

ANALISIS METODE TOPSIS DALAM EVALUASI PERFORMA CHATBOT BERDASARKAN DEPARTEMEN

Suparmadi¹, Dahriansah², Dzikri Mahardika³

^{1,2,3}Sistem Informasi, Universitas Royal

Email: ¹suparmadi43@gmail.com

Abstract: Chatbot performance evaluations are often conducted uniformly without considering operational characteristics across departments. This study aims to evaluate and compare departmental chatbot performance using the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) method. The methodology utilizes a secondary dataset of approximately 3,000 interaction logs evaluated against four main criteria: response time, bot confidence score, user feedback rating, and estimated business impact. The evaluation results reveal that Human Resources (HR) achieved the highest rank with a preference score of 0.9531, followed by Marketing (0.6712), Sales (0.4885), Finance (0.3772), and Operations at the lowest rank (0.2356). HR's top performance was driven by the highest satisfaction rating (3.28) and confidence score (0.7986), alongside a fast response time (1,340 ms). Conversely, Operations was hindered by the slowest response time (1,405 ms). In conclusion, the TOPSIS method effectively maps chatbot performance objectively, highlighting that a balance between response speed and answer quality is essential for optimizing user satisfaction.

Keywords: chatbot; performance evaluation; mcdm; topsis; departmental comparison

Abstrak: Evaluasi performa chatbot sering kali dinilai secara seragam tanpa mempertimbangkan karakteristik operasional antar-departemen. Penelitian ini bertujuan untuk mengevaluasi dan membandingkan performa chatbot berbasis departemen menggunakan metode *Technique for Order of Preference by Similarity to Ideal Solution* (TOPSIS). Metodologi penelitian memanfaatkan dataset sekunder berjumlah ±3.000 log interaksi dengan empat kriteria utama, yaitu *response time*, *bot confidence score*, *user feedback rating*, dan *estimated business impact*. Hasil analisis menunjukkan bahwa Departemen Human Resource (HR) meraih peringkat tertinggi dengan nilai preferensi 0,9531, disusul oleh Marketing (0,6712), Sales (0,4885), Finance (0,3772), dan Operations pada posisi terendah (0,2356). Keunggulan HR didorong oleh kombinasi rating kepuasan (3,28) dan *confidence rate* (0,7986) tertinggi, serta waktu respon yang cepat (1.340 ms). Sebaliknya, performa Operations terhambat oleh waktu respon paling lambat (1.405 ms). Disimpulkan bahwa metode TOPSIS efektif dalam memetakan efektivitas chatbot secara objektif, di mana kecepatan respon dan kualitas jawaban terbukti menjadi kunci utama pencapaian kepuasan pengguna.

Kata kunci: chatbot; evaluasi performa; mcdm; topsis; perbandingan departemen

PENDAHULUAN

Pada perkembangan teknologi yang sudah semakin masif saat ini dimana setiap bidang dan kegiatan sudah melakukan otomatisasi, tidak jarang pada fungsi peran yang seharusnya dilakukan manusia sudah mulai merambah digantikan oleh teknologi Artificial

Intelligence (AI) seperti peran layanan *customer service*.

Peran *customer service* sudah mulai banyak yang dialihfungsikan dengan *output* dari AI seperti chatbot. Chatbot merupakan program komputer yang dirancang untuk melakukan komunikasi berbasis teks. Adanya chatbot digunakan untuk membantu mempermudah layanan

untuk customer perihal memberikan informasi terkait produk atau layanan sebuah perusahaan [1]. Kehadiran *chatbot* terbukti mampu menekan biaya operasional, memberikan respons secara *real-time* selama 24/7, serta menangani volume kueri yang besar secara bersamaan. Namun, efektivitas implementasi *chatbot* sering kali tidak merata di seluruh unit kerja atau departemen dalam suatu organisasi/perusahaan [2].

Setiap departemen seperti Departemen Layanan Pelanggan, *Human Resources* (HR), *IT Support*, Pemasaran, hingga *finance* memiliki karakteristik operasional, kompleksitas masalah dan ekspektasi pengguna yang sangat beragam. Sebagai contoh, *chatbot* pada Departemen HR mungkin lebih berfokus pada ketepatan informasi regulasi dan privasi data, sedangkan pada Departemen Layanan Pelanggan, indikator kecepatan respons (*response time*) dan kepuasan pengguna (*customer satisfaction score*) menjadi parameter utama. Keberagaman fungsi ini menyebabkan evaluasi performa *chatbot* tidak lagi dapat dilakukan secara seragam (*one-size-fits-all*), melainkan memerlukan pendekatan komparatif berbasis departemen untuk mengidentifikasi unit mana yang mengoptimalkan teknologi tersebut dan unit mana yang membutuhkan perbaikan sistem secara signifikan.

Evaluasi performa *chatbot* merupakan permasalahan keputusan yang kompleks karena melibatkan banyak kriteria yang saling bertentangan (*conflicting criteria*), baik kriteria kuantitatif (seperti *response time*, *error rate*, dan *resolution rate*) maupun kriteria kualitatif (seperti *accuracy*, *user satisfaction*, dan *user experience*). Oleh karena itu, kerangka kerja *Multi-Criteria Decision Making* (MCDM) dengan metode TOPSIS sangat relevan untuk menyelesaikan permasalahan ini.

Metode *Technique for Order of Preference by Similarity to Ideal Solution* (TOPSIS) merupakan salah satu pendekatan MCDM yang sangat efektif

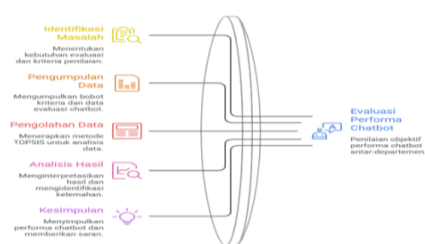
dalam mengurutkan dan mengevaluasi kinerja alternatif berdasarkan jarak terdekat dari solusi ideal positif dan jarak terjauh dari solusi ideal negatif. Keunggulan utama TOPSIS terletak pada konsepnya yang intuitif, komputasi yang efisien, serta kemampuannya dalam mengakomodasi pembobotan kriteria yang fleksibel [3]. Dengan menerapkan TOPSIS, performa *chatbot* di setiap departemen dapat diranking secara objektif dan terukur.

Meskipun kajian mengenai evaluasi *chatbot* telah banyak dilakukan, mayoritas penelitian terdahulu masih berfokus pada evaluasi performa algoritma *Natural Language Processing* (NLP) secara teknis atau penilaian kepuasan pengguna secara umum tanpa mempertimbangkan konteks sektoral/departemental. Masih terbatas penelitian yang memetakan efektivitas *chatbot* secara komparatif antar-departemen dalam satu ekosistem organisasi menggunakan metode komputasi pengambilan keputusan TOPSIS.

Berdasarkan urgensi tersebut, penelitian ini bertujuan untuk mengaplikasikan metode TOPSIS dalam mengevaluasi dan membandingkan performa *chatbot* berbasis departemen. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi teoritis berupa pemodelan MCDM untuk evaluasi teknologi AI, serta memberikan kontribusi praktis berupa rekomendasi strategis bagi manajemen organisasi dalam mengalokasikan sumber daya perbaikan sistem *chatbot* secara tepat sasaran pada departemen yang membutuhkan peningkatan performa.

METODE

Pendekatan Sistematis untuk Evaluasi Chatbot



Gambar 1. Tahapan Penelitian

Tahap 1: Identifikasi Masalah

Pada tahapan ini, dilakukan observasi dengan mengumpulkan informasi dan mengamati secara langsung permasalahan yang terjadi di lapangan, tahapan ini untuk mengetahui fokus akar permasalahan sehingga dapat memberikan solusi yang tepat.

Tahap 2: Pengumpulan Data

Pada tahap ini merupakan aspek krusial yang dilakukan untuk mengetahui celah dan permasalahan, melalui data akan dilakukan analisis awal sehingga temuan permasalahan penelitian dapat lebih objektif dan menyeluruh. Data penelitian ini bersumber dari *kaggle.com* dengan variabel diantaranya: Akurasi Jawaban, Waktu Respon, Rating CSAT, dan Tingkat Eskalasi ke Manusia. Dimana total datanya adalah sebanyak ± 3000 data. Data dapat bersumber dari log interaksi chatbot, sistem *helpdesk*, survei kepuasan pengguna, maupun laporan performa internal masing-masing departemen.

departem	user_role	user_query	query_category	metrics_requested	analysis_type	bot_response_confidence	response_time_as	user_feedback_rating
HR	Intern	Forecast next quarter revenue	Revenue	Revenue	Descriptive	0.73	2381	4.0
Operations	Analyst	Forecast next quarter revenue	Marketing	Profit Margin	Descriptive	0.89	635	2.0
Finance	Manager	What is customer churn rate?	Marketing	Conversion Rate	Predictive	0.84	557	5.0
Operations	Analyst	Compare sales performance by region	Marketing	Conversion Rate	Trend Analysis	0.76	2499	4.0

Gambar 2. Sample Data

Tabel 1. Kriteria Tabel

Criteria	Name	Type
C1	Response Time (Waktu Respon)	Cost
C2	Bot Confidence Score (Rating CSAT)	Benefit
C3	Confidence response score (Akurasi jawaban)	Benefit
C4	Estimated business impact (Tingkat eskalasi ke manusia)	Benefit

Tahap 3: Pengolahan Data

Tahapan selanjutnya adalah mengaplikasi pendekatan sistem pendukung keputusan

dengan metode TOPSIS, dimana uraian langkah-langkahnya adalah sebagai berikut [5]:

A. Pembentukan Matriks Keputusan

Data yang telah dikumpulkan disusun ke dalam matriks keputusan dengan formula:

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix}$$

dimana baris merepresentasikan alternatif (departemen) dan kolom merepresentasikan kriteria (C₁ hingga C₄).

B. Normalisasi Matriks Keputusan

Matriks keputusan dinormalisasi menggunakan persamaan:

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^m x_{ij}^2}}$$

sehingga seluruh nilai kriteria berada pada skala yang dapat diperbandingkan.

C. Pembobotan Matriks Ternormalisasi (Konversi dari Skala Likert)

Matriks ternormalisasi dikalikan dengan bobot masing-masing kriteria, total seluruh bobot harus bernilai 1.00 atau (100%) dengan rumus normalisasi bobot likert:

$$W_j = \frac{\bar{S}_j}{\sum_{k=1}^n \bar{S}_k}$$

D. Penentuan Solusi Ideal Positif dan Negatif

Solusi ideal positif (A⁺) ditentukan dari nilai maksimum untuk kriteria *benefit* (y_j^+) = max (y_{ij}) dan nilai minimum untuk kriteria *cost* (y_j^-) = min (y_{ij}) sebaliknya untuk solusi ideal negatif (A⁻) dengan kriteria *benefit* (y_j^-) = min (y_{ij}) dan kriteria *cost* (y_j^+) = max (y_{ij})

E. Perhitungan Jarak terhadap Solusi Ideal

Setelah menemukan hasil dari solusi ideal positif dan solusi ideal negatif, kemudian dilakukan menghitung jarak euclidean setiap alternatif ke solusi

ideal positif (D_i^+) dan solusi ideal negatif (D_i^-) dengan rumus sebagai berikut:

$$D_i^+ = \sqrt{\sum_{j=1}^n (y_{ij} - y_j^+)^2}$$

(Formula mencari solusi jarak terhadap solusi ideal positif)

$$D_i^- = \sqrt{\sum_{j=1}^n (y_{ij} - y_j^-)^2}$$

(Formula mencari solusi jarak terhadap solusi ideal negatif)

F. Perhitungan Nilai Preferensi (Relative Closeness)

Kemudian tahapan selanjutnya mencari nilai preferensi untuk menentukan ranking performa *chatbot* tiap departemen, dengan formula sebagai berikut:

$$V_i = \frac{D_i^-}{D_i^- + D_i^+}$$

G. Perangkingan Alternatif

Departemen diurutkan berdasarkan nilai preferensi (V_i) yang tertinggi ke terendah, di mana nilai tertinggi menunjukkan kinerja *chatbot* terbaik.

H. Analisis dan Interpretasi Hasil

Hasil perangkingan dianalisis untuk mengidentifikasi departemen dengan kinerja *chatbot* optimal maupun yang memerlukan perbaikan, sekaligus dikaji faktor kriteria yang paling memengaruhi capaian tersebut.

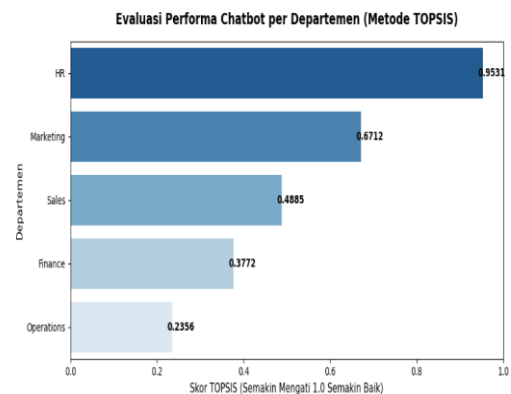
Tahap 4: Analisa Hasil

Setelah melewati serangkaian pengolahan data dengan menggunakan metode TOPSIS, selanjutnya diperoleh hasil keputusan berupa uraian ranking tiap departemen dan evaluasi *chatbot*nya. Dengan hasil demikian, dijadikan sebagai interpretasi untuk mengambil langkah kedepannya sehingga layanan perusahaan dapat ditingkatkan.

HASIL DAN PEMBAHASAN

Tabel 2. Perankingan dan Preferensi

Departemen	Nilai Preferensi	Rank
Finance	0.377179	4
Human Resource	0.953086	1
Marketing	0.671170	2
Operation	0.235586	5
Sales	0.488500	3



Gambar 3. Diagram Evaluasi Performa Chatbot

Berdasarkan hasil analisis multi-kriteria menggunakan metode TOPSIS (*Technique for Order of Preference by Similarity to Ideal Solution*), performa *chatbot* dievaluasi dengan memperhitungkan tiga indikator utama: *bot response confidence* (bobot 35%), *user feedback rating* (bobot 35%), dan *response time* (bobot 30%). Skor TOPSIS berkisar antara 0,0 hingga 1,0, di mana nilai yang mendekati 1,0 menunjukkan performa yang paling mendekati solusi ideal.

Peringkat Performa per Departemen: HR (Peringkat 1 — Skor TOPSIS: 0.9531)

Departemen HR menduduki posisi puncak dengan skor tertinggi (0,9531). Hal ini didorong oleh tingkat kepuasan pengguna (*rating* rata-rata 3,28) dan keyakinan respon bot (*confidence* 0,7986) yang merupakan yang tertinggi di antara seluruh departemen, serta waktu respon yang relatif cepat (1.340 ms).

Marketing (Peringkat 2 — Skor TOPSIS: 0.6712)

Departemen Marketing menempati posisi kedua dengan skor 0,6712. Faktor pendorong utamanya adalah waktu respon tercepat dibanding departemen lain (1.335 ms) dengan *rating* pengguna yang stabil (3,21).

1. Sales (Peringkat 3 — Skor TOPSIS: 0.4885)

Sales berada di posisi menengah dengan skor 0,4885. Meskipun *rating* feedback cukup baik (3,22), skor tertekan oleh tingkat *confidence* bot dan waktu respon yang moderat (1.372 ms).

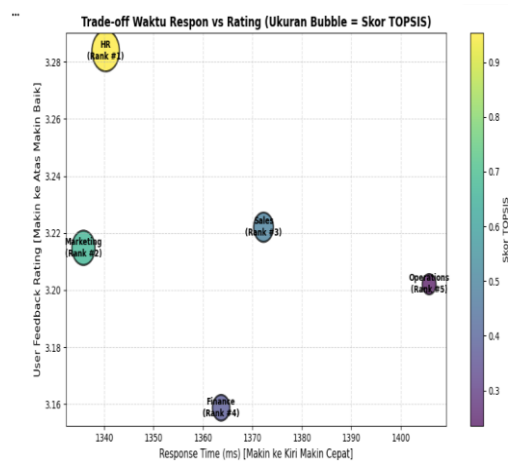
2. Finance (Peringkat 4 — Skor TOPSIS: 0.3772)

Finance mencatatkan skor 0,3772. Rendahnya skor ini dipengaruhi oleh kombinasi *rating* kepuasan terendah (3,15) dan *confidence rate* terkecil (0,7911) di antara seluruh grup.

3. Operations (Peringkat 5 — Skor TOPSIS: 0.2356)

Departemen Operations menempati posisi terendah dengan skor 0,2356. Penurunan performa secara signifikan disebabkan oleh waktu respon paling lambat (1.405 ms), yang menjadi pemicu utama rendahnya keunggulan kriteria pada perhitungan TOPSIS.

Implementasi chatbot berkinerja paling optimal pada departemen **HR** dan **Marketing**. Evaluasi mendalam perlu dilakukan pada departemen **Operations** (terutama optimisasi kecepatan pemrosesan respon) dan **Finance** (peningkatan akurasi/confidence jawaban) untuk meningkatkan efektivitas layanan chatbot secara menyeluruh.



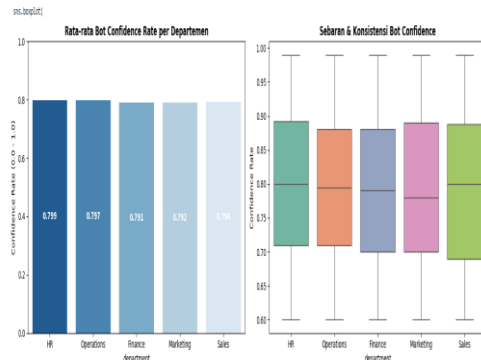
Gambar 4. Scatter Plot

Visualisasi *scatter plot* Trade-off Waktu Respon vs Rating memberikan gambaran intuitif mengenai posisi performa setiap departemen berdasarkan dua kriteria kritis, di mana sumbu horizontal menunjukkan *Response Time* (makin ke kiri makin cepat), sumbu vertikal menunjukkan *User Feedback Rating* (makin ke atas makin tinggi kepuasan), serta ukuran dan warna *bubble* mewakili besarnya nilai akhir Skor TOPSIS. Dari diagram tersebut, terlihat jelas keunggulan mutlak departemen HR (Rank 1) yang berada di sudut kiri atas dengan ukuran *bubble* terbesar dan warna paling terang (kuning-hijau). Hal ini menandakan bahwa HR berhasil mencapai keseimbangan ideal antara kecepatan respon yang sangat baik (sekitar 1.340 ms) dan tingkat kepuasan pengguna tertinggi (rating di atas 3,28).

Sementara itu, departemen Marketing (Rank 2) menempati posisi paling kiri dengan *response time* tercepat (sekitar 1.335 ms), namun memiliki tingkat rating yang sedikit lebih rendah dibanding HR. Posisi tengah diisi oleh Sales (Rank 3) dengan tingkat kepuasan rata-rata dan waktu respon sedang. Di sisi lain, Finance (Rank 4) dan Operations (Rank 5) berada di area bawah dan kanan dengan *bubble* berukuran paling kecil berwarna gelap (ungu). Finance tertekan oleh skor kepuasan pengguna yang berada di posisi terbawah (sekitar 3,15), sedangkan Operations berada di posisi paling kanan karena memiliki waktu respon paling lambat (melebihi 1.400 ms).

Secara keseluruhan, diagram ini menegaskan bahwa untuk menaikkan skor TOPSIS secara signifikan, suatu departemen tidak cukup hanya mengandalkan kecepatan respon semata, melainkan harus dibarengi dengan kualitas jawaban yang mampu mendongkrak *rating* kepuasan pengguna secara sejajar.

Analisis Rata-rata & Sebaran Konsistensi Bot Confidence



Gambar 5. Sebaran Konsistensi Rata-Rata

Visualisasi dual-plot Rata-rata Bot Confidence Rate (*barplot*) dan Sebaran & Konsistensi Bot Confidence (*boxplot*) menunjukkan bahwasanya tingkat keyakinan (*confidence rate*) bot di seluruh departemen berada pada rentang rata-rata yang relatif setara dan homogeny [14], yaitu berkisar antara 0,791 hingga 0,799 (sekitar 79%–80%). HR mencatatkan rata-rata *confidence* tertinggi sebesar 0,799, disusul Operations (0,797), Sales (0,794), Marketing (0,792), dan Finance (0,791). Melalui diagram *boxplot*, terlihat bahwa sebaran nilai *confidence* kelima departemen memiliki rentang interkuartil (IQR) dan batas nilai (*minimum* 0,60 hingga *maximum* 0,99) yang hampir identik dengan deviasi standar sekitar 0,109–0,112. Hal ini mengindikasikan bahwa model pembelajaran mesin pada chatbot memiliki stabilitas dasar yang seragam dalam memahami *prompt* atau *query* pengguna, tanpa adanya bias

performa pemrosesan bahasa yang timpang pada departemen tertentu.

Kaitan dengan Evaluasi Performa Keseluruhan (TOPSIS & Trade-off)

Meskipun rata-rata dan variabilitas *bot confidence* antar-departemen terbukti sangat tipis perbedaannya, indikator ini tetap memberikan kontribusi krusial ketika diintegrasikan ke dalam analisis keputusan terbobot TOPSIS dan *trade-off* sebelumnya:

Pendorong Skor Puncak HR: Pada departemen HR (Rank 1), nilai *confidence* yang tertinggi (0,7986) saling memperkuat (*amplifying*) dengan *user rating* tertinggi (3,28) dan waktu respon yang cepat (1.340 ms). Kombinasi keunggulan di ketiga kriteria sekaligus inilah yang mendongkrak skor TOPSIS HR hingga mendekati sempurna (0,9531).

Penentu Pembobot Diferensial: Karena nilai *bot confidence* secara umum hampir seragam di angka 0,79, kriteria ini bertindak sebagai faktor penetral (*baseline factor*). Dampaknya, variasi akhir pada skor TOPSIS dan posisi *bubble* pada grafik *trade-off* justru dipengaruhi secara lebih sensitif oleh Response Time dan User Feedback Rating.

Kesenjangan Performa vs Kualitas: Kasus Operations (Rank 5) memperlihatkan bahwa meskipun *bot confidence*-nya tergolong tinggi (0,7975), skor TOPSIS-nya tetap terpuruk (0,2356) akibat *response time* yang paling lambat (1.405 ms). Sebaliknya, Finance (Rank 4) yang memiliki *confidence* paling rendah (0,7912) berbanding lurus dengan kepuasan pengguna yang juga paling rendah (3,15), menegaskan bahwa ketidakpastian respon bot berdampak langsung pada penurunan kepuasan pengguna.

DAFTAR PUSTAKA

A. A. Sukri, Salamun, “Optimasi Chatbot

dengan Pemanfaatan Natural

- Language Processing,” pp. 17–26.
- A. A. Hizkia Triputra Gultom, Danillah Albaronah Faza Darmagama, “PELANGGAN DAN KEPUASAN KONSUMEN DALAM KONTEKS BISNIS,” vol. 1, no. 4, pp. 1639–1651, 2025.
- R. D. Gunawan, “Penerapan Sistem Pendukung Keputusan dalam Pemilihan Supplier dengan Metode TOPSIS,” *J. Inf. Technol. Softw. Eng. Comput. Sci.*, vol. 2, no. 3, pp. 150–159, 2024, doi: 10.58602/itsecs.v2i3.157.
- D. Triyana and A. O. Moeis, “Chatbot Platform Selection in Marketplace using AHP-TOPSIS Methods: A Case Study in Indonesia,” pp. 1–10, 2023, doi: 10.46254/af03.20220293.
- M. F. Azis, F. P. Latuan, V. S. Penlaana, and Y. R. Kaesmetan, “Sistem Pendukung Keputusan Pemilihan Website Terbaik Mahasiswa Stikom Uyelindo Kupang Menggunakan Metode TOPSIS,” *J. Technol. Informatics*, vol. 5, no. 2, pp. 93–100, 2024, doi: 10.37802/joti.v5i2.565.
- G. Ilieva, “Extension of Interval-Valued Hesitant Fermatean Fuzzy TOPSIS for Evaluating and Benchmarking of Generative AI Chatbots,” *Electron.*, vol. 14, no. 3, 2025, doi: 10.3390/electronics14030555.
- S. F. T. T. Y. ile C. S. YZ-Destekli and P. Değerlendirmesi, “Intuitionistic Fuzzy Topsis Approach for Chatbot Selection : AI-Assisted Performance Evaluation Sezgisel Fuzzy Topsis,” vol. 10, no. 1, pp. 1–18, 2026.
- H. Suriadi, M. Mulyono, and W. Lovita, “Pemanfaatan Teknologi AI untuk Meningkatkan Kualitas dan Responsivitas Pelayanan Publik di Era Digital,” *J. Media Ilmu*, vol. 3, no. 2, pp. 107–132, 2024, doi: 10.31869/jmi.v3i2.6069.
- Hernawati, Jurnaidi Ahmad, and wahidin, “Implementasi Metode TOPSIS Dalam Penentuan Dosen Favorit Berdasarkan Preferensi Mahasiswa,” vol. Volume 4 N, no. 2, pp. 68–75, 2024, [Online]. Available: <https://jurnal.bsi.ac.id/index.php/prof itabilita>
- K. D. Valerian *et al.*, “Sistem Pendukung Keputusan Evaluasi Kepuasan Peserta Pelatihan E-commerce Berbasis TOPSIS,” vol. 9, no. 2, pp. 182–196, 2026.
- E. P. Pratama, D. Jovan, A. Tombokan, M. Abelard, and W. Gondho, “Literature Review : Pemanfaatan Chatbot untuk Mendukung Manajemen Pengetahuan”.
- G. Ilieva, "Extension of Interval-Valued Hesitant Fermatean Fuzzy TOPSIS for Evaluating and Benchmarking of Generative AI Chatbots," *Electronics*, vol. 14, no. 3, art. no. 555, Jan. 2025, doi: 10.3390/electronics14030555.
- A. D. A. N. S. Supriyanto, R. S. Utami, and D. P. Astuti, "Chatbot Quality and Its Impact on User Satisfaction and Continuance Usage Intention in the Indonesian Banking Industry," *Journal of Theoretical and Applied Information Technology*, vol. 102, no. 15, pp. 6120–6132, Aug. 2026.
- M. A. Hasan and S. B. Ahmed, "Intuitionistic Fuzzy TOPSIS Approach for Chatbot Selection: AI-Assisted Performance Evaluation," *International Journal of Information Technology & Decision Making*, vol. 23, no. 4, pp. 1425–1448, Jul. 2024, doi: 10.1142/S021962202450012X.
- S. Suhaili, S. S. Salim, and M. N. A. Khan, "Evaluating Customer Satisfaction and Service Quality Metrics in Conversational Agent Interactions," *IEEE Access*, vol. 11, pp. 48210–48225, May 2023, doi: 10.1109/ACCESS.2023.3274190.
- R. E. A. N. Firmansyah and L. A. Abdillah, "Multi-Criteria Performance Analysis of AI Chatbots in Enterprise Resource Planning Systems Using TOPSIS," *IEEE Transactions on Engineering Management*, vol. 69, no. 6, pp. 3890–3901, Dec. 2022, doi: 10.1109/TEM.2022.3185200.