

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Perkembangan teknologi informasi dan kecerdasan buatan telah mendorong berbagai institusi, termasuk institusi pendidikan, untuk memanfaatkan sistem berbasis digital dalam meningkatkan kualitas layanan kepada pengguna. Salah satu teknologi yang berkembang pesat dan banyak diadopsi dalam layanan informasi adalah *chatbot*, yaitu sistem perangkat lunak yang mampu mensimulasikan percakapan dengan pengguna secara otomatis melalui antarmuka teks (Jurafsky & Martin, 2023). *Chatbot* telah terbukti efektif dalam berbagai *domain* layanan, mulai dari layanan pelanggan bisnis, layanan kesehatan, hingga layanan informasi di lingkungan pendidikan tinggi. Penelitian oleh (Rizki et al., 2023) membuktikan bahwa penerapan *chatbot* pada layanan penerimaan mahasiswa baru mampu menjawab pertanyaan calon mahasiswa secara otomatis dan meringankan beban staf dalam melayani pertanyaan berulang.

Dalam konteks pendidikan tinggi, salah satu layanan yang memiliki tingkat permintaan informasi paling tinggi dari masyarakat adalah layanan Penerimaan Mahasiswa Baru (PMB) (Rizki et al., 2023). Calon mahasiswa dan orang tua kerap membutuhkan informasi yang cepat, akurat, dan dapat diakses kapan saja mengenai program studi yang tersedia, biaya kuliah, jalur pendaftaran, persyaratan, akreditasi, serta jadwal dan prosedur pendaftaran. Kebutuhan ini tidak selalu dapat dipenuhi

secara optimal apabila layanan informasi masih bergantung pada respons manual dari staf tanpa dukungan sistem otomatis (Nuzul Hikmah et al., 2022).

STMIK AMIKOM Surakarta merupakan perguruan tinggi berbasis teknologi informasi yang berlokasi di Sukoharjo, Jawa Tengah. Dalam mengelola layanan informasi Penerimaan Mahasiswa Baru (PMB), STMIK AMIKOM Surakarta telah menggunakan dua kanal utama, yaitu *website* resmi berbasis *WordPress* yang aktif digunakan sebagai media informasi publik, serta layanan *WhatsApp* yang dikelola secara manual oleh staf Humas PMB. *Website WordPress* kampus menjadi pintu pertama bagi calon mahasiswa untuk mencari informasi seputar PMB, sementara *WhatsApp* digunakan sebagai saluran komunikasi langsung antara calon mahasiswa dan staf PMB untuk pertanyaan yang lebih spesifik.

Berdasarkan hasil observasi dan wawancara yang dilakukan dengan staf Humas PMB STMIK AMIKOM Surakarta, ditemukan bahwa proses layanan informasi PMB saat ini sepenuhnya dilakukan secara manual. Staf PMB menerima dan menjawab pertanyaan dari calon mahasiswa satu per satu melalui *WhatsApp*, baik pada jam kerja maupun di luar jam operasional. Tidak terdapat sistem otomatis yang dapat merespons pertanyaan calon mahasiswa secara langsung ketika staf tidak sedang bertugas.

Permasalahan utama yang teridentifikasi dari hasil wawancara adalah tingginya volume pertanyaan yang diterima staf PMB setiap harinya. Berdasarkan keterangan staf Humas PMB, pada masa aktif penerimaan mahasiswa baru periode Januari hingga Juni 2026, staf menerima rata-rata 30 hingga 50 pesan per hari

melalui *WhatsApp*. Sebagian besar dari pesan tersebut berisi pertanyaan berulang seputar informasi standar seperti biaya kuliah, program studi yang tersedia, jalur pendaftaran, persyaratan dokumen, dan jadwal gelombang penerimaan informasi yang sebenarnya telah tersedia di *website WordPress* kampus. Selain itu, hasil wawancara yang dilakukan terhadap sejumlah mahasiswa baru semester 2 menunjukkan bahwa sebagian dari mereka merasa respons staf PMB melalui *WhatsApp* tidak selalu cepat, terutama di luar jam kerja dan pada akhir pekan, sehingga mereka merasa kurang puas dengan layanan yang diberikan. Kondisi ini menyebabkan staf PMB harus memberikan jawaban yang sama secara berulang kepada calon mahasiswa yang berbeda, sehingga mengurangi efisiensi kerja dan meningkatkan beban operasional, terutama pada periode puncak penerimaan mahasiswa baru.

Selain itu, berdasarkan hasil observasi langsung terhadap *website WordPress* STMIK AMIKOM Surakarta, ditemukan bahwa *website* tersebut belum dilengkapi dengan fitur layanan informasi interaktif yang memungkinkan calon mahasiswa memperoleh jawaban secara langsung dan otomatis. Calon mahasiswa yang mengunjungi *website* pada malam hari atau akhir pekan tidak dapat memperoleh respons segera karena tidak ada sistem yang beroperasi di luar jam kerja staf PMB. Perlu ditegaskan bahwa proses komunikasi saat ini dilakukan melalui *WhatsApp* secara manual oleh staf, sedangkan sistem *chatbot* yang akan dikembangkan dalam penelitian ini diimplementasikan langsung pada *website WordPress* kampus sehingga calon mahasiswa dapat memperoleh jawaban

otomatis tanpa harus beralih ke aplikasi lain. Kondisi ini berpotensi menurunkan tingkat kepuasan calon mahasiswa dan mengurangi peluang konversi pendaftar.

Berdasarkan permasalahan tersebut, solusi yang relevan adalah mengembangkan sistem FAQ *Retrieval Chatbot* berbasis *WordPress* yang mampu menjawab pertanyaan calon mahasiswa secara otomatis selama 24 jam berdasarkan basis pengetahuan *Frequently asked questions* (FAQ) PMB yang telah disiapkan. *Chatbot* ini akan diintegrasikan langsung ke dalam *website WordPress* yang sudah aktif digunakan kampus, sehingga calon mahasiswa dapat mengakses informasi PMB secara otomatis tanpa perlu meninggalkan halaman *website*. Pendekatan serupa telah terbukti berhasil diterapkan di berbagai perguruan tinggi; penelitian oleh (Fathan Fauzan et al., 2024) membuktikan bahwa *chatbot* berbasis *website* mampu meningkatkan efektivitas pelayanan informasi PMB dan secara signifikan mengurangi beban kerja admin dalam menjawab pertanyaan yang sama secara berulang.

Pendekatan FAQ *Retrieval Chatbot* dipilih karena lebih konsisten, mudah divalidasi kontennya oleh staf PMB, dan tidak memerlukan infrastruktur *server* yang besar. Berbeda dengan pendekatan *deep learning* seperti FastText dan LSTM yang membutuhkan sumber daya komputasi besar dan proses pelatihan model yang kompleks (Yusron & Komarudin, 2024), pendekatan berbasis *Information retrieval* terbukti lebih ringan dan sesuai untuk implementasi pada lingkungan *website* berbasis *shared hosting* seperti yang digunakan STMIK AMIKOM Surakarta. Untuk menentukan jawaban yang paling relevan dari basis data FAQ, sistem menggunakan metode TF-IDF (*Term Frequency-Inverse Document Frequency*)

untuk merepresentasikan teks pertanyaan dan dokumen FAQ sebagai *vektor numerik* berbobot, kemudian *Cosine similarity* digunakan untuk mengukur tingkat kemiripan antara *vektor* pertanyaan pengguna dengan setiap *vektor* dokumen FAQ yang tersedia. Penelitian oleh (Setiawan et al., 2023) membuktikan bahwa kombinasi TF-IDF dan *Cosine similarity* mampu meningkatkan relevansi jawaban *chatbot* berbasis *retrieval* secara signifikan dengan tingkat komputasi yang ringan. Penelitian oleh (Nuzul Hikmah et al., 2022) juga membuktikan bahwa metode TF-IDF efektif diterapkan pada *chatbot* layanan informasi di lingkungan perguruan tinggi.

Meskipun demikian, pendekatan FAQ *Retrieval* berbasis TF-IDF dan *Cosine similarity* memiliki keterbatasan bawaan yang perlu diatasi. Metode ini bersifat kaku karena hanya mampu mencocokkan kata-kata yang secara leksikal sama atau mirip, sehingga rentan terhadap variasi penyebutan pertanyaan oleh pengguna yang berbeda. Selain itu, pada kasus pertanyaan standar yang sifatnya tetap dan terprediksi, proses kalkulasi *vektor* secara penuh sebenarnya tidak diperlukan dan menghasilkan *overhead* komputasi yang tidak efisien. Tantangan lain adalah pengguna sering melakukan kesalahan pengetikan (*typo*) saat mengajukan pertanyaan, yang dapat menyebabkan sistem gagal menemukan kecocokan meskipun maksud pertanyaan sebenarnya sudah terdapat dalam basis data FAQ.

Untuk mengatasi keterbatasan tersebut, sistem dikembangkan menggunakan mekanisme pemrosesan bertahap yang menggabungkan *Rule-based*, TF-IDF dan *Cosine similarity*, *Fuzzy Matching*, serta mekanisme *Escalation*. *Rule-*

based digunakan pada tahap awal untuk menangani *intent* tertentu secara deterministik. Apabila tidak ditemukan kecocokan *Rule-based*, pertanyaan diproses menggunakan TF-IDF dan *Cosine similarity* untuk menentukan FAQ dengan tingkat kemiripan tertinggi. Pertanyaan dengan tingkat kemiripan yang rendah selanjutnya dapat diproses menggunakan *Fuzzy Matching* sebagai mekanisme *fallback* untuk menangani variasi atau kesalahan pengetikan. Apabila sistem tetap tidak memperoleh kecocokan yang memadai, pertanyaan diarahkan pada mekanisme eskalasi kepada pihak PMB.

Lapis kedua menggunakan TF-IDF dan *Cosine similarity* untuk menangani pertanyaan yang diketikkan secara bebas oleh pengguna dalam bentuk teks. Penggunaan arsitektur *Hybrid* seperti ini sejalan dengan temuan (Rizki et al., 2023) yang menunjukkan bahwa kombinasi pendekatan berbasis aturan dan pemrosesan teks mampu menghasilkan *chatbot* PMB yang lebih andal dan responsif. Selain itu, untuk meningkatkan ketahanan sistem terhadap kesalahan penulisan, proses *PreProcessing* dilengkapi dengan metode *Fuzzy Matching* agar variasi ejaan tetap dapat dikenali dengan benar.

Justifikasi pemilihan metode ini didasarkan pada kemampuannya mengukur kemiripan string secara karakter per karakter sehingga variasi ejaan dan kesalahan pengetikan ringan tetap dapat dikenali oleh sistem dengan tepat, tanpa memerlukan kecocokan leksikal yang sempurna (Cohen et al., 2003) . Selain itu, bagi calon mahasiswa yang terbiasa bertanya melalui *WhatsApp*, sistem ini menyediakan tombol tindak lanjut (*follow-up*) ke *WhatsApp* staf Humas PMB di bagian akhir

percakapan *chatbot*, sehingga mereka tetap dapat diarahkan ke kanal komunikasi yang sudah dikenal apabila pertanyaan tidak terjawab oleh sistem otomatis.

Berdasarkan uraian permasalahan di atas, penelitian ini bertujuan untuk mengimplementasikan sistem FAQ *Retrieval Chatbot* layanan informasi PMB dengan menggunakan metode TF-IDF dan *Cosine similarity* yang diintegrasikan langsung ke dalam *website WordPress* STMIK AMIKOM Surakarta melalui metode pengembangan *Rapid Application Development* (RAD). Sistem ini dirancang untuk membantu staf Humas PMB STMIK AMIKOM Surakarta dalam mengurangi beban menjawab 30 hingga 50 pertanyaan berulang setiap harinya, sekaligus membantu calon mahasiswa memperoleh informasi PMB secara cepat, akurat, dan tersedia selama 24 jam melalui *website* kampus. Dengan tersedianya sistem *chatbot* otomatis pada *website WordPress* yang sudah aktif digunakan, diharapkan pengalaman calon mahasiswa dalam mengakses informasi PMB menjadi lebih baik, serta efisiensi kerja staf PMB meningkat secara signifikan.

1.2. Rumusan Masalah

Berdasarkan latar belakang dan permasalahan yang telah diuraikan, rumusan masalah dalam penelitian ini disusun untuk menjadi dasar dalam perancangan, implementasi, pengujian, serta evaluasi sistem FAQ *Retrieval Chatbot* Penerimaan Mahasiswa Baru di STMIK AMIKOM Surakarta. Rumusan masalah dalam penelitian ini adalah sebagai berikut:

- a. Bagaimana merancang dan mengimplementasikan sistem FAQ *Retrieval Chatbot* layanan informasi Penerimaan Mahasiswa Baru menggunakan metode

TF-IDF dan *Cosine similarity* yang terintegrasi pada *website WordPress* STMIK AMIKOM Surakarta?

- b. Bagaimana menerapkan mekanisme pemrosesan bertahap yang menggabungkan *Rule-based*, TF-IDF dan *Cosine similarity*, *Fuzzy Matching*, serta mekanisme eskalasi untuk menangani pertanyaan standar, pertanyaan teks bebas, dan kesalahan pengetikan pengguna?
- c. Bagaimana tingkat akurasi jawaban sistem FAQ *Retrieval Chatbot* yang dikembangkan berdasarkan hasil pengujian terhadap variasi pertanyaan layanan informasi PMB STMIK AMIKOM Surakarta?
- d. Bagaimana tingkat kelayakan dan kepuasan pengguna terhadap sistem FAQ *Retrieval Chatbot* PMB yang dikembangkan berdasarkan hasil uji pengguna menggunakan instrumen *USE Questionnaire* dengan skala Likert?

1.3. Batasan Masalah

Agar penelitian lebih terarah dan pembahasan tidak meluas di luar tujuan yang telah ditetapkan, diperlukan batasan masalah yang menentukan ruang lingkup penelitian dan sistem yang dikembangkan. Batasan masalah dalam penelitian ini adalah sebagai berikut:

- a. Sistem *chatbot* dikembangkan berbasis *website WordPress* yang sudah aktif digunakan oleh STMIK AMIKOM Surakarta sebagai media informasi PMB.
- b. Sistem menggunakan mekanisme pemrosesan bertahap berupa *Rule-based* → TF-IDF dan *Cosine similarity* → *Fuzzy Matching* → *Escalation*.

- c. Data FAQ yang digunakan mencakup informasi PMB STMIK AMIKOM Surakarta meliputi biaya kuliah, program studi yang tersedia, jalur dan persyaratan pendaftaran, akreditasi, prospek kerja lulusan, lokasi kampus, serta jadwal gelombang penerimaan.
- d. Bahasa yang digunakan dalam sistem *chatbot* adalah Bahasa Indonesia dan Bahasa Inggris.
- e. Sistem dilengkapi dengan fitur pengumpulan identitas calon mahasiswa (*lead collection*) berupa nama, nomor *WhatsApp*, dan program studi yang diminati, sebagai data pendukung proses tindak lanjut (*follow-up*) secara manual oleh staf Humas PMB.
- f. Sistem dilengkapi fitur eskalasi ke admin apabila pertanyaan pengguna tidak ditemukan jawabannya dalam basis data FAQ.
- g. Pengguna sistem *chatbot* ini adalah calon mahasiswa baru STMIK AMIKOM Surakarta yang mengakses informasi PMB melalui *website* kampus, serta staf Humas PMB selaku pengelola sistem dan penerima data *lead collection*.

1.4. Tujuan Penelitian

Berdasarkan rumusan masalah yang telah ditetapkan, penelitian ini memiliki tujuan yang berkaitan dengan perancangan, penerapan metode, pengujian performa, serta evaluasi sistem FAQ *Retrieval Chatbot* PMB. Tujuan penelitian ini adalah sebagai berikut:

- a. Merancang dan mengimplementasikan sistem FAQ *Retrieval Chatbot* layanan informasi Penerimaan Mahasiswa Baru menggunakan metode TF-IDF dan

Cosine similarity yang terintegrasi pada *website WordPress* STMIK AMIKOM Surakarta, sehingga calon mahasiswa dapat memperoleh informasi PMB secara otomatis selama 24 jam.

- b. Menerapkan mekanisme pemrosesan bertahap yang menggabungkan *Rule-based*, TF-IDF dan *Cosine similarity*, *Fuzzy Matching*, serta *Escalation* untuk meningkatkan kemampuan sistem dalam menangani pertanyaan standar, pertanyaan bebas, variasi penulisan, dan kesalahan pengetikan pengguna.
- c. Mengukur tingkat akurasi jawaban sistem FAQ *Retrieval Chatbot* yang dikembangkan berdasarkan hasil pengujian terhadap variasi pertanyaan layanan informasi PMB STMIK AMIKOM Surakarta, guna memastikan sistem mampu memberikan jawaban yang relevan dan tepat kepada calon mahasiswa.
- d. Mengukur tingkat kelayakan dan kepuasan pengguna terhadap sistem FAQ *Retrieval Chatbot* PMB yang dikembangkan menggunakan instrumen *USE Questionnaire* dengan skala Likert, guna memastikan sistem bermanfaat, mudah digunakan, mudah dipelajari, dan memberikan kepuasan kepada pengguna.

1.5. Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat baik secara teoritis maupun praktis bagi pengembangan ilmu pengetahuan, institusi, pengelola layanan PMB, serta calon mahasiswa. Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

Manfaat Teoritis:

Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan ilmu pengetahuan pada bidang *Information retrieval*, *Natural language Processing*, dan pengembangan *chatbot* berbasis *website*.

Manfaat Praktis:

- a. Bagi STMIK AMIKOM Surakarta
Membantu meningkatkan kualitas layanan informasi PMB melalui sistem *chatbot* otomatis.
- b. Bagi Humas PMB
Mengurangi beban pelayanan pertanyaan berulang serta membantu pengumpulan data calon mahasiswa.
- c. Bagi Calon Mahasiswa
Mempermudah akses informasi PMB secara cepat dan real-time.