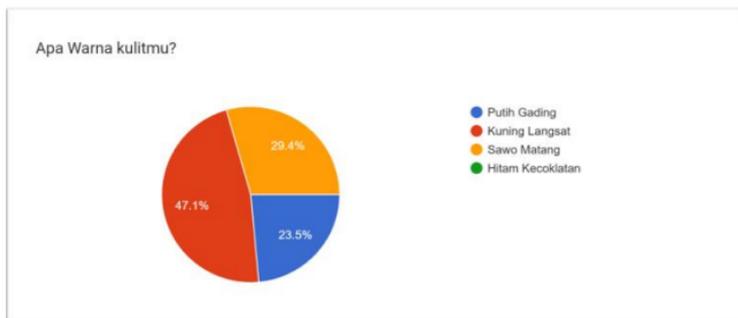
Kemudian, untuk memperoleh nilai prediksi shade foundation perlu dilakukan adalah mendefinisikan indikator untuk setiap keadaan. Oleh karena itu, dibuat aturan logika fuzzy, dengan aturan sebagai berikut:

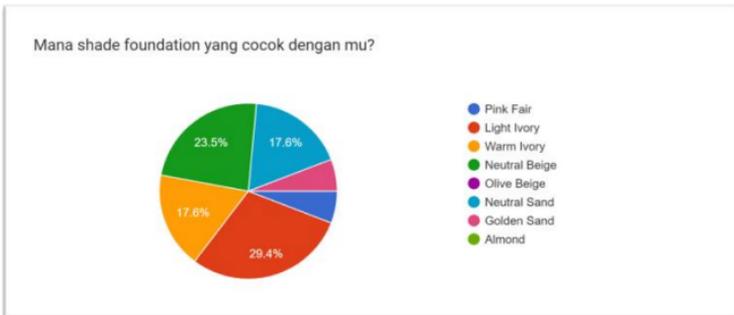
- a. IF undertone == 'C' AND warna_kulit == 'p' THEN shade_foundation == 'Pink Fair'
- b. IF undertone == 'C' and warna_kulit == 'K' THEN shade_foundation == 'Light Ivory'
- c. IF undertone == 'C' and warna_kulit == 'S' THEN shade_foundation == 'Olive Beige'

- d. IF undertone == 'N' and warna_kulit == 'P' THEN shade_foundation == 'Warm Ivory'
- e. IF undertone == 'N' and warna_kulit == 'K' THEN shade_foundation == 'Light Ivory'
- f. IF undertone == 'N' and warna_kulit == 'S' THEN shade_foundation == 'Neutral Sand'
- g. IF undertone == 'N' and warna_kulit == 'C' THEN shade_foundation == 'Golden Sand'
- h. IF undertone == 'W' and warna_kulit == 'P' THEN shade_foundation == 'Neutral Beige'
- i. IF undertone == 'W' and warna_kulit == 'K' THEN shade_foundation == 'Warm Ivory'
- j. IF undertone == 'W' and warna_kulit == 'S' THEN shade_foundation == 'Golden Sand'
- k. IF undertone == 'W' and warna_kulit == 'C' THEN shade_foundation == 'Almond'

3. Hasil dan Pembahasan

Dari pengambilan data yang dilakukan menggunakan google form dihasilkan diagram lingkaran sebagai berikut:





Dari hasil yang sudah diketahui, dibuat tabel data uji yang disesuaikan dengan data diatas. Pengujian terhadap sistem ini dilakukan dengan uji coba memasukkan data uji ke dalam sistem, beberapa data uji yang dipakai tercantum pada Tabel 4.

Tabel 4. Tabel Uji Untuk Pengujian Prediksi Shade Foundation

Warna Kulit	Undertone	Shade Foundation
Putih Gading	Cool	Pink Fair
Putih Gading	Cool	Pink Fair
Kuning Langsat	Cool	Light Ivory
Kuning Langsat	Cool	Light Ivory
Kuning Langsat	Cool	Warm ivory
Kuning Langsat	Cool	Light Ivory
Kuning Langsat	Cool	Neutral Beige
Sawo Matang	Cool	Olive Beige
Sawo Matang	Cool	Neutral Beige
Sawo Matang	Netral	Warm ivory
Kuning Langsat	Netral	Light ivory
Kuning Langsat	Netral	Warm Ivory
Kuning Langsat	Netral	Light Ivory
Sawo Matang	Netral	Warm Ivory
Sawo Matang	Netral	Neutral Sand
Sawo Matang	Netral	Warm Ivory
Kuning Langsat	Netral	Light Ivory
Kuning Langsat	Netral	Light ivory
Sawo Matang	Netral	Olive Beige
Kuning Langsat	Netral	Light Ivory
Hitam Kecoklatan	Netral	Golden Sand
Sawo Matang	Netral	Neutral Sand
Sawo Matang	Netral	Olive Beige
Sawo Matang	Warm	Warm Ivory

Warna Kulit	Undertone	Shade Foundation
Sawo Matang	Warm	Olive Beige
Sawo Matang	Warm	Warm Ivory
Sawo Matang	Warm	Warm Ivory
Kuning Langsung	Warm	Warm ivory
Hitam Kecoklatan	Warm	Golden Sand
Hitam Kecoklatan	Warm	Almond

Berdasarkan tabel uji data, diperoleh prediksi shade foundation dengan metode Mamdani dan Sugeno sebagai berikut:

Tabel 5. Tabel Hasil Pengujian Menggunakan Mamdani

Data Aktual	Data Prediksi
Pink Fair	Pink Fair
Pink Fair	Pink Fair
Light Ivory	Light Ivory
Light Ivory	Light Ivory
Warm ivory	Light Ivory
Light Ivory	Light Ivory
Neutral Beige	Light Ivory
Olive Beige	Olive Beige
Neutral Beige	Olive Beige
Warm ivory	Neutral Sand
Light ivory	Light Ivory
Warm Ivory	Light Ivory
Light Ivory	Light Ivory
Warm Ivory	Neutral Sand
Neutral Sand	Neutral Sand
Warm Ivory	Neutral Sand
Light Ivory	Light Ivory
Light ivory	Light Ivory
Olive Beige	Neutral Sand
Light Ivory	Light Ivory
Golden Sand	Golden Sand
Neutral Sand	Neutral Sand
Olive Beige	Neutral Sand
Warm Ivory	Golden Sand
Olive Beige	Golden Sand
Warm Ivory	Golden Sand
Warm Ivory	Golden Sand
Warm ivory	Warm Ivory

Data Aktual	Data Prediksi
Golden Sand	Almond
Almond	Almond

Tabel 6. Tabel Hasil Pengujian menggunakan Sugeno

Data Aktual	Data Prediksi
Pink Fair	Pink Fair
Pink Fair	Pink Fair
Light Ivory	Light Ivory
Light Ivory	Light Ivory
warm ivory	Light Ivory
Light Ivory	Light Ivory
Neutral Beige	Light Ivory
Olive Beige	Olive Beige
Neutral Beige	Olive Beige
Warm ivory	Neutral Sand
Light ivory	Light Ivory
Warm Ivory	Light Ivory
Light Ivory	Light Ivory
Warm Ivory	Neutral Sand
Neutral Sand	Neutral Sand
Warm Ivory	Neutral Sand
Light Ivory	Light Ivory
Light ivory	Light Ivory
Olive Beige	Neutral Sand
Light Ivory	Light Ivory
Golden Sand	Golden Sand
Neutral Sand	Neutral Sand
Olive Beige	Neutral Sand
Warm Ivory	Golden Sand
Olive Beige	Golden Sand
Warm Ivory	Warm Ivory
Golden Sand	Almond
Almond	Almond

Dari tabel hasil prediksi diatas, diperoleh bahwa shade foundation sangat dipengaruhi oleh warna kulit dan undertone. Namun terdapat perbedaan mencolok antara data aktual dengan hasil prediksi. Hal ini dapat terjadi karena nilai fuzzy tidak seefektif penglihatan manusia. Meskipun menggunakan parameter yang sama, apa yang dipahami fuzzy (aturan fuzzy) yang didapat setelah pengamatan masih dapat sangat berbeda dengan kenyataan.

4. Kesimpulan

Dari hasil penelitian sistem prediksi shade foundation menggunakan pendekatan fuzzy, dengan menggunakan metode mamdani dan sugeno, diperoleh kesimpulan sebagai berikut:

- a. Metode Mamdani dan Sugeno menghasilkan hasil yang kurang akurat namun memiliki potensi untuk dikembangkan lebih lanjut. Tabel menunjukkan kesesuaian sekitar 50% untuk Mamdani dan 48% untuk Sugeno. Meskipun hasil yang didapat kurang akurat, dapat dilihat bahwa Mamdani menghasilkan nilai yang lebih detail dan fleksibel, sedangkan Sugeno menghasilkan nilai yang saklek.
- b. Dilihat dari kesimpulan sebelumnya, dapat disimpulkan lagi bahwa metode Mamdani memiliki akurasi yang lebih baik dibandingkan dengan metode sugeno. Hal ini juga bisa dilihat melalui perbandingan tabel hasil dan tabel uji.

Daftar Pustaka

- [1] N. W. Pardede, "Implementasi Fuzzy Logic Menentukan Jurusan di Smk Dengan Menggunakan Metode Sugeno," UPB Repository. Accessed: Jun. 19, 2024. [Online]. Available: <http://repository.upbatam.ac.id/2407/>
- [2] M. Kastina and M. Silalahi, "Logika Fuzzy Metode Mamdani Dalam Sistem Keputusan Fuzzy Produksi Menggunakan Matlab," JIK: Jurnal Ilmu Komputer, vol. 1, no. 02, Jan. 2016, doi: 10.47007/komp.v1i2.1871.
- [3] S. K. Singh, V. Abolghasemi, and M. H. Anisi, "Fuzzy Logic with Deep Learning for Detection of Skin Cancer," Applied Sciences, vol. 13, no. 15, p. 8927, Aug. 2023, doi: 10.3390/app13158927.
- [4] H. Almubarak, R. Stanley, W. Stoecker, and R. Moss, "Fuzzy Color Clustering for Melanoma Diagnosis in Dermoscopy Images," Information, vol. 8, no. 3, p. 89, Jul. 2017, doi: 10.3390/info8030089.
- [5] F. Afra, "3 Contoh Metode Penelitian yang Bisa Diterapkan dalam Penelitian Ilmiah," detikcom. Accessed: Jun. 19, 2024. [Online]. Available: <https://www.detik.com/edu/detikpedia/d-6945247/3-contoh-metode-penelitian-yang-bisa-diterapkan-dalam-penelitian-ilmiah>
- [6] "Fortuna, D. R. F., Rahmadi, I. F., Ardyananda, D. A., & Sari, A. P. (2023). Analisis Prediksi Intensitas Cahaya Lampu Menggunakan Metode Fuzzy Mamdani dan Sugeno. Seminar Nasional Informatika Bela Negara (SANTIKA), 3, 2747-0563.," Bing. [Online]. Available:
- [7] S. Nurdini, G. W. Nurcahyo, and J. Santony, "Analisis Perkiraan Jumlah Produksi Tahu Menggunakan Metode Fuzzy Sugeno," Jurnal Sistem Informasi dan Teknologi, pp. 18–23, Sep. 2019, doi: 10.37034/jsisifotek.v1i3.4.
- [8] R. Bakri, A. N. Rahma, I. Suryani, and Y. Sari, "Penerapan Logika Fuzzy Dalam Menentukan Jumlah Peserta Bpjs Kesehatan Menggunakan Fuzzy Inference System Sugeno," Jurnal Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika, vol. 1, no. 3, pp. 182–192, Dec. 2020, doi: 10.46306/lb.v1i3.38.
- [9] S. L. M. Sitio, "Penerapan Fuzzy Inference System Sugeno untuk Menentukan Jumlah Pembelian Obat (Studi Kasus: Garuda Sentra Medika)," Jurnal Informatika Universitas Pamulang, vol. 3, no. 2, p. 104, Jun. 2018, doi: 10.32493/informatika.v3i2.1522.

Segmentasi Pengguna Spotify Berdasarkan Preferensi Musik dengan Algoritma K-Means Clustering

Kadek Bisma Dharmasena^{a1}, Cokorda Pramatha^{a2}

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

Jln. Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, 08261, Bali, Indonesia

¹bisma2412@gmail.com

²cokorda@unud.ac.id

Abstract

Music streaming platforms like Spotify have become integral to the daily lives of millions globally, offering personalized listening experiences. However, managing a vast music catalog to present relevant content to each user remains a challenge. This study explores the application of the K-means clustering algorithm to segment Spotify users based on their music preferences. The goal is to group users into clusters with similar tastes to enhance targeted marketing and user engagement. We utilized a secondary dataset of trending Spotify songs and their attributes from 2023. Through data preprocessing, feature selection, and normalization, we prepared the data for clustering. The optimal number of clusters was determined using the Elbow Method, resulting in six distinct clusters. Each cluster represents unique music preferences, analyzed through metrics such as danceability, energy, and popularity. The findings demonstrate that K-means clustering effectively identifies user segments, providing insights for improving personalized recommendations and marketing strategies. This research underscores the potential of machine learning in optimizing user experiences on music streaming platforms.

Keywords: Spotify, Elbow Method, K-Means Clustering, Music, User Segmentation

1. Pendahuluan

Platform streaming musik seperti Spotify telah menjadi bagian integral dari kehidupan sehari-hari bagi jutaan pengguna di seluruh dunia. Dengan akses tak terbatas ke jutaan lagu dari berbagai genre dan artis, Spotify menawarkan pengalaman mendengarkan musik yang disesuaikan dengan preferensi individual pengguna. Namun, dalam mengelola katalog musik yang luas ini, Spotify dihadapkan pada tantangan untuk menyajikan konten yang relevan dan menarik bagi setiap pengguna. Salah satu cara untuk mengatasi tantangan ini adalah dengan menggunakan teknik segmentasi pengguna. Segmentasi pengguna memungkinkan Spotify untuk memahami preferensi musik dari setiap pengguna dan menyajikan konten yang lebih sesuai dengan selera mereka. Dalam artikel ini, kami akan mengeksplorasi penerapan algoritma K-Means Clustering untuk melakukan segmentasi pengguna Spotify berdasarkan preferensi musik mereka. K-Means Clustering adalah salah satu teknik yang populer dalam analisis data yang memungkinkan kita untuk mengelompokkan data ke dalam kategori yang berbeda berdasarkan kemiripan karakteristik. Penerapan algoritma K-Means Clustering pada data preferensi musik pengguna Spotify diharapkan akan memungkinkan kita untuk mengidentifikasi kelompok-kelompok pengguna dengan preferensi musik yang serupa. Dengan demikian, artikel ini tidak hanya akan menjelaskan konsep dasar dari algoritma K-Means Clustering, tetapi juga akan menunjukkan bagaimana teknik ini dapat diterapkan secara praktis untuk meningkatkan pengalaman pengguna dalam platform streaming musik seperti Spotify.

2. Metode Penelitian

2.1. Pengumpulan Data

Data yang digunakan dalam penelitian kali ini adalah dataset sekunder yang berupa dataset lagu-lagu spotify beserta atributnya yang trending pada tahun 2023 dari website Kaggle.

2.2. Spotify

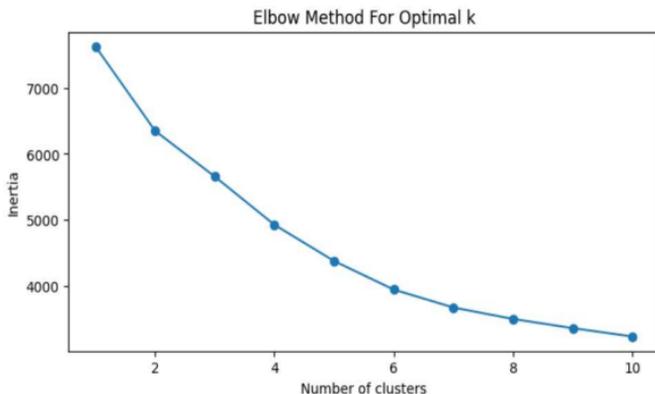
Spotify adalah layanan streaming musik yang populer di seluruh dunia, menawarkan akses ke jutaan lagu dan podcast dari berbagai genre dan artis. Diluncurkan pada tahun 2008, Spotify telah menjadi platform utama bagi pecinta musik yang ingin menikmati beragam pilihan musik kapan saja dan di mana saja. Pengguna dapat mendengarkan musik secara gratis dengan iklan atau berlangganan layanan premium untuk menikmati pengalaman bebas iklan, kualitas audio lebih baik, dan fitur tambahan lainnya. Spotify dikenal karena kemampuannya untuk menyediakan pengalaman mendengarkan yang personal. Dengan menggunakan algoritma canggih, Spotify menganalisis kebiasaan mendengarkan pengguna dan memberikan rekomendasi musik yang disesuaikan dengan preferensi masing-masing individu. Fitur seperti "Discover Weekly" dan "Release Radar" membantu pengguna menemukan musik baru yang sesuai dengan selera mereka. Selain itu, Spotify memungkinkan pengguna untuk membuat, berbagi, dan mengikuti playlist, serta melihat apa yang sedang didengarkan teman-teman mereka. Ini menambahkan elemen sosial yang memperkaya pengalaman mendengarkan musik. Setiap akhir tahun, fitur "Spotify Wrapped" memberikan ringkasan visual tentang kebiasaan mendengarkan pengguna selama tahun tersebut, menyoroti artis, lagu, dan genre yang paling sering diputar. Dengan berbagai fitur inovatif dan beragam pilihan musik, Spotify terus menjadi platform yang menarik bagi berbagai segmen pengguna dengan preferensi musik yang berbeda-beda.

2.3. Preprocessing Data

Tahap preprocessing dilakukan agar data menjadi lebih siap untuk diolah dan menjadi lebih mudah untuk digunakan. Dalam penelitian ini, tahap preprocessing mencakup penyesuaian tipe data, pemilihan fitur, lalu normalisasi. Pada tahap ini tidak dilakukan data cleaning, karena berdasarkan fitur yang dipilih tidak ada data kosong ataupun data yang tidak terdapat nilai. Fitur dari setiap lagu yang dipilih untuk penelitian ini adalah 'bpm', 'danceability_%', 'valence_%', 'energy_%', 'acousticness_%', 'instrumentalness_%', 'liveness_%', dan 'speechiness_%'. Sedangkan untuk fitur popularitasnya terdapat 'songs', 'streams', 'in_spotify_playlists', dan 'in_spotify_charts'. Lalu normalisasi perlu dilakukan terlebih dahulu untuk menyamakan rentang nilai atau domain setiap atribut sehingga mencapai rentang nilai 0, 1 sebelum melanjutkan ke tahap klusterisasi [1].

2.4. Elbow Method

Elbow Method merupakan metode yang digunakan dalam analisis kluster, khususnya K-Means Clustering, untuk menentukan jumlah kluster yang optimal. Ilustrasi nilai K pada metode Elbow dengan K-Means adalah grafik hubungan antara kluster dengan penurunan kesalahan, di mana nilai K yang meningkat akan membuat grafik tersebut menurun secara perlahan hingga mencapai nilai K yang stabil [2]. Metode ini melibatkan penerapan algoritma K-Means dengan berbagai jumlah kluster (K) yang berbeda dan menghitung Within-Cluster Sum of Squares (WCSS) untuk setiap K. WCSS adalah jumlah kuadrat jarak dari setiap titik data ke pusat kluster terdekatnya, yang mengukur seberapa baik kluster tersebut meminimalkan variabilitas dalam kluster. Hasil WCSS kemudian dipetakan dalam grafik dengan sumbu y sebagai WCSS dan sumbu x sebagai jumlah kluster K [3]. Bentuk grafik ini biasanya menunjukkan penurunan tajam di awal, diikuti dengan penurunan yang semakin landai, membentuk bentuk mirip siku atau "elbow." Titik di mana penurunan WCSS mulai melambat secara signifikan menandakan jumlah kluster yang optimal. Dengan memilih K pada titik "elbow" ini, kita dapat menentukan jumlah kluster yang memberikan keseimbangan terbaik antara kompleksitas model dan kemampuan untuk menjelaskan variabilitas data. Berikut ini merupakan grafik yang didapat dari elbow method.



Gambar 1. Grafik Elbow Method

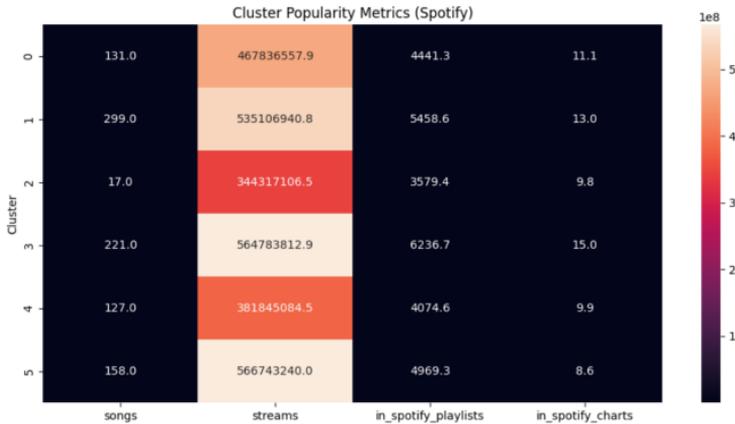
2.5. Algoritma K-Means Clustering

Klasterisasi merujuk pada proses mengelompokkan entitas seperti rekaman, pengamatan, atau objek-objek, berdasarkan kesamaan mereka untuk membentuk kelas atau kelompok [4]. Salah satu contoh algoritma klasterisasi adalah K-Means Clustering. Algoritma K-Means Clustering beroperasi dengan cara memisahkan data ke dalam kelompok-kelompok yang berbeda, tergantung pada kesamaan atribut yang dimiliki [5]. Tujuannya adalah untuk meminimalkan jumlah variabilitas dalam setiap kelompok, dengan menempatkan titik pusat (centroid) yang mewakili setiap kelompok sedekat mungkin dengan anggota-anggotanya. Cara kerja algoritma ini dimulai dengan inisialisasi, di mana titik-titik awal dipilih secara acak sebagai pusat kelompok. Setiap titik data dalam dataset kemudian diberikan label yang sesuai dengan kelompok terdekat (berdasarkan jarak Euclidean, misalnya) dari titik pusat yang diinisialisasi sebelumnya. Selanjutnya, pusat baru untuk setiap kelompok dihitung sebagai rata-rata dari semua titik data yang termasuk dalam kelompok tersebut. Langkah ini diulangi sampai tidak ada perubahan dalam label kelompok dari satu iterasi ke iterasi berikutnya, atau hingga mencapai jumlah iterasi maksimum yang ditentukan sebelumnya. Algoritma K-Means Clustering cenderung cepat dan efisien dalam menangani dataset besar, meskipun sensitif terhadap inisialisasi titik-titik pusat awalnya. Oleh karena itu, seringkali disarankan untuk menjalankan algoritma beberapa kali dengan inisialisasi yang berbeda dan memilih hasil terbaik berdasarkan evaluasi. Evaluasi hasil Klasterisasi dapat dilakukan menggunakan metrik seperti inersia (jumlah total kuadrat jarak antara setiap titik data dan pusat kelompoknya), serta metrik eksternal seperti indeks Inertia atau indeks Silhouette untuk mengevaluasi kualitas klasterisasi. Dengan memahami langkah-langkah dasar dan cara kerja Algoritma K-Means Clustering, peneliti dapat menggunakannya secara efektif untuk menganalisis dan mengelompokkan data, termasuk dalam kasus segmentasi pengguna Spotify berdasarkan preferensi musik.

3. Hasil dan Pembahasan

Penelitian ini akan menguji beberapa fitur dari tiap lagu dan membandingkannya dengan tingkat popularitasnya, lalu model akan menciptakan klasterisasi terhadap lagu-lagu tersebut dengan algoritma K-Means Clustering. Setelah tahap preprocessing dilakukan, dan hasil dari Elbow Method didapatkan, peneliti mengambil keputusan akan menggunakan klaster atau 'K' sebanyak 6, lalu menjalankan klasterisasi berdasarkan fitur-fitur yang disimpan dalam variabel "features".

Berikut ini adalah gambaran heatmap dari hasil klusterisasi yang dibandingkan dengan popularity metrics.



Gambar 2. Heatmap Klusterisasi

Lalu berikut ini adalah data rata rata dari setiap kluster yang dihasilkan, beserta jumlah lagu yang ada pada tiap kluster.

Tabel 1. Hasil Rata-rata Klusterisasi

Cluster	in_spotify_playlists	in_spotify_charts	streams	bpm	Danceability_%	valence_%
0	4441.297710	11.106870	4.678366e+08	116.969466	67.656489	55.946565
1	5458.645485	13.030100	5.351069e+08	116.066890	77.200669	69.976589
2	3579.352941	9.823529	3.443171e+08	122.764706	60.352941	32.235294
3	6236.696833	14.968326	5.647838e+08	134.972851	58.443439	35.696833
4	4074.582677	9.889764	3.818451e+08	127.314961	74.181102	51.559055
5	4969.259494	8.626582	5.667432e+08	118.158228	53.879747	36.563291

energy_%	acousticness_%	instrumentalness_%	liveness_%	speechiness_%	songs
73.305344	17.000000	0.068702	43.473282	8.068702	131.0
71.337793	19.882943	0.314381	12.829431	6.989967	299.0
58.647059	31.000000	57.411765	14.117647	5.411765	17.0
66.778281	13.425339	0.963801	13.990950	6.330317	221.0
62.559055	26.259843	0.078740	15.118110	31.346457	127.0
41.930380	68.259494	1.297468	16.291139	6.556962	158.0

Dari tabel 1 dapat dilihat bahwa tiap kluster mempunyai variasi nilai yang beragam dan signifikan dalam preferensi musik di setiap kluster. Kluster 0 memiliki danceability dan energy yang tinggi dengan acousticness rendah, menunjukkan preferensi untuk musik yang energik dan mudah

untuk menari. Kluster 1, yang merupakan kluster terbesar, menyukai musik upbeat dan positif dengan nilai danceability dan valence yang tinggi. Kluster 2, yang paling sedikit jumlah lagunya, cenderung menyukai musik instrumental dengan instrumentality tinggi dan valence rendah. Kluster 3 memiliki popularitas tertinggi dalam metrik Spotify dan menunjukkan preferensi untuk musik dengan tempo cepat namun tidak selalu upbeat. Kluster 4 menyukai musik akustik yang tetap bisa digunakan untuk menari, dengan nilai acousticness dan danceability yang tinggi. Terakhir, kluster 5 menunjukkan preferensi untuk musik yang lebih tenang dan akustik, dengan acousticness tertinggi dan energy terendah. Data ini membantu memahami preferensi musik dari setiap kluster dan dapat digunakan untuk strategi pemasaran dan pengembangan konten yang lebih tepat sasaran.

4. Kesimpulan

Berdasarkan penelitian ini, dapat diperoleh kesimpulan bahwa segmentasi pengguna spotify berdasarkan preferensi musiknya dengan metode K-Means Clustering telah berhasil dilakukan. Dengan memahami preferensi musik dari setiap kluster, data tersebut dapat digunakan baik untuk strategi pemasaran maupun meningkatkan pengalaman pengguna yang lebih relevan, menarik dan tepat sasaran

Daftar Pustaka

- [1] I. G. A. M. Pratama, L. G. Astuti, I. M. Widiartha, I. G. N. C. P. Anom, C. R. A. Pramatha, and I. D. M. B. A. Darmawan, "Diagnosis Penyakit Ginjal Kronis dengan Algoritma C4.5, K-Means dan BPSO," 2022.
- [2] M. A. Syakur, B. K. Khotimah, E. M. S. Rochman, and B. D. Satoto, "Integration K-Means Clustering Method and Elbow Method for Identification of the Best Customer Profile Cluster," in IOP Conference Series: Materials Science and Engineering, Institute of Physics Publishing, Apr. 2018. doi: 10.1088/1757-899X/336/1/012017.
- [3] D. M. Saputra, D. Saputra, and L. D. Oswari, "Advances in Intelligent Systems Research," 2020.
- [4] Z. Nabila, A. Rahman Isnain, and Z. Abidin, "Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means," Jurnal Teknologi dan Sistem Informasi (JTSI), vol. 2, no. 2, p. 100, 2021, [Online]. Available: <http://iim.teknokrat.ac.id/index.php/JTSI>
- [5] C. Hafidz Ardana et al., "Segmentasi Pelanggan Penjualan Online Menggunakan Metode K-means Clustering," 2024.

Halaman ini sengaja dibiarkan kosong

Pengujian Prototipe Sistem Jasa Pengembangan Aplikasi Menggunakan Metode System Usability Scale

I Kadek Agus Wijaya Kusuma^{a1}, Gst. Ayu Vida Matrika Giri^{a2}

^aProgram Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹aguswkusuma2424@gmail.com
²vida@unud.ac.id (Corresponding Author)

Abstract

The utilization of websites to spread information related to services is commonly used by service providers, including those in application development services. In developing a website, designing an interactive interface becomes an important aspect. This has an impact on the effectiveness and efficiency of the website, as well as the user experience. Therefore, the participation of users in designing the website, considering their needs and issues they face is essential. To ensure the website meets user's standards and expectations, usability testing is conducted on the website prototype. This research uses the System Usability Scale (SUS) method. The SUS scores from the testing serve as a benchmark to evaluate the effectiveness, efficiency, and optimization of the user experience of the website prototype. To achieve high SUS scores, the website prototype must be capable of providing effective and efficient solutions to user's issues while also supplying their needs. A website prototype with high scores can be considered to have met user's standards and expectations, therefore it is ready for implementation.

Keywords: *Prototype, User Interface, User Experience, Usability Testing, System Usability Scale*

1. Pendahuluan

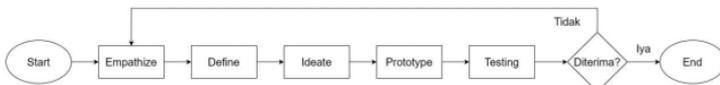
Di era digital ini, manusia hidup berdampingan dengan kemajuan teknologi. Banyak aktivitas manusia yang dipermudah semenjak hadirnya teknologi. Aktivitas yang sebelumnya dilakukan dengan manual pelan-pelan akan dilakukan dengan bantuan teknologi, sering kali menggunakan komputer dan internet. Hal ini memicu perubahan dari bentuk fisik ke bentuk digital. Transisi dari penggunaan atau proses manual ke dalam bentuk digital sering disebut dengan istilah digitalisasi. Proses digitalisasi sangat mudah dirasakan karena banyak aspek yang dipermudah sebab keberlangsungan digitalisasi. Salah satu aspek yang diuntungkan setelah terjadinya digitalisasi yaitu aspek perdagangan. Pelaku UMKM dapat menggaet lebih banyak pembeli melalui media sosial dan situs web [1]. Aplikasi menjadi salah satu komponen yang penting sebagai penunjang digitalisasi. Kini aplikasi digunakan dalam berbagai sektor dan kebutuhan, mulai dari untuk berkomunikasi, hiburan, bertransaksi, pendidikan, administrasi, penyebaran informasi, dan masih banyak lagi. Penggunaan aplikasi yang meluas ini tentunya meningkatkan permintaan aplikasi di pasaran. Hal tersebut memicu kemunculan hingga meledaknya eksistensi jasa pengembangan aplikasi. Jasa ini diperuntukan kepada individu atau instansi yang ingin mendapatkan aplikasi dengan mudah, cepat, memperhatikan kualitas, dan tidak perlu mempekerjakan tim pengembang yang akan membutuhkan sumber daya lebih banyak. Jasa pengembangan aplikasi tentunya perlu memasarkan jasa mereka agar dilirik oleh calon pembeli. Berbagai cara dapat dilakukan untuk memasarkan jasa, salah satunya melalui situs web. Dalam merancang situs web yang optimal untuk memasarkan jasa, kita perlu memerhatikan antarmuka pengguna dan pengalaman pengguna. Antarmuka pengguna menjadi jembatan atau penghubung antara sistem dengan pengguna sistem yang terdiri dari komponen seperti warna, bentuk, dan tulisan [2]. Komponen tersebut harus dirancang dengan menarik, konsisten, dan dapat bekerja serta berinteraksi dengan pengguna untuk menciptakan

pengalaman pengguna yang optimal. Selain itu, antarmuka yang selaras dan interaktif memengaruhi persentase kemungkinan pengguna untuk mengunjungi sistem kembali sehingga berdampak terhadap traffic sistem. Pengalaman pengguna mendeskripsikan perasaan pengguna ketika sebelum, sedang, dan setelah menggunakan sistem [3]. Antarmuka dan pengalaman pengguna tidak dapat dipisahkan karena saling melengkapi satu sama lain [4]. Antarmuka pengguna dan pengalaman pengguna yang optimal dapat diraih dengan memerhatikan pemilihan dan penggunaan komponen tersebut. Riset perlu dilakukan sebelum merancang sistem kepada calon pemilik sistem dan calon pengguna sistem untuk menjaga merek pemilik sistem dan pengguna merasa nyaman selama menggunakan sistem. Kepuasan pengguna dalam menggunakan sistem ataupun prototipe sistem dapat diketahui melalui pengujian fungsionalitas atau usability testing pada sistem ataupun prototipe sistem. Penelitian ini menggunakan prototipe sistem berupa situs web jasa pengembangan aplikasi. Situs berisi informasi mengenai jasa yang ditawarkan dan pemilik usaha dengan tujuan pemasaran digital. Prototipe situs web akan diuji fungsionalitasnya untuk mengetahui efektivitas dan efisiensi dari prototipe serta kepuasan pengguna dari pengalaman menggunakan prototipe sebelum diimplementasikan. Uji Pengujian dilakukan dengan metode System Usability Scale (SUS) dengan harapan mampu menjadi evaluasi dalam meningkatkan kegunaan sistem dan pengalaman pengguna sehingga sistem dapat berjalan dengan optimal ketika diluncurkan. Selain itu melalui penelitian ini diharapkan dapat menjadi inspirasi penelitian lain dalam menggunakan metode yang sama untuk pengujian sistemnya. Metode System Usability Scale (SUS) telah diterapkan dalam beberapa penelitian sebelumnya. Penelitian [2] menerapkan metode SUS pada aplikasi Android Course Online untuk meningkatkan pengalaman pengguna. Hasil penelitian tersebut berupa skor SUS sebesar 78,3 dengan level marginal high berdasarkan sisi acceptability ranges, posisi OK pada sisi adjektif, dan grade B pada sisi grade scale. Skor tersebut didapat dengan melibatkan 30 orang mahasiswa sebagai respon. Kemudian pada penelitian [5] menerapkan metode SUS tdengan melibatkan 32 responden berupa tenaga pendidik SMA Swasta Wiyata Dharma. Penelitian ini menguji platform Google Classroom dan memperoleh skor SUS sebesar 68,91 dengan rating GOOD dan letter grade B. Terdapat juga penelitian [6] yang menggunakan skor SUS sebagai metode evaluasi website Dinas Pendidikan Provinsi Riau. Penelitian ini melibatkan responden sebanyak 96 orang dan memperoleh skor SUS sebesar 51,87 dengan menempati posisi OK untuk kategori adjektif, kemudian huruf F untuk kategori grade scale, dan marginal low untuk kategori acceptability ranges. Website yang diuji dapat diterima tetapi tergolong rendah sehingga perlu dilakukan perbaikan atau evaluasi pengembangan kedepannya.

2. Metode Penelitian

2.1 Design Thinking

Design thinking merupakan metode yang dapat digunakan dalam perancangan sistem. Metode ini dilakukan dengan memahami kebutuhan calon pengguna dan visualisasi rancangan sistem melalui prototipe sehingga terciptanya sistem yang memfasilitasi dan mempercepat implementasi solusi. Design thinking terdiri atas beberapa proses di antaranya yaitu empathize atau proses memahami permasalahan dan kebutuhan pengguna, define atau proses mendefinisikan permasalahan dan kebutuhan pengguna, ideate atau proses merancang ide kreatif untuk menyelesaikan permasalahan pengguna, prototype atau proses membangun rancangan atau memvisualisasikan sistem, dan testing atau proses menguji prototipe yang telah dibangun [7]. Adapun tahapan penelitian menggunakan metode design thinking yang dilakukan divisualisasikan pada Gambar 1.



Gambar 1. Tahapan Penelitian [7]

Gambar 1 menjelaskan tahapan pengembangan sistem dari awal hingga akhir. Namun pada penelitian ini fokus terhadap proses testing prototipe situs web jasa pengembangan aplikasi. Proses testing menggunakan prototipe berguna untuk menilai kelayakan sistem sebelum diluncurkan ataupun sebagai bahan evaluasi untuk proses pengembangan sistem kedepannya.

2.2 Usability Testing

Usability dapat diartikan sebagai proses mengoptimalkan interaksi antara pengguna dan sistem sehingga terbentuk hubungan yang interaktif. Usability menjadi aspek yang penting dalam membangun sistem karena memengaruhi pengalaman pengguna dan impresi pertama pengguna terhadap sistem nantinya. Melalui usability testing, sistem diharapkan dapat memastikan sistem memenuhi kebutuhan pengguna, serta pengguna dapat menggunakan sistem dengan lancar, nyaman, dan puas. Usability testing juga mampu mengurangi biaya perbaikan sistem setelah diluncurkan.

2.3 System Usability Scale

System Usability Scale atau SUS menjadi salah satu metode yang dapat digunakan untuk mengukur usability sistem. SUS berupa kuesioner yang digunakan sebagai alat ukur penilaian usability suatu produk. Kuesioner SUS terdiri dari sepuluh pertanyaan yang berkaitan dengan sistem dengan lima poin skala dari sangat setuju hingga sangat tidak setuju. Poin skala sangat setuju (SS) berbobot 5, poin skala setuju (S) berbobot 4, poin skala ragu-ragu (RR) berbobot 3, poin skala tidak setuju (TS) bernilai 2, poin skala sangat tidak setuju (STS) berbobot 1 [8]. Adapun pertanyaan kuesioner SUS dijelaskan pada Tabel 1.

Tabel 1. Pertanyaan Kuesioner SUS [8]

No	Pertanyaan	Poin Skala				
		SS	S	RR	TS	STS
1	Saya akan menggunakan sistem ini	5	4	3	2	1
2	Saya merasa desain sistem ini rumit untuk digunakan	5	4	3	2	1
3	Saya merasa desain sistem ini mudah digunakan	5	4	3	2	1
4	Saya membutuhkan bantuan dari orang lain dalam menggunakan sistem ini	5	4	3	2	1
5	Saya merasa fitur-fitur sistem ini berjalan dengan semestinya	5	4	3	2	1
6	Saya merasa ada banyak hal yang tidak konsisten pada sistem ini	5	4	3	2	1
7	Saya merasa orang lain akan memahami cara menggunakan sistem ini dengan cepat	5	4	3	2	1
8	Saya merasa sistem ini membingungkan ketika digunakan	5	4	3	2	1
9	Saya merasa percaya diri dalam menggunakan sistem ini	5	4	3	2	1
10	Saya perlu membiasakan diri terlebih dahulu sebelum menggunakan sistem ini	5	4	3	2	1

Dari jawaban responden akan dilakukan perhitungan untuk menentukan skor SUS. Adapun ketentuan dalam perhitungan skor SUS yaitu pada setiap pertanyaan bernomor ganjil (1,3,5,7,9) yang berupa pernyataan positif berbobot poin atau skor dikurangi 1. Kemudian pada setiap pertanyaan bernomor genap (2,4,6,8,10) yang berupa pernyataan negatif berbobot 5 dikurangi poin atau skor. Skor SUS didapat dengan menjumlahkan nilai dari proses tersebut kemudian hasilnya dikalikan dengan 2,5. Skor akhir SUS merupakan nilai rata-rata dari jumlah skor SUS keseluruhan responden. Skor SUS memiliki rentang nilai dari 0 hingga 100 [6].

Formula dari perhitungan skor SUS dapat dilihat pada rumus 1 untuk mencari skor SUS dari tiap responden dan rumus 2 untuk mencari skor SUS akhir.

$$\text{Skor SUS} = ((P1 - 1) + (5 - P2) + (P3 - 1) + (5 - P4) + (P5 - 1) + (5 - P6) + (P7 - 1) + (5 - P8) + (P9 - 1) + (5 - P10)) \times 2,5 \quad (1) [6]$$

$$\text{Skor SUS akhir} = \frac{\text{Jumlah skor SUS}}{\text{Jumlah responden}} \quad (2) [6]$$

Sebagai interpretasi skor SUS, dilakukan pembobotan melalui peringkat sebanyak 5 untuk kategori adjektif dan grade scale atau nilai huruf, dimulai dari excellent jika skor SUS lebih besar dari 80,3, good jika skor SUS di antara 69 dan 80,3, ok jika skor SUS adalah 68, poor jika skor SUS di antara 51 dan 67, dan awful jika skor SUS di bawah 51 [5]. Interpretasi skor SUS berdasarkan kategori adjektif dan nilai huruf dijelaskan pada Tabel 2.

Tabel 2. Interpretasi Skor SUS Kategori Adjektif dan Nilai Huruf [5]

Skor SUS	Nilai Huruf	Adjektif
>80,3	A	Excellent
69-80,3	B	Good
68	C	OK
51-67	D	Poor
<51	F	Awful

Di sisi lain terdapat pembobotan berupa tingkat penerimaan sistem yang diuji atau acceptability ranges atau rentang penerimaan yang terdiri dari 4 rentang di antaranya acceptable jika skor SUS di antara 70 hingga 100, marginal high jika skor SUS berada di antara 63 dan 70, marginal low jika skor SUS berada di antara 51 hingga 62, dan not acceptable jika skor SUS di antara 0 dan 50 [6]. Istilah tersebut juga dapat disimbolkan dengan huruf A hingga D dan F. Interpretasi skor SUS berdasarkan kategori rentang penerimaan dijelaskan pada Tabel 3.

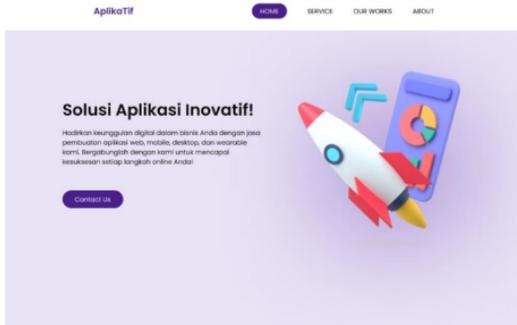
Tabel 3. Interpretasi Skor SUS Kategori Rentang Penerimaan [6]

Skor SUS	Rentang Penerimaan
70-100	Acceptable
63-69	Marginal High
51-62	Marginal Low
0-50	Not Acceptable

3. Hasil dan Diskusi

3.1 Antarmuka Prototipe

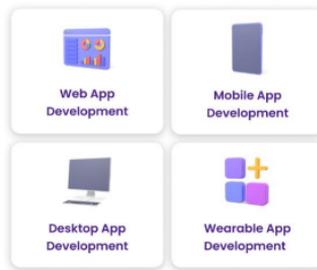
Antarmuka prototipe dibuat berdasarkan hasil proses design thinking yang telah dilalui. Dari proses empathize menghasilkan informasi berupa permasalahan dan kebutuhan calon pengunjung sistem yaitu informasi pada sistem yang sulit dicerna, terlalu bertele-tele, dan tampilan pada sistem yang terlalu ramai objek. Dari proses define menghasilkan definisi dari permasalahan dan kebutuhan calon pengunjung sistem berupa pain point yaitu penyampaian informasi pada sistem yang tidak efektif dan tampilan pada sistem yang memecah fokus pengunjung. Dari proses ideate menghasilkan solusi permasalahan yang kreatif dan inovatif yaitu penyampaian informasi pada sistem dibuat dengan ringkas dan interaktif serta tampilan pada sistem dirancang dengan primary color yang pastel dan kontras warna antar objek tetap terjaga sehingga setiap teks maupun objek tampak dengan jelas.



Gambar 2. Halaman Utama Prototipe

Gambar 2 merupakan halaman utama dari prototipe jasa pengembangan aplikasi. Pengunjung akan diarahkan ke halaman utama ketika baru mengakses situs web. Halaman utama menjadi impresi pertama pengunjung situs web sehingga perlu diperhatikan dalam penggunaan komponen berupa warna, teks, objek, penempatan komponen, dan juga informasi yang disampaikan. Pada halaman utama menampilkan informasi berupa nama situs web, tagline atau slogan untuk memikat pengunjung situs web, dan informasi singkat mengenai jasa yang ditawarkan. Terdapat juga bar navigasi untuk mempermudah dan mempercepat pengunjung meraih informasi yang spesifik serta tombol kontak sebagai pintu komunikasi antara pengunjung dengan admin. Tombol kontak akan mengarahkan pengunjung ke platform komunikasi yang ramai digunakan oleh target market atau sasaran konsumen jasa sehingga tidak merepotkan pengunjung bahkan mengecilkan peluang mereka untuk mencoba jasa yang ditawarkan.

Our Services

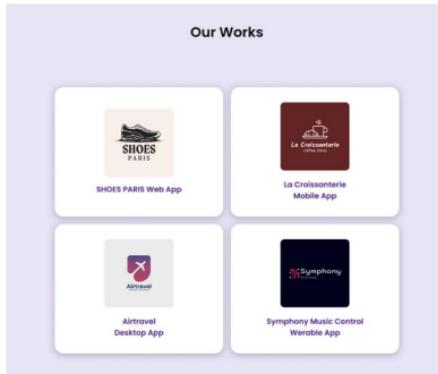


Gambar 3. Halaman Layanan Singkat Prototipe



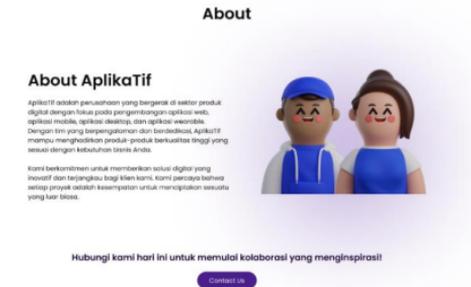
Gambar 4. Halaman Layanan Detail Prototipe

Gambar 3 dan 4 menampilkan layanan yang disediakan oleh jasa. Informasi ditampilkan dimulai dari informasi singkat jasa seperti pada Gambar 3 dan informasi detail jasa seperti pada Gambar 4. Penyampaian informasi jasa dibuat interaktif dengan menjadikan setiap card pada informasi singkat jasa menjadi tombol yang akan mengarahkan pengunjung ke informasi detail jasa. Melalui interaksi yang ditawarkan harapannya dapat memberikan pengalaman pengguna yang unik dan berkesan. Hal ini mampu memicu pengunjung untuk mengingat situs web jasa karena menawarkan hal yang dirasa menarik.



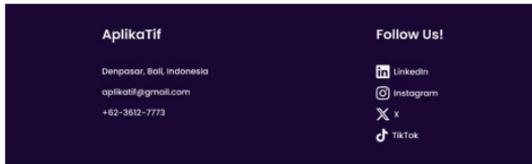
Gambar 5. Halaman Proyek Prototipe

Gambar 5 menampilkan proyek yang telah dikerjakan oleh jasa tersebut. Halaman ini bertugas untuk meningkatkan kepercayaan pengunjung terhadap jasa. Proyek ditampilkan melalui card yang berisi logo dan nama aplikasi sehingga pengunjung dapat dengan mudah mencari aplikasi tersebut.



Gambar 6. Halaman Tentang Pemilik Jasa Prototipe

Gambar 6 menampilkan tentang pemilik jasa pengembangan aplikasi. Halaman ini membantu mendekatkan pengunjung dengan pemilik jasa. Pada halaman ini juga menampilkan kalimat persuasif untuk menggunakan jasa dan didampingi dengan tombol kontak sebagai aksi yang diharapkan kepada pengunjung.



Gambar 7. Footer Prototipe

Gambar 7 merupakan footer dari prototipe situs web jasa pengembangan aplikasi. Informasi berupa lokasi, e-mail, dan nomor telepon kantor pengembangan jasa ditampilkan di sisi kiri sedangkan media sosial jasa pengembangan aplikasi ditampilkan di sisi kanan. Informasi ditampilkan untuk mempermudah pengunjung menggali informasi lebih jauh mengenai jasa dan pemilik jasa. Hal ini menjadi salah satu upaya pendekatan dengan pengunjung dan meningkatkan kepercayaan dan potensi pengunjung untuk menggunakan jasa yang ditawarkan.

3.2 Hasil System Usability Scale

Usability testing pada penelitian ini dilakukan dengan melibatkan 10 responden dengan kualifikasi berupa usia 18-40 tahun yang dianggap sebagai rentang usia produktif dan aktif menggunakan internet. Rentang usia tersebut juga dianggap berpotensi sebagai calon pengunjung sistem. Pada kuesioner usability testing dilampirkan tautan untuk mengakses prototipe dari sistem sehingga mempermudah responden untuk menjangkau dan merasakan pengalamannya menggunakan prototipe. Kuesioner diawali dengan pertanyaan berupa identitas diri seperti nama, umur, dan pekerjaan dan dilanjutkan dengan 10 pertanyaan kuesioner SUS seperti pada tabel 1. Responden menjawab seluruh pertanyaan berdasarkan pengalaman mereka setelah menggunakan prototipe. Setelah mendapatkan jawaban dari 10 responden, dilakukan perhitungan skor SUS tiap respondennya dengan rumus 1. Skor SUS tiap responden kemudian dijumlahkan dan dibagi 10 yang merupakan jumlah responden pada penelitian ini untuk mencari nilai akhir SUS sebagaimana dijelaskan pada rumus 2. Berikut rekapitulasi dari 10 jawaban responden beserta skor SUS tiap responden dan hasil akhir skor SUS yang dipaparkan pada Tabel 4.

Tabel 4. Hasil Pengujian SUS

Responden	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	Skor SUS
R1	4	3	4	2	4	2	3	2	4	4	65
R2	5	4	5	2	5	4	5	3	5	5	67,5
R3	5	1	5	2	4	1	5	1	5	3	90
R4	4	2	5	1	5	1	5	2	3	1	87,5
R5	3	2	4	2	4	2	4	1	4	4	70
R6	5	2	4	3	4	1	4	1	5	2	82,5
R7	5	1	5	1	5	4	3	2	4	2	80
R8	4	3	4	1	4	3	4	2	3	1	72,5
R9	2	1	5	2	5	1	5	2	5	3	82,5
R10	5	2	4	1	5	2	2	1	3	3	75
	Jumlah										772,5
	Rata-Rata										77,25

Tabel 4 menjelaskan bahwa lebih banyak responden merasa akan menggunakan sistem ini, mudah untuk digunakan, berjalan dengan semestinya, orang lain akan mudah memahami atau melakukan navigasi pada sistem, dan percaya diri ketika menggunakan sistem. Responden beranggapan tidak setuju terhadap pernyataan sistem yang rumit untuk digunakan, terdapat banyak hal yang tidak konsisten, kebingungan ketika menggunakan sistem. Selain itu, lebih banyak responden merasa netral dan setuju dalam pernyataan perlu membiasakan diri terlebih dahulu sebelum menggunakan sistem. Dari perhitungan yang dilakukan, prototipe sistem menghasilkan skor SUS sebesar 77,25. Skor tersebut masuk ke dalam kategori good yang disimbolkan dengan huruf B dan masuk dalam kategori acceptable dari sisi acceptability ranges. Kategori tersebut memungkinkan prototipe lanjut ke proses implementasi situs web tanpa mengulangi proses design thinking.

4. Kesimpulan

Dalam membangun sebuah sistem perlu memerhatikan setiap proses dan komponen yang berkaitan dengan proses tersebut. Seperti pada proses usability testing kita perlu memerhatikan antarmuka prototipe untuk mengetahui efektivitas dan efisiensi prototipe sebelum diimplementasikan. Antarmuka prototipe harus dirancang interaktif dengan pengguna sistem karena antarmuka berperan sebagai jembatan antara sistem dan pengguna sistem. Efektifitas dan efisiensi prototipe dinilai dari kemampuan sistem dalam memecahkan permasalahan serta memenuhi kebutuhan pengguna. Hal tersebut menjadi kunci dari keberhasilan rancangan sistem karena mampu mendapatkan pengalaman pengguna yang optimal. Pada penelitian ini, dilakukan pengujian prototipe situs web jasa pengembangan aplikasi dengan metode System Usability Scale untuk mengetahui efektifitas dan efisiensi dari rancangan sistem. Sebelumnya telah dilakukan perancangan sistem dengan metode design thinking yang menjadi alur dari penelitian. Optimalisasi pengalaman pengguna diinterpretasikan melalui skor SUS yang diperoleh. Prototipe situs web jasa pengembangan sistem memperoleh skor SUS 77,25 dengan kategori good yang dapat disimbolkan dengan huruf B. Melalui skor yang diperoleh, disimpulkan bahwa prototipe mampu menyediakan kebutuhan pengguna, menyelesaikan permasalahan pengguna, dan memberikan pengalaman pengguna yang optimal sehingga dapat diimplementasikan ke tahap berikutnya.

Daftar Pustaka

- [1] Sunarsi D, "Digitalisasi UMKM," 2020.
- [2] W. Buana and B. Nurina Sari, "Analisis User Interface Meningkatkan Pengalaman Pengguna Menggunakan Usability Testing pada Aplikasi Android Course," vol. 5, no. 2, pp. 91–97, 2022, [Online]. Available: <http://e-journal.unipma.ac.id/index.php/doubleclick>
- [3] R. Hartson and P. Pyla, "The UX Book Agile UX Design for a Quality User Experience," 2019.
- [4] A. Hussain and E. O. C. Mkpjoju, "The Effect of Responsive Web Design on The User Experience with Laptop and Smartphone Devices," 2015. [Online]. Available: www.jurnalteknologi.utm.my
- [5] S. Hadini Marpaung, "Pengukuran Usability Google Classroom dari Sisi Tenaga Pengajar SMA Wiyata Dharma Medan," Jurnal Pendidikan Tambusai, vol. 6, no. 2, pp. 9658–9668, 2022.
- [6] S. Aisyah et al., "Evaluasi Usability Website Dinas Pendidikan Provinsi Riau Menggunakan Metode System Usability Scale," Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi, vol. 7, no. 2, pp. 125–132, 2021, [Online]. Available: <https://disdik.riau.go.id>
- [7] M. Wahyu Sanjaya and A. Febriandirza, "Penerapan Metode Design Thingking Terhadap Peningkatan Pengalaman Pengguna Pada Sistem Akademik Uhamka," JOISIE Journal of Information System and Informatics Engineering, vol. 7, no. 1, pp. 7–16, 2023.
- [8] A. Sidik, S. Sn, M. Ds, U. Islam, K. Muhammad, and A. Al-Banjari, "Penggunaan System Usability Scale (SUS) Sebagai Evaluasi Website Berita Mobile," 2018.

Halaman ini sengaja dibiarkan kosong

Klasifikasi Tingkat Keparahan Kecelakaan Lalu Lintas Menggunakan Random Forest Classifier

I Gusti Ngurah Bagus Lanang Purbhawa^{a1}, I Gede Arta Wibawa^{a2}

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹purbhawa.2208561108@student.unud.ac.id
²gede.arta@unud.ac.id

Abstract

Traffic accidents are a common problem that often occurs. Many factors cause and determine the severity of traffic accidents. These factors can include road conditions, weather, light conditions, driver age, and the cause of the accident. In this study, researchers will try to apply the Random Forest method to classify the severity of traffic accidents. The Random Forest method was chosen because of its excellent ability to handle high-dimensional data and tolerance for overfitting. The dataset used in this research was taken from Kaggle, consisting of 12316 records and 32 features covering various attributes related to traffic accidents. Before applying random forest, it is necessary to carry out a preprocessing stage on the dataset to remove irrelevant features, fill in empty values and divide the data into training and testing data. The results of this research show that Random Forest can produce a good level of in classifying the severity of traffic accidents with 92% accuracy. This shows the potential of this method as a useful tool in the analysis and prediction of traffic accidents. Therefore, this research makes a significant contribution to efforts to improve road safety.

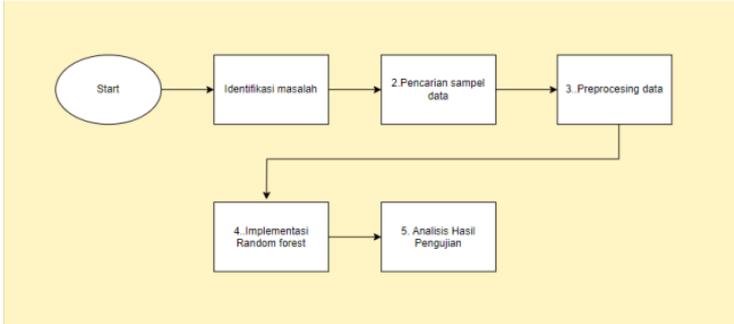
Keywords: *Random Forest Classifier, Traffic accident, Machine learning, Data Classification, Supervised Learning*

1. Pendahuluan

Kecelakaan lalu lintas merupakan masalah serius yang mempengaruhi keselamatan masyarakat dan mobilitas di jalan raya. Setiap tahunnya, ribuan kecelakaan terjadi di jalan raya, menyebabkan kerugian besar baik dalam hal korban jiwa maupun kerugian materi. Menurut laporan Organisasi Kesehatan Dunia jumlah kematian akibat kecelakaan lalu lintas lebih dari 1,25 juta orang dan setiap tahunnya kecelakaan non-fatal menimpa lebih dari 20–50 juta orang [1].hingga dengan memahami faktor-faktor yang mempengaruhi keparahan kecelakaan lalu lintas merupakan hal penting untuk merancang strategi pencegahan yang lebih efektif.Tingkat keparahan dalam kecelakaan lalu lintas dapat dipengaruhi oleh hal seperti, kondisi lingkungan sekitar, misalnya jalan yang licin, hujan, jalanan yang berbatu hingga faktor internal seperti pengalaman mengemudi hingga umur pengemudi. Dalam konteks ini, analisis data dapat menjadi alat yang sangat berguna untuk memahami pola dan tren kecelakaan lalu lintas. Dengan memanfaatkan teknik analisis data dan machine learning, kita dapat mengklasifikasikan faktor-faktor apa yang memiliki pengaruh terhadap tingkat keparahan kecelakaan, sehingga dapat dilakukan tindakan pencegahan yang lebih tepat dan efektif.Algoritma Machine Learning dalam kasus ini dapat menemukan pola tersembunyi untuk memprediksi apakah tingkat keparahan kecelakaan itu fatal, serius, atau ringan [2]. Random forest merupakan algoritma yang sesuai dengan permasalahan klasifikasi ini yang dimana dalam beberapa kasus yang menanganis data yang lebih kompleks dengan baik, termasuk data dengan banyak fitur dan non-linearitas jika dibandingkan dengan algoritma lain seperti SVM, algoritma random forest dapat menghasilkan nilai akurasi, presisi, dan recall yang lebih tinggi [3].

2. Metode Penelitian

2.1 Alur Penelitian



Gambar 1. Alur Penelitian

Adapun alur kerangka dari penelitian klasifikasi jenis kecelakaan berdasarkan keparahannya, yang memiliki penjelasan sebagai berikut:

a. Identifikasi masalah

Merupakan tahapan awal dari penelitian, di mana peneliti mengidentifikasi permasalahan yang akan di carikan solusi, dalam kasus ini peneliti akan mengklasifikasi faktor-faktor dalam pengaruh keparahan dalam kecelakaan lalu lintas

b. Pencarian sampel data

Pencarian data dilakukan secara sekunder yang dimana datanya diambil dari kaggle, yang dimana datanya data tingkat keparahan kecelakaan lalu lintas, data yang diambil memiliki format csv dengan 32 features di dalamnya

c. Preprocessing data

Preprocessing merupakan proses penting sebelum melakukan pemodelan dengan Random Forest, di tahap ini akan menghilangkan nilai yang tidak perlu, sehingga data yang diuji merupakan data yang baik.

d. Implementasi Random Forest

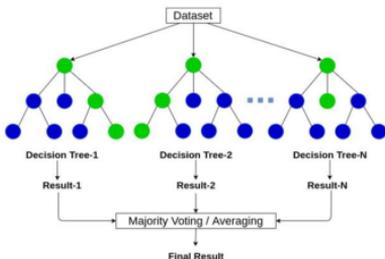
Pengimplementasian random forest akan menggunakan python dan google collabs sebagai tools.

e. Analisis Hasil Pengujian

Setelah dilakukan implementasi kita perlu melakukan evaluasi apakah hasil dari pengujian mendapatkan akurasi yang baik sehingga dapat menjadi solusi untuk penyelesaian kasus ini

2.2 Random Forest Classification

Random Forest



Gambar 2. Random forest klasifikasi

Klasifikasi dapat diartikan sebagai suatu pendekatan yang dikenal supervised algorithm dalam ranah data science. Tentu pada teknik klasifikasi memerlukan data yang berlabel [4]. Tipe data pada classification memerlukan data label, yang dapat berupa kategori biner, multi-kelas, nilai numerik, atau diekstrak dari teks, gambar, atau audio. Tanpa adanya label dalam data, maka algoritma classification seperti random forest tidak dapat belajar untuk mengidentifikasi pola dan membuat prediksi yang akurat. Algoritma classification belajar dari data berlabel untuk mengidentifikasi pola dan membuat prediksi, dan beberapa algoritma umum termasuk SVM, KNN, Decision Trees, dan Random Forests. Random Forest merupakan algoritma ensemble learning yang kuat dan populer, digunakan untuk menyelesaikan berbagai tugas machine learning. Random Forest dipakai untuk masalah regresi dan klasifikasi dengan kumpulan data yang berukuran besar [5]. Algoritma ini bekerja dengan membangun banyak pohon keputusan acak dari subset data yang berbeda, dan hasil akhir dikonsensuskan untuk menghasilkan prediksi yang lebih akurat dan robust. Keunggulan utama Random Forest terletak pada kemampuannya menangani kumpulan data besar dengan efisien tanpa mengalami overfitting

3. Hasil dan Diskusi

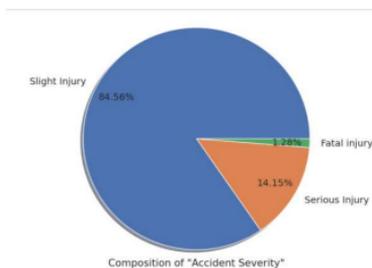
Berdasarkan Alur penelitian dan kajian literatur, berikut adalah hasil yang diperoleh.

3.1. Pencarian Sampel data

Penelitian ini menggunakan data tingkat keparahan kecelakaan lalu lintas dari Kaggle.com yang di-publish oleh Saurabh Shanane. dataset ini terdiri dari 12316 baris/records dan 32 features/columns.

```
Index(['Time', 'Day_of_week', 'Age_band_of_driver', 'Sex_of_driver',  
      'Educational_level', 'Vehicle_driver_relation', 'Driving_experience',  
      'Type_of_vehicle', 'Owner_of_vehicle', 'Service_year_of_vehicle',  
      'Defect_of_vehicle', 'Area_accident_occured', 'Times_of_pedians',  
      'Road_allignment', 'Types_of_junction', 'Road_surface_type',  
      'Road_surface_conditions', 'Light_conditions', 'Weather_conditions',  
      'Type_of_collision', 'Number_of_vehicles_involved',  
      'Number_of_casualties', 'Vehicle_movement', 'Casualty_class',  
      'Sex_of_casualty', 'Age_band_of_casualty', 'Casualty_severity',  
      'Work_of_casualty', 'Fitness_of_casualty', 'Pedestrian_movement',  
      'Cause_of_accident', 'Accident_severity'],  
      dtype='object')
```

Gambar 3. Kolom pada dataset



Gambar 4. Komposisi Accident Severity pada data

Berdasarkan gambar diatas dapat dilihat kalau adanya ketidakseimbangan label tingkat keparahan kecelakaan lalu lintas. pada label fatal injury memiliki persentase yang sangat rendah hanya sekitar 1%. kita perlu menyeimbangkan records pada masing-masing label pada tahap preprocessing dan menghapus fitur-fitur yang tidak perlu.

3.2 Preprocessing Data

Data Preprocessing merupakan tahapan awal dalam data mining, biasanya dilakukan melalui cara eliminasi data yang tidak sesuai.

a. Mengecek nilai null pada kolom

```
▶ null_df = df.isnull().sum().sort_values(ascending=False).to_frame()
null_df.columns= ["No of Null values"]
null_df["% of Null values"] = round(null_df["No of Null values"]/len(df)*100,2)
null_df[null_df["No of Null values"] > 0]
```

Gambar 5. Kode mengecek nilai null

Setelah dilakukan pengecekan null pada setiap kolom kita akan mencari kolom nilai yang memiliki lebih 2500 null

b. Mengecek baris yang duplikat

```
[62] ### Checkin for the duplicate values in the dataset
df.duplicated().sum()

0
```

Gambar 6. Mengecek baris yang duplikat

Pada Dataset tingkat keparahan kecelakaan lalu lintas tidak ditemukannya ada baris yang bernilai sama atau duplikat

c. Mengisi nilai yang kosong

```
#for categorical values we can replace the null values with the Mode of it
for i in categorical:
    df[i].fillna(df[i].mode()[0],inplace=True)
```

Gambar 7. Mengisi kolom kategorikal kosong

potongan kode tersebut akan mengganti nilai-nilai yang hilang (null) dalam fitur-fitur dengan modus (nilai yang paling sering muncul) dari setiap fitur tersebut.

d. Label Encoding

#	Column	Non-Null Count	Dtype
0	Day_of_week	12316 non-null	int64
1	Age_band_of_driver	12316 non-null	int64
2	Sex_of_driver	12316 non-null	int64
3	Educational_level	12316 non-null	int64
4	Vehicle_driver_relation	12316 non-null	int64
5	Driving_experience	12316 non-null	int64
6	Type_of_vehicle	12316 non-null	int64
7	Owner_of_vehicle	12316 non-null	int64
8	Area_accident_occured	12316 non-null	int64
9	Lanes_or_Medians	12316 non-null	int64
10	Road_allignment	12316 non-null	int64
11	Types_of_Junction	12316 non-null	int64
12	Road_surface_type	12316 non-null	int64
13	Road_surface_conditions	12316 non-null	int64
14	Light_conditions	12316 non-null	int64
15	Weather_conditions	12316 non-null	int64
16	Type_of_collision	12316 non-null	int64
17	Vehicle_movement	12316 non-null	int64
18	Casualty_class	12316 non-null	int64
19	Sex_of_casualty	12316 non-null	int64
20	Age_band_of_casualty	12316 non-null	int64
21	Casualty_severity	12316 non-null	int64
22	Pedestrian_movement	12316 non-null	int64
23	Cause_of_accident	12316 non-null	int64

Gambar 8. Label encoding

mengubah nilai-nilai dalam fitur kategorikal menjadi nilai numerik menggunakan metode Label Encoding. Ini dilakukan setelah analisis chi-kuadrat dilakukan pada dataset asli untuk memilih fitur-fitur yang paling relevan.

e. Fitur Selection

	Features	Fscore	Pvalues
14	Light_conditions	16.082824	0.000322
20	Age_band_of_casualty	13.778413	0.001019
16	Type_of_collision	10.096323	0.009421
1	Age_band_of_driver	8.915392	0.011589
12	Road_surface_type	6.994806	0.030276
4	Vehicle_driver_relation	5.345345	0.069067
5	Driving_experience	4.499679	0.105416
8	Area_accident_occured	3.616540	0.163937
9	Lanes_or_Medians	3.291615	0.103824
18	Casualty_class	3.216860	0.200202
23	Cause_of_accident	3.193666	0.202537
11	Types_of_junction	3.086487	0.213687
17	Vehicle_movement	2.200712	0.332753
15	Weather_conditions	1.149345	0.562889
7	Owner_of_vehicle	1.104282	0.575722
6	Type_of_vehicle	1.077671	0.583427

Gambar 9. Table score nilai fitur

Nilai-nilai ini dapat digunakan untuk menilai signifikansi setiap fitur dalam memprediksi variabel target. Semakin kecil nilai p, semakin signifikan hubungan antara fitur dan variabel targetnya.

f. Dummy Variable

```

dummy = pd.get_dummies(df[['Age_band_of_driver', 'vehicle_driver_relation', 'driving_experience',
                           'area_accident_occured', 'lanes_or_medians', 'types_of_junction', 'road_surface_type',
                           'light_conditions', 'weather_conditions', 'type_of_collision', 'vehicle_movement',
                           'casualty_class', 'age_band_of_casualty', 'cause_of_accident']], drop_first=True)

dummy.head()

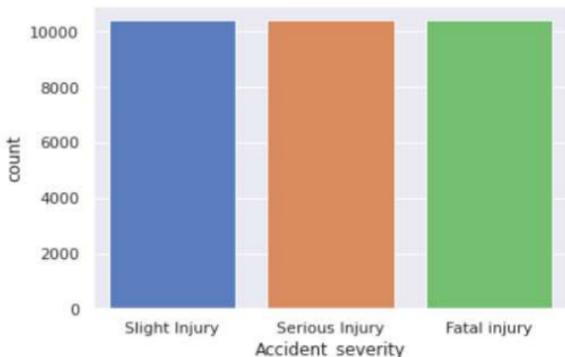
```

	Age_band_of_driver_31-39	Age_band_of_driver_40-49	Age_band_of_driver_50-59	Age_band_of_driver_unknown	vehicle_driver_relation_other	vehicle_driver_relation_ownership	vehicle_driver_relation_ride
0	False	False	False	False	False	False	False
1	True	False	False	False	False	False	False
2	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False

Gambar 10. Output Dummy Variable

Setiap fitur kategorikal akan diubah menjadi kolom2 baru berdasarkan jumlah data unik pada kolom nya. Dengan menggunakan variabel dummy, Anda dapat lebih memahami bagaimana fitur kategorikal mempengaruhi variabel target.

g. Oversampling



Gambar 11. Diagram data setelah oversampling

Oversampling adalah teknik preprocessing data yang digunakan untuk meningkatkan jumlah contoh dari kelas minoritas dalam dataset yang tidak seimbang. Dataset tidak seimbang terjadi ketika terdapat perbedaan signifikan dalam jumlah contoh antara kelas-kelas di dalamnya. Hal ini dapat menyebabkan model machine learning menjadi bias terhadap kelas mayoritas dan menghasilkan performa yang buruk untuk kelas minoritas. dalam dataset ini kelas dengan label fatal injury memiliki jumlah yang sangat kecil sehingga perlu dilakukan oversampling

h. Split Train and test data

```
#converting data to training data and testing data
from sklearn.model_selection import train_test_split
#splitting 70% of the data to training data and 30% of data to testing data
x_train,x_test,y_train,y_test=train_test_split(xo,yo,test_size=0.30,random_state=42)

print(x_train.shape,x_test.shape,y_train.shape,y_test.shape)

(21871, 104) (9374, 104) (21871,) (9374,)
```

Gambar 12. Split data

Langkah terakhir dalam tahap preprocessing adalah membagi dataset kita menjadi data latih dan data uji, disini data akan dibagi menjadi data latih dan data uji, dengan perbandingan 70% banding 3%.

3.3 Random Forest Classifier

Pada tahap ini, algoritma Random Forest akan diimplementasikan untuk mengklasifikasikan data keparahan kecelakaan lalu lintas. Data ini terdiri dari 3 label, yaitu fatal injury, Serious Injury, dan Slight Injury. Random Forest bekerja dengan membangun banyak pohon keputusan acak dari subset data yang berbeda. RandomizedSearchCV disini digunakan untuk melakukan tuning parameter model Random Forest, dengan tujuan menemukan kombinasi parameter terbaik yang menghasilkan kinerja optimal. Data yang digunakan dibagi menjadi data pelatihan (training data) dan data pengujian (test data).

```

Best parameters found: {'criterion': 'gini', 'max_depth': None, 'min_samples_split': 3, 'n_estimators': 189}
Best cross-validation accuracy: 0.9176537618394409
    
```

	precision	recall	f1-score	support
Fatal injury	0.99	0.99	0.99	3126
Serious Injury	0.89	0.89	0.89	3144
Slight Injury	0.89	0.89	0.89	3104
accuracy			0.92	9374
macro avg	0.92	0.92	0.92	9374
weighted avg	0.92	0.92	0.92	9374

Gambar 13. Akurasi random forest

Berdasarkan hasil tuning parameter, berikut adalah kombinasi parameter terbaik yang ditemukan adalah criterion: 'entropy', max_depth: None, min_samples_split: 5, n_estimators: 195.



Gambar 14. Confusion matrix display

Hasil akhir dari penelitian ini, Random Forest menunjukkan akurasi yang tinggi dalam mengklasifikasikan tingkat keparahan kecelakaan lalu lintas, yaitu 92%. Hal ini menunjukkan bahwa model mampu memprediksi kategori kecelakaan dengan tepat Implementasi Random Forest diharapkan dapat menghasilkan akurasi klasifikasi yang tinggi untuk membantu memahami pola dan faktor-faktor yang berkontribusi pada tingkat keparahan kecelakaan lalu lintas.

4. Kesimpulan

Hasil akhir dari penelitian ini dapat mengklasifikasi data keparahan kecelakaan lalu lintas dengan baik, hal itu dibuktikan dari akurasi yang didapat pada pengimplementasian algoritma random forest yang dimana menyentuh angka 92%. Algoritma ini menghasilkan akurasi tinggi dan robust terhadap data dengan membangun banyak pohon keputusan acak dan voting mayoritas sebagai prediksi akhir. model ini dapat digunakan untuk mendukung pengambilan keputusan dalam upaya

peningkatan keselamatan lalu lintas. Hasil evaluasi menunjukkan bahwa model ini sangat efektif dalam mengidentifikasi kecelakaan dengan tingkat keparahan fatal, serta cukup andal dalam mengklasifikasikan kecelakaan dengan tingkat keparahan serius dan ringan. dengan adanya penelitian ini diharapkan dapat membantu banyak orang baik para ahli hingga pihak terkait dalam mengambil keputusan pencegahan kecelakaan lalu lintas yang tepat.

Daftar Pustaka

- [1] Tadesse Kebede Bahiru, Dheeraj Kumar Singh, and Engdaw Ayalew Tessfaw, "Comparative Study on Data Mining Classification Algorithms for Predicting Road Traffic Accident Severity," Apr. 2018, doi: <https://doi.org/10.1109/iccict.2018.8473265>.
- [2] S. Malik, Hesham El Sayed, Manzoor Ahmed Khan, and Muhammad Jalal Khan, "Road Accident Severity Prediction — A Comparative Analysis of Machine Learning Algorithms," Dec. 2021, doi: <https://doi.org/10.1109/gcaiot53516.2021.9693055>.
- [3] Ayu Aina Nurkhaliza and Arie Wahyu Wijayanto, "Perbandingan Algoritma Klasifikasi Support Vector Machine dan Random Forest pada Prediksi Status Indeks Mitigasi dan Kesiapsiagaan Bencana (IMKB) Satuan Kerja BPS di Indonesia Tahun 2020," *Jurnal Informatika Universitas Pamulang*, vol. 7, no. 1, pp. 54–59, 2022, doi: <https://doi.org/10.32493/informatika.v7i1.16117>.
- [4] "Teknik pre-processing dan classification dalam data science," *Master of Industrial Engineering*, 2019. <https://mie.binus.ac.id/2022/08/26/teknik-pre-processing-dan-classification-dalam-data-science> (accessed May 10, 2024).
- [5] Ilham Adriansyah, Muhammad Diemas Mahendra, Errissya Rasywir, and Yovi Pratama, "Perbandingan Metode Random Forest Classifier dan SVM Pada Klasifikasi Kemampuan Level Beradaptasi Pembelajaran Jarak Jauh Siswa," *Bulletin of Informatics and Data Science*, vol. 1, no. 2, pp. 98–98, Nov. 2022, doi: <https://doi.org/10.61944/bids.v1i2.49>.

Halaman ini sengaja dibiarkan kosong

Analisis Sentimen Ulasan Traveloka Menggunakan Metode Naïve Bayes Classifier dan Information Gain

Kadek Yuni Suratri¹, I Gede Santi Astawa²

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹yunisuratri@gmail.com
²santi.astawa@unud.ac.id

Abstract

In the increasingly rapid digital era, Traveloka is present as an online travel agency that makes it easier for users to order and plan trips. Reviews left by users can reflect the user's experience in using the platform. Indirectly, reviews can also reflect user satisfaction. Therefore, it is important to carry out sentiment analysis of existing reviews so that you can improve service quality. This research examines the performance of the Information Gain feature selection in classifying the sentiment of Traveloka application reviews using the Naïve Bayes method. The research results show that classification using the Naïve Bayes model obtained an accuracy of 83%, precision of 81%, and recall of 98%. Meanwhile, classification with feature selection obtained an accuracy of 79%, precision of 76%, and recall of 100%. This shows that the feature selection performance has not been able to increase the accuracy value.

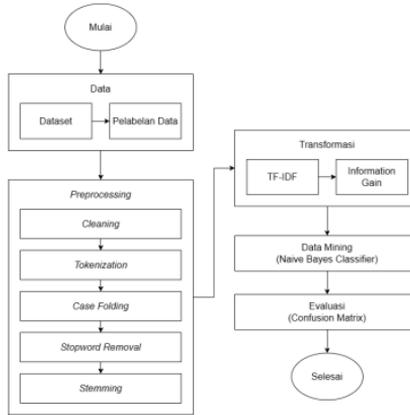
Keywords: Sentiment Analysis, Reviews, Traveloka, Naïve Bayes Classifier, TF-IDF, Information Gain

1. Pendahuluan

Di era digital yang berkembang semakin pesat, Traveloka hadir sebagai salah satu online travel agency yang memudahkan pengguna dalam pemesanan dan merencanakan perjalanan. Dilansir dari situs similarweb, total kunjungan pada situs Traveloka terhitung sampai bulan April 2024 telah mencapai sebesar 22.3 juta. Sebanyak 50 juta pengguna telah mengunduh aplikasi Traveloka pada Google Play Store. Hal ini menandakan suatu kepercayaan besar terhadap platform ini. Ulasan yang ditinggalkan oleh pengguna dapat mencerminkan pengalaman pengguna dalam menggunakan platform. Ulasan yang diberikan dapat berupa feedback, mulai dari pujian atas kemudahan dalam penggunaan platform hingga kritik terhadap segi layanan yang ditawarkan, masalah teknis. Secara tidak langsung, ulasan juga dapat mencerminkan kepuasan pengguna. Oleh karena itu, penting melakukan analisis sentimen terhadap ulasan yang ada sehingga dapat meningkatkan kualitas pelayanan. Dengan adanya peningkatan volume ulasan, diperlukan juga adanya suatu sistem untuk menganalisis sentimen secara otomatis. Analisis sentimen merupakan proses menganalisis teks dengan tujuan memperoleh informasi berupa sentimen baik positif, negatif ataupun netral di dalamnya. Analisis sentimen termasuk ke dalam salah satu bidang Natural Language Processing (NLP) yang berfokus pada identifikasi, ekstraksi, dan mengubah informasi subjektif dalam teks. Algoritma Naïve Bayes Classifier (NBC) merupakan salah satu algoritma yang populer dan sering terpilih dalam melakukan analisis sentimen. Hal ini karena algoritma NBC sederhana, mampu mengatasi missing value, mengelola atribut yang beragam, gangguan dalam data, memiliki kinerja yang baik sehingga cocok digunakan pada dataset dengan jumlah besar [1]. Dalam analisis sentimen, ekstraksi fitur memegang peranan penting sebagai jembatan antara data raw dan algoritma Machine Learning. Pada proses ini kompleksitas dan dimensi dari data teks yang sangat besar dapat berkurang, sehingga model dapat fokus pada elemen-elemen yang paling berpengaruh terhadap sentimen. Selain itu, ekstraksi fitur dapat membantu dalam mengeliminasi noise yang dapat merusak akurasi prediksi. Pada penelitian ini akan menguji performa seleksi fitur Information Gain (IG) dalam mengoptimalkan performa algoritma NBC.

2. Metode Penelitian

Penelitian ini mengimplementasikan metode Knowledge Discovery in Database (KDD) yang meliputi beberapa tahapan seperti pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola ataupun hubungan pada dataset yang berukuran besar [2].



Gambar 1. Tahap Penelitian

Penelitian ini diawali dengan tahap mengumpulkan data. Data yang telah dikumpulkan akan diberi label. Selanjutnya akan dilakukan tahap preprocessing teks. Kemudian data yang telah bersih akan masuk ke tahap transformasi. Tahapan transformasi diawali dengan penghitungan ekstraksi fitur menggunakan Term Frequency-Inverse Document Frequency (TF-IDF) dan dilanjutkan ke proses seleksi fitur dengan Information Gain (IG). Berikutnya masuk ke tahap klasifikasi data dengan metode Naive Bayes Classifier. Terakhir merupakan tahap evaluasi dengan confusion matrix. Hal ini dilakukan untuk mengukur validitas dari tahap klasifikasi.

2.1. Dataset

reviewId	userName	userName	content	score	thumbsUpCount	reviewCount	replyContent	replyDate	aspiration	
0	9451333a	64-MuA-Supria	https://play.Oh...	5	0	0.1861111	22/10/23 5:30	Hi, thank you!	22/10/23 5:12	0.1861111
1	61526420	no-ny-htng	https://play.Lasah-kel...	5	0	0.1861111	22/10/23 5:26	Hi Kak, semar	20/10/23 5:12	0.1861111
2	10746205	81-Sigit-Murny	https://play.Selaku-paka...	5	0	0.1861111	22/10/23 4:50	Hi Kak, teri	22/10/23 5:12	0.1861111
3	63d3d310	64-Rokhan-Akhan	https://play.mmmalab...	5	0	0.1861111	22/10/23 4:50	Hi Kak, semar	22/10/23 5:12	0.1861111
4	1046421b	10-Lulu-Ahm	https://play.Sip...	5	0	0.1861111	22/10/23 4:53	Thank you fo	22/10/23 5:12	0.1861111
5	1043247a	6-Sumati	https://play.The best app...	5	0	0.1861111	22/10/23 4:49	Hi, thank you!	22/10/23 5:12	0.1861111
6	1451514f	64-Rokhan-Akhan	https://play.Maklaima...	5	0	0	22/10/23 4:49	Hi Kak, teri	22/10/23 4:42	0.1861111
7	1390488b	67-Iqah-Ruhani	https://play.Limuk-car...	5	0	0	22/10/23 4:33	Hi Kak, semar	22/10/23 4:42	0.1861111
8	1a8b710b	64-Ah-Paripal	https://play.Mengamam...	5	0	0	22/10/23 4:33	Hi Kak, teri	22/10/23 4:42	0.1861111
9	105050ed	81-Nanda-Khana	https://play.Akhal-Anda...	5	0	0.1786111	22/10/23 4:31	Hi Kak, teri	22/10/23 4:42	0.1786111
10	12a0271c	10-William-They	https://play.Masa-Stepen...	5	0	0.1861111	22/10/23 3:06	Hi Kak, semar	22/10/23 3:23	0.1861111
11	10716c1b	10-Joseph-Oh-Itung	https://play.Trenada-ka...	5	0	0.1861111	22/10/23 2:27	Hi Kak, teri	22/10/23 4:42	0.1861111
12	10808220	10-Lili	https://play.Pengangan...	5	0	0.1861111	22/10/23 2:27	Hi Kak, semar	20/10/23 5:12	0.1861111
13	11722400	10-perti-Isabella	https://play.Oh...	5	0	0.1861111	22/10/23 0:50	Hi, thank you!	22/10/23 5:12	0.1861111
14	12249749	64-Sam-Isagah	https://play.Sipag-Ingat...	5	0	0.1861111	22/10/23 0:50	Hi Kak, semar	22/10/23 5:12	0.1861111
15	10445400	64-Melinda-ya	https://play.Sihir-gemp...	5	0	0.1861111	22/10/23 0:50	Hi Kak, teri	22/10/23 5:12	0.1861111
16	1046104d	64-Dan-Wanika	https://play.Murung-pom...	1	0	0.1861111	22/10/23 2:53	Hi Di, semar	22/10/23 2:23	0.1861111
17	10214616	64-Luci-M-ama	https://play.Ternak-kasih...	5	0	0	22/10/23 2:50	Hi Kak, teri	22/10/23 2:53	0.1861111
18	10217414	64-Muhammad	https://play.The di-perk...	1	0	0.1861111	22/10/23 2:49	Hi Kak, semar	22/10/23 2:23	0.1861111
19	1023045d	10-Sahel-Suhani	https://play.Akham-kulit...	5	0	0.1861111	22/10/23 18:04	Hi Kak, teri	22/10/23 18:12	0.1861111
20	12170648	10-Toni-Singih	https://play.Sudat...	1	0	0.1861111	22/10/23 15:52	Hi Semar...	22/10/23 18:12	0.1861111
21	10392463	10-Rendi-Durna	https://play.Sipag...	5	0	0.1786111	22/10/23 15:52	Hi Kak, teri	22/10/23 18:12	0.1786111
22	10706810	64-Ahmad-Dinar	https://play.Sihir-gem...	5	0	0.1861111	22/10/23 15:50	Hi Kak, semar	22/10/23 15:12	0.1861111
23	10132516	10-Daddy-Kahk	https://play.Sipag-semang...	5	0	0	22/10/23 14:51	Hi Kak, teri	22/10/23 18:12	0.1861111
24	1030791a	10-Fanni-Fiahy	https://play.Mulid-dan-g...	5	0	0	22/10/23 13:53	Hi Kak, teri	22/10/23 18:12	0.1861111
25	10186496	64-Agus-Pagi	https://play.Sipag...	5	0	0.1854187	22/10/23 12:26	Hi Semar...	22/10/23 12:42	0.1854187
26	10349210	10-Obay-Utomo	https://play.Akhal-Anda...	1	4	0.1854187	22/10/23 12:26	Hi Kak, semar	10/11/23 18:30	0.1854187

Gambar 2. Dataset

Penelitian ini menggunakan dataset yang telah tersedia di situs Kaggle. Proses pengambilan data dilakukan dengan mengunduh secara langsung dari situs. Total dataset berjumlah 177.159, namun pada penelitian hanya digunakan sebanyak 1.501 data. Pemberian label pada dataset dilakukan dengan bahasa pemrograman Python dan menjadikan score sebagai indikator pelabelan. Ulasan akan dikelompokkan ke dalam 3 kelas sentimen yaitu sentiment positif, negatif dan netral.

2.2. Preprocessing Text

Teks preprocessing merupakan proses yang dilakukan sebelum data teks diolah menggunakan algoritma NLP dengan tujuan pembersihan dan penyusunan ulang data teks agar terstruktur, sehingga lebih mudah diolah oleh algoritma. Tahap pertama dari preprocessing yaitu cleaning, membersihkan teks dari segala elemen yang tidak relevan seperti menghapus URL, karakter khusus, angka, dan sebagainya. Dilanjutkan pada tahap kedua yaitu tokenization, memecah kalimat menjadi sebuah kata atau yang disebut token. Dilanjutkan pada tahap ketiga yaitu case folding, mengubah keseluruhan huruf menjadi huruf kecil. Tahap keempat yaitu stopword removal, menghapus kata yang tidak diperlukan seperti kata konjungsi. Tahap terakhir yaitu stemming, mengubah kata ke bentuk dasar dengan cara menghilangkan imbuhan dari kata tersebut.

2.3. TF-IDF (Term Frequency-Inverse Document Frequency)

TF-IDF merupakan algoritma dari ekstraksi fitur untuk mengukur seberapa penting sebuah kata dalam dokumen. Nilai TF-IDF didapat dengan mengalikan nilai dari Term Frequency (TF) dan Inverse Document Frequency (IDF). TF menghitung seberapa sering kemunculan suatu kata pada sebuah dokumen sedangkan IDF menghitung seberapa jarang/tidak sering munculnya suatu kata di dokumen lain [3]. Adapun representasi sistematis untuk menghitung pembobotan setiap token t di dokumen d sebagai berikut:

$$tfidf_t = f_{t,d} \times \log \frac{N}{df_t} \quad (1)$$

$tfidf_t$ merupakan bobot dari kata t di dokumen d , $f_{t,d}$ merupakan kemunculan kata t dalam dokumen d , N merupakan total dari dokumen, dan df_t merupakan banyaknya dokumen yang mengandung kata- t .

2.4. Information Gain

Information gain (IG) merupakan teknik yang digunakan dalam seleksi fitur dengan menggunakan metode penilaian untuk fitur kategori dan juga untuk pembobotan atribut yang berkelanjutan (kontinu) yang kemudian diubah menjadi nilai diskrit menggunakan entropi maksimal [4]. Suatu entropi digunakan untuk mengukur jumlah informasi yang diperlukan untuk mengkodekan suatu kelas [5]. IG dari suatu fitur diukur dengan menghitung jumlah informasi yang diperoleh dari prediksi kategori dengan mempertimbangkan keberadaan atau ketiadaan fitur tersebut dalam dokumen. Adapun representasi sistematisnya sebagai berikut:

$$\text{InfoGain}(S, A) = \text{Entropy}(S) - \sum_{v \in \text{Value}(A)} \frac{|S_v|}{|S|} \text{Entropy}(S_v) \quad (2)$$

$$\text{Entropy}(S) = - \sum \frac{|S_i|}{|S|} \log_2 \frac{|S_i|}{|S|} \quad (3)$$

S merupakan jumlah seluruh fitur, A merupakan kategori, S_v merupakan jumlah sampel untuk nilai v , v merupakan nilai yang mungkin untuk kategori A , S_i merupakan fitur ke i dan $\text{Value}(A)$ merupakan himpunan nilai-nilai yang mungkin untuk kategori A . Fitur yang dipilih merupakan fitur yang memiliki IG yang berbeda dari nol dan lebih besar dari nilai threshold tertentu. Konsep dibalik penggunaan IG untuk menentukan fitur yang paling signifikan dalam menjelaskan kategori.

2.5. Algoritma Naïve Bayes Classifier

Algoritma Naive Bayes Classifier (NBC) merupakan sebuah model klasifikasi dengan menggunakan prinsip Naive Bayes dalam memprediksi kelas dari suatu data berdasarkan fitur-fiturnya. Naive Bayes Classifier sendiri merupakan metode pembelajaran probabilistik di mana setiap kata memiliki probabilitas kemunculan yang independen. Ini berarti bahwa nilai atribut kategori tidak saling mempengaruhi atau dipengaruhi oleh nilai atribut lainnya [6].

2.6. Evaluasi

Tahap terakhir merupakan tahap evaluasi dari performa model klasifikasi dengan confusion matrix. Confusion matrix merupakan sebuah tabel dengan informasi di dalamnya tentang jumlah prediksi keseluruhan benar dan salah yang dilakukan oleh model untuk setiap kelas dari data. Hal yang akan dihitung adalah nilai accuracy, recall dan precision.

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \quad (4)$$

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \quad (5)$$

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \quad (6)$$

TP merupakan jumlah data positif yang diprediksikan benar sebagai nilai positif, FN merupakan jumlah data positif yang salah diklasifikasikan sebagai nilai negatif, FP merupakan jumlah data negatif yang salah diklasifikasikan sebagai nilai positif, dan TN merupakan jumlah data negatif yang diprediksi benar sebagai nilai negatif.

3. Hasil dan Diskusi

3.1. Preprocessing

Dilakukan tahap preprocessing pada data sebelum memasuki tahap berikutnya. Pada tahap preprocessing, data akan dibersihkan dengan menghilangkan noise melalui penghapusan URL, simbol, angka, dan lain-lain. Dataset yang akan di preprocessing berjumlah 1.232 karena telah melalui proses penghilangan redundansi.

Tabel 1. Hasil Preprocessing

Tahap preprocessing	Hasil
Dataset	Selalu pakai Traveloka untuk pembelian tiket dan hotel kemana pun... 🍷
Cleaning	Selalu pakai Traveloka untuk pembelian tiket dan hotel kemana pun
Tokenization dan Case Folding	selalu, pakai,traveloka,untuk,pembelian,tiket,dan,hotel,kemana,pun
Stopword Removal	pakai,traveloka,pembelian,tiket,hotel,kemana
Stemming	pakai traveloka beli tiket hotel mana

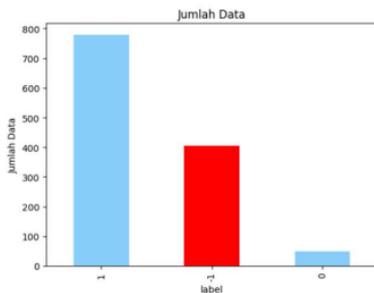
3.1.1. Labeling

Pada labeling, akan dilakukan pengklasifikasian sentimen ke dalam 3 kelas kata yaitu sentimen positif, negatif, dan netral [7].

```
#labeling
def pelabelan(skor):
    if skor < 3:
        return '-1'
    elif skor == 3:
        return '0'
    elif skor > 3 :
        return '1'
data['label'] = data['skor'].apply(pelabelan)
data.head(10)
```

Gambar 3. Labeling

Pada gambar 3, labeling dilakukan berdasarkan dengan score atau rating yang diberikan oleh pengguna. Apabila skor yang diberikan bernilai < 3 (lebih kecil dari 3), maka sentimen akan diklasifikasikan menjadi sentimen negatif. Skor == 3 merupakan sentimen netral, dan skor > 3 (kurang dari 3) merupakan sentimen positif. Didapat sebanyak 405 data dengan label negatif, 779 data dengan label positif, dan 48 data dengan label negatif.



Gambar 4. Persebaran Label

Tabel 2. Contoh Data Label

Data	Label
aplikasi eror	Negatif
aplikasi curang rubah tanggal sistem buruk	Negatif
aplikasi digital mesan pesawat hotel murah	Positif
aplikasi nya keren bantu cari hotel tiket pesawat	Positif
kk sy pesan hotel tp udah sy batal knp sy ttp bayar ya	Netral
Aplikasi lot	Netral

3.1.2. Pembobotan Kata

Pada tahap pembobotan kata menggunakan TF-IDF (Term Frequency – Inverse Document Frequency). Sebelum dilakukan TF-IDF, dataset akan dibagi menjadi data train dan data test dengan rasio perbandingan berturut-turut 80:20. Jumlah data train sebanyak 985 dan jumlah data test sebanyak 247.

Confusion Matrix:

$$\begin{bmatrix} 39 & 0 & 42 \\ 3 & 0 & 7 \\ 0 & 0 & 156 \end{bmatrix}$$

Gambar 7. Confusion Matrix Naïve Bayes + IG

Berdasarkan gambar 6, diperoleh accuracy sebesar 83%, precision sebesar 81%, dan recall sebesar 98%. Hasil klasifikasi perolehan dengan model Naïve Bayes dapat dilihat pada gambar 8.

Classification Report:				
	precision	recall	f1-score	support
-1	0.88	0.63	0.73	81
0	0.00	0.00	0.00	10
1	0.81	0.98	0.89	156
accuracy			0.83	247
macro avg	0.56	0.54	0.54	247
weighted avg	0.80	0.83	0.80	247

Gambar 8. Klasifikasi dengan Naïve Bayes

Berdasarkan gambar 7, diperoleh accuracy sebesar 79%, precision sebesar 76%, dan recall sebesar 100%. Hasil klasifikasi perolehan dengan model Naïve Bayes dan seleksi fitur IG dapat dilihat pada gambar 9.

Classification Report:				
	precision	recall	f1-score	support
-1	0.93	0.48	0.63	81
0	0.00	0.00	0.00	10
1	0.76	1.00	0.86	156
accuracy			0.79	247
macro avg	0.56	0.49	0.50	247
weighted avg	0.79	0.79	0.75	247

Gambar 9. Klasifikasi Naïve Bayes dan Information Gain

Pada tabel 3 merupakan perbandingan performa model antara Naïve Bayes dan Naïve Bayes + IG berdasarkan akurasi.

Tabel 3. Perbandingan Performa Model Berdasarkan Accuracy

Metode	Akurasi
Naïve Bayes	83%
Naïve Bayes + IG	79%

Berdasarkan tabel perbandingan hasil performa, didapat bahwa performa kinerja dari model Naïve Bayes lebih besar dibandingkan dengan performa kinerja dengan menggunakan tambahan seleksi fitur.

4. Kesimpulan

Berdasarkan penelitian yang telah dilakukan terhadap penggunaan seleksi fitur dengan menggunakan metode Naïve Bayes, performa seleksi fitur belum cukup baik jika dibandingkan dengan tidak menggunakan seleksi fitur. Diperoleh nilai akurasi sebesar 79%, dimana seleksi fitur belum mampu meningkatkan akurasi klasifikasi. Adapun hal yang dapat diperbaiki untuk

penelitian kedepannya yaitu pada pelabelan dapat meminta bantuan ahli untuk menganalisis supaya lebih akurat, dan menggunakan dataset dalam jumlah persebaran label yangimbang.

Daftar Pustaka

- [1] Q. A'yunyah et al., "Implementasi Algoritma Naïve Bayes Classifier (NBC) untuk Klasifikasi Penyakit Ginjal Kronik," *J. Sist. Komput. dan Inform.*, vol. 4, no. 1, p. 72, 2022, doi: 10.30865/json.v4i1.4781.
- [2] D. S. O. Panggabean, E. Buulolo, and N. Silalahi, "Penerapan Data Mining Untuk Memprediksi Pemesanan Bibit Pohon Dengan Regresi Linear Berganda," *JURIKOM (Jurnal Ris. Komputer)*, vol. 7, no. 1, p. 56, 2020, doi: 10.30865/jurikom.v7i1.1947.
- [3] I. Syahrohimi, S. D. Saputra, and R. W. Saputra, "Pilpres 2024 Twitter Menggunakan," vol. 12, no. 2, 2024.
- [4] I. Maulida, A. Suyatno, and H. R. Hatta, "Seleksi Fitur Pada Dokumen Abstrak Teks Bahasa Indonesia Menggunakan Metode Information Gain," *J. SIFO Mikroskil*, vol. 17, no. 2, pp. 249–258, 2016, doi: 10.55601/jsm.v17i2.379.
- [5] D. Abadi, "Perbandingan algoritme feature selection information gain dan symmetrical uncertainty pada data ketahanan pangan," 2013.
- [6] A. Isnanda, Y. Umaidah, and J. H. Jaman, "Implementasi Naïve Bayes Classifier Dan Information Gain Pada Analisis Sentimen Penggunaan E-Wallet Saat Pandemi," *J. Teknol. Inform. dan Komput.*, vol. 7, no. 2, pp. 144–153, 2021, doi: 10.37012/jtik.v7i2.648.
- [7] N. R. Siahaan, R. Y. Tiffany, and S. R. E. Sinaga, "Analisis Sentimen Ulasan Aplikasi Media Sosial Whatsapp Menggunakan Metode Naive Bayes Classifier," *J. Ilm. Betrik*, vol. 14, no. 02, pp. 343–354, 2023, [Online]. Available: <https://ejournal.pppmitpa.or.id/index.php/betrik/article/view/104%0Ahttps://ejournal.pppm itpa.or.id/index.php/betrik/article/download/104/76>

Perancangan Desain Antarmuka Aplikasi Soul Notes dengan Metode Design Thinking Berbasis Mobile

Ni Komang Purnami¹, I Gede Surya Rahayuda²

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹purnami.2208561071@student.unud.ac.id
²igedesuryarahayuda@unud.ac.id

Abstract

This study aimed to design an interface for the Soul Notes application using design thinking methodology, focusing on mobile-based applications. The study employed a qualitative approach, involving 5 participants who were asked to use the application and provide feedback. The results showed that the participants found the application easy to use and appreciated its features, such as mood tracking and relaxation techniques. The study also found that the participants valued the application's ability to provide a platform for expressing emotions freely and securely. The results of the study were analyzed using the Single Ease Question (SEQ) method, which indicated a high level of ease of use, with an average score of 6.52 out of 7. The study's findings suggest that the Soul Notes application can effectively support mental well-being and emotional management, and that its design should prioritize user experience and security.

Keywords: Design Thinking, UI/UX, Journaling, SEQ

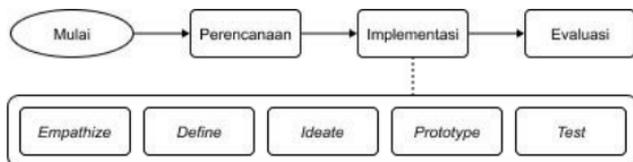
1. Pendahuluan

Pada era modern ini, tuntutan hidup semakin meningkat dengan cepat, memaksa individu untuk beradaptasi dengan beragam tekanan dari pekerjaan, hubungan sosial, dan ekspektasi diri sendiri. Akibatnya, tidak sedikit individu merasa tertekan dan cenderung kesulitan dalam menyampaikan serta mengelola perasaan dengan baik. Dalam kebanyakan kasus, tekanan ini dapat berdampak negatif pada kesejahteraan mental dan emosional. Survei Kesehatan mental nasional yang mengukur angka gangguan mental pada remaja 10-17 tahun yaitu Indonesia National Adolescent Mental Health Survey (I-NAMHS) menunjukkan satu dari tiga remaja Indonesia mempunyai masalah kesehatan mental sementara satu dari dua puluh remaja memiliki gangguan mental [1]. Perasaan yang sulit disampaikan sering kali diekspresikan melalui beberapa cara seperti bercerita dengan orang terdekat, berdoa, meditasi dan lain sebagainya. Selain itu, journaling juga menjadi cara yang positif untuk mengekspresikan diri. Journaling merupakan aktivitas mencatat ide, pemikiran, emosi, dan peristiwa dalam hidup menggunakan tulisan di buku, komputer, atau gambar. Jurnal pribadi dapat menjadi alat yang bermanfaat untuk mengekspresikan perasaan, mengendalikan emosi, mengenal diri sendiri, serta meredakan stres dan rasa cemas [2]. Media journaling yang biasa digunakan yaitu buku cacatan. Tetapi, dengan perkembangan teknologi, kegiatan journaling dapat juga dilakukan pada ponsel pintar tanpa memerlukan buku dan alat tulis. Terciptanya platform yang aman dan ramah bagi pengguna untuk mengekspresikan perasaan, pemikiran, dan pengalaman sehari-hari secara konsisten perlu memperhatikan aspek User Interface (UI) dan User Experience (UX). UI (User Interface) diartikan sebagai interaksi antara sistem dan pengguna melalui perintah seperti menggunakan konten dan memasukkan data. Sementara itu, UX (User Experience) diartikan sebagai pengalaman yang berkaitan dengan reaksi, persepsi, perilaku, emosi, dan pikiran pengguna saat menggunakan sistem [3]. Ada beberapa metode yang digunakan untuk merancang antarmuka dari suatu aplikasi salah satunya metode design thinking. Metode Design Thinking adalah pendekatan yang berorientasi pada manusia untuk inovasi dan pemecahan masalah yang melibatkan pengguna dalam setiap tahapnya. Metode ini terdiri dari lima langkah antara lain empathize, define, ideate, prototype, dan test [4]. Pada tahap test, SEQ (Single Ease Question) digunakan untuk menguji

desain yang dirancang. Berdasarkan uraian permasalahan tersebut, penelitian ini akan membuat perancangan desain antarmuka aplikasi mobile Soul Notes dengan metode design thinking. Soul Notes merupakan aplikasi journaling yang memungkinkan pengguna untuk mencatat pikiran, perasaan, pengalaman, atau kejadian dalam kehidupan mereka sehari-hari. Penelitian ini bertujuan untuk menciptakan desain antarmuka aplikasi Soul Notes yang ramah pengguna.

2. Metode Penelitian

Ada beberapa tahapan yang dilakukan pada penelitian ini yaitu perencanaan, implementasi, dan evaluasi. Pada tahap perencanaan, studi literatur diperlukan untuk mendapatkan referensi yang relevan untuk kepentingan penelitian. Metode design thinking diterapkan pada tahap implementasi. Tahapan evaluasi dilakukan untuk memperbaiki rancangan desain antarmuka berdasarkan pengujian yang dilakukan.



Gambar 1. Alur Penelitian

2.1. Perencanaan

Pada tahap perencanaan, studi literatur akan dilakukan dengan mengumpulkan referensi untuk kepentingan penelitian. Gagasan diperoleh dari jurnal dan website yang relevan dengan topik penelitian.

2.2. Implementasi

Pada tahap implementasi, metode design thinking akan diterapkan dalam beberapa langkah antara lain empathize, define, ideate, prototype, dan test.

2.2.1. Empathize

Empathize merupakan kegiatan untuk menggali permasalahan yang dihadapi pengguna [6]. Tahap ini melibatkan pengumpulan data dan informasi yang komprehensif tentang pengguna, termasuk preferensi, kebiasaan, dan tantangan yang mereka hadapi. Ini dilakukan melalui berbagai metode seperti wawancara mendalam, observasi langsung, dan analisis data pengguna. Pada penelitian ini data dan informasi dikumpulkan melalui survei dengan menyebarkan kuisioner. Selain itu, desainer juga berusaha untuk merasakan dan memahami pengalaman pengguna secara emosional, sosial, dan psikologis, sehingga mereka dapat melihat dunia melalui sudut pandang pengguna dan memahami kebutuhan serta motivasi mereka secara mendalam. Hasil dari tahap empathize ini adalah pemahaman yang lebih dalam tentang pengguna, memungkinkan desainer untuk mengidentifikasi peluang inovasi dan menciptakan pengalaman yang lebih relevan dan memuaskan bagi pengguna. Dengan memahami dengan baik siapa pengguna, apa yang mereka butuhkan, dan bagaimana mereka berinteraksi dengan produk atau layanan yang sedang dirancang, desainer dapat memastikan bahwa desain UI/UX yang dihasilkan memenuhi kebutuhan pengguna secara optimal.

2.2.2. Define

Tahap "Define" pada metode design thinking perancangan UI/UX adalah tahap ketiga dalam proses desain thinking. Pada tahap ini, informasi yang dikumpulkan selama tahap "Empathize" digunakan untuk memahami lebih lanjut kebutuhan dan masalah pengguna. Informasi ini

digunakan untuk mengidentifikasi dan memformulasikan masalah yang akan dipecahkan melalui desain UI/UX. Tahap ini memungkinkan tim desainer untuk memahami lebih dalam apa yang diperlukan pengguna dan bagaimana mereka berinteraksi dengan produk atau layanan yang sedang dirancang. Dalam tahap "Define", tim desainer harus mengidentifikasi dan memformulasikan masalah yang akan dipecahkan dengan cara mengumpulkan informasi yang relevan dan memahami lebih lanjut kebutuhan pengguna. Hal ini dilakukan dengan cara menganalisis data yang dikumpulkan selama tahap "Empathize" dan mengidentifikasi pattern dan kebutuhan yang ditemukan. Dengan demikian, tim desainer dapat membuat definisi yang jelas dan spesifik dari masalah yang akan dipecahkan, serta memahami lebih lanjut bagaimana pengguna berinteraksi dengan produk atau layanan yang sedang dirancang. Dalam beberapa jurnal, tahap "Define" juga diterjemahkan sebagai "penentuan masalah" atau "pengidentifikasi kebutuhan". Dalam beberapa kasus, tahap ini melibatkan analisis data yang lebih lanjut dan pengumpulan informasi tambahan untuk memahami lebih lanjut kebutuhan pengguna. Dengan demikian, tim desainer dapat membuat desain yang lebih efektif dan efisien yang memenuhi kebutuhan pengguna [1][2][3][4].

2.2.3. Ideate

Tahap "Ideate" pada metode design thinking perancangan UI/UX adalah tahap keempat dalam proses desain thinking. Pada tahap ini, tim desainer menggunakan informasi yang dikumpulkan selama tahap "Define" untuk mengembangkan ide-ide yang dapat memecahkan masalah yang ditemukan. Tahap "Ideate" ini memungkinkan tim desainer untuk berpikir kreatif dan mencari solusi yang inovatif untuk memenuhi kebutuhan pengguna. Dalam tahap "Ideate", tim desainer harus berpikir luar biasa dan mencari ide-ide yang dapat memenuhi kebutuhan pengguna. Hal ini dilakukan dengan cara mengumpulkan ide-ide yang diperoleh dari analisis data dan diskusi tim. Dengan demikian, tim desainer dapat membuat daftar ide yang luas dan beragam, yang kemudian dapat diuji dan diperbaiki melalui tahap "Prototype" dan "Test". Dalam beberapa jurnal, tahap "Ideate" juga diterjemahkan sebagai "pembuatan ide" atau "pembuatan solusi". Dalam beberapa kasus, tahap ini melibatkan analisis data yang lebih lanjut dan pengumpulan informasi tambahan untuk memahami lebih lanjut kebutuhan pengguna. Dengan demikian, tim desainer dapat membuat desain yang lebih efektif dan efisien yang memenuhi kebutuhan pengguna.

2.2.4. Prototype

Tahap "Prototype" pada metode design thinking perancangan UI/UX adalah tahap kelima dalam proses desain thinking. Pada tahap ini, tim desainer menggunakan ide-ide yang dikembangkan selama tahap "Ideate" untuk membuat sebuah prototipe yang dapat diuji dan diperbaiki. Prototipe ini digunakan untuk menguji dan memvalidasi desain yang telah dibuat, serta untuk mendapatkan umpan balik yang lebih spesifik dari pengguna. Dalam tahap "Prototype", tim desainer harus membuat sebuah prototipe yang dapat diuji secara efektif dan efisien. Hal ini dilakukan dengan cara mengumpulkan ide-ide yang diperoleh dari analisis data dan diskusi tim, serta mengembangkan prototipe yang dapat memenuhi kebutuhan pengguna. Dengan demikian, tim desainer dapat membuat desain yang lebih efektif dan efisien yang memenuhi kebutuhan pengguna. Dalam beberapa jurnal, tahap "Prototype" juga diterjemahkan sebagai "pembuatan prototipe" atau "pembuatan model". Dalam beberapa kasus, tahap ini melibatkan analisis data yang lebih lanjut dan pengumpulan informasi tambahan untuk memahami lebih lanjut kebutuhan pengguna. Dengan demikian, tim desainer dapat membuat desain yang lebih efektif dan efisien yang memenuhi kebutuhan pengguna.

2.2.5. Test

Tahap "Test" dalam metode design thinking perancangan UI/UX adalah tahap terakhir dalam proses desain thinking. Pada tahap ini, tim desainer menguji dan memvalidasi desain yang telah dibuat melalui prototipe yang telah dibuat sebelumnya. Tujuan dari tahap "Test" adalah untuk memastikan bahwa desain yang dibuat memenuhi kebutuhan pengguna dan dapat digunakan secara efektif dan efisien.

Pada tahap "Test", tim desainer harus menguji desain yang telah dibuat dengan cara melakukan pengujian secara langsung kepada pengguna. Hal ini dilakukan dengan cara memberikan prototipe kepada pengguna dan meminta mereka untuk menggunakan aplikasi tersebut dan memberikan umpan balik tentang bagaimana mereka berinteraksi dengan aplikasi tersebut. Dengan demikian, tim desainer dapat memperbaiki dan meningkatkan desain yang telah dibuat untuk memenuhi kebutuhan pengguna lebih baik. Dalam beberapa jurnal, tahap "Test" juga diterjemahkan sebagai "pengujian" atau "evaluasi". Dalam beberapa kasus, tahap ini melibatkan analisis data yang lebih lanjut dan pengumpulan informasi tambahan untuk memahami lebih lanjut kebutuhan pengguna. Dengan demikian, tim desainer dapat membuat desain yang lebih efektif dan efisien yang memenuhi kebutuhan pengguna. Metode Single Ease Question (SEQ) adalah metode yang digunakan untuk menilai kemudahan menyelesaikan tugas atau skenario. Dalam metode ini, responden diminta untuk menilai kemudahan tugas menggunakan skala Likert dengan tujuh poin, mulai dari "sangat sulit" hingga "sangat mudah".

Tabel 1. Skala SEQ

Skor	Keterangan
1	Sangat sulit
2	Sulit
3	Cukup sulit
4	Netral
5	Cukup mudah
6	Mudah
7	Sangat mudah

Tabel 2. Instrumen Penelitian

Rentang Skor	Keterangan
1 – 1,9	Sangat sulit
2 – 2,9	Sulit
3 – 3,9	Cukup sulit
4 – 4,9	Netral
5 – 5,9	Cukup mudah
6 – 6,9	Mudah
7	Sangat mudah

Metode Single Ease Question (SEQ) biasanya digunakan dalam penelitian untuk mengetahui persepsi responden tentang kemudahan menyelesaikan tugas atau skenario. Cara kerja metode ini adalah dengan merancang kuesioner yang menampilkan tugas atau skenario kepada responden, lalu setelah responden menyelesaikan tugas atau skenario tersebut, mereka menilai kemudahan menyelesaikannya menggunakan skala Likert yang disediakan. Dengan demikian, metode ini dapat memberikan wawasan berharga tentang persepsi responden tentang kemudahan tugas atau skenario dan dapat digunakan dalam berbagai konteks penelitian.

2.3. Evaluasi

Tahap evaluasi diperlukan untuk mengetahui tingkat keberhasilan suatu program, meningkatkan kemungkinan tercapainya suatu tujuan atau inisiatif, memastikan sumber daya yang bermanfaat, mengidentifikasi apa dan mengapa suatu kegiatan tidak berjalan sesuai dengan rencana, serta

untuk meningkatkan efektivitas dan produktivitas. Dengan demikian, evaluasi membantu dalam menentukan kebijaksanaan dalam usaha memperbaiki pendidikan atau kegiatan lainnya.

3. Hasil dan Pembahasan

3.1. Empathize

3.1.1. Survei

Survei dilaksanakan dengan mengirimkan kuesioner melalui Google Form kepada individu yang memiliki sebanyak 5 responden berpartisipasi dalam survei ini, dengan usia berkisar antara 20 hingga 21 tahun. Hasil yang diharapkan dari survei tersebut adalah pemahaman yang lebih baik tentang kebutuhan dan harapan pengguna terhadap aplikasi Soul Notes. Ini akan membantu dalam merancang aplikasi yang lebih sesuai dengan kebutuhan pengguna, meningkatkan peluang kesuksesan, dan memberikan wawasan penting untuk pengembangan selanjutnya.

3.1.2. Affinity Diagram

Affinity diagram berikut berasal dari hasil kuesioner yang diisi oleh 5 orang responden.



Gambar 2. Affinity Diagram

3.2. Define

3.2.1. User Journey Map



Gambar 3. User Journey Map

3.2.2. User Persona



Gambar 4. User Persona

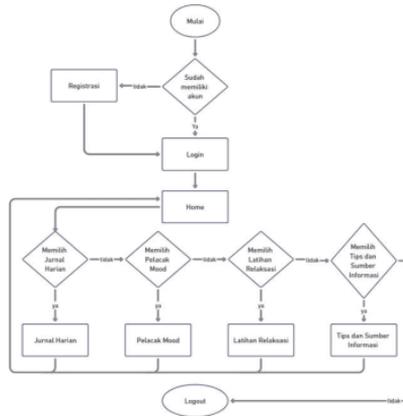
3.2.3. Problem Statement dan Solusi

Tabel 3. Problem Statement dan Solusi

Problem Statement	Solusi
Bagaimana aplikasi dapat membantu individu mengelola stres dan kecemasan sehari-hari dengan cara yang efektif?	Membuat aplikasi yang memiliki fitur pelacak mood dan latihan relaksasi untuk membantu individu mengelola stres dan kecemasan sehari-hari secara terstruktur.
Bagaimana aplikasi dapat memenuhi kebutuhan individu dalam mengekspresikan perasaan secara bebas dan mencurahkan segalanya?	Membuat aplikasi yang memiliki fitur draw atau mewarnai, serta platform untuk mengekspresikan perasaan secara bebas dan mencurahkan segalanya.
Bagaimana aplikasi dapat mempertahankan privasi dan keamanan pengguna dalam menggunakan aplikasi?	Membuat aplikasi yang tentunya secure dalam mengatasi stress keseharian dan memiliki fitur pelacakan mood yang kurang, serta memberikan konten relaksasi seperti musik atau suara alam serta tips harian untuk membantu pengguna mengelola stres dan kecemasan sehari-hari.
Bagaimana aplikasi dapat membantu individu mengatasi stres atau kecemasan sehari-hari dengan cara yang efektif?	Membuat aplikasi yang memiliki fitur untuk menguraikan stres yang tidak bisa dilampiaskan secara langsung, serta fitur-fitur lain yang menghasilkan warna dan memberikan konten relaksasi seperti musik atau suara alam serta tips harian untuk membantu pengguna mengelola stres dan kecemasan sehari-hari

3.3. Ideate

3.3.1. User Flow



Gambar 5. User Flow

3.3.2. Moodboard

a. Warna

Kode warna #B983FF dapat digunakan untuk memberikan sentuhan feminin atau lembut, sementara #666666 cocok untuk menonjolkan teks atau elemen desain dengan kontras yang kuat. Warna #E7E8EA memberikan kesan bersih dan modern, sementara #FFFFFF cocok digunakan untuk menciptakan kesan minimalis atau kontras yang kuat dengan warna lain dalam desain visual. Dengan memadukan kode warna ini secara bijak, Anda dapat menciptakan desain visual yang menarik dan sesuai dengan kebutuhan komunikasi visual yang ingin Anda sampaikan.



Gambar 6. User Flow

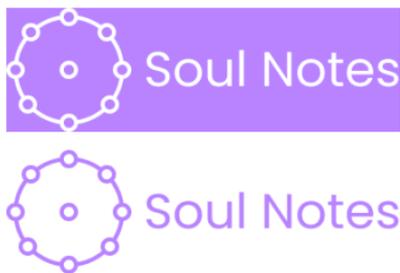
b. Tipografi

Jenis font Poppins, yang memiliki karakteristik bersih, modern, dan mudah dibaca, sangat cocok digunakan dalam desain UI/UX. Kelegaan dan keluwesan font ini membuatnya menjadi pilihan yang solid untuk tampilan antarmuka pengguna yang segar dan mudah dibaca, sehingga sering digunakan dalam desain aplikasi dan situs web. Selain itu, kemampuannya mempertahankan kejelasan pada berbagai ukuran teks membuat Poppins menjadi pilihan yang populer dalam desain web dan identitas merek yang bersih dan modern.

c. Logo

Logo "Soul Notes" untuk aplikasi jurnal pribadi mobile memiliki makna yang terkait dengan konsep refleksi diri dan introspeksi. Bentuk lingkaran yang bulat mewakili kesatuan dan

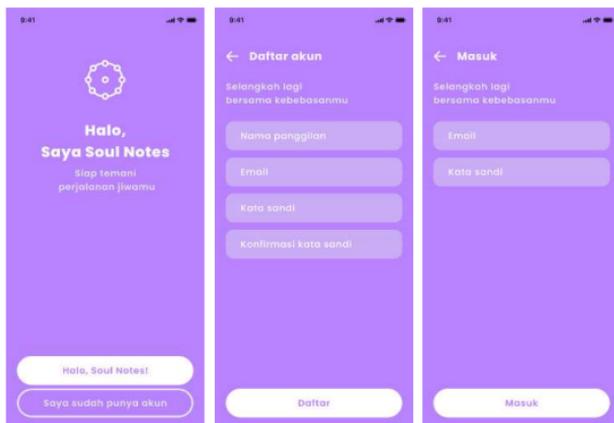
keseluruhan diri, sedangkan delapan lingkaran kecil di sekitar lingkaran mewakili berbagai aspek kehidupan dalam yang direkam dan dipahami melalui proses jurnal. Delapan lingkaran mewakili delapan emosi dasar, seperti kegembiraan, kecewaan, marah, takut, kaget, antusias, percaya, dan tidak suka. Interpretasi ini menekankan aspek emosional jurnal, yang penting untuk memahami dan mengelola perasaan. Selain itu, delapan lingkaran mewakili berbagai tahapan kehidupan, seperti masa kanak-kanak, remaja, dewasa, dan tua. Interpretasi ini menekankan ide bahwa jurnal adalah proses seumur hidup, memungkinkan individu untuk memahami dan tumbuh melalui berbagai tahapan kehidupan. Logo "Soul Notes" efektif mewakili esensi aplikasi, yaitu memberikan platform untuk individu memahami, memantau, dan tumbuh melalui proses jurnal. Desain logo yang sederhana dan elegan membuatnya mudah diingat dan diidentifikasi, membuatnya sebagai alat branding yang efektif untuk aplikasi.



Gambar 7. Logo Soul Notes

3.4. Prototype

3.4.1. Halaman Daftar Akun dan Masuk



Gambar 8. Halaman Daftar Akun dan Masuk

3.4.2. Halaman Utama



Gambar 9. Halaman Utama

3.5. Test

Tabel 4. Hasil Test SEQ (Single Ease Question)

Skor	Skor Responden					Rata-Rata Skor
	R1	R2	R3	R4	R5	
Q1	7	7	7	7	6	6.8
Q2	7	6	7	6	6	6.4
Q3	6	7	6	7	5	6.2
Q4	5	6	7	7	6	6.2
Q5	7	7	7	7	7	7
Hasil (Rata-Rata)						6.52

4. Kesimpulan

Berdasarkan hasil pengujian yang telah dilakukan dalam perancangan desain antarmuka aplikasi Soul Notes dengan metode design thinking, dapat disimpulkan bahwa aplikasi Soul Notes dapat digunakan dengan mudah oleh responden. Hasil pengujian menggunakan SEQ (Single Ease Question) yang diujikan kepada 5 responden menunjukkan rata-rata skor 6.52 dari 7, yang menunjukkan bahwa aplikasi Soul Notes memenuhi kebutuhan dan preferensi responden dalam mengelola kesejahteraan mental dan emosi.

Daftar Pustaka

- [1] Gloriabarus, "Hasil Survei I-NAMHS: Satu dari Tiga Remaja Indonesia Memiliki Masalah Kesehatan Mental." Universitas Gadjah Mada, Oct. 24, 2022.

- <https://uqm.ac.id/id/berita/23086-hasil-survei-i-namhs-satu-dari-tiga-remaja-indonesia-memiliki-masalah-kesehatan-mental> (accessed May 07, 2024)
- [2] M. Cristy P Dame, "Manfaat Journaling bagi Kesehatan Mental yang Sayang untuk Dilewatkan," Alodokter, Apr. 29, 2022. <https://www.alodokter.com/manfaat-journaling-bagi-kesehatan-mental-yang-sayang-untuk-dilewatkan> (accessed May 06, 2024).
- [3] M. Multazam, I. V. Papatungan, and B. Suranto, "Perancangan User Interface dan User Experience pada Placeplus menggunakan pendekatan User Centered Design," Automata, vol. 1, no. 2, pp. 234–241, Jun. 2020.
- [4] A. Swarnadwitya, "Design Thinking: Pengertian, Tahapan dan Contoh Penerapannya.," School of Information Systems, Mar. 17, 2020. <https://sis.binus.ac.id/2020/03/17/design-thinking-pengertian-tahapan-dan-contoh-penerapannya/> (accessed May 10, 2024).
- [5] Salmaa, "Studi Literatur: Pengertian, Ciri, Teknik Pengumpulan Datanya," Penerbit Deepublish, Mar. 17, 2023. <https://penerbitdeepublish.com/studi-literatur/> (accessed May 10, 2024).

Analisis Prediktif Bitcoin dengan Metode SVM serta Pembobotan TIF-IDF Berbasis Data Narrative

Danendra Darmawansyah^{a1}, I Gusti Agung Gede Arya Kadyanan^{a2}

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹danendracool08@gmail.com
²gungde@unud.ac.id

Abstract

The cryptocurrency market has experienced significant volatility in recent years, making it challenging for investors to make informed decisions. This study aims to develop a predictive model for cryptocurrency price increases using TF-IDF (Term Frequency-Inverse Document Frequency) and SVM (Support Vector Machine) based on narrative data. Narrative data, such as news articles and social media posts, can provide valuable insights into investor sentiment and market trends. The proposed model extracts relevant features from narrative data using TF-IDF and employs SVM to classify cryptocurrency price movements into positive, negative, or neutral categories. Experimental results demonstrate the effectiveness of the proposed model in predicting cryptocurrency price increases, with an accuracy of over 70%. The findings suggest that narrative data can be a valuable source of information for cryptocurrency price prediction and that TF-IDF and SVM are effective methods for analyzing narrative data.

Keywords: Cryptocurrency, Price Prediction, TF-IDF, SVM, Narrative Data

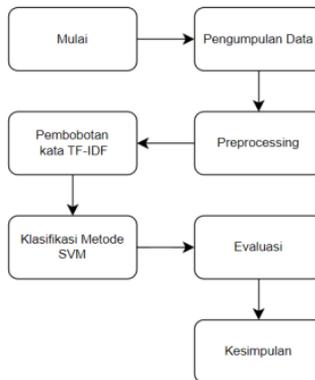
1. Pendahuluan

Fiat merupakan mata uang modern yang saat ini digunakan oleh seluruh dunia, mata uang ini dikeluarkan secara resmi oleh setiap pemerintah dari masing-masing negara. Mata uang fiat berfungsi sebagai alat tukar resmi yang menangani peran perekonomian suatu negara dari unit moneter seperti menyimpan nilai, menyediakan akun numerik, dan memfasilitasi pertukaran. Namun, mata uang fiat tidak terkait dengan Cadangan fisik seperti emas dan perak, sehingga mata uang tersebut berisiko kehilangan nilai akibat inflasi atau menjadi tidak ada harganya jika terjadi hiperinflasi[1]. Maka dari itu, Bitcoin hadir sebagai solusi mata uang modern yang memiliki fungsi digital seperti mata uang fiat namun juga dapat sebagai perlindungan terhadap inflasi yang bisa terjadi pada mata uang fiat. Hal ini terjadi karena token kripto memiliki jumlah yang terbatas, berbeda dengan uang fiat yang pasokannya tidak terbatas. Dunia crypto mempunyai salah satu keunikan yaitu sistemnya menganut desentralisasi yang memungkinkan setiap server saling terkoneksi dan mempunyai peranan yang sama. Crypto adalah mata uang virtual yang keamanannya dijamin dengan kriptografi. Kriptografi membuat uang kripto tidak mungkin dipalsukan atau dibelanjakan secara ganda. Jadi, meskipun digunakan secara virtual, tidak mungkin ada pemalsuan yang merugikan penggunaannya. Cryptocurrency didukung oleh teknologi bernama blockchain. Teknologi inilah yang menjamin keamanan transaksi secara online meskipun tanpa menggunakan campur tangan pihak ketiga. Uang kripto dilindungi berbagai algoritma dan enkripsi dan kriptografi yang mengacu pada teknologi blockchain [2]. Pasar mata uang kripto telah mengalami volatilitas yang signifikan dalam beberapa tahun terakhir, sehingga sulit bagi investor untuk membuat keputusan yang tepat[3]. Hal ini disebabkan oleh berbagai faktor, seperti spekulasi, regulasi, dan sentimen investor. Sentimen investor, khususnya, memainkan peran penting dalam menentukan pergerakan harga mata uang kripto. Sentimen positif dapat mendorong harga naik, sedangkan sentimen negatif dapat menyebabkan harga turun[4].

Penelitian ini bertujuan untuk mengembangkan model prediksi kenaikan harga mata uang kripto menggunakan metode SVM (Support Vector Machine) dengan pembobotan TF-IDF (Term Frequency-Inverse Document Frequency) berdasarkan data naratif. Data naratif, seperti artikel berita dan posting media sosial, dapat memberikan wawasan berharga tentang sentimen investor dan tren pasar. Model yang diusulkan mengekstrak fitur yang relevan dari data naratif menggunakan TF-IDF dan menggunakan SVM untuk mengklasifikasikan pergerakan harga mata uang kripto ke dalam kategori positif atau negatif.

2. Metode Penelitian

Penelitian akan melalui beberapa tahapan yaitu mulai dari pengumpulan data, preprocessing, pembobotan kata TF-IDF, klasifikasi metode SVM, evaluasi, dan langkah terakhir adalah kesimpulan dari penelitian ini. Untuk gambaran dari tahapan penelitian dapat dilihat pada gambar 1 dibawah.



Gambar 1. Tahapan Penelitian

2.1. Pengumpulan Data

Proses pengambilan data melibatkan pengumpulan data pendukung, yang diambil selama proses tersebut. Beberapa cara yang dilakukan adalah melalui studi literatur dan mempelajari beberapa referensi yang berkaitan dengan penelitian. Dataset diambil dengan proses crawling melalui Google Colab dengan menggunakan Twitter Harvest. Dataset dalam bentuk format file csv. Untuk penjelasan lebih lanjut dapat dilihat pada table 1 dibawah.

conversation_id_str	created_at	favorite_count	full_text	id_str	image_url
1780509508629897597	Wed Apr 17 08:12:01 +0000 2024	0	\$BTC LONG market order SL 59309 TP1(10%) 6923...	1780509508629897597	NaN
1780508646490980489	Wed Apr 17 08:08:36 +0000 2024	0	Curiga banget Mei 2024 bakal berdarah-darah ma...	1780508646490980489	https://pbs.twimg.com/tweet_video_thumb/GLWJKq...
1780246978489917951	Wed Apr 17 08:03:03 +0000 2024	0	@yunepto kalo kejsian kyk blackswan tetap pal...	1780507249624142240	NaN
1780461881150304721	Wed Apr 17 06:58:47 +0000 2024	2	@EverythingAjay Ajay bhai yhe khal Crypto Mai...	1780491076660523091	NaN
1780465144666915147	Wed Apr 17 05:15:44 +0000 2024	0	Di bawah lampu kereta bawah tanah yang bertram...	1780465144666915147	https://pbs.twimg.com/media/GLV7m_tacAA_l9t.jpg

Gambar 2. Dataset yang Digunakan

Data yang digunakan diatas berjumlah 1001 baris dan 20 kolom data, didalamnya terdapat keyword untuk memperoleh hasil yang dibutuhkan pada penelitian ini. Dataset yang digunakan adalah data yang diperoleh pada tahun 2023.

2.2. Metode

2.2.1. Analisis Sentimen

Analisis sentimen bertujuan untuk mengetahui, mengekstrak, dan menjalankan informasi tekstual secara otomatis untuk menemui keterangan sentimental dalam mengungkapkan pemikiran. Sentiment analysis digunakan untuk melihat pendapat atau kesamaan opini terhadap sebuah persoalan[5].

2.2.2. Text Mining

Text mining merupakan proses ekstraksi berita dari data asal yang belum terstruktur. Data yang belum terstruktur akan diolah memakai Teknik serta metode tertentu membuat berita yang bermanfaat untuk pengguna. Text mining adalah Teknik yang dipergunakan untuk menangani persoalan classification, clustering, information extraction, dan information retrieval [6].

2.2.3. Preprocessing

Preprocessing adalah proses untuk menormalkan istilah. Hal ini dilakukan agar data yang diterima dengan baik. Mengumpulkan data opini dari media sosial Twitter tidak boleh identik dengan kata baku, kata yang tidak ada dalam kamus, atau bahasa daerah yang digunakan atau dihilangkan. Untuk mengembalikan sejumlah teks ke teks alami dengan mengeliminasi ekspresi atipikal agar dapat meminimalkan noise pada tahap selanjutnya, diperlukan pre-processing atau normalisasi untuk mengatasi hal ini[7].

2.2.4. Term Frequency-Inverse Document Frequency

Salah satu proses pembobotan setiap kata agar bisa mengoptimalkan kemampuan analisis sentiment pada proses text mining. Term Frequency (tf(w,d)) dianggap memiliki proporsi kepentingan sesuai total kemunculannya dalam teks atau dokumen. Inverse Document Frequency (IDF) merupakan metode pembobotan token yang berfungsi untuk memonitor kemunculan token dalam himpunan teks. Berikut merupakan rumus dalam mencari TF-IDF:

$$tf(t, d) = \frac{f_{t,d}}{\sum_{t' \in d} f_{t',d}} \quad (1)$$

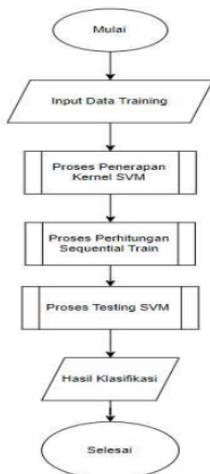
$$TF = \frac{\text{Number of times the term appears in the document}}{\text{Total number of terms in the document}}$$

$$idf(t, D) = \log \frac{N}{|\{d \in D : t \in d\}|} \quad (2)$$

$$tfidf(t, d, D) = tf(t, d) \cdot idf(t, D) \quad (3)$$

2.2.5. Support Vector Machine

SVM adalah Teknik untuk memprediksi klasifikasi dan regresi. SVM merupakan algoritma pembelajaran mesin yang bekerja berdasarkan prinsip Structural Risk Minimalization (SRM) dengan tujuan untuk menemukan hyperlane optimal yang memisahkan dua kelas dalam ruang input. Untuk alur penggunaan algoritma SVM dapat dilihat pada gambar 3 dibawah ini.



Gambar 3. Alur Algoritma SVM

2.2.6. Evaluasi

Evaluasi bertujuan untuk menilai hasil uji coba system yang dibuat apakah telah sesuai dengan hasil yang diinginkan. Pemilihan model dan pencarian parameter terbaik menggunakan Grid Search, yang bertujuan untuk mengoptimalkan kinerja model SVM (Support Vector Machine) dengan TF-IDF sebagai fitur representasi. Proses ini terdiri dari beberapa langkah yaitu mendefinisikan parameter grid, melakukan pencarian parameter terbaik dengan k-fold cross validation, serta melatih dan mengevaluasi model dengan parameter terbaik yang ditemukan.

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+FP+FN+TN)} \quad (4)$$

$$\text{Precision (P)} = \frac{TP}{(TP+FP)} \quad (5)$$

$$\text{Recall (R)} = \frac{TP}{(TP+FN)} \tag{6}$$

$$F - 1 \text{ Score} = 2 \times \frac{(R \times P)}{(R+P)} \tag{7}$$

Keterangan:

- TP : True Positives
- TN : True Negatives
- FN : False Negatives
- FP : False Positives

3. Hasil dan Pembahasan

Penelitian ini menggunakan metode SVM dengan pembobotan TF-IDF untuk menganalisa tren harga pada token Bitcoin berbasis data naratif yang diambil melalui proses crawling menggunakan Twitter Harvest. Dataset yang digunakan dibagi menjadi 80% data training dan 20% data testing. Dataset yang diperoleh telah melewati preprocessing agar data dapat dibaca sebagai data positif atau negatif. Setelah itu dilakukan K-fold cross validation untuk menilai performa model dari data yang diberikan.

3.1. Preprocessing

3.1.1. Labeling

Data yang telah diperoleh melalui crawling akan dilanjutkan dengan proses labeling yang bertujuan untuk dapat menentukan sentimen terhadap token Bitcoin.

in_reply_to_screen_name	lang	location	quote_count	reply_count	retweet_count	tweet_url	user_id_str	username
1ETQWJpg	NaN	in indonesia	0	0	0	https://x.com/juragantrading/status/1791080173779165622	148161125947293962	juragantrading
blockvestglobal	in	Jakarta	0	2	0	https://x.com/blockvestglobal/status/1792080832303779177	111471040401228608	blockvestglobal
NaN	in	Indonesia	2	53	26	https://x.com/anggaandinata/status/1792727449677246594	134943969	anggaandinata
iGAAbOwApg	NaN	in Indonesia	0	2	4	https://x.com/PintuID/status/178904036836835797	763262916817288728	PintuID
hidatufkri	in	Tangerang, Indonesia	0	1	0	https://x.com/hidatufkri/status/1790372432961851725	629625730	hidatufkri

Gambar 4. Sebelum Labeling

tweet_url	user_id_str	username	cleaning	hapusEmoji	replaceTOM	caseFolding	label	sentiment	score
status/1700414611...	1484413667708676102	cryptodailyid	BREAKING naik data\$	BREAKING naik data\$	BREAKING naik data\$	BREAKING naik data\$	POSITIVE	[[label: 'POSITIVE', 'score': 0.9947491288181]]	0.994749
status/1730949337...	1702269721767190528	connedion_id	mempertahankan perdagangan sideways dengan anti...	NEGATIVE	[[label: 'NEGATIVE', 'score': 0.99723188181]]	0.997231			
status/17917040919...	1107742918952083462	bitcoin_brown	lamesyoyuareno Muh Wakanda	lamesyoyuareno Muh Wakanda	lamesyoyuareno Muh Wakanda	lamesyoyuareno Muh Wakanda	NEGATIVE	[[label: 'NEGATIVE', 'score': 0.976173818111]]	0.976174
79221289648119842	439326281	ukikahr	Bitcoin adalah penipuan yang pada akhirnya aka...	NEGATIVE	[[label: 'NEGATIVE', 'score': 0.999657985548]]	0.999658			
status/17914657650...	1133128842209628096	brahm632931	longuppella US\$y e z UwaGYHMLUV ofNGKGLUQ vb...	NEGATIVE	[[label: 'NEGATIVE', 'score': 0.995218]]	0.995218			
stat/1792699718048...	238091066	investordID	Harga Bitcoin Merak ke Level US\$ Ribu ETH ...	Harga Bitcoin Merak ke Level US\$ Ribu ETH ...	Harga Bitcoin Merak ke Level US\$ Ribu ETH ...	Harga Bitcoin Merak ke Level US\$ Ribu ETH ...	NEGATIVE	[[label: 'NEGATIVE', 'score': 0.9931919]]	0.993191
66status/179119805...	1390889451099197441	Buatom92248936	Senator Cynthia Lumina menyerukan persetujuan ...	NEGATIVE	[[label: 'NEGATIVE', 'score': 0.881730556488]]	0.881731			
status/179262947926...	1496570216847241217	magalata_xyz	Bitcoin mencapai US\$ pada Mei	NEGATIVE	[[label: 'NEGATIVE', 'score': 0.990741]]	0.990741			

Gambar 5. Sesudah Labeling

Proses ini akan mengeluarkan hasil sentiment dari dataset yang sudah diperoleh. Proses labeling menggunakan Transformer Python untuk mendapatkan hasil sentiment positif atau negatif.

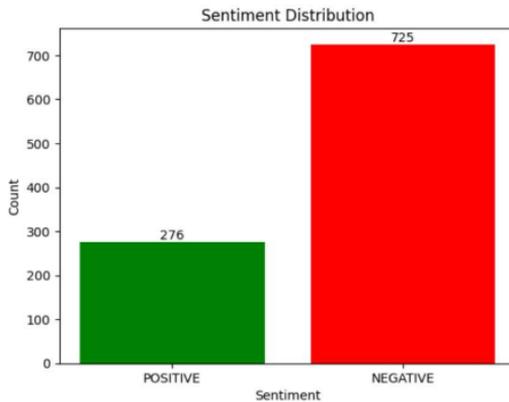
3.1.2. Cleansing pipeline

Pada proses ini, dataset yang digunakan akan melalui proses pembersihan kata agar tersisa kata penting yang dapat dianalisa sentimennya. Bagian ini meliputi cleaning karakter, emoji, Stopword, caseFolding, Tokenizing, formalisasi, dan stemming.

cleaning	hapusEmoji	replaceTOM	caseFolding	label	sentiment	score	tokenizing	formalisasi
Dulu Bitcoin masih rb rupiah orang masih ga m...	Dulu Bitcoin masih rb rupiah orang masih ga m...	Dulu Bitcoin masih rb rupiah orang masih ga m...	dulu bitcoin masih rb rupiah orang masih ga m...	NEGATIVE	[[label: 'NEGATIVE', 'score': 0.998764753341]]	0.998765	[dulu, bitcoin, masih, rb, rupiah, orang, masih, ga, m...]	[dulu, bitcoin, masih, rb, rupiah, orang, masi...]
Market Overview Bitcoin hari ini diperdagangka...	NEGATIVE	[[label: 'NEGATIVE', 'score': 0.988911854287...]]	0.988912	[market, overview, bitcoin, hari, ini, diperda...]	[market, overview, bitcoin, hari, ini, diperda...]			
klo negara eropa mata uangnya ganti euro buat...	klo negara eropa mata uangnya ganti euro buat...	klo negara eropa mata uangnya ganti euro buat...	klo negara eropa mata uangnya ganti euro buat...	NEGATIVE	[[label: 'NEGATIVE', 'score': 0.979081392288...]]	0.979081	[klo, negara, eropa, mata, uangnya, ganti, eur...]	[klo, negara, eropa, mata, uangnya, ganti, eur...]
ylad bRey dari namanya aja udh beda bitcoin...	ylad bRey dari namanya aja udh beda bitcoin...	ylad bRey dari namanya aja udh beda bitcoin...	ylad brey dari namanya aja udh beda bitcoin...	POSITIVE	[[label: 'POSITIVE', 'score': 0.966544508934...]]	0.966545	[ylad, brey, dari, namanya, aja, udh, beda, bi...]	[ylad, brey, dari, namanya, aja, udh, beda, b...]

Gambar 6. Hasil Cleansing Pipeline

Gambar diatas merupakan hasil dari proses cleansing pipeline pada dataset. Diperoleh hasil teks setelah cleaning, clear emoji, replaceTOM, caseFolding, tokenizing, dan formalisasi.



Gambar 8. Grafik Sentimen

Gambar diatas merupakan hasil visualisasi data dalam bentuk grafik yang berfokus pada sentimen untuk dibandingkan agar dapat dianalisa dengan baik terkait tren token Bitcoin. Terlihat bahwa tren negatif lebih banyak daripada tren positif yang berarti tren toko token Bitcoin belum menunjukkan hasil positif pada investor crypto pada tahun 2023 sesuai dengan dataset yang diambil.

3.4. Evaluasi dan Perbandingan Data dengan Market

3.4.1. Hasil Evaluasi

Berikut merupakan hasil evaluasi yang menghasilkan nilai akhir seperti precision, recall, f1-score, dan akurasi k-fold cross validation. Pada penelitian ini hasil yang didapatkan dari SVM dengan TF-IDF dapat dilihat pada gambar 8 dibawah ini.

```
Parameter terbaik: {'C': 1, 'gamma': 'scale', 'kernel': 'linear'}
Akurasi terbaik: 0.7762499999999999

Classification Report for TF-IDF with Best Parameters:
              precision    recall  f1-score   support

   NEGATIVE      0.77      0.92      0.83        143
   POSITIVE      0.60      0.31      0.41         58

 accuracy              0.74        201
 macro avg              0.68      0.61      0.62        201
 weighted avg          0.72      0.74      0.71        201
```

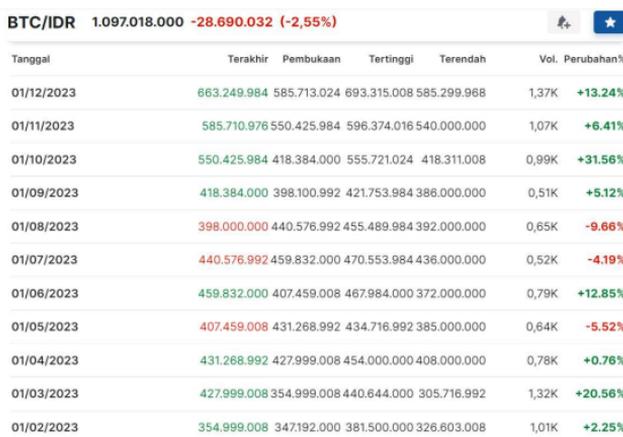
Gambar 9. Hasil SVM dengan TF-IDF

Dari hasil SVM dengan TF-IDF sehingga mendapatkan evaluasi berupa nilai C terbaik ada pada nilai 1, lalu gamma yang digunakan yaitu pada scale, dan kernel yang digunakan yaitu linear. Sehingga dari testing yang dilakukan didapatkan nilai precision 77%, recall 92%, dan f1-score

83% pada sentiment negatif. Sedangkan pada sentiment positif menghasilkan nilai precision 60%, recall 31%, dan f1-score 41%. Sehingga hasil akurasi akhir yang didapatkan yaitu sebesar 74%.

3.4.2. Perbandingan Data dengan Market

Setelah mendapatkan hasil akhir dari analisa yang dilakukan dengan menggunakan metode SVM serta pembobotan TF-IDF diperoleh hasil negatif yang melambung tinggi jika dibandingkan dengan hasil positif. Hal ini berbanding lurus dengan market crypto Bitcoin yang terjadi pada tahun 2023 sesuai dengan data yang diperoleh. Bitcoin mengalami penurunan yang signifikan terjadi pada tahun 2023 hingga mengalami 50% penurunan harga jual. Data market crypto Bitcoin dapat dilihat pada gambar 9 dibawah ini.



BTC/IDR		1.097.018.000		-28.690.032 (-2,55%)			
Tanggal	Terakhir	Pembukaan	Tertinggi	Terendah	Vol.	Perubahan%	
01/12/2023	663.249.984	585.713.024	693.315.008	585.299.968	1,37K	+13.24%	
01/11/2023	585.710.976	550.425.984	596.374.016	540.000.000	1,07K	+6.41%	
01/10/2023	550.425.984	418.384.000	555.721.024	418.311.008	0,99K	+31.56%	
01/09/2023	418.384.000	398.100.992	421.753.984	386.000.000	0,51K	+5.12%	
01/08/2023	398.000.000	440.576.992	455.489.984	392.000.000	0,65K	-9.66%	
01/07/2023	440.576.992	459.832.000	470.553.984	436.000.000	0,52K	-4.19%	
01/06/2023	459.832.000	407.459.008	467.984.000	372.000.000	0,79K	+12.85%	
01/05/2023	407.459.008	431.268.992	434.716.992	385.000.000	0,64K	-5.52%	
01/04/2023	431.268.992	427.999.008	454.000.000	408.000.000	0,78K	+0.76%	
01/03/2023	427.999.008	354.999.008	440.644.000	305.716.992	1,32K	+20.56%	
01/02/2023	354.999.008	347.192.000	381.500.000	326.603.008	1,01K	+2.25%	

Gambar 10. Harga Jual Bitcoin Tahun 2023

Sesuai dengan gambar data diatas bahwa pada tahun 2023 menunjukkan harga jual Bitcoin sedang dalam kondisi Bearish Market. Bearish Market merupakan kondisi pasar melemah diakibatkan banyak investor yang menjual aset Bitcoin mereka sehingga harga jual mengalami penurunan.

4. Kesimpulan

Berdasarkan hasil dari penelitian yang telah dilaksanakan, metode SVM dengan pembobotan TF-IDF merupakan metode yang baik untuk digunakan dalam analisis prediktif tren market Bitcoin berbasis data naratif. Hal ini dibuktikan dengan hasil perbandingan data antara hasil pengujian dengan data real yang diperoleh. Data naratif merupakan salah satu bentuk analisa market crypto yang sering digunakan, sehingga metode ini sesuai dengan analisa data naratif yang dijalankan. Metode ini menghasilkan nilai akurasi 74% yang berarti sudah baik untuk digunakan dalam analisa. Pada penelitian selanjutnya diharapkan dapat meningkatkan akurasi dari data yang diperoleh sehingga mendapatkan analisa yang lebih baik terhadap tren market crypto Bitcoin.

Daftar Pustaka

- [1] J. Chen, "Uang Fiat: Apa Artinya, Cara Kerja, Contoh, Kelebihan & Kekurangan", Investopedia, 28 Februari 2024, [Online]. Available: <https://www.investopedia.com/terms/f/fiatmoney.asp> [Diakses: 11 Mei 2024].

- [2] A. L. Hidayah, "Yuk, Berkenalan dengan Kripto!", Kementerian Keuangan Republik Indonesia, 11 April 2023, [Online]. Available: <https://www.dikn.kemenkeu.go.id/kpknl-tangerang1/baca-artikel/16059/Yuk-Berkenalan-dengan-Kripto.htm> [Diakses: 11 Mei 2024].
- [3] M. M. Karim, M. H. Ali, L. Yarovaya, M. H. Uddin, S. Hammoudeh, "Return-volatility relationships in cryptocurrency markets: Evidence from asymmetric quantiles and non-linear ARDL approach", *International Review of Financial Analysis*, vol. 90, November 2023.
- [4] E. Akyildirim, A. F. Aysan, O. Cepni, O. Serbest, "Sentiment matters: the effect of news-media on spillovers among cryptocurrency returns", *The European Journal of Finance*, Juli 2022.
- [5] M. W. Berry and J. Kogan, "Text Mining Application and Theory", Wiley; 1st Edition, 2010
- [6] R. Feldman and J. Sanger, "The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data", Cambridge University Press; 1st edition, 2006.
- [7] F. Tala, "A Study of Stemming Effects on Information Retrieval in Bahasa Indonesia", Institute for Logic, Language and Computation: Universiteit van Amsterdam, 2003.

Identifikasi Mekar dan Kuncupnya Bunga Sedap Malam Menggunakan Convolutional Neural Network

Kadek Bakti Pramanayoga St^{a1}, I Gusti Agung Gede Arya Kadyanan^{a2}

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹kdkbaktip18sep04@gmail.com
²gungde@unud.ac.id

Abstract

*The utilization of technology can aid humans across various sectors, including agriculture. This study harnesses one such technology to identify a particular agricultural commodity, tuberose flowers (*Polianthes tuberosa*), based on their blooming using Convolutional Neural Network (CNN). The CNN method can assist farmers in harvesting tuberose flowers by distinguishing between bloomed and budding flowers. In this research, a dataset comprised of 600 primary data points captured via smartphones is utilized, divided into 420 training sets and 180 testing sets. Three scenarios are tested, involving training epochs of 10, 15, and 20. The testing results indicate that the first scenario achieves an accuracy score of approximately 82.44%, falling below the 85% threshold. Meanwhile, the second and third scenarios achieve accuracies of approximately 91.20% and 92%, respectively.*

Keywords: *Classify, Polianthes tuberosa, Deep Learning, Convolutional Neural Network, Bloom Level, Accuracy*

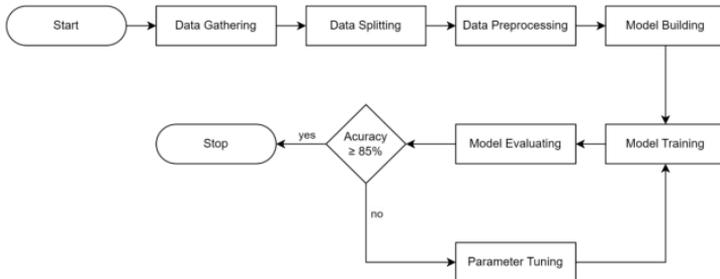
1. Pendahuluan

Indonesia merupakan negara agraris dengan sektor pertanian sebagai salah satu sektor yang berperan penting dalam kehidupan masyarakat. Dalam sektor pertanian, salah satu komoditas tanaman yang dimanfaatkan masyarakat Indonesia adalah komoditas tanaman bunga potong seperti bunga sedap malam (*Polianthes tuberosa*). Bunga sedap malam merupakan bunga yang memiliki ciri khas berwarna putih, memiliki aroma harum yang khas, dan memiliki tangkai yang panjang [1]. Bunga sedap malam dimanfaatkan sebagai hiasan sekaligus pengharum ruangan di hotel, rumah sakit, restoran, vila, dan rumah tangga [1]. Produksi bunga sedap malam di Indonesia khususnya di Bali sejak tahun 2018 mengalami peningkatan yang signifikan. Pada tahun 2022, produksi bunga sedap malam di Bali adalah sebanyak 120.400 tangkai dan mengalami peningkatan sebanyak 681,19% dari tahun 2018 [2]. Dari data tersebut, diperlukan teknologi yang dapat membantu petani dalam proses pemanenan bunga sedap malam. Salah satu teknologi yang dapat dimanfaatkan untuk membantu proses pemanenan bunga sedap malam yaitu deep learning. Dengan deep learning, komputer dapat mengetahui bunga yang siap panen dengan mengidentifikasi bunga yang sudah mekar dan masih kuncup. Di antara metode deep learning yang ada, Convolutional Neural Network (CNN) adalah metode yang memiliki hasil yang baik dalam mendeteksi data citra [3]. Karena hal tersebut, metode ini sangat cocok digunakan untuk mendeteksi kemekaran bunga. Metode CNN dapat dilatih untuk mengenali perbedaan bunga yang sudah mekar dan bunga yang masih kuncup, serta dapat menghindari adanya human error saat pemanenan bunga. Penelitian terdahulu menunjukkan potensi CNN dalam mengidentifikasi gambar seperti penelitian oleh A. Arkadia, dkk. pada tahun 2021 yang mencapai akurasi 94,6% pada data testing dalam klasifikasi buah mangga badami berdasarkan tingkat kematangannya [4]. Berdasarkan penelitian terdahulu, penelitian ini bertujuan untuk mengembangkan model yang mampu mengidentifikasi bunga sedap malam antara bunga yang mekar dan kuncup menggunakan metode CNN. Melalui penelitian ini, diharapkan dapat membantu petani bunga sedap malam dalam proses pemanenan.

2. Metode Penelitian

Metode deep learning yang digunakan dalam penelitian ini adalah metode Convolutional Neural Network (CNN) untuk mengidentifikasi bunga yang sudah mekar dan bunga yang masih kuncup dengan arsitektur CNN yang disusun oleh penulis.

2.1. Alur Penelitian

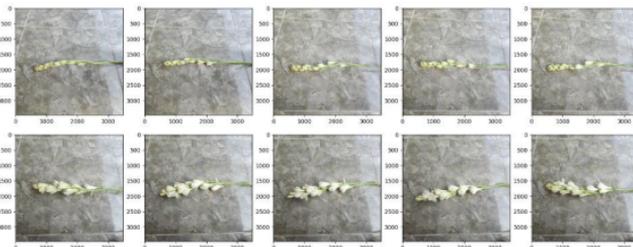


Gambar 1. Alur Penelitian

Pelaksanaan penelitian dimulai dengan proses pengumpulan data yaitu data bunga mekar dan bunga kuncup. Selanjutnya data melalui proses pemisahan menjadi data training dan data testing. Kemudian data melalui proses preprocessing untuk menormalisasi data. Selanjutnya membuat arsitektur CNN yang akan dilatih. Setelah model dilatih, dilanjutkan dengan evaluasi model untuk mengetahui performa model yang dilatih. Jika performa model buruk, maka parameter model akan disesuaikan lalu model kembali dilatih hingga mendapatkan performa yang baik pada proses evaluasi sehingga model siap digunakan.

2.2. Data Gathering

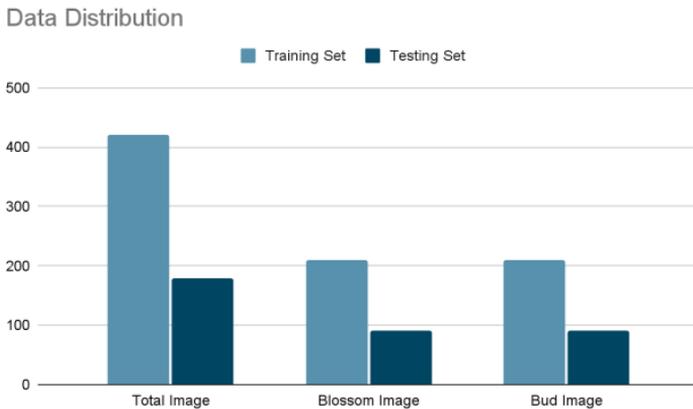
Data yang digunakan dalam penelitian ini adalah data primer yang diambil menggunakan kamera handphone. Bunga yang digunakan berasal dari kelompok petani bunga sedap malam di Desa Tunjuk, Kecamatan Tabanan, Kabupaten Tabanan, Bali. Dataset berisi data citra bunga sedap malam yang dibagi menjadi kelas bunga mekar dan bunga kuncup. Masing-masing kelas terdapat 300 data yang berbentuk citra RGB. Jadi total data pada dataset sebanyak 600 data. Data citra yang didapatkan dari proses data gathering memiliki dimensi 3472 x 3472 pixel. Berikut merupakan contoh dari data yang digunakan dalam penelitian.



Gambar 2. Contoh Data Citra Bunga Sedap Malam Kuncup dan Mekar

2.3. Data Splitting

Dataset citra bunga sedap malam dipisah menjadi dua bagian yaitu training set dan testing set. Perbandingan antara training set dan testing set yang digunakan dalam data splitting ini yaitu 7:3. Training set memiliki 420 data citra sedangkan testing set memiliki 180 data citra. Berikut merupakan grafik yang menunjukkan hasil distribusi data saat proses data splitting.



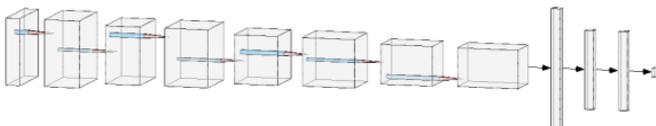
Gambar 3. Distribusi Data

2.4. Data Preprocessing

Dataset citra bunga sedap malam melalui proses preprocessing yang diawali dengan melakukan normalisasi dengan cara rescaling sebesar $1/255$ sehingga dari rentang nilai awal 0 sampai dengan 255 menjadi antara 0 sampai dengan 1. Kemudian dilakukan resizing sehingga ukuran data citra menjadi 200×200 pixel. Setelah data melewati proses resizing, untuk training set dilakukan proses flipping secara horizontal dan vertikal serta rotating antara -20° sampai dengan 20° .

2.5. Model Building

Convolution Neural Network terdiri dari lapisan-lapisan seperti input layer, hidden layers, dan output layer [5]. Dalam Penelitian ini lapisan tersembunyi terdiri dari convolutional layers, pooling layers, dan fully connected layers. Berikut merupakan arsitektur dari model Convolution Neural Network.



Gambar 4. Arsitektur Convolution Neural Network

Dari arsitektur diatas, berikut merupakan ringkasan model yang akan digunakan untuk proses pelatihan *Convolution Neural Network*.

Tabel 1. Ringkasan Model

Layer (Activation)	Output Shape	Param
Conv2D (ReLu)	(None, 198, 198, 16)	448
MaxPooling2D	(None, 99, 99, 16)	0
Conv2D (ReLu)	(None, 97, 97, 32)	4640
MaxPooling2D	(None, 48, 48, 32)	0
Conv2D (ReLu)	(None, 46, 46, 64)	18496
MaxPooling2D	(None, 23, 23, 64)	0
Conv2D (ReLu)	(None, 21, 21, 128)	73856
MaxPooling2D	(None, 10, 10, 128)	0
Flatten	(None, 12800)	0
Dense (ReLu)	(None, 512)	6554112
Dropout	(None, 512)	0
Dense (Softmax)	(None, 2)	1026

Dapat dilihat pada Tabel 1. terdapat 5 jenis lapisan yang berbeda, yaitu:

a. Convolution Layer

Pada convolution layer akan dilakukan operasi konvolusi antara matriks citra input dengan matriks-matriks filter. Filter-filter tersebut akan digeser ke seluruh permukaan citra sehingga akan mengeluarkan output matriks feature map [6]. Berikut merupakan persamaan untuk mencari ukuran feature map:

$$n_{out} = \left(\frac{n_{in} - k + 2p}{s} \right) + 1 \tag{1}$$

dimana:

- n_{out} : Ukuran matriks feature map
- n_{in} : Ukuran matriks input
- k : Ukuran matriks filter
- p : Ukuran padding
- s : Stride

Berikut merupakan persamaan operasi konvolusi:

$$FM[i]_{j,k} = \left(\sum_m \sum_n N_{[j-m \ k-n]} F_{[m,n]} \right) + bF \tag{2}$$

dimana:

- $FM[i]_{j,k}$: Matriks feature map ke-1
- N : Matriks citra input
- F : Matriks filter konvolusi
- bF : Nilai bias pada filter
- j, k : Posisi pixel pada matriks citra input
- m, n : Posisi pixel pada matriks filter konvolusi

- b. Pooling Layer
Pooling atau subsampling adalah pengurangan ukuran dari suatu matriks. Terdapat dua jenis pooling layer yang sering digunakan yaitu max pooling dan average pooling. Sesuai dengan namanya, perbedaan antara max pooling dan average pooling adalah nilai yang diambil, yaitu max pooling mengambil nilai maksimal sedangkan average pooling mengambil nilai rata-rata [7]. Pada penelitian ini, jenis pooling layer yang digunakan adalah max pooling.
- c. Flatten Layer
Flatten layer merupakan layer yang digunakan untuk mengubah feature map yang diterima dari pooling layer ke dalam format yang dapat dimengerti oleh dense layer [8].
- d. Dense Layer
Dense layer merupakan lapisan yang terdiri dari unit-unit yang terhubung dengan unit pada lapisan sebelumnya. Dense layer dapat dianggap sebagai fully connected layer. Lapisan ini digunakan pada tahap akhir dari sebuah model, yaitu setelah melakukan proses pengolahan data pada lapisan-lapisan sebelumnya[8].
- e. Dropout Layer
Dropout layer merupakan lapisan yang secara acak menetapkan unit masukan ke nilai 0 dengan frekuensi laju di setiap langkah selama proses pelatihan. Hal ini dilakukan untuk menghindari terjadinya overfitting. Input yang tidak diubah menjadi nilai 0 akan ditingkatkan skalanya sebesar $1 / (1 - \text{rate})$ sehingga jumlah input-nya tidak berubah [8].

Selain 5 jenis lapisan yang ada, terdapat 2 jenis Activation Layer yang digunakan yaitu:

- a. ReLu Layer
ReLU (Rectification Linear Unit) merupakan fungsi yang digunakan untuk mengenalkan nonlinieritas dan meningkatkan representasi dari model [7]. Fungsi Aktivasi ReLU yaitu:

$$f(x) = \max(0, x) \quad (3)$$

- b. Softmax Layer
Fungsi aktivasi softmax digunakan untuk mendapatkan output dari model. Fungsi ini menghasilkan suatu nilai yang diinterpretasi sebagai probabilitas yang belum dinormalisasi untuk setiap kelas [7]. Karena penelitian ini mengidentifikasi 2 kelas yang berbeda, maka fungsi aktivasi softmax yang digunakan memiliki 2 neuron. Berikut merupakan persamaan dari fungsi softmax.

$$y_{ijk} = \frac{e^{x_{ijk}}}{\sum_{t=1}^D e^{x_{ijt}}} \quad (4)$$

dimana:

y : Vektor yang memiliki nilai 0 atau 1

x : Vektor yang nilai hasil dari lapisan fully connected terakhir

2.6. Model Training

Proses pelatihan merupakan proses yang penting. Model dilatih menggunakan training set dengan jumlah epoch yaitu:

- a. 10 Epoch
- b. 15 Epoch
- c. 20 Epoch

2.7. Model Evaluation

Evaluasi model diperlukan untuk menilai kinerja model yang telah dibangun. Evaluasi model dalam penelitian ini yaitu menghitung nilai akurasi untuk memberikan gambaran tentang kinerja model dalam mengidentifikasi data. Berikut merupakan rumus dari perhitungan akurasi model.

$$\text{Akurasi} = \frac{\sum \text{Data Uji Benar}}{\text{Total Uji}} \times 100\% \quad (5)$$

3. Hasil dan Diskusi

3.1. Hasil Pelatihan dan Evaluasi Model

Pada model Convolution Neural Network terdapat 3 skenario pengujian, yaitu:

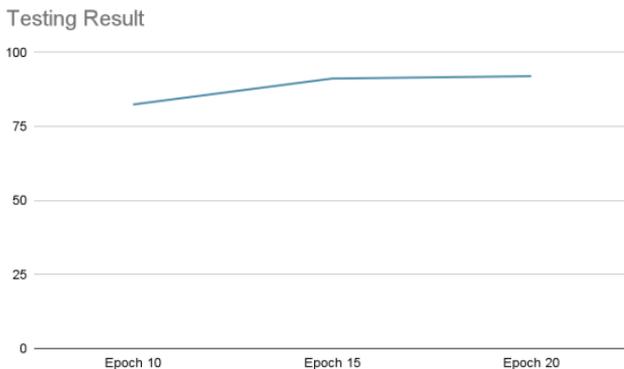
- Pengujian pada pelatihan model dengan 10 epoch
- Pengujian pada pelatihan model dengan 15 epoch
- Pengujian pada pelatihan model dengan 20 epoch

Ketiga skenario ini digunakan pada setiap pengujian 140 testing set. Hasil dari pengujian akan dihitung untuk mendapatkan akurasi model. Berikut merupakan tabel hasil pengujian model.

Tabel 2. Hasil Pengujian

No.	Skenario	Akurasi
1	Pengujian pada pelatihan model dengan 10 epoch	82,44%
2	Pengujian pada pelatihan model dengan 15 epoch	91,20%
3	Pengujian pada pelatihan model dengan 20 epoch	92%

Dari tabel 2. tersebut, berikut merupakan grafik hasil pengujian model.



Gambar 5. Grafik hasil pengujian

Dari hasil pengujian ketiga skenario tersebut, pada pengujian pelatihan model dengan 10 epoch mendapatkan akurasi kurang dari 85% yaitu 82,44%. Sedangkan, pengujian pada pelatihan

model dengan 15 epoch mendapatkan akurasi 91,20% dan pengujian pada pelatihan model dengan 20 epoch mendapatkan akurasi 92%.

4. Kesimpulan

Berdasarkan hasil pengujian pada identifikasi bunga sedap malam berdasarkan mekar kuncupnya bunga menggunakan Convolutional Neural Network (CNN) yang telah dilakukan dan telah diuraikan dalam penelitian ini, maka dapat ditarik kesimpulan bahwa pada pengujian skenario pertama, pengujian pada pelatihan model dengan 10 epoch, mendapatkan nilai akurasi kurang dari 85% yaitu 82,44%. Hal itu menunjukkan pelatihan model dengan 10 epoch memiliki hasil yang kurang baik dibandingkan dengan pengujian pada skenario kedua dan ketiga. Pada pengujian skenario kedua dan ketiga, pengujian pada pelatihan model dengan 15 epoch mendapatkan nilai akurasi lebih dari 85% yaitu 91,20% dan pengujian pada pelatihan model dengan 20 epoch mendapatkan nilai akurasi 92%. Maka dapat disimpulkan pelatihan model dengan 15 epoch dan 20 epoch pada identifikasi bunga sedap malam berdasarkan mekar kuncupnya bunga menggunakan metode Convolutional Neural Network (CNN) dapat bekerja dengan baik.

Daftar Pustaka

- [1] N. W. C. P. A. Putri, B. Ahmadi H., dan K. A. B. Sadyasmara, "Distribusi dan Perbaikan Pasca Panen Bunga Potong Sedap Malam (*Polianthes tuberosa*) dari Petani Desa Tunjuk, Tabanan ke Denpasar," *J. Rekyasa dan Manajemen Agroindustri*, vol. 8, no. 2, pp. 301–309, 2020.
- [2] Badan Pusat Statistik Indonesia, "Produksi Tanaman Hias Menurut Provinsi dan Jenis Tanaman," Badan Pusat Statistik, 2023, [Online]. Available: www.bps.go.id.
- [3] I G. D. Dwijayanti, I P. F. T. Mahendra, I. L. Simarmata, G. A. V. M. Giri, "Klasifikasi Kualitas Buah dengan Menggunakan Convolutional Neural Network (Studi Kasus: Dataset Fresh and Rotten Classification)," *J. Nasional Teknologi Informasi dan Aplikasinya*, vol. 2, no. 2, pp. 429-442, 2024.
- [4] A. Arkadia, S. A. Damayanti, dan D. S. Prasvita, "Klasifikasi Buah Mangga Badami untuk Menentukan Tingkat Kematangan dengan Metode CNN," *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)*, ISBN 978-623-93343-4-5, pp. 158-165, 2021.
- [5] M. R. Alwanda, R. P. K. Ramadhan, dan D. Alamsyah, "Implementasi Metode Convolutional Neural Network Menggunakan Arsitektur LeNet-5 untuk Pengenalan Doodle," *Jurnal Algoritme*, vol. 1, no. 1, pp. 45-56, 2020.
- [6] A. Hibatullah, I. Maliki, "Penerapan Metode Convolutional Neural Network pada Pengenalan Pola Citra Sandi Rumpuk," [Online]. Available: elibrary.unikom.ac.id
- [7] M. H. Romarion, E. Ihsanto, T. M. Kadarina, "Sistem Hitung dan Klasifikasi Objek dengan Metode Convolutional Neural Network," *Jurnal Teknologi Elektro*, vol. 11, no. 2, pp. 108-114, 2020.
- [8] Keras, "Keras Layers API," [Online]. Available: <https://keras.io/api/layers>

Halaman ini sengaja dibiarkan kosong

Penerapan Model Ontologi dalam Perkembangan Game Digital

I Komang Maheza Yudistia¹, I Gusti Ngurah Anom Cahyadi Putra²

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹yudistia.2208561115@student.unud.ac.id
²anom.cp@unud.ac.id

Abstract

This research develops an ontology for interactive digital games using the Methontology methodology, which includes the stages of specification, knowledge acquisition, conceptualization, integration, implementation, and evaluation. The development process involved using the Protégé tool to build and manage the ontology. The result is an ontology consisting of 32 classes, 12 property objects, 0 property data. This ontology was tested using SPARQL queries on the Jena Fuseki web, demonstrating its ability to support more effective classification and knowledge management. The application of this ontology improves data interoperability and the gaming experience and enables the development of more structured and complex applications. This research confirms the benefits of ontologies in enriching data structure and integration in the interactive digital game domain.

Keywords: Digital game ontology, Methontology, Protégé.

1. Pendahuluan

Dalam konteks industri hiburan yang terus berkembang, pengembangan permainan digital interaktif menjadi fokus penting. Karena rumitnya menciptakan pengalaman bermain game yang imersif, peran ontologi semantik menjadi semakin penting. Ontologi menyediakan kerangka struktural untuk mendefinisikan konsep, hubungan, dan batasan dalam domain tertentu. Hal ini memungkinkan pengembang untuk mengembangkan pemahaman yang lebih dalam tentang aspek-aspek utama permainan, tetapi juga memfasilitasi integrasi dan interoperabilitas data yang lebih baik. Metodologi Methontology dikenal karena pendekatannya yang mendetail terhadap pengembangan ontologi, metodologi ini telah menjadi pilihan utama untuk membangun ontologi yang konsisten dan akurat. Langkah-langkah sistematis mulai dari definisi hingga evaluasi membantu memastikan bahwa ontologi yang dihasilkan memenuhi kebutuhan dan tujuan yang ditetapkan. Tujuan dari penelitian ini adalah untuk mengembangkan ontologi yang didedikasikan untuk permainan digital interaktif. Ontologi ini berfokus pada perbaikan struktur dan integrasi data serta memberikan pengalaman bermain game yang lebih terstruktur kepada pengguna, sehingga diharapkan dapat memberikan landasan yang kuat bagi pengembangan aplikasi game yang lebih kompleks dan memuaskan. Penelitian ini diharapkan dapat memberikan bukti nyata bahwa penerapan ontologi pada pengembangan game digital dapat memberikan manfaat nyata, termasuk peningkatan kualitas pengalaman bermain game dan kemampuan pengelolaan data yang lebih efisien.

1.1. Game Digital

Game digital adalah permainan yang dimainkan di perangkat elektronik seperti komputer, konsol game, ponsel pintar, atau tablet. Game-game ini sering kali menampilkan interaksi pemain dengan antarmuka pengguna grafis yang menggambarkan dunia virtual. Terlepas dari genre dan kompleksitasnya, game digital memiliki beberapa elemen umum, seperti tujuan yang jelas, aturan yang jelas, tantangan, dan interaksi pemain. Pengembangan game digital melibatkan beberapa proses kompleks, mulai dari desain konsep, pengembangan grafis dan audio hingga pengkodean

perangkat lunak. Game digital dapat memiliki berbagai genre, termasuk aksi, petualangan, strategi, game olahraga, dan banyak lagi, yang masing-masing menawarkan pengalaman bermain game yang unik. Selain hiburan, game digital juga menjadi platform penting untuk pendidikan, pelatihan, dan kolaborasi. Mereka dapat digunakan untuk menyampaikan pesan pendidikan, mengembangkan keterampilan kognitif dan motorik serta memfasilitasi kerja sosial dan tim.

1.2. Ontologi

Ontologi merupakan kosakata yang terdiri dari kalimat-kalimat yang mendefinisikan konsep, hubungan dan batasan suatu disiplin ilmu tertentu [1]. Hal ini sangat berguna untuk interoperabilitas data karena dapat dilakukan dengan lebih efisien dan efektif. Penggunaan ontologi memiliki beberapa keunggulan, seperti kemampuan mendeskripsikan domain pengetahuan secara eksplisit dengan menyediakan struktur konsep hierarki yang menjelaskan domain dan hubungannya. Penyusunan ontologi yang terstruktur dengan baik, memahami makna dan hubungan antar konsep, serta mengorganisir informasi secara sistematis [4]. Untuk pemahaman yang berbeda mengenai informasi terstruktur dan penggunaan kembali area data, misalnya kita ingin membuat ontologi yang luas, kita dapat mengembangkan ontologi yang sudah ada dan mengintegrasikannya dengan beberapa ontologi lain yang relevan dengan ontologi yang dibangun [2-3].

1.3. SPARQL

SPARQL adalah perintah atau bahasa yang digunakan untuk mengakses query terhadap model data suara format data RDF. SPARQL juga dapat dianggap sebagai bahasa untuk mengakses data tertaut menggunakan titik akhir untuk menciptakan hubungan antara satu informasi. Bahasa SPARQL dianggap mirip dengan bahasa SQL, dengan sintaks yang sama dan tujuan yang berbeda [6].

1.4. Protégé

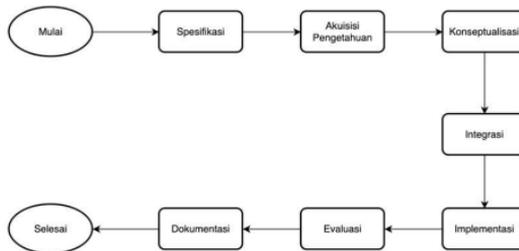
Protégé adalah alat yang dapat membuat domain ontologi dan menanyakannya menggunakan SPARQL [5]. Protégé dikembangkan menggunakan bahasa pemrograman Java dan format penyimpanan seperti OWL, RDF, XML, Turtle, Manchester OWL, JSON-LD, LaTeX dan OBO. Fungsionalitas alat anak didik dapat diakses melalui antarmuka pengguna grafis (GUI) yang menampilkan tab untuk setiap komponen dan fungsi standar [7].

1.5. Methontology

Metode yang digunakan dalam penelitian ini adalah metode. Untuk membuat model ontologi, yang terbaik adalah mendeskripsikan setiap aktivitas yang akan dilakukan secara terpisah. Salah satu aspek terpenting dari metode ini adalah kemampuan untuk menentukan kebutuhan informasi spesifik, mengidentifikasi konsep dan hubungan yang relevan, dan merancang struktur ontologi yang konsisten dan akurat [8]. Dalam penelitian ini, pendekatan metodologis digunakan untuk membuat model ontologi umum. Metodologi ini memberikan seperangkat pedoman untuk mengelola aktivitas yang menjadi ciri proses pengembangan ontologi. Langkah-langkah proses konvensional adalah sebagai berikut [9].

2. Metode Penelitian

Metode yang digunakan pada penelitian ini adalah Methontology. Methontology adalah metode terstruktur yang dapat digunakan untuk membangun ontologi dari awal. Metode ini mencakup serangkaian aktivitas, teknik, dan hasil yang diproduksi oleh eksekusi dari setiap aktivitas menggunakan tekniknya masing-masing. Methontology sangat merekomendasikan penggunaan ontologi yang telah ada. Adapun tahapan dari Methontology sebagai berikut.



Gambar 1. Alur Metode Penelitian

2.1. Spesifikasi

Pada bagian spesifikasi implementasi model ontologi pada permainan digital interaktif bertujuan untuk menghasilkan dokumen spesifikasi ontologi, baik formal, semi formal maupun informal, yang ditulis dalam bahasa asli. Metode ini menggunakan serangkaian pernyataan perantara atau pertanyaan lanjutan. Tujuan dari bagian otentikasi ini adalah untuk mendefinisikan secara jelas konsep yang akan dimodelkan dalam ontologi permainan digital interaktif dan hubungan antara konsep-konsep tersebut. Dokumen visi ini berfungsi sebagai panduan untuk langkah selanjutnya dalam implementasi ontologi, memastikan bahwa struktur ontologi yang dibuat memenuhi persyaratan dan tujuan yang ditetapkan. Dalam proses penamaan, berbagai jenis informasi dan pertanyaan lanjutan digunakan untuk membantu memodelkan ontologi dengan cara yang efisien dan efektif yang sesuai dengan konteks pengembangan game digital interaktif. Oleh karena itu, ontologi yang dibangun tidak hanya mendukung pengalaman game yang kaya dan terstruktur, tetapi juga meningkatkan integrasi antara berbagai elemen game.

2.2. Akuisisi Pengetahuan

Dalam konteks penerapan model ontologi pada permainan digital interaktif, fase perolehan pengetahuan memegang peranan penting dalam pengembangan ontologi. Ini adalah tahap yang terjadi bersamaan dengan tahap otentikasi, dan merupakan tahap dimana data dan informasi terkait dikumpulkan dan dianalisis untuk mengidentifikasi konsep dan hubungan yang diperlukan dalam ontologi permainan. Data dan informasi yang dihasilkan selama proses visualisasi digunakan untuk merancang konsep dan hubungan antar konsep dalam ontologi. Sebagian besar pengumpulan informasi dilakukan bersamaan dengan bagian garis besar ini. Tahap implementasi merupakan langkah selanjutnya dalam pengembangan ontologi. Pada fase ini, konsep dan hubungan yang didefinisikan dan dirancang pada fase otentikasi diimplementasikan dalam bentuk ontologi. Proses implementasi melibatkan penentuan struktur ontologi, penetapan karakteristik dan hubungan antar konsep, dan pengujian keseluruhan ontologi untuk memastikan konsistensi dan akurasi. Seiring berjalannya proses pengembangan ontologi, aktivitas fase perolehan pengetahuan menurun seiring dengan beralihnya fokus ke fase implementasi, yang mengimplementasikan struktur ontologi yang telah direncanakan sebelumnya. Oleh karena itu, tahap perolehan pengetahuan dan tahap penamaan diselesaikan untuk menciptakan landasan yang kokoh bagi implementasi ontologi dalam permainan digital interaktif.

2.3. Konseptualisasi

Dalam konteks penerapan model ontologi pada permainan digital interaktif, bagian teoretis memainkan peran penting dalam menciptakan model konseptual yang mewakili pengetahuan domain permainan dan mendeskripsikan masalah serta solusi menggunakan kata-kata yang didefinisikan di bagian definisi. Tugas utamanya adalah membuat kamus komprehensif yang mencakup konsep, contoh, kata kerja dan objek yang berkaitan dengan permainan digital interaktif yang dikembangkan. Kamus komprehensif menangkap dan mendeskripsikan semua

pengetahuan tentang domain game, menunjukkan arti dan relevansi setiap konsep dan memastikan pemahaman lengkap tentang domain game sebagai model. Oleh karena itu, pada tahap konseptual, pemahaman yang jelas dan konsisten terhadap seluruh konsep dan istilah yang digunakan akan membantu menciptakan landasan yang kokoh bagi pengembangan ontologi permainan digital interaktif.

2.4. Integrasi

Dalam konteks penerapan model ontologi pada permainan digital interaktif, tahap integrasi berperan penting dalam mempercepat pengembangan ontologi dengan menggunakan definisi ontologi yang ada dan mengintegrasikannya ke dalam ontologi yang dibuat. Hal ini memungkinkan pengembangan ontologi untuk menggunakan struktur dan konsep yang telah didefinisikan dalam ontologi lain tanpa memulai dari awal. Proses integrasi mempercepat pengembangan ontologi dengan mengurangi redundansi dan meningkatkan interoperabilitas antar ontologi. Dengan menggunakan definisi ontologi yang ada, langkah ini membantu memperkaya ontologi baru dengan pengetahuan yang tervalidasi dan teruji dari sumber terpercaya. Oleh karena itu, komponen integrasi berperan penting dalam memperkaya dan memperluas ontologi permainan digital interaktif, serta menggunakan sumber daya nyata yang ada untuk meningkatkan efisiensi pengembangan.

2.5. Implementasi

Dalam konteks penggunaan model ontologi dalam permainan digital interaktif, tahap implementasi merupakan langkah penting untuk mengelola ontologi yang dirancang dan memungkinkannya dijalankan di komputer. Bagian ini mencakup semua proses mulai dari identifikasi hingga instalasi yang telah dilakukan sebelumnya. Tahap implementasi melibatkan modifikasi dan implementasi ontologi yang dibuat menggunakan perangkat lunak Protégé atau alat lain yang sesuai. Proses implementasi ini melibatkan penerjemahan konsep, hubungan, dan aturan yang didefinisikan dalam ontologi ke dalam struktur yang dapat dijalankan di komputer. Ontologi yang dibuat melalui implementasi ini dapat digunakan dalam lingkungan komputasi seperti pengambilan informasi, integrasi data, dan aplikasi lain yang berkaitan dengan permainan digital interaktif. Fase implementasi memastikan bahwa ontologi yang dibuat dapat diimplementasikan sesuai dengan kebutuhan konteks pengembangan game.

2.6. Evaluasi

Dalam konteks penggunaan model ontologi dalam permainan digital interaktif, tahap evaluasi memainkan peran penting dalam memastikan kualitas dan keberhasilan ontologi dan lingkungan komputasi terkait. Ada dua proses utama dalam tinjauan ini: verifikasi dan validasi. Verifikasi mengacu pada proses teknis yang memastikan kebenaran ontologi, lingkungan komputasi, dan dokumen yang terkait dengan kerangka referensi di setiap tahapan dan antar tahapan siklus hidup. Hal ini mencakup tinjauan teknis implementasi ontologi dan perangkat lunak terkait, termasuk pengujian fungsional dan kompatibilitas dengan persyaratan yang ditentukan. Verifikasi, di sisi lain, memastikan bahwa ontologi, lingkungan komputasi, dan dokumentasi sesuai dengan sistem yang ingin diwakilinya. Proses ini melibatkan evaluasi kecukupan ontologi yang dikembangkan terhadap kebutuhan dan tujuan yang telah ditentukan sebelumnya dalam konteks pengembangan game digital interaktif. Secara umum, tujuan dari tahap evaluasi adalah untuk memverifikasi bahwa ontologi dan lingkungan komputasi terkait telah dibangun dengan benar, sesuai dengan persyaratan yang ditetapkan, dan bahwa persyaratan sistem dapat dipenuhi. Oleh karena itu, tahap evaluasi merupakan langkah penting untuk menjamin keberhasilan implementasi ontologi dalam pengembangan game digital interaktif.

2.7. Dokumentasi

Ini adalah langkah terakhir. Dengan kata lain, proses dokumentasi dilakukan dalam kode ontologi, dalam dokumen bahasa alami yang dilampirkan pada definisi formal, dan dalam makalah yang diterbitkan dalam prosiding konferensi dan jurnal manajemen, pertanyaan-pertanyaan penting ontologi dirumuskan.

3. Hasil dan Diskusi

Dalam penelitian ini, mengembangkan sebuah ontologi yang berfokus pada domain Game Digital. Proses pengembangan ontologi ini melibatkan serangkaian tahapan, mulai dari spesifikasi hingga dokumentasi, dengan tujuan untuk memudahkan pemahaman dan pengembangan aplikasi game digital interaktif.

3.1. Spesifikasi

Tujuan dari fase spesifikasi ini adalah untuk menghasilkan dokumen spesifikasi ontologi yang dapat berbentuk formal, semi-formal, atau informal, ditulis dalam bahasa alami (natural language). Metode ini menggunakan satu set representasi yang berada di tengah-tengah (menengah) atau menggunakan pertanyaan kompetensi.

- a. Domain : Game
- b. Tanggal : 4 Mei 2024
- c. Dikonsep oleh : I Komang Maheza Yudistia
- d. Dilaksanakan oleh : I Komang Maheza Yudistia
- e. Tujuan : Membangun Model Ontologi untuk memudahkan klasifikasi Game
- f. Tingkat Formalitas : Semi-Formal
- g. Ruang Lingkup : Game
- h. Sumber Pengetahuan : Internet, buku, dan jurnal

3.2. Akuisisi Pengetahuan

Dalam proses pengembangan ontologi ini, sebagian besar akuisisi pengetahuan dilakukan pada tahap pemrosesan dengan persyaratan spesifikasi saat proses pengembangan ontologi. Pada tahap akuisisi pengetahuan ontologi pengembangan menggunakan teknik sebagai berikut.

- a. Melakukan studi literatur melalui sumber informasi tentang game digital yang ada.
- b. Analisis teks informal dan formal untuk mengidentifikasi konsep-konsep utama dan struktur ontologi.
- c. Pengumpulan data dari berbagai sumber, termasuk internet, buku, jurnal, dan game digital yang ada.

3.3. Konseptualisasi

Konseptualisasi bertujuan untuk menyusun domain pengetahuan dalam bentuk konseptual serta pemeliharaan dan pengelolaan pengetahuan yang diperoleh dalam proses akuisisi pengetahuan. Setelah model konseptual dibuat, proses berlanjut untuk mengubah model konseptual menjadi model formal untuk diimplementasikan dalam bahasa ontologi. Membangun domain pengetahuan mencakup konsep, instance, verba, dan properti. Jadi Kamus mendefinisikan dan mengumpulkan pengetahuan domain yang berguna, yang dapat digunakan dan diimplementasikan sebagai kelas dan subkelas.

3.4. Integritas

Pertimbangan untuk menggunakan kembali definisi yang ada dalam ontologi, yaitu meninjau penggunaan bahasa untuk menghindari kesalahan saat menentukan hubungan. Oleh karena itu dilakukan langkah-langkah untuk menggabungkan atau mengintegrasikan ontologi yang sudah ada ke dalam ontologi yang baru dibuat. Tujuannya adalah untuk memastikan bahwa materi baru dan perubahan pada dokumen yang sudah ada dibuat sesuai dengan persyaratan yang sama.

3.5. Implementasi

Dalam mengimplementasikan model ontologi, peneliti menggunakan aplikasi Protégé 5.6.3 dalam pengembangan ontologi. Protégé adalah perangkat lunak yang dikembangkan oleh Stanford Center for Biomedical Informatics Research di Stanford University School of Medicine.

Protégé adalah perangkat lunak untuk membantu mengembangkan ontologi berdasarkan sistem pengetahuan dasar. Setiap bagian ontologi didefinisikan sesuai dengan hasil dari setiap tahapan tugas dalam metode Methontology. Rancangan konseptual yang telah dilakukan kemudian diformalkan menggunakan aplikasi Protégé 5.6.3 Ontografi dan dapat dihasilkan model ontologi yang dibangun pada laporan ini.



Gambar 2. Classes dari ontologi Game Digital

Dalam pembuatan ontologi Game ini penulis menghasilkan 32 kelas. Setiap kelas dalam ontologi telah memiliki relasi dengan setiap individual yang disebut perpanjangan kelas.



Gambar 3. Object properties dari ontologi Game Digital

Objek properti yang dihasilkan dalam ontologi Game adalah 12 objek properti. Objek properti merupakan properti yang menghubungkan individu dengan individu lainnya.



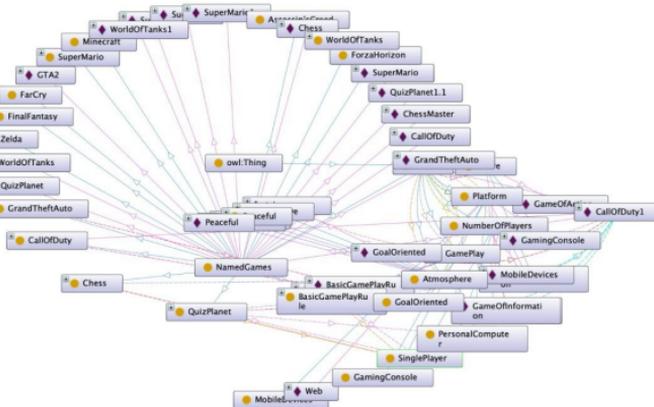
Gambar 4. Data properties dari ontologi Game Digital

Data properti yang dihasilkan dari ontologi Game sebanyak 0 data properti. Data properti adalah properti dari tiap individu yang memiliki hubungan dengan nilai data tersebut.

- ◆ Adventurous
- ◆ Aggressive
- ◆ BasicGamePlayRule
- ◆ CallOfDuty
- ◆ CallOfDuty1
- ◆ Chess
- ◆ ChessMaster
- ◆ GameOfAction
- ◆ GameOfInformation
- ◆ GamingConsole
- ◆ GoalOriented
- ◆ GrandTheftAuto
- ◆ GTA2
- ◆ MobileDevices
- ◆ Multiplayer
- ◆ Peaceful
- ◆ PersonalComputer
- ◆ QuizPlanet
- ◆ QuizPlanet1.1
- ◆ SinglePlayer
- ◆ SuperMario

Gambar 5. Individuals dari ontologi Game Digital

Individu dalam kelas diperpanjang disebut instance.



Gambar 6. Ontograf dari ontologi Game Digital

3.6. Evaluasi

Pada tahap ini, ontologi yang telah dibuat akan diuji. Tahap evaluasi ini melibatkan penggunaan SPARQL query pada web jena fuseki. Pertanyaan-pertanyaan yang telah disiapkan akan diubah menjadi query SPARQL untuk mendapatkan hasil yang terkait dengan ontologi yang telah dibuat. Hasil dari query dapat dilihat pada Gambar 6. Pada Gambar 6, terdapat contoh SPARQL query yang dieksekusi, yaitu sebagai berikut:

namedGames
CallOfDuty
GrandTheftAuto
QuizPlanet
SuperMario1
SuperMario
CallOfDuty1
GTA2
ChessMaster
WorldOfTanks
Sudoku1
Sudoku
Chess

Gambar 7. Hasil SPARQL query

3.7. Dokumentasi

Hasil dokumentasi dari penelitian pengembangan ontologi pengembangan game digital berupa tulisan yang tertuang dalam laporan ini.

4. Kesimpulan

Berdasarkan pembahasan yang telah dikemukakan, maka dapat disimpulkan sebagai berikut. Penelitian ini berhasil mengembangkan ontologi untuk game digital interaktif dengan menggunakan metodologi Methontology, yang mencakup tahap spesifikasi, akuisisi pengetahuan, konseptualisasi, integrasi, implementasi, dan evaluasi. Ontologi yang dihasilkan terdiri dari 32 kelas, 12 objek properti, 0 data properti, serta diuji menggunakan SPARQL query. Dengan menggunakan alat Protégé, ontologi ini dapat memfasilitasi klasifikasi dan pengelolaan pengetahuan secara lebih efektif, mendukung interoperabilitas data, dan meningkatkan pengalaman bermain game digital interaktif. Penelitian ini menunjukkan bahwa penerapan ontologi dalam domain game dapat meningkatkan struktur dan integrasi data, memungkinkan pengembangan aplikasi yang lebih kompleks dan terstruktur.

Daftar Pustaka

- [1] I. M. Ditha and H. Wiranata, "Pengembangan Web Semantik Silsilah Keluarga Kawitan Nararya Dalem Benciluk Tegeh Kori Dengan Metode Forward Chaining Dan Backward Chaining," vol. 8, no. 1, pp. 1–11, 2019.
- [2] C. R. A. Pramatha, "Assembly the Semantic Cultural Heritage Knowledge," J. Ilmu Komput., vol. 11, no. 2, p. 83, 2018. DOI: 10.24843/jik. 2018.v11.i02.p03.
- [3] C. Pramatha, J. G. Davis, and 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, pp. 560–571, 2018. DOI: 10.1007/978-3-030-01762-0_49.
- [4] N. M. A. Wiradhanta and I. B. M. Mahendra, "Perancangan Ontologi Semantik: Representasi Digital Oleh-Oleh Khas Bali," Jurnal Nasional Teknologi Informasi dan Aplikasinya, vol. 1, no. 3, p. 925, 2023.
- [5] P. R. Ganeswara and C. R. A. Pramatha, "Ontology-based Approach for Klungkung Royal Family," JELIKU (Jurnal Elektron. Ilmu Komput. Udayana), vol. 8, no. 4, p. 497, 2020. DOI:

- 10.24843/jlk.2020.v08.i04.p16.
- [6] P. D. Bangsa and I. Hermawan, "Jurnal Teknologi Terpadu," *J. Teknol. Terpadu*, vol. 7, no. 1, pp. 15–22, 2021, [Online]. Available: <https://media.neliti.com/media/publications/493730-water-ph-and-turbidity-control-system-in-0a553e14.pdf>.
- [7] P. I. Nugroho, B. Priyambadha, and N. Y. Setiawan, "Sistem Pencarian Koleksi Laporan Skripsi Dan PKL dengan Teknologi Web Semantik (Studi Kasus: Ruang Baca Fakultas Ilmu Komputer Universitas Brawijaya)," *J. Pengemb. Teknol. Inf. dan Ilmu Komputer*, vol. 2, no. 9, pp. 3440–3444, 2018.
- [8] K. D. P. Novianti, "Implementasi Methontology Untuk Pembangunan Model," *J. TEKNOIF*, vol. 4, no. 1, pp. 40–47, 2016. [Online]. Available: <https://ejournal.itp.ac.id/index.php/tinformatika/article/view/588/424>.
- [9] P. R. Ganeswara and C. R. A. Pramatha, "Ontology-based Approach for Klungkung Royal Family," *JELIKU (Jurnal Elektron. Ilmu Komput. Udayana)*, vol. 8, no. 4, p. 497, 2020. DOI: 10.24843/jlk.2020.v08.i04.p16.

Halaman ini sengaja dibiarkan kosong

Klasifikasi Mood pada Musik Pop dan Jazz dengan Menggunakan Mel Frequency Cepstral Coefficients dan K-Nearest Neighbor

I Gusti Bagus Putrawan^{a1}, I Ketut Gede Suhartana^{a2}

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹putrawan.2208561133@student.unud.ac.id
²kg.suhartana@unud.ac.id

Abstract

This research discusses mood classification in pop and jazz music using Mel Frequency Cepstral Coefficients (MFCC) and the K-Nearest Neighbor (KNN) algorithm. The dataset used consists of 900 songs with mood labels angry, happy, relaxed, and sad obtained from Kaggle. The data was processed by extracting 13 MFCC features and then continuing with classification using KNN. The research results show that the best accuracy reaches 64% with K=9. Accuracy at K=7 obtained a value of 60%, while at K=11 an accuracy of 58% was obtained. Evaluation was carried out using accuracy, precision, recall and f1-score metrics, with the best results found at K=9. This research emphasizes the importance of selecting K parameters for optimizing mood classification models.

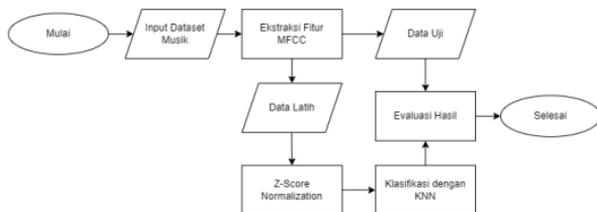
Keywords: Mood Classification, MFCC, K-Nearest Neighbor, Music Emotion Recognition

1. Pendahuluan

Musik merupakan seni yang menggabungkan irama, melodi, dan harmoni yang tercipta melalui kombinasi nada atau suara. Saat ini, musik telah menjadi simbol ekspresi emosional manusia. Misalnya, ketika seseorang merasa sedih, mereka cenderung mendengarkan musik yang melankolis, sedangkan saat bahagia, mereka memilih musik yang sesuai dengan suasana hati mereka. Musik telah menjadi bagian integral dari kehidupan manusia yang tidak dapat dipisahkan, mencerminkan dan mempengaruhi perasaan dan suasana hati dalam berbagai situasi. Penelitian terdahulu yang melakukan klasifikasi mood pada musik oleh Maulana memiliki dataset berjumlah 200 dari masing-masing 50 dataset mood, yaitu contentment, depression, exuberance, dan, anxiety. Pada penelitian ini dijelaskan bahwa menggunakan mfcc diperoleh akurasi sebesar 87,67 % dengan pemodelan Backpropagation Neural Network (BPNN) menggunakan 1 hidden layer 256 neuron, 0,4 dropout, 1000 epoch, learning rate 0.001, frame 40ms, dan overlap 40% dengan durasi 60 detik [1]. Pada penelitian yang berjudul Sistem Identifikasi Arti Tangisan Bayi Menggunakan Metode MFCC, DWT, dan KNN pada Raspberry Pi yang dilakukan oleh Prasetyo, dengan mengklasifikasikan suara bayi ke dalam 5 kelas, yaitu pada saat kondisi bayi merasa tidak nyaman, lapar, masuk angin, sendawa, dan mengantuk. Jumlah keseluruhan yang digunakan dalam penelitian ini adalah sebanyak 74 data dengan 4 data bukan tangisan bayi dengan perbandingan data latih sebesar 50, data uji sebesar 20. Diperoleh akurasi terbaik 90% dengan menggunakan frame sebanyak 512 data per frame, 39 fitur MFCC, dan klasifikasi menggunakan KNN dengan nilai K=5 [2].

2. Metode Penelitian

Adapun tahapan-tahapan yang dilakukan dalam penelitian adalah sebagai berikut.



Gambar 1. Flowchart klasifikasi Musik Berdasarkan Mood

2.1 Music Emotion Recognition (MER)

Dalam musik, suasana hati adalah keadaan emosi yang bertahan lama. Suasana hati berbeda dengan emosi sederhana karena suasana hati kurang spesifik, kurang intens, dan kecil kemungkinannya dipicu oleh rangsangan atau peristiwa tertentu. Mendengarkan sebuah lagu dapat memberikan suasana hati bagi pendengarnya tergantung dari musik yang didengarkannya. MER atau Music Emotion Recognition merupakan salah satu cabang pengetahuan dari Music Information Retrieval yang mempelajari tentang pengenalan proses emosi atau mood pada musik. Terdapat banyak perumusan yang dilakukan oleh ahli psikologi di zamannya. Oleh karena itu, dalam penelitian ini akan digunakan pengklasifikasian mood pada music yang terbagi ke dalam 4 kuadran, yaitu happy (kuadran I), angry (kuadran II), sad (kuadran III), dan relax (kuadran IV) berdasarkan definisi dari Robert E. Thayer dalam artikelnya yang berjudul *The Biopsychology of Mood and Arousal* tahun 1989 [3].



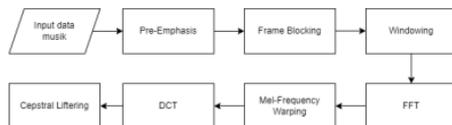
Gambar 2. Klasifikasi Mood Berdasarkan Definisi Robert E. Thayer [3]

2.2 Dataset

Dalam penelitian, digunakan dataset musik berjumlah 900 data yang sudah diberi label angry, happy, relax, dan sad dengan format .mp3 yang diperoleh dari kaggle yang berjudul 4Q audio emotion dataset (Russell's model) [4]. Dalam penelitian, dataset musik yang digunakan akan berfokus pada dataset lagu pop dan jazz barat dengan jumlah data sebesar 900. Dataset yang digunakan sebagai data latihan dan data uji masing-masing adalah 85:15.

2.3 Mel Frequency Cepstral Coefficient (MFCC)

Data audio yang sudah dilakukan preprocessing dilanjutkan dengan melakukan ekstraksi fitur MFCC. Secara umum, proses ekstraksi fitur MFCC dibagi ke dalam beberapa tahapan, antara lain:



Gambar 3. Tahapan-Tahapan Ekstraksi Fitur MFCC

2.3.1 Pre-Emphasis

Pre-emphasis adalah proses untuk menerima dan mempertahankan amplitudo berfrekuensi tinggi agar meningkatkan kualitas dari sinyal audio. Pre-emphasis memiliki beberapa fungsi, diantaranya adalah untuk mengurangi noise pada sinyal, meningkatkan resolusi frekuensi sinyal, dan menyeimbangkan spektrum dari sinyal audio [5]. Pre-emphasis dirumuskan dalam persamaan berikut.

$$y(n) = s(n) - \alpha s(n - 1) \tag{1}$$

Keterangan:

$y(n)$ = sinyal setelah dilakukan pre-emphasis
 $s(n)$ = sinyal awal
 α = konstanta pre-emphasis dimana biasanya digunakan pada rentang $0,95 < \alpha < 0,99$

2.3.2 Frame Blocking

Pada tahapan ini disebut tahapan segmentasi, yaitu sinyal audio akan dibagi menjadi beberapa frame yang lebih kecil dimana masing-masing frame mewakili sejumlah sampel audio. Sifat dari frame ini saling tumpang tindih atau overlap untuk menghindari kehilangan informasi yang terdapat pada ujung frame. Panjang ukuran frame yang digunakan adalah 25ms dengan panjang tiap frame bernilai sama.

2.3.3 Windowing

Proses windowing digunakan untuk mengurangi terjadinya efek aliasing setelah sinyal diproses pada proses frame blocking. Aliasing adalah sinyal baru dengan frekuensi yang berbeda dari frekuensi sinyal asli yang dapat menyebabkan sinyal terputus [6]. Pada tahap ini, sinyal dari proses frame blocking dikalikan dengan nilai Hamming windowing. Nilai dari Hamming windowing dituliskan dalam persamaan berikut dengan N merupakan jumlah sampel setiap frame.

$$w_n = 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right), \quad 0 \leq n \leq N-1 \tag{2}$$

Pada persamaan windowing dituliskan sebagai berikut dengan x_n adalah frame ke- n .

$$y_n = x_n \cdot w_n \tag{3}$$

2.3.4 Fast Fourier Transform

FFT atau Fast Fourier Transform digunakan untuk mengubah domain waktu menjadi domain frekuensi pada frekuensi setiap frame [7]. Dalam tahapan ini, dihasilkan spektrum atau spektral yang merepresentasikan amplitudo dari berbagai frekuensi di dalam sinyal audio. FFT dirumuskan dalam persamaan berikut

$$y_k = \sum_{n=0}^{N-1} y_n e^{-\frac{2\pi i k n}{N}}, \quad n = 0, 1, 2, \dots, N - 1 \quad (4)$$

2.3.5 Mel Frequency Warping

Sinyal yang sudah melalui proses FFT selanjutnya akan melewati proses Mel Frequency Warping dengan menggunakan filterbank. Filterbank adalah bentuk filter yang dirancang untuk menentukan ukuran energi dari suatu frekuensi tertentu dalam domain frekuensi. Dalam proses ini, sinyal hasil FFT dikelompokkan menggunakan filter triangular filter file. Setiap nilai FFT kemudian dikalikan dengan nilai filter yang sesuai dan hasil perkalian dijumlahkan. Tujuan dari proses tersebut adalah untuk mendapatkan representasi energi dari sinyal suara dalam skala mel, yang lebih sesuai dengan persepsi pendengaran manusia [7]. Frekuensi suara diukur menggunakan skala mel, yang bersifat linier untuk frekuensi di bawah 1000 Hz dan logaritmik untuk frekuensi di atas 1000 Hz.

2.3.6 Discrete Cosine Transform

DCT (Discrete Cosine Transform) adalah langkah terakhir dari proses utama ekstraksi fitur MFCC. DCT digunakan untuk mengonversi spektrum magnitudo menjadi domain waktu agar dapat direpresentasikan lebih baik dan akan menghasilkan Mel-Frequency Coefficient Cepstrum dalam bentuk vektor [6]. Jumlah koefisien dari cepstral yang digunakan dalam penelitian umumnya berjumlah 13. DCT dirumuskan dalam bentuk persamaan berikut.

$$C_m = \sum_{k=1}^N \cos \left[n \left(k - \frac{1}{2} \right) \frac{\pi}{N} \right] \log E_k, \quad n = 1, 2, \dots, L \quad (5)$$

Keterangan:

- C_m = koefisien MFCC ke-k
- E_k = output dari Mel Frequency Warping
- N = jumlah Mel-Frequency Coefficient Cepstrum
- L = jumlah Cepstral Coefficient

2.3.7 Cepstral Liftering

Cepstral Liftering diimplementasikan ketika diperoleh hasil dari DCT dalam bentuk fitur cepstral dengan menggunakan persamaan window sehingga memperoleh hasil model yang lebih baik. Persamaan windows dirumuskan sebagai berikut.

$$w_n = 1 + 2 \sin \left(\frac{n\pi}{L} \right) \quad (6)$$

Keterangan:

- W_n = Koefisien lifting untuk Cepstral Coefficient ke-nn = index ke-n Cepstral Coefficient
- L = jumlah Cepstral Coefficient

2.4 Z-Score Normalization

Standard Scaling atau disebut juga dengan Z-Score Normalization adalah metode yang digunakan dalam mengatur data berdasarkan nilai rata-rata atau mean dan standar deviasi dari data. Dataset dinormalisasi dengan mengubah skala pada distribusi nilai rata-rata fitur menjadi 0 dan untuk standar deviasi fitur menjadi 1 [8]. Persamaan dari Z-Score Normalization dirumuskan sebagai berikut.

$$x_{\text{Scaled}} = \frac{x - \mu}{\sigma} \quad (7)$$

2.5 K-Nearest Neighbor (KNN)

Algoritma KNN biasa digunakan dalam supervised learning dimana model mempelajari dari label yang sudah diberikan sebelumnya. KNN sendiri bekerja dengan cara mengklasifikasikan data yang baru dikenali berdasarkan mayoritas dari nilai k-tetangga terdekat. KNN menyimpan semua data latih dan label diikuti dengan menentukan nilai K atau jumlah tetangga terdekat yang akan digunakan untuk mengklasifikasikan data baru. Pada setiap data baru yang muncul, KNN menghitung jarak antara data tersebut dengan setiap data dalam data latih menggunakan metode seperti Euclidean, Manhattan, dan Minkowski. Dilanjutkan dengan pemilihan K data latih dengan jarak terdekat sebagai tetangga terdekat [9]. KNN dengan euclidean distance function dirumuskan dalam persamaan berikut.

$$d(x_i, y_i) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (8)$$

2.6 Confusion Matrix

Pada pengevaluasian model dilakukan dengan menggunakan confusion matrix yang bertujuan melakukan pengelompokan berdasarkan empat kelas yang terdiri atas True Positive (TP), False Positif (FP), False Negative (FN), dan True Negative (TN). TP terjadi ketika model dapat memprediksi suatu peristiwa yang sesuai dengan kondisi yang sesungguhnya. FP menunjukkan prediksi yang benar, akan tetapi pada kondisi yang sebenarnya bernilai salah. TN menunjukkan prediksi bernilai salah tepat dengan keadaan sesungguhnya yang bernilai salah. FN memprediksi kejadian bernilai salah dimana peristiwa sesungguhnya bernilai benar. Berdasarkan output Confusion Matrix sehingga diperoleh akurasi [10]. Berdasarkan definisi tersebut, confusion matrix dapat dirumuskan sebagai berikut.

Tabel 1. Confusion Matrix

Predicted Label		Actual Label	
		Positive	Negative
Positive	True Positive (TP)	False Positive (FP)	
	False Negative (FN)	True Negative (TN)	

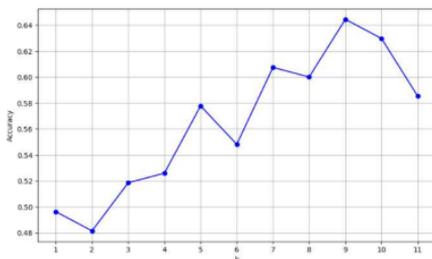
Akurasi adalah perbandingan antara akumulasi data TP dan TN dengan keseluruhan data sehingga dapat dirumuskan dalam persamaan berikut.

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + FP + TN} \quad (9)$$

3. Hasil dan Pembahasan

3.1 Klasifikasi dengan K-Nearest Neighbor dan MFCC

Dataset yang telah terkumpul terdiri dari 900 musik dengan genre pop yang berbeda. Dilanjutkan dengan melakukan ekstraksi fitur pada dataset dengan menggunakan MFCC sehingga tiap data musik berisi 13 fitur audio MFCC. Data musik tersebut akan dipisahkan menjadi data latih dan data uji dengan perbandingan sebesar 85:15. Proses klasifikasi data latih menggunakan K- Nearest Neighbor dan pengujian dengan metrik akurasi, presisi, recall, dan f1-score.

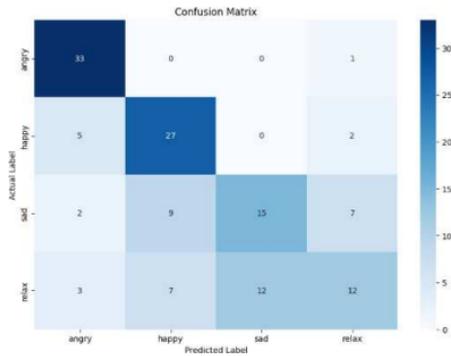


Gambar 4. Hasil Perbandingan Nilai K pada K-Nearest Neighbor

	precision	recall	f1-score	support
angry	0.77	0.97	0.86	34
happy	0.63	0.79	0.70	34
relax	0.55	0.35	0.43	34
sad	0.56	0.45	0.50	33
accuracy			0.64	135
macro avg	0.62	0.64	0.62	135
weighted avg	0.62	0.64	0.62	135

Gambar 5. Hasil Akurasi, Presisi, Recall, dan F1-Score dengan K=9

Berdasarkan Gambar 5, diperoleh hasil terbaik dari model K-Nearest Neighbor dengan parameter K = 9, dimana pemilihan K berpengaruh pada tingkat akurasi, akan tetapi ketika parameter K > 9, menghasilkan model yang kurang optimal dikarenakan akurasinya mulai menurun. Nilai akurasi dari data uji yang diprediksi benar oleh model sebesar 64%, yaitu sebanyak 87 data dan prediksi salah sebesar 36% dengan jumlah data 48 data. Mood marah diklasifikasikan benar dengan jumlah 33 data, mood happy dengan jumlah 27 data, mood sedih dengan jumlah 15 data, dan mood santai dengan jumlah 12 data.



Gambar 6. Evaluasi dengan Confusion Matrix

4. Kesimpulan

Klasifikasi mood pada lagu pop dan jazz dilakukan menggunakan fitur MFCC dan K-Nearest Neighbor. Evaluasi akurasi menggunakan metrik akurasi, presisi, recall, dan f1-score. Akurasi terbaik diperoleh dengan perbandingan data latih dengan data uji adalah 85:15. Klasifikasi menggunakan K-Nearest Neighbor dan 13 fitur MFCC memberikan akurasi sebesar 64% dengan parameter K = 9, akurasi 60% pada K = 7, dan 58% pada K = 11. Jumlah mood yang diprediksi benar sebanyak 87 data sedangkan jumlah yang diprediksi salah sebanyak 48 data.

Referensi

- [1] P. I. Maulana, A. Aranta, F. Bimantoro, and I. G. Andika, "Klasifikasi Mood Musik Berdasarkan Mel Frequency Cepstral Coefficients Dengan Backpropagation Neural Network," *Jurnal Resistor (Rekayasa Sistem Komputer)*, vol. 5, no. 1, pp. 72–85, Apr. 2022, doi: 10.31598/jurnalresistor.v5i1.1089.
- [2] Y. Yohannes and R. Wijaya, "Klasifikasi Makna Tangisan Bayi Menggunakan CNN Berdasarkan Kombinasi Fitur MFCC dan DWT," *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 8, no. 2, pp. 599–610, Jun. 2021, doi: 10.35957/jtisi.v8i2.470.
- [3] R. E. Thayer, *The Biopsychology of Mood and Arousal*. Oxford University PressNew York, NY, 1990. [Online]. Available: <http://dx.doi.org/10.1093/oso/9780195068276.001.0001>
- [4] R. Panda, R. Malheiro, and R. P. Paiva, "Novel Audio Features for Music Emotion Recognition," *IEEE Transactions on Affective Computing*, vol. 11, no. 4, pp. 614–626, Oct. 2020, doi: 10.1109/taffc.2018.2820691.
- [5] S. Y. Yusdiantoro and T. B. Sasongko, "Implementasi Algoritma MFCC dan CNN dalam Klasifikasi Makna Tangisan Bayi," *Indonesian Journal of Computer Science*, vol. 12, no. 4, Aug. 2023, doi: 10.33022/ijcs.v12i4.3243.
- [6] S. P. Dewi, A. L. Prasasti, and B. Irawan, "Analysis of LFCC Feature Extraction in Baby Crying Classification using KNN," in *2019 IEEE International Conference on Internet of Things and Intelligence System (IoTaIS)*, Nov. 2019. Accessed: May 24, 2024. [Online]. Available: <http://dx.doi.org/10.1109/iotais47347.2019.8980389>
- [7] I. D. G. y A. Wibawa and I. D. M. B. A. Darmawan, "Implementation of audio recognition using mel frequency cepstrum coefficient and dynamic time warping in wirama praharsini," *Journal of Physics: Conference Series*, vol. 1722, no. 1, p. 012014, Jan. 2021, doi: 10.1088/1742-6596/1722/1/012014.
- [8] P. I. Maulana, A. Aranta, F. Bimantoro, and I. G. Andika, "Klasifikasi Mood Musik Berdasarkan Mel Frequency Cepstral Coefficients Dengan Backpropagation Neural

- Network," *Jurnal Resistor (Rekayasa Sistem Komputer)*, vol. 5, no. 1, pp. 72–85, Apr. 2022, doi: 10.31598/jurnalresistor.v5i1.1089.
- [9] I. N. Y. T. Giri, L. A. A. Rahning Putri, G. A. V. Mastrika Giri, I. G. N. Anom Cahyadi Putra, I. M. Widiartha, and I. W. Supriana, "Music Genre Classification Using Modified K-Nearest Neighbor (MK-NN)," *JELIKU (Jurnal Elektronik Ilmu Komputer Udayana)*, vol. 10, no. 3, p. 261, Feb. 2022, doi: 10.24843/jjk.2022.v10.i03.p02.
- [10] J. Davis and M. Goadrich, "The relationship between Precision-Recall and ROC curves," in *Proceedings of the 23rd international conference on Machine learning - ICML '06*, 2006. [Online]. Available: <http://dx.doi.org/10.1145/1143844.1143874>

Penerapan Enkripsi dan Dekripsi Dokumen Data UMKM Menggunakan Algoritma ChaCha20-Poly1305

I Made Chandra Widjaya^{a1}, I Komang Ari Mogi^{a2}

^aProgram Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹madecandra55@gmail.com
²arimogi@unud.ac.id

Abstract

Despite the increasing importance of data security in safeguarding sensitive information, this study addresses the potential risks associated with unauthorized access to critical data. Employing the ChaCha20-Poly1305 algorithm, the research focuses on implementing encryption and decryption processes for Small and Medium Enterprise (SME) documents, supplemented by key derivation from AES-256 for enhanced security. A nonce Initialization Vector (IV) is generated using the ChaCha20-Poly1305 algorithm, with users inputting secret keys for encryption. The system then encrypts the data using the ChaCha-Poly1305 algorithm and derives keys from AES using SHA-256 hashing. For decryption, users input the encrypted document into the program, along with the previously used key. The system design employs a simple web-based application, with the ChaCha20-Poly1305 cryptography algorithm implemented in PHP. The study successfully tests the ChaCha20-Poly1305 algorithm, and the program exhibits secure decryption processes, evidenced by consistent byte sizes of tested SME documents.

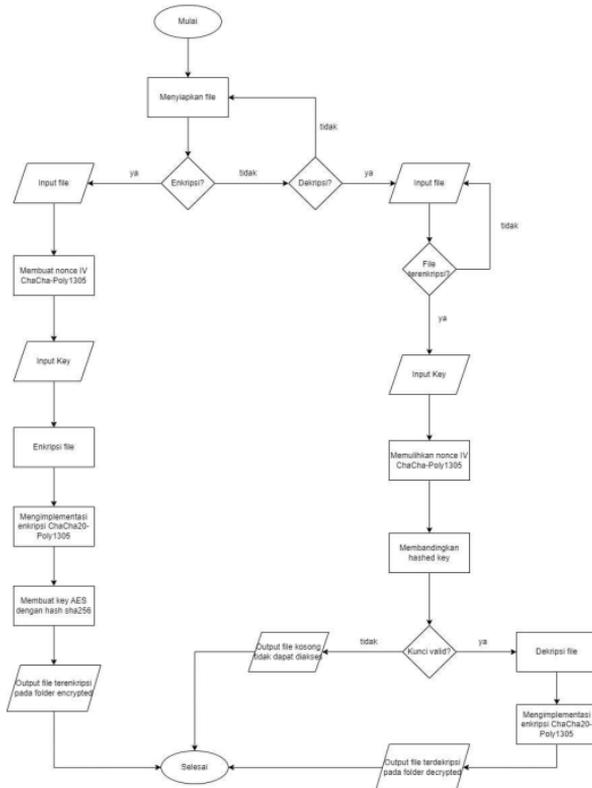
Keywords: Cryptography, ChaCha20-Poly1305, File Security, Encryption, Decryption

1. Pendahuluan

Dalam periode perkembangan teknologi yang semakin canggih, informasi yang tersebar dalam dunia maya menjadi semakin rentan. Meski demikian, keamanan data dianggap sebagai hal yang krusial untuk melindungi informasi sensitif, mengingat potensi dampak negatifnya jika data penting disadap oleh pihak yang tidak berwenang[1]. Salah satu cara untuk mengamankan informasi sensitif adalah melalui penggunaan teknik enkripsi dan dekripsi. Enkripsi menjadi salah satu proses mengubah teks menjadi format yang tidak dapat dibaca atau sulit dipahami. Sedangkan dekripsi menjadi proses mengembalikan teks tersebut ke dalam bentuk semula. Dengan menerapkan teknik ini, informasi yang dikirimkan melalui internet atau disimpan dalam database dapat dijaga keamanannya meskipun rentan terhadap serangan pihak yang tidak bertanggung jawab. Penerapan bentuk enkripsi dan dekripsi teks dapat dilakukan juga pada dokumen-dokumen penting. Dokumen-dokumen tersebut mungkin berisi informasi rahasia perusahaan, data keuangan sensitif, atau bahkan dokumen pribadi seperti surat-surat elektronik dan file-file identitas. Dengan menggunakan teknik enkripsi yang tepat, dokumen-dokumen tersebut dapat diamankan dari akses yang tidak sah, baik saat disimpan dalam sistem penyimpanan digital maupun saat dikirimkan melalui jaringan internet. Selain itu, penerapan dekripsi yang hanya dapat dilakukan oleh pihak yang berwenang juga dapat memberikan tingkat keamanan tambahan terhadap dokumen-dokumen penting tersebut[2]. Kriptografi menjadi sangat populer dengan berbagai macam algoritma yang ada. Beberapa diantaranya terdapat algoritma AES-256, RSA, Diffie Helman, ChaCha20, dan masih banyak lagi. Penelitian ini menggunakan ChaCha20-Poly1305 sebagai algoritma untuk menerapkan proses enkripsi dan dekripsi pada dokumen data UMKM. Tidak hanya itu, dalam penerapannya menggunakan key turunan dari AES-256. Maka dari itu, diharapkan penerapan kriptografi ini dapat menjaga dengan baik kepentingan setiap dokumen atau file data terkait [3].

2. Metode Penelitian

2.1 Desain Sistem



Gambar 1. Flowchart Sistem

Untuk alur cara kerja sistem, diawali dengan menyiapkan dokumen file penting Usaha Mikro Kecil Menengah (UMKM). Setelah itu, masuk ke dalam program untuk melakukan proses enkripsi dengan memasukkan dokumen. Program akan membuat *nonce Initialization Vector (IV)* dari algoritma ChaCha20-Poly1305. Selanjutnya, pengguna menginputkan kunci rahasia dan melakukan proses enkripsi. Pada tahap ini, sistem akan melakukan proses enkripsi dengan algoritma ChaCha-Poly1305 dan membuat kunci turunan dari algoritma *Advanced Encryption Standard (AES)* dengan *hash* SHA-256. Setelah proses enkripsi selesai, program otomatis mengubah *environment* disesuaikan dengan ekstensi dokumen, lalu menyimpannya pada *folder encrypted*.

Jika pemilik dokumen ingin melakukan dekripsi, maka pemilik harus menggunakan dokumen yang telah dienkripsi dan masuk ke dalam program untuk di dekripsi. Di dalam program, diharuskan untuk memasukkan kunci yang telah dimasukkan sebelumnya untuk dokumen pada saat di enkripsi, lalu sistem akan memulihkan *nonce Initialization Vector (IV)* dari algoritma ChaCha20-Poly1305. Selanjutnya, program membandingkan kunci yang telah di *hash* sebelumnya. Jika kunci tidak *valid*, maka program hanya akan memberikan *output* berupa dokumen kosong atau tidak dapat diakses. Jika kunci *valid*, maka proses dekripsi dilanjutkan dan menyimpan dokumen terdekripsi yang sama seperti awalnya di dalam *folder decrypted*.

2.2 Pemrosesan Awal

Proses awal sistem dengan pembuatan *nonce Initialization Vector (IV)* menggunakan ChaCha20-Poly1305, program mengambil langkah-langkah yang cermat untuk memastikan keamanan. ChaCha20-Poly1305 digunakan sebagai konstruksi kriptografi yang menggabungkan keamanan *stream cipher* ChaCha20 dengan autentikasi pesan Poly1305 untuk melindungi data dokumen. Saat proses enkripsi dimulai, program secara otomatis menghasilkan *nonce IV* yang unik setiap kali dengan menggunakan fungsi acak yang kuat. *Nonce IV* ini membuat nilai acak yang hanya digunakan satu kali dan berperan penting dalam mencegah serangan repetisi yang mungkin terjadi. Selanjutnya, *nonce IV* ini digunakan bersama dengan kunci rahasia dan data yang akan dienkripsi, memastikan keamanan dan keunikan setiap proses enkripsi. Sementara itu, terkait dengan koneksi antara kunci AES256 dan hash SHA-256, sistem juga mengimplementasi langkah-langkah yang efisien dalam pengelolaan kunci enkripsi. Sebelum proses enkripsi dimulai, pengguna diminta untuk memasukkan kunci rahasia yang akan digunakan. Kunci ini kemudian diberikan fungsi *hash SHA-256* untuk menghasilkan kunci turunan yang akan digunakan dalam proses enkripsi AES. *Hashing SHA-256* mengubah kunci rahasia menjadi nilai *hash* yang panjang dan unik, memastikan keamanan kunci yang digunakan dalam proses enkripsi[4]. Hasil *hash SHA-256* ini kemudian menjadi kunci untuk algoritma AES, yang akan digunakan dalam proses enkripsi data. Dengan demikian, program memastikan bahwa kunci enkripsi yang digunakan dalam proses AES unik dan aman.

2.3 Enkripsi ChaCha20-Poly1305

Proses enkripsi dengan ChaCha20-Poly1305 adalah sebuah algoritma kriptografi simetris yang terdiri dari dua komponen utama, yaitu ChaCha20 untuk enkripsi data dan Poly1305 untuk otentikasi pesan[5]. Pertama, ChaCha20 mengambil kunci enkripsi dan *nonce* (nomor sekali pakai) sebagai *input*, kemudian menghasilkan *keystream*. *Keystream* ini kemudian digunakan untuk melakukan enkripsi data dengan operasi XOR. Selanjutnya, Poly1305 digunakan untuk menghasilkan *tag* otentikasi pesan yang unik untuk setiap pesan yang dienkripsi. Proses ini melibatkan penggunaan polinomial pada *keystream* dan kunci enkripsi. *Tag* otentikasi ini memastikan bahwa pesan tidak diubah oleh pihak yang tidak sah selama proses pengiriman. Proses enkripsi dengan ChaCha20-Poly1305 terjamin keamanannya karena ChaCha20 adalah algoritma *stream cipher* yang aman, sementara Poly1305 adalah fungsi *hash* yang kuat untuk otentikasi pesan. Kombinasi dari kedua komponen ini memberikan keamanan yang kuat dan efisien untuk proses enkripsi data.

2.4 Dekripsi ChaCha20-Poly1305

Proses dekripsi dengan ChaCha20-Poly1305 mirip dengan proses enkripsinya, tetapi dilakukan secara terbalik. Pertama, *tag* otentikasi pesan diverifikasi menggunakan algoritma Poly1305 dengan menggunakan kunci enkripsi yang sama dan *nonce* yang sama yang digunakan pada saat enkripsi. Jika *tag* otentikasi tidak cocok, proses dekripsi dihentikan karena pesan mungkin telah dimanipulasi. Setelah verifikasi *tag* otentikasi berhasil, *keystream* dihasilkan menggunakan algoritma ChaCha20 dengan kunci enkripsi yang sama dan *nonce* yang sama seperti pada proses enkripsi. *Keystream* kemudian digunakan untuk melakukan operasi XOR pada pesan terenkripsi, menghasilkan pesan asli. Dalam proses dekripsi, sangat penting untuk menggunakan *nonce* yang sama seperti yang digunakan pada saat enkripsi dan memastikan bahwa kunci enkripsi tetap rahasia. Dengan menggunakan *nonce* yang sama, proses dekripsi dapat

menghasilkan *keystream* yang identik dengan yang digunakan pada saat enkripsi, sehingga memungkinkan untuk mendapatkan pesan asli dari pesan terenkripsi dengan benar[6].

3. Hasil dan Pembahasan

3.1 Rancangan Sistem

Rancangan sistem pada penelitian ini menggunakan aplikasi berbasis *website* sederhana. Bentuk algoritma kriptografi ChaCha20-Poly1305 dibuat dalam bahasa pemrograman PHP. Program juga menggunakan *webserver* Apache pada XAMPP untuk dijalankan. Perancangan sistem ini memiliki tujuan untuk menguji algoritma dapat digunakan pada proses enkripsi dan dekripsi dokumen data UMKM. Penelitian ini menggunakan contoh data Surat Keterangan Tempat Usaha (SKTU) sebagai data UMKM.



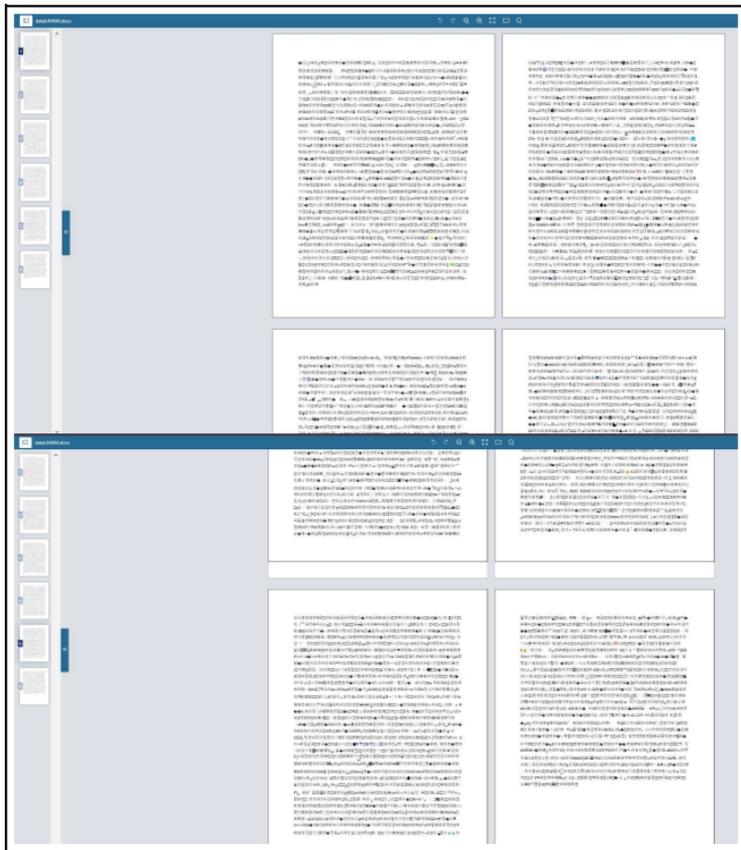
Gambar 2. Tampilan Sistem

3.2 Uji Coba Sistem

a. Proses Enkripsi Dokumen UMKM

Proses uji coba sistem ini dilakukan untuk mengetahui apakah dokumen terkait dienkripsi dengan menggunakan algoritma ChaCha20-Poly1305 merujuk pada tahapan-tahapan berikut:

- Menyiapkan dokumen UMKM dengan ekstensi *.docx* yang akan dienkripsi.
- Masukkan dokumen ke dalam program enkripsi.
- Program akan menggunakan algoritma ChaCha20-Poly1305 untuk membuat *nonce Initialization Vector (IV)*.
- Memasukkan kunci rahasia untuk proses enkripsi.
- Proses enkripsi ChaCha20-Poly1305 akan dilakukan.
- Program akan membuat kunci turunan dari algoritma *Advanced Encryption Standard (AES)* dengan *hash* SHA-256.
- Setelah proses enkripsi selesai, program secara otomatis mengunduh *environment* yang disesuaikan dengan ekstensi dokumen.
- Dokumen yang telah dienkripsi disimpan dalam folder yang ditetapkan, yaitu *folder "encrypted"*.



b. Proses Dekripsi Dokumen UMKM

- Uji coba menggunakan dokumen yang telah dienkripsi untuk proses dekripsi.
- Dokumen dienkripsi dimasukkan ke dalam program dekripsi.
- Program meminta untuk memasukkan kunci yang digunakan saat proses enkripsi.
- Sistem memulihkan *nonce Initialization Vector (IV)* dari algoritma ChaCha20-Poly1305.
- Program membandingkan kunci yang telah di-hash sebelumnya dengan kunci yang dimasukkan.
- Jika kunci tidak valid, program hanya akan memberikan output berupa dokumen kosong atau tidak dapat diakses.
- Jika kunci valid, proses dekripsi dilanjutkan.
- Dokumen terdekripsi disimpan dalam folder yang ditetapkan, yaitu folder "decrypted".

Tabel 2. Proses Dekripsi

Dokumen UMKM Setelah Dekripsi

PEMERINTAH KODTA DENPASAR
KECAMATAN DENPASAR SELATAN
DESA SANUR KAJA

Alamat : Jl. By Pass Ngurah Rai No. 59
<http://sanurkaja.denpasar.kota.go.id>

Telp. (0361) 287388 Kode Pos 80227
sanurkaja@gmail.com

SURAT KETERANGAN TEMPAT USAHA

Nomor : 323/Pel.Um/TV/2021

Yang bertanda tangan di bawah ini, Perbekel Desa Sanur Kaja, Kecamatan Denpasar Selatan, Kota Denpasar, berdasarkan surat pengantar Kepala Dusun *Belong* Nomor *226/KD-BLGI/2021* Tanggal : *20 April 2021* bahwa :

Nama Pemohon/ : IKOMANG TRISNA DHARMITA
Penanggung Jawab
Alamat Pemohon/ : JL. HANGTUAH GG. MAWAR NO. 22 DPS,
Penanggung Jawab BR/LINK.BELONG, Kode Pos : 80227, Sanur Kaja,
Denpasar Selatan, Kota Denpasar, Bali
Nama Perusahaan : Warung Made
Jenis Usaha : Berdagang Sembako

Sepanjang pengetahuan kami dan sampai surat ini dikeluarkan memang benar usaha tersebut beralamat di :

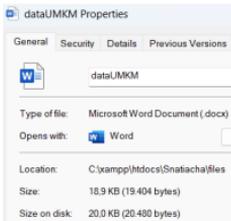
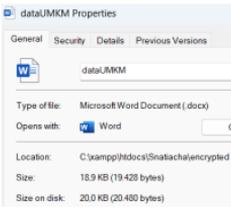
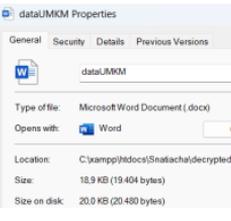
Jalan : Jl. Hangtuh Gg Mawar No. 22
Dusun/Lingkungan : Belong
Kelurahan / Desa : Sanur Kaja
Kecamatan : Denpasar Selatan
Kabupaten/Kota : Kota Denpasar
Provinsi : Bali

Demikian surat keterangan ini kami buat dengan sebenarnya agar dapat dipergunakan untuk *Keperluan Administrasi BPUM Kota Denpasar*

Denpasar, 20 April 2021
Perbekel Desa Sanur Kaja

I MADE SUDANA

Tabel 3. Perbandingan Besaran File

Default	Encrypted	Decrypted
		
Besaran: 19.404 bytes	Besaran: 19.428 bytes	Besaran: 19.404 bytes

4. Kesimpulan

Berdasarkan hasil dan bahasan yang dijelaskan, penelitian ini telah berhasil menguji coba algoritma ChaCha20-Poly1305 dengan kunci *hash* SHA-256 pada enkripsi dan dekripsi dokumen UMKM. Program dapat melindungi dengan baik data dokumen UMKM menggunakan *nonce Initialization Vector* (IV) yang dibangkitkan secara acak untuk dapat melakukan proses enkripsi. Program juga menunjukkan hasil proses dekripsi yang aman dan baik dilihat dari tidak berubahnya ukuran *bytes* dokumen UMKM yang diuji. Dengan penelitian ini, diharapkan untuk penelitian selanjutnya agar dapat mengembangkan algoritma sejenis pada pengamanan data.

Daftar Pustaka

- [1] M. Azhari, D. I. Mulyana, F. J. Perwitosari, dan F. Ali, "Implementasi Pengamanan Data pada Dokumen Menggunakan Algoritma Kriptografi Advanced Encryption Standard (AES)," *Jurnal Pendidikan Sains dan Komputer*, vol. 2, no. 01, pp. 163–171, Mar. 2022, doi: 10.47709/jpsk.v2i01.1390.
- [2] F. P. Utama, G. Wijaya, R. Faurina, and A. Vatesria, "Implementasi Algoritma AES 256 CBC, BASE 64, Dan SHA 256 dalam Pengamanan dan Validasi Data Ujian Online," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 10, no. 5, pp. 945–954, Oct. 2023, doi: 10.25126/jtiik.20231056558.
- [3] Djong, Handrian Saputra, and Siswanto Siswanto. "Implementasi Kriptografi Dengan Menggunakan Metode RC4 dan AES-256 Untuk Mengamankan File Dokumen pada PT Varnion Technology Semester." *Prosiding Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)*. Vol. 1. No. 1. 2022.
- [4] G. D. M. Zulma, H. B. Seta, dan T. Yuniati, "Implementasi Algoritma Aes Dan Bcrypt Untuk Pengamanan File Dokumen," *Informatik: Jurnal Ilmu Komputer*, vol. 18, no. 2, p. 163, Aug. 2022, doi: 10.52958/itfk.v18i2.4667.
- [5] R. Serrano, C. Duran, M. Sarmiento, C.-K. Pham, and T.-T. Hoang, "*ChaCha20-Poly1305 Authenticated Encryption with Additional Data for Transport Layer Security 1.3*," *Cryptography*, vol. 6, no. 2, p. 30, 2022, doi: 10.3390/cryptography6020030.
- [6] A. M. ElRashidy, M. H. Abd AlAzeem, A. A. Abd ElHafez, and M. F. Abo Sree, "*ChaCha20-AES Combined Algorithm with 512 Bits of Security*," *2024 6th International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE)*. 2024, pp. 1–6. doi: 10.1109/REEPE60449.2024.10479797.

Halaman ini sengaja dibiarkan kosong

Klasifikasi Genre Musik Menggunakan Metode Support Vector Machine Dengan Multi-Kernel

I Gusti Agung Istri Agrivina Shyta Devi^{a1}, I Made Widiartha^{a2}

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹gunggekshyta@email.com
²madewidiartha@unud.ac.id

Abstract

Music is a universal art that reflects cultural diversity and individual preferences through various genres. This research explores music genre classification using Support Vector Machine (SVM) with multi-kernel methods. The SVM algorithm, known for its effectiveness in handling complex datasets, is employed to classify music genres based on audio features. The research utilizes the GTZAN dataset, comprising 10 music genres, and extracts audio features from WAV files. After normalization and data splitting, SVM models are trained and evaluated. Results indicate a significant accuracy improvement after hyperparameter tuning, with the best models achieving accuracies of 88.92% for the polynomial kernel and 89.32% for the RBF kernel.

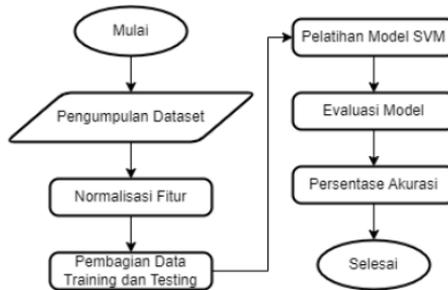
Keywords: Classification, Music Genre, SVM, Kernel Function

1. Pendahuluan

Musik merupakan sebuah seni yang universal, keberagaman genre musik merupakan cerminan dari keberagaman budaya dan selera dari setiap individu [1]. Klasifikasi genre musik adalah salah satu aplikasi umum dari analisis data audio. Hal ini melibatkan pengelompokan atau pengkategorian lagu-lagu berdasarkan karakteristik seperti tempo, melodi, harmoni, dan instrumen yang digunakan. Klasifikasi genre musik merupakan langkah penting dalam memahami dan mengorganisir keanekaragaman musik yang ada. Dengan membagi lagu-lagu ke dalam genre-genre yang berbeda, kita dapat lebih mudah mengidentifikasi dan menemukan jenis musik yang memenuhi keinginan dan preferensi kita. Selain itu, klasifikasi genre juga memungkinkan para pengamat musik, peneliti, dan profesional industri musik untuk menganalisis tren, memahami preferensi audiens, dan mengembangkan rekomendasi musik yang lebih tepat. Dalam era digital saat ini, klasifikasi genre musik juga menjadi kunci untuk pengembangan teknologi. Dari pemrosesan sinyal digital hingga kecerdasan buatan, pemahaman tentang genre musik membuka pintu untuk inovasi yang tak terbatas. Misalnya, algoritma yang mampu mengidentifikasi dan merekomendasikan musik berdasarkan preferensi pengguna dapat meningkatkan pengalaman mendengarkan musik secara signifikan. Salah satu algoritma yang telah terbukti efektif dalam tugas klasifikasi adalah Mesin Vektor Pendukung (Support Vector Machine/SVM). SVM menawarkan keunggulan dalam menangani masalah klasifikasi dengan banyak fitur dan kumpulan data yang kompleks, seperti yang sering ditemui dalam analisis data audio. Dalam jurnal penelitian dengan judul "Klasifikasi Genre Musik Menggunakan Teknik Pembelajaran Mesin" menyatakan bahwa jika dibandingkan dengan K-NN nilai akurasi dari SVM lebih tinggi yaitu 71%, sementara dengan K-NN sebesar 69,6% [6]. Selain itu, menurut Santosa dan Erwin SVM dapat mengatasi permasalahan multiclass [7].

2. Metode Penelitian

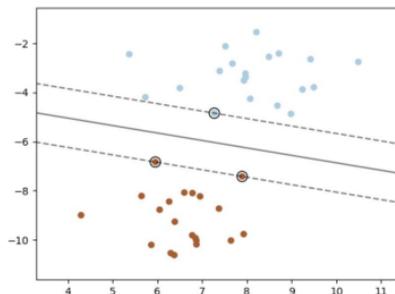
Dalam penelitian ini, ada prosedur yang harus dijalankan dalam mengklasifikasikan genre musik. Hal ini dimulai dari pengumpulan dataset sampai proses pelatihan model dan evaluasi.



Gambar 2. Flowchart Alur Penelitian

2.1 Support Vector Machine

Algoritma pembelajaran mesin Support Vector Machine (SVM) dapat digunakan untuk tugas regresi dan klasifikasi. Konsep Minimisasi Risiko Struktural (SRM) adalah dasar dari prinsip kerja SVM. Tujuan SRM adalah untuk menemukan Hyperplane yang memiliki kemampuan untuk memisahkan ruang input menjadi dua kelas. Teori SVM dimulai dengan mengklasifikasikan contoh-contoh dalam ruang linier yang dapat dibedakan oleh hyperplane dan dikelompokkan berdasarkan kelasnya. Ide dasar SVM dimulai dengan permasalahan klasifikasi biner yang memerlukan kumpulan data pelatihan yang terdiri dari sampel positif dan negatif. SVM akan berupaya mengoptimalkan pemilihan hyperplane sebagai pemisah antara kedua kelas dengan maksud untuk memaksimalkan margin antara kelas-kelas tersebut [4]. Satu atau lebih hyperplane dibangun dalam ruang yang besar atau tak terbatas oleh support vector machine. Ruang ini dapat digunakan untuk melakukan tugas seperti klasifikasi, regresi, atau lainnya. Hyperplane yang memiliki jarak terdekat ke titik data latih terdekat dari setiap kelas melakukan pemisahan yang efektif; biasanya, semakin besar jarak, semakin kecil kesalahan generalisasi yang dihasilkan oleh pengklasifikasi. Pada gambar 1 di bawah ini menunjukkan fungsi keputusan untuk masalah yang dapat dipisahkan secara linear, dengan tiga sampel di batas margin, yang disebut "support vectors" [4].



Gambar 1. Support vectors [4]

2.2 Kernel Function

Kernel function adalah metode yang digunakan untuk mengambil data sebagai input dan mengubahnya menjadi bentuk yang diperlukan untuk memproses data [5]. "Kernel" digunakan karena serangkaian fungsi matematika yang digunakan dalam Mesin Vektor Pendukung memberikan kesempatan untuk memanipulasi data. Jadi, kernel function pada umumnya mengubah set data latih sehingga permukaan keputusan non-linear dapat diubah menjadi persamaan linear dalam ruang dimensi yang lebih tinggi. Pada dasarnya, Kernel function mengembalikan hasil perkalian dalam dua titik dalam dimensi fitur standar. Adapun rumus standar dari kernel function [5]:

$$K(\underline{x}) = 1, \text{ if } \|\underline{x}\| \leq 1 \quad (1)$$

$$K(\underline{x}) = 0, \text{ otherwise} \quad (2)$$

Terdapat beberapa kernel function yang umum digunakan dalam pengklasifikasian yaitu, kernel linear, RBF, dan polynominal. Berikut ini adalah persamaan dari ketiga function kernel:

Rumus perhitungan dari kernel linear [3].

$$K(x, x_i) = x \cdot x_i \quad (3)$$

Rumus perhitungan dari kernel RBF [3].

$$K(x, x_i) = \exp(-\gamma * \|x - x_i\|^2) \quad (4)$$

Rumus perhitungan dari kernel polynomial [3].

$$K(x, x_i) = [y * (x, x_i) + \text{coed}]^d \quad (5)$$

2.3 Data

Dataset yang digunakan merupakan data sekunder yang diperoleh dari situs web <https://www.kaggle.com> yang dikenal sebagai "GTZAN dataset – Music Genre Clasification". Terdapat 10 genre musik yang diklasifikasikan yaitu, blues, classical, country, disco, hiphop, jazz, metal, pop, reggae, dan rock. Jumlah dari masing – masing genre adalah sebanyak 100 data untuk setiap genre musik dengan format WAV. Dalam dataset juga terdapat file CSV "features_3_sec.csv" yang berisi hasil ekstraksi fitur dari setiap audio musik, data tersebut adalah data yang digunakan dalam penelitian ini.

```
Index(['length', 'chroma_stft_mean', 'chroma_stft_var', 'rms_mean', 'rms_var',  
      'spectral_centroid_mean', 'spectral_centroid_var',  
      'spectral_bandwidth_mean', 'spectral_bandwidth_var', 'rolloff_mean',  
      'rolloff_var', 'zero_crossing_rate_mean', 'zero_crossing_rate_var',  
      'harmony_mean', 'harmony_var', 'perceptr_mean', 'perceptr_var', 'tempo',  
      'mfcc1_mean', 'mfcc1_var', 'mfcc2_mean', 'mfcc2_var', 'mfcc3_mean',  
      'mfcc3_var', 'mfcc4_mean', 'mfcc4_var', 'mfcc5_mean', 'mfcc5_var',  
      'mfcc6_mean', 'mfcc6_var', 'mfcc7_mean', 'mfcc7_var', 'mfcc8_mean',  
      'mfcc8_var', 'mfcc9_mean', 'mfcc9_var', 'mfcc10_mean', 'mfcc10_var',  
      'mfcc11_mean', 'mfcc11_var', 'mfcc12_mean', 'mfcc12_var', 'mfcc13_mean',  
      'mfcc13_var', 'mfcc14_mean', 'mfcc14_var', 'mfcc15_mean', 'mfcc15_var',  
      'mfcc16_mean', 'mfcc16_var', 'mfcc17_mean', 'mfcc17_var', 'mfcc18_mean',  
      'mfcc18_var', 'mfcc19_mean', 'mfcc19_var', 'mfcc20_mean', 'mfcc20_var',  
      'label'],  
      dtype='object')
```

Gambar 3. Data Frame File CSV

2.4 Normalisasi Fitur

Proses ini terdiri dari dua langkah utama yaitu normalisasi fitur dan pembagian data menjadi set pelatihan, validasi, dan pengujian. Langkah pertama adalah normalisasi fitur, langkah ini dilakukan dengan menggunakan "StandardScaler" dari "sklearn.preprocessing". Fungsi dari "StandardScaler" adalah menghitung rata-rata dan standar deviasi dari setiap fitur "x" dan mentransformasikannya sehingga setiap fitur memiliki rata – rata nol dan varians satu. Hasil dari normalisasi akan disimpan dalam "x_scaled". Langkah kedua adalah pembagian data menjadi beberapa subset untuk pelatihan, validasi, dan pengujian. Pembagian ini dilakukan dalam dua tahap menggunakan "train_test_split" dari pustaka "sklearn.model_selection". Pada tahap pertama data dibagi menjadi dua bagian yaitu 70% untuk data pelatihan dan validasi ("X_train_val" dan "y_train_val") dan 30% untuk data pengujian awal ("X_test_initial" dan "y_test_initial"). Lalu pada tahap kedua, data pengujian awal dibagi menjadi dua bagian yaitu 50% untuk validasi ("X_val" dan "y_val") dan 50% untuk pengujian akhir ("X_test_final" dan "y_test_final").

2.5 Pelatihan Model SVM

Pelatihan model SVM dalam penelitian ini dimulai dengan inisialisasi model SVM untuk setiap kernel yang didefinisikan dalam daftar "kernels", yaitu "linear", "poly", dan "rbf". Pada setiap iterasi dalam loop "for", model SVM diinisialisasi dengan kernel spesifik menggunakan "SVC(kernel=kernel)". Setelah model diinisialisasi, model akan dilatih dengan menggunakan metode "fit" pada data pelatihan "X_train" dan labelnya "y_train". Proses pelatihan ini melibatkan penyesuaian parameter model untuk memaksimalkan kinerja pada data pelatihan sehingga model dapat memisahkan atau mengklasifikasikan data dengan baik berdasarkan kernel yang digunakan. Selanjutnya, dilakukan tuning hyperparameter untuk masing – masing kernel menggunakan Grid Search. Parameter grid didefinisikan dalam dictionary "param_grid" mencakup berbagai nilai untuk hyperparameter "c" pada semua kernel, serta "degree" untuk kernel polinomial dan "gamma" untuk kernel rbf. Grid Search dilakukan dengan "GridSearchCV", yang melibatkan pelatihan model dengan setiap kombinasi hyperparameter yang mungkin dan mengevaluasi kinerjanya menggunakan cross-validation dengan 5 lipatan ("cv=5"). Kombinasi hyperparameter terbaik dipilih berdasarkan akurasi tertinggi pada data validasi.

2.6 Evaluasi

Proses evaluasi model SVM dilakukan setelah melalui tahap pelatihan dan tuning hyperparameter. Evaluasi dimulai dengan menggunakan model terbaik yang ditemukan melalui Grid Search untuk masing-masing kernel: linear, polynomial (poly), dan Radial Basis Function (rbf). Setelah model terbaik dipilih, evaluasi dimulai dengan menggunakan data validasi ("X_val" dan "y_val"). Model SVM digunakan untuk membuat prediksi pada data validasi, dan hasil prediksi ini dibandingkan dengan label sebenarnya ("y_val"). Metrik evaluasi yang digunakan dalam kode tersebut adalah akurasi, yang mengukur persentase prediksi yang benar dari total prediksi yang dibuat. Hasil akurasi dari prediksi pada data validasi kemudian dicatat untuk setiap kernel dan disimpan dalam sebuah dictionary best_model_accuracies. Dengan mengevaluasi model menggunakan data validasi yang tidak digunakan selama proses pelatihan atau tuning hyperparameter, kita dapat menilai secara obyektif kinerja model yang dihasilkan. Evaluasi ini memungkinkan kita untuk membandingkan kinerja berbagai kernel SVM setelah tuning hyperparameter dan memilih kernel yang memberikan hasil terbaik untuk tugas klasifikasi yang diberikan.

3. Hasil dan Diskusi

Tabel 1. Nilai Akurasi

Kernel	Persentase
linear	77.30%

Kernel	Persentase
polynomial	79.77%
RBF	85.25%

Tabel diatas adalah nilai akurasi sebelum dilakukannya tuning hyperparameter. Nilai akurasi yang diperoleh untuk masing-masing kernel SVM memberikan gambaran awal tentang kinerja model dengan menggunakan parameter default. Akurasi untuk kernel linear tercatat sebesar 77.30%, menunjukkan bahwa model ini mampu mengklasifikasikan data dengan tingkat akurasi yang cukup baik tanpa penyetelan lebih lanjut. Kernel polynomial menghasilkan akurasi sebesar 79.77%, yang sedikit lebih tinggi dibandingkan kernel linear, menunjukkan bahwa kernel ini bisa menangkap beberapa pola non-linear dalam data dengan lebih efektif. Kernel RBF mencatat akurasi tertinggi sebesar 85.25%, menunjukkan keunggulannya dalam menangani data yang mungkin memiliki pola yang lebih kompleks dan non-linear.

3.1. Hasil Akurasi

Tabel 2. Nilai Akurasi Tuning Hyperparameter

Kernel	Persentase
linear	76.84%
polynomial	88.92%
RBF	89.32%

Tabel diatas ini merupakan hasil nilai akurasi setelah tuning hyperparameter menggunakan Grid Search, hasil nilai akurasi menunjukkan perubahan yang signifikan dibandingkan dengan akurasi sebelum tuning. Akurasi untuk kernel linear sedikit menurun dari 77.30% menjadi 76.84%, yang mungkin disebabkan oleh overfitting atau perubahan parameter yang tidak cocok untuk data ini. Namun, kernel polynomial dan RBF menunjukkan peningkatan akurasi yang substansial. Akurasi kernel polynomial meningkat dari 79.77% menjadi 88.92%, dan akurasi kernel RBF meningkat dari 85.25% menjadi 89.32%. Peningkatan akurasi pada kernel polynomial dan RBF menunjukkan bahwa tuning hyperparameter berhasil menemukan kombinasi parameter yang lebih optimal untuk model, sehingga model dapat lebih baik dalam mengklasifikasikan data pada set validasi.

4. Kesimpulan

Berdasarkan hasil pengujian, klasifikasi genre musik menggunakan metode Support Vector Machine (SVM) dengan multi-kernel menunjukkan kinerja yang sangat baik. Dengan menggunakan dataset GTZAN yang terdiri dari 10 genre musik dan ekstraksi fitur audio, model SVM berhasil mencapai akurasi yang signifikan pada data pengujian. Sebelum tuning hyperparameter, akurasi terbaik yang dicapai oleh kernel RBF adalah sebesar 85.25%. Setelah tuning hyperparameter, terjadi peningkatan akurasi yang substansial, terutama pada kernel polynomial dan RBF. Kernel polynomial meningkat dari 79.77% menjadi 88.92%, dan kernel RBF meningkat dari 85.25% menjadi 89.32%, menunjukkan bahwa tuning hyperparameter berhasil menemukan kombinasi parameter yang lebih optimal untuk model, sehingga model dapat lebih baik dalam mengklasifikasikan data pada set validasi.

Daftar Pustaka

- [1] F.Aulia, "Keberagaman Genre Musik yang Bisa Memecah Belah Penikmatnya", 28 April 2024. [Online]. Available: <https://geotimes.id/komentar/keberagaman-genre-musik-yang-bisa-memecah-belah-penikmatnya/>. [8 Mei 2024]

- [2] Ardiansyah, and Meilina. "Klasifikasi Genre Musik Menggunakan Metode Support Vector Machine." *Jurnal Ilmiah Fasikom* 5, no. 2, 81–86, 2016
- [3] Ren, Yili, Fuxiang Hu and Hongping Miao. "The optimization of kernel function and its parameters for SVM in well-logging." 2016 13th International Conference on Service Systems and Service Management (ICSSSM), 1-5, 2016
- [4] BINUS. "Support Vector Machine Algorithm", 14 Februari 2022. [Online]. Available: <https://sis.binus.ac.id/2022/02/14/support-vector-machine-algorithm/>. [7 Mei 2024]
- [5] N.Fury. "Major Kernel Function in Support Vector Machine (SVM)", 7 Februari 2022. [Online]. Available: <https://www.geeksforgeeks.org/major-kernel-functions-in-support-vector-machine-svm/>. [7 Mei 2024]
- [6] G. A. V. M. Giri and M. L. Radhitya, "Klasifikasi Genre Musik Menggunakan Teknik Pembelajaran Mesin", *Jutik*, vol. 9, no. 1, 2023.
- [7] Jaya, E., Santosa. "Klasifikasi Genre Musik Menggunakan Metode Support Vector Machine". *Jurnal Ilmiah Fasikom* 81–82, 2016

Analisa Rancangan Desain Antarmuka Aplikasi LibrarySense Menggunakan System Usability Scale

Gagas Pradipta Jatmiko¹, I Putu Gede Hendra Suputra²

¹Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹gagasjatmiko16@gmail.com
²hendrasuputra@unud.ac.id

Abstract

The comfort of the library plays a crucial role in supporting learning activities. Temperature that is too hot or too cold, high levels noise, or the presence of smoke can affect the focus and effectiveness of learning for library users. LibrarySense is an application that utilizes wireless sensor network to monitor the environmental conditions of the library in real-time. The application is designed to assist library managers in improving operational efficiency and facility management, as well as enhancing user experience. The System Usability Scale (SUS) method is employed to evaluate the usability of the application. Literature review is conducted to understand the context of implementing SUS, while user interface design is carried out through wireframing and high-fidelity design. Usability testing is conducted using SUS questionnaires distributed to respondents. The research findings indicate that the LibrarySense application achieves an average SUS score of 81.74, indicating a high level of user satisfaction with usability and user experience. This suggests that the LibrarySense application has the potential to enhance the quality of the library environment and support managers in making more informed decisions.

Kata Kunci: System Usability Scale, Wireless Sensor Network, Library Monitoring System, User Interface, Usability

1. Pendahuluan

Kenyamanan perpustakaan sangat penting dalam menentukan pengalaman belajar pengunjung. Suhu yang terlalu panas atau dingin, tingkat kebisingan yang tinggi, atau adanya asap dapat tingkat fokus dan efektivitas belajar dari para pengguna perpustakaan [1]. Dengan memanfaatkan teknologi sensor nirkabel kita dapat memantau kondisi lingkungan perpustakaan secara real-time melalui aplikasi mobile. Sensor-sensor yang terintegrasi dalam aplikasi dapat mengukur suhu (sensor DHT11), tingkat kebisingan (MAX4466), kualitas udara (MQ-2), dan mendeteksi asap (MQ-2). Aplikasi LibrarySense dapat meningkatkan efisiensi operasional dan pengelolaan fasilitas perpustakaan. Pengelola perpustakaan dapat secara real-time memantau kondisi perpustakaan dimanapun dan kapanpun. Pengelola perpustakaan juga akan menerima notifikasi jika terdapat peningkatan suhu yang tidak diinginkan, tingkat kebisingan yang tinggi, atau deteksi asap, sehingga memungkinkan mereka untuk mengambil tindakan preventif dengan cepat untuk menangani masalah tersebut sebelum menjadi lebih serius. Penelitian ini bertujuan dalam merancang sebuah desain antarmuka pengguna berbasis IOS pada aplikasi pemantauan kenyamanan pada sebuah perpustakaan LibrarySense. Dimana pada antarmuka ini akan ditampilkan informasi real-time tentang kenyamanan pada perpustakaan berdasarkan sensor-sensor jaringan nirkabel yang telah dipasang pada perpustakaan. Selain itu, aplikasi ini juga menyediakan data historis tentang kondisi lingkungan perpustakaan. Dengan demikian, aplikasi ini tidak hanya membantu meningkatkan kualitas lingkungan perpustakaan, tetapi juga mendukung pengelola dalam mengambil keputusan yang lebih terinformasi dan tepat waktu. System Usability Scale membuat peneliti dapat mengumpulkan data yang kuat tentang bagaimana pengguna menilai berbagai aspek desain antarmuka, termasuk kejelasan informasi, efisiensi navigasi, dan keseluruhan kepuasan pengguna. Analisis rancangan desain antarmuka

menggunakan SUS akan memberikan wawasan yang berharga bagi pengembang dalam mengidentifikasi area yang memerlukan perbaikan atau peningkatan, sehingga memungkinkan mereka untuk membuat keputusan desain yang lebih terinformasi dan efektif.

2. Metode Penelitian

Penelitian ini akan menggunakan metode SUS (System Usability Scale) sebagai metode penelitian. System Usability Scale adalah alat pengukuran yang digunakan dalam mengevaluasi kegunaan dari berbagai macam perangkat lunak, situs web, aplikasi, dan lainnya. System Usability Scale terdiri dari sepuluh buah pertanyaan yang dirancang untuk menilai aspek-aspek kegunaan pada suatu sistem. Selain itu System Usability Scale juga akan menilai kejelasan, kemudahan penggunaan, dan kepuasan pengguna.



Gambar 1. Alur Penelitian

2.1 Studi Literatur

Pada penelitian ini studi literatur menjadi landasan dalam memahami konteks penerapan System Usability Scale. Studi literatur membantu peneliti dalam memperoleh pemahaman yang mendalam tentang System Usability Scale. Studi literatur adalah aktivitas penelitian dengan menggunakan teknik pengumpulan informasi dan data dengan sumber-sumber informasi yang relevan [2]. Studi literatur dilakukan dengan menganalisis berbagai sumber jurnal ilmiah, buku, dan riset terkait. Dengan demikian, peneliti dapat memanfaatkan dan mengimplementasikan metode System Usability Scale dengan memastikan validitas dan reliabilitas hasil evaluasi kegunaan sistem yang dihasilkan [3].

2.2 Perancangan Wireframe

Pada tahap selanjutnya setelah studi literatur dilakukan, dilakukanlah perancangan berupa rancangan desain antarmuka aplikasi. Perancangan akan berupa sebuah representasi visual sederhana yang disebut wireframe. Dimana wireframe nantinya akan dibuat dengan figma. Figma adalah platform desain kolaboratif yang memungkinkan tim desain untuk bekerja bersama secara online, membuat wireframe dengan cepat dan efisien, serta memfasilitasi diskusi dan iterasi desain secara timbal balik. Wireframe dapat digunakan untuk merancang tata letak dan struktur dasar dari antarmuka aplikasi, menentukan posisi dan fungsi elemen-elemen utama, serta menggambarkan alur interaksi pengguna secara kasar [4].

2.3 Pembuatan High-Fidelity

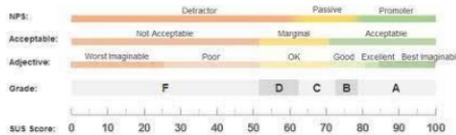
Pada tahap selanjutnya akan dibuat sebuah desain berdasarkan wireframe sebelumnya. Wireframe akan dikembangkan menjadi sebuah high fidelity. High Fidelity (hi-fi) adalah desain yang mendekati tingkat akhir atau final dari sebuah program atau aplikasi. Desain high fidelity berupa detail-detail visual lengkap, seperti warna, tipografi, dan elemen-elemen grafis lain dengan tingkat ketelitian tinggi [5]. Rancangan ini nantinya akan digunakan pada uji kelayakan kepada pengguna. Pembuatan high fidelity juga akan menggunakan figma.

2.4 Pengujian Usability dengan Metode System Usability Scale

Pada tahap terakhir akan dilakukan pengujian usability dengan menggunakan metode system usability scale. Pertama peneliti akan menentukan skala pertanyaan system usability scale yang akan digunakan dalam kuisisioner evaluasi. Skala ini akan terdiri dari sepuluh pertanyaan yang akan dievaluasi oleh responden berdasarkan tingkat setuju hingga tidak setuju. Setelah itu, peneliti akan menyebarkan kuisisioner kepada responden terkait. Kuisisioner akan diberikan setelah responden menggunakan prototipe aplikasi. Data yang dikumpulkan akan dianalisis untuk

mendapatkan gambaran tentang persepsi pengguna pada aplikasi [6]. Aturan dalam menghitung skor rata-rata system usability scale adalah sebagai berikut:

- a. Pertanyaan nomor ganjil akan mendapatkan skor dari skor pengguna yang dikurangi 1.
- b. Pertanyaan nomor genap akan mendapatkan skor dari mengurangi nilai 5 dengan skor pengguna.
- c. Hasil dari konversi akan dijumlahkan untuk setiap responden. Setelah itu akan dilakukan perkalian dengan 2.5 agar mendapatkan rentang nilai antara 0 hingga 100 [7].



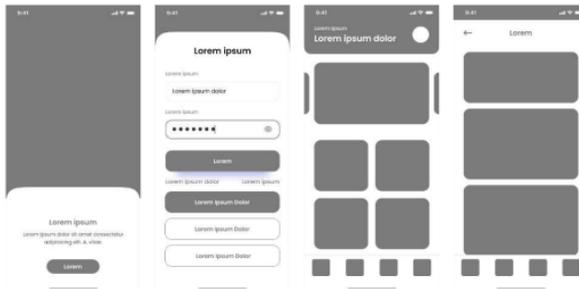
Gambar 2. Tingkatan Skor SUS

Pada gambar (2) dapat dilihat tingkatan-tingkatan skor penilaian yang ada pada metode System Usability Scale. Dimana tingkatan-tingkatan tersebut akan menjadi dasar penilaian untuk hasil rata-rata skor System Usability Scale berdasarkan hasil kuesioner [8].

3. Hasil dan Diskusi

3.1 Wireframe

Berikut ini adalah hasil desain wireframe untuk aplikasi LibrarySense:



Gambar 3. Wireframe LibrarySense



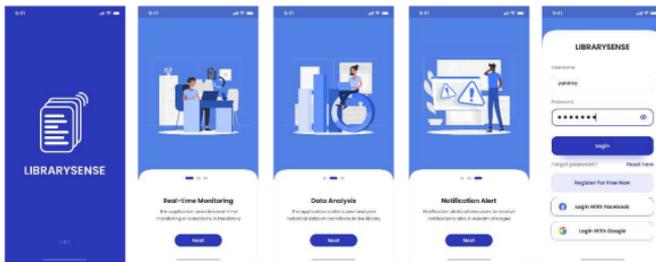
Gambar 4. Wireframe LibrarySense

Pada gambar 3 dan 4 dapat dilihat tampilan desain wireframe dari aplikasi LibrarySense. Dimana wireframe adalah sebuah sketsa atau diagram dasar yang menunjukkan struktur dan tata letak elemen-elemen utama, seperti tombol, menu, dan area konten. Dimana pada gambar 3 dan 4 terdapat desain wireframe dari login page, home page, connect page, notification page, dan profile page. Wireframe ini digunakan sebagai dasar kerangka dalam membuat desain HighFidelity dari aplikasi LibrarySense.

3.2 High Fidelity

Berdasarkan elemen-elemen dasar wireframe yang telah diciptakan, selanjutnya akan dikembangkan menjadi sebuah High Fidelity (Hi-Fi). Dimana Hi-Fi menjadi sebuah desain tingkat tinggi yang mendekati desain akhir dari aplikasi LibrarySense ini. Berikut adalah desain Hi-Fi dari aplikasi ini:

a. Login Page



Gambar 5. High Fidelity Login Page

Pada gambar 5, terlihat desain high fidelity dari halaman login. Di halaman antarmuka login, pengguna akan melihat tampilan awal yang menampilkan logo aplikasi LibrarySense. Selanjutnya, pengguna akan diberikan informasi singkat tentang layanan yang disediakan oleh LibrarySense. Selanjutnya pengguna diminta untuk memasukkan kredensial autentikasi, seperti nama pengguna dan kata sandi, atau menggunakan metode

otentikasi lain seperti login melalui akun facebook atau google. Setelah kredensial pengguna diverifikasi, pengguna dapat diarahkan ke halaman beranda.

b. Home Page



Gambar 6. High Fidelity Home Page

Pada gambar 6 dapat dilihat tampilan high fidelity dari home page. Dimana ini merupakan tampilan utama pada aplikasi LibrarySense. Pada halaman ini pengguna dapat melihat informasi-informasi detail tentang suhu ruangan, kebisingan, kualitas udara, dan asap pada perpustakaan. Pengguna juga dapat melihat detail lebih dengan melihat diagram yang menggambarkan tren yang terjadi pada perpustakaan. Pengguna juga dapat melihat kondisi perpustakaan pada hari ini, kemarin, satu minggu terakhir, satu bulan terakhir, hingga satu tahun terakhir.

c. Connect Page



Gambar 7. High Fidelity Connect Page

Pada gambar 7 dapat dilihat tampilan high fidelity dari connect page. Dimana pada halaman ini pengguna dapat memantau sensor-sensor yang terhubung telah terhubung dengan aplikasi LibraryConnect. Selain itu pada halaman ini pengguna juga dapat menambahkan sensor baru yang ingin dihubungkan dengan aplikasi.

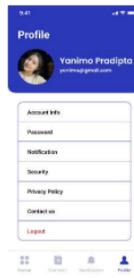
d. Notification Page



Gambar 8. High Fidelity Notification Page

Gambar 8 memperlihatkan tampilan high fidelity dari notification page aplikasi LibraySense. Dimana pada halaman ini pengguna akan diberikan informasi tentang riwayat notifikasi yang masuk ke aplikasi.

e. Profile Page



Gambar 9. High Fidelity Profile Page

Pada gambar 9 terlihat desain high fidelity dari profile page, dimana pada halaman ini pengguna dapat melihat informasi detail tentang nama pengguna, email pengguna, informasi akun, menu pergantian password, pengaturan notifikasi, pengaturan keamanan, privacy policy, contact us, dan menu logout.

3.3 Hasil Pengujian System Usability Scale

Berikut adalah 10 pertanyaan system usability scale yang akan digunakan dalam mengukur aspek kegunaan pada prototipe aplikasi ini.

Table 1. 10 Pertanyaan System Usability Scale

No	Pertanyaan System Usability Scale
Q1	Saya berpikir akan menggunakan sistem ini lagi

Q2	Saya merasa sistem ini sulit digunakan
Q3	Saya merasa sistem ini mudah digunakan
Q4	Saya membutuhkan bantuan dalam menggunakan sistem ini
Q5	Saya merasa fitur sistem berjalan semestinya
Q6	Saya merasa ada banyak hal yang tidak konsisten
Q7	Saya merasa orang lain akan memahami cara menggunakan sistem ini dengan cepat
Q8	Saya merasa sistem ini membingungkan
Q9	Saya merasa tidak ada hambatan dalam menggunakan sistem ini
Q10	Saya perlu membiasakan diri dahulu dalam menggunakan sistem ini

Berdasarkan kuesioner yang telah disebarakan diperoleh hasil sebagai berikut:

Tabel 2. Hasil Kuesioner

Responden	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
R1	3	2	3	2	3	3	3	3	3	3
R2	5	2	5	4	4	3	4	3	5	2
R3	4	1	5	1	5	1	3	2	5	2
R4	4	1	4	1	4	1	4	1	4	1
R5	4	4	4	4	4	1	3	3	3	5
R6	5	1	5	1	5	2	5	1	5	2
R7	5	1	5	3	5	2	5	2	5	2
R8	4	2	4	2	4	2	4	2	4	2
R9	4	2	5	4	4	2	4	2	4	4
R10	3	2	4	1	4	3	5	2	5	1
R11	4	4	5	1	4	2	5	1	5	3
R12	5	2	4	3	5	2	4	2	4	2
R13	3	3	4	3	5	2	4	3	3	4
R14	4	2	4	2	4	1	4	1	4	3
R15	3	3	3	4	3	3	3	3	2	4
R16	5	1	5	1	5	1	5	1	5	1
R17	5	1	5	1	5	1	5	1	5	1
R18	5	1	5	1	5	1	5	1	5	1
R19	5	2	5	2	5	1	5	1	5	1
R20	5	1	5	2	5	1	5	2	5	1
R21	5	2	5	1	4	1	5	1	5	2
R22	5	2	5	1	5	1	5	1	4	1
R23	5	1	5	1	5	1	5	1	5	1

Data hasil dari kuesioner yang telah disebarakan akan diproses berdasarkan aturan analisis metode System Usability Scale. Dimana data tersebut akan menghasilkan tabel berikut.

Tabel 3. Skor SUS

Responden	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Total	Skor SUS
R1	2	3	2	3	2	2	2	2	2	2	22	55
R2	4	3	4	1	3	2	3	2	4	3	29	72.5
R3	3	4	4	4	4	4	2	3	4	3	35	87.5
R4	3	4	3	4	3	4	3	4	3	4	35	87.5
R5	3	1	3	1	3	4	2	2	2	0	21	52.5
R6	4	4	4	4	4	3	4	4	4	3	38	95
R7	4	4	4	2	4	3	4	3	4	3	35	87.5
R8	3	3	3	3	3	3	3	3	3	3	30	75
R9	3	3	4	1	3	3	3	3	3	1	27	67.5
R10	2	3	3	4	3	2	4	3	4	4	32	80
R11	3	1	4	4	3	3	4	4	4	2	32	80
R12	4	3	3	2	4	3	3	3	3	3	31	77.5
R13	3	2	3	2	4	3	3	2	2	1	25	62.5
R14	4	3	3	3	3	4	3	4	3	2	32	80
R15	2	2	2	1	2	2	2	2	1	1	17	42.5
R16	4	4	4	4	4	4	4	4	4	4	40	100
R17	4	4	4	4	4	4	4	4	4	4	40	100
R18	4	4	4	4	4	4	4	4	4	4	40	100
R19	4	3	4	3	4	4	4	4	4	4	38	95
R20	4	4	4	3	3	4	4	3	4	4	37	92.5
R21	4	3	4	4	4	4	4	4	4	3	38	95
R22	4	3	4	4	4	4	4	4	3	4	38	95
R23	4	4	4	4	4	4	4	4	4	4	40	100
Total Keseluruhan												1880
Rata-Rata												81.73913

Berdasarkan hasil perhitungan skor System Usability Scale dari 23 repositen yang telah mengisi kuesioner didapatkan nilai rata-rata skor 81.73913 Berdasarkan nilai rata-rata tersebut kita dapat memperoleh tingkatan skor System Usability Scale sebagai berikut:

- Pada tingkatan NPS aplikasi mendapatkan tingkat Promoter
- Pada tingkatan Acceptability aplikasi mendapatkan tingkat Acceptable
- Pada tingkatan Adjectives aplikasi mendapatkan tingkat Excellent
- Pada tingkatan Grade aplikasi mendapatkan tingkat A

4. Kesimpulan

Berdasarkan penelitian yang telah dilakukan dapat disimpulkan bahwa penggunaan System Usability Scale (SUS) membantu dalam mengukur kegunaan dan kepuasan pengguna terhadap aplikasi. Desain antarmuka aplikasi LibrarySense mendapatkan skor rata-rata SUS yang cukup tinggi, yaitu sebesar 81.74. Dimana LibrarySense pada tingkat NPS mendapatkan tingkat Promoter, tingkat Acceptability mendapatkan tingkat Acceptable, tingkat Adjectives mendapatkan tingkat Excellent, dan tingkat mendapatkan tingkat A. Hal ini menunjukkan bahwa pengguna cenderung merasa puas dengan kegunaan dan pengalaman pengguna yang disajikan oleh aplikasi.

Daftar Pustaka

- [1] R. Vira Zwagery dan R. Santia Dewi, Pengaruh Kebisingan Terhadap Daya Ingat Pada Remaja. 2019.
- [2] A. Prayudi dan S. Yapis Dompou, "Studi Literatur : Penggunaan Model Analogi dalam Proses Pembelajaran," 2023. [Online]. Available: <http://journal.ainarapress.org/index.php/ainj>
- [3] K. A. C. Nanda dan I. P. G. H. Suputra, "Analisa UI dalam Perancangan Aplikasi Ekspor Hasil Laut 'FishGo' dengan Metode SUS.," Jurnal Nasional Teknologi Informasi dan Aplikasinya, vol. 1, no. 3, pp. 881-892, Mei 2023.
- [4] I. Angelica dan C. Nas, "Design UI / UX Prototype Aplikasi Pemesanan Produk Dimsky Berbasis Mobile Dengan Menggunakan Figma."
- [5] I. B. Karo Sekali, C. E. J. C. Montolalu, dan S. A. Widiara, "Perancangan UI/UX Aplikasi Mobile Produk Fashion Pria pada Toko Celcius di Kota Manado Menggunakan Design Thinking," Jurnal Ilmiah Informatika dan Ilmu Komputer (JIMA-ILKOM), vol. 2, no. 2, pp. 53-64, Sep. 2023, doi: 10.58602/jima-ilkom.v2i2.17.
- [6] I. K. Wardani, P. Utomo, A. Budiman, dan D. N. Amadi, "Pemanfaatan Metode Design Thinking dan Pengujian SUS untuk UI/UX Aplikasi Home Care Madiun Berbasis Android," Journal of Computer and Information Systems Ampera, vol. 4, no. 2, 2023, doi: 10.51519/journalcisa.v4i2.399.
- [7] M. Dermawan Mulyodiputro, V. Yoga, dan P. Ardhana, "Pengujian Usability Sistem Informasi Akademik (SISKA) Universitas Qamarul Huda Badaruddin Menggunakan System Usability Scale (SUS)," 2023.
- [8] M. Rafid Pratama, J. Umam, dan R. Yakok, "Usability Testing pada Aplikasi iJateng Menggunakan Metode System Usability Scale," Manajemen dan Teknologi Informasi, vol. 2, no. 1, 2024.

Halaman ini sengaja dibiarkan kosong

Pengaruh Penanganan Ketidakseimbangan Kelas pada Prediksi Cacat Perangkat Lunak dengan Teknik Oversampling

I Gusti Agung Ramananda Wira Dharma^{a1}, I Wayan Santiyasa^{a2}

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹agungramananda@gmail.com
²santiyasa@unud.ac.id

Abstract

Software defect prediction plays a vital role in SDLC testing by identifying modules prone to defects. However, imbalanced class distributions, where defect (minority) samples are outnumbered by non-defect ones, can hinder model performance. This study investigates the impact of oversampling techniques (SMOTE, ADASYN) on Naive Bayes classification for defect prediction. While the base Naive Bayes model achieved good overall accuracy (83%), it struggled with defect class recall (30%). Applying SMOTE and ADASYN improved recall (40% and 38%, respectively) but slightly lowered accuracy (77% and 80%). Future work will explore feature selection and deep learning approaches for potentially better performance.

Keywords: Software Defect Prediction, Classification, Naive Bayes, Oversampling, SMOTE, ADASYN

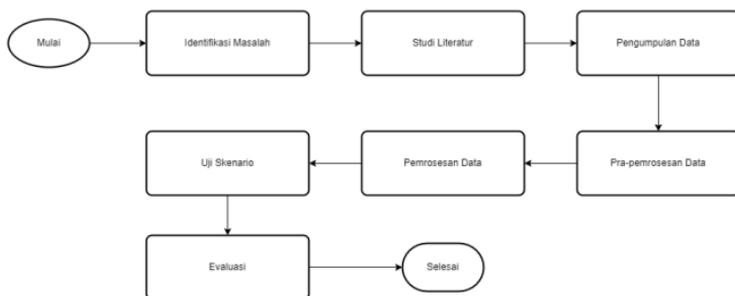
1. Pendahuluan

Prediksi cacat perangkat lunak merupakan salah satu hal yang paling membantu dalam fase pengujian SDLC (Software Development Life Cycle) dengan mengidentifikasi modul-modul yang rawan cacat dan memerlukan pengujian ektesif. Dengan cara ini, sumber daya pada pengujian dapat digunakan secara efisien tanpa melanggar batasan-batasan [1]. Saat ini, algoritma machine learning yang umum digunakan adalah K-Means, K Nearest Neighbors (KNN), Naive Bayes, Decision Tree, Support Vector Machine (SVM), dan Random Forest (RF) [2]. Efisiensi model prediksi pada cacat perangkat lunak sangat ditentukan oleh distribusi kelas dari data pelatihan [1]. Distribusi kelas digambarkan sebagai jumlah instance dari setiap kelas dalam dataset pelatihan. Jika jumlah instance yang termasuk dalam satu kelas jauh lebih banyak daripada jumlah instance yang termasuk dalam kelas lain, maka masalah tersebut dikenal sebagai masalah ketidakseimbangan kelas [3]. Kelas dengan lebih banyak instance disebut kelas mayoritas dan kelas dengan instance yang lebih sedikit disebut kelas minoritas [1]. Pada prediksi cacat perangkat lunak kelas non-cacat merupakan kelas mayoritas dan kelas cacat merupakan kelas minoritas. Ketidakseimbangan kelas dapat menyebabkan model yang tidak praktis dalam prediksi cacat perangkat lunak, karena kebanyakan kasus akan diprediksi sebagai non-cacat rawan [4]. Terdapat beberapa metode yang diusulkan untuk menangani masalah ketidakseimbangan kelas dari berbagai penelitian diantaranya, yakni pada penelitian Liu et al [5] menggunakan tiga jenis teknik sampling seperti resampling, spread subsampling dan SMOTE, dan lima jenis pengklasifikasi machine learning seperti C4.5, Naive Bayes, Logistic Regression, Support Vector Machine and Deep Learning untuk prediksi cacat. Diambil kesimpulan bahwa sampling sangat berpengaruh besar terhadap hasil prediksi cacat karena menghasilkan distribusi data yang berbeda untuk pelatihan model. Dan pada penelitian [6] mengusulkan metode yaitu Naive Bayes dengan integrasi teknik oversampling SMOTE dengan seleksi fitur Information Gain untuk solusi dari ketidakseimbangan kelas. Metode yang diusulkan meningkatkan kinerja dari Naive Bayes pada prediksi cacat perangkat lunak dengan diperoleh hasil nilai rata-rata AUC pada

model adalah 0.798, sedangkan model Naive Bayes yang memiliki nilai rata-rata AUC adalah 0.753.

Dari hasil penelitian-penelitian diatas, penelitian ini dilakukan dengan tujuan untuk mengetahui pengaruh dari penanganan ketidakseimbangan kelas dengan menggunakan teknik oversampling. Teknik oversampling yang diterapkan diantaranya SMOTE dan ADASYN dengan menguji performa dari model klasifikasi Naive Bayes. Dilakukan pengujian beberapa skenario pada data yang telah dilakukan oversampling dan yang tidak dilakukan oversampling. Evaluasi menggunakan beberapa metrik dan dilakukan perbandingan pada beberapa pengujian berdasarkan skenario dan metode-metode yang diterapkan pada model klasifikasi yang digunakan.

2. Metode Penelitian



Gambar 1. Alur Penelitian

Dalam penelitian yang dilakukan untuk membandingkan kinerja model klasifikasi antara menggunakan SMOTE dengan tanpa melakukan oversampling data pada software defect prediction ini dibagi menjadi 6 tahapan seperti yang dapat dilihat pada Gambar 1. Berikut penjelasan dari alur penelitian tersebut:

- a. Identifikasi Masalah
Pada tahapan dilakukan identifikasi latar belakang permasalahan yang akan diangkat pada penelitian. Tahapan ini sangat menentukan tahapan-tahapan selanjutnya seperti topik studi literatur yang akan dipelajari, data yang akan digunakan, dan metode yang akan digunakan pada penelitian.
- b. Studi Literatur
Tahapan ini mencari dan mempelajari literatur dari artikel pada jurnal seperti penelitian-penelitian terkait atau penelitian sebelumnya tentang topik yang sama dan buku yang berkaitan dengan persoalan dan metode yang digunakan untuk pendekatan teori yang akan digunakan pada penelitian.
- c. Pengumpulan Data
Pada tahapan ini bertujuan untuk mencari dan mengumpulkan data yang digunakan untuk membandingkan kinerja model klasifikasi pada penelitian ini. data yang digunakan adalah data sekunder yang diperoleh dari NASA Software Defect Dataset yang sudah dibersihkan. Dataset tersebut memiliki ketidakseimbangan kelas antara kelas minoritas (cacat) dan kelas mayoritas (non-cacat).
- d. Pra-pemrosesan Data
Pada tahapan ini dilakukan pembersihan pada data yang kosong sebelum dilakukan pemrosesan data dan normalisasi data untuk mengubah nilai menjadi kisaran 0 sampai 1.

- e. Pemrosesan Data
 Tahapan ini melakukan resampling data, yakni dengan oversampling data untuk penanganan untuk ketidakseimbangan data antara kelas minoritas (cacat) dan kelas mayoritas (non-cacat) menggunakan SMOTE (Synthetic Minority Over-sampling Technique), dan ADASYN (Adaptive Synthenic Sampling).
- f. Uji Skenario
 Skenario pengujian diatur menjadi beberapa skenario, yakni menerapkan oversampling dan yang tidak menerapkan overampling dengan menguji performa model klasifikasi Naive Bayes.
- g. Evaluasi
 Perhitungan performa model Naive Bayes yang digunakan berfokus pada masalah ketidakseimbangan kelas. Pada penelitian ini menggunakan Confusion Matrix sebagai evaluasi dengan melihat hasil recall dari kelas minoritas.

2.1. Data Penelitian

Dataset yang digunakan dalam penelitian ini adalah dataset NASA MDP Repository. Dataset tersebut memiliki ketidakseimbangan kelas antara kelas mayoritas (non-cacat) dan kelas minoritas (cacat). Dataset terdiri dari dua belas dataset dan dipilih 7 dataset diantaranya, yaitu CM1, JM1, KC1, dan PC1. Penjelasan dataset dapat dilihat dari tabel berikut.

Tabel 1. Dataset NASA MDP

Dataset	Sistem	Bahasa	Jumlah Atribut	Jumlah Instance
CM1	Instrumen pesawat ruang angkasa	C++	22	327
JM1	Sistem prediksi pendaratan secara real-time	C	22	10885
KC1	Manajemen penyimpanan data Lapangan	C++	22	2109
PC1	Perangkat lunak penerbangan satelit yang mengorbit bumi	C	38	1109

Dibawah ini merupakan contoh isi dataset CM1 :

Tabel 2. Isi Dataset CM1

Dataset	No	LOC_BLANK	BRANCH_COUNT	CALL_PAIRS	...	Defective
CM1	1	9	5	3	...	False
	2	19	3	2	...	False
	3	0	9	0	...	False

	327	10	17	1	...	False

2.2. SMOTE (Synthetic Minority Over-sampling Technique)

SMOTE [7] berdasar bahwa ruang fitur dari instance kelas minoritas serupa. Untuk setiap x^l instance yang terdapat pada kelas minoritas, SMOTE mencari tetangga terdekatnya dengan K -Nearest Neighbors dan satu tetangga secara acak dipilih sebagai x' . Selanjutnya nomer acak δ diantara 0 dengan 1 dihasilkan. Sample baru x_{new} dibuat :

$$x_{new} = x^l + (x' - x^l) \times \delta \tag{1}$$

2.3. ADASYN (Adaptive Synthetic)

ADASYN [8] menggunakan bobot distribusi untuk instance yang terdapat pada kelas minoritas berdasarkan pada tingkat kesulitan pembelajaran model. Sample yang dihasilkan dari instance pada kelas minoritas yang sulit dipahami dibandingkan dengan instance kelas minoritas yang lebih mudah dipahami. Algoritma ADASYN adalah sebagai berikut:

- a. Menentukan nilai parameter dari ADASYN, yaitu nilai dari maksimal toleran d_{th} , ketidakseimbangan kelas dan nilai level keseimbangan β
- b. Menghitung derajat keseimbangan

$$d = m_{minority} / m_{majority} \quad (2)$$

- c. Menghitung banyaknya instance sample sintetis yang akan dibuat untuk kelas minoritas

$$G = m_{majority} - m_{minority} \times \beta \quad (3)$$

- d. Menghitung rasio berdasarkan K-Nearest Neighbors menggunakan Euclidean distance

$$r_i = \frac{d_i}{K}, i = 1, \dots, m_{minority} \quad (4)$$

- e. Normalisasi r_i sehingga r'_i adalah distribusi kerapatan

$$r'_i = \frac{r_i}{\sum_{i=1}^{m_{minority}} r_i} \quad (5)$$

- f. Menghitung banyaknya instance sample sintetis yang perlu dibangkitkan untuk setiap instance kelas minoritas

$$g_i = r_i \times G \quad (6)$$

- g. Pembangkitan sample sintetis

$$s_i = x_i + (x_{z_i} - x_i) \times \lambda \quad (7)$$

Setiap instance pada kelas minoritas x_i , dipilih satu instance secara acak x_{z_i} menggunakan K-Nearest Neighbors dan λ adalah bilangan acak antara 0 dan 1.

2.4. Naïve Bayes

Naïve Bayes [9] merupakan metode dengan teknik prediksi probabilitas dengan penerapan teorema bayes yang mana diantara suatu fitur dengan fitur lain dalam satu data itu tidak saling berikatan. Metode ini merupakan salah satu bentuk sederhana untuk klasifikasi, persamaan dari Naïve Bayes dapat dilihat pada persamaan (8).

$$P_{(x|y)} = \frac{P_{(y|x)}P_{(x)}}{P_{(y)}} \quad (8)$$

Keterangan :

- x = Hipotesis data y
- y = Data kelas yang belum diketahui
- $P_{(x|y)}$ = Probabilitas hipotesis x berdasarkan kondisi y
- $P_{(x)}$ = Probabilitas hipotesis x
- $P_{(y|x)}$ = Probabilitas hipotesis y berdasarkan kondisi x
- $P_{(y)}$ = Probabilitas hipotesis y

2.5. Confusion Matrix

Confusion matrix adalah matrix yang menyatakan klasifikasi jumlah data uji yang benar dan jumlah data uji yang salah [10].

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Gambar 1. Confusion Matrix

Keterangan:

TP (True Positive) = Jumlah instance dari kelas 1 yang benar diklasifikasikan sebagai kelas 1
TN (True Negative) = Jumlah instance dari kelas 0 yang benar diklasifikasikan sebagai kelas 0
FP (False Positive) = Jumlah instance dari kelas 0 yang salah diklasifikasikan sebagai kelas 1
FN (False Negative) = Jumlah instance dari kelas 1 yang salah diklasifikasikan sebagai kelas 0

Persamaan confusion matrix untuk menghitung accuracy, recall, precision, dan f1-score dapat dilihat pada persamaan (9) dan (10).

$$\text{accuracy} = \frac{TP+TN}{\text{Total}} \tag{9}$$

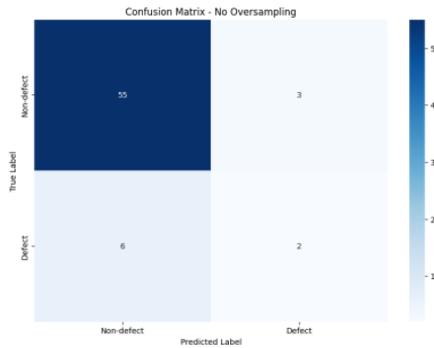
$$\text{recall} = \frac{TP}{TP+FN} \tag{10}$$

3. Hasil dan Diskusi

Penelitian ini menggunakan algoritma Naïve Bayes sebagai model klasifikasi. Dataset yang digunakan akan dibagi terlebih dahulu menjadi data latih dan data uji dengan pembagian skema 80% untuk data latih dan 20% untuk data uji. Dilakukan tiga scenario pengujian, diantaranya menggunakan model Naïve Bayes tanpa oversampling, Naïve Bayes dengan SMOTE, dan Naïve Bayes dengan ADASYN. Berikut merupakan hasil performa dari masing-masing model.

3.1. Hasil Performa Naïve Bayes

Berikut merupakan confusion matrix dari model Naïve Bayes tanpa menerapkan teknik oversampling pada salah satu dataset yang digunakan.



Gambar 2. Confusion Matrix dari Naive Bayes Tanpa Oversampling pada Dataset CM1

Dari confusion matrix diatas dapat dilihat bahwa model kurang sensitif dalam mengidentifikasi kelas minoritas atau cacat dan cenderung memberikan prediksi yang salah pada sampel yang sebenarnya cacat. Berikut merupakan tabel hasil akurasi dan hasil recall pada kelas cacat dari model pada seluruh dataset yang digunakan.

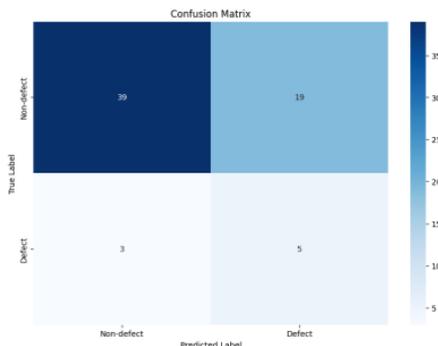
Tabel 3. Hasil Performa Model Naive Bayes

Dataset	Akurasi	Recall
CM1	0.86	0.25
JM1	0.78	0.19
KC1	0.78	0.34
PC1	0.89	0.43
Rata-Rata	0.83	0.30

Dari hasil dapat dilihat bahwa model memiliki akurasi yang tinggi dengan rata-rata 82%, akan tetapi model memiliki recall terhadap kelas minoritas yang rendah dengan rata-rata 30%.

3.2. Hasil Performa Naive Bayes + SMOTE

Berikut merupakan confusion matrix dari model Naive Bayes dengan menerapkan teknik oversampling SMOTE pada salah satu dataset yang digunakan.



Gambar 3. Confusion Matrix dari Naive Bayes+SMOTE pada Dataset CM1

Dari confusion matrix diatas dapat dilihat bahwa model masih kurang sensitif dalam mengidentifikasi kelas minoritas atau cacat dengan 5 sampel yang diprediksi cacat dengan benar dan 3 sampel yang seharusnya cacat tapi diprediksi tidak cacat. Berikut merupakan tabel hasil akurasi dan hasil recall pada kelas cacat dari model pada seluruh dataset yang digunakan.

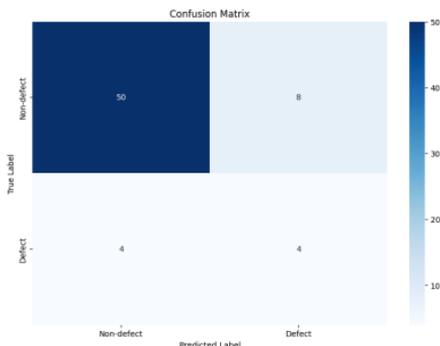
Tabel 4. Hasil Performa Model Naive Bayes+SMOTE

Dataset	Akurasi	Recall
CM1	0.67	0.62
JM1	0.78	0.22
KC1	0.76	0.36
PC1	0.88	0.43
Rata-Rata	0.77	0.40

Dari hasil dapat dilihat bahwa model memiliki akurasi yang tinggi dengan rata-rata 82%, akan tetapi model memiliki recall terhadap kelas minoritas yang rendah dengan rata-rata 40%.

3.3. Hasil Performa Naive Bayes + ADASYN

Berikut merupakan confusion matrix dari model Naive Bayes dengan menerapkan teknik oversampling ADASYN pada salah satu dataset yang digunakan.



Gambar 4. Confusion Matrix dari Naive Bayes+ADASYN pada Dataset CM1

Dari confusion matrix diatas dapat dilihat bahwa model kurang baik dalam mengidentifikasi kelas minoritas atau cacat dengan 5 sampel yang diprediksi cacat dengan benar dan 5 sampel yang seharusnya cacat tapi diprediksi tidak cacat. Berikut merupakan tabel hasil akurasi dan hasil recall pada kelas cacat dari model pada seluruh dataset yang digunakan.

Tabel 5. Hasil Performa Model Naive Bayes+ADASYN

Dataset	Akurasi	Recall
CM1	0.82	0.50
JM1	0.78	0.24
KC1	0.75	0.38
PC1	0.88	0.43
Rata-Rata	0.80	0.38

Dari hasil dapat dilihat bahwa model memiliki akurasi yang tinggi dengan rata-rata 80%, akan tetapi model memiliki recall terhadap kelas minoritas yang rendah dengan rata-rata 38%.

4. Kesimpulan

Berdasarkan hasil yang didapat, model Naive Bayes memiliki performa yang cukup baik dalam mengklasifikasikan data cacat perangkat lunak, dengan rata-rata akurasi mencapai 83%. Namun, model ini menunjukkan kelemahan dalam mengidentifikasi kelas minoritas (cacat), dengan rata-rata recall hanya 30%. Hal ini berarti model sering kali gagal memprediksi sampel yang sebenarnya cacat. Penerapan teknik oversampling SMOTE dan ADASYN menunjukkan peningkatan terhadap recall akan tetapi menurunkan akurasi secara keseluruhan. SMOTE meningkatkan rata-rata recall kelas minoritas menjadi 40% dan akurasi rata-rata 77% sedangkan ADASYN, di sisi lain, menghasilkan rata-rata recall 38% dan akurasi rata-rata 80%. Pada penelitian selanjutnya diharapkan dapat menerapkan seleksi fitur atau ekstraksi fitur dan optimasi pada algoritma klasifikasi yang digunakan atau menerapkan deep-learning sehingga mendapatkan performa dari model yang lebih baik.

Daftar Pustaka

- [1] I. Arora, V. Tetarwal, and A. Saha, "Open issues in software defect prediction," in *Procedia Computer Science*, Elsevier B.V., 2015, pp. 906–912. doi: 10.1016/j.procs.2015.02.161.
- [2] X. Dong, Y. Liang, S. Miyamoto, and S. Yamaguchi, "Ensemble learning based software defect prediction," *Journal of Engineering Research (Kuwait)*, vol. 11, no. 4, pp. 377–391, Dec. 2023, doi: 10.1016/j.jer.2023.10.038.
- [3] M. Galar, A. Fernandez, E. Barrenechea, H. Bustince, and F. Herrera, "A review on ensembles for the class imbalance problem: Bagging-, boosting-, and hybrid-based approaches," *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, vol. 42, no. 4, pp. 463–484, Jul. 2012. doi: 10.1109/TSMCC.2011.2161285.
- [4] T. M. Khoshgoftaar and K. Gao, "Feature selection with imbalanced data for software defect prediction," in *8th International Conference on Machine Learning and Applications, ICMLA 2009, 2009*, pp. 235–240. doi: 10.1109/ICMLA.2009.18.
- [5] Y. Liu, W. Zhang, G. Qin, and J. Zhao, "A comparative study on the effect of data imbalance on software defect prediction," in *Procedia Computer Science*, Elsevier B.V., 2022, pp. 1603–1616. doi: 10.1016/j.procs.2022.11.349.
- [6] S. A. Putri and R. S. Wahono, "Integrasi SMOTE dan Information Gain pada Naive Bayes untuk Prediksi Cacat Software," *Journal of Software Engineering*, vol. 1, no. 2, 2015, [Online]. Available: <http://journal.ilmukomputer.org>
- [7] Z. Zheng, Y. Cai, Y. Li, Z. Zheng, Y. Cai, and Y. Li, "Oversampling Method For Imbalanced Classification," 2015.
- [8] H. He, Y. Bai, E. A. Garcia, and S. Li, "ADASYN: Adaptive synthetic sampling approach for imbalanced learning," in *Proceedings of the International Joint Conference on Neural Networks*, 2008, pp. 1322–1328. doi: 10.1109/IJCNN.2008.4633969.
- [9] M. E. Lasulika, "Komparasi Naive Bayes, Support Vector Machine Dan K-Nearest Neighbor Untuk Mengetahui Akurasi Tertinggi Pada Prediksi Kelancaran Pembayaran Tv Kabel," *ILKOM Jurnal Ilmiah*, vol. 11, no. 1, pp. 11–16, May 2019, doi: 10.33096/ilkom.v11i1.408.11-16.
- [10] D. Normawati and S. A. Prayogi, "Implementasi Naive Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," 2021.

Halaman ini sengaja dibiarkan kosong

Analisis Performa Algoritma K-Nearest Neighbor dalam Klasifikasi Penyakit Tumor Otak

Komang Gede Bagus Devit Aditiya^{*1}, I Wayan Santiyasa^{*2}

^{*}Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹aditiya.2208561073@student.unud.ac.id
²santiyasa@unud.ac.id

Abstract

Brain tumor disease poses a significant health challenge globally, including in Indonesia. Detecting brain tumors early is crucial for effective treatment. In this study, we investigated the performance of the K-Nearest Neighbor (KNN) algorithm in classifying brain tumor disease using brain image data. Our findings reveal that the choice of K value significantly impacts the KNN algorithm's performance. The highest accuracy of 81% was achieved with K=3, while the lowest accuracy of 66% occurred at K=7. On average, across all scenarios, the accuracy was 72.8%. These results underscore the importance of selecting the appropriate K value for optimal classification accuracy in brain tumor disease using the KNN algorithm.

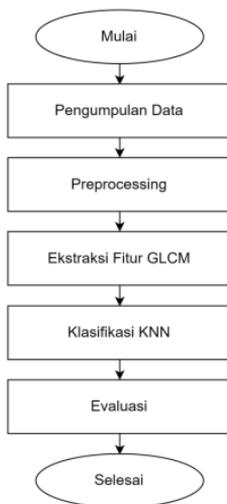
Keywords: Brain Tumor, Classifier, K-Nearest Neighbor, Grayscale, Accuracy

1. Pendahuluan

Tumor otak merupakan penyakit serius dan semakin mengkhawatirkan di seluruh dunia, termasuk Indonesia. Tumor otak bisa berasal dari pertumbuhan abnormal di otak itu sendiri (tumor primer) atau akibat penyebaran tumor (metastasis) ke bagian tubuh lain. Berbeda dengan tumor lainnya, tumor otak memiliki ciri khas yang unik. Meskipun tumor otak mungkin jinak dari sudut pandang histologis, tumor otak sering kali menjadi ganas karena lokasinya di dekat struktur vital atau di rongga tertutup yang sulit dijangkau. Tumor otak dikenal sebagai tumor paling ganas kedua setelah tumor darah (leukemia) [1], dan jumlah penderita tumor otak di Indonesia terus meningkat dalam beberapa tahun terakhir. Oleh karena itu, deteksi dini sangat penting dan perlu dilakukan sebelum tumor otak berkembang menjadi stadium lebih lanjut. Penting bagi tenaga medis dan masyarakat untuk terus mewaspadai gejala dan perubahan yang terjadi. Menghadapi tantangan ini, penting untuk mengembangkan metode klasifikasi yang efektif untuk diagnosis dan pengobatan tumor otak. Beberapa algoritma klasifikasi yang dapat digunakan antara lain K-nearest neighbours, support vector machine, dan convolutional neural network. Salah satu metode yang sering digunakan dalam klasifikasi adalah K-Nearest Neighbors (K-NN). Metode KNN merupakan metode klasifikasi data yang mencocokkan data uji dengan data latih berdasarkan jarak mayoritas terdekat. KNN merupakan metode klasifikasi sederhana yang mampu menampung data pelatihan dalam jumlah besar [2]. Penelitian sebelumnya tentang klasifikasi tumor otak pada tahun 2020 berjudul "Klasifikasi tumor otak menggunakan metode support vector machine pada gambar resonansi magnetik" [3]. Penelitian ini menggunakan metode support vector machine (SVM) untuk mengklasifikasikan data gambar. Data set yang digunakan adalah data gambar dengan format .jpg yang berjumlah 100 buah data. Penelitian ini mencapai akurasi maksimum sebesar 76% dan akurasi rata-rata sebesar 70%. Dengan menguji keakuratan penelitian ini, penulis berharap dapat menggunakan metode berbeda untuk mengklasifikasikan tumor otak. Perbedaan penelitian ini dengan penelitian sebelumnya terletak pada pendekatan algoritmik yang diterapkan. Penelitian ini menggunakan metode K tetangga terdekat dan menggunakan ekstraksi fitur grey level co-occurrence matriks (GLCM) untuk dataset yang berbeda. Hasil akhir dari penelitian ini adalah tingkat akurasi metode K-NN dalam mengklasifikasikan data citra ke dalam label tumor dan non tumor.

2. Metode Penelitian

Pada tahapan ini akan dijelaskan urutan langkah-langkah atau proses yang akan penulis lakukan selama melakukan penelitian. Input penelitian ini berupa data gambar tumor otak, kemudian data tersebut dibagi menjadi data latih dan data uji, serta dibagi menjadi dua label yaitu gambar dengan tumor otak dan gambar tanpa tumor. Data citra akan menggunakan Gray Level Co-occurrence Matrix (GLCM) untuk mengekstraksi fitur. Output yang akan dihasilkan merupakan evaluasi algoritma KNN dengan ekstraksi fitur GLCM. Proses penelitian ditunjukkan pada Gambar 1.



Gambar 1. Alur Penelitian

2.1. Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yaitu "Brain Tumor Classification" yang diperoleh dari website kaggle.com. Datanya berupa file berformat ".jpg" dengan pixel yang bervariasi. Data yang digunakan sebanyak 270 data gambar, dan data tersebut dibagi menjadi 2 label yaitu 135 data "glioma_tumor" dan 135 data "no_tumor". Kumpulan data dibagi menjadi data latih dan data uji. Data pelatihan menggunakan 90% dataset, sedangkan data pengujian menggunakan 10% sisanya.

2.2. Preprocessing

Pada tahap ini, data citra yang telah terkumpul akan diubah menjadi *grayscale* untuk menyederhanakan representasi warna. Selain itu, semua data akan disesuaikan ukurannya (*resize*) menjadi 200 x 200 piksel. Hal ini dilakukan untuk memastikan konsistensi ukuran dan memudahkan proses analisis, sehingga memungkinkan perbandingan yang lebih adil antara data yang berbeda.

2.3. Ekstraksi Fitur GLCM

GLCM (Gray Level Co-occurrence Matrix) adalah sebuah array yang mencatat seberapa sering pasangan piksel dengan tingkat keabuan tertentu muncul pada jarak (d) dan sudut (θ) tertentu. Biasanya, sudut yang dipertimbangkan adalah 0° , 45° , 90° , 135° , 180° , dan seterusnya, sementara jarak antar piksel umumnya diatur menjadi 1 piksel atau 2 piksel [4]. Dalam penelitian ini, empat atribut turunan dari GLCM digunakan dengan persamaan berikut:

a. Kontras

Kontras menggambarkan variabilitas intensitas antara nilai piksel yang paling terang dan paling gelap dalam pasangan piksel yang berdekatan. Perhitungan nilai kontras dapat dilakukan menggunakan persamaan nomor 1.

$$CON = \sum_{i,j} i - j^2 p_{i,j} \tag{1}$$

Keterangan:

i = baris ke- i dalam tingkat keabuan

j = kolom ke- j dalam tingkat keabuan

$p_{i,j}$ = probabilitas keabuan pada baris ke- i dan kolom ke- j

b. Entropi

Entropi mencerminkan tingkat ketidakteraturan dari distribusi nilai keabuan dalam citra. Ketika entropi memiliki nilai yang tinggi, ini menunjukkan bahwa struktur citra memiliki transisi nilai keabuan yang lebih teratur. Sebaliknya, nilai entropi yang rendah menandakan bahwa citra memiliki struktur yang kurang teratur atau lebih bervariasi dalam distribusi nilai keabuannya. Nilai Entropi dapat dilakukan menggunakan persamaan nomor 2.

$$EN = \sum_{i,j} p_{i,j} (-\ln(p_{i,j})) \tag{2}$$

Keterangan:

$p_{i,j}$ = Peluang keabuan baris ke - i , kolom ke - j

$\ln p_{i,j}$ = Logaritma natural dari $p_{i,j}$

c. Homogenitas

Homogenitas digunakan untuk mengevaluasi sejauh mana variasi intensitas dalam citra bersifat homogen atau seragam. Perhitungan nilai homogenitas (H) dapat dilakukan menggunakan persamaan nomor 3.

$$H = \sum_{i,j} \frac{p_{i,j}}{1+|i-j|} \tag{3}$$

Keterangan:

i = baris ke- i dalam tingkat keabuan

j = kolom ke- j dalam tingkat keabuan

d. Energi

Energi merupakan metrik yang mengindikasikan seberapa fokusnya pasangan piksel dengan tingkat keabuan tertentu di dalam matriks. Perhitungan energi (E) dapat dilakukan dengan menggunakan persamaan yang dinyatakan dalam nomor 4.

$$E = \sum_{i,j} (p_{i,j})^2 \tag{4}$$

i = baris ke- i dalam tingkat keabuan

j = kolom ke- j dalam tingkat keabuan

$p_{i,j}$ = probabilitas keabuan pada baris ke- i dan kolom ke- j

2.4. Klasifikasi K-NN

K-Nearest Neighbor (K-NN) merupakan algoritma klasifikasi yang memanfaatkan kedekatan antara data uji dengan data lainnya. Pada algoritma K-NN, ketika data mempunyai dimensi q maka dapat diukur jarak antara data tersebut dengan data lainnya. Jarak ini digunakan sebagai indikator kedekatan atau kemiripan antara data uji dan data latih. Nilai K pada K-NN mengacu pada jumlah data yang dipertimbangkan selama proses klasifikasi yang paling dekat dengan data uji [5]. Berikut adalah langkah-langkah dalam menghitung metode K-NN:

- Tetapkan nilai K ke jumlah tetangga terdekat yang akan dipertimbangkan.
- Hitung jarak Euclidean kuadrat antara data yang akan diklasifikasikan (contoh kueri) dan setiap sampel data yang ada.
- Urutkan objek berdasarkan jarak Euclidean, dari yang terkecil hingga yang terbesar.
- Kumpulkan kategori dari tetangga terdekat.
- Dengan menggunakan kelas mayoritas dari tetangga terdekat, nilai terhitung dari sebuah instance query dapat diprediksi.

2.5. Evaluasi

Pada penelitian ini, evaluasi akan dihasilkan setelah melakukan pengujian dengan menggunakan *Confusion Matrix*.

Tabel 1. Confusion Matrix

Label	Terklarifikasi Positif	Terklarifikasi Negatif
Positif	TP	FN
Negatif	FP	TN

Keterangan:

- True Positive (TP) : Jumlah total data positif yang diklasifikasikan dengan benar.
- True Negative (TN) : Jumlah total data negatif yang diklasifikasikan dengan benar.
- False Positive (FP) : Jumlah total data positif yang seharusnya negatif, tetapi salah diklasifikasikan sebagai positif.
- False Negative (FN) : Jumlah total data negatif yang seharusnya positif, tetapi salah diklasifikasikan sebagai negatif.

Dari Confusion Matrix, nilai Akurasi dapat dihitung menggunakan persamaan berikut.

$$Accuracy = \frac{TP + TN}{(TP + FP + TN + FN)} \quad (5)$$

3. Hasil dan Diskusi

Penelitian ini bertujuan untuk mengklasifikasi tumor otak menggunakan algoritma K-nearest neighbour (KNN) dengan ekstraksi fitur Gray Level Co-Occurrence Matrix (GLCM) yang diimplementasikan dalam bahasa pemrograman Python. Sebelum melakukan proses klasifikasi, data citra dari tumor otak akan diekstraksi fitur menggunakan GLCM. Fitur yang digunakan dari data citra, yaitu kontras, entropi, homogenitas, dan energy. Proses klasifikasi tumor otak dilakukan dengan mencari label terdekat berdasarkan jarak antara data uji dengan data latih yang ada, kemudian menentukan kelas mayoritas dari tetangga terdekat. Selain itu, penelitian ini bertujuan untuk mengevaluasi dampak pemilihan fitur parameter K pada algoritma K-Nearest Neighbor terhadap hasil klasifikasi. Nilai K yang dieksperimen pada penelitian ini antara lain $k=1$, $k=3$, $k=5$, $k=7$, dan $k=9$.

3.1. Hasil Ekstraksi Fitur GLCM

Pada tahap ekstraksi fitur menggunakan Gray Level Co-occurrence Matrix (GLCM), empat atribut utama yang digunakan adalah, kontras, entropi, homogenitas, dan energi. Masing-masing fitur ini dihitung berdasarkan distribusi spasial intensitas piksel dalam citra grayscale. Kontras mengukur variasi intensitas, entropi menilai ketidakteraturan, homogenitas menunjukkan keseragaman, dan energi mencerminkan konsentrasi pasangan piksel tertentu. Hasil ekstraksi fitur ini kemudian digunakan sebagai input dalam algoritma K-Nearest Neighbor (KNN) untuk klasifikasi tumor otak. Berikut adalah hasil dari ekstraksi fitur menggunakan GLCM.

No	energi_0	homogenity_0	entropy_0	contrast_0	energi_05	homogenity_05	entropy_05	contrast_05	energi_10	homogenity_10	entropy_10	contrast_10	energi_15	homogenity_15	entropy_15	contrast_15	label
0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
11	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
16	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
17	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
18	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
19	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
20	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
21	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
22	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
23	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
24	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
25	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
26	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
27	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
28	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
29	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
30	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
31	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
32	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
33	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
34	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
35	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
36	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
37	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
38	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
39	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
40	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
41	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
42	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
43	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
44	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
45	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
46	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
47	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
48	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
49	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal
50	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Normal

Gambar 2. Hasil Ekstraksi Fitur dengan GLCM

3.2. Pengujian Model

3.2.1. Pengujian dengan Nilai K Bernilai 1

Pengujian yang menggunakan parameter K=1, hasilnya menunjukkan rata-rata akurasi sebesar 77%, dengan 21 data terklasifikasi dengan benar dan 6 data terklasifikasi secara tidak tepat. Detail hasil pengujian dapat dilihat pada tabel 2.

Tabel 2. Hasil Pengujian Nilai K = 1

Kelas	Terklarifikasi Positif	Terklarifikasi Negatif
Positif	10	4
Negatif	2	11

3.2.2. Pengujian dengan Nilai K Bernilai 3

Pengujian yang menggunakan parameter K=3, hasilnya menunjukkan rata-rata akurasi sebesar 81%, dengan 22 data terklasifikasi dengan benar dan 5 data terklasifikasi secara tidak tepat. Detail hasil pengujian dapat dilihat pada tabel 3.

Tabel 3. Hasil Pengujian Nilai K = 3

Kelas	Terklarifikasi Positif	Terklarifikasi Negatif
Positif	11	3
Negatif	2	11

3.2.3. Pengujian dengan Nilai K Bernilai 5

Tabel 4. Hasil Pengujian Nilai K = 5

Kelas	Terklarifikasi Positif	Terklarifikasi Negatif
Positif	9	5
Negatif	3	10

3.2.4. Pengujian dengan Nilai K Bernilai 7

Pengujian yang menggunakan parameter K=7, hasilnya menunjukkan rata-rata akurasi sebesar 66%, dengan 18 data terklasifikasi dengan benar dan 9 data terklasifikasi secara tidak tepat. Detail hasil pengujian dapat dilihat pada tabel 5.

Tabel 5. Hasil Pengujian Nilai K = 7

Kelas	Terklarifikasi Positif	Terklarifikasi Negatif
Positif	8	6
Negatif	3	10

3.2.5. Pengujian dengan Nilai K Bernilai 9

Pada pengujian yang menggunakan parameter K=9, hasilnya menunjukkan rata-rata akurasi sebesar 70%, dengan 19 data terklasifikasi dengan benar dan 8 data terklasifikasi secara tidak tepat. Detail hasil pengujian dapat dilihat pada tabel 6.

Tabel 6. Hasil Pengujian Nilai K = 9

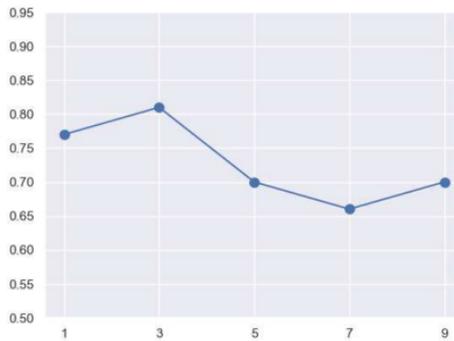
Kelas	Terklarifikasi Positif	Terklarifikasi Negatif
Positif	9	5
Negatif	3	10

3.3. Hasil Klasifikasi

Berdasarkan pengujian yang sudah dilakukan, pemilihan nilai K berpengaruh pada performa algoritma KNN. Akurasi maksimum yang tercapai adalah 81% saat menggunakan nilai K=3, sementara akurasi minimum adalah 66% dengan nilai K=7. Rata-rata akurasi dari semua skenario adalah 72,8%, sebagaimana tercantum dalam Tabel 7. Terlihat bahwa semakin besar nilai K, akurasi cenderung menurun, menunjukkan bahwa nilai K memengaruhi akurasi. Oleh karena itu, disarankan untuk mencapai akurasi optimal dengan menggunakan nilai K=3.

Tabel 7. Rekapitan Hasil Pengujian

No	Pengujian	K	Akurasi
1	Pengujian 1	1	77%
2	Pengujian 2	3	81%
3	Pengujian 3	5	70%
4	Pengujian 4	7	66%
5	Pengujian 5	9	70%
	Rata-rata		72,8%



Gambar 3. Grafik Akurasi Pengujian

4. Kesimpulan

Berdasarkan penelitian yang telah dilakukan, penulis dapat mengimplementasikan algoritma K-Nearest Neighbor (KNN) untuk melakukan klasifikasi penyakit tumor otak menggunakan data citra otak. Hasil pengujian menunjukkan bahwa nilai K (jumlah tetangga terdekat) memiliki dampak yang sangat signifikan terhadap performa algoritma KNN dalam proses klasifikasi tumor otak. Penulis menemukan bahwa performa algoritma KNN terbaik dalam klasifikasi tumor otak diperoleh pada nilai $K=3$ dengan akurasi sebesar 81%. Sebaliknya, nilai akurasi terendah terjadi saat nilai $K=7$, dengan akurasi mencapai 66%. Performa algoritma KNN dengan nilai K lebih dari 3 cenderung menurun dalam melakukan klasifikasi tumor otak. Secara rata-rata, akurasi dari seluruh skenario pengujian adalah sebesar 72,8%. Penelitian ini menemukan pentingnya pemilihan nilai K yang tepat dalam melakukan klasifikasi menggunakan algoritma *K-Nearest Neighbor*.

Daftar Pustaka

- [1] S. Heranurweni, B. Destyningtias, and A. K. Nugroho, "Klasifikasi Pola Image pada Pasien Tumor Otak Berbasis Jaringan Syaraf Tiruan (Studi Kasus Penanganan Kuratif Pasien Tumor Otak)," *eLEKTRIKAL*, Vol. 10, No. 2, 2018.
- [2] A. Yuliani, A. Labellapansa, and A. Yulianti, "Klasifikasi Citra Daun Kelapa Sawit Yang Terkena Dampak Hama Menggunakan Metode K-Nearest Neighbor," *Seminar Nasional Informatika Medis*, 2019
- [3] A. S. Febrianti, T. A. Sardjono, and A. F. Babgei, "Klasifikasi Tumor Otak pada Citra Magnetic Resonance Image dengan Menggunakan Metode Support Vector Machine," *Jurnal Teknik ITS*, Vol. 9, No. 1, 2020.
- [4] D. P. Pamungkas, "Ekstraksi Citra Menggunakan Metode GLCM dan KNN untuk Identifikasi Jenis Anggrek (Orchidaceae)," *Innovation in Research of Informatics*, Vol. 1, No. 2, 2019.
- [5] A. A. D. Halim, S. Anraeni, "Analisis Klasifikasi Dataset Citra Penyakit Pneumonia Menggunakan Metode K-Nearest Neighbor (KNN)," *Indonesian Journal of Data and Science*, Vol. 2, No. 1, 2021.

Halaman ini sengaja dibiarkan kosong

Implementasi Algoritma Yolo untuk Deteksi Kebusukan pada Sayur Kembang Kol

Alexander Ibrahim^{a1}, I Wayan Supriana^{a2}

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹alexander772004@gmail.com
² wayan.supriana@unud.ac.id

Abstract

This research utilizes the YOLOv8 algorithm to detect spoilage in cauliflower vegetables. Image data was collected from Google, processed using Roboflow, and tested using Google Colab. The study results indicate an accuracy of 59%, recall of 58%, and MAP of 60%. The YOLOv8 algorithm significantly contributes to image recognition and visual data processing. Additionally, the article discusses the application of the YOLOv8 algorithm for object detection in 360-degree panoramic images. The training process was conducted to recognize objects in the images, and evaluation was performed using a confusion matrix and mAP50. The evaluation results demonstrate the model's good performance in object recognition. Several references cited in the article are also included.

Keywords: You Only Look Once (YOLO), Google Colab, Roboflow

1. Pendahuluan

Sayur kembang kol merupakan salah satu komoditas pertanian yang populer di Indonesia. Namun, kembang kol mudah mengalami kebusukan, yang dapat menyebabkan kerugian ekonomi yang signifikan bagi petani dan pedagang. Oleh karena itu, penting untuk mengembangkan metode yang efektif untuk mendeteksi kebusukan pada kembang kol sesegera mungkin. Algoritma YOLOv8 telah menjadi salah satu metode yang populer dalam deteksi objek dalam gambar, dengan kemampuannya yang efisien dan andal. Dalam penelitian ini, algoritma YOLOv8 digunakan untuk dua tujuan yang berbeda namun terkait. Pertama, penelitian ini bertujuan untuk mendeteksi kebusukan pada sayur kembang kol, sementara penelitian kedua fokus pada deteksi objek dalam gambar panorama 360 derajat. Kedua penelitian ini mengevaluasi kinerja model menggunakan metrik yang berbeda, seperti akurasi, recall, dan Mean Average Precision (MAP), serta confusion matrix dan mAP50. Hasil dari kedua penelitian menunjukkan bahwa algoritma YOLOv8 memberikan kontribusi yang signifikan dalam pengenalan objek dalam gambar dan pemrosesan data visual. Dengan demikian, penelitian ini memberikan wawasan yang berharga tentang penggunaan algoritma YOLOv8 dalam konteks deteksi objek pada berbagai jenis gambar.

2. Metode Penelitian

Penelitian ini bersifat eksperimental dengan objek penelitian yaitu kembang kol yang diperoleh dengan cara mengambil foto secara langsung dari Google. Citra yang telah didapatkan selanjutnya akan diproses melalui beberapa tahapan seperti anotasi, pra- pemrosesan, augmentasi (pelebaran dataset), serta mengekspor dataset ke dalam format YOLOv8. Proses tersebut dilakukan secara otomatis dengan memanfaatkan platform roboflow. Setelah dataset ter-ekspor ke dalam format YOLOv8 langkah selanjutnya masuk ke platform Google Collab memanggil model YOLOv8 dengan cara menjalankan code cloud, kemudian mengupload show donload code. Kemudian, dilanjutkan ke beberapa proses yaitu pelatihan, pendeteksian dan evaluasi dataset. Output dari seluruh proses ini adalah pemodelan dataset dengan nilai confusion metric seperti akurasi, recall, precision, serta f1 score.

Berikut merupakan tahapan dalam metodologi penelitian yang dijelaskan melalui diagram alur penelitian dan dapat dilihat pada Gambar 1.



Gambar 1. Diagram Alur Penelitian

a. Studi Literatur

Pada tahap ini mempelajari berbagai referensi atau rujukan yang mendukung penelitian bersumber dari internet, jurnal, tugas akhir, penelitian dan modul -modul yang berkaitan dengan penerapan algoritma YOLOv8 untuk sistem deteksi pada dataset baru citra kembang kol serta penelitian - penelitian sejenis.

b. Pengumpulan Gambar

Tahap ini merupakan persiapan yang harus dilakukan sebelum melakukan penelitian. Tahap yang dimaksud adalah pengumpulan gambar sayur kembang kol. Data yang tersaji merupakan data primer yang diambil dari mesin Google. Gambar yang diambil terdiri dari 3 tingkat yaitu segar, mulai membusuk, dan juga busuk.

c. Pemrosesan Data

Setelah data mentah dikumpulkan, langkah selanjutnya adalah tahap pemrosesan data. Data yang sudah ada dimasukkan ke dalam roboflow untuk menjalani beberapa proses seperti:

- Anotasi gambar atau pemberian nama gambar berdasarkan tingkatannya yaitu segar, mulai membusuk, dan juga busuk.
- Pra pemrosesan data dilakukan dengan mengekstrak bagian - bagian yang tidak penting seperti latar belakang yang tidak relevan, objek orientasi otomatis sekaligus mengubah ukuran menjadi 640 x 640 piksel.
- Ekspor data ke dalam format YOLOv8

d. Penerapan Algoritma YOLOv8

Setelah data melewati proses augmentasi, tahap selanjutnya adalah melakukan proses penerapan algoritma YOLOv8 pada data tindakan ini merupakan proses training data yang ada untuk mengupload file Lantana zip dan data yaml ke dalam codingan Google Collab lalu melakukan proses runing di dalam coding setelah itu menginput nilai epoch berapa kali data akan dilatih. Setelah hasil nilai keluar lanjut ke proses pendeteksian dengan cara memasukan gambar yang ingin dideteksi ke codingan dalam Google Collab setelah itu model akan segera membaca dan menampilkan hasil deteksi. Setelah objek benar maka tahap selanjutnya melihat dan mencocokkan hasilnya. Jika objek yang ada sudah terdeteksi sama dengan jenis warna pada lantana camara untuk tiap klasifikasi atau kelasnya, maka dapat dikatakan proses training nya berhasil

e. Evaluasi Model pada Validasi Data

Setelah data melewati proses augmentasi, tahap selanjutnya adalah melakukan proses penerapan algoritma YOLOv8 pada data tindakan ini merupakan proses training data yang ada untuk mengupload show donload code ke dalam codingan Google Collab lalu melakukan proses runing di dalam coding, setelah itu menginput nilai epoch berapa kali data akan dilatih. Setelah hasil nilai keluar lanjut ke proses pendeteksian dengan cara memasukan gambar yang ingin dideteksi ke codingan dalam Google Collab setelah itu model akan segera membaca dan menampilkan hasil deteksi. Setelah objek benar maka tahap selanjutnya melihat dan mencocokkan hasilnya. Jika objek yang ada sudah terdeteksi sama dengan jenis warna pada lantana camara untuk tiap klasifikasi atau kelasnya, maka dapat dikatakan proses trainingnya berhasil.

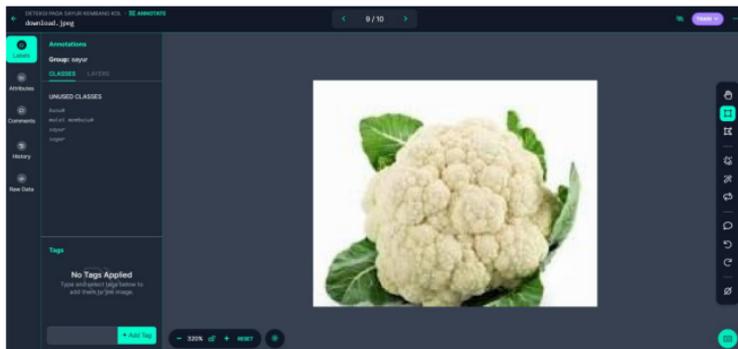
f. Kesimpulan

Tahap ini model melakukan prediksi pada validasi data, dan hasil prediksi bisa dibandingkan dengan label sebenarnya. Evaluasi ini memberikan wawasan tentang seberapa baik model dapat melakukan prediksi pada data yang belum pernah dilihat sebelumnya hasil ini dapat dilihat dari hasil best epoch dengan nilai rata-rata accuracy, precision, recall dan MAP yang dihasilkan oleh Google Collab.

3. Hasil dan Diskusi

3.1 Implementasi YOLOv8

Tahap Implementasi YOLO dimulai dengan mempersiapkan dataset jenis sayur yaitu kembang kol dengan total gambar sebanyak 10 gambar. Gambar yang dikumpulkan kemudian dimasukan ke dalam roboflow dilakukan beberapa perlakuan seperti anotasi gambar atau pemberian nama kelas pada objek, hal ini dapat dilihat pada Gambar 2



Gambar 2. Contoh Anotasi Objek

Setelah melalui tahap anotasi lanjut ke tahap preprocessing di mana tahapan ini dilakukan dengan cara mengekstrak bagian - bagian yang tidak penting seperti latar belakang yang tidak relevan, objek orientasi diatur otomatis sekaligus mengubah ukuran menjadi 640 x 640 piksel. Seperti yang diperlihatkan pada Gambar 3.

```
# Pip install method (recommended)

!pip install ultralytics==8.0.196

from IPython import display
display.clear_output()

import ultralytics
ultralytics.checks()
```

Gambar 3. Tampilan source code untuk mengunduh YOLOv8

Setelah mengunduh YOLOv8 ke dalam Google Collab, selanjutnya mengunduh Show Download Code yang didapatkan di Platform roboflow. Hal ini perlu dilakukan agar memungkinkan Google Collab untuk mengakses ke data atau berkas yang diperlukan untuk melatih ataupun menguji model. Hal ini dapat dilihat pada Gambar 4 dimana Show Download Code.

```
!pip install roboflow

from roboflow import Roboflow
rf = Roboflow(api_key="4A9p4v14qAt2CxYGgXxb")
project = rf.workspace("alex-yxgvp").project("deteksi-pada-sayur-kembang-kol")
version = project.version(1)
dataset = version.download("yolov8")
```

Gambar 4. Tampilan Show Download Code

Langkah selanjutnya masuk ke tahap pelatihan data yang dapat dilihat pada Gambar 5, dimana tahap yolo task=detect mode=train model=YOLOv8s.pt data={dataset.location}/data.yaml epochs=25 imgsz=800 plots=True, bagian ini menggunakan program YOLOv8 untuk melaksanakan training data. model=YOLOv8s.pt, menentukan model awal yang akan digunakan untuk pelatihan. Dalam hal ini, YOLOv8s.pt adalah model YOLOv8 versi ringan yang sudah disediakan. data={dataset.location}/data.yaml, menentukan lokasi file konfigurasi data pelatihan Anda. Ganti {dataset.location} dengan nama folder aktual dataset Anda. File.yaml ini berisi informasi penting tentang dataset, seperti lokasi gambar, label objek, dan lainnya. epochs=25, menentukan jumlah iterasi pelatihan. Model akan dilatih selama 25 iterasi. imgsz=800, menentukan ukuran gambar yang akan digunakan selama pelatihan. Dalam hal ini, gambar akan diubah ukurannya menjadi 800x800 piksel. plots=True, menampilkan grafik pelatihan selama proses berlangsung.

```
%cd {HOME}

!yolo task=detect mode=train model=yolov8s.pt data={dataset.location}/data.yaml epochs=25 imgsz=800 plots=True
```

Gambar 5. Tampilan Source Code Untuk Training Data

```
data.yaml X
1 names:
2 - busuk
3 - mulai membusuk
4 - segar
5 nc: 3
6 roboflow:
7   license: CC BY 4.0
8   project: deteksi-pada-sayur-kembang-kol
9   url: https://universe.roboflow.com/alex-yxgvp/deteksi-pada-sayur-kembang-kol/dataset/1
10  version: 1
11  workspace: alex-yxgvp
12 test: test/images
13 train: train/images
14 val: valid/images
15
```

Gambar 6. Tampil Data Yaml

Setelah melakukan training data, langkah berikutnya melakukan pendeteksian objek seperti terlihat pada Gambar 7. Proses ini diperlukan untuk menguji apakah data train berhasil dikenali oleh model atau tidak. Proses ini dilakukan dengan menjalankan perintah python task=detect. Untuk %cd {HOME}, Ini akan mengubah direktori kerja ke direktori home Anda di Google Colab. ! yolo task=detect mode=predict model={HOME}/runs/detect/train/weights/best.pt conf=0.25 source={dataset.location}/test/images save=True, bagian ini menggunakan program YOLOv8 untuk melakukan deteksi objek. Deteksi objek, mencari dan mengenali objek tertentu dalam gambar. Mode prediksi, menggunakan model yang sudah dilatih sebelumnya untuk memprediksi objek dalam gambar baru. model={HOME}/runs/detect/train/weights/best.pt, menentukan lokasi file bobot model terlatih Anda. Bobot model, parameter yang dipelajari model selama proses pelatihan dan berisi informasi penting untuk melakukan deteksi objek. conf=0.25, mengatur ambang kepercayaan minimum. Model hanya akan menampilkan deteksi dengan skor kepercayaan minimal 25% (0.25). Skor kepercayaan, Nilai antara 0 dan 1 yang menunjukkan seberapa yakin model terhadap deteksi objek. source= {dataset.location}/test/images, Menentukan lokasi gambar yang ingin Anda deteksi objeknya. Ganti {dataset.location} dengan nama folder aktual dataset Anda. save=True, menyimpan hasil deteksi ke dalam gambar baru. Gambar ini akan memiliki kotak pembatas yang digambar di sekitar objek yang terdeteksi beserta labelnya.

```
%cd {HOME}
!yolo task=detect mode=predict model={HOME}/runs/detect/train/weights/best.pt conf=0.25 source={dataset.location}/test/images save=True
```

Gambar 7. Tampilan Source Code Untuk Menjalankan Deteksi Objek

3.2 Performance Metric

Evaluasi pelatihan dengan epoch atau iterasi 50 dapat dilihat pada Gambar 8, dimana pada ketiga kelas dengan nilai image 360 merujuk pada gambar panorama 360 derajat. Gambar panorama 360 derajat adalah gambar yang mencakup seluruh lingkaran visual sehingga dapat dilihat ke segala arah secara vertikal maupun horizontal. Untuk jumlah instances pada setiap kelas berbeda-beda karena model deteksi objek dilatih untuk mengenali dan menempatkan objek-objek yang muncul dalam gambar panorama 360 derajat sehingga dari masing-masing kelas secara berturut-turut model mampu mengenali sebanyak 3, 1, 2, dimana hasil ini bergantung pada karakteristik objek, kerumitan dataset, ukuran dataset pelatihan, kualitas anotasi, parameter pelatihan seperti jumlah Epoch dapat mempengaruhi jumlah instances. Untuk mAP50 (mean average precision) atau metrik yang mencakup precision dan recall pada bagian ambang batas, mAP dihitung dengan cara menghitung nilai rata-rata dari precision sehingga mAP pada

rentang nilai ambang deteksi 0,5 hingga 0,95 (mAP50-95) dapat memberikan gambaran lebih tentang evaluasi kinerja model secara menyeluruh di berbagai tingkat ambang deteksi, sehingga nilai mAP yang didapat pada masing-masing kelas di atas 0.22%.

```

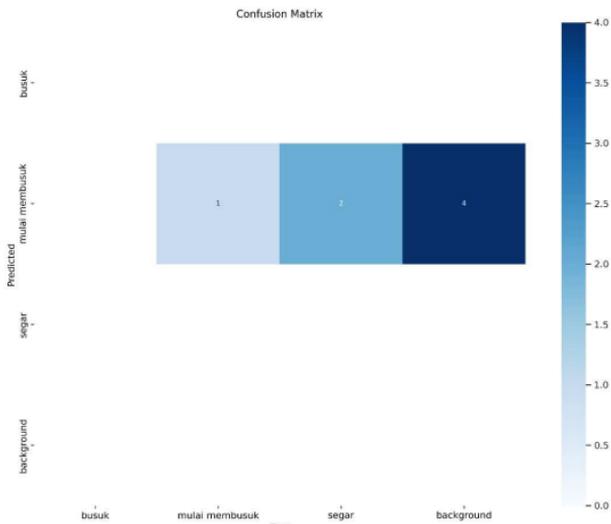
yolo task-detect mode-val model=[KEM]/run/detect/train/weights/best.pt data=[dataset.location]/data.yaml

!./detect
INFO: Validating model [KEM]... Python 3.10.12 torch 2.2.1 torchvision 0.18.0 (CUDA 12, 12580MB)
model summary (fused): 108 layers, 1122024 parameters, 0 gradients, 36.4 GFLOPs
val: Scanning [dataset.location]/deteksi pada sayur-kembang-kol 1/valid/labels-cache... 1 images, 0 backgrounds, 0 corrupt: 100% 1/1 [00:00<, 117/s]
/usr/lib/python3.10/site-packages/processing/paper_task.py:66: RuntimeWarning: os.fork() was called, os.fork() is incompatible with multithreaded code, and fork is multithreaded, so this
self.pid = os.fork()
class:      images  instances  mAP@F  R  mAP50  mAP50-95  SPS  1/1 [00:00<@0MB, 1.101/s]
  all       1         1      0.249  0.497  0.451  0.242
  mulai membusuk  1         1      0.497  0.995  0.497  0.266
  segar     1         1         0      0.228  0.186

Speed: 6.2ms preprocess, 30.0ms inference, 0.0ms loss, 100.0ms postprocess per image
Results saved to run/detect/val
Learn more at https://docs.ultralytics.com/models/yol
    
```

Gambar 8. Tampilan Parameter Hasil Training Dengan Google Colab

Hasil evaluasi confusion matrix yang digunakan untuk membandingkan hasil prediksi model dengan ground truth (kebenaran sebenarnya) yang didapat dari perlakuan train data train data, Terdapat 3 jenis yaitu segar, mulai membusuk dan juga busuk. Untuk yang bernilai true positive (TP), model membuat prediksi benar sesuai dengan kenyataan ditandai dengan Predicted dan True sama, masing-masing kelas bernilai 1, 2, dan 4. Hal ini dapat dilihat pada Gambar 9 tampilan confusion matrix.



Gambar 9. Funfusion Matrix

4. Kesimpulan

Penelitian ini menggunakan algoritma YOLOv8 untuk deteksi objek dalam gambar dengan hasil yang menjanjikan. Penelitian pertama fokus pada deteksi kebusukan pada sayur kembang kol dengan akurasi 59%, recall 58%, dan MAP 60%. Sementara itu, penelitian ini juga membahas deteksi objek dalam gambar panorama 360 derajat dengan evaluasi menggunakan confusion

matrix dan mAP50, menunjukkan kinerja model yang baik. Algoritma YOLOv8 terbukti memberikan kontribusi penting dalam pengenalan gambar dan pemrosesan data visual dalam kedua konteks penelitian.

Daftar Pustaka

- [1] A. Muliantara and N. A. S. ER, "Pengembangan Pengklasifikasi jenis Tanaman Menggunakan pendekatan backpropagation Dan Nguyen-Widrow," *Jurnal Ilmu Komputer*, <https://ojs.unud.ac.id/index.php/jik/article/view/18779> (accessed May 6, 2024).
- [2] Bagus Janapriya, A. A. G. (2023). Pengenalan Jenis Rambu Lalu Lintas menggunakan Metode YOLO V5. *JELIKU (Jurnal Elektronik Ilmu Komputer Udayana)*, 11(4). <https://doi.org/10.24843/jik.2023.v11.i04.p32>.
- [3] Anggotra, P., & Muliantara, A. (2023). Deteksi Relief Candi Borobudur Menggunakan Metode Template Matching. *JELIKU (Jurnal Elektronik Ilmu Komputer Udayana)*, 12(1). <https://doi.org/10.24843/jik.2023.v12.i01.p03>.
- [4] Keraf, M. P. S., Sinlae, A. A. J., & Batarius, P. (2024). Pemodelan Warna Pada Dataset Baru Citra Bunga Lantana Camara Menggunakan Algoritma Yolov5. *Jurnal Mnemonic*, 7(1), 108-117.
- [5] Saputra, D. H., & Imran, B. (2023). Object Detection Untuk Mendeteksi Citra Buah-Buahan Menggunakan Metode Yolo. *Jurnal Kecerdasan Buatan dan Teknologi Informasi*, 2(2), 70-80.
- [6] Lusiana, L., Wibowo, A., & Dewi, T. K. (2023). Implementasi Algoritma Deep Learning You Only Look Once (YOLOv5) Untuk Deteksi Buah Segar Dan Busuk. *Paspalum: Jurnal Ilmiah Pertanian*, 11(1), 123-130.
- [7] HWB, N., Mailoa, E., & Purnomo, H. D. (2020). Deteksi Buah untuk Klasifikasi Berdasarkan Jenis dengan Algoritma CNN Berbasis YOLOv3. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 4(3), 476-481.
- [8] Wulandari, Dwi, et al. "Implementasi Algoritma YOLOv5 untuk Deteksi Kerusakan pada Buah Mangga." *Jurnal Sains dan Teknologi Agroindustri* 8.1 (2023): 1-10.
- [9] Sari, Rini Puspita, et al. "Implementasi Algoritma YOLO untuk Deteksi Kerusakan pada Daun Jeruk." *Jurnal Agroteknologi Universitas Lampung* 16.1 (2021): 1-8.

Halaman ini sengaja dibiarkan kosong

Deteksi Hate Speech pada Unggahan Media Sosial dengan Naive Bayes Menggunakan Seleksi Fitur Chi-Square

Putu Steven Belva Chan^{a1}, Ida Ayu Gde Suwiprabayanti Putra^{a2}

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana

Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia

¹stevenbelvachann@gmail.com

²agsuwiprabayantiputra@unud.ac.id

Abstract

In the digital age, social media's pervasive use has revolutionized global communication but also introduced challenges like hate speech. This study proposes a Multinomial Naive Bayes model optimized with Chi-square feature selection to detect hate speech efficiently from large-scale social media data. Leveraging machine learning, this approach aims to combat harmful content by identifying relevant text features crucial for distinguishing hate speech from non-hate speech. The study utilizes TF-IDF for feature extraction and Chi-square for feature selection, showing significant performance improvements in hate speech detection. The Chi-square feature selection model yielded average precision, recall, F1-score, and accuracy values of 92%, 92%, 91%, and 92% respectively. In contrast, the model without feature selection achieved values of 89%, 89%, 88%, and 89% for the same metrics. Results demonstrate enhanced accuracy, precision, recall, and F1-score across various hate speech categories.

Keywords: Hate Speech, Naive Bayes, TF-IDF, Chi-square

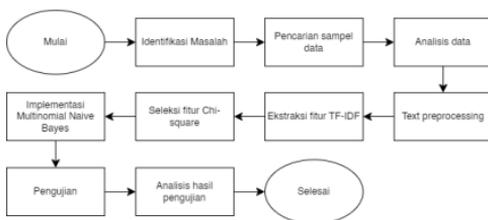
1. Pendahuluan

Di era digital saat ini, eksistensi media sosial telah membawa dampak signifikan dalam memfasilitasi komunikasi antar individu di seluruh dunia. Namun, bersamaan dengan kebebasan berekspresi yang ditawarkan oleh platform media sosial, kita juga menghadapi tantangan baru terkait penyalahgunaan informasi dan konten yang merugikan, seperti hate speech. Penggunaan media sosial secara massal menghasilkan volume data yang besar dan kompleks, membentuk apa yang kita kenal sebagai Big Data [1]. Data-data ini mencakup beragam informasi termasuk teks, gambar, video, dan lainnya yang diposting setiap detik. Di tengah kebebasan berbicara, kita juga melihat peningkatan dalam kasus hate speech, yang merupakan bentuk ujaran yang menghasut, menyerang, atau mengancam individu atau kelompok berdasarkan karakteristik seperti ras, agama, orientasi seksual, atau latar belakang budaya. Pentingnya penanggulangan hate speech di masyarakat semakin terasa, terutama mengingat dampaknya yang merusak kedamaian sosial dan kohesi komunitas. Hate speech dapat memicu konflik, memperburuk polarisasi sosial, dan menciptakan lingkungan online yang tidak aman dan tidak ramah [2]. Oleh karena itu, diperlukan pendekatan yang efektif untuk mendeteksi dan mengatasi hate speech secara proaktif di platform media sosial. Dalam menghadapi tantangan penanggulangan hate speech secara efektif, salah satu solusi yang ditawarkan adalah menggunakan metode klasifikasi teks berbasis machine learning, seperti Naive Bayes yang dioptimalkan dengan seleksi fitur Chi-square. Naive Bayes adalah algoritma klasifikasi yang populer dalam analisis teks karena sederhana dan efisien. Berdasarkan penelitian sebelumnya oleh Kurnia, dkk, didapatkan bahwa algoritma Naive Bayes mengungguli akurasi algoritma CNN [3]. Sementara itu, seleksi fitur digunakan untuk memilih subset optimal fitur untuk konstruksi model. Proses seleksi fitur menghitung skor dari setiap fitur yang mungkin berdasarkan teknik seleksi fitur tertentu, dan kemudian mengidentifikasi 'K' fitur terbaik[4]. Chi-square merupakan salah satu seleksi fitur yang

dapat membantu mengidentifikasi fitur-fitur teks yang paling relevan dan informatif untuk membedakan antara konten hate speech dan non-hate speech. Pada penelitian ini, penguji akan mengembangkan model deteksi hate speech menggunakan algoritma Multinomial Naive Bayes dengan seleksi fitur Chi-square yang dapat mengolah data dalam skala besar (Big Data) dari media sosial dengan efisien dan akurat. Dengan memanfaatkan teknik-teknik machine learning dan analisis data, diharapkan penelitian ini dapat memberikan kontribusi dalam upaya melindungi lingkungan online dari konten yang merugikan dan mempromosikan komunikasi yang inklusif dan bermartabat di dunia digital saat ini.

2. Metode Penelitian

2.1. Desain Penelitian



Gambar 1. Bagan Alur Penelitian

Penelitian ini diawali dengan mengidentifikasi terkait masalah yang ingin diselesaikan kemudian dilanjutkan dengan pencarian sampel data terhadap teks postingan di media sosial di kaggle.com yang nantinya akan dilakukan proses klasifikasi. Setelah data terkumpul, tahapan selanjutnya adalah melakukan proses text preprocessing. Kemudian, akan dilakukan perhitungan ekstraksi fitur menggunakan Term Frequency-Inverse Document Frequency (TF-IDF). Setelah ekstraksi fitur, selanjutnya adalah melakukan pelatihan model Multinomial Naive Bayes dengan menggunakan fitur yang telah diekstraksi dan diseleksi pada tahap selanjutnya. Tahap terakhir adalah melakukan analisis terhadap hasil pengujian yang didapat dari pengujian dengan model yang telah dibuat.

2.2. Pengumpulan Data

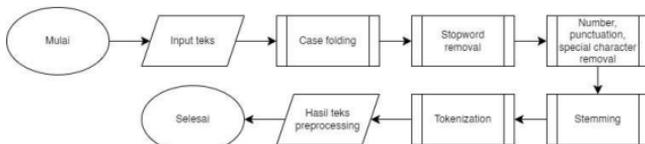
Data yang digunakan pada penelitian ini adalah data sekunder yang diambil dari Kaggle yaitu data hate speech dan abusive teks twitter. Data ini terdiri dari 13.169 data teks dan setiap record data dikelompokkan ke dalam 12 kategori ujaran kebencian. Berdasarkan target ujaran kebencian yaitu individu dan kelompok. Berdasarkan isi ujaran kebencian yaitu ras, agama, fisik, dan gender. Berdasarkan tingkat ujaran kebencian yaitu lemah, sedang, dan kuat.

Tabel 1. Gambaran Dataset

id	text	HS	Abusive	HS_Individual	HS_Race	HS_Moderate	...
1	Kalimat 1	1	0	1	1	0	...
2	Kalimat 2	0	1	0	0	1	...
3	Kalimat 3	1	0	0	1	1	...

2.3. Text Preprocessing

Text preprocessing adalah tahap awal yang biasa dilakukan saat melakukan klasifikasi teks. Text preprocessing adalah prosedur mengolah data tekstual tergantung pada beberapa kriteria untuk memperoleh teks murni[5]. Tahap ini akan menghasilkan teks yang siap digunakan ke dalam model machine learning. Tahap preprocessing pada penelitian ini ditunjukkan oleh gambar 2.



Gambar 2. Tahap Preprocessing Teks

Tahapan pertama dalam text preprocessing adalah melakukan case folding, yaitu mengubah semua huruf menjadi huruf kecil atau lowercase. Setelah dilakukan case folding, tahapan selanjutnya adalah menghilangkan karakter selain alfabet seperti tanda baca, angka, dan karakter spesial. Kemudian, akan dilakukan proses stemming pada data teks. Stemming dalam preprocessing teks adalah proses untuk memperoleh kata dasar (stem) dari sebuah kata dengan menghilangkan prefiks dan afiks kata tersebut[5]. Tahapan terakhir adalah melakukan tokenisasi seluruh kata pada setiap kalimat pada data teks.

2.4. Ekstraksi Fitur

Metode ekstraksi fitur yang digunakan dalam penelitian ini adalah Term Frequency-Inverse Document Frequency (TF-IDF). Pada klasifikasi teks, kata-kata yang terdapat pada dokumen merupakan fitur dari dokumen tersebut. Salah satu algoritma yang digunakan untuk mengekstraksi fitur ini adalah TF-IDF. Algoritma ini mencatat nilai yang menunjukkan seberapa penting sebuah kata pada dokumen. Prinsip kerja algoritma ini adalah jika sebuah kata sering muncul dalam sebuah dokumen tapi tidak pada dokumen lainnya, maka kata tersebut dapat dikatakan sebagai pembeda yang baik untuk klasifikasi. Nilai TF-IDF didapatkan dengan menghitung TF (term frequency) yang dikalikan dengan IDF (inverse document frequency).

Term frequency (TF) adalah proses untuk menghitung jumlah kemunculan sebuah kata dalam dokumen. Persamaan matematis dari TF dapat dilihat pada persamaan 1[6].

$$TF(t, d) = \frac{\text{jumlah term dalam dokumen } d}{\text{jumlah kata dalam dokumen } d} \quad (1)$$

Inverse Document Frequency (IDF) adalah proses untuk menghitung nilai yang menunjukkan seberapa penting suatu term secara global dalam suatu koleksi dokumen. IDF memberi bobot lebih pada term yang sering muncul pada satu dokumen tertentu tapi jarang muncul pada dokumen secara global. Persamaan matematis dari IDF dapat dilihat pada persamaan 2[6].

$$IDF(t) = \log\left(\frac{N}{df(t)}\right) \quad (2)$$

Dimana N adalah jumlah total dokumen dalam koleksi dan $df(t)$ adalah jumlah dokumen yang mengandung term t dalam koleksi.

Rumus TF-IDF menggabungkan Term Frequency (TF) dan Inverse Document Frequency (IDF) untuk memberikan bobot pada setiap term dalam suatu dokumen dalam konteks seluruh koleksi dokumen[6].

$$TFIDF(t, d) = TF(t, d) \times IDF(t) \quad (3)$$

2.5. Multinomial Naive Bayes

Multinomial Naive Bayes merupakan model pembelajaran probabilitas yang berasal dari teori Bayes dan tergolong dalam model supervised learning. Model ini banyak digunakan dalam Pengolahan Bahasa Alami (NLP) dan beroperasi berdasarkan prinsip frekuensi kata, yang mengukur kemunculan kata-kata dalam sebuah dokumen. Model ini menjelaskan dua aspek: kehadiran atau ketiadaan sebuah kata dalam teks, dan frekuensi kemunculan kata tersebut dalam dokumen. Persamaan Multinomial Naive Bayes dapat dilihat pada persamaan 4.

$$P(C|D) = P(C) \prod_{i=1}^n P(W_i|C) \quad (4)$$

Probabilitas kelas C muncul dalam dokumen D, dilambangkan sebagai $P(C|D)$, diturunkan dari persamaan 5, yang menghitung total jumlah kata dalam dokumen, yang dilambangkan sebagai Nc.

$$P(C) = \frac{Nc}{N} \quad (5)$$

Persamaan 6 menghitung probabilitas kelas C, $P(C)$, dengan membagi total jumlah dokumen oleh jumlah dokumen kelas C, Nc.

$$P(w_i|C) = \frac{\text{count}(w_i|C)+1}{\text{count}(C)+|V|} \quad (6)$$

$P(w_i|c)$ adalah probabilitas bersyarat kata ke-i dalam kelas C. $\text{count}(w_i|C)$ adalah frekuensi kata ke-i dalam kelas C. $\text{count}(C)$ adalah total jumlah kata dalam kelas C, sedangkan $|V|$ adalah jumlah total kata unik di semua kelas.

2.6. Chi-square

Chi-square adalah salah satu algoritma seleksi fitur berbasis statistik yang paling efisien[7]. Formula Chi-Square terkait dengan fungsi seleksi fitur berbasis informasi yang mencoba untuk memahami bahwa fitur terbaik t_k untuk kategori c_i adalah yang paling berbeda dalam himpunan contoh positif dan negatif dari kelas c_i [8]. Persamaan matematis Chi-square dapat dilihat pada persamaan 7.

$$\text{Chi-square}(t_k, c_i) = \frac{N(AD-CB)^2}{(A+C)(B+D)(A+B)(C+D)} \quad (7)$$

Dimana N adalah jumlah total dokumen dalam dataset, A adalah jumlah dokumen dalam kategori c_i yang mengandung fitur t_k , B adalah jumlah dokumen yang mengandung fitur t_k dalam kategori lain, C adalah jumlah dokumen dalam kategori c_i yang tidak mengandung fitur t_k , D adalah jumlah dokumen yang tidak mengandung fitur t_k dalam kategori lain.

2.7. Evaluasi

Evaluasi dilakukan dengan melakukan testing pada model yang telah dibangun. Pada penelitian ini, dataset yang telah melalui proses *preprocessing* akan dibagi menjadi data training dan data testing masing-masing sebanyak 80% dan 20%. Kemudian dilakukan ekstraksi fitur dengan Term Frequency-Inverse Document Frequency (TF-IDF) dan seleksi fitur dengan algoritma Chi-square. Dua jenis model dibangun, satu model menggunakan seleksi fitur dan satu model lainnya tanpa seleksi fitur. Dari skenario tersebut, akan dilakukan testing pada masing-masing model untuk melakukan prediksi pada teks di setiap kategori. Dari hasil testing akan dilihat perbandingan performa model yang menggunakan seleksi fitur dan tanpa seleksi fitur. Performa yang ditinjau dalam penelitian ini adalah tingkat accuracy, recall, precision, dan F1-score. Perhitungan accuracy, recall, precision, dan F1-score dapat dihitung menggunakan rumus berikut[9]:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (8)$$

No.	Proses	Hasil
5	Tokenization	["film", "bagus", "kesan", "jalan", "cerita"]

3.3. Ekstraksi fitur

Setelah melalui tahap preprocessing, selanjutnya akan dilakukan ekstraksi fitur dengan algoritma Term Frequency-Inverse Frequency Document (TF-IDF). Untuk melakukan ekstraksi fitur TF-IDF, penulis menggunakan bantuan modul TfidfVectorizer dari library scikit-learn dan menggunakan bahasa pemrograman Python.

```
tfidf = TfidfVectorizer()
tfidf.fit(train_data['text'])

train_X_vec = tfidf.transform(train_data['text'])
val_X_vec = tfidf.transform(val_data['text'])
```

Gambar 4. Implementasi TF-IDF

3.4. Chi-square

Setelah dilakukan ekstraksi fitur TF-IDF, dapat dilakukan seleksi fitur dengan algoritma Chi-square, menggunakan bantuan modul SelectKBest dan chi2 dari library scikit-learn dengan jumlah *top feature* sebanyak 5000 fitur.

```
chi2_features = SelectKBest(chi2, k=5000)
X_kbest_features = chi2_features.fit_transform(X_tfidf_array, y)
```

Gambar 5. Implementasi Chi-square

3.5. Pemodelan Multinomial Naive Bayes

Model Multinomial Naive Bayes dibangun dengan bantuan modul MultinomialNB dari library scikit-learn menggunakan bahasa pemrograman Python. Model akan dilatih dengan data training dari ekstraksi fitur TF-IDF tanpa seleksi fitur Chi-square dan dengan seleksi fitur Chi-square.

```
mnb = MultinomialNB()
mnb.fit(train_X_vec, train_y)

mnb_chi = MultinomialNB()
mnb_chi.fit(train_X_vec_chi, train_y)
```

Gambar 6. Implementasi Model MNB

3.6. Performa Model

Tabel 3. Perbandingan Performa Model

	Tanpa seleksi fitur Chi-square				Dengan seleksi fitur Chi-square			
	Precision	Recall	F1-Score	Accurac y	Precision	Recall	F1-Score	Accurac y
HS	81%	81%	81%	81%	89%	89%	89%	89%
Abusive	87%	87%	87%	87%	92%	92%	92%	92%

	Tanpa seleksi fitur Chi-square				Dengan seleksi fitur Chi-square			
	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy
HS_Individual	79%	80%	78%	80%	86%	84%	82%	84%
HS_Group	88%	89%	86%	89%	90%	90%	88%	90%
HS_Religion	95%	95%	94%	95%	96%	96%	96%	96%
HS_Race	96%	96%	94%	96%	97%	97%	96%	97%
HS_Physical	98%	98%	97%	98%	98%	98%	97%	98%
HS_Gender	95%	98%	96%	98%	98%	98%	97%	98%
HS_Other	82%	82%	81%	82%	87%	87%	85%	87%
HS_Weak	79%	80%	78%	80%	85%	84%	81%	84%
HS_Moderate	87%	88%	85%	88%	90%	90%	87%	90%
HS_Strong	96%	97%	95%	97%	98%	98%	97%	98%
Rata-rata	89%	89%	88%	89%	92%	92%	91%	92%

Berdasarkan Tabel 3, dapat dilihat bahwa nilai precision, recall, F1-score, dan accuracy dari model Naive Bayes dengan mengimplementasikan seleksi fitur Chi-square di setiap kategori yang ada lebih besar dibandingkan model Naive Bayes tanpa seleksi fitur Chi-square. Nilai akurasi tanpa seleksi fitur di bawah 90% mengalami kenaikan yang cukup signifikan setelah dilakukan seleksi fitur. Data ini membuktikan bahwa performa model Naive Bayes dengan ekstraksi fitur TF-IDF dan seleksi fitur Chi-square lebih baik dari model tanpa seleksi fitur Chi-square.

4. Kesimpulan

Penelitian ini menunjukkan bahwa penggunaan metode Multinomial Naive Bayes yang dioptimalkan dengan seleksi fitur Chi-square mampu meningkatkan performa dalam mendeteksi hate speech pada media sosial. Model dengan seleksi fitur Chi-square memberikan nilai rata-rata precision, recall, F1-score, dan accuracy masing-masing sebesar 89%, 89%, 88%, dan 89%. Sementara itu, model tanpa seleksi fitur memberikan nilai masing-masing sebesar 92%, 92%, 91%, dan 92%. Hasil evaluasi menunjukkan peningkatan signifikan dalam metrik evaluasi seperti precision, recall, F1-score, dan accuracy pada berbagai kategori ujaran kebencian setelah penerapan seleksi fitur Chi-square.

Daftar Pustaka

- [1] N. A. Ghani, S. Hamid, I. A. T. Hashem, E. Ahmed, "Social media big data analytics: A survey", *Computers in Human Behavior*, vol. 101, p. 417, 2019.
- [2] D. J. Ningrum, Suryadi, D. E. C. Wardhana, "Kajian Ujaran Kebencian Di Media Sosial", *Jurnal Ilmiah KORPUS*, vol. 2, no. 3, p. 243, 2018.
- [3] Y. Kurnia, E. D. Kusuma, L. W. Kusuma, Suwitno, W. Apringdi, "Perbandingan Naive Bayes dan CNN yang Dioptimasi PSO pada Identifikasi Berita Hoax Politik Indonesia", *Binary Digital - Technology*, vol. 6, no. 3, p. 346, 2024.
- [4] S. T. Ikram, A. K. Cherukuri, "Intrusion detection model using fusion of chi-square feature selection and multi class SVM", *Journal of King Saud University - Computer and Information Sciences*. vol. 29, p. 463, 2017.
- [5] M. F. Karaca, "Effects of Preprocessing on Text Classification in Balanced and Imbalanced Datasets", *Ksii Transactions on Internet and Information Systems*, vol. 18, no. 3, p. 595, 2024.
- [6] K. P. Harmandini, K. Muslim L, "Analysis of TF-IDF and TF-RF FeatureExtraction on Product Review Sentiment", *Sinkron: Jurnal dan Penelitian Teknik Informatika*, vol. 8, no. 2, p. 933, 2024.
- [7] C. Jin, T. Ma, R. Hou, M. Tang, Y. Tian, A. Al-Dhelaan, & M. Al-Rodhaan, "Chi-Square Statistics Feature Selection Based on Term Frequency and Distribution for Text Categorization", *IETE Journal of Research*, vol. 61, no. 4, pp. 351-362, 2015.
- [8] S. Bahassine, A. Madani, M. Al-Sarem, M. Kissi, "Feature selection using an improved Chi-

- square for Arabic text classification", *Journal of King Saud University - Computer and Information Sciences*. vol. 32, no. 2, pp. 225–231, 2020.
- [9] Mas'udah E, Wahyuni ED, and Anjani A, "Analisis sentimen: Pemindahan ibu kota Indonesia pada twitter", *Jurnal Informatika dan Sistem Informasi*, vol. 1, no. 2, pp. 397-401, 2020.

Rancang Model Ontologi dalam Representasi Digital Loloh Cemcem Penglipuran

Putu Chandra Mayoni, Ida Bagus Gede Dwidasmara

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
1mayoni.2208561111@student.unud.ac.id
dwidasmara@unud.ac.id

Abstract

Bali is renowned for its diverse cultural heritage, including traditional herbal drinks like Loloh Cemcem from Penglipuran Village. This immunity-boosting beverage dates back to the Majapahit era but gained popularity during the COVID 19 pandemic. Despite being recorded as a Communal Intellectual Property, information about Loloh Cemcem is not well-structured, hindering knowledge dissemination. This research aims to address this issue by designing an ontology model to represent knowledge about Loloh Cemcem digitally. The Methodology approach is employed, consisting of specifying, knowledge acquisition, conceptualization, integration, implementation using Protégé software, evaluation through SPARQL queries, and documentation. The ontology model captures concepts, properties, relationships, and individuals related to Loloh Cemcem, facilitating structured information access and preservation of this traditional beverage's knowledge.

Keywords: Ontologi, Loloh Cemcem Penglipuran, Methodology, Protégé, SPARQL Query

1. Pendahuluan

Bali merupakan salah satu pulau yang terletak di Indonesia, pulau ini terkenal dengan julukan Pulau Dewata. Bali dengan keanekaragaman budaya, tradisi, adat istiadat, dan kesenian menjadikan daerah ini banyak digemari oleh masyarakat Indonesia. Tak hanya wisatawan dalam negeri, wisatawan luar pun tertarik akan keanekaragaman yang dimiliki Bali, hal ini terlihat pada peningkatan jumlah pengunjung yang kian melonjak. Hal tersebut dapat dilihat dari peningkatan jumlah pengunjung, baik dari dalam negeri maupun mancanegara di setiap tahunnya [3]. Salah satu kabupaten di Bali yang terkenal adalah Bangli. Kabupaten Bangli tidak terlepas dari objek wisatanya yang sangat indah yaitu Desa Penglipuran. Di tahun 2016, desa Penglipuran telah terpilih menjadi desa terbersih di dunia [1]. Tak hanya itu, desa penglipuran juga terkenal akan minuman berkhasiatnya yaitu loloh cemcem. Loloh cemcem merupakan jamu tradisional yang memiliki manfaat salah satunya untuk menjaga imunitas tubuh, hal ini terbukti dengan penelitian yang dilakukan sebelumnya oleh Cahyaningrum dan Sudaryanti pada tahun 2021. Penelitian tersebut membuktikan bahwa, pada masa COVID-19 loloh cemcem banyak dikonsumsi oleh masyarakat untuk menjaga imunitas tubuh mereka agar dapat menangkal masuknya virus ke dalam tubuh (Cahyaningrum & Sudaryanti, 2021). Loloh cemcem telah ada sejak zaman Majapahit dan mulai dikenal oleh masyarakat umum di tahun 2014 sampai dengan 2015. Minuman loloh ini diproduksi oleh 5 unit keluarga secara besar[5]. Loloh cemcem telah tercatat sebagai salah satu bagian dari Kekayaan Intelektual Komunal (KIK) khususnya jenis Pengetahuan Tradisional (PT) [2]. Meskipun telah tercatat pada Kekayaan Intelektual Komunal (KIK), namun informasi mengenai loloh cemcem belum direpresentasikan secara terstruktur. Hal ini menyebabkan para wisatawan kesulitan dalam mencari informasi mengenai loloh cemcem. Untuk itu diperlukan sebuah solusi yang dapat membantu masyarakat dalam mendapat informasi dan pengetahuan yang terstruktur dan sistematis. Ontologi merupakan sebuah metode yang digunakan untuk menguraikan informasi secara semantik dan mampu dalam mengorganisasikan penempatan pada sekumpulan sumber informasi secara terstruktur. Ontologi mampu dalam memfasilitasi pengembangan sistem manajemen pengetahuan yang memungkinkan transisi dari

pendekatan yang berfokus pada dokumen menuju pendekatan yang berorientasi pada pengetahuan. Dengan ontologi, pengetahuan dan informasi dapat direpresentasikan dalam bentuk yang saling terhubung, sehingga memudahkan proses penggabungan dan pemanfaatan kembali pengetahuan secara baik. Selain itu, ontologi juga merupakan tulang punggung atau pondasi dari sebuah web semantik[8]. Dengan adanya ontologi pada web semantik, dapat menjadikan sebuah website menjadi lebih pintar serta memungkinkan komputer untuk menangkap makna atau arti di balik data yang diperoleh dan tidak hanya membaca data mentah saja [8-9]. Untuk membangun sebuah model ontologi maka diperlukan sebuah metode yang tepat. Dalam penelitian ini, metode yang akan digunakan yaitu Methodology. Methodology ini memiliki kelebihan untuk mendeskripsikan setiap aktivitas secara rinci serta dapat melakukan *reusability* ontologi yang telah disusun sebelumnya untuk pengembangan sistem atau aplikasi lanjutan [10]. Penelitian ini bertujuan untuk mengatasi permasalahan pada informasi yang belum terstruktur dengan baik melalui penerapan rancangan model Ontologi pada representasi informasi dan pengetahuan mengenai loloh cemcem Penglipuran.

1.1 Loloh Cemcem

Loloh merupakan minuman tradisional yang terbuat dari rempah-rempah dan dikonsumsi oleh masyarakat untuk menyegarkan dan menjaga kesehatan tubuh mereka. Terdapat salah satu loloh yang terkenal di kabupaten Bangli yaitu loloh cemcem. Loloh Cemcem ini merupakan loloh khas tradisional Penglipuran yang terbuat dari rempah-rempah seperti daun cemcem, gula aren, asam, kelapa muda, cabai dan garam [6]. Pada proses pembuatannya tidak harus menggunakan mesin namun juga dapat dibuat dengan cara yang manual. Cara pembuatannya dapat dilakukan dengan mencuci terlebih dulu daun cemcem yang didapat hingga bersih selanjutnya menumbuk daun cemcem hingga menjadi hancur kemudian akan disaring untuk memisahkan antara daun dan sari-sari yang sudah didapatkan. Selanjutnya daun cemcem tersebut akan dicampurkan dengan air dan bumbu rujak yang sebelumnya sudah dibuat hingga mendidih [11]. Loloh cemcem tidak hanya memiliki satu varian saja, namun terdapat 5 varian yaitu manis, asam, pahit, pedas dan asin. Untuk varian manis pada loloh ini dibuat dengan menambahkan gula sedangkan untuk varian asam akan ditambahkan buah asam atau sering disebut dengan tamarin. Varian pahit diperoleh langsung dari daun cemcem itu sendiri sedangkan untuk varian asin terbuat dari tambahan garam pada loloh cemcem [12]. Bukan hanya varian yang beraneka ragam, loloh cemcem juga memiliki manfaat yang beragam yaitu menjaga imunitas tubuh, menjaga kesehatan tubuh, meredakan panas dalam, menambah nafsu makan, membantu mengatasi permasalahan buang air besar dan air kecil, menyegarkan tubuh, membantu dalam menjaga pencernaan agar tetap lancar dan menurunkan tekanan darah[4,12,16]. Berbagai pengetahuan mengenai loloh cemcem tersebut sangat penting untuk diketahui oleh masyarakat agar minuman tradisional ini tetap dapat lestari dari budaya kita serta membantu dalam memperkenalkan lebih luas lagi mengenai loloh khas Penglipuran ini kepada seluruh masyarakat Indonesia dan luar negara, karena minuman tradisional ini dapat dikonsumsi oleh siapa saja dan memiliki berbagai macam manfaat untuk tubuh manusia.

1.2 Ontology

Ontologi merupakan sebuah metode untuk merepresentasikan domain pengetahuan tertentu secara jelas dan eksplisit. Ontologi mendefinisikan konsep-konsep dalam domain tersebut dengan memberikan makna, properti dan hubungan antar konsep. Melalui pendefinisian ini, pengetahuan dalam satu domain dapat dikumpulkan dan terstruktur membentuk sebuah basis pengetahuan. Sehingga ontologi memungkinkan pengetahuan di suatu bidang dapat dimodelkan dan direpresentasikan dalam format yang dapat dimengerti dan diproses oleh komputer [8]. Ontologi biasanya dimanfaatkan untuk mendukung implementasi web semantik, agar komputer dapat mengerti konteks dan makna dari input pengguna.

1.3 Web Semantik

Web Semantik merupakan model generasi ketiga pada halaman *website*, yang memiliki keunggulan lebih jika dibandingkan dengan generasi yang sebelumnya. Web semantik mampu untuk menghasilkan dan membaca informasi dengan skala yang luas [10]. Konsep dari web

semantik adalah "web of data". Konsep tersebut menerapkan ontologi web pada sebuah halaman *website* yang mengutamakan penggabungan data [9]. Tujuan dari web semantik yaitu untuk memudahkan komputer dalam memahami makna dari data yang diperoleh. Dengan begitu pengguna (manusia) dan komputer dapat saling bekerja sama. Selain itu, web semantik dapat menjadikan sebuah *website* menjadi lebih pintar dikarenakan didalamnya terdapat sebuah basis pengetahuan dengan bentuk ontologi sebagai inti. Teknologi web semantik bertujuan untuk mengoptimalkan performa server yang dimanfaatkan untuk melakukan pemuatan serta mengoperasikan aplikasi bisnis pada suatu situs *website* [9].

1.4 Methodology

Methodology adalah metode yang digunakan membangun sebuah model ontologi. Kelebihan dari methodology yaitu dapat mendeskripsikan setiap aktivitas secara rinci dan dapat melakukan *reusability* ontologi yang telah disusun sebelumnya untuk pengembangan sistem atau aplikasi lanjutan[10].

1.5 SPARQL

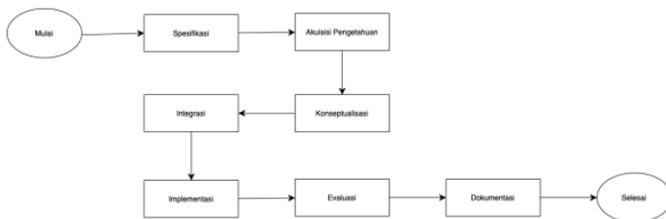
SPARQL Protocol and RDF Query Language merupakan standar yang diluncurkan oleh W3C untuk melakukan pertukaran query dan bahasa untuk mendapatkan data dari sumber website yang ada pada dokumen RDF (Resource Description Framework) dan OWL (Web Ontology Language) [8].

1.6 Protégé

Protégé merupakan perangkat yang digunakan untuk merancang ontologi dan melakukan query dengan mengimplementasikan SPARQL dan memudahkan pengguna dalam pengembangan prototipe. Pada Protégé, pengguna dapat mengakses berbagai fungsi yang tersedia melalui antarmuka grafis (GUI) yang menampilkan tab-tab terpisah untuk setiap bagian dan fitur utamanya [13].

2. Metode Penelitian

Pada penelitian ini, metode yang digunakan adalah Methodology. Methodology adalah metode yang digunakan untuk membangun sebuah model ontologi dan untuk mendeskripsikan setiap aktivitas secara rinci. Adapun tahapan-tahapan dari methodology, yaitu sebagai berikut [10,15].



Gambar 1. Tahapan Methodology

2.1 Spesifikasi

Spesifikasi dilakukan pada penelitian ini untuk mendapatkan sebuah dokumen yang menjabarkan spesifikasi ontologi yang akan dibangun. Dokumen spesifikasi ini dapat ditulis dalam format formal, semi-formal dan informal dengan menggunakan bahasa alami. Pada tahap ini, metode yang digunakan dapat berupa serangkaian representasi dengan menerapkan pertanyaan kompetensi atau tingkat keformalan menengah.

2.2 Akuisisi Pengetahuan

Selanjutnya adalah tahapan akuisisi yang pada sebagian kegiatan ini telah dilakukan secara bersamaan dengan tahap spesifikasi di awal. Namun, peranan tahap akuisisi ini akan secara bertahap berkurang dengan berjalannya proses pengembangan ontologi menuju tahapan selanjutnya. Tahap ini berisi mengenai pengumpulan informasi dan pengetahuan terkait dengan Loholoh Cmcem Penglipuran dari berbagai sumber seperti wawancara, internet dan jurnal.

2.3 Konseptualisasi

Pada tahap konseptualisasi, sebuah model konseptual yang merepresentasikan pengetahuan atau informasi dalam domain tertentu akan dibentuk. Model ini akan menggambarkan permasalahan dan solusi dengan menggunakan istilah yang telah diidentifikasi pada tahap yang sebelumnya yaitu tahap spesifikasi. Hal yang selanjutnya akan dilakukan adalah membangun Glossary of Terms yang lengkap terkait dengan domain tersebut. Glossary ini akan mengumpulkan seluruh pengetahuan domain yang memiliki kemungkinan dan memberikan definisi untuk setiap istilah.

2.4 Integrasi

Dalam tahap integrasi ini, proses melibatkan pemanfaatan definisi-definisi ontologi yang sudah ada sebelumnya dan mengkombinasikannya dengan ontologi lain yang relevan. Hal ini dilakukan untuk agar pengembangan ontologi baru tidak harus dimulai dari awan, namun dapat memanfaatkan komponen-komponen ontologi yang telah dibangun sebelumnya.

2.5 Implementasi

Tahapan ini merupakan pengimplementasian desain ontologi yang dirancang dengan menggunakan perangkat lunak yaitu Protégé.

2.6 Evaluasi

Pada tahap ini akan dilakukan evaluasi ontologi yang telah dibangun, memastikan bahwa ontologi telah memenuhi kebutuhan dan tujuan yang telah ditetapkan pada tahap spesifikasi. Selanjutnya akan dilakukan perbaikan atau penyesuaian apabila diperlukan berdasarkan umpan balik dari evaluasi.

2.7 Dokumentasi

Tahapan terakhir yaitu dokumentasi yang dapat dilakukan dalam beberapa cara proses pendokumentasian yaitu, artikel-artikel yang dipublikasikan di konferensi, jurnal ilmiah yang membahas isu-isu penting yang muncul dari ontologi yang telah dihasilkan, di dalam ontologi itu sendiri dan teks berbahasa alami yang dilampirkan pada definisi formal ontologi. Tahapan dokumentasi ini harus menyediakan dokumentasi yang jelas dan terstruktur sehingga dapat memudahkan pemahaman dan pengguna ontologi untuk di masa yang mendatang.

3. Hasil dan Diskusi

Dalam penelitian ini, sebuah ontologi dirancang dengan domain Loholoh Cmcem Penglipuran. Adapun hasil yang diperoleh dari setiap tahapan-tahapan pada metode penelitian yang telah dilakukan, antara lain.

3.1. Spesifikasi

Tahap spesifikasi ini terkait ontologi yang telah dibangun sebelumnya, berikut merupakan deskripsi pada ontologi "Loholoh Cmcem Penglipuran".

- a. Domain : Loholoh Cmcem Penglipuran

- b. Tanggal : 6 Mei 2024
- c. Tujuan : Merancang ontologi pada representasi lolah
- d. Dirancang oleh : Putu Chandra Mayoni
- e. Dilaksanakan oleh : Putu Chandra Mayoni
- f. Metode : Methodology
- g. Perangkat Lunak : Protégé
- h. Kategori Formalitas : Formal
- i. Ruang Lingkup : Loloh Cemcem Penglipuran
- j. Sumber Pengetahuan : Jurnal, wawancara dan internet.

3.2. Akuisisi Pengetahuan

Tahap Akuisisi Pengetahuan pada minuman tradisional ini menggunakan teknik sebagai berikut.

- a. Mencari dan mengidentifikasi pengetahuan dan informasi pada internet dan jurnal terkait dengan Loloh Cemcem penglipuran.
- b. Melaksanakan wawancara dengan para ahli yang membuat Loloh Cemcem Penglipuran untuk mendapatkan informasi yang lengkap untuk proses perancangan ontologi.
- c. Mencari dan Mengidentifikasi struktur yang akan digunakan melalui studi literatur terkait dengan Ontologi Web semantik.

3.3. Konseptual

Pada tahap konseptualisasi, sebuah model konseptual yang merepresentasikan pengetahuan atau informasi dalam domain tertentu akan dibentuk. Setelah model konseptual berhasil dibentuk, metodologi akan berlanjut dengan mengubah model konseptual tersebut menjadi model yang formal yang nantinya akan diterapkan dengan menggunakan bahasa ontologi. Berikut merupakan kelas dan sub kelas yang terbentuk untuk domain Loloh Cemcem Penglipuran.

3.4. Integrasi

Tahap Integrasi bertujuan untuk memilih ontologi yang tepat sehingga nantinya hasil yang diperoleh sesuai dengan harapan. Tahapan ini memanfaatkan ontologi yang sebelumnya telah dibentuk untuk di integrasi untuk menyesuaikan dengan domain Loloh Cemcem Penglipuran.

3.5. Implementasi

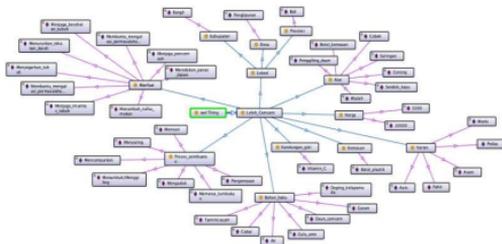
Pada pengembangan model ontologi menerapkan perangkat lunak yaitu Protégé untuk membuat setiap tahapan pada metodologi seperti kelas ontologi pada loloh cemcem, relasi antar kelas, data property, individuals class dan OntoGraf pada ontologi Loloh Cemcem Penglipuran.



Gambar 2. Class Ontology Loloh Cemcem Penglipuran

Pada Gambar 2. merupakan kelas ontologi pada domain Loloh Cemcem Penglipuran yang terdiri dari 9 Kelas dan 3 subkelas dari kelas Lokasi.

untuk subclass bahan baku terdiri dari 6 individuals, harga terdiri sebanyak 2 individuals, kandungan gizi terdiri 1 individuals, kemasan terdiri dari 1 individuals, lokasi terdiri dari 3 individuals, manfaat terdiri dari 9 individuals, proses terdiri dari 7 individuals dan varian terdiri sebanyak 5 individuals' class. Individuals class ini akan terhubung dengan object property dan subclass dari Loloh Cemcem.



Gambar 6. OntoGraf pada ontologi Loloh Cemcem Penglipuran

Pada Gambar 6 merupakan rancangan dari OntoGraf pada ontologi loloh cemcem yang terdiri dari class ontology, object class atau class relation, data property dan individual's class. Tanda garis yang berwarna biru merupakan relasi antara class dengan subclass. Sedangkan untuk garis yang berwarna pink keunguan merupakan relasi antara subclass dengan individual.

3.6. Evaluasi

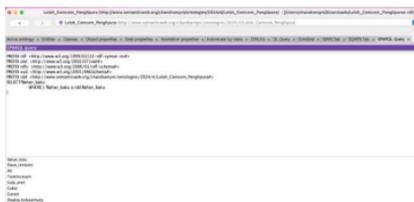
Tahap evaluasi dilakukan untuk mengidentifikasi hasil dari ontologi yang telah dibangun dan memastikan bahwa ontologi sudah sesuai dengan yang diharapkan. Pada tahap ini, akan digunakan query SPARQL yang terdapat pada perangkat lunak protégé 5.5.0. Adapun evaluasi yang akan dilakukan yaitu untuk mengetahui lokasi, manfaat, bahan baku, alat, kandungan gizi, proses pembuatan, varian, dan harga dari Loloh Cemcem Penglipuran. Pada proses ini, akan dibuat sebuah pertanyaan mengenai loloh cemcem dan akan diubah menjadi SPARQL Query sesuai dengan rumus yang telah dibuat pada perangkat lunak protégé.

Tabel 1. Pertanyaan Evaluasi Model Ontologi Loloh Cemcem Penglipuran

Pertanyaan
Apa saja bahan baku yang terdapat pada Loloh Cemcem?
Apa saja alat yang diperlukan dalam pembuatan Loloh Cemcem?
Apa saja varian dari Loloh Cemcem?
Apa saja manfaat dari minuman tradisional Loloh Cemcem?
Berapa saja harga jual dari produk minuman Loloh Cemcem?
Apa saja tahapan yang dapat dilakukan dalam pembuatan minuman tradisional Loloh cemcem?
Dari kabupaten manakah minuman tradisional Loloh Cemcem ini berasal/diproduksi?

- a. Untuk evaluasi pertama akan mencari tahu dengan menggunakan pertanyaan “apa saja bahan baku yang terdapat pada loloh cemcem”. Pertanyaan yang akan diuji pada perangkat lunak protégé tersebut akan diubah terlebih dahulu ke dalam rumus SPARQL Query, sehingga menjadi :

```
SELECT ?Bahan_baku  
WHERE { ?Bahan_baku a rdd:Bahan_baku  
}
```



Gambar 7. Hasil query bahan baku loloh cemcem

Pada gambar diatas, telah ditampilkan informasi mengenai apa saja bahan baku yang terdapat pada pembuatan loloh cemcem. Dari hasil informasi tersebut, diperoleh bahwa pada pembuatan loloh cemcem terdiri dari 8 bahan baku.

- b. Selanjutnya akan mengajukan pertanyaan mengenai “apa saja alat yang diperlukan dalam pembuatan loloh cemcem” akan diubah ke dalam rumus SPARQL Query menjadi:

```
SELECT ?Alat  
WHERE { ? Alat a rdd:Alat  
}
```

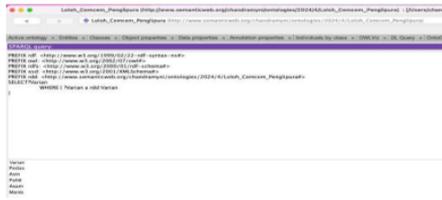


Gambar 8. Hasil query alat pembuatan loloh cemcem

Pada gambar diatas, merupakan hasil dari query yang telah diuji ke dalam SPARQL Query. Dari hasil tersebut, diperoleh informasi bahwa terdapat sebanyak 8 buah alat yang diperlukan pada pembuatan loloh cemcem.

- c. Pertanyaan yang selanjutnya akan diuji adalah “apa saja varian dari loloh cem cem”. Berikut merupakan query yang akan diuji:

```
SELECT ?Varian  
WHERE { ? Varian a rdd:Varian  
}
```

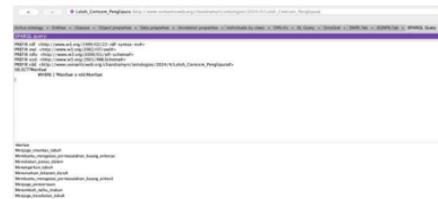


Gambar 9. Hasil query varian loloh cemcem

Gambar diatas merupakan hasil dari query, varian yang dimiliki oleh loloh cemcem. Dari hasil informasi diatas diperoleh untuk varian loloh cemcem terdiri dari 5 varian.

- d. Pertanyaan yang berikutnya akan diuji adalah "apa saja manfaat dari minuman tradisional loloh cemcem". Dari pertanyaan tersebut, akan diubah menjadi query.

```
SELECT?Manfaat
WHERE {? Manfaat a rdd:Manfaat
}
```

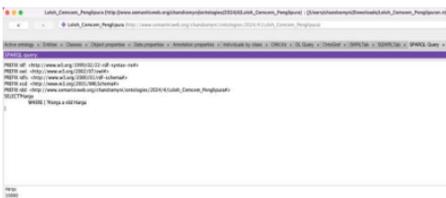


Gambar 10. Hasil query manfaat loloh cemcem

Gambar diatas merupakan hasil query mengenai manfaat yang dapat diperoleh dari mengkonsumsi loloh cemcem. Dari hasil tersebut, diperoleh informasi untuk manfaat pada loloh cemcem terdapat sebanyak 9 manfaat yang baik untuk kesehatan tubuh.

- e. Pertanyaan selanjutnya, yang akan diuji adalah "berapa saja harga jual dari produk minuman loloh cemcem". Dari pertanyaan tersebut, akan diubah menjadi query.

```
SELECT?Harga
WHERE {? Harga a rdd:Harga
}
```



Gambar 11. Hasil query harga loloh cemcem

Gambar diatas merupakan hasil query mengenai harga jual dari produk minuman loh cecem. Dari hasil tersebut, diperoleh informasi untuk harga dari minuman loh cecem yaitu 5000 dan 10000.

- f. Pertanyaan selanjutnya yang akan diuji adalah "apa saja tahapan yang dapat dilakukan dalam pembuatan minuman tradisional loh cecem". Dari pertanyaan tersebut, akan diubah menjadi query.

```
SELECT?Proses_pembuatan  
WHERE {? Proses_pembuatan a rdd:Proses_pembuatan  
}
```



Gambar 12. Hasil query proses pembuatan loh cecem

Pada gambar diatas merupakan hasil dari query mengenai pertanyaan proses yang dapat dilakukan dalam pembuatan loh cecem. Dari hasil pengujian tersebut diperoleh informasi bahwa proses pembuatan dapat dilakukan dengan 7 tahapan.

- g. Pertanyaan selanjutnya yang akan diuji adalah "Dari kabupaten manakah minuman tradisional loh cecem ini berasal/diproduksi?" Dari pertanyaan tersebut, akan diubah menjadi query.

```
SELECT?Kabupaten  
WHERE {? Kabupaten a rdd:Kabupaten  
}
```



Gambar 13. Hasil query kabupaten asal loh cecem

Pada gambar diatas merupakan hasil dari query mengenai pertanyaan kabupaten yang memproduksi loh cecem. Dari hasil query menampilkan bahwa loh cecem berasal dan diproduksi di kabupaten Bangli.

3.7. Dokumentasi

Tahap dokumentasi yang dihasilkan pada penelitian ini adalah jurnal dengan judul "Rancang Model Ontologi Dalam Representasi Digital Loh Cecem Pengilpuran".

4. Kesimpulan

Dari proses penelitian yang telah dilakukan, telah diperoleh sebuah model ontologi yang terdiri dari konsep, property, relation dan individuals dari Loloh Cemcem Penglipuran. Ontologi ini terdiri dari 9 class yaitu Alat, Bahan_baku, harga, kandungan gizi, kemasan, lokasi, manfaat, proses pembuatan dan varian. Serta relation pada object property sebanyak 9, data property sebanyak 12 dan individuals sebanyak 42. Evaluasi pada penelitian ini dilakukan dengan menggunakan SPARQL Query untuk menguji dan memastikan bahwa ontologi telah memenuhi kebutuhan, tujuan dan sesuai dengan yang diharapkan. Berdasarkan hasil pengujian tersebut menunjukkan bahwa ontologi pada Loloh Cemcem dapat menampilkan informasi dengan baik dan sesuai dengan query yang diajukan. Dari model ontologi ini diharapkan dapat memfasilitasi akses informasi terstruktur dan pelestarian pengetahuan tentang minuman tradisional Loloh Cemcem.

Daftar Pustaka

- [1] M, Zamroni, "Desa Penglipuran Raih Penghargaan Desa Wisata Terbaik Tahun 2023 Versi UNWTO", TunasHijau, 2 November 2023, [Online]. Tersedia: <https://tunashijau.id/2023/11/desa-penglipuran-raih-penghargaan-desa-wisata-terbaik-2023-versi-unwto/#:~:text=Desa%20Penglipuran%20yang%20berada%20di.mendapatkan%20penghargaan%20bergenesi%20tingkat%20dunia>. [Diakses: 21 April 2024].
- [2] I.M.Mertawan, "Paon dan Loloh Cemcem Penglipuran Bali Tercatat sebagai Kekayaan Intelektual di Kemenkumham RI," Bali Express, 24 November 2023, [Online]. Tersedia: <https://baliexpress.jawapos.com/bali/673321242/paon-dan-loloh-cemcem-penglipuran-bali-tercatat-sebagai-kekayaan-intelektual-di-kemenkumham-ri> [Diakses: 21 April 2024].
- [3] I.G.N.Y.P.Negara, and T.Wahab, "Perancangan Identitas Visual Kemasan Loloh Cemcem Meme Nyampuh Di Desa Penglipuran Provinsi Bali," e-Proceeding of Art & Design, vol.10, no.2, pp. 1987, April 2023.
- [4] P.L.Cahyaningrum, and N.L.G.Sudaryati, "Budaya Minum Loloh Sebagai Upaya Meningkatkan Imunitas Tubuh Di Masa Pandemi COVID 19," E-Jurnal Widya Kesehatan, vol.3, no.2, pp.19-24, October 2021.
- [5] Z.Zellini, N.M.Wiasti, and I.N.Sama, "Ekonomi Kreatif: Pengembangan Loloh Cemcem Di Desa Penglipuran, Kabupaten Bangli, Bali," Jurnal Humanis, Fakultas Ilmu Budaya Unud, vol.19, no.1 pp. 334- 335, May 2017.
- [6] I.G.B.N.P.Putra, G.D.Larasdiputra, I.G.S.Pratama and A.P.Putra, "Pemerdayaan Masyarakat Pada Kelompok Usaha Loloh Cemcem," International Journal Of Community Service Learning, vol.4, no.4, pp.305, November 2020.
- [8] Y.F.Badron, F.Agus, and H.R.Hatta, "Studi Tentang Pemodelan Ontologi Web Semantik Dan Prospek Penerapan Pada Bibliografi Artikel Jurnal Ilmiah," Prosiding Seminar Ilmu Komputer dan Teknologi Informasi, vol.2, no.1, pp.164-168, March 2017.
- [9] Himawan, T.W.Harjanti, R.Supriati, and H.Setiyani, "Evolusi Penggunaan Teknologi Web 3.0: Semantic Web," Journal Of Information System, Graphics, Hospitality and Technology, vol.2, no.2, pp. 54-57, November 2020.
- [10] I.M.S.Gunawan, and D.D.L.G.Astuti, "Rancangan Ontologi Semantik: Representasi Digital Kain Songket Bali," Jurnal Nasional Teknologi Informasi dan Aplikasinya, vol.1, no.4, pp.118-1190, August 2023.
- [11] Cerakan Kebudayaan Bali, "Loloh Cemcem Minuman Khas Desa Penglipuran," Diskominfo Provinsi Bali, 20 Maret 2023, [Online]. Tersedia: <https://cerakan.baliprov.go.id/detail/loloh-cemcem-minuman-khas-desa-penglipuran-1656572427>. [Diakses: 27 April 2024].
- [12] L.E.Wahyuni, and I.K.A.K.Stiawan, "Loloh Cemcem Berkhasiat Khas Penglipuran Bali."
- [13] K.K.Noviyanti, L.G.Astuti, "Pengembangan Model Ontologi Pada Domain Bimbingan Belajar," Jurnal Elektronik Ilmu Komputer Udayana, vol.12, no.1, pp.136-137, August 2023.

Halaman ini sengaja dibiarkan kosong

Simulasi IoT Pemantauan Tanaman Lidah Buaya Berbasis Algoritma Fuzzy

Bayu Yudistira Ramadhan^{a1}, Luh Arida Ayu Rahning Putri^{a2}

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹ramadhan.2208561085@student.unud.ac.id
²rahningputri@unud.ac.id

Abstract

In the digital era, technological advancements have entered the Industrial Revolution 4.0, and the agricultural sector in Indonesia is adapting to this revolution. Using a fuzzy algorithm, this research simulates an IoT-based monitoring system for Aloe vera plants. The system aims to assist farmers in real-time and accurately monitor Aloe vera plant conditions. Aloe vera has a narrow optimal temperature range of 16-33°C and an ideal soil moisture range of 40-75% for optimal growth. By implementing fuzzy logic, a mathematical concept that is easy to understand, the system can accurately map input conditions to output decisions. The simulation uses MATLAB and the Tinkercad website to design a fuzzy logic system that controls a water pump and fan based on soil moisture and temperature inputs from SEN1 and LM35 sensors, respectively. The fuzzy rules maintain the ideal conditions for Aloe vera growth, reducing energy consumption and water waste. The results demonstrate the high accuracy of the fuzzy logic system in making control decisions for maintaining optimal growing conditions.

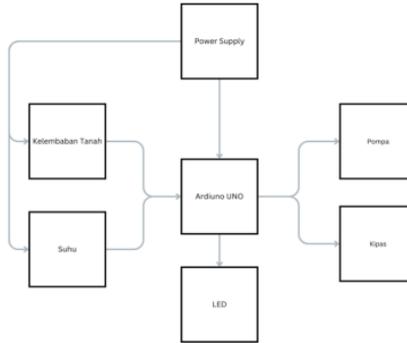
Keywords: Internet of Things, Fuzzy Logic, Aloe Vera, Soil Moisture, Temperature

1. Pendahuluan

Pada era digital saat ini, perkembangan teknologi sudah mulai memasuki era revolusi industri 4.0, sektor pertanian pun mulai membenahi sektor industri yang ada di Indonesia [1]. Salah satu penerapannya adalah sistem pemantauan tanaman lidah buaya berbasis IoT dan algoritma fuzzy. Sistem ini diharapkan dapat membantu petani dalam memantau kondisi tanaman lidah buaya mereka secara real-time dan akurat. Tanaman Lidah Buaya Tanaman lidah buaya (Aloe vera) memiliki rentang suhu optimal untuk pertumbuhan yang sempit, yaitu antara 16-33°C[2]. Di luar jangkauan suhu tersebut, tanaman akan berhenti tumbuh, mengalami dehidrasi, layu, dan menjadi lebih rentan terhadap sinar matahari yang berlebihan [3]. Selain suhu, kelembaban tanah juga memegang peranan penting dalam menjaga kesuburan dan pertumbuhan tanaman lidah buaya. Kelembaban tanah yang ideal untuk tanaman ini berkisar antara 40-75% [4]. Dengan menerapkan logika fuzzy, yang merupakan dasar dari penalaran fuzzy, konsep matematisnya menjadi lebih mudah dimengerti, sehingga mempelajarinya tidak membutuhkan waktu yang lama [5]. Logika fuzzy mampu menunjukkan sejauh mana suatu nilai dianggap salah atau benar. Ketika memetakan ruang input ke dalam ruang output, logika fuzzy dapat diandalkan untuk memberikan hasil yang akurat dan berguna [6]. Himpunan fuzzy adalah kumpulan elemen yang memiliki dua atribut: variabel bahasa, yang menggambarkan kondisi dengan menggunakan bahasa alami seperti panas, dingin, cepat, lambat, dan sebagainya; serta variabel numerik, yang mengukur ukuran suatu variabel dalam bentuk angka[5]. Berdasarkan permasalahan diatas, maka pada penelitian ini akan disimulasikan logika fuzzy untuk sistem pengendalian suhu optimal tanaman lidah buaya 16-33°C dan pengendalian kelembaban tanah maksimal nilai 75%. Penggunaan logika fuzzy dalam sistem pengendalian dapat mengurangi energi yang terpakai selama pemeliharaan tanaman dan mengurangi pemborosan sumber daya air [3]. Tujuan yang ingin dicapai merupakan merancang logika fuzzy yang sesuai untuk sistem pengendalian kelembaban tanah dan suhu yang khusus untuk tanaman lidah buaya.

2. Metode Penelitian

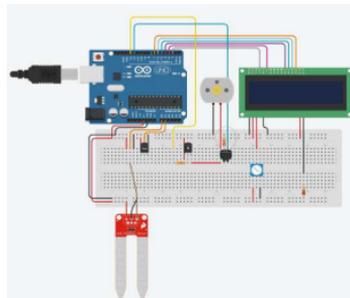
Metode yang digunakan dalam simulasi IoT logika fuzzy ini adalah dengan menggunakan aplikasi MATLAB yang memanfaatkan metode logika fuzzy untuk pengambilan keputusan. Keputusan yang diambil adalah untuk menentukan apakah kondisi suhu dan kelembaban tanah sudah memadai untuk menyediakan nutrisi yang cukup bagi tanaman lidah buaya dan menggunakan website Tinkercad untuk menyimulasikan alat yang akan dipakai. Data diambil secara acak dengan membedakan antara penelitian sebelumnya dan mengambil rata-rata dari penelitian sebelumnya mengenai suhu dan kelembaban tanah pada tanaman lidah buaya



Gambar 1. Diagram Blok Sistem

Pada Gambar 1. Ditunjukkan bahwa terdapat sebuah perangkat yang bertugas untuk mengontrol suhu dan kelembaban tanah, yaitu pompa dan kipas. Suhu dan kelembaban tanah berperan sebagai input yang akan diproses oleh logika fuzzy untuk menentukan keputusan, sedangkan pompa dan kipas bertindak sebagai output fungsi untuk menerapkan keputusan yang dihasilkan oleh logika fuzzy. LED bertugas untuk menunjukkan data yang dihasilkan oleh sensor kelembaban tanah dan sensor suhu. Pada website Tinkercad terdapat penggunaan beberapa alat atau komponen digunakan, yang berfungsi sebagai berikut:

- Penggunaan sensor SEN1 sebagai pendeteksi kelembaban tanah.
- Penggunaan sensor LM35 sebagai pendeteksi suhu udara.
- Penggunaan Arduino Uno sebagai alat yang memproses kerja output dari sensor SEN1 dan LM35.
- Penggunaan DC Motor yang diasumsikan sebagai pompa air
- Penggunaan lampu bohlam yang diasumsikan sebagai kipas.
- Penggunaan LCD sebagai display atau penampil data yang berupa karakter dan dalam bentuk bilangan bit tertentu.

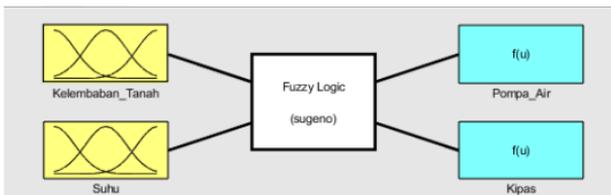


Gambar 2. Skema Rangkaian Sistem

Pada skema diagram yang ditunjukkan pada Gambar 2, ada penggunaan kabel untuk menghubungkan berbagai komponen agar bisa bekerja satu sama lain. Dalam diagram tersebut, variasi warna kabel digunakan khususnya warna merah dan hitam untuk membedakan komponen yang dihubungkan ke ground, tegangan positif, dan tegangan negatif. Dalam hal ini, kabel merah merepresentasikan tegangan positif dan kabel hitam sebagai ground.

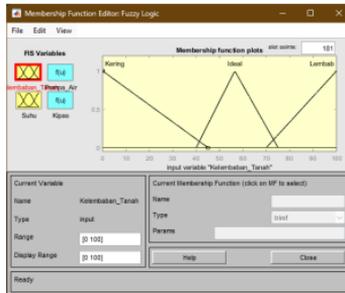
3. Hasil dan Pembahasan

Dalam penelitian ini, logika fuzzy digunakan sebagai metode pengambilan keputusan untuk menentukan durasi waktu yang diperlukan dalam mempertahankan kelembaban tanah dan suhu pada tingkat yang diinginkan. Hal ini bertujuan untuk menjaga pertumbuhan tanaman lidah buaya secara optimal. Percobaan dilakukan dengan variabel linguistik pada himpunan fuzzy. Nilai variabel linguistik pada himpunan fuzzy dibuat berdasarkan tiga kondisi output untuk pompa air dan kipas.

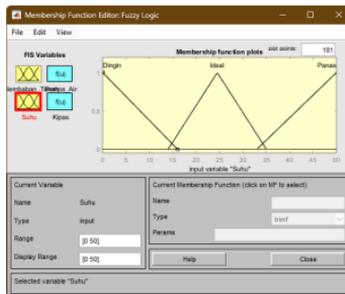


Gambar 3. Sistem logika Fuzzy

Arduino Uno yang digunakan dalam perancangan ini merupakan komponen utama yang menjalankan sistem pengendalian sebagai pemroses perintah input dari sensor kelembaban tanah dan sensor suhu. Rancangan Arduino diperoleh dari website Tinkercad. Data diproses oleh Arduino Uno sesuai dengan perancangan logika fuzzy yang telah dirancang untuk mengendalikan output berupa pompa air dan kipas. Arduino Uno kemudian mengaktifkan atau menonaktifkan pompa air dan kipas secara otomatis. Masukan atau input keanggotaan dari suhu dan kelembapan tanah dijelaskan sebagai berikut.



Gambar 4. Input Kelembaban Tanah



Gambar 5. Input Suhu

Dengan input suhu dan kelembapan tersebut, dihasilkan output berupa aktivasi pompa air dan kipas sebagai berikut.



Gambar 6. Output Pompa Air



Gambar 7. Output Kipas Angin

Setelah menentukan kondisi input berupa suhu dan kelembapan tanah serta output berupa pompa dan kipas yang berguna untuk mengendalikan kondisi input dari suhu dan kelembapan. Berikut ini aturan logika fuzzy yang digunakan pada simulasi ini untuk menentukan suatu kondisi dari output ketika diberikan input. Pada simulasi ini, sensor SEN1 digunakan sebagai sensor kelembapan dan LM35 sebagai sensor suhu sebagai sensor input dalam menghasilkan output berupa parameter suhu dan kelembapan yang dapat dipertahankan sesuai dengan karakteristik perkembangan dan pertumbuhan tanaman lidah buaya.

Tabel 1. Rules Logika Fuzzy

No	Input		Output	
	Kelembaban Tanah	Suhu	Pompa Air	Kipas
1	Lembab	Panas	Mati	Lama
2	Lembab	Ideal	Mati	Mati
3	Lembab	Dingin	Mati	Mati
4	Ideal	Panas	Cepat	Lama
5	Ideal	Ideal	Cepat	Mati
6	Ideal	Dingin	Mati	Mati
7	Kering	Panas	Lama	Lama
8	Kering	Ideal	Lama	Mati
9	Kering	Dingin	Lama	Mati

Untuk memperdalam pemahaman tentang manfaat implementasi sistem monitoring dengan logika fuzzy pada budidaya tanaman lidah buaya, studi komparatif dapat dilakukan dengan membandingkan hasil panen yang dibudidayakan menggunakan sistem ini dengan metode tradisional. Hal ini akan memberikan gambaran jelas tentang efektivitas dan peningkatan yang diperoleh dengan menggunakan teknologi modern. Evaluasi keandalan sensor suhu dan kelembapan tanah yang digunakan juga penting untuk memastikan akurasi data yang dihasilkan, sehingga algoritma fuzzy logic dapat mengambil keputusan dengan optimal. Selanjutnya, perbandingan hasil panen tanaman lidah buaya dari area yang dikelola dengan bantuan sistem monitoring dan area yang tidak menggunakan teknologi serupa dapat memperkuat pemahaman tentang dampak nyata implementasi sistem ini. Diskusi juga dapat mencakup aspek ekonomi, keberlanjutan budidaya, serta tingkat adopsi teknologi oleh petani tanaman lidah buaya. Dengan membahas aspek-aspek tersebut, pemahaman yang komprehensif tentang dampak

implementasi sistem monitoring dengan logika fuzzy pada budidaya tanaman lidah buaya dapat diperoleh.

4. Kesimpulan

Simulasi yang telah dilakukan menunjukkan bahwa metode logika fuzzy dalam sistem pengambilan keputusan atau sistem pengendali memiliki akurasi yang tinggi, sehingga sangat tepat digunakan untuk mendapatkan keputusan terbaik dalam sistem pengendali. Oleh karena itu, logika fuzzy banyak diaplikasikan dalam berbagai kondisi pengaturan atau pengendalian suatu sistem. Sistem pengendalian kelembaban tanah dan suhu dirancang dengan sistem otomatis menggunakan logika fuzzy metode Sugeno. Metode Sugeno cocok diterapkan untuk sistem pengendalian karena menggunakan fungsi If-Then dalam menghasilkan output berupa logika fuzzy. Simulasi yang dilakukan menunjukkan bahwa logika fuzzy mampu mengambil keputusan dengan sangat baik dan memiliki akurasi yang sangat tinggi untuk sistem kontrol atau sistem pengendali, terutama dengan menggunakan sensor SEN1 sebagai sensor kelembaban dan LM35 sebagai sensor suhu sebagai input untuk menghasilkan output berupa parameter suhu dan kelembaban tanah yang dapat dipertahankan sesuai dengan karakteristik perkembangan dan pertumbuhan ideal tanaman lidah buaya.

Daftar Pustaka

- [1] W. F. Wardhani, "Peran Politik Pertanian Dalam Pembangunan Pertanian Menghadapi Era Revolusi Industri 4.0 di Sektor Pertanian," *JISIPOL | Jurnal Ilmu Sosial dan Ilmu Politik*, vol. 3, no. 2, 2019.
- [2] D. O. Ramadhani, G. N. Fatih, A. Muhibah, A. El Maula, D. Tami, dan ..., "Pemanfaatan Tumbuhan Apotek Hidup Sebagai Obat Tradisional Masyarakat di Gunungpati," *Jurnal Potensial*, vol. 3, no. 1, hlm. 105–116, 2024.
- [3] Nasron, Suroso, dan A. R. Putri, "Perancangan Logika Fuzzy Untuk Sistem Pengendali Kelembaban Tanah dan Suhu Tanaman," *Jurnal Media Informatika Budidarm*, vol. 3, no. 4, hlm. 307–312, 2019, doi: 10.30865/mib.v3i4.1245.
- [4] A. Sari W.A dan R. Khana, "Rancang Bangun Sistem Monitoring Tanaman Lidah Buaya Berbasis IOT," *Jurnal Kajian Teknik Elektro*, vol. 6, no. 1, 2022, doi: 10.52447/jkte.v6i1.5732.
- [5] M. I. Salim, "Simulasi Logika Fuzzy Pada Pengatur Sensor Suhu dan Kelembaban Tanah Tanaman," *SinarFe7*, hlm. 136–139, 2021.
- [6] A. Siti, "Analisis komparasi metode tsukamoto dan sugeno dalam prediksi jumlah siswa baru," *Jurnal Teknologi Informasi dan Komunikasi*, vol. 8, no. 2, 2016.

Analisa Perancangan Sistem Rekomendasi Makanan Untuk Defisit Kalori “Calcraft” Melalui Evaluasi UI/UX

Putu Ananda Darma Wiguna^{a1}, Luh Gede Astuti^{a2}

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹wiguna.2208561099@student.unud.ac.id
²lg.astuti@unud.ac.id

Abstract

This research raises issues related to unbalanced diets in modern society that cause obesity problems. As a solution, calorie deficit was chosen as a program to solve this obesity problem. Developed an application called “Calcraft” which is integrated with a recommendation system to provide food selection recommendations by considering the calculation of BMI (Body Mass Index) and BMR (Basal Metabolic Rate). Application evaluation is carried out using the System Usability Scale (SUS) method, with a focus on user interface and user experience. Not only that, but the user’s understanding of the knowledge also gained from the features on Calcraft is one of the things the author hopes for. Of the 35 respondents who filled out the SUS questionnaire, “Calcraft” obtained a total score of 2907.5, with an average of 83.07142857. These results indicate very good quality, corresponding to the “A” or “Excellent” category in SUS interpretation. Thus, this app has great potential to assist individuals in running a calorie deficit program effectively and conveniently.

Keywords: calories deficit, recommendation system, system usability scale, user interface, user experience

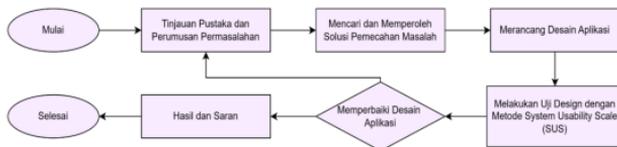
1. Pendahuluan

Pola hidup masyarakat sekarang ini, cenderung mengarah pada kebiasaan makan bebas yang tidak seimbang dengan asupan kalori melebihi kebutuhan tubuh. Gaya hidup yang kurang aktif, suka bermalasan-malasan atau dikenal sebagai sedentary lifestyle menjadi suatu masalah yang cukup serius dalam kesehatan global, termasuk Indonesia. Hal-hal tersebut dipicu dari berbagai faktor seperti urbanisasi, perubahan pola konsumsi makanan serta kemajuan teknologi yang sangat mempengaruhi aktivitas fisik sehari-hari. Jika pola hidup seperti ini tidak diselesaikan dengan sungguh-sungguh, tentu akan memberikan dampak buruk dan berpotensi pada peningkatan angka penyakit obesitas atau overweight yang berdampak buruk terhadap kesehatan global. Obesitas menjadi salah satu faktor risiko utama dalam beberapa penyakit kronis, contohnya seperti diabetes tipe II, penyakit jantung hingga kanker. Data Riskesdas pada tahun 2023 mencatat bahwa terdapat sebesar 15,3% kasus obesitas di Indonesia. Yang dimana ini terus berlanjut dan mengacu pada data di tahun 2018, kasusnya meningkat menjadi sebesar 21,8% [1]. Disisi lain, laporan dari World Health Organization (WHO) pada tahun 2014 mengungkapkan bahwa lebih dari 1,9 miliar orang dewasa berusia 18 tahun ke atas mengalami kelebihan berat badan, dengan lebih dari 600 juta diantaranya menderita obesitas. Sekitar 39% populasi dewasa menghadapi masalah kelebihan berat badan, dan 13% diantaranya mengalami obesitas. Selain itu, terdapat juga 41 juta anak di bawah usia 5 tahun yang mengalami kelebihan berat badan, dengan 10% diantaranya menderita obesitas. [2]. Melihat tingginya angka prevalensi obesitas ini, solusi yang dapat dilakukan yaitu dengan melakukan edukasi dan memberikan ilmu pengetahuan terkait dengan perhitungan kalori dalam pemilihan makan serta penerapan program defisit kalori. Perhitungan penting dilakukan dengan memperhatikan Body Mass Index (BMI) dan Basal Metabolic Rate (BMR) seseorang. BMI (Body Mass Index) merupakan metode pengukuran untuk mengevaluasi kondisi berat badan seseorang, apakah berada dalam kondisi normal atau tidak (underweight atau overweight) [3]. Sementara itu, BMR

(Basal Metabolic Rate) adalah tingkat metabolisme dasar, yaitu jumlah energi yang digunakan oleh tubuh untuk menjalankan fungsi-fungsi dasar organ-organ tubuh [4]. Dengan mengetahui hasil perhitungan BMI dan BMR, maka akan mempermudah penentuan jenis makanan yang sesuai untuk dikonsumsi pada saat melakukan program defisit kalori. Defisit kalori merupakan kondisi yang terjadi ketika jumlah kalori yang dikonsumsi lebih sedikit daripada yang dibakar oleh tubuh. Ini merupakan konsep dalam usaha menurunkan berat badan. Kata kalori disini merujuk pada jumlah energi yang terdapat dalam makanan maupun minuman yang dikonsumsi [5]. Meskipun sering disebut sebagai diet di kalangan masyarakat, namun sebenarnya defisit kalori secara spesifik berbeda dengan diet. Diet biasanya lebih merujuk pada pola makan yang diikuti seseorang dalam mencapai level kebugaran tertentu. Sedangkan defisit kalori merupakan konsep yang lebih berfokus dalam mencapai penurunan berat badan. Program defisit kalori tidak hanya berfokus pada pengurangan asupan kalori semata, tetapi juga memperhatikan pengeluaran energi tubuh. Jika seseorang memiliki pengetahuan yang minim dalam menjalani program ini, makan dapat berpotensi berbahaya dan dapat menyebabkan kekurangan gizi. Dalam era teknologi yang berkembang begitu pesat saat, muncul sebuah gagasan dari penulis terkait pengembangan aplikasi yang terintegrasi dengan sistem rekomendasi untuk memilih makanan dengan kalori yang sesuai untuk mendukung penerapan program defisit kalori. Dalam konteks ini, penulis berfokus pada sebuah rancangan aplikasi rekomendasi pemilihan makanan untuk mendukung program defisit kalori yang bernama "Calcraft". Calcraft diambil dari dua kata bahasa asing yaitu "Calorie" dan "Craft" yang bermakna pada keahlian dalam pemilihan makanan serta pengukuran energi yang dibutuhkan tubuh. Perancangan aplikasi ini akan dilakukan dengan memperhatikan antarmuka (user interface) yang intuitif dan pengalaman pengguna (user experience) yang memiliki energi positif. Selain itu, ide pengembangan aplikasi ini juga mendukung upaya global dalam mencapai Tujuan Pembangunan Berkelanjutan (Sustainable Development Goals/SDGs), yang dalam hal ini berfokus pada poin SDG 3 tentang "Kehidupan Sehat dan Sejahtera" [6].

2. Metode Penelitian

Penelitian ini melibatkan beberapa tahapan metodologi untuk memperoleh hasil yang diinginkan. Tahap pertama dimulai dengan melakukan tinjauan pustaka dengan mengumpulkan beberapa informasi yang bersumber dari jurnal atau artikel terkait dan selanjutnya barulah mulai untuk merumuskan permasalahan. Setelah kedua hal tersebut dilakukan, alur selanjutnya yaitu mencari solusi pemecahan masalah. Pada penelitian ini solusi yang diusulkan adalah merancang sebuah antarmuka pengguna untuk sistem rekomendasi berbasis mobile. Kemudian dilakukan perancangan desain antarmuka aplikasi dengan melakukan beberapa pertimbangan yang sesuai dalam memecahkan masalah. Setelah itu, pengujian akan dilakukan dengan menerapkan metode System Usability Scale (SUS). Metode ini menyertakan penyebaran kuesioner kepada sejumlah responden, dan hasilnya akan dijumlahkan untuk menghasilkan nilai SUS. Berdasarkan nilai SUS tersebut, dilakukan perbaikan pada desain, serta peninjauan ulang terhadap literatur dan permasalahan yang telah dirumuskan. Jika hasilnya sesuai harapan, akan disusun laporan dan rekomendasi terkait dengan penelitian ini.



Gambar 1. Alur Metodologi Penelitian

2.1 System Usability Scale (SUS)

System Usability Scale atau kerap dikenal dengan nama SUS merupakan alat ukur yang digunakan untuk melakukan evaluasi terhadap tingkat kegunaan (usability) sebuah sistem. Pertama kali diciptakan oleh John Brooke pada tahun 1986, SUS diimplementasikan untuk mengukur kegunaan dari berbagai jenis produk, termasuk software, hardware, aplikasi mobile, dan website [7]. Penting untuk diketahui bahwa nilai SUS tidak menunjukkan kegunaan suatu sistem secara mutlak, namun mengukur kegunaan yang dirasakan oleh seorang pengguna. SUS menggunakan sebuah kuesioner dalam mengumpulkan data responden. Kuesioner SUS terdiri dari 10 pertanyaan dengan 5 pilihan opsi jawaban. Pada metode SUS ini, responden dapat memilih jawaban dari rentang skala atau poin 1 yaitu "Sangat Tidak Setuju", skala 2 "Tidak Setuju", skala 3 "Netral", skala 4 "Setuju" dan skala 5 "Sangat Setuju". Nilai berupa angka ini, dalam rentang 0-100 ini mencerminkan tingkat kegunaan yang diuji. Namun, perlu diketahui bahwa skala tersebut tidak setara dengan sebuah persentase, misalnya nilai 65 pada SUS tidak sama dengan 65%, melainkan lebih mendekati 41%.

Metode SUS memiliki aturan perhitungan skor yang harus diikuti, berikut adalah aturannya:

1. Untuk pertanyaan nomor ganjil, skala atau skor yang dijawab oleh responden dikurangi dengan nilai 1 (skor responden -1).
2. Untuk pertanyaan nomor genap, skor akhir 5 dikurangi dengan skala atau skor responden (5 - skor responden).
3. Hasil dari proses nomor 1 dan 2 akan dikalikan dengan 2,5.

Aturan perhitungan skala atau skor di atas hanya dilakukan untuk satu responden. Kemudian, langkah tambahan selanjutnya yaitu masing-masing skor responden akan dijadikan rata-untuk mendapatkan total hasil uji SUS [8]. Skor tersebut dapat dibuatkan dalam bentuk rumus:

$$\text{SUS Score} = (((R1 - 1) + (5 - R2) + (R3 - 1) + (5 - R4) + (R5 - 1) + (5 - R6) + (R7 - 1) + (5 - R8) + (R9 - 1) + (5 - R10)) \times 2,5)$$

Dalam metode SUS terdapat interpretasi rentang nilai yang dapat ditunjukkan melalui sebuah indeks huruf (grade). Dalam pengujian SUS, jika total skor rata-rata keseluruhan > 68, sistem akan dianggap memenuhi standar untuk pengembangan lebih lanjut. Sebaliknya, jika skor < 68, sistem perlu dilakukan pengujian ulang. Berikut merupakan tabel interpretasi nilai SUS [9]:

Tabel 2. Interpretasi Nilai SUS

SUS Score	Index (Grade)	Adjective Rating
>80.3	A	Excellent
68 - 80.3	B	Good
68	C	Okay
51-68	D	Poor
<51	E	Awful

2.2 Pengumpulan Data

Data dalam penelitian ini dikumpulkan dengan cermat dan sistematis melalui penyebaran kuesioner yang berisikan 10 pertanyaan sesuai dengan metode System Usability Scale (SUS). Kuesioner disebarikan secara online melalui platform Google Form yang didalamnya berisikan skala berupa angka untuk mengukur evaluasi desain. Pilihan skala yang dimaksudkan yaitu mulai dari skala 1 "Sangat tidak setuju" sampai skala 5 "Sangat setuju". Dalam proses pengisian kuesioner ini, syarat responden yang dapat mengisi kuesioner ini adalah individu yang memiliki rentang usia minimal 15 tahun ke atas. Adapun latar belakang yang dimiliki oleh para responden, mulai dari siswa, mahasiswa, desainer grafis, gym enthusiast dan diet enthusiast. Total

responden yang terlibat dalam pengisian kuesioner yang mem mencapai angka 35 responden dengan rincian 63.9% (22 orang) responden laki-laki dan 36.1% (13 orang) responden perempuan. Di bawah ini terdapat daftar 10 pertanyaan survei yang diterapkan menggunakan metode System Usability Scale (SUS):

Tabel 1. Pertanyaan Survei System Usability Scale (SUS)

No	Pertanyaan	STS	TS	N	S	SS
1	Saya berpikir akan menggunakan aplikasi "Calcraft" lagi	0	0	0	0	0
2	Saya merasa aplikasi "Calcraft" rumit untuk digunakan	0	0	0	0	0
3	Saya merasa aplikasi "Calcraft" mudah digunakan	0	0	0	0	0
4	Saya membutuhkan bantuan dari orang lain atau teknisi dalam menggunakan aplikasi "Calcraft"	0	0	0	0	0
5	Saya merasa fitur-fitur dalam aplikasi "Calcraft" berjalan dengan semestinya	0	0	0	0	0
6	Saya merasa ada banyak hal yang tidak konsisten (tidak serasi pada aplikasi ini)	0	0	0	0	0
7	Saya rasa merasa orang lain akan memahami cara menggunakan aplikasi "Calcraft"	0	0	0	0	0
8	Saya merasa aplikasi "Calcraft" ini membingungkan	0	0	0	0	0
9	Saya merasa tidak ada hambatan dalam menggunakan aplikasi ini "Calcraft"	0	0	0	0	0
10	Saya perlu membiasakan diri terlebih dahulu sebelum menggunakan aplikasi "Calcraft"	0	0	0	0	0

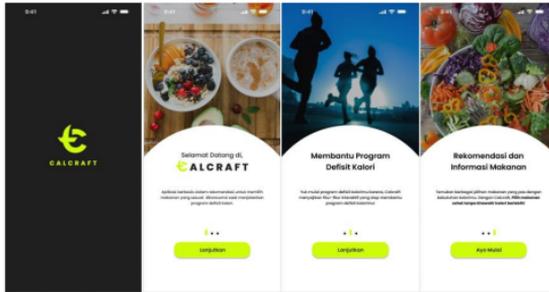
Semua survei di atas sangat penting digunakan dalam menguji pemahaman tentang nilai usability dari desain antar muka pengguna dan pengalaman pengguna terkait dengan rancangan aplikasi "Calcraft". Dengan adanya tanggapan dari para responden, tahap evaluasi akan lebih mudah untuk dilakukan dengan memperoleh wawasan terkait seberapa efektif dan efisien aplikasi "Calcraft" ini dalam memenuhi kebutuhan dan harapan pengguna

3. Hasil dan Pembahasan

Dari proses penelitian dan tinjauan pustaka yang telah diuraikan, didapatkan hasil implementasi dan skor hasil pengujian rancangan aplikasi sistem rekomendasi makanan untuk program defisit kalori "Calcraft" dengan menggunakan metode System Usability Scale (SUS), sebagai berikut:

3.1. Implementasi Desain

a. Tampilan Splash dan Onboarding Screen



Gambar 2. Tampilan Splash dan Onboarding Screen

Pada Gambar 2, dapat dilihat itu merupakan tampilan splash screen dan onboarding screen yang berisikan pengenalan awal dari aplikasi. Dimulai dari logo Calcraft hingga deskripsi singkat apa itu Calcraft dan apa saja layanan yang mampu diberikan.

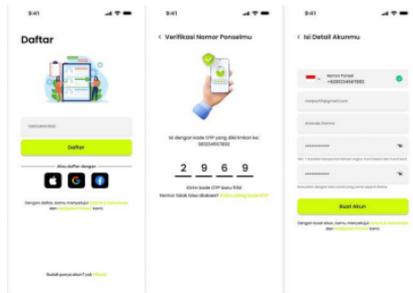
b. Tampilan Halaman Masuk



Gambar 3. Tampilan Halaman Masuk

Pada Gambar 3, dapat dilihat itu merupakan tampilan halaman masuk. Pada halaman masuk ini pengguna diarahkan untuk dapat memasukkan nomor telepon atau email beserta password jika sudah pernah mendaftar akun sebelumnya. Pengguna juga dapat masuk melalui akun media sosial lainnya.

c. Tampilan Halaman Daftar dan Verifikasi Email/Nomor Ponsel



Gambar 4. Tampilan Halaman Daftar dan Verifikasi Email/Nomor Ponsel

Pada Gambar 4, dapat dilihat itu merupakan tampilan halaman daftar. Pada halaman daftar ini pengguna dapat memasukkan email atau nomor ponsel, kemudian secara otomatis akan dikirimkan kode OTP, lalu pengguna diarahkan mengisi identitas diri secara lengkap untuk mendaftar akun. Sama juga seperti tampilan masuk, pengguna dapat langsung mendaftar dengan akun media sosial lainnya.

d. Tampilan Halaman Home/Beranda



Gambar 5. Tampilan Halaman Home/Beranda

Pada Gambar 5, dapat dilihat itu merupakan tampilan halaman home. Pada halaman home ini berisi shortcut untuk menuju melihat detail makanan terpopuler, menu cari serta dapat melihat informasi terkait apa itu defisit kalori. Pada halaman ini juga dilengkapi dengan menu bar untuk mengarahkan pengguna menuju fitur atau menu lainnya.

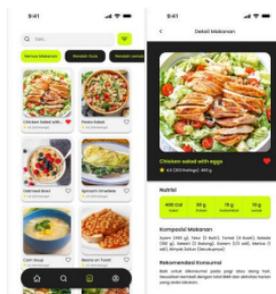
e. Tampilan Halaman Pencarian Rekomendasi Makanan



Gambar 6. Tampilan Halaman Pencarian Rekomendasi Makanan

Pada Gambar 6, dapat dilihat itu merupakan tampilan dari halaman pencarian rekomendasi makanan. Halaman ini dapat dikatakan sebagai halaman inti atau utama dari gagasan pembuatan aplikasi ini. Pada halaman ini pengguna dapat mencari rekomendasi makanan yang ingin di cari dengan menginputkan jenis kelamin, umur, tinggi badan dan berat badan. Hal tersebut dilakukan untuk mencari tahu terlebih dahulu perhitungan BMI dan BMR pengguna. Perhitungan ini merupakan bagian yang sangat penting dilakukan agar pemberian rekomendasi makanan tidak melenceng. Pada halaman ini, tidak hanya rekomendasi makanan yang ditampilkan, aplikasi ini juga memberikan informasi terkait perhitungan pembagian kalori pengguna yang sesuai.

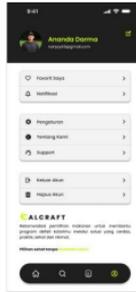
f. Tampilan Halaman Daftar dan Detail Makanan



Gambar 7. Tampilan Halaman Daftar dan Detail Makanan

Pada Gambar 7, menampilkan tampilan untuk halaman daftar makanan dan detail makanan. Pada bagian daftar makanan, pengguna nantinya dapat memfilter makanan sesuai kategori yang diinginkan. Untuk detail makanan menampilkan informasi tentang nutrisi dan komposisi bahan makanan, serta berisi rekomendasi pengkonsumsian.

g. Tampilan Halaman Profile Pengguna



Gambar 8. Tampilan Halaman Profile Pengguna

Pada Gambar 8, menampilkan tampilan profile yang juga di dalamnya berisi beberapa pilihan menu seperti pengaturan, notifikasi, daftar favorit hingga keluar dan hapus akun.

3.2. Hasil Pengujian dengan Metode System Usability Scale (SUS)

Dari hasil penyebaran kuesioner yang telah dilakukan, diperoleh data dengan total 35 data. Data ini selanjutnya akan diproses dengan melakukan perhitungan sesuai aturan dari metode System Usability Scale (SUS). Kemudian akan hasil perhitungan data tersebut akan dicari skor rata-ratanya untuk mengetahui tingkat usability perancangan UI/UX aplikasi "Calcraft". Berikut merupakan hasil yang diperoleh:

Tabel 3. Perhitungan Skor Rata-Rata System Usability Scale (SUS)

No	Responden	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	Jumlah	Skor SUS
1	R1	5	1	5	1	5	1	5	1	5	1	40	100
2	R2	5	2	5	2	5	1	5	1	5	2	37	92.5
3	R3	5	1	5	1	5	1	5	1	5	1	40	100
4	R4	5	1	5	2	5	1	5	2	5	1	38	95
5	R5	5	1	4	3	4	2	4	1	5	4	31	77.5
6	R6	3	3	3	3	3	3	3	3	3	3	20	60
7	R7	3	2	4	1	4	3	4	3	4	5	25	62.5
8	R8	5	1	5	2	5	1	4	1	5	3	36	90
9	R9	5	1	5	4	5	1	5	1	5	5	33	82.5
10	R10	4	2	4	3	4	2	4	2	3	3	27	67.5
11	R11	5	2	4	3	4	2	4	2	5	3	30	75
12	R12	5	1	5	1	5	1	4	1	5	4	36	90
13	R13	4	2	4	2	4	2	4	1	4	2	31	77.5
14	R14	5	1	5	1	5	1	5	1	5	2	39	97.5
15	R15	4	2	4	2	4	2	4	2	4	3	29	72.5

No	Responden	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	Jumlah	Skor SUS
16	R16	4	2	4	2	4	2	4	3	4	5	26	65
17	R17	5	1	5	1	5	1	5	1	5	2	39	97.5
18	R18	5	1	4	1	5	2	5	2	5	3	35	87.5
19	R19	5	2	4	3	5	1	4	1	4	2	33	82.5
20	R20	4	2	3	4	5	1	5	1	5	3	31	77.5
...	R...
35	R35	5	1	4	1	4	2	4	1	4	1	35	87.5
Skor Akhir SUS												83.07142857	

Dari hasil pengamatan tabel diatas, dapat dideskripsikan bahwa total skor SUS yang didapatkan sebesar 2907.5 dengan skor rata-rata yang diperoleh juga yaitu sebesar 83,07142857. Jika nilai rata-rata melampaui angka 68, maka sistem dapat dikatakan sudah memenuhi standar layak untuk dikembangkan. Dengan demikian, sesuai dengan pernyataan dan tabel interpretasi nilai SUS diatas, aplikasi rekomendasi pemilihan makanan untuk program defisit kalori "Calcraft" sudah layak dan sesuai aturan untuk dapat dilakukan pengembangan, yang dimana juga dengan skor rata-rata sebesar 83,07142857, aplikasi Calcraft sudah termasuk ke kategori "A" atau "Excellent" sesuai dengan perhitungan metode System Usability Scale (SUS) yang diimplementasikan.

4. Kesimpulan

Keresahan terkait pola hidup yang kurang baik dengan banyaknya pengkonsumsian makanan sembarang menjadi suatu permasalahan dalam penelitian ini. Solusi yang ditawarkan yaitu dengan melakukan sebuah program defisit kalori sesuai dengan knowledge base yang merujuk pada perhitungan BMI (Body Mass Index) dan BMR (Basal Metabolic Rate). Selanjutnya gagasan dari penulis adalah mengembangkan sebuah aplikasi yang terintegrasi dengan sistem rekomendasi untuk pemilihan makanan khusus untuk menunjang program defisit kalori. Metode System Usability Scale (SUS) digunakan sebagai alat untuk mengukur keberhasilan aplikasi tersebut, yang dinamakan "Calcraft". Dengan 35 responden mengisi kuesioner yang mencakup 10 pertanyaan SUS, total nilai yang diperoleh adalah 2907.5, dengan rata-rata 83.07142857. Menurut interpretasi tabel SUS, skor ini termasuk kategori "A" atau "Excellent". Dari hasil dan pembahasan ini, dapat disimpulkan bahwa aplikasi "Calcraft" layak untuk dilanjutkan dan dikembangkan di masa depan. Untuk pengembangan selanjutnya, dapat dilakukan pengimplementasian beberapa metode atau algoritma yang mampu dalam mendukung pengembangan sistem rekomendasi ini contohnya seperti algoritma Collaborative Filtering atau Content Based Filtering. Metode-metode ini dapat digunakan sebagai sarana untuk meningkatkan akurasi rekomendasi makanan yang dilakukan dengan mempelajari preferensi dan pola makan pengguna dari waktu ke waktu.

Daftar Pustaka

- [1] CNN Indonesia. "Obesitas Meningkat di Indonesia, Kenaikan Income Jadi Biang Kerok," 25 July 2023. [Online]. Available: <https://www.cnnindonesia.com/gaya-hidup/20230725090151-255-977373/obesitas-meningkat-di-indonesia-kenaikan-income-jadi-biang-kerok>. [6 May 2024].
- [2] Kementerian Kesehatan RI, "Factsheet Obesitas Kit Informasi Obesitas", 8 February 2018. [Online]. Available: <https://p2ptm.kemkes.go.id/dokumen-ptm/factsheet-obesitas-kit-informasi-obesitas>. [6 May 2024]
- [3] L A Arini and I K Wijana, "Korelasi Antara Body Mass Index (BMI) Dengan Blood Pressure (BP) Berdasarkan Ukuran Antropometri Pada Atlet," Jurnal Kesehatan Perintis (Perintis's Health Journal), vol. 7, no. 1, pp. 32-40, 2020.

- [4] N N S Yuliani and E Trinovita, "Korelasi Usia Metabolik Terhadap Indeks Massa Tubuh," *Jurnal Surya Medika*, vol. 5, no. 2, pp 35-41, 2020.
- [5] T. Saras, "Kalori: Panduan Lengkap Tentang Energi dalam Nutrisi," Tiram Media, 2023.
- [6] UN Sustainable Development Goals, "Sustainable Development Goals", [Online]. Available: <https://sdgs.un.org/goals>.
- [7] R. A. Alfaresy and C. I. Ratnasari, "Website Evaluation of The Faculty of Industrial Technology Universitas Islam Indonesia Using the System Usability Scale Method," *Jurnal Riset Informatika*, vol. 5, no. 3, pp. 285-294, 2023. [Online]. Available: <https://doi.org/10.34288/jri.v5i3.220>. [6 May 2024].
- [8] I. A. G. R. W. Astari and I. N. T. Putra, "Analisis Sistem Informasi Kemdikbud pada SD Negeri 2 Dawan Klod Dengan System Usability Scale," *JIKO (Jurnal Inform. dan Komputer)*, vol. 4, no. 1, pp. 23-30, 2021.
- [9] W. Buana and B. N. Sari, "Analisis User Interface Meningkatkan Pengalaman Pengguna Menggunakan Usability Testing pada Aplikasi Android Course," vol. 5, no. 2, pp. 91-97, 2022.

Analisis dan Klasifikasi Genre Musik Menggunakan Algoritma STFT dan Random Forest

Merry Royanti Manalu^{a1}, Made Agung Raharja^{a2}

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana

Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia

¹manalu.2208561069@student.unud.ac.id

²made.agung@unud.ac.id

Abstract

This research analyzes the classification of music genres using the Short Time Fourier Transform (STFT) algorithm. The main objective is to identify the effectiveness of STFT, along with the Random Forest classification algorithm, in distinguishing music genres based on their spectral characteristics. The STFT method is utilized to transform audio signals into a spectral representation within a short time window. The extracted spectral features are then fed into the Random Forest classification algorithm to classify different music genres. This research involves the use of representative datasets from various music genres for performance evaluation. Experimental results show that using STFT as a feature and employing the Random Forest classification algorithm in the process are able to provide satisfactory results in distinguishing music genres, with an accuracy of 86%. These findings demonstrate the potential of STFT, in combination with Random Forest, as a useful tool in music analysis and automatic classification of music genres.

Keywords: *Classification, Music Genres, Spectral Characteristics, Short-Time Fourier Transform (STFT), Random Forest*

1. Pendahuluan

Musik adalah salah satu bentuk seni yang paling universal dan bervariasi di dunia. Menurut KBBI, musik merupakan nada atau suara yang disusun sedemikian rupa sehingga mengandung irama, lagu, dan keharmonisan (terutama yang menggunakan alat-alat yang dapat menghasilkan bunyi-bunyi itu) [1]. Dengan variasi genre musik yang beragam, analisis genre musik telah menjadi subjek penting dalam bidang Music Information Retrieval (MIR). Pengklasifikasian genre musik menjadi langkah awal yang sangat penting dalam upaya untuk memahami dan mengelompokkan berbagai macam musik yang tersedia. Pelabelan terhadap musik merupakan salah satu hal yang penting dalam Music Information Retrieval, karena kunci utama dari Music Information Retrieval adalah klasifikasi berdasarkan genre, artis atau yang lainnya. Pada era digital seperti saat ini, penambahan musik pada internet menjadi sangat banyak dan cepat oleh karena itu sangat tidak efektif apabila menggunakan pelabelan genre secara manual [2]. Dalam upaya untuk mengotomatisasikan proses pengklasifikasian genre musik, banyak penelitian telah dilakukan untuk mengembangkan algoritma komputer yang dapat mengidentifikasi genre musik secara otomatis dari sinyal audio. Salah satu pendekatan yang telah terbukti efektif adalah penggunaan metode Short Time Fourier Transform (STFT) dalam mengekstraksi fitur-fitur spektral dari sinyal audio. Penelitian ini bertujuan untuk menganalisis kinerja klasifikasi genre musik dengan memanfaatkan algoritma STFT dan Random Forest. Dengan mentransformasikan sinyal audio ke dalam domain frekuensi-waktu menggunakan STFT, penelitian ini dapat mengekstraksi fitur-fitur spektral yang mencerminkan karakteristik unik dari setiap genre musik. Selanjutnya, fitur-fitur ini akan dimasukkan ke dalam algoritma klasifikasi Random Forest untuk mengidentifikasi dan mengklasifikasikan genre musik secara otomatis. Alasan penulis menggunakan metode STFT adalah karena kemampuannya untuk menganalisis sinyal yang berubah-ubah seiring waktu, seperti musik. STFT menghasilkan representasi frekuensi-waktu yang memungkinkan melihat

bagaimana frekuensi sinyal berubah selama durasi tertentu. Hal ini sangat penting dalam musik karena memungkinkan kita menangkap detail frekuensi yang bervariasi seiring waktu, yang merupakan karakteristik penting dari berbagai genre musik. Algoritma Random Forest dipilih sebagai model klasifikasi karena kelebihanannya dalam menangani dataset yang kompleks dan beragam. Random Forest, yang merupakan kumpulan dari pohon keputusan, memiliki kemampuan untuk mengurangi overfitting dan meningkatkan generalisasi melalui mekanisme voting. Algoritma ini juga efektif dalam menangani data dengan banyak fitur dan dapat memberikan estimasi pentingnya fitur, sehingga membantu dalam memahami kontribusi masing-masing fitur dalam proses klasifikasi. Kombinasi metode STFT dan Random Forest menawarkan pendekatan yang kuat dan efisien untuk klasifikasi genre musik. STFT menyediakan fitur-fitur spektral yang kaya dan representatif, sementara Random Forest memberikan kinerja klasifikasi yang andal dan interpretatif.

2. Metode Penelitian

2.1 Studi Literatur

Pada tahapan ini, penulis mengeksplorasi penelitian terbaru yang membahas teknik-teknik ekstraksi fitur audio, klasifikasi genre musik, serta aplikasi Short Time Fourier Transform (STFT) dalam konteks analisis musik. Melalui pemahaman mendalam tentang kerangka kerja konseptual dan temuan terbaru dalam literatur, penelitian ini akan membentuk landasan teoritis yang kuat untuk analisis klasifikasi genre musik. Selain itu, penelitian ini juga mempertimbangkan penerapan metode klasifikasi Random Forest, yang merupakan salah satu algoritma yang efektif dalam pengklasifikasian data kompleks, untuk meningkatkan kinerja model klasifikasi dalam konteks ini. Proses metode STFT melibatkan beberapa langkah utama. Pertama, sinyal audio dibagi menjadi segmen-segmen pendek yang saling tumpang tindih. Panjang segmen dan jumlah tumpang tindih ini ditentukan sesuai kebutuhan analisis. Kemudian, setiap segmen diubah dari domain waktu ke domain frekuensi menggunakan Transformasi Fourier, menghasilkan spektrum frekuensi untuk setiap segmen. Hasil dari setiap segmen ini digabungkan untuk membentuk representasi waktu-frekuensi dari sinyal asli, yang menunjukkan bagaimana distribusi frekuensi berubah seiring waktu. Representasi ini memungkinkan penangkapan detail perubahan frekuensi yang tidak dapat ditangkap oleh Transformasi Fourier biasa, menjadikannya sangat cocok untuk menganalisis musik yang memiliki variasi frekuensi dinamis. Selanjutnya, dilakukan klasifikasi menggunakan metode Random Forest, yang terdiri dari beberapa tahap penting. Pertama, algoritma membentuk banyak pohon keputusan dari subset acak data pelatihan. Setiap pohon dibuat menggunakan subset fitur dan sampel data yang dipilih secara acak. Selama tahap pelatihan, setiap pohon keputusan dilatih untuk memprediksi kelas target menggunakan subset data tersebut, dan belajar membuat keputusan berdasarkan fitur-fitur yang ada. Untuk melakukan prediksi, setiap pohon dalam hutan memberikan suaranya (prediksinya), dan prediksi akhir adalah kelas yang mendapatkan suara terbanyak dari semua pohon. Selain itu, Random Forest juga mampu memberikan estimasi pentingnya setiap fitur dalam proses klasifikasi, yang membantu memahami kontribusi masing-masing fitur terhadap hasil akhir.

2.2 Pengumpulan Data

Pada tahapan ini, penulis menggunakan data sekunder yang diperoleh dari platform Kaggle, yaitu "GTZAN Dataset" tersedia dalam format .zip dan terdiri dari 1000 file audio dari 10 genre, masing-masing berdurasi 30 detik. Dataset tersebut juga berisi gambar spektrogram dari setiap audio, juga file CSV berisi atribut dari setiap audio. Atribut-atribut yang terdapat dalam dataset yang digunakan seperti, panjang lagu, statistik spektrum warna, sinyal RMS, pusat spektral, lebar pita spektral, serta koefisien MFCC. Dataset ini telah terverifikasi dan dirinci berdasarkan genre musiknya, memungkinkan peneliti untuk melakukan analisis klasifikasi genre dengan menggunakan algoritma Short Time Fourier Transform (STFT) dan Random Forest.

Tabel 1. Contoh Dataset

filename	length	chroma_stft_mean	rms_mean	rms_var
blues.00000.wav	661794	0.350088	0.130228	0.002827
blues.00001.wav	661794	0.340914	0.095948	0.002373
blues.00002.wav	661794	0.363637	0.17557	0.002746
blues.00003.wav	661794	0.404785	0.141093	0.006346
blues.00004.wav	661794	0.308526	0.091529	0.002303
—	—	—	—	—

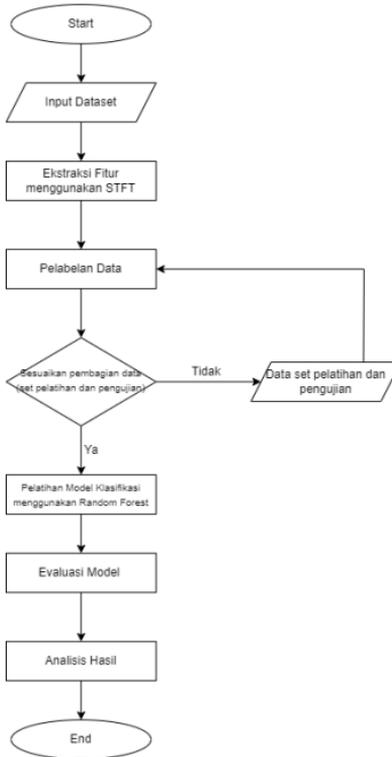
2.3 Ekstraksi Fitur

Pada tahapan ini, akan dilakukan proses ekstraksi fitur dari data audio menggunakan metode Short Time Fourier Transform (STFT). STFT digunakan untuk mentransformasikan sinyal audio menjadi domain frekuensi-waktu, memungkinkan untuk mengekstraksi informasi spektral dari setiap jendela waktu. Rumus STFT dinyatakan sebagai berikut:

$$X(m, \omega) = \sum_{n=0}^{N-1} x(n)w(n-m)e^{-j\omega n} \quad (1)$$

di mana $X(m, \omega)$ adalah koefisien STFT pada waktu ke- m dan frekuensi ω , $x(n)$ adalah sinyal audio, $w(n)$ adalah fungsi jendela, N adalah panjang jendela, dan ω adalah frekuensi sudut. Setelah ekstraksi fitur dengan STFT, hasilnya akan digunakan sebagai masukan untuk model klasifikasi. Pada tahap ini, metode Random Forest akan diterapkan untuk melatih model klasifikasi berdasarkan fitur-fitur yang diekstraksi.

2.4 Flowchart Sistem



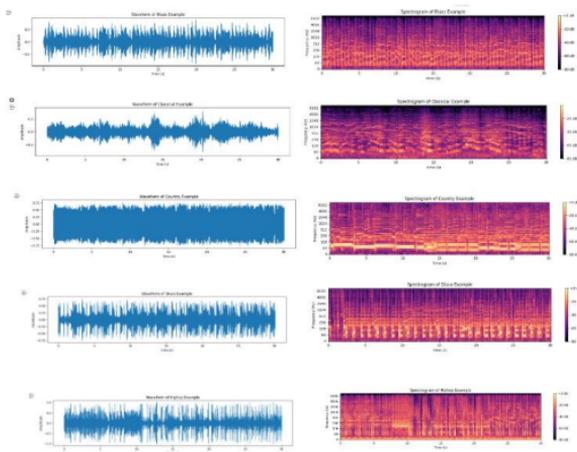
Gambar 1. Flowchart Alur Sistem

Sistem ini dimulai dengan input data yang sebelumnya sudah didapatkan dari platform Kaggle. Setelah dataset disiapkan untuk analisis, langkah selanjutnya adalah ekstraksi fitur menggunakan metode Short Time Fourier Transform (STFT). STFT digunakan untuk mentransformasikan sinyal audio ke dalam domain frekuensi-waktu, sehingga fitur-fitur spektral dapat diekstraksi. Setelah fitur-fitur diekstraksi, tahap berikutnya adalah pelabelan data. Selanjutnya, data dibagi menjadi set pelatihan dan set pengujian. Pada tahap ini, keputusan diambil berdasarkan apakah pembagian data sudah memenuhi kriteria yang ditetapkan. Jika ya, maka data akan lanjut ke tahap klasifikasi menggunakan algoritma Random Forest. Namun, jika pembagian data belum optimal, maka data tersebut akan kembali ke tahap pelabelan untuk dilakukan koreksi. Setelah data dibagi dengan benar dan model klasifikasi dilatih, langkah selanjutnya adalah evaluasi model. Model akan diuji menggunakan set pengujian, dan metrik evaluasi performa seperti akurasi, presisi, recall, dan F1-score akan dihitung. Jika performa model memenuhi kriteria yang ditetapkan, maka analisis hasil dilakukan untuk mengevaluasi kinerja sistem dan menginterpretasi hasil klasifikasi. Ketika semua tahap telah selesai dieksekusi,

sistem akan mencapai akhir (End). Pada setiap tahap, keputusan diambil berdasarkan hasil dari langkah sebelumnya, dan jika diperlukan, tindakan koreksi dilakukan untuk meningkatkan kinerja sistem.

3. Hasil dan Pembahasan

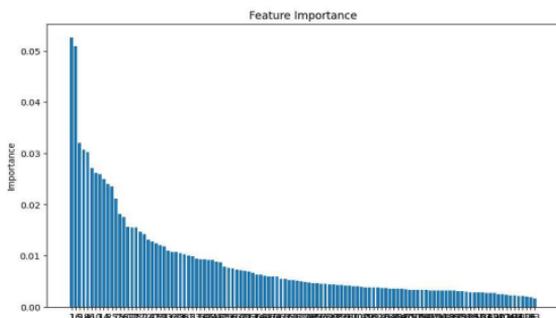
Program analisis dan klasifikasi genre musik ini menggunakan algoritma Short-Time Fourier Transform (STFT) untuk ekstraksi fitur dan algoritma Random Forest untuk proses klasifikasinya. Tujuan dari penelitian ini adalah mengembangkan model yang dapat secara akurat mengidentifikasi genre musik dari data audio. Dalam bagian ini, akan dipresentasikan berbagai visualisasi dan metrik evaluasi untuk menilai performa model yang telah dibangun. Tahap pertama yang dilakukan adalah analisis terhadap waveform dan spektrogram dari lima genre musik: blues, classical, country, disco, dan hiphop. Visualisasi ini bertujuan untuk memberikan pemahaman yang lebih mendalam tentang pola frekuensi dan distribusi energi dalam setiap genre musik. Dengan menampilkan waveform, dapat diamati fluktuasi amplitudo dari sinyal audio, sementara spektrogram memberikan informasi tentang bagaimana energi sinyal audio didistribusikan pada berbagai frekuensi sepanjang waktu. Analisis ini penting untuk mengidentifikasi karakteristik unik dari setiap genre yang dapat digunakan dalam proses klasifikasi. Berikut adalah sampel hasil visualisasi waveform dan spektrogram untuk masing-masing genre yang dianalisis.



Gambar 2. Sampel Waveform Dan Spektrogram Dari Lima Genre

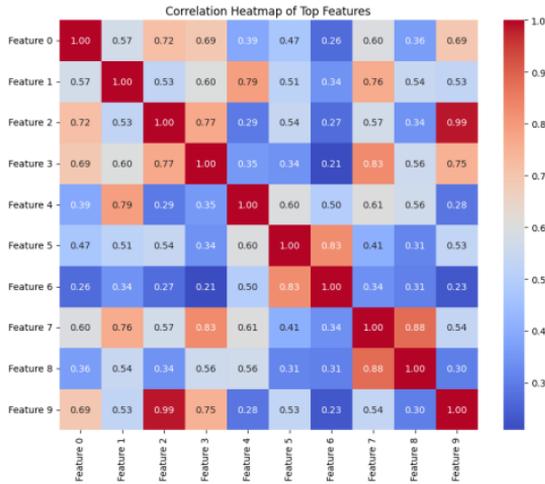
Gambar 2 menampilkan waveform dan spektrogram untuk sampel lagu dari masing-masing genre yang dianalisis, yaitu blues, classical, country, disco, dan hiphop. Waveform menggambarkan perubahan amplitudo sinyal audio dari waktu ke waktu, sedangkan spektrogram menunjukkan distribusi frekuensi dari sinyal audio tersebut. Dengan mengamati kedua representasi ini, dapat dipahami bagaimana pola frekuensi dan energi bervariasi di antara genre-genre musik yang berbeda. Genre blues menunjukkan pola frekuensi yang khas dengan fluktuasi amplitudo yang menonjol pada frekuensi rendah hingga menengah. Genre classical memperlihatkan pola yang lebih kompleks dan beragam dengan distribusi frekuensi yang luas, menunjukkan energi yang tersebar dari frekuensi rendah hingga tinggi. Genre country memiliki karakteristik yang mirip dengan genre blues, namun dengan pola frekuensi yang lebih terstruktur. Genre disco menunjukkan energi yang kuat pada frekuensi menengah hingga tinggi,

mencerminkan ritme dan tempo cepat yang khas dari genre ini. Sementara itu, genre hiphop memperlihatkan pola amplitudo yang berulang dengan energi yang terkonsentrasi pada frekuensi rendah, sesuai dengan karakteristik bass-heavy dari genre ini. Tahap selanjutnya adalah, sinyal audio dari setiap lagu dibagi menjadi segmen-segmen pendek yang saling tumpang tindih, kemudian setiap segmen diubah dari domain waktu menjadi domain frekuensi menggunakan Transformasi Fourier. Proses ini menghasilkan representasi waktu-frekuensi dari sinyal asli, yang memungkinkan penangkapan detail perubahan frekuensi seiring waktu. Fitur-fitur ini kemudian digunakan sebagai input untuk proses klasifikasi menggunakan algoritma Random Forest. Setelah fitur-fitur STFT diekstraksi, langkah selanjutnya adalah membentuk model klasifikasi menggunakan algoritma Random Forest. Algoritma ini dipilih karena kemampuannya yang baik dalam menangani dataset yang kompleks dan beragam. Model Random Forest dibentuk dengan menggunakan parameter default, yang mencakup jumlah pohon keputusan (estimators) dan kedalaman maksimum pohon (max depth). Parameter-parameter ini dapat dioptimalkan lebih lanjut untuk meningkatkan performa model. Setelah model Random Forest dibentuk, langkah berikutnya adalah melatih model tersebut dengan menggunakan data latih (training data) yang terdiri dari fitur-fitur STFT dan label genre musik. Proses pelatihan ini memungkinkan model untuk belajar dari data dan membentuk pohon keputusan yang digunakan untuk memprediksi genre musik berdasarkan fitur-fitur yang ada.



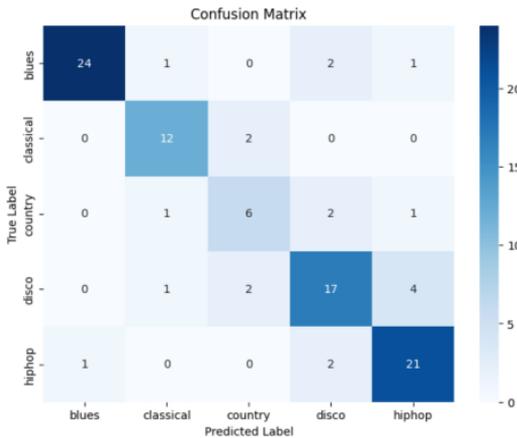
Gambar 3. Feature Importance

Tahap selanjutnya adalah analisis pentingnya fitur-fitur yang digunakan dalam model klasifikasi dengan algoritma Random Forest. Feature importance dihasilkan oleh model Random Forest, yang menunjukkan kontribusi setiap fitur dalam proses klasifikasi genre musik. Gambar 3 menunjukkan ranking fitur berdasarkan tingkat kepentingannya. Fitur-fitur yang paling penting memberikan kontribusi terbesar terhadap akurasi model dalam mengklasifikasikan genre musik. Analisis ini membantu memahami fitur mana yang paling berpengaruh dan bagaimana model menggunakan informasi tersebut untuk membuat prediksi.



Gambar 4. Correlation Heatmap

Untuk memahami hubungan antara berbagai fitur yang digunakan dalam model, diperlukan heatmap korelasi dari sepuluh fitur teratas berdasarkan ranking feature importance. Heatmap ini menunjukkan korelasi antara setiap pasangan fitur, yang membantu dalam mengidentifikasi fitur-fitur yang mungkin redundant atau yang sangat berkaitan satu sama lain. Korelasi tinggi antara beberapa fitur menunjukkan bahwa ada fitur yang menyampaikan informasi yang mirip, yang dapat berguna dalam proses optimasi model untuk mengurangi redundansi dan meningkatkan efisiensi model.



Gambar 5. Confusion Matrix

Untuk mengevaluasi kinerja model dalam mengklasifikasikan genre musik, digunakan confusion matrix. Confusion matrix di atas memberikan gambaran visual tentang bagaimana model yang telah dibuat memprediksi kelas dibandingkan dengan label yang sebenarnya. Setiap baris dalam matriks menunjukkan contoh dari kelas sebenarnya, sementara setiap kolom menunjukkan contoh dari kelas yang diprediksi oleh model. Dengan menggunakan confusion matrix, dapat diidentifikasi seberapa sering model salah mengklasifikasikan genre tertentu dan seberapa akurat model dalam memprediksi setiap genre.

Akurasi Model : 0.86
 Laporan Klasifikasi :

	precision	recall	f1-score	support
blues	0.93	0.96	0.95	28
classical	0.87	0.93	0.90	14
country	0.86	0.60	0.71	10
disco	0.83	0.79	0.81	24
hiphop	0.81	0.88	0.84	24
accuracy			0.86	100
macro avg	0.86	0.83	0.84	100
weighted avg	0.86	0.86	0.86	100

Gambar 6. Laporan Klasifikasi

Dalam evaluasi klasifikasi yang dilakukan, model yang dikembangkan berhasil mencapai tingkat akurasi sebesar 86%. Hasil ini menunjukkan bahwa model memiliki kemampuan yang baik dalam mengklasifikasikan genre musik berdasarkan fitur-fitur yang diekstraksi. Dengan nilai akurasi sebesar ini, model dapat secara tepat memprediksi sebagian besar genre musik yang ada dalam dataset. Tingkat presisi, recall, dan F1-score yang terdapat dalam laporan klasifikasi memberikan gambaran lebih rinci tentang kinerja model untuk setiap kelas genre.

4. Kesimpulan

Penelitian ini berhasil melakukan analisis dan klasifikasi genre musik menggunakan algoritma STFT dan Random Forest. Model klasifikasi yang dikembangkan mencapai tingkat akurasi sebesar 86%, menunjukkan kemampuannya dalam memprediksi genre musik dengan tepat berdasarkan fitur-fitur yang diekstraksi. Hasil evaluasi juga mengungkapkan variasi performa model di antara kelas genre musik yang berbeda. Dengan demikian, penelitian ini memberikan kontribusi dalam memahami dan mengimplementasikan teknik analisis musik yang efektif untuk klasifikasi genre, dengan potensi untuk diterapkan dalam berbagai aplikasi dalam industri musik dan rekayasa audio.

Daftar Pustaka

- [1] KBBI. "Kamus Besar Bahasa Indonesia". [Online] Tersedia di: <https://kbbi.kemdikbud.go.id/>. [Diakses pada 5 Mei 2024]
- [2] Thoriq. Nurhaidir, Widodo. and Bambang. Prasetya. Adhi, "Klasifikasi Genre Musik Menggunakan Algoritma Naive Bayes Classifier Untuk Layanan Streaming Youtube" *Jurnal PINTER*, vol. 7, no. 1, p. 1-6, 2023.
- [3] Hilarius. C, Catherine. Olivia. S, Lanny. W. Pandjaitan dan Maria. Angela. Kartwidjaja, "Klasifikasi Genre Musik Dengan Menggunakan Metode Machine Learning" *Jurnal Ilmiah Teknik Elektro*, vol. 21, no. 1, p. 71-82, 2023.
- [4] I. Putra, I. Gusti. Kadyanan, "Pengenalan Nada Piano dengan Algoritma Short Time Fourier Transform (STFT)" *Jurnal Nasional Teknologi Informasi dan Aplikasinya*, vol. 2, no. 2, p. 409-413, 2024.
- [5] Muhammad. Sarofi, Irhamah dan Adatul. Makurromah, "Identifikasi Genre Musik dengan Menggunakan Metode Random Forest" *Jurnal Sains Dan Seni ITS*, vol. 9, no. 1, p. D79-D86, 2020.
- [6] Sally. Lutfiani, Triando. Hamonangan. Saragih, Friska. Abadi, Mohammad. Reza. Faisal dan Dwi. Kartini, "Perbandingan Metode Extreme Gradient Boosting Dan Metode Decision Tree Untuk Klasifikasi Genre Musik" *JIP (Jurnal Informatika Polinema)*, vol. 9, no. 4, p. 373-382, 2023.
- [7] Danny. Lionel, Rudy. Adipranata dan Endang. Setyati, "Klasifikasi Genre Musik Menggunakan Metode Deep Learning Convolutional Neural Network dan Mel-Spektrogram" *JURNAL INFRA*, vol. 7, no. 1, p. 1-5, 2019.
- [8] Gst. Giri, Made. Leo. Radhitya, "Klasifikasi Genre Musik Menggunakan Teknik Pembelajaran Mesin" *Jurnal Teknologi Informasi dan Komputer*, vol. 9, no. 1, p. 1-9, 2023.
- [9] Yisti. Via, Intan. Yuniar. Purbasari dan Aditya. Putra. Pratama, "Analisa Algoritma Convolution Neural Network (Cnn) Pada Klasifikasi Genre Musik Berdasar Durasi Waktu" *SCAN Jurnal Teknologi Informasi dan Komunikasi*, vol. XVII, no. 1, p. 35-41, 2022.
- [10] PPandu. Prasetyo, I. Gede. Wijaya dan Ario. Yudo. Husodo, "Klasifikasi Genre Musik Menggunakan Metode Mel Frequency Cepstrum Coefficients (Mfcc) Dan K-Nearest Neighbors Classifier" *JTIKA*, vol. 1, no. 2, p. 189-197, 2019.
- [11] cCarl. Wairata, Ericks. Rahmat. Swedia dan Margi. Cahyanti, "Pengklasifikasian Genre Musik Indonesia Menggunakan Convolutional Neural Network" *Jurnal SEBATIK*, vol. 25, no. 1, p. 255-261, 2021.

Halaman ini sengaja dibiarkan kosong

Pengamanan Gambar dengan Metode Cipher Block Chaining

Made Yayang Eka Prananda¹⁾, I Gusti Ngurah Anom Cahyadi Putra²⁾

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹yayangp32@gmail.com
²anom.cp@unud.ac.id

Abstract

Data security is something that everyone needs to maintain their privacy, for example images. Images are very familiar to everyone, because we can capture various moments in the form of images. But there are some people who steal someone's image to misuse it for personal gain. Image pixel encryption using cryptography can be an alternative in securing the image. In this study, the cipher block chaining method is used to encrypt each pixel. The results of using this method are quite satisfactory, the resulting image becomes more abstract. The encrypted image can also be decrypted into the original image. However, the resulting pattern is still the same as the previous image pattern. So the use of this method may be more effective when combined with other cryptographic methods.

Keywords: *Cipher Block Chaining, Encrypted, Decrypted, Photo, Pixel.*

1. Pendahuluan

Keamanan data merupakan suatu hal yang dibutuhkan semua orang untuk menjaga privasinya. Untuk itu banyak cara dilakukan agar privasi tersebut dapat terjaga dengan aman. Privasi seseorang dapat berupa teks, gambar maupun video. Gambar sangat familiar di dengar oleh semua orang, karena kita dapat mengabadikan berbagai momen dalam bentuk gambar. Mulai dari berlibur, bersama keluarga, bersama orang tersayang dapat kita abadikan momennya dalam bentuk gambar untuk dapat dikenang suatu hari nanti. Namun ada saja beberapa oknum yang memanfaatkan hal ini, dengan menggunakan gambar privasi untuk disalahgunakan. Seperti mengambil gambar privasi tersebut untuk disebarakan demi mendapat keuntungan tersendiri. Oleh karena itu dalam menyimpan gambar privasi kita harus bisa mengamatkannya. Mengamankan gambar dapat dilakukan dengan berbagai cara, seperti menyimpan di folder tersembunyi, memberi proteksi kepada folder tempat menyimpan gambar, dan lainnya. Namun ada satu cara yang dapat memberi proteksi yang sangat tinggi, yaitu dengan mengacak acak warna gambar tersebut dengan menggunakan algoritma kriptografi. Sehingga gambar asli yang sebelumnya sangat indah menjadi gambar yang sangat tidak jelas atau abstrak. Banyak algoritma kriptografi yang sangat efisien dalam mengenkripsi setiap pixel gambar tersebut salah satunya Cipher Block Chaining. Cipher Block Chaining digunakan karena merupakan salah satu algoritma modern yang menggunakan block cipher dan dalam melakukan enkripsi menggunakan inialisasi variabel yang variabel terkait. Sehingga apabila ada salah satu pixel yang di rusak, pixel selanjutnya akan menjadi lebih tidak jelas. Diharapkan dengan menggunakan Cipher Block Chaining dalam melakukan enkripsi setiap pixel sebuah gambar, dapat membuat seseorang mengamankan gambarnya lebih aman.

2. Metode Penelitian

2.1 Algoritma Cipher Block Chaining

Menurut Menezes et al (1997) algoritma Cipher Block Chaining (CBC) merupakan salah satu metode kriptografi yang berbasis pada block, pada metode ini mempunyai kelebihan setiap block

ciphertext bergantung tidak hanya pada block plaintextnya tetapi juga pada seluruh block plaintext sebelumnya. Mode operasi Cipher Block Chaining (CBC) merupakan salah satu mode operasi block cipher yang menggunakan vektor inialisasi (initialization vector/IV) dengan ukuran tertentu (ukurannya sama dengan satu blok plaintext). Pada mode operasi ini plaintext dibagi menjadi beberapa blok, kemudian masing-masing blok dienkripsi dengan ketentuan blok plaintext pertama dienkripsi lebih dahulu. Sebelum dienkripsi, plaintext di-XOR dengan IV. Lalu, hasil XOR tersebut dienkripsi hingga menghasilkan ciphertext. Selanjutnya, ciphertext tersebut digunakan sebagai IV untuk proses penyandian blok plaintext selanjutnya. Algoritma CBC merupakan kriptografi simetris jadi kunci digunakan dalam proses enkripsi dan juga dekripsi adalah kunci yang sama. Namun untuk enkripsi CBC memerlukan inialisasi variabel sebagai tambahan dalam melakukan enkripsi, untuk sistematis proses enkripsi CBC dapat dilihat pada gambar 2.1. Contoh dalam melakukan proses enkripsi, yaitu:

Proses Enkripsi

Plaintext: Nan
Key: ya
IV: 01100101

Ubah plaintext dan kunci ke bentuk biner:

Plaintext

- N: 01001110
- a: 01100001
- n: 01101110

Kunci

- y: 01111001
- a: 01100001

Atur plaintext dengan tiap kunci:

Plaintext N a n

Kunci y a y

Langkah-Langkah Enkripsi:

- C1 diperoleh dari:**
 $P1 \oplus C0 = 01001110 \oplus 01100101 = 00101011P1$ \oplus plus $C0 = 01001110$ \oplus plus $01100101 = 00101011$
Enkripsi hasil tersebut (C1) dengan fungsi E:
 $C1 \oplus K1 = 00101011 \oplus 01111001 = 01010010C1$ \oplus plus $K1 = 00101011$ \oplus plus $01111001 = 01010010$
Geser hasilnya satu bit ke kiri: 1010010010100100
- C2 diperoleh dari:**
 $P2 \oplus C1 = 01100001 \oplus 01010010 = 11000101P2$ \oplus plus $C1 = 01100001$ \oplus plus $01010010 = 11000101$
Enkripsi hasil tersebut (C2) dengan fungsi E:
 $C2 \oplus K2 = 11000101 \oplus 01100001 = 10100100C2$ \oplus plus $K2 = 11000101$ \oplus plus $01100001 = 10100100$
Geser hasilnya satu bit ke kiri: 0100100101001001
- C3 diperoleh dari:**
 $P3 \oplus C2 = 01101110 \oplus 01001001 = 00100111P3$ \oplus plus $C2 = 01101110$ \oplus plus $01001001 = 00100111$

Enkripsi hasil tersebut (C3) dengan fungsi E:
 $C3 \oplus K3 = 001001111 \oplus 01111001 = 01011110C3$ \oplus plus K3 = 00100111 \oplus plus 01111001 = 01011110
Geser hasilnya satu bit ke kiri: 101110010111100

Hasil Enkripsi

Plaintext	Binary	Ciphertext
N	01001110	10100100
a	01100001	01001001
n	01101110	10111100

Proses Dekripsi

Ciphertext: 10100100 01001001 10111100

Key: ya

IV: 01100101

Ubah kunci ke bentuk biner:

Kunci

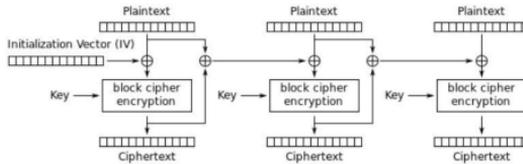
- y: 01111001
- a: 01100001

Langkah-Langkah Dekripsi:

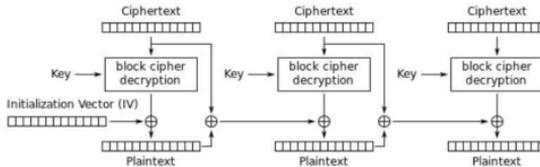
- P1 diperoleh dari:**
Dekripsi C1 dengan fungsi D:
Geser ciphertext satu bit ke kanan: 0101001001010010
 $C1 \oplus K1 = 010100100 \oplus 01111001 = 00101011C1$ \oplus plus K1 = 01010010 \oplus plus 01111001 = 00101011
Hasil dekripsi tersebut di XOR dengan C0C0:
 $P1 \oplus C0 = 00101011 \oplus 01100101 = 01001110P1$ \oplus plus C0 = 00101011 \oplus plus 01100101 = 01001110
- P2 diperoleh dari:**
Dekripsi C2 dengan fungsi D:
Geser ciphertext satu bit ke kanan: 1010010010100100
 $C2 \oplus K2 = 10100100 \oplus 01100001 = 11000101C2$ \oplus plus K2 = 10100100 \oplus plus 01100001 = 11000101
Hasil dekripsi tersebut di XOR dengan C1C1:
 $P2 \oplus C1 = 11000101 \oplus 10100100 = 01100001P2$ \oplus plus C1 = 11000101 \oplus plus 10100100 = 01100001
- P3 diperoleh dari:**
Dekripsi C3 dengan fungsi D:
Geser ciphertext satu bit ke kanan: 010111001011110
 $C3 \oplus K3 = 01011110 \oplus 01111001 = 00100111C3$ \oplus plus K3 = 01011110 \oplus plus 01111001 = 00100111
Hasil dekripsi tersebut di XOR dengan C2C2:
 $P3 \oplus C2 = 00100111 \oplus 01001001 = 01101110P3$ \oplus plus C2 = 00100111 \oplus plus 01001001 = 01101110

Hasil Dekripsi

Ciphertext	Binary	Plaintext
10100100	01001110	N
01001001	01100001	a
10111100	01101110	n



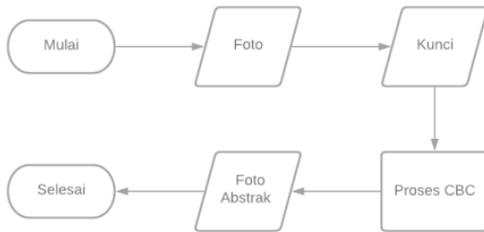
Gambar 2.1 Flowchart Proses Enkripsi Cipher Block Chaining



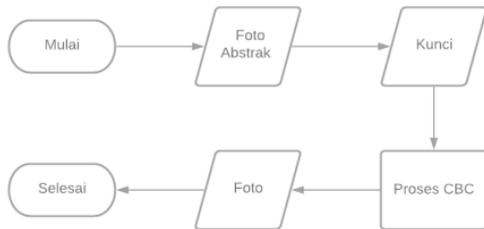
Gambar 2.2 Flowchart Proses Dekripsi Cipher Block Chaining

2.2 Flowchart

Pada penelitian ini, akan menggunakan Cipher Block Chaining sebagai enkripsi maupun dekripsi gambar yang ingin diamankan. Namun format gambar yang digunakan dibatasi dengan hanya menggunakan format png. Untuk proses melakukan enkripsi dimulai dengan pengguna menginput gambar dan kata sandi untuk mengenkripsi, lalu program akan melakukan enkripsi setiap pixel, outputnya berupa gambar baru yang sudah tidak terlihat jelas atau abstrak. Flowchart dalam melakukan proses enkripsi dapat dilihat pada gambar 2.3. Untuk proses melakukan dekripsi dimulai dengan pengguna menginput gambar abstrak dan kata sandi yang digunakan dalam melakukan enkripsi, lalu program akan melakukan dekripsi setiap pixel, outputnya berupa gambar privasi yang sebelumnya di enkripsi. Untuk flowchart proses dekripsi dapat dilihat pada gambar 2.4.



Gambar 2.3 Flowchart Enkripsi



Gambar 2.4 Flowchart Dekripsi

3. Hasil dan Pembahasan

Pada penelitian ini menggunakan beberapa gambar sebagai objek untuk di enkripsi. Gambar yang digunakan dibatasi dengan format png. Hasil enkripsi beserta kunci yang digunakan dapat dilihat pada tabel 3.1. Dimana ketika di dekripsi menggunakan kunci yang sama akan menghasilkan gambar asli yang di enkripsi.

Tabel 3.1 Hasil Enkripsi dan Dekripsi

Hasil Enkripsi	Kunci	Hasil Dekripsi
	harimau	
	anjing	

Hasil Enkripsi	Kunci	Hasil Dekripsi
	winniehepooh	
	earthstar	
	tyranosaurus	
	hapiness crocodile	
	elephanthead	

Pada penelitian ini dapat dilihat penggunaan Cipher Block Cipher dalam mengenkripsi pixel pada gambar masih belum maksimal. Dikarenakan secara kasat mata, hasil enkripsi masih berbentuk gambar aslinya, sehingga untuk mengamankan gambar tersebut masih kurang apabila menggunakan metode Cipher Block Chaining.

4. Kesimpulan

Pada penelitian ini dapat disimpulkan bahwa, penggunaan kriptografi dalam mengamankan sebuah gambar dengan cara mengenkripsinya sudah menghasilkan hasil yang cukup memuaskan. Penggunaan Cipher Block Chaining juga sudah menghasilkan gambar dengan warna yang berantakan. Gambarnya pun dapat dikembalikan ke gambar aslinya. Namun hasil enkripsi dengan metode ini masih kurang maksimal dikarenakan gambar hasil enkripsi masih memiliki pola yang sama dengan gambar aslinya. Ada baiknya apabila metode ini di kombinasikan dengan metode kriptografi modern lainnya untuk menghasilkan gambar enkripsi yang lebih abstrak. Sehingga tingkat keamanan gambar tersebut semakin tinggi.

Daftar Pustaka

- [1] H. Sahara, "Implementasi Pengamanan Pesan Chatting menggunakan Metode Vigenere Cipher dan Cipher Block Chaining," MEANS (Media Inf. Anal. dan Sist., vol. 3, no. 2, pp. 173–178, 2018.
- [2] K. Informatika. "Perhitungan Manual Cipher Block Chaining Untuk Kriptografi Teks" available:
<http://www.kitainformatika.com/2019/05/perhitungan-manual-algoritma-chiper.html>.
(accessed Oct. 01 2021).
- [3] D. Naufal, "Penjelasan Mengenai Cipher Block Chaining (CBC) dan Message Authentication Code (MAC), 29 November 2018." available:
<https://medium.com/@dimas.naufal11/penjelasan-mengenai-cipher-block-chaining-cbc-dan-message-authentication-code-mac-39074d11f816>. (accessed Oct. 01 2021).

Halaman ini sengaja dibiarkan kosong

Pengembangan Model Ontologi pada Domain Oleh-Oleh Khas Bali

Putu Ardi Sudarmika^{a1}, Nguhrah Agus Sanjaya ER^{a2}

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹ardisudarmika94@email.com
²agus_sanjaya@unud.ac.id

Abstract

Bali's typical souvenirs not only represent a high cultural heritage but also serve as a primary attraction for tourists. However, detailed information on these souvenirs remains limited, making it challenging for users to choose. To enhance accessibility, the concept of representing Bali's souvenir data in ontology-based knowledge emerged. The aim is to provide references to tourists by regularly structuring an ontology model. Knowledge about Bali's souvenirs is represented using RDF in triple concepts of subject, predicate, and object, accessible through SPARQL. The development process employs the methodology methodology, encompassing specification, knowledge acquisition, conceptualization, integration, implementation, evaluation, and documentation phases. The outcome is an ontology model featuring 4 classes, 3 Object Properties, data properties, 34 ontographs, and 28 individuals or instances, offering regularly relevant information about Bali's souvenirs. This method utilizes Protégé 5.5.0 software and a search-based application for testing its efficacy.

Keywords: Bali, Ontology, methodology, RDF, SPARQL

1. Pendahuluan

Bali adalah salah satu destinasi wisata terkemuka di Indonesia, yang terkenal dengan kekayaan budaya dan seni tradisionalnya. Salah satu aspek yang tak terpisahkan dari pengalaman wisata di Bali adalah oleh-oleh khasnya yang beragam. Pulau ini menjadi daya tarik bagi wisatawan domestik maupun mancanegara yang ingin menikmati keunikan dan keramahan masyarakatnya. Namun, dalam menghadapi perkembangan industri pariwisata yang pesat, tantangan muncul bagi wisatawan yang baru pertama kali mengunjungi Bali, seringkali sulit untuk mendapatkan informasi yang lengkap dan akurat tentang oleh-oleh khas Bali yang menarik. Informasi yang mereka dapatkan mungkin hanya dari cerita teman atau dari agen perjalanan wisata yang tidak selalu memberikan detail yang cukup [1]. Dalam era digital saat ini, pendekatan baru yang memanfaatkan teknologi informasi menjadi semakin penting dalam memberikan solusi untuk memenuhi kebutuhan akan informasi yang akurat dan dapat diandalkan bagi para wisatawan. Dengan memanfaatkan teknologi informasi berbasis internet, calon wisatawan akan sangat terbantu di dalam melakukan pencarian informasi oleh oleh khas bali [2]. Salah satu pendekatan yang menjanjikan adalah penggunaan pemodelan ontology, yang memungkinkan representasi pengetahuan yang terstruktur dan interoperabilitas data di dalam domain yang kompleks seperti oleh-oleh khas Bali [1]. Penelitian ini bertujuan untuk mengembangkan sebuah model ontology yang memetakan secara sistematis pengetahuan tentang oleh-oleh khas Bali. Dengan memanfaatkan konsep ontology, diharapkan bahwa model yang dihasilkan akan memberikan landasan yang kokoh untuk pengorganisasian informasi, memfasilitasi pencarian yang lebih efisien, dan meningkatkan pengalaman wisatawan dalam memilih dan memperoleh oleh-oleh khas Bali yang sesuai dengan preferensi dan kebutuhan mereka. Meskipun penelitian terkait dalam domain ini telah menyoroti pentingnya penggunaan ontology dalam berbagai konteks, mulai dari representasi pengetahuan di bidang pariwisata hingga sistem pencarian informasi. Namun, penelitian ini lebih spesifik yang berfokus pada pengembangan ontology untuk oleh-oleh khas Bali masih terbatas. Oleh karena itu, kontribusi penelitian ini diharapkan dapat melengkapi

dan memperkaya literatur yang ada, serta memberikan wawasan yang berharga bagi pengembangan sistem informasi pariwisata yang lebih canggih dan efektif di masa depan.

1.1 Oleh-Oleh Khas Bali

Oleh-oleh khas Bali adalah berbagai macam produk yang dijual di Bali, Indonesia, yang biasanya dibeli oleh wisatawan sebagai kenang-kenangan atau sebagai hadiah untuk keluarga dan teman di rumah. Oleh-oleh khas Bali sangat beragam dan dapat berupa pakaian, jajanan, aksesoris, makanan, minuman, atau souvenir yang unik dan khas dari Bali. Oleh-oleh khas Bali ini biasanya dibuat dengan bahan-bahan lokal dan memiliki ciri khas yang berbeda dengan produk serupa di tempat lain [3]. Oleh-oleh khas Bali ini dapat dijadikan sebagai buah tangan yang unik dan berharga bagi wisatawan yang berkunjung ke Bali, serta dapat juga dijadikan sebagai hadiah yang berarti untuk orang-orang yang disayangi.

1.2 Ontology

Ontology adalah representasi konseptual dari berbagai aspek dalam domain tertentu. Ini merupakan cara untuk menggambarkan pengetahuan tentang serangkaian konsep dan hubungan di dalam suatu bidang informasi. Ontology dapat digunakan untuk menyajikan informasi secara semantik, mengatur, dan memetakan sumber daya informasi dengan cara yang terstruktur dan terorganisir. Dengan ontology, kita dapat menggambarkan sebuah teori tentang objek dan hubungan antara mereka [4]. Ontology memiliki definisi yang beragam dalam literatur kecerdasan buatan. Pada awalnya, ontology merujuk pada konsep filosofis, namun dalam konteks kecerdasan buatan dan web, ontology digunakan untuk menggambarkan konsep domain dan relasinya. Ontology menjelaskan teori tentang suatu objek dan keterkaitannya dengan objek lain. Ontology menjadi elemen penting dalam berbagai aplikasi, seperti sistem agen, sistem pengelolaan pengetahuan, dan platform e-commerce. Selain itu, ontology juga dapat digunakan untuk menghasilkan bahasa alami, mengintegrasikan informasi cerdas, menyediakan akses berbasis semantik, dan mengekstrak informasi dari teks. Ontology juga dapat diterapkan dalam berbagai aplikasi lain yang memiliki pengetahuan yang tertanam secara eksplisit. Neches dan rekannya memberikan definisi tentang Ontology yaitu bahwa: "Ontology merupakan definisi dari pengertian dasar dan relasi vokabulari dari sebuah area sebagaimana aturan dari kombinasi istilah dan relasi untuk mendefinisikan vokabulari"[5]. Ada beberapa tujuan dalam pengembangan ontology, antara lain:

- a. Berbagi pemahaman umum tentang struktur informasi antara pengguna atau agen perangkat lunak.
- b. Memfasilitasi penggunaan kembali pengetahuan domain.
- c. Menyatakan asumsi-asumsi domain secara eksplisit.
- d. Memisahkan pengetahuan domain dari pengetahuan operasional.
- e. Menganalisis domain pengetahuan.

Secara teknis, ontology direpresentasikan dalam bentuk kelas, properti, slot, dan instans, sesuai dengan deskripsi eksplisit dari konsep dalam domain pengetahuan yang telah dijelaskan sebelumnya.

1.3 Komponen Ontology

Ontology terdiri dari beberapa elemen, di antaranya:

- a. Individual (Instances)
Individual (Instances) merupakan elemen dasar dalam ontology yang dapat berupa objek nyata maupun abstrak.
- b. Classes
Suatu kelas menggambarkan konsep-konsep yang ada dalam domain tertentu. Kelas biasanya merupakan kumpulan objek-objek yang memiliki kesamaan. Kelas dapat memiliki subkelas yang mewakili konsep yang lebih spesifik daripada kelas utamanya.
- c. Attributes

Objek-objek dalam ontology dapat dijelaskan lebih lanjut dengan memberikan atribut pada objek tersebut. Setiap atribut memiliki nama dan nilai, dan digunakan untuk menyimpan informasi tambahan tentang objek yang bersangkutan.

d. Relasi

Relasi menjadi elemen penting dalam ontology karena menggambarkan hubungan antara objek-objek yang ada. Kemampuan ontology dalam mendeskripsikan relasi antar objek adalah salah satu keunggulannya. Kumpulan relasi ini mendefinisikan semantik atau makna dari entitas tertentu [6].

1.4 Methontology

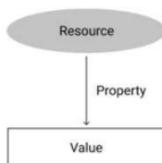
Methontology adalah suatu metodologi yang dirancang oleh sekelompok peneliti di Universitas Politécnica de Madrid. Metodologi ini memfasilitasi pengembangan ontology pada tingkat pengetahuan dengan menyediakan proses pengembangan ontology, siklus hidup berbasis pengembangan prototipe, serta teknik-teknik untuk menjalankan setiap aktivitas yang terkait. Methontology juga memiliki kemampuan untuk melakukan siklus hidup ontology berdasarkan pada pengembangan prototipe, yang memungkinkan untuk menambah, mengubah, dan menghapus elemen-elemen pada setiap versi terbaru (prototipe).

1.5 Web Semantik

Secara mendasar, semantic web adalah proses yang bertujuan mengubah data menjadi pengetahuan dengan menggunakan metode yang dapat dimengerti oleh mesin. Hal ini mencakup kemampuan mesin untuk menggambarkan hubungan antar data serta membuat representasi grafik dari data tersebut. Tujuan dari pengembangan semantic web adalah untuk memungkinkan mesin tidak hanya menampilkan informasi, tetapi juga untuk melakukan proses otomatisasi dan integrasi, sehingga informasi tersebut dapat digunakan kembali dalam berbagai aplikasi. Tim Berners-Lee dan timnya, sebagai pendiri W3C (World Wide Web Consortium), menegaskan bahwa inti dari konsep semantic web adalah penerapan ontology untuk mengubah informasi menjadi bentuk-bentuk basis pengetahuan yang dapat dipahami oleh mesin[6]. Beberapa komponen yang telah dibangun diantaranya adalah RDF (Resource Description Framework) serta OWL (Ontology Web Language). Komponen utama web semantic didasarkan pada komponen lainnya yang telah dibangun oleh W3C (World Wide Web Consortium) yaitu XML (Extensible Markup Language), URI (Uniform Resource Identifier), Maupun HTTP (Hypertext Transfer Protocol).

1.6 RDF

RDF adalah format data yang digunakan untuk membentuk konsep triple dalam bentuk subjek predikat dan objek pada sebuah informasi. RDF juga dikenal dengan standar untuk mendeskripsikan suatu data. Konsep RDF mirip dengan ERD tetapi dengan kegunaan yang berbeda. RDF digunakan dalam pembuatan web semantik sementara ERD dilakukan dalam membuat basis data relasional [7].



Gambar 1. RDF

1.7 SPARQL Query

SPARQL adalah perintah atau bahasa yang digunakan untuk mengakses query pada sebuah model data semantik pada format data RDF. SPARQL bisa disebut juga sebagai bahasa untuk mengakses linked data dengan penggunaan end point untuk dapat menghasilkan relasi antara satu informasi dengan yang lainnya. Bahasa SPARQL dianggap setara dengan SQL dengan memiliki format sintak yang serupa, hanya saja dalam penggunaannya berbeda[7]. Struktur query SPARQL terdiri dari beberapa klausa yang mencakup:

- Prefix: Mekanisme untuk menentukan awalan singkat untuk URI, memungkinkan singkatan dalam kueri.
- Select: Menentukan variabel-variabel yang akan dikembalikan sebagai hasil kueri.
- From: Menentukan jalur atau URL dari dokumen RDF yang akan di query.
- Where: Menyatakan sederetan pola triple yang harus dimiliki oleh setiap hasil kueri yang valid, sesuai dengan format RDF.
- Filter: Digunakan untuk mencocokkan nilai tertentu dari variabel yang di query.
- Optional: Menyatakan bahwa jika suatu komponen tidak memiliki data, tetap dikembalikan sebagai hasil kueri.
- Union: Menyimpan data dari dua pola RDF yang berbeda ke dalam variabel yang sama.
- Distinct: Memastikan bahwa setiap kombinasi hasil query adalah unik.
- Order by: Mengurutkan hasil query berdasarkan nilai variabel tertentu.

1.8 Protégé

PROTEGE merupakan tool ontology dengan platform Open source untuk membangun domain model dan aplikasi knowledge based. Protege mengimplementasikan struktur pemodelan knowledge dan dapat memvisualisasikan hasil ontologi dalam berbagai format. Sebuah ontologi menggambarkan konsep dan hubungan yang penting dalam domain khusus, yang menyediakan kosakata dalam domain tersebut. Dalam beberapa tahun terakhir, ontologi telah diadopsi dalam bisnis atau komunitas ilmiah seperti scientific knowledge portal, manajemen informasi, integrasi informasi, e-commerce dan semantic web service [8].

2. Metode Penelitian

Dalam penelitian ini, digunakan pendekatan Methontology untuk merancang model ontology. Methontology merupakan metode yang menghasilkan model ontology dengan memperhatikan keterkaitan antara objek-objek. Berikut adalah Langkah-langkah yang diterapkan dalam penerapan metodologi tersebut:



Gambar 2. Alur Metode Penelitian

2.1. Analisis Kebutuhan

Tahap ini bertujuan untuk mengevaluasi kebutuhan sistem yang mencakup penggunaan data, penelitian dari sumber yang ada, dan alat yang digunakan. Tahap ini terbagi menjadi dua bagian, yaitu analisis kebutuhan fungsional yang mengidentifikasi fungsi sistem dan kebutuhan non-fungsional yang memperhatikan elemen-elemen pendukung dalam pelaksanaan system seperti perangkat keras laptop atau komputer dan perangkat lunak seperti aplikasi Protégé, dan vscode.

2.2. Pengumpulan Data

Data yang dimanfaatkan dalam studi ini adalah data yang terkait dengan oleh-oleh khas Bali yang diperoleh dan dikumpulkan dari berbagai sumber internet dan jurnal. Setelah itu, data mengenai oleh-oleh khas Bali yang terkumpul diolah menggunakan aplikasi Protégé untuk membangun model ontology.

2.3. Spesifikasi

Tujuan dari tahap spesifikasi adalah untuk menciptakan dokumen spesifikasi ontology yang bersifat informal, semi-formal, atau formal, yang disusun dalam bahasa alami, dan mungkin menggunakan berbagai bentuk representasi perantara atau pertanyaan kompetensi.

2.4. Akuisisi Pengetahuan

Tahap perolehan pengetahuan adalah fase yang terpisah dalam proses pengembangan ontology. Sumber-sumber pengetahuan diperoleh dari para ahli, literatur, gambar, serta ontology lainnya dengan menerapkan teknik seperti sesi diskusi bebas gagasan, analisis teks secara formal dan informal, dan perangkat untuk mengumpulkan pengetahuan.

2.5. Pembangunan Model Ontology

Metode yang dipakai untuk mengembangkan model ontology adalah Methontology. Methontology merupakan salah satu Teknik untuk Pembangunan model ontology yang memiliki keunggulan dalam menjelaskan setiap Langkah kegiatan. Dalam konteks ini, Methontology memiliki kapabilitas untuk memanfaatkan Kembali ontology yang telah dibangun sebelumnya untuk pengembangan sistem selanjutnya. Tahapan dalam metode Methontology meliputi kegiatan perencanaan, akuisisi pengetahuan, dokumentasi, dan evaluasi. Kegiatan perencanaan dilakukan pada tahap awal pengembangan proyek. Fase meliputi spesifikasi, konseptualisasi, formalisasi, integrasi, implementasi, dan pemeliharaan.

2.6. Implementasi

Pada tahap ini, konsep ontology yang telah disusun akan diterapkan. Implementasi ontology memerlukan penggunaan lingkungan yang mendukung meta-ontology serta ontology yang dipilih.

2.7. Pengujian dan Evaluasi Sistem

Proses pengujian dan evaluasi model ontology dilakukan dengan menyusun serangkaian pertanyaan yang terkait dengan fitur-fitur model ontology, serta persepsi kegunaan dan kemudahan penggunaan yang dirasakan oleh pengguna. Data jawaban yang diperoleh dari proses pengujian model ontology kemudian diolah dan dianalisis untuk memperoleh pemahaman yang lebih baik. Untuk pertanyaan yang terkait dengan pengujian fitur-fitur model ontology, data diolah dengan memberikan penandaan pada setiap jawaban yang diberikan.

2.8. Dokumentasi

Pada tahap ini, belum ada standar yang disepakati mengenai cara dokumentasi ontology. Tahap dokumentasi umumnya melibatkan kode untuk ontology, penggunaan bahasa alami, dan penulisan makalah yang nantinya akan dipublikasikan dalam prosiding dan jurnal. Dokumentasi ini berguna untuk menetapkan pertanyaan kunci terkait ontology yang sedang dikembangkan.

3. Hasil dan Diskusi

3.1 Spesifikasi

Pada tahap ini adalah spesifikasi terkait dengan ontology yang telah dibangun, yang merupakan deskripsi dari ontology "Oleh-Oleh Khas Bali":

- a. Domain : Ontology Oleh-Oleh Khas Bali.
- b. Tanggal : 9 Mei 2024.
- c. Dirancang oleh : Putu Ardi Sudarmika.
- d. Diimplementasikan oleh : Putu Ardi Sudarmika.
- e. Level Formalitas : Semi Formal.
- f. Ruang Lingkup : Fokus pada Oleh-Oleh Khas Bali.
- g. Sumber Pengetahuan : Berdasarkan informasi dari Internet dan studi literatur.

Pada langkah ini, kami memberikan spesifikasi terkait ontology yang telah dikembangkan, yang merinci deskripsi dari ontology "Oleh-Oleh Khas Bali":

3.2 Akuisisi Pengetahuan

Akuisisi pengetahuan merupakan proses yang berbeda dalam pengembangan ontology. Ini melibatkan menganalisis teks informal dan memahami konsep-konsep utama dalam manual. Dalam penelitian ini, tahap akuisisi pengetahuan mencakup mengidentifikasi pengetahuan dan struktur yang digunakan melalui studi literatur. Data yang diperoleh dari penelitian ini merupakan hasil kumpulan informasi dari berbagai sumber yang memiliki pengetahuan tentang oleh oleh khas bali.

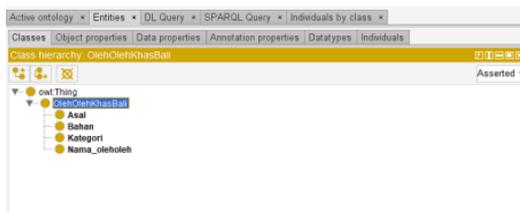
3.3 Konseptualisasi

Pada tahapan ini, fokusnya adalah merancang kerangka konseptual yang akan menjelaskan masalah yang dihadapi serta solusi yang akan diterapkan. Selama proses ini, kami menyusun daftar lengkap istilah yang mencakup konsep, contoh, tindakan, dan properti yang terkait dengan domain oleh-oleh khas Bali.

3.4 Integrasi

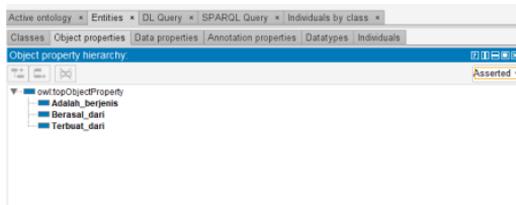
Pada tahap ini, tujuannya adalah untuk melakukan evaluasi ulang terhadap definisi yang digunakan dalam ontology yang sedang dibangun, dengan tujuan untuk memastikan bahwa tidak ada kesalahan yang terjadi dalam pembentukan hubungan antar entitas. Proses ini melibatkan peninjauan kembali definisi dan konsep yang telah diatur sebelumnya, dengan tujuan untuk memastikan konsistensi dan akurasi data representasi relasi antar entitas

3.5 Implementasi



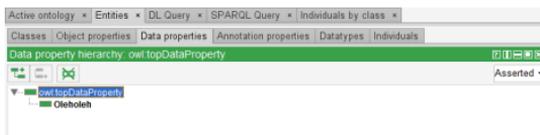
Gambar 3. Class dari Ontologi Oleh Oleh Khas Bali

Ontologi Pemodelan oleh oleh khas Bali terdiri dari empat kelas. Setiap kelas dalam ontology memiliki kaitan dengan individu yang dikenal sebagai ekstensi kelas tersebut.



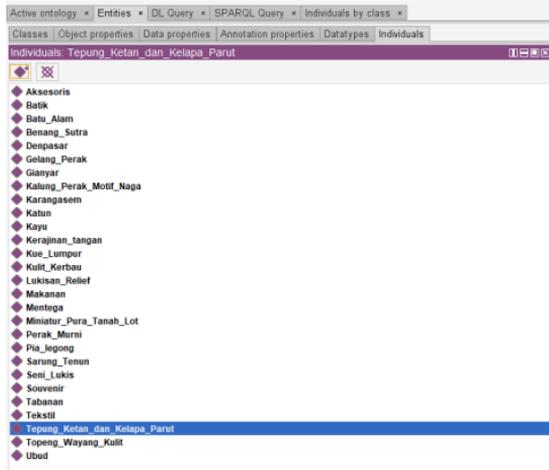
Gambar 4. Object Properties dari Ontologi Oleh Oleh Khas Bali

Ontologi domain oleh oleh khas Bali memiliki tiga Object Property. Object Property merupakan properti yang menghubungkan satu individu dengan individu lainnya.



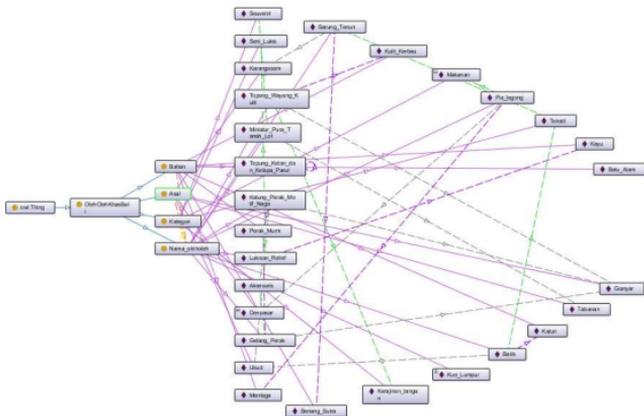
Gambar 5. Data Properties dari dari Ontologi Oleh Oleh Khas Bali

Data properti yang terdapat dalam ontologi domain oleh oleh khas Bali. Data Properties digunakan untuk mengaitkan instansi dengan nilai tipe data seperti teks.



Gambar 6. . Individu yang digunakan dalam Ontologi Oleh Oleh Khas Bali.

Ini adalah entitas yang dimanfaatkan dalam ontologi oleh oleh khas Bali. Entitas yang terdapat dalam kelas disebut instansi.



Gambar 7. Ontograf dari dari Ontologi Oleh Oleh Khas Bali

Ini menunjukkan contoh keterkaitan semantik yang mengilustrasikan setiap kelas, properti objek, dan entitas yang dibuat dalam ontologi oleh oleh khas Bali. Keterkaitan tersebut diwakili dalam bentuk gambar pada ontograf.

3.6 Evaluasi

Konseptualisasi dalam konteks oleh-oleh khas Bali bertujuan untuk membangun pengetahuan yang meliputi berbagai konsep, contoh produk, aktivitas, dan atribut terkait. Dengan demikian, pembuatan ontology yang terstruktur dapat mengidentifikasi dan mengumpulkan semua aspek penting dari domain oleh-oleh Bali yang bermanfaat dan dapat diterapkan dalam bentuk kelas dan subkelas. Pada tahap ini, penulis melakukan evaluasi terhadap ontology yang telah dibuat. Evaluasi dilakukan melalui penggunaan query SPARQL pada aplikasi Protégé 5.5 untuk menjawab pertanyaan tertentu. Penulis membuat PREFIX baru yang disebut "rdd" yang merujuk ke alamat ontology IRI yang sesuai dengan alamat di aplikasi Protégé 5.5. Pertanyaan yang diuji adalah "Apa saja jenis oleh-oleh yang termasuk dalam kategori makanan?". Berikut adalah hasil dari query SPARQL yang dilakukan.

```

Active ontology • Entities • DL Query • SPARQL Query • Refresh by click
SPARQL query
PREFIX rdd <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl <http://www.w3.org/2002/07/owl#>
PREFIX rdfs <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd <http://www.w3.org/2001/XMLSchema#>
PREFIX rdd <http://www.semanticscience.org/ontology/olehOlehKhasBali#>
SELECT *
WHERE { ?nama_nsthehah rdd:Beras rdd:isiGanyu }

Nama_nsthehah
Kategori_Parasi_MHR_7899
Tipe_nsthehah_PuH4
Ganyu_Parasi
    
```

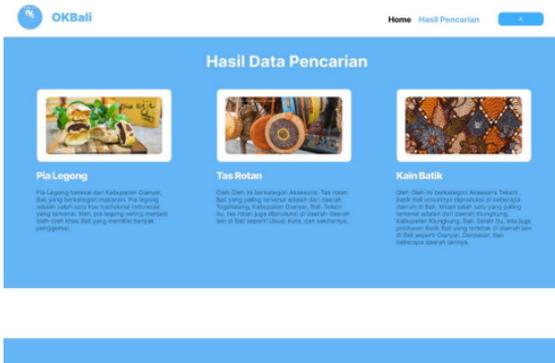
Gambar 8. Hasil SPARQL query

Ini adalah tampilan hasil query yang muncul dalam pertanyaan yang diajukan, menggunakan sintaks PREFIX rdd. Kueri ini mencari oleh-oleh yang berasal dari Denpasar. Hasil SPARQL yang dihasilkan termasuk individu seperti Kalung Perak Motif Naga, Topeng Wayang Kulit, dan Gelang Perak.

3.7 Dokumentasi



Gambar 9. Halaman pencarian



Gambar 10. Hasil data pencarian

4. Kesimpulan

Berdasarkan hasil dan pembahasan yang telah dilakukan, pemodelan ontologi untuk oleh-oleh khas Bali telah berhasil dibangun. Proses pemodelan ontologi ini menggunakan aplikasi Protege 5.5.0 dengan metode Methontology, menghasilkan 4 kelas, 3 Properti Objek, properti data, 34 ontografi, dan 28 individual atau instansi untuk setiap kelas. Selain itu, penelitian ini juga melibatkan tahap evaluasi atau pengujian terhadap model yang diajukan dengan menyajikan beberapa pertanyaan kepada pengguna untuk mengakses informasi tentang oleh-oleh khas Bali.

Dengan demikian, diharapkan model ontologi yang dihasilkan mampu secara sistematis memberikan informasi mengenai oleh-oleh khas Bali. Selain itu, pengembangan ontologi untuk oleh-oleh khas Bali dapat menjadi dasar bagi pengembangan lebih lanjut terkait sistem yang berkaitan dengan oleh-oleh tersebut.

Daftar Pustaka

- [1] C. Pramatha, "Pengembangan Ontologi Tujuan Wisata Bali Dengan Pendekatan Kulkul Knowledge Framework," SINTECH (Science and Information Technology) Journal, vol. 3, no. 2, pp. 77–89, 2020.
- [2] D. P. Githa and D. P. S. Putri, "Rancang Bangun Sistem Rekomendasi Destinasi Wisata di Bali," SINTECH (Science and Information Technology) Journal, vol. 2, no. 2, pp. 81–90, 2019.
- [3] N. W. Rustiarini and A. A. D. Widyani, "Pembinaan Aspek Manajemen pada Kelompok Usaha Oleh-Oleh Khas Bali," Jurnal Bakti Saraswati (JBS), vol. 4, no. 2, 2015.
- [4] G. E. Rupilu, "Desain Ontologi Sederhana Perpustakaan Online," Jurnal Sosial Teknologi, vol. 4, no. 2, pp. 120–123, 2024.
- [5] K. W. Triyoga, D. E. Cahyani, and S. W. Sihwi, "Pembangunan Ontology Berbasis Metode Methontology Untuk Domain Tuberculosis," in Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi), 2019, pp. 47–54.
- [6] Y. S. Baskoro, H. Hafsa, and H. Jayadianti, "Representasi Pengetahuan Dalam Semantik Ontologi Pada Domain Kriminalitas Kepolisian Sektor Depok Timur Daerah Istimewa Yogyakarta," Telematika: Jurnal Informatika dan Teknologi Informasi, vol. 15, no. 1, pp. 13–29, 2018.
- [7] P. D. Bangsa and I. Hermawan, "Rancang Bangun Sistem Repositori Berbasis Web Semantik Menggunakan Ontologi," Jurnal Teknologi Terpadu, vol. 6, no. 2, pp. 58–65, 2020.
- [8] A. Irawan, "Pengembangan Portal Berita Menggunakan Teknologi Web Semantic Pada Banten Tv," ProTekInfo (Pengembangan Riset dan Observasi Teknik Informatika), vol. 1, pp. 76–79, 2014.

p-ISSN

ISSN 2986-3929



9 772986 392004

e-ISSN

ISSN 3032-1948



9 773032 194009