

Implementasi Case Base Reasoning Sebagai Diagnosis Penyakit Kesehatan Mental Pada Data Aplikasi Alodokter

I Wayan Supriana^{a1}, I Wayan Adhi Surya Gemilang^{a2}

^aProgram Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana
Jl. Raya Kampus Unud Bukit Jimbaran, Badung, Indonesia

¹cool.gemilang@gmail.com

²wayan.supriana@unud.ac.id

Abstract

Kesehatan psikis atau kesehatan mental merujuk pada keadaan di mana seseorang memiliki kesejahteraan yang tercermin dari kemampuannya untuk mengenali dan memahami potensi dirinya, menghadapi tekanan kehidupan normal, dan bekerja dengan produktif memanfaatkan semua potensi yang dimiliki. Identifikasi dini kesehatan mental sangat penting dilakukan, salah satu model yang dapat digunakan adalah Case-Based Reasoning (CBR). CBR pada penelitian ini memodelkan proses diagnosis penyakit mental dengan memanfaatkan data yang tersedia dari platform kesehatan digital Alodokter. Alodokter menyediakan informasi medis yang terstruktur, termasuk gejala, riwayat medis, dan diagnosis dari berbagai kasus. Dengan menggunakan pendekatan CBR, sistem secara efisien mencari kasus serupa untuk setiap pasien baru berdasarkan gejala yang dilaporkan. Dengan menggunakan TF-IDF untuk representasi kasus dan K-Means untuk Indexing kasus, hasil dari penelitian ini berupa model Case-Based Reasoning yang dapat mendiagnosis penyakit psikis berdasarkan input user, dengan nilai similaritas sebesar 83%.

Keywords: Case-Based Reasoning, Penyakit Psikis, K-Means, Diagnosis, TF-IDF

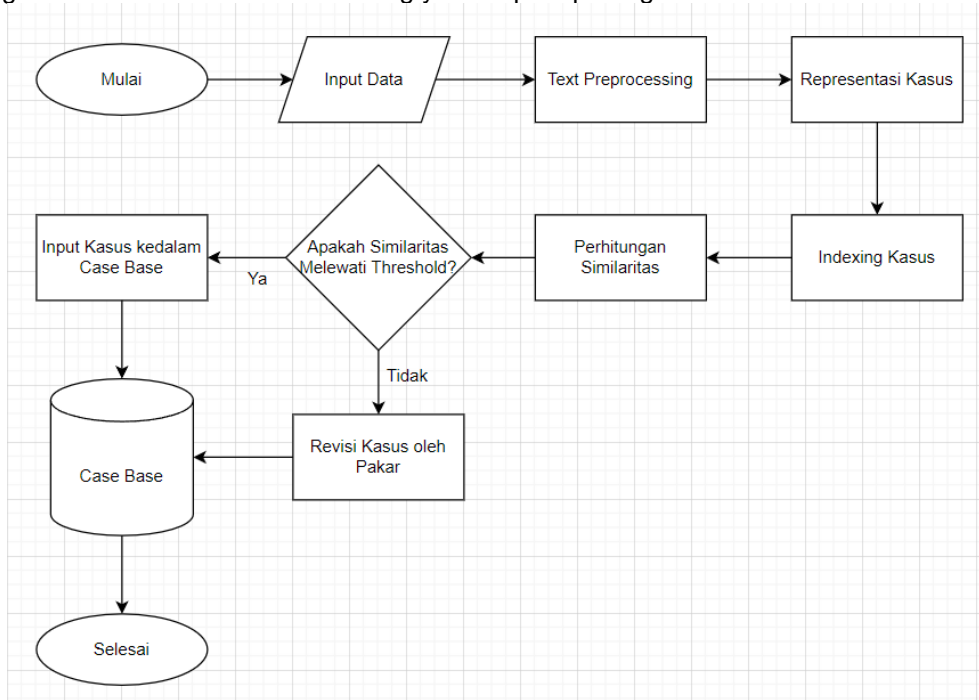
1. Pendahuluan

Kesehatan psikis atau kesehatan mental merujuk pada keadaan di mana seseorang memiliki kesejahteraan yang tercermin dari kemampuannya untuk mengenali dan memahami potensi dirinya, menghadapi tekanan kehidupan normal, dan bekerja dengan produktif memanfaatkan semua potensi yang dimiliki. Seseorang dengan kesehatan mental yang baik mampu berpikir dengan jelas dan positif ketika menghadapi berbagai tantangan dan kesulitan hidup, serta mampu memberikan kontribusi yang berarti bagi komunitas dan lingkungannya. Keadaan mental yang baik ditandai dengan ketenangan dan ketenangan, dengan demikian, seseorang dapat menghargai orang lain dan lingkungan sekitarnya serta menikmati kehidupan sehari-hari. Namun, sayangnya, kesehatan mental sering kali diabaikan, karena banyak orang hanya memprioritaskan kesehatan fisik. [2].

Sistem Case-Based Reasoning (CBR) yang dikembangkan dalam konteks ini bertujuan untuk meningkatkan proses diagnosis penyakit mental dengan memanfaatkan data yang tersedia dari platform kesehatan digital Alodokter[2]. Alodokter menyediakan informasi medis yang terstruktur, termasuk gejala, riwayat medis, dan diagnosis dari berbagai kasus. Dengan menggunakan pendekatan CBR, sistem secara efisien mencari kasus serupa untuk setiap pasien baru berdasarkan gejala yang dilaporkan [7]. Informasi dari kasus-kasus serupa ini digunakan untuk merumuskan diagnosis awal yang personal dan mendetail. Selanjutnya, sistem dapat merevisi diagnosis berdasarkan perkembangan informasi tambahan atau respons terhadap pengobatan, sambil terus menyimpan dan memperbarui basis pengetahuan dengan kasus-kasus baru. Pendekatan ini tidak hanya meningkatkan efisiensi diagnosis, tetapi juga memastikan kontinuitas perawatan yang optimal dengan memanfaatkan pengalaman masa lalu secara efektif untuk mendukung pengambilan keputusan klinis yang lebih baik [3].

2. Metode Penelitian

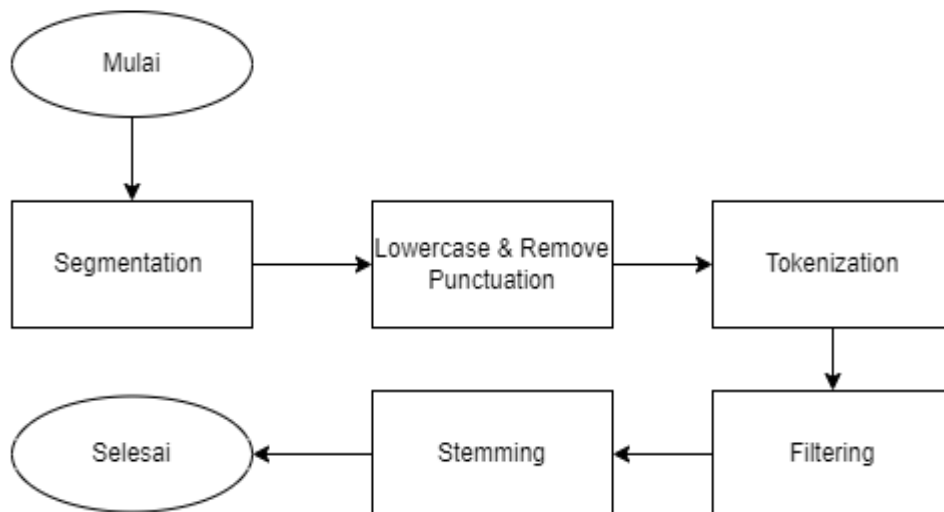
Metode penelitian dimulai dengan mengumpulkan informasi atau data serta melakukan observasi terhadap data yang diperoleh. Metode ini menjelaskan rancangan penelitian, mencakup berbagai aspek seperti prosedur dan langkah-langkah yang harus dilaksanakan, durasi penelitian, sumber data, serta cara pengumpulan, pengolahan, dan analisis data. [1]. Metode penelitian yang dilakukan penulis membangun model Case Based Reasoning yaitu seperti pada gambar 1.



Gambar 1. Metode Penelitian

2.1. Text Preprocessing

Tahap pertama yang dilakukan adalah melakukan text preprocessing terhadap data yang sudah didapatkan. Pada tahap ini, semua data yang didapatkan akan melewati proses segmentation, lowercase dan remove punctuation, tokenizing, filtering dan stemming.



Gambar 2. Text Preprocessing

2.2. Representasi Kasus dan Indexing

Pada tahap ini, dilakukan dua model ekstraksi fitur terhadap data yang digunakan untuk dapat merepresentasikan kasus. Data melewati proses ekstraksi fitur penyakit berdasarkan referensi gejala dan nama penyakit yang didapatkan melalui website Alodokter, kemudian diberikan bobot dan dimasukkan ke dalam kolom disease. Setelah penentuan bobot dari kasus, dilakukan tahap Indexing yang bertujuan untuk mencari kasus serupa dalam basis data dengan cepat dan mudah. Indexing yang digunakan dalam penelitian ini adalah K-Means clustering. K-means adalah algoritma yang tergolong dalam unsupervised learning. Algoritma ini berfungsi untuk mengelompokkan data ke dalam beberapa kluster dan dapat memproses data yang tidak memiliki label kategori.

Algoritma K-Means Clustering juga termasuk metode non-hirarki. Algoritma ini mengelompokkan data ke dalam beberapa kluster, di mana data dalam satu kluster memiliki karakteristik yang serupa, namun berbeda dengan data yang berada di kluster lain. Untuk menentukan jumlah cluster yang akan digunakan, method Elbow digunakan untuk menghitung inertia berdasarkan jumlah cluster yang digunakan untuk mencari jumlah optimal cluster yang akan digunakan.

2.3. Retrieve Kasus

Untuk tahap retrieve kasus digunakan perhitungan cosine similarity, Pertama, kasus baru yang masuk dikonversi menjadi representasi vektor TF-IDF, memungkinkan ekstraksi fitur-fitur penting dari teks, seperti gejala atau deskripsi kondisi. Setelah itu, dengan menggunakan model K-Means yang telah dilatih sebelumnya, kasus baru diberi label kluster yang sesuai berdasarkan representasi TF-IDF-nya. Selanjutnya, algoritma mencari kasus-kasus serupa dalam kluster yang sama dengan menghitung similaritas kosinus antara kasus baru dan setiap kasus dalam kluster menggunakan fungsi cosine_similarity dari scikit-learn [5]. Kasus-kasus dengan similaritas tertinggi dipilih berdasarkan urutan nilai similaritas kosinusnya. Pendekatan ini tidak hanya meningkatkan efisiensi dalam menemukan kasus-kasus serupa, tetapi juga memastikan bahwa pengambilan keputusan diagnosa didasarkan pada pengalaman dari kasus-kasus sebelumnya yang relevan. Dibawah ini merupakan contoh hasil dari tahap retrieve kasus yang telah dilakukan.

2.4. Reuse Kasus

Pada tahap ini kasus baru yang sudah ditentukan kedekatannya pada tahap retrieve, yaitu dengan menggunakan kembali solusi pada basis kasus. Teknik reuse sebuah kasus dalam konteks kasus baru berfokus pada dua hal utama: perbedaan antara kasus yang sudah ada dengan kasus baru, serta bagian mana dari kasus yang diambil (retrieve case) yang dapat dimanfaatkan pada kasus baru. Terdapat dua pendekatan untuk re-use kasus yang sudah ada: menggunakan kembali solusi dari kasus yang ada (transformational reuse) atau memanfaatkan metode dari kasus tersebut untuk mengembangkan solusi baru [4].

2.5. Revise Kasus

Pada tahap revise, diberikan sebuah batasan threshold untuk similaritas kasus baru dengan kasus lama, jika nilai similaritas melebihi threshold tersebut, maka kasus baru akan langsung dimasukkan ke dalam case base, jika nilai similaritas kurang dari threshold tersebut, maka kasus tersebut akan dimasukkan ke dalam sebuah file csv, yang dimana digunakan untuk menyimpan kasus-kasus yang gagal melewati threshold dan akan direvisi oleh pakar untuk dimasukkan ke dalam case base utama.

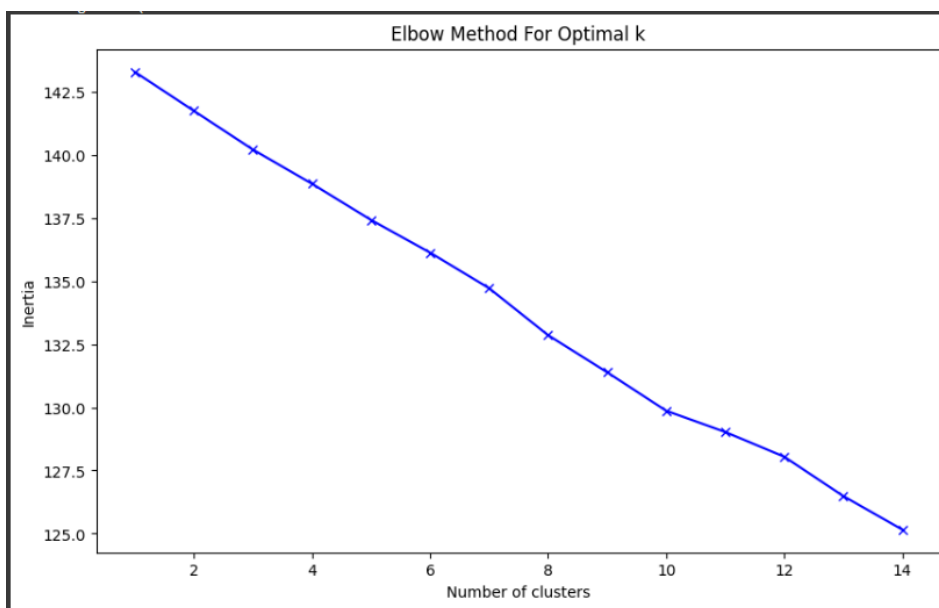
2.6. Retain Kasus

Tahap ini memakai solusi yang terakhir sebagai bagian dari kasus baru. Pada tahap ini terjadi suatu proses penggabungan dari solusi kasus yang baru yang benar ke knowledge/case base yang telah ada.

3. Hasil dan Pembahasan

3.1. Clusterisasi Data

Untuk menentukan jumlah cluster yang akan digunakan, method Elbow digunakan untuk menghitung inertia berdasarkan jumlah cluster yang digunakan untuk mencari jumlah optimal cluster yang akan digunakan. Berikut merupakan grafik dari method Elbow yang dilakukan.



Gambar 3. Grafik Elbow Method

Dari gambar diatas tampak bahwa inertia terus berkurang tanpa “elbow point” yang jelas. Namun, terdapat perbedaan yang mencolok antara 6 dan 8 cluster. Sehingga jumlah cluster yang akan digunakan adalah 6, dengan memperhitungkan jumlah total data yang dimiliki yaitu 150.

3.2. Perhitungan Similaritas

Perhitungan Cosine Similarity kesamaan kosinus adalah ukuran yang digunakan untuk mengukur seberapa mirip dua vektor non-nol dalam ruang produk dalam. Ini sering digunakan dalam analisis teks dan pemrosesan bahasa alami untuk mengukur seberapa mirip dua dokumen atau fragmen teks. Kesamaan kosinus mengukur kosinus dari sudut antara dua vektor dalam ruang multi-dimensi. Nilai kosinus 0° adalah 1, yang berarti vektor tersebut identik, sedangkan nilai kosinus 90° adalah 0, yang berarti vektor tersebut ortogonal atau sepenuhnya berbeda [4]. Rumus untuk kesamaan kosinus adalah:

$$\frac{A \cdot B}{\|A\| \cdot \|B\|} \quad (1)$$

Di mana:

- $A \cdot B$ adalah hasil dot product dari vektor A dan B
- $\|A\|$ dan $\|B\|$ adalah magnitudo (atau panjang) dari vektor A dan B

Langkah-langkah untuk menghitung cosine similarity adalah sebagai berikut: Pertama, representasikan data teks sebagai vektor. Ini dapat dilakukan dengan metode seperti TF-IDF (Term Frequency-Inverse Document Frequency), di mana setiap dokumen direpresentasikan sebagai vektor dalam ruang multi-dimensi. Kedua, hitung hasil kali dot dari dua vektor tersebut. Ketiga, hitung magnitudo dari masing-masing vektor. Terakhir, bagi hasil kali dot dengan hasil kali magnitudo dari kedua vektor tersebut untuk mendapatkan skor kesamaan.

3.2. Evaluasi Model

Setelah model dijalankan dengan contoh kasus seperti dibawah ini:

Tabel 1. Contoh Kasus Baru

Judul Kasus	Pertanyaan Kasus
"penyakit skizofrenia paranoid tidak sembuh"	"siang dok, dok kakek saya yang berusia 63 tahun merasa terancam jiwanya takut oleh sekumpulan orang dianggap ancaman yang ternyata hanya halusinasi"

Dari kasus diatas akan didapatkan hasil seperti gambar dibawah ini:

```

Title \
1      penyakit skizofrenia paranoid sembuh
0      penyakit skizofrenia pengobatannya
38     takut konsultasi psikolog psikiat lingkungan m...
40     takut kehilangan kesadaran jantung berdetak ke...
62     melamun tremor dada sesak kondisi fisik melema...

Question \
1      siang dok dok keluhan kakek usia 63 kakek tera...
0      selamat sore dok curhat kakek kakek tidur mala...
38     halo dok laki laki umur 19 mengen mengendalika...
40     permisi dok usia 18 semenjak nenek meningg tak...
62     permisi doksudah â 2 mengalami aneh contohnya ...

Answer \
1      alo selamat siang skizofrenia paranoid salah j...
0      alo selamat siang skizofrenia gangguan mental ...
38     alo terimakasih pertanyaannya gangguan kejiwaa...
40     alo terima kasih pertanyaannya alodokt keluhan...
62     alo terima kasih alodokt dr ainul keluhan alam...

Diseases cluster
1      [Skizofrenia, PTSD] 5
0      [Skizofrenia] 5
38     [Depresi, Skizofrenia, Bipolar, Gangguan Kepri... 5
40     [Depresi, Bipolar, Kontrol Impuls, Gangguan Ti... 5
62     [] 5
[0.83275226 0.36161967 0.08706966 0.06972725 0.0694797 ]

```

Gambar 4. Hasil Perhitungan Similaritas Kasus Baru

Dari hasil diatas bisa dilihat hasil dari perhitungan similaritas tertinggi adalah 83%. Kasus baru tersebut dibandingkan dengan kasus yang diambil dari cluster 5 dengan nilai similaritas setelah yang tertinggi yaitu 36%, 8.7%, 6.9%, 6.9%.

3.3. Revisi Kasus

Pada tahap revisi kasus akan digunakan threshold sebesar 70%, yang dimana jika nilai perhitungan similaritas melewati threshold maka kasus akan langsung dimasukkan kedalam case base, jika perhitungan similaritas tidak melewati threshold tersebut maka kasus tersebut akan dipindahkan ke dataset baru yang kosong untuk di revisi oleh pakar, dibawah ini merupakan bentuk dari dataset tersebut.

Tabel 2. Contoh Kasus Baru

No	Title	Question	Answer	Diseases	Cluster
1	Contoh Judul Kasus 1	Contoh Pertanyaan Kasus 1	Need Revision		

4. Kesimpulan

Dari penelitian diatas bisa disimpulkan bahwa kesehatan psikis atau kesehatan mental adalah kondisi di mana seseorang memiliki kesejahteraan yang dapat dilihat dari kemampuannya menyadari potensi diri, menghadapi tekanan kehidupan normal, dan bekerja dengan produktif memanfaatkan semua potensi yang dimiliki. Sistem Case-Based Reasoning (CBR) yang dikembangkan dalam konteks ini bertujuan untuk meningkatkan proses diagnosa penyakit mental dengan memanfaatkan data yang tersedia dari platform kesehatan digital Alodokter. Hasil dari perhitungan similaritas kasus yang dilakukan sebesar 83% sehingga bisa dikatakan model Case-Based Reasoning yang dibangun dapat dengan presisi menghitung similaritas kasus yang mirip dengan kasus-kasus yang ada di basis kasus.

References

- [1] V. A. Afeanpah, S. A. S. Mola, and A. Fanggidae, "Case Based Reasoning untuk Mendiagnosa Jenis Gangguan Jiwa Menggunakan Metode Dempster Shafer," *Jurnal Komputer dan Informatika*, vol. 10, no. 1, pp. 9–17, Mar. 2022, doi: 10.35508/jicon.v10i1.6326.
- [2] Alodokter. (2024). Alodokter. Diambil 23 Juni 2024, dari <https://www.alodokter.com/>
- [3] Dwinta Ananda, Sumarna Indra, Ilman Kadori, "Metode Case Based Reasoning Untuk Mendiagnosa Gangguan Psikologi Pada RSD Gunung Jati Cirebon" *Journal of Information and Technology*, Vol. 17, No. 1, Februari 2024, pp. 1-13
- [4] Minarni, I. Warman, and W. Handayani, "Case- Based Reasoning (CBR) Pada Sistem Pakar Identifikasi Hama dan Penyakit Tanaman Singkong Dalam Usaha Meningkatkan Produktivitas Tanaman Pangan," *J. TEKNOIF*, vol. 5, no. 1, pp. 41–47, 2017.
- [5] Supriana, I. W, Prebiana K. D. 2020. "Metode Penalaran Berbasis Kasus (Case Base Reasoning) Dalam Menentukan Kelayakan Sokola Perawat", *Jurnal Rekayasa Sistem Komputer*, Vol.3, No.1, 2020.
- [6] E. Seniwati, "Case Base Reasoning Menggunakan Algoritma Bayesian untuk Penentuan Pemberian Beras Miskin," *J. Mantik Penusa*, vol. 2, no. 2, pp. 152–156, 2018.
- [7] .C. S. Fatoni and F. D. Noviantha, "Case Base Reasoning Diagnosis Penyakit Diferi dengan Algoritma K-Nearest Neighbor," *Creat. Inf. Technol. J.*, vol. 4, no. 3, pp. 220–232, 2017.