PENERAPAN NATURAL LANGUAGE PROCESSING PADA HELPDESK IT UNTUK OTOMATISASI RESPON AWAL

Irvan Ronaldi Napitupulu[1)], Sukma Evadini[1)], Nadya Satya Handayani[2)]

[1]Teknologi Rekayasa Perangkat Lunak, Politeknik Negeri Batam
[2]Teknik Informatika, Politeknik Negeri Batam

E-mail: sukma@polibatam.ac.id

Abstract

IT Helpdesk services are tasked with resolving a wide range of technology-related issues faced by users within an organization. Because the initial handling of tickets still depends heavily on IT personnel, response times can become suboptimal, especially as ticket volume grows. This research aims to apply Natural Language Processing to automate the first-line response to user complaints by leveraging a Knowledge Base. The Prototyping method is used, built on a microservice architecture. The processing pipeline consists of text preprocessing, synonym normalization, Term Frequency-Inverse Document Frequency (TF-IDF) weighting, document representation via the Vector Space Model, and similarity computation using Cosine Similarity to produce the most relevant solution recommendation (Top-1 Retrieval). A Confidence Threshold of 20% is also applied to screen out low-similarity recommendations, ensuring that such tickets are automatically escalated to the IT team. Robustness testing showed an overall relevance rate of 96%, demonstrating the system's ability to handle short-query variations, synonym differences, and minor typos though it remains limited in interpreting paraphrased text and variations in sentence meaning.

Keywords: *Helpdesk, Natural Language Processing, TF-IDF, Cosine Similarity, Robustness Testing*

PENDAHULUAN

Layanan Helpdesk IT memegang peran penting dalam menangani beragam kendala teknologi informasi yang dihadapi pengguna di instansi/organisasi. Semakin intensifnya pemanfaatan teknologi informasi turut mendorong lonjakan volume permintaan bantuan yang masuk ke tim IT, yang pada akhirnya memperlambat respon awal dan menurunkan efisiensi layanan secara keseluruhan. Menyikapi hal ini, tidak sedikit organisasi yang mulai menerapkan sistem Helpdesk Ticketing guna mencatat, memantau, dan mengelola tiket secara lebih terstruktur (Tarigan dkk, 2022).

Kendati sistem Helpdesk Ticketing telah memberikan kemudahan dalam pengelolaan tiket, mayoritas sistem yang ada masih menempatkan staf IT sebagai pihak yang memberikan respon permulaan terhadap setiap permintaan perbaikan yang masuk. Akibatnya, pengguna kerap harus menanti giliran hingga tiketnya diproses oleh petugas, apalagi ketika volume tiket sedang tinggi. Untuk menjawab tantangan ini, penerapan Natural Language Processing (NLP) menjadi salah satu pendekatan yang memungkinkan

sistem memberikan pilihan solusi secara otomatis dengan memanfaatkan Knowledge Base yang telah dimiliki perusahaan (Fauzan dkk, 2024).

Sejumlah penelitian terdahulu telah mengangkat tema pengembangan sistem Helpdesk maupun penerapan NLP. Nugroho dan Fianty (2023) menitikberatkan pembahasan pada peningkatan mutu layanan melalui kerangka kerja ITIL, sementara Junadi (2023) beserta Liharja dkk (2022) merancang sistem Helpdesk berbasis web yang dilengkapi fitur pengelolaan tiket. Sementara itu, Fridayanto dan Wibowo (2025) mengaplikasikan metode TF-IDF dan Cosine Similarity untuk memberikan jawaban otomatis berlandaskan NLP, kendati penerapannya belum terhubung secara utuh dengan sistem bantuan tiket. Merujuk pada penelitian tersebut, tampak masih adanya celah penelitian, yakni belum hadirnya sistem yang memadukan otomatisasi respon permulaan berbasis NLP dengan manajemen sekaligus eskalasi tiket pada satu kesatuan sistem.

Berangkat dari problematika tersebut, penelitian ini dirancang untuk menerapkan NLP sistem Helpdesk IT dalam rangka mengotomatisasi respon permulaan atas keluhan pengguna, dengan memanfaatkan metode Term Frequency–Inverse Document Frequency (TF-IDF), Vector Space Model (VSM), dan Cosine Similarity. Pengembangan sistem dilakukan melalui metode Prototyping dengan arsitektur microservice dengan Laravel berperan sebagai sistem inti dan Flask menjadi fitur pendukung NLP (Purwanto, 2023).

Kontribusi yang ditawarkan penelitian ini adalah terwujudnya sistem Helpdesk IT yang memadukan proses pembuatan tiket dengan modul NLP untuk menghadirkan rekomendasi solusi permulaan secara otomatis berbasis Knowledge Base. Selain mengukur performa sistem melalui metrik akurasi, penelitian ini turut melakukan Robustness Testing guna menilai sejauh mana sistem mampu menghadapi variasi input pengguna. Melalui pendekatan ini, penelitian tidak hanya berfokus pada tingkat akurasi sistem, tetapi juga menelaah ketahanann kondisi pemakaian yang mendekati situasi nyata.

METODE PENELITIAN

A. Arsitektur dan Perancangan Sistem

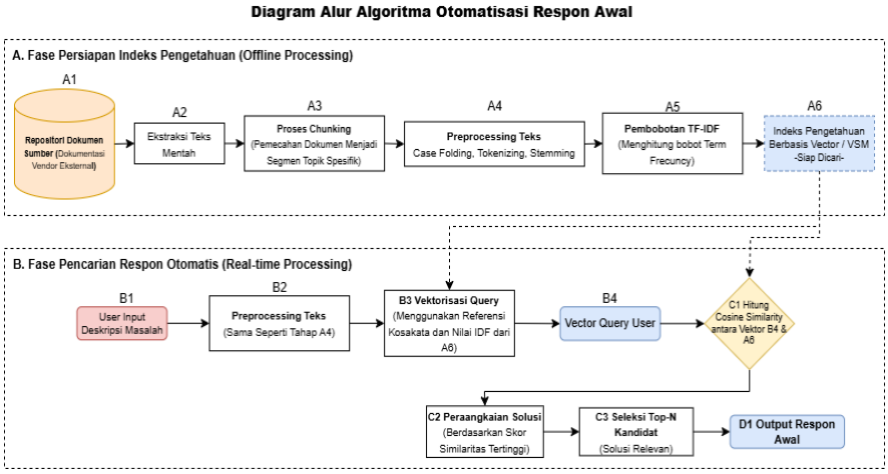
Sistem ini dibangun dengan menerapkan arsitektur microservice yang memisahkan secara jelas antara aplikasi utama dan modul NLP. Arsitektur tersebut tersusun atas empat komponen inti, yakni: Aplikasi Web berbasis Laravel, Layanan NLP yang dibangun menggunakan framework Flask (Python), Basis Data MySQL sebagai

media penyimpanan data, serta REST API yang berfungsi sebagai jembatan komunikasi antara aplikasi Laravel dan layanan Flask (Saepuloh, 2021).

B. Pemrosesan Natural Language Processing (NLP)

Proses NLP dalam sistem ini terbagi ke dalam dua tahapan utama, yaitu pemrosesan secara offline dan pemrosesan real-time. Pada tahap offline processing, sebanyak 50 topik dokumen Knowledge Base dalam format PDF yang dihimpun dari staf IT diekstraksi terlebih dahulu, kemudian melalui serangkaian proses berupa chunking, text preprocessing, hingga pembobotan dengan metode Term Frequency-Inverse Document Frequency (TF-IDF) guna membangun indeks dokumen berbasis Vector Space Model (VSM).

Sementara itu, pada tahap real-time processing, setiap deskripsi masuk yang dikirimkan oleh pengguna diolah melalui tahapan preprocessing yang serupa (Ahmed dkk, 2023), mencakup Case Folding, Tokenizing, dan Stemming.Tahap praproses bertujuan untuk mengurangi variasi kata yang tidak diperlukan sehingga proses pencocokan dapat dilakukan secara lebih efektif (Evadini, 2022).



Gambar 1. Tahapan Pemrosesan NLP

Untuk membuktikan praproses yang diterapkan pada sistem, dilakukan pengujian menggunakan kalimat masukan pengguna: "Gak bisa konek wifi". Hasil dari setiap tahapan praproses ditunjukkan pada Tabel 1.

Tabel 1
Sampel *Knowledge Base*

Tahap	Kategori
Teks Asli	Gak bisa konek wifi
Case Folding	gak bisa konek wifi

Tokenizing	[gak, bisa, konek, wifi]
Synonim Normalization	[tidak, bisa, terhubung, wifi]
Stopword Removal	[tidak, bisa, terhubung, wifi]
Stemming	[tidak, bisa, hubung, wifi]
Hasil Akhir	"tidak bisa hubung wifi"

Berdasarkan hasil pada sistem berhasil melakukan normalisasi terhadap kata tidak baku yang umum digunakan pengguna, seperti kata "gak" yang diubah menjadi "tidak". Proses ini membantu sistem dalam memahami berbagai variasi penulisan pertanyaan sehingga pencarian solusi pada basis pengetahuan dapat dilakukan secara lebih konsisten. Setelah praproses teks selesai dilakukan, setiap dokumen pada basis pengetahuan direpresentasikan dalam bentuk numerik menggunakan TF-IDF.

Untuk memperlihatkan proses pembobotan TF-IDF, digunakan empat dokumen sampel hasil praproses sebagai ilustrasi perhitungan.

a. Dokumen 1 : tidak bisa login sistem
b. Dokumen 2 : password akun kunci
c. Dokumen 3 : lupa password tidak bisa masuk
d. Dokumen 4 : tidak bisa akses folder

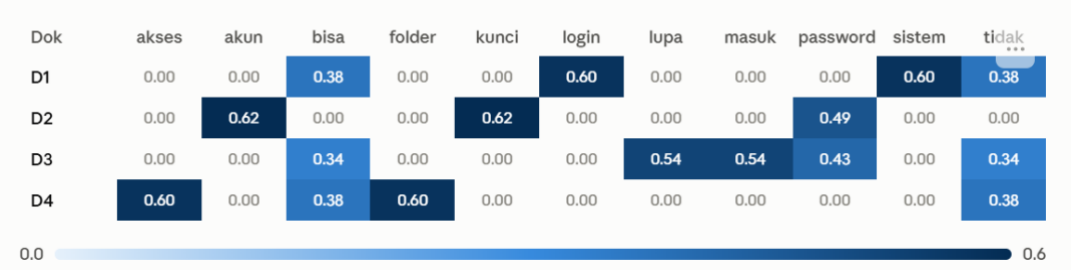
Selanjutnya sistem melakukan proses ekstraksi kosakata untuk membentuk himpunan kata yang akan digunakan sebagai fitur pada proses pembobotan TF-IDF.

```
D:\laragon\www\Ticketing-TA\ai-service>python test_tfidf.py
============================================================
1. PROSES INIT TF-IDF VECTORIZER
============================================================
Vocabulary (Daftar Kata yang diekstrak):
['akses' 'akun' 'bisa' 'folder' 'kunci' 'login' 'lupa' 'masuk' 'password'
 'sistem' 'tidak']


============================================================
2. MATRIKS TF-IDF (Term Frequency - Inverse Document Frequency)
============================================================
           akses    akun    bisa  folder   kunci  login    lupa   masuk  password  sistem   tidak
Dokumen 1  0.000  0.0000  0.3804   0.000  0.0000  0.596  0.0000  0.0000    0.0000   0.596  0.3804
Dokumen 2  0.000  0.6176  0.0000   0.000  0.6176  0.000  0.0000  0.0000    0.4869   0.000  0.0000
Dokumen 3  0.000  0.0000  0.3443   0.000  0.0000  0.000  0.5394  0.5394    0.4253   0.000  0.3443
Dokumen 4  0.596  0.0000  0.3804   0.596  0.0000  0.000  0.0000  0.0000    0.0000   0.000  0.3804
```

Gambar 2. Vocabulary Hasil Ekstraksi TF-IDF

Berdasarkan hasil ekstraksi, sistem menghasilkan 11 kata unik yang digunakan sebagai fitur representasi dokumen, yaitu: akses, akun, bisa, folder, kunci, login, lupa, masuk, password, sistem, dan tidak. Setelah vocabulary terbentuk, sistem menghitung bobot TF-IDF untuk setiap kata pada masing-masing dokumen.



Gambar 3. Hasil Pembobotan TF-IDF

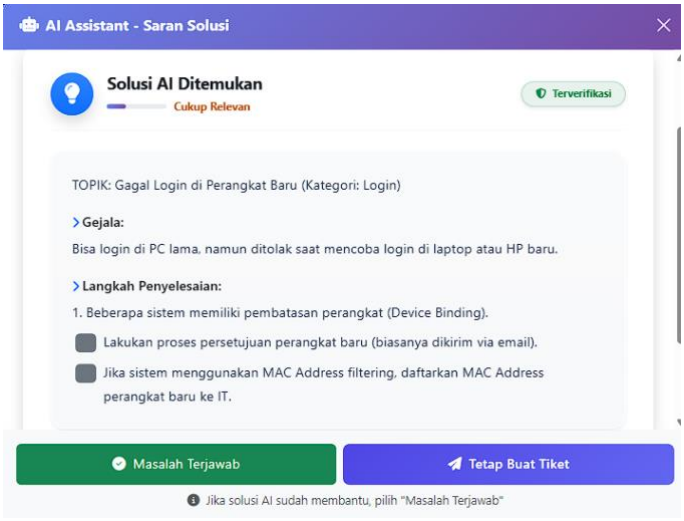
Pemberian bobot TF-IDF tersebut kemudian diterjemahkan ke dalam bentuk Vector Space Model (VSM), yaitu representasi matematis dokumen dalam bentuk vektor numerik. Pada model ini, setiap dokumen direpresentasikan sebagai kumpulan nilai bobot TF-IDF yang menggambarkan karakteristik dokumen berdasarkan kata-kata penyusunnya. Hasil representasi vektor inilah yang selanjutnya digunakan pada proses pencarian solusi otomatis menggunakan Cosine Similarity.

Cosine Similarity digunakan untuk mengukur sudut antara vektor pertanyaan pengguna. Nilai kemiripan berada pada rentang 0 sampai 1, di mana nilai yang semakin mendekati 1 menunjukkan tingkat kemiripan yang semakin tinggi. Selanjutnya, sistem menghitung nilai Cosine Similarity terhadap 51 dokumen pengetahuan yang terdapat pada basis pengetahuan. Berikut adalah hasil pencarian sistem untuk mendapatkan 5 dokumen dengan nilai kemiripan tertinggi.

Tabel 2
Nilai Similarity

Peringkat	Topik Solusi	Nilai Similarity
1	Gagal Login di Perangkat Baru	0,3126
2	Lupa Password	0,2994
3	Website Internal	0,2055
4	Session Expired	0,2054
5	Tidak Bisa Login	0,2014

Berdasarkan hasil perhitungan Cosine Similarity, sistem memberikan peringkat solusi berdasarkan tingkat kemiripan kata antara pertanyaan pengguna dan dokumen pada basis pengetahuan.



Gambar 4. Hasil Pencarian Solusi Menggunakan Cosine Similarity

C. Mekanisme *Confidence Threshold*

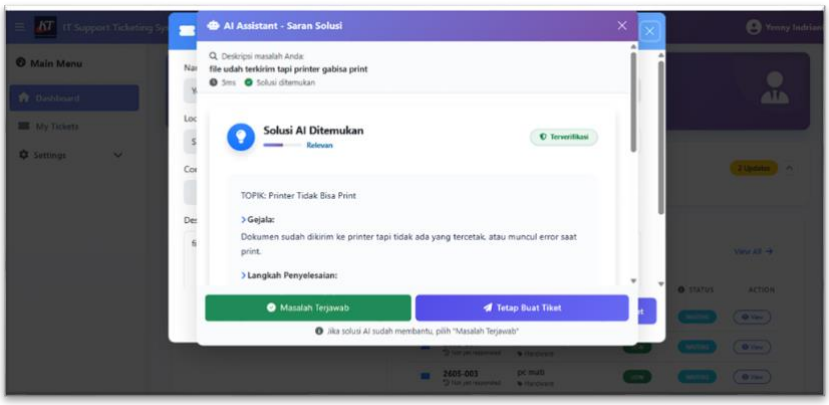
Untuk meningkatkan performa rekomendasi, sistem mengimplementasikan mekanisme Confidence Threshold dengan nilai 20%. Nilai ambang batas dipilih setelah melalui proses uji coba dan titik keseimbangan threshold diperoleh pada angka 20%. Jika threshold terlalu tinggi maka sistem menjadi terlalu ketat, sehingga banyak tiket yang sebenarnya punya solusi identik di riwayat tetap dianggap tidak cukup mirip. Sebaliknya, jika threshold terlalu rendah maka sistem menjadi terlalu permisif, sehingga tiket dengan topik berbeda pun dipaksakan mendapat rekomendasi solusi yang tidak relevan.

HASIL DAN PEMBAHASAN

A. Implementasi Sistem

Dokumen Knowledge Base yang digunakan terdiri dari 50 topik permasalahan yang mencakup beberapa kategori, seperti masalah akun pengguna, perangkat keras, perangkat lunak, printer, jaringan, email, dan akses sistem. Selanjutnya data uji digunakan untuk mengevaluasi kemampuan sistem dalam menemukan solusi yang relevan (Supiyanto & Sriyono, 2023).

` Untuk membuat tiket baru, pengguna cukup mengisi kategori beserta deskripsi permasalahan yang dihadapi. Sebelum tiket tersebut diteruskan ke Tim IT, sistem akan lebih dulu menelusuri kemungkinan solusi pada Knowledge Base melalui modul NLP. Selanjutnya, hasil rekomendasi yang diperoleh beserta nilai similarity-nya akan ditampilkan kepada pengguna.



Gambar 5. Rekomendasi Solusi Menggunakan NLP

B. Evaluasi Kinerja Sistem menggunakan *Robustness Testing*

Robustness Testing dilaksanakan guna mengukur sejauh mana sistem mampu menghadapi berbagai variasi masukan dari pengguna.

Tabel 3
Robustness Testing

Kategori	Relevansi
Typographical Robustness	100%
Synonym Robustness	100%
Paraphrase Robustness	80%
Query Reformulation	100%
Incomplete Query	100%
Rata-rata Relevansi	96%

Dari hasil pengujian, sistem terbukti mampu mempertahankan tingkat relevansi hingga 96%. Seluruh kategori pengujian mencatatkan nilai 100%, kecuali pada kategori Paraphrase Robustness.

SIMPULAN

Penelitian ini berhasil mengimplementasikan NLP pada sistem Helpdesk IT guna mengotomatisasi pemberian respon permulaan atas keluhan yang disampaikan pengguna. Melalui pendekatan TF-IDF dan Cosine Similarity, sistem mampu menghasilkan rekomendasi solusi yang cukup andal, sebagaimana ditunjukkan oleh hasil Robustness Testing dengan tingkat relevansi mencapai 96%. Selain itu, penerapan mekanisme fConfidence Threshold sebesar 20% turut berkontribusi dalam meningkatkan kemampuan sistem, sebab hanya rekomendasi yang mencapai ambang batas kemiripan tertentu yang akan disampaikan kepada pengguna.

DAFTAR PUSTAKA

Ahmed, S., Singh, M., Doherty, B., Ramlan, E., Harkin, K., Bucholc, M., Coyle, D. (2023). Knowledge-based Intelligent System for IT Incident DevOps. *IEEE/ACM International Workshop on Cloud Intelligence & AIOps (AIOps)*. 1-7.

Anam, M. K., Efendi, A., Yuliana, D., Syahputra, I., Sari, A, K., Handayani, N, S., Lestari, T. P., Evadini, S., Fadillah, N., Hadisugelar, D. (2026). Optimization of Convolutional Neural Network Architecture forAccurate Banana Ripeness Classification. *JOIV: International Journal on Informatics Visualization*, 10(4), 1604-1611.

Evadini, S. (2022). Analisis Faktor Risiko Kematian dengan Penyakit Komorbid COVID-19 menggunakan Algoritma ECLAT. *JIDT : Jurnal Informasi dan Teknologi*. 4(1). 52-57

Fauzan, M. F., Imanda, R., & Hasbi, M. A. (2024). Designing a chatbot with NLP technology in a website-based new student admission information system. *Journal of Applied Informatics and Computing*, 8(2), 358–366.

Fridayanto, T. M., & Wibowo, A. P. W. (2025). Design of an automatic help desk Response module using natural language processing. *Brilliance: Research of Artificial Intelligence*, 5(1), 551-558.

Junadi, B. (2024). Rancangan sistem informasi IT helpdesk berbasis web pada PT Nusa Raya Cipta Tbk. Jutisi: *Jurnal Ilmiah Teknik Informatika dan Sistem Informasi*, 13(2), 1434-1443.

Kupiyalova, A., Satybaldiyeva, R., Aiaskarov, S. (2020). Semantic search using Natural Language Processing. *IEEE 22nd Conference on Business Informatics (CBI)*. 96-100

Liharja, Y., Sari, A. O., & Satriansyah, A. (2022). Rancangan bangun sistem informasi helpdesk IT support berbasis website. *Jurnal Informatika dan Teknologi*, 5(1), 157-166.

Nugroho, A. F. J., & Fianty, M. I. (2023). Streamlining IT help desk and incident management: Harnessing the power of the ITIL framework for enchanced efficiency in IT service. *Journal of Information System and Informatics*, 5(2), 683-695.

Purwanto, T. (2023). Analisa perbandingan kinerja REST API dengan framework Flask, Laravel, dan Express.js. Scientia Sacra: *Jurnal Sains, Teknologi dan Masyarakat*, 3(4), 49–55.

Saepuloh, M. (2021). Perancangan sistem informasi manajemen helpdesk berbasis web dengan framework CodeIgniter dan MySQL. *Jurnal Teknik Informatika dan Sistem Informasi (JATISI)*, 8(4), 2261–2276.

Supiyanto., Sriyono. (2023). Metode *Cosine Similarity* Untuk Mendeteksi Kemiripan Pada Dokumen Teks. *SAINS Jurnal MIPA dan Pengajarannya*. 1(1). 1-7.

Tarigan, R., Kusosi, I., & Usri, A. (2022). Perancangan aplikasi helpdesk ticketing system pada PT Indonesia Nippon Seiki. *Jurnal SISFOKOM (Sistem Informasi dan Komputer)*, 11(1), 9-18.

Wahyuddin, M. N., & Putri, D. P. (2024). Designing a helpdesk ticketing system to improve the efficiency of IT support services on an XYZ company-based website using the Laravel framework. Matrix: *Jurnal Manajemen Teknologi dan Informatika*, 14(1), 9–24.