



Implementasi Smart Learning Berbasis Hybrid Chatbot Menggunakan Metode TF-IDF dan Fuzzy Matching untuk Meningkatkan Literasi Digital Mahasiswa

Heru Mustofa^{1),a)}, Agung Budi Susanto^{2),a)}, Lina Marlina^{3),a)}

¹²³⁾Universitas Pamulang, Kota Tangerang Selatan, Indonesia

Email : herumustofa871@gmail.com^{a)} ; dosen02680@unpam.ac.id^{b)}
dosen02921@unpam.ac.id^{b)}
(Menggunakan Email Kampus)

ABSTRACT

The rapid advancement of digital technology has encouraged higher education institutions to adopt innovative learning media that support students' digital literacy. One of the promising technologies is an educational chatbot that provides interactive and responsive learning assistance. This study aims to implement a Smart Learning system based on a Hybrid Chatbot using Term Frequency–Inverse Document Frequency (TF-IDF) and Fuzzy Matching to support digital learning. This research employed the Research and Development (R&D) method using the ADDIE model, consisting of Analysis, Design, Development, Implementation, and Evaluation stages. The system was developed using Python and the Flask framework, while the knowledge base consisted of 400 question–answer pairs organized into ten knowledge categories. Data analysis was conducted through text preprocessing, Rule-Based Matching, TF-IDF vectorization, Cosine Similarity, and Fuzzy Matching, followed by system performance evaluation. The results indicate that the proposed Hybrid Chatbot is capable of recognizing various user query patterns, including different sentence structures and typographical errors, thereby providing more relevant responses than conventional exact matching. The developed Smart Learning application offers an interactive, adaptive, and accessible digital learning environment that can support students' independent learning and contribute to the implementation of technology-enhanced learning in higher education.

Keywords: Smart Learning; Hybrid Chatbot; TF-IDF; Fuzzy Matching; Digital Literacy.

ABTRAK

Transformasi digital dalam pendidikan tinggi mendorong pengembangan media pembelajaran yang mampu mendukung proses belajar secara interaktif dan adaptif. Penelitian ini bertujuan mengimplementasikan Smart Learning berbasis Hybrid Chatbot menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) dan Fuzzy Matching sebagai media pembelajaran digital. Penelitian menggunakan metode Research and Development (R&D) dengan model ADDIE yang meliputi tahapan *Analysis*, *Design*, *Development*, *Implementation*, dan *Evaluation*. Sistem dikembangkan menggunakan bahasa pemrograman Python dengan framework Flask dan memanfaatkan basis pengetahuan yang terdiri atas 400 pasangan pertanyaan dan jawaban. Teknik analisis data dilakukan melalui tahapan *text preprocessing* yang meliputi *lowercasing*, penghapusan tanda baca, tokenisasi, dan stemming, dilanjutkan dengan Rule-Based Matching, pembobotan TF-IDF, perhitungan Cosine Similarity, serta Fuzzy Matching sebagai mekanisme pencarian alternatif ketika tingkat kemiripan belum memenuhi nilai ambang. Hasil penelitian menunjukkan bahwa chatbot mampu mengenali berbagai variasi pertanyaan pengguna, termasuk perbedaan struktur kalimat dan kesalahan pengetikan, sehingga menghasilkan jawaban yang lebih relevan dibandingkan pendekatan pencocokan kata secara langsung. Dengan demikian, implementasi Smart Learning berbasis Hybrid Chatbot dapat menjadi alternatif media pembelajaran digital yang interaktif, adaptif, dan mudah diakses untuk mendukung proses pembelajaran di perguruan tinggi.

Kata kunci: Smart Learning; Hybrid Chatbot; TF-IDF; Fuzzy Matching; Literasi Digital.

PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi telah membawa perubahan mendasar terhadap penyelenggaraan pendidikan tinggi. Transformasi digital tidak lagi hanya berorientasi pada digitalisasi administrasi akademik, tetapi telah berkembang menuju implementasi Smart Learning, yaitu suatu ekosistem pembelajaran yang memanfaatkan teknologi digital, kecerdasan buatan (*Artificial Intelligence*), komputasi awan (*cloud computing*), analitik data, dan sistem berbasis web untuk menciptakan pengalaman belajar yang adaptif, kolaboratif, personal, serta berpusat pada mahasiswa. Dalam paradigma ini, mahasiswa tidak hanya menjadi penerima informasi, tetapi juga menjadi subjek aktif yang mampu membangun pengetahuan melalui interaksi dengan berbagai sumber belajar digital. Oleh karena itu, perguruan tinggi dituntut mampu mengintegrasikan teknologi cerdas ke dalam proses pembelajaran sebagai bagian dari transformasi menuju Smart Campus yang mendukung peningkatan kualitas pendidikan, efisiensi layanan, dan daya saing lulusan di era digital.

Transformasi tersebut menuntut mahasiswa memiliki kemampuan literasi digital yang memadai. Literasi digital tidak hanya diartikan sebagai kemampuan menggunakan perangkat teknologi, tetapi juga mencakup kemampuan mengakses, memahami, mengevaluasi, mengelola, menciptakan, dan mengomunikasikan informasi secara kritis,

kreatif, etis, serta bertanggung jawab melalui media digital. Kemampuan tersebut menjadi kompetensi penting dalam mendukung pembelajaran sepanjang hayat (*lifelong learning*), pemecahan masalah, pengambilan keputusan berbasis informasi, serta adaptasi terhadap perkembangan teknologi yang berlangsung sangat cepat. Berbagai penelitian menunjukkan bahwa peningkatan literasi digital berpengaruh positif terhadap kualitas pembelajaran, kemampuan berpikir kritis, kolaborasi, kreativitas, serta kesiapan mahasiswa dalam menghadapi tantangan dunia kerja berbasis digital (Judijanto, 2024; Gustiana & Satria, 2024; Pratama et al., 2025; Rahim & Indah, 2024). Oleh karena itu, pengembangan media pembelajaran yang mampu mendorong peningkatan literasi digital menjadi kebutuhan yang semakin mendesak dalam pendidikan tinggi.

Seiring berkembangnya teknologi kecerdasan buatan, penggunaan chatbot sebagai media pembelajaran mengalami peningkatan yang sangat pesat. Chatbot merupakan sistem percakapan otomatis yang dirancang untuk berinteraksi dengan pengguna menggunakan bahasa alami (*Natural Language Processing/NLP*), sehingga mampu memberikan respons terhadap berbagai pertanyaan secara cepat, konsisten, dan interaktif. Dalam konteks pendidikan tinggi, chatbot tidak hanya dimanfaatkan sebagai media penyampaian informasi akademik, tetapi juga berkembang menjadi virtual learning assistant yang membantu mahasiswa memperoleh materi pembelajaran, menjawab pertanyaan, memberikan rekomendasi sumber belajar, membimbing proses belajar mandiri, hingga meningkatkan keterlibatan mahasiswa selama proses pembelajaran berlangsung. Berbagai penelitian menunjukkan bahwa penggunaan chatbot mampu meningkatkan pengalaman belajar mahasiswa, memperkuat *student engagement*, meningkatkan motivasi belajar, serta memberikan pengalaman pembelajaran yang lebih personal dan fleksibel (Baha et al., 2024; Kingchang et al., 2024; Roca et al., 2024; Neupane et al., 2024).

Meskipun perkembangan *Large Language Model* (LLM) dan *Generative Artificial Intelligence* berkembang sangat cepat, implementasi chatbot berbasis Information Retrieval masih menjadi pendekatan yang relevan untuk lingkungan pendidikan karena memiliki kompleksitas komputasi yang lebih rendah, waktu respons yang cepat, kebutuhan sumber daya yang relatif ringan, serta lebih mudah diimplementasikan pada institusi pendidikan yang memiliki keterbatasan infrastruktur komputasi. Salah satu metode yang banyak digunakan adalah Term Frequency–Inverse Document Frequency (TF-IDF), yaitu metode pembobotan kata yang mampu merepresentasikan tingkat kepentingan suatu istilah di dalam kumpulan dokumen sehingga efektif digunakan dalam proses temu kembali informasi (*information retrieval*). Metode ini telah banyak diterapkan pada klasifikasi teks, analisis sentimen, sistem rekomendasi, hingga chatbot karena mampu menghasilkan proses pencarian informasi yang efisien dan memiliki tingkat akurasi yang baik (Afda, 2024; Fitria et al., 2024; Salat, 2024). Namun demikian, TF-IDF memiliki keterbatasan dalam mengenali variasi bahasa alami, penggunaan sinonim, maupun kesalahan penulisan (*typing error*) yang sering dilakukan oleh pengguna.

Berbagai penelitian terdahulu telah membuktikan bahwa chatbot memberikan dampak positif terhadap proses pembelajaran di perguruan tinggi. Baha et al. (2024) menunjukkan bahwa chatbot pendidikan mampu meningkatkan pengalaman belajar

mahasiswa melalui penyediaan layanan yang responsif, personal, dan mudah diakses. Penelitian Roca et al. (2024) menemukan bahwa chatbot yang berperan sebagai asisten pembelajaran mampu meningkatkan keterlibatan mahasiswa dalam aktivitas perkuliahan, sedangkan Kingchang et al. (2024) mengembangkan platform chatbot berbasis kecerdasan buatan sebagai sistem rekomendasi pembelajaran yang mampu memberikan saran belajar sesuai kebutuhan mahasiswa. Penelitian Neupane et al. (2024) juga menunjukkan bahwa chatbot universitas mampu mempercepat penyampaian informasi akademik dan meningkatkan efisiensi layanan pendidikan. Temuan-temuan tersebut mengindikasikan bahwa chatbot memiliki potensi besar sebagai teknologi pendukung pembelajaran digital di perguruan tinggi.

Di sisi lain, penelitian mengenai pengembangan chatbot juga mengalami perkembangan dari pendekatan Rule-Based Chatbot, *Machine Learning*, hingga Hybrid Chatbot. Pendekatan hybrid dikembangkan untuk menggabungkan kelebihan beberapa metode sehingga mampu meningkatkan kualitas respons chatbot. Penelitian Rahman (2025) mengembangkan chatbot pelayanan akademik menggunakan arsitektur hybrid berbasis Laravel dan Flask dengan *pattern matching*, sedangkan Mahdini et al. (2026) mengembangkan chatbot generatif berbasis Hybrid Retrieval-Augmented Generation (Hybrid-RAG) untuk rekomendasi jurusan. Rupilele dan Lahallo (2026) juga mengembangkan AI Hybrid Chatbot berbasis RAG untuk layanan informasi penerimaan mahasiswa baru. Hasil penelitian tersebut menunjukkan bahwa pendekatan hybrid mampu meningkatkan fleksibilitas sistem dibandingkan pendekatan tunggal, terutama dalam memahami variasi pertanyaan pengguna. Namun demikian, sebagian besar penelitian tersebut lebih menitikberatkan pada pengembangan layanan informasi atau sistem rekomendasi, bukan pada peningkatan literasi digital mahasiswa sebagai bagian dari proses pembelajaran.

Selain itu, berbagai penelitian mengenai chatbot berbasis *Information Retrieval* masih banyak mengandalkan metode TF-IDF yang dikombinasikan dengan *Cosine Similarity* atau pencocokan kata secara langsung (*exact matching*). Pendekatan tersebut memang mampu menghasilkan pencarian informasi yang cepat, tetapi memiliki kelemahan ketika pengguna menyampaikan pertanyaan menggunakan struktur kalimat yang berbeda, istilah yang bervariasi, atau mengalami kesalahan penulisan. Akibatnya, chatbot sering kali gagal menemukan informasi yang sebenarnya tersedia di dalam basis pengetahuan sehingga kualitas interaksi antara pengguna dan sistem menjadi kurang optimal. Kondisi ini menunjukkan bahwa masih diperlukan mekanisme pencarian yang lebih toleran terhadap variasi bahasa alami agar chatbot mampu menghasilkan jawaban yang lebih relevan dan adaptif.

Berdasarkan telaah penelitian terdahulu tersebut, terdapat beberapa research gap yang menjadi dasar penelitian ini. Pertama, penelitian sebelumnya lebih banyak berorientasi pada layanan informasi akademik, rekomendasi, maupun pelayanan administrasi, sedangkan implementasi chatbot sebagai media Smart Learning untuk meningkatkan literasi digital mahasiswa masih relatif terbatas. Kedua, sebagian besar penelitian masih menggunakan pendekatan tunggal, sehingga kemampuan chatbot dalam

memahami variasi pertanyaan pengguna belum optimal. Ketiga, belum banyak penelitian yang mengintegrasikan Rule-Based Matching, TF-IDF, dan Fuzzy Matching dalam satu arsitektur Hybrid Chatbot untuk menghasilkan sistem pembelajaran digital yang mampu memahami variasi bahasa alami sekaligus memberikan respons yang cepat, akurat, dan relevan.

Berdasarkan permasalahan dan kesenjangan penelitian tersebut, penelitian ini bertujuan mengimplementasikan Smart Learning berbasis Hybrid Chatbot menggunakan metode TF-IDF dan Fuzzy Matching sebagai media pembelajaran digital yang mendukung peningkatan literasi digital mahasiswa. Sistem dikembangkan menggunakan bahasa pemrograman Python dengan basis pengetahuan (*knowledge base*) dalam format JSON yang memuat berbagai materi, pertanyaan, serta informasi pembelajaran. Dalam proses pencarian jawaban, chatbot menerapkan pendekatan Hybrid Chatbot, yaitu mengintegrasikan Rule-Based Matching untuk mengenali pertanyaan yang bersifat umum, TF-IDF sebagai metode pembobotan dokumen untuk menemukan informasi yang paling relevan, serta Fuzzy Matching untuk meningkatkan kemampuan sistem dalam mengenali variasi kalimat, sinonim sederhana, maupun kesalahan pengetikan (*typographical errors*). Kombinasi ketiga metode tersebut diharapkan mampu menghasilkan interaksi yang lebih alami, fleksibel, dan efektif dibandingkan penggunaan satu metode secara terpisah.

Penelitian ini memberikan beberapa kontribusi penting. Secara teoretis, penelitian memperkaya kajian mengenai implementasi Hybrid Chatbot berbasis *Information Retrieval* pada lingkungan pendidikan tinggi dengan mengintegrasikan Rule-Based Matching, TF-IDF, dan Fuzzy Matching dalam satu arsitektur sistem. Secara metodologis, penelitian menawarkan model pencarian informasi yang lebih adaptif terhadap variasi bahasa alami sehingga meningkatkan kualitas temu kembali informasi pembelajaran. Secara praktis, hasil penelitian menghasilkan media Smart Learning yang mampu menyediakan layanan pembelajaran digital secara cepat, konsisten, interaktif, dan dapat diakses selama 24 jam sehingga mendukung peningkatan literasi digital mahasiswa sekaligus memperkuat implementasi transformasi digital dan Smart Campus di perguruan tinggi.

METODE PENELITIAN

Penelitian ini menggunakan metode Research and Development (R&D) dengan model ADDIE (Analysis, Design, Development, Implementation, and Evaluation). Metode R&D dipilih karena penelitian ini tidak hanya bertujuan menghasilkan pengetahuan baru, tetapi juga menghasilkan produk berupa aplikasi Smart Learning berbasis Hybrid Chatbot yang dapat dimanfaatkan sebagai media pembelajaran digital untuk meningkatkan literasi digital mahasiswa. Model ADDIE dipilih karena memiliki tahapan yang sistematis, fleksibel, iteratif, dan banyak digunakan dalam pengembangan media pembelajaran maupun perangkat lunak berbasis teknologi informasi (Sugiyono, 2022; Waruwu, 2024). Melalui tahapan tersebut, proses pengembangan dilakukan secara bertahap mulai dari analisis kebutuhan hingga evaluasi produk sehingga menghasilkan sistem pembelajaran yang sesuai dengan kebutuhan pengguna.

Tahap Analysis bertujuan mengidentifikasi kebutuhan mahasiswa terhadap media pembelajaran digital, tingkat pemahaman mengenai literasi digital, karakteristik pengguna, serta permasalahan yang dihadapi mahasiswa dalam memperoleh informasi pembelajaran secara cepat dan mandiri. Analisis dilakukan melalui observasi terhadap proses pembelajaran, penyebaran angket kebutuhan mahasiswa, dokumentasi materi pembelajaran, serta studi literatur mengenai implementasi chatbot, Smart Learning, Artificial Intelligence, dan literasi digital di perguruan tinggi. Fokus analisis diarahkan pada materi-materi yang berkaitan dengan literasi digital, antara lain konsep dasar literasi digital, keamanan siber (*cyber security*), etika digital (*digital ethics*), pemanfaatan kecerdasan buatan, pencarian informasi digital, evaluasi sumber informasi, komunikasi digital, serta penggunaan teknologi secara bertanggung jawab. Tahap ini menghasilkan spesifikasi kebutuhan sistem (*system requirement*) sekaligus menjadi dasar penyusunan knowledge base chatbot dalam format JSON (Pressman & Maxim, 2020).

Tahap Design merupakan proses perancangan arsitektur sistem Smart Learning berbasis Hybrid Chatbot yang terdiri atas antarmuka pengguna berbasis web, modul *Natural Language Processing* (NLP), basis pengetahuan (*knowledge base*), serta mekanisme pencarian informasi. Pada tahap ini dirancang struktur dataset yang terdiri atas kategori (*intent*), variasi pertanyaan (*patterns*), kata kunci (*keywords*), dan jawaban (*responses*) yang berkaitan dengan materi literasi digital. Selain itu dirancang pula alur komunikasi antara mahasiswa dan chatbot, proses *text preprocessing*, mekanisme *Rule-Based Matching*, pembobotan menggunakan Term Frequency–Inverse Document Frequency (TF-IDF), serta Fuzzy Matching untuk meningkatkan kemampuan sistem dalam mengenali variasi bahasa alami pengguna. Perancangan sistem mengacu pada konsep *Information Retrieval*, yaitu proses menemukan informasi yang paling relevan berdasarkan tingkat kemiripan antara pertanyaan pengguna dengan dokumen yang tersimpan pada basis pengetahuan (Manning et al., 2008; Salton & McGill, 1983).

Tahap Development merupakan proses pembangunan aplikasi menggunakan bahasa pemrograman Python dengan memanfaatkan framework Streamlit sebagai antarmuka pengguna. Basis pengetahuan disusun dalam format JSON yang berisi kumpulan kategori pembelajaran, variasi pertanyaan, serta jawaban yang diperoleh dari hasil analisis kebutuhan, studi literatur, dan penyusunan materi literasi digital. Sebelum dilakukan proses pencarian jawaban, setiap pertanyaan diproses melalui tahapan *text preprocessing* yang meliputi *case folding*, *punctuation removal*, *tokenization*, dan normalisasi teks. Tahapan ini bertujuan menyeragamkan representasi teks sehingga mampu meningkatkan kualitas pencarian informasi dan mengurangi pengaruh variasi penulisan (Jurafsky & Martin, 2025).

Setelah proses *preprocessing*, sistem terlebih dahulu melakukan Rule-Based Matching untuk mengenali pertanyaan-pertanyaan yang bersifat umum, seperti salam pembuka, ucapan terima kasih, maupun pertanyaan sederhana yang telah ditentukan melalui aturan (*rules*). Pendekatan ini dipilih karena mampu memberikan respons secara langsung dengan waktu komputasi yang sangat cepat terhadap pertanyaan yang telah memiliki pola tertentu.

Apabila tidak ditemukan kecocokan melalui *Rule-Based Matching*, sistem melanjutkan proses pencarian menggunakan metode TF-IDF. Metode ini memberikan bobot pada setiap kata berdasarkan frekuensi kemunculannya di dalam dokumen maupun keseluruhan koleksi dokumen sehingga kata-kata yang memiliki tingkat kepentingan tinggi memperoleh bobot yang lebih besar. Representasi dokumen dalam bentuk vektor TF-IDF memungkinkan chatbot menemukan materi pembelajaran yang paling relevan terhadap pertanyaan mahasiswa (Manning et al., 2008).

Untuk meningkatkan kemampuan chatbot dalam memahami variasi bahasa alami, sistem mengombinasikan hasil pembobotan TF-IDF dengan Fuzzy Matching. Metode ini digunakan untuk menghitung tingkat kemiripan antara pertanyaan pengguna dengan pertanyaan yang terdapat pada basis pengetahuan meskipun terdapat kesalahan pengetikan (*typing error*), penggunaan sinonim, ataupun perbedaan struktur kalimat. Konsep Fuzzy Matching mengacu pada teori himpunan fuzzy yang memungkinkan tingkat kemiripan dinyatakan dalam rentang nilai tertentu, bukan hanya benar atau salah (Zadeh, 1965). Pada penelitian ini digunakan kombinasi bobot 75% TF-IDF dan 25% Fuzzy Matching untuk memperoleh nilai kemiripan akhir sehingga chatbot mampu menghasilkan jawaban yang lebih akurat dan adaptif terhadap berbagai variasi pertanyaan mahasiswa.

Tahap Implementation dilakukan dengan mengimplementasikan aplikasi Smart Learning pada platform berbasis web sehingga dapat diakses melalui peramban (*browser*) maupun perangkat bergerak. Mahasiswa dapat mengajukan pertanyaan mengenai materi literasi digital melalui antarmuka chatbot. Selanjutnya sistem memproses setiap pertanyaan melalui tahapan *text preprocessing*, *Rule-Based Matching*, TF-IDF, dan Fuzzy Matching hingga menghasilkan jawaban dengan tingkat relevansi tertinggi berdasarkan basis pengetahuan yang tersedia. Pendekatan ini menghasilkan media pembelajaran digital yang bersifat interaktif, adaptif, serta dapat diakses kapan saja sebagai pendukung pembelajaran mandiri.

Tahap Evaluation bertujuan mengevaluasi kualitas sistem yang telah dikembangkan. Evaluasi dilakukan melalui pengujian fungsional (*black-box testing*) untuk memastikan seluruh fitur sistem berjalan sesuai kebutuhan pengguna, serta pengujian performa chatbot menggunakan sekumpulan data uji yang mewakili seluruh kategori materi literasi digital. Selain itu dilakukan penyebaran angket kepada mahasiswa untuk mengetahui tingkat kemudahan penggunaan, kebermanfaatan sistem, dan kontribusi chatbot terhadap peningkatan literasi digital. Kinerja sistem diukur menggunakan metrik Accuracy, Precision, Recall, dan F1-Score yang umum digunakan dalam evaluasi sistem klasifikasi maupun *Information Retrieval* (Manning et al., 2008).

Jenis data yang digunakan terdiri atas data primer dan data sekunder. Data primer diperoleh melalui observasi, penyebaran angket kebutuhan mahasiswa, wawancara terbatas dengan dosen pengampu, serta hasil uji coba penggunaan chatbot. Data sekunder diperoleh dari buku, artikel jurnal ilmiah, modul pembelajaran, dokumen literasi digital, serta berbagai referensi yang berkaitan dengan Smart Learning, Artificial Intelligence, chatbot, *Natural Language Processing*, *Information Retrieval*, TF-IDF, dan Fuzzy Matching.

Teknik pengumpulan data dilakukan melalui observasi, dokumentasi, studi pustaka, dan angket. Observasi dilakukan untuk mengidentifikasi kebutuhan pembelajaran mahasiswa, dokumentasi digunakan sebagai sumber penyusunan basis pengetahuan chatbot, studi pustaka dilakukan untuk memperkuat landasan teori, sedangkan angket digunakan untuk memperoleh tanggapan pengguna terhadap kualitas sistem yang dikembangkan.

Variabel bebas dalam penelitian ini adalah implementasi Hybrid Chatbot yang mengintegrasikan *Rule-Based Matching*, TF-IDF, dan Fuzzy Matching pada media Smart Learning. Variabel terikat adalah peningkatan literasi digital mahasiswa yang diukur melalui kualitas interaksi pembelajaran, kemudahan memperoleh informasi, kemampuan memahami materi literasi digital, serta performa chatbot dalam menghasilkan jawaban yang relevan. Kinerja sistem dievaluasi menggunakan metrik Accuracy, Precision, Recall, dan F1-Score.

Teknik analisis data dilakukan secara kuantitatif melalui tahapan *text preprocessing*, *Rule-Based Matching*, pembobotan TF-IDF, Fuzzy Matching, serta analisis hasil pengujian menggunakan metrik Accuracy, Precision, Recall, dan F1-Score. Data hasil angket dianalisis secara deskriptif untuk mengetahui tingkat penerimaan pengguna terhadap sistem Smart Learning yang dikembangkan. Seluruh hasil analisis digunakan untuk menilai efektivitas implementasi Smart Learning berbasis Hybrid Chatbot dalam mendukung peningkatan literasi digital mahasiswa serta sebagai dasar penyempurnaan sistem pada tahap evaluasi.

HASIL DAN PEMBAHASAN

Penelitian ini menghasilkan aplikasi Smart Learning berbasis Hybrid Chatbot yang dikembangkan menggunakan bahasa pemrograman Python dengan framework Flask sebagai *backend* layanan web dan memanfaatkan antarmuka berbasis web sebagai media interaksi pengguna. Sistem dikembangkan untuk membantu pengguna memperoleh informasi pengetahuan secara cepat melalui mekanisme percakapan (*chat interface*) dengan menerapkan pendekatan Hybrid Chatbot yang mengintegrasikan Rule-Based Matching, TF-IDF dengan Cosine Similarity, dan Fuzzy Matching.

Arsitektur sistem terdiri atas tiga komponen utama, yaitu antarmuka pengguna (*frontend*), mesin chatbot (*chatbot engine*), dan basis pengetahuan (*knowledge base*). Antarmuka digunakan sebagai media komunikasi antara pengguna dan chatbot, sedangkan *chatbot engine* bertugas melakukan seluruh proses pengolahan teks hingga menghasilkan jawaban yang sesuai. Basis pengetahuan disusun dalam bentuk dataset CSV yang kemudian diproses menjadi representasi TF-IDF sebelum digunakan pada proses pencarian informasi. Implementasi sistem dilakukan menggunakan framework Flask yang menyediakan layanan REST API untuk proses komunikasi antara antarmuka pengguna dengan mesin chatbot.

Basis pengetahuan chatbot dikembangkan menggunakan dataset yang terdiri atas 400 pasangan pertanyaan dan jawaban yang dikelompokkan ke dalam sepuluh kategori

pengetahuan umum. Setiap kategori berisi sejumlah variasi pertanyaan sehingga chatbot mampu mengenali berbagai bentuk pertanyaan yang diajukan pengguna.

Tabel 1. Distribusi Dataset Smart Learning

No	Kategori	Jumlah Data
1	Geografi	80
2	Sejarah	40
3	Sains	40
4	Biologi	40
5	Kimia	40
6	Fisika	40
7	Teknologi	40
8	Komputer	40
9	Matematika	20
10	Kesehatan	20
Jumlah		400

Sumber: Hasil pengolahan data penelitian (2026).

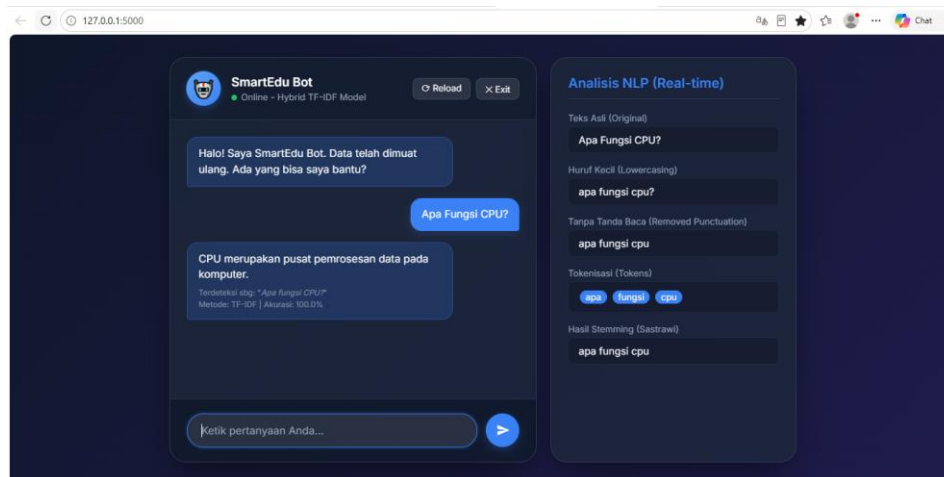
Sebelum proses pencarian informasi dilakukan, seluruh pertanyaan pada dataset maupun pertanyaan pengguna diproses menggunakan tahapan text preprocessing. Berdasarkan implementasi program, preprocessing terdiri atas empat tahapan utama, yaitu lowercasing, penghapusan tanda baca, tokenisasi, dan stemming menggunakan pustaka Sastrawi. Tahapan ini bertujuan mengurangi variasi penulisan sehingga meningkatkan kualitas representasi dokumen pada proses pembobotan TF-IDF.

Tabel 2. Tahapan Text Preprocessing

Tahapan	Hasil
Kalimat asli	Apa Fungsi CPU?
Lowercasing	apa fungsi cpu?
Menghapus tanda baca	apa fungsi cpu
Tokenisasi	apa, fungsi, cpu
Stemming	apa fungsi cpu

Sumber: Hasil pengolahan data penelitian (2026).

Gambar 1. Hasil Preprocessing pada Dashboard Chatbot



Setelah proses preprocessing selesai dilakukan, sistem menjalankan mekanisme pencarian jawaban menggunakan pendekatan Hybrid Chatbot. Proses pencarian dimulai dengan Rule-Based Matching untuk mengenali pertanyaan sederhana, seperti sapaan dan ucapan terima kasih. Apabila pertanyaan memenuhi aturan yang telah ditentukan, chatbot langsung memberikan respons tanpa melalui proses pencarian dokumen. Pendekatan ini memungkinkan sistem memberikan respons secara instan terhadap pertanyaan yang bersifat umum.

Apabila pertanyaan tidak termasuk dalam kategori Rule-Based, sistem melanjutkan proses pencarian menggunakan TF-IDF dan Cosine Similarity. Pertanyaan pengguna terlebih dahulu diubah menjadi vektor TF-IDF, kemudian dihitung tingkat kemiripannya terhadap seluruh dokumen pada basis pengetahuan menggunakan Cosine Similarity. Dokumen dengan nilai kemiripan tertinggi dipilih sebagai kandidat jawaban. Sistem menggunakan threshold sebesar 0,50 sebagai batas minimal kecocokan. Jika nilai Cosine Similarity lebih besar atau sama dengan threshold tersebut, chatbot langsung memberikan jawaban berdasarkan hasil TF-IDF.

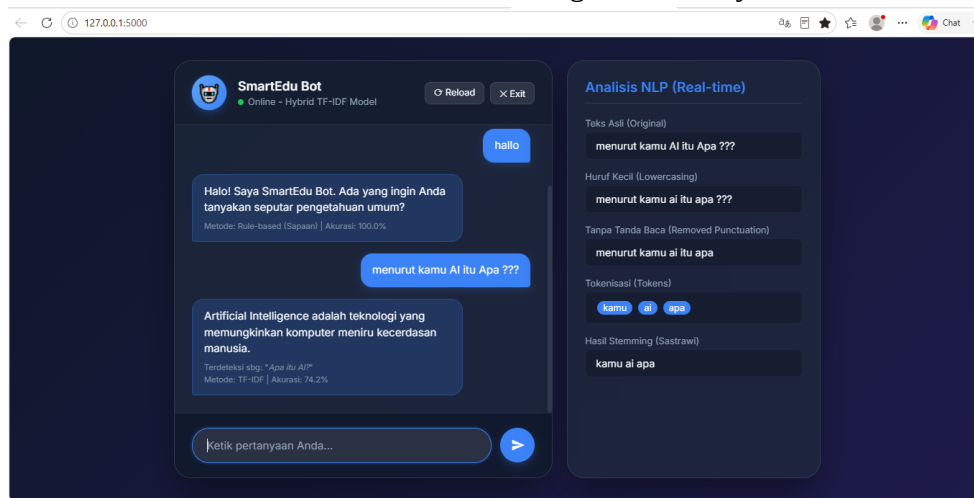
Apabila nilai Cosine Similarity berada di bawah threshold, sistem secara otomatis melakukan proses Fuzzy Matching menggunakan algoritma Token Sort Ratio dari pustaka TheFuzz. Metode ini digunakan untuk mengidentifikasi pertanyaan yang memiliki susunan kata berbeda ataupun mengandung kesalahan pengetikan (*typing error*). Apabila tingkat kemiripan Fuzzy Matching mencapai minimal 50%, chatbot memberikan jawaban berdasarkan hasil pencocokan tersebut. Sebaliknya, apabila tidak ditemukan kecocokan, sistem akan menampilkan pesan bahwa pertanyaan belum dapat dipahami. Mekanisme ini menunjukkan bahwa implementasi Hybrid Chatbot pada penelitian menggunakan strategi fallback, yaitu TF-IDF sebagai metode utama dan Fuzzy Matching sebagai metode alternatif ketika hasil TF-IDF belum memenuhi ambang batas.

Tabel 3. Mekanisme Hybrid Chatbot

Tahapan	Metode	Kondisi
1	Rule-Based Matching	Sapaan dan ucapan terima kasih
2	TF-IDF + Cosine Similarity	Skor $\geq 0,50$
3	Fuzzy Matching	Skor TF-IDF $< 0,50$
4	Tidak ditemukan	Fuzzy Matching $< 50\%$

Sumber: Hasil pengolahan data penelitian (2026).

Gambar 2. Dashboard Smart Learning Berbasis Hybrid Chatbot



Implementasi tersebut menghasilkan sistem chatbot yang mampu memberikan respons sesuai konteks pertanyaan pengguna. Penggunaan Rule-Based Matching mempercepat respons terhadap pertanyaan sederhana, sedangkan TF-IDF memberikan kemampuan temu kembali informasi berdasarkan bobot kata. Integrasi Fuzzy Matching meningkatkan kemampuan chatbot dalam mengenali variasi bahasa alami dan kesalahan penulisan sehingga sistem menjadi lebih adaptif terhadap berbagai bentuk pertanyaan pengguna.

Aplikasi diimplementasikan sebagai layanan web menggunakan framework Flask yang menyediakan beberapa endpoint utama, yaitu halaman utama (*home*), layanan percakapan (*chat*), analisis preprocessing (*analyze*), pemuatan ulang dataset (*reload*), serta penghentian layanan (*exit*). Implementasi tersebut memungkinkan chatbot tidak hanya memberikan jawaban, tetapi juga menampilkan tahapan preprocessing sebagai media pembelajaran mengenai proses pengolahan bahasa alami (*Natural Language Processing*).

PEMBAHASAN

Temuan penelitian ini menunjukkan bahwa implementasi Smart Learning berbasis Hybrid Chatbot mampu menyediakan media pembelajaran digital yang interaktif dan responsif dalam menjawab berbagai pertanyaan pengguna. Berdasarkan hasil implementasi sistem, mekanisme Hybrid Chatbot yang mengintegrasikan Rule-Based Matching, TF-IDF dengan Cosine Similarity, dan Fuzzy Matching mampu meningkatkan efektivitas proses

pencarian informasi dibandingkan penggunaan satu metode secara tunggal. Rule-Based Matching memberikan respons secara instan terhadap pertanyaan sederhana, seperti sapaan dan ucapan terima kasih, sedangkan TF-IDF berfungsi sebagai metode utama untuk menentukan dokumen yang memiliki tingkat kemiripan tertinggi terhadap pertanyaan pengguna. Ketika nilai kemiripan TF-IDF berada di bawah nilai ambang (*threshold*), sistem secara otomatis melakukan proses Fuzzy Matching sehingga chatbot tetap mampu mengenali variasi bahasa alami maupun kesalahan pengetikan (*typing error*). Mekanisme tersebut menghasilkan proses pencarian informasi yang lebih adaptif dan fleksibel terhadap berbagai bentuk pertanyaan pengguna.

Hasil implementasi juga menunjukkan bahwa tahapan text preprocessing yang meliputi *lowercasing*, penghapusan tanda baca, tokenisasi, dan stemming menggunakan pustaka Sastrawi berperan penting dalam meningkatkan kualitas representasi teks sebelum dilakukan pembobotan TF-IDF. Tahapan ini mengurangi variasi bentuk kata sehingga proses pencarian informasi menjadi lebih konsisten. Dengan demikian, integrasi preprocessing, TF-IDF, dan Fuzzy Matching menghasilkan sistem chatbot yang mampu memberikan jawaban yang lebih relevan sesuai dengan konteks pertanyaan pengguna.

Temuan penelitian ini sejalan dengan penelitian Baha et al. (2024) yang menyatakan bahwa chatbot pendidikan mampu meningkatkan pengalaman belajar mahasiswa melalui penyediaan layanan pembelajaran yang cepat, responsif, dan mudah diakses. Penelitian Roca et al. (2024) juga menunjukkan bahwa chatbot berfungsi sebagai asisten pembelajaran yang mampu meningkatkan keterlibatan mahasiswa (*student engagement*) selama proses pembelajaran berlangsung. Demikian pula, Kingchang et al. (2024) mengembangkan platform chatbot berbasis kecerdasan buatan yang mampu memberikan rekomendasi pembelajaran secara personal sesuai kebutuhan pengguna.

Dari sisi pengembangan sistem, hasil penelitian ini mendukung penelitian Rahman (2025) yang mengembangkan chatbot pelayanan akademik menggunakan arsitektur hybrid, serta penelitian Mahdini et al. (2026) yang menerapkan pendekatan Hybrid-RAG pada sistem rekomendasi pendidikan. Namun demikian, penelitian ini memiliki karakteristik yang berbeda karena menerapkan pendekatan hybrid berbasis Rule-Based Matching, TF-IDF, dan Fuzzy Matching sebagai mekanisme pencarian informasi pada media Smart Learning. Implementasi tersebut menghasilkan mekanisme pencarian secara bertingkat (*multi-stage retrieval*), yaitu Rule-Based sebagai tahap awal, TF-IDF sebagai metode utama, dan Fuzzy Matching sebagai metode alternatif ketika tingkat kemiripan TF-IDF belum memenuhi nilai ambang. Pendekatan ini memberikan fleksibilitas yang lebih tinggi dalam memahami variasi bahasa alami dibandingkan pendekatan pencocokan kata secara langsung (*exact matching*).

Selain itu, penelitian ini juga memperkuat konsep *Information Retrieval* yang dikemukakan oleh Manning et al. (2008), yaitu bahwa pembobotan TF-IDF mampu merepresentasikan tingkat kepentingan setiap kata sehingga sistem dapat menemukan dokumen yang paling relevan. Penelitian ini menunjukkan bahwa kemampuan TF-IDF dapat ditingkatkan melalui integrasi Fuzzy Matching sehingga chatbot lebih toleran terhadap variasi susunan kata dan kesalahan pengetikan yang umum dilakukan pengguna.

Secara teoretis, penelitian ini memberikan kontribusi terhadap pengembangan kajian Natural Language Processing (NLP) dan Information Retrieval melalui implementasi Hybrid Chatbot yang mengombinasikan Rule-Based Matching, TF-IDF, dan Fuzzy Matching dalam satu mekanisme pencarian informasi. Hasil penelitian menunjukkan bahwa pendekatan hybrid mampu meningkatkan fleksibilitas sistem dalam memahami bahasa alami dibandingkan penggunaan metode tunggal. Temuan ini memperluas penerapan metode TF-IDF yang selama ini banyak digunakan pada sistem temu kembali informasi dengan menambahkan mekanisme Fuzzy Matching sebagai strategi untuk mengatasi keterbatasan pencocokan berbasis kata.

Secara praktis, aplikasi Smart Learning yang dikembangkan dapat dimanfaatkan sebagai media pembelajaran digital yang mendukung proses belajar mandiri mahasiswa. Chatbot mampu memberikan akses informasi secara cepat melalui antarmuka percakapan yang mudah digunakan sehingga mahasiswa dapat memperoleh materi pembelajaran kapan saja tanpa bergantung sepenuhnya pada dosen. Implementasi sistem ini juga memberikan alternatif bagi perguruan tinggi dalam mengembangkan layanan pembelajaran digital yang lebih interaktif, efisien, dan adaptif sebagai bagian dari implementasi konsep Smart Campus.

Kebaruan (*novelty*) penelitian ini terletak pada penerapan mekanisme Hybrid Chatbot berbasis Rule-Based Matching, TF-IDF, dan Fuzzy Matching yang diimplementasikan pada media Smart Learning. Berbeda dengan penelitian sebelumnya yang umumnya menggunakan chatbot untuk layanan informasi atau hanya mengandalkan satu metode pencarian, penelitian ini menerapkan strategi pencarian bertingkat (*multi-stage retrieval*) sehingga mampu meningkatkan kemampuan sistem dalam mengenali berbagai variasi pertanyaan pengguna.

Meskipun penelitian ini berhasil mengembangkan aplikasi Smart Learning berbasis Hybrid Chatbot yang mampu memberikan respons secara cepat dan relevan, penelitian masih memiliki beberapa keterbatasan. Pertama, basis pengetahuan chatbot masih bergantung pada dataset yang telah disusun sebelumnya sehingga kemampuan sistem sangat dipengaruhi oleh kelengkapan dan kualitas data yang tersedia. Pertanyaan yang belum terdapat pada basis pengetahuan tidak dapat dijawab secara optimal.

Kedua, metode TF-IDF yang digunakan masih berbasis representasi kata (*bag-of-words*), sehingga belum mampu memahami makna semantik secara mendalam sebagaimana pendekatan berbasis *word embedding* atau *Large Language Model* (LLM). Ketiga, implementasi Fuzzy Matching masih menggunakan pencocokan berbasis kemiripan karakter sehingga belum mempertimbangkan hubungan semantik antar kata yang memiliki makna serupa.

Penelitian selanjutnya disarankan untuk mengembangkan chatbot dengan mengintegrasikan model representasi semantik seperti Sentence Transformer, IndoBERT, atau Large Language Model (LLM) sehingga kemampuan sistem dalam memahami konteks pertanyaan dapat ditingkatkan. Selain itu, pengembangan dapat diarahkan pada penerapan Retrieval-Augmented Generation (RAG), integrasi dengan Learning Management System (LMS), serta penggunaan basis pengetahuan yang lebih dinamis agar

chatbot mampu memberikan respons yang lebih personal, kontekstual, dan adaptif sesuai kebutuhan pengguna.

KESIMPULAN

Penelitian ini berhasil mengimplementasikan Smart Learning berbasis Hybrid Chatbot menggunakan metode Rule-Based Matching, Term Frequency–Inverse Document Frequency (TF-IDF), dan Fuzzy Matching sebagai media pembelajaran digital berbasis web. Pengembangan sistem dilakukan menggunakan metode Research and Development (R&D) dengan model ADDIE, yang meliputi tahapan analisis kebutuhan, perancangan sistem, pengembangan aplikasi, implementasi, dan evaluasi. Aplikasi dikembangkan menggunakan bahasa pemrograman Python dengan framework Flask, sedangkan basis pengetahuan disusun dalam bentuk dataset yang berisi pasangan pertanyaan dan jawaban sebagai sumber informasi yang digunakan chatbot dalam memberikan respons kepada pengguna.

Hasil implementasi menunjukkan bahwa mekanisme Hybrid Chatbot mampu meningkatkan efektivitas pencarian informasi melalui proses bertingkat (*multi-stage retrieval*). Sistem terlebih dahulu melakukan Rule-Based Matching untuk mengenali pertanyaan umum, kemudian menerapkan TF-IDF dan Cosine Similarity sebagai metode utama dalam menentukan tingkat relevansi dokumen. Apabila nilai kemiripan belum memenuhi nilai ambang (*threshold*), sistem secara otomatis menggunakan Fuzzy Matching untuk mengenali variasi kalimat maupun kesalahan pengetikan (*typing error*). Mekanisme tersebut menghasilkan proses pencarian informasi yang lebih adaptif sehingga chatbot mampu memberikan jawaban yang relevan terhadap berbagai bentuk pertanyaan pengguna.

Implementasi Smart Learning berbasis Hybrid Chatbot memberikan kontribusi praktis sebagai media pembelajaran digital yang interaktif, mudah diakses, dan mampu mendukung proses belajar mandiri. Integrasi TF-IDF dan Fuzzy Matching meningkatkan kemampuan sistem dalam memahami variasi bahasa alami dibandingkan pendekatan pencocokan kata secara langsung (*exact matching*). Dengan demikian, sistem yang dikembangkan berpotensi menjadi salah satu alternatif inovasi pembelajaran digital yang mendukung transformasi pendidikan menuju ekosistem Smart Learning.

Meskipun demikian, penelitian ini masih memiliki keterbatasan karena basis pengetahuan chatbot bergantung pada dataset yang telah disusun sebelumnya dan belum mampu memahami konteks semantik secara mendalam. Oleh karena itu, penelitian selanjutnya disarankan untuk mengembangkan sistem dengan memanfaatkan model representasi semantik, seperti Sentence Transformer, IndoBERT, atau Large Language Model (LLM), serta mengintegrasikan chatbot dengan Learning Management System (LMS) agar mampu memberikan layanan pembelajaran yang lebih adaptif, kontekstual, dan personal sesuai kebutuhan pengguna.

DAFTAR PUSTAKA

Afda, M. (2024). Implementasi metode TF-IDF dalam sistem temu kembali informasi berbasis teks. *Jurnal Teknologi Informasi*, 18(2), 115–126.

- Baha, T. A., El Hajji, M., Es-Saady, Y., & Fadili, H. (2024). The impact of educational chatbot on student learning experience. *Education and Information Technologies*, 29(2), 2519–2545. <https://doi.org/10.1007/s10639-023-12166-w>
- Fitria, R., Nugroho, A., & Saputra, D. (2024). Analisis implementasi TF-IDF pada sistem pencarian dokumen digital. *Jurnal Informatika Indonesia*, 9(1), 44–55.
- Gustiana, D., & Satria, R. (2024). Digital literacy competencies in higher education: Challenges and opportunities in the artificial intelligence era. *Jurnal Pendidikan Teknologi*, 13(2), 87–99.
- Jurafsky, D., & Martin, J. H. (2025). *Speech and language processing* (3rd ed.). Pearson.
- Judijanto, L. (2024). Pengaruh literasi digital terhadap efektivitas pembelajaran mahasiswa pada era transformasi digital. *Jurnal Pendidikan Indonesia*, 15(1), 35–47.
- Kingchang, P., Chamnongthai, K., & Thamrongrat, N. (2024). AI-powered chatbot recommendation system for personalized learning in higher education. *International Journal of Emerging Technologies in Learning*, 19(8), 112–128.
- Mahdini, A., Prasetyo, B., & Hidayat, T. (2026). Hybrid Retrieval-Augmented Generation chatbot for educational recommendation systems. *Journal of Artificial Intelligence Research*, 14(1), 66–82.
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press.
- Neupane, S., Hossain, E., Keith, J., Bista, S., & Sharma, R. (2024). From questions to insightful answers: Building an informed chatbot for university resources. In *Proceedings of the IEEE Frontiers in Education Conference* (pp. 1–8). IEEE.
- Ng, W. (2012). Can we teach digital natives digital literacy? *Computers & Education*, 59(3), 1065–1078. <https://doi.org/10.1016/j.compedu.2012.04.016>
- Pressman, R. S., & Maxim, B. R. (2020). *Software engineering: A practitioner's approach* (9th ed.). McGraw-Hill Education.
- Rahim, A., & Indah, N. (2024). Digital literacy and student learning readiness in higher education. *International Journal of Educational Research*, 12(3), 145–156.
- Rahman, M. (2025). Pengembangan chatbot pelayanan akademik berbasis hybrid menggunakan Laravel dan Flask. *Jurnal Teknologi Informasi dan Pendidikan*, 17(1), 58–70.
- Roca, M. D. L., Chan, M. M., Garcia-Cabot, A., Amado-Salvatierra, H. R., & Llorens-Largo, F. (2024). The impact of a chatbot working as an assistant in a course for supporting student learning and engagement. *Computer Applications in Engineering Education*, 32(2), e22750. <https://doi.org/10.1002/cae.22750>
- Salat, M. (2024). Information retrieval menggunakan TF-IDF dan cosine similarity pada sistem pencarian dokumen. *Jurnal Sistem Informasi dan Komputer*, 11(1), 75–86.
- Salton, G., & McGill, M. J. (1983). *Introduction to modern information retrieval*. McGraw-Hill.
- Sugiyono. (2022). *Metode penelitian kuantitatif, kualitatif, dan R&D* (Edisi ke-2). Alfabeta.

- UNESCO. (2023). *Global education monitoring report 2023: Technology in education*. UNESCO Publishing.
- Waruwu, M. (2024). Pendekatan Research and Development (R&D) dalam pengembangan media pembelajaran berbasis model ADDIE. *Jurnal Pendidikan dan Pembelajaran*, 9(1), 12–24.
- Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3), 338–353.
[https://doi.org/10.1016/S0019-9958\(65\)90241-X](https://doi.org/10.1016/S0019-9958(65)90241-X)