JURNAL TEKNIK INFORMATIKA DAN MULTIMEDIA Vol 6, No.1, Mei 2026, pp. 348-366
p-ISSN : 2808-909X (print)
e-ISSN : 2798-2505 (online)
http://journal.politeknik-pratama.ac.id/index.php/JTIM · page 348
DOI : https://doi.org/10.51903/informatika.v6i1.1645

PENERAPAN *NATURAL LANGUAGE PROCESSING* ALGORITME *SBERT* PADA APLIKASI *CHATBOT* PENERIMAAN MAHASISWA BARU UNIVERSITAS PAMULANG

Aji Sakti Ardiansyah[1*],Kahfi Heryandi Suradiradja[2]
[1] Universitas Pamulang
addres, telp/fax of institution/affiliation, e-mail: aji.sakty12323@gmail.com
[2] Universitas Pamulang
addres, telp/fax of institution/affiliation, e-mail: dosen01514@unpam.ac.id

ARTICLE INFO

Article history:

Received 10 December 2026
Received in revised form 24 January 2026
Accepted 8 Mei 2026
Available online 15 Mei 2026

ABSTRACT

The New Student Admission (PMB) process at Pamulang University currently relies on one-way information channels such as Instagram and brochures only at designated posts. This condition limits access to information for prospective students, especially when they require quick and detailed answers regarding registration requirements, schedules, and tuition fees. Furthermore, the large number of applicants in each admission period makes manual inquiry services inefficient and prone to delays. Therefore, a technology-based solution must provide interactive, fast, and accurate information services. This study aims to develop a Chatbot based on Natural Language Processing (NLP) using the Sentence-BERT (SBERT) algorithm to support the PMB process at the Informatics Engineering Program at Pamulang University. The Chatbot is designed to provide accurate and relevant automated responses to prospective students' inquiries regarding PMB. The SBERT algorithm enhances the Chatbot's ability to understand sentence meanings and deliver contextual responses. The software development adopts the Waterfall model, encompassing requirements analysis, system design, implementation, and testing phases. The expected outcome is a Chatbot system with-high response accuracy that improves

Received December 10, 2025; Accepted Mei 8, 2026; Available online Mei 15, 2026

p-ISSN : 2808-909X e-ISSN : 2798-2505

efficiency and accessibility of PMB information services. This research is expected to contribute to the implementation of NLP technology in academic environments.

Keywords: Chatbot, Informatics Engineering, Natural Language Processing, New Student Admission, Sentence-BERT.

Abstrak

Proses Penerimaan Mahasiswa Baru (PMB) di Universitas Pamulang masih mengandalkan media informasi satu arah, seperti Instagram dan brosur di pos-pos tertentu. Kondisi ini menimbulkan keterbatasan akses informasi bagi calon mahasiswa, terutama ketika mereka membutuhkan jawaban cepat dan detail mengenai persyaratan pendaftaran, jadwal, maupun biaya kuliah. Selain itu, banyaknya jumlah pendaftar di setiap periode penerimaan membuat layanan informasi manual menjadi tidak efisien dan rentan terhadap keterlambatan penyampaian informasi. Oleh karena itu, diperlukan solusi berbasis teknologi yang mampu memberikan layanan informasi interaktif, cepat, dan akurat. Penelitian ini bertujuan untuk mengembangkan *Chatbot* aplikasi berbasis *Natural Language Processing* (NLP) menggunakan algoritme *Sentence-BERT* (SBERT) untuk mendukung proses PMB Program Studi Teknik Informatika Universitas Pamulang. *Chatbot* ini dirancang untuk memberikan respon otomatis yang akurat dan relevan terhadap pertanyaan calon mahasiswa terkait informasi PMB. Pendekatan SBERT digunakan untuk meningkatkan kemampuan *Chatbot* dalam memahami makna kalimat dan memberikan respons kontekstual. Metode pengembangan perangkat lunak yang digunakan adalah model *Waterfall*, yang mencakup tahapan analisis kebutuhan, perancangan sistem, implementasi, dan pengujian. Hasil yang diharapkan adalah *Chatbot* yang mampu menjawab pertanyaan dengan tingkat akurasi tinggi dan dapat meningkatkan efisiensi dan aksesibilitas layanan informasi PMB. Penelitian ini diharapkan memberikan kontribusi dalam implementasi teknologi NLP di lingkungan akademik.

Kata Kunci: *Chatbot*, *Natural Language Processing*, Penerimaan Mahasiswa Baru, *Sentence-BERT*, Teknik Informatika.

1. PENDAHULUAN

Proses penerimaan mahasiswa baru (PMB) adalah merupakan salah satu kegiatan penting dalam institusi perguruan tinggi yang membutuhkan pelayanan informasi yang cepat dan juga akurat untuk calon mahasiswa. Informasi PMB mencakup berbagai aspek, seperti persyaratan pendaftaran, jadwal seleksi, serta biaya kuliah [1], [2] Apabila informasi tersebut tidak tersampaikan dengan baik, maka dapat menimbulkan kebingungan bagi calon mahasiswa dan berpotensi menurunkan kualitas layanan institusi pendidikan.

Saat ini, layanan informasi PMB di Universitas Pamulang bergantung pada media satu arah seperti Instagram dan brosur di pos tertentu, yang membatasi interaksi langsung untuk jawaban cepat dan detail [3]. Layanan tanya jawab manual oleh admin juga kurang efisien dalam menangani ribuan pertanyaan pada periode pendaftaran, menyebabkan keterlambatan dan penurunan efisiensi.

p-ISSN : 2808-909X e-ISSN : 2798-2505

Berdasarkan kondisi tersebut, Program Studi Teknik Informatika Universitas Pamulang masih mengandalkan layanan manual tanpa implementasi *chatbot* khusus berbasis *Natural Language Processing* (NLP) untuk PMB. Hal ini menjadi urgensi penelitian untuk menerapkan teknologi kecerdasan buatan guna mengoptimalkan layanan interaktif, cepat, akurat, dan efisien.

Chatbot berbasis *NLP* menawarkan solusi otomatisasi informasi, meningkatkan efisiensi dan kepuasan pengguna dengan respons *real-time* 24 jam terhadap pertanyaan umum [4], [5], [6]. Pendekatan yang digunakan adalah *Algoritme Sentence-BERT* (SBERT), dipilih karena kemampuannya memahami makna kalimat melalui representasi vektor efisien, mendukung *chatbot* yang lebih cerdas dan kontekstual [7], [8].

Penelitian terdahulu menunjukkan potensi *chatbot* NLP dalam layanan akademik: [9] mencapai akurasi pencocokan semantik 89% dengan SBERT; [10] meningkatkan pemahaman konteks hingga 80-85% menggunakan BERT; [11] mengimplementasikan ANN untuk layanan akademik dengan kepuasan 93%; [12] menerapkan ANN pada *chatbot* sekolah untuk efisiensi informasi; dan [13] memperoleh akurasi *semantic similarity* 87% dengan SBERT. Namun, penerapan SBERT khusus untuk PMB di perguruan tinggi Indonesia masih terbatas.

Penelitian ini bertujuan merancang, mengembangkan, dan menguji *chatbot* berbasis NLP dengan SBERT untuk respons akurat pada PMB, diterapkan di aplikasi Telegram Program Studi Teknik Informatika Universitas Pamulang, menggunakan metode *Waterfall* dan pengujian Black-Box.

2. TINJAUAN PUSTAKA

2.1. Landasan Teori

2.1.1. *Natural Languange Processing* (NLP)

NLP adalah cabang dari kecerdasan buatan yang memungkinkan komputer memahami dan menghasilkan bahasa manusia. Komponen utama NLP meliputi tokenisasi, parsing, dan representasi semantik [14]. *Natural Language Processing* (NLP) merupakan kemampuan komputer dalam memproses kalimat-kalimat yang digunakan dalam percakapan sehari-hari oleh manusia [15]. Teknologi ini sangat cocok dikembangkan dalam aplikasi *chatbot* karena memungkinkan sistem untuk memahami pesan dan memberikan respons secara cepat dan tepat [4]. Selain itu, NLP juga mampu mengidentifikasi kesalahan penulisan, seperti kesalahan ejaan atau tata bahasa [16], sehingga *chatbot* tetap dapat memahami dan menafsirkan pesan meskipun terdapat kesalahan dalam *input* pengguna. Sebagai contoh *ChatGPT, Glide Apps Customer Service,* dan *Auto Correct* pada *handphone* ataupun PC adalah beberapa aplikasi yang menerapkan NLP sebagai medianya.

2.1.2. Sentence-BERT (SBERT)

SBERT adalah pengembangan dari model BERT yang dioptimalkan untuk menghasilkan *embedding* kalimat. SBERT menggunakan arsitektur *siamese network* untuk membandingkan kemiripan semantik antar kalimat, cocok untuk aplikasi *chatbot* [9]. *Sentence-BERT* (SBERT) berperan penting dalam melakukan proses pemetaan dan pencocokan antar kalimat berdasarkan makna semantik [4]. SBERT mengubah setiap kalimat menjadi representasi vektor (*embedding*) berdimensi tinggi, kemudian menghitung

p-ISSN : 2808-909X e-ISSN : 2798-2505

tingkat kemiripan antar vektor tersebut menggunakan *cosine similarity* [9]. Dua kalimat dikatakan memiliki makna yang mirip atau relevan apabila nilai *cosine similarity* yang dihasilkan mendekati 1, dengan nilai 1 menunjukkan kemiripan sempurna, sedangkan nilai 0 menunjukkan tidak ada kemiripan sama sekali. Dengan demikian, semakin tinggi nilai *cosine similarity*, semakin besar kemungkinan bahwa kedua kalimat memiliki makna yang serupa [5].

$$cosine_{similarity} = \frac{A.B}{||A||||B||}$$

Keterangan:

A	= Kalimat Pembanding 1
B	= Kalimat Pembanding 2
A·B	= Jumlah hasil perkalian antara elemen-elemen yang bersesuaian dari vektor A dan B
||A|| dan ||B||	= Menunjukkan panjang dari masing-masing vektor
||A||||B||	= Hasil kali dari panjang vektor A dan B

Karakteristik dari model *Sentence Transformers* (SBERT), yaitu [9], [17]:

1) *Pre-trained Language Model*
   SBERT menyediakan akses ke lebih dari 10.000 model *pre-trained* yang tersedia di *Hugging Face*, termasuk model-model terdepan dari *leaderboard Massive Text Embeddings Benchmark* (MTEB).
2) Kemampuan *Encoding* dan *Similarity*
   Pengguna dapat dengan mudah memuat model dan menghitung *embedding* untuk kalimat tunggal atau pasangan teks. Selain itu, SBERT mendukung perhitungan skor kesamaan semantik secara langsung.
3) *Sentence Transformer*
   Memuat model dan memanggil `model.encode()` untuk menghitung *embedding* (misalnya, *embedding* berukuran 384 dimensi). Kemudian, `model.similarity()` dapat digunakan untuk menghitung skor kesamaan antar *embedding*.
4) Efisiensi Pencarian
   SBERT mendukung *semantic search* yang lebih cepat dan akurat karena *embedding* bisa di indeks menggunakan *approximate nearest neighbor* (ANN).

2.1.3. *Chatbot*

Chatbot adalah perangkat lunak yang dirancang untuk berinteraksi dengan pengguna melalui teks atau suara. *Chatbot* berbasis *NLP* dapat memahami maksud dari pengguna dan memberikan respons yang relevan [18]. Menurut [9] dalam penelitian tentang *Sentence-BERT, chatbot* modern yang didukung oleh model *transformer* seperti SBERT unggul dalam pemahaman semantik, memungkinkan penanganan yang akurat terhadap pertanyaan kompleks dalam aplikasi *real-time*.

2.1.4. *Cosine Similarity*

Cosine similarity merupakan metrik fundamental dalam *Natural Language Processing* (NLP) yang mengukur kesamaan antara dua vektor non-nol melalui kosinus sudut di antara keduanya [19]. Teknik ini unggul untuk tugas pencocokan maksud dan pengambilan respons di *chatbot* berbasis SBERT, karena efektif menangkap nuansa kontekstual meskipun variasi kata, khususnya pada pertanyaan berbahasa Indonesia tentang proses penerimaan universitas [14]. Cosine similarity lebih baik daripada jarak Euclidean pada

p-ISSN : 2808-909X e-ISSN : 2798-2505

representasi teks jarang, mendukung penilaian kesamaan akurat untuk aplikasi real-time di *chatbot* pendidikan tinggi [9].

2.1.5. Precision, Recall dan F1-Score

a) Precision

Precision merupakan metrik evaluasi yang mengukur proporsi prediksi positif yang benar-benar tepat dari seluruh prediksi positif yang dihasilkan model. *Precision* dihitung dengan rumus:

$$Precision = \frac{TP}{TP + FP}$$

Keterangan:
TP = True Positive
FP = False Positive

Dalam konteks *chatbot* berbasis *NLP* dan *SBERT* untuk penerimaan mahasiswa baru di Universitas Pamulang, *precision* tinggi memastikan bahwa respons yang diberikan terhadap pertanyaan calon mahasiswa tentang persyaratan pendaftaran atau jadwal tes memang relevan, sehingga mengurangi kesalahan informasi yang dapat menyesatkan pengguna [14]. Metrik ini krusial ketika prioritas utama adalah meminimalkan jawaban salah yang berpotensi menimbulkan ketidakpercayaan terhadap sistem.

b) Recall

Recall merupakan metrik yang mengukur kemampuan model dalam mendeteksi seluruh kasus positif yang sebenarnya dari total kasus positif yang ada. Rumusnya yaitu:

$$Recall = \frac{TP}{TP + FN}$$

Keterangan:
TP = True Positive
FN = False Negative

Pada aplikasi SBERT untuk *chatbot* penerimaan mahasiswa, *recall* yang baik memungkinkan sistem menangkap semua variasi pertanyaan terkait biaya kuliah atau dokumen administrasi meskipun difrasakan berbeda oleh calon mahasiswa dari berbagai daerah di Indonesia (Reimers & Gurevych, 2019). Tingginya *recall* mencegah sistem melewatkan pertanyaan penting, sehingga layanan tetap komprehensif walaupun menghadapi keragaman bahasa lokal.

c) F1-Score

F1-score adalah *harmonic mean* dari *precision* dan *recall* yang memberikan keseimbangan tunggal antara keduanya. *F1-score* dihitung menggunakan rumus:

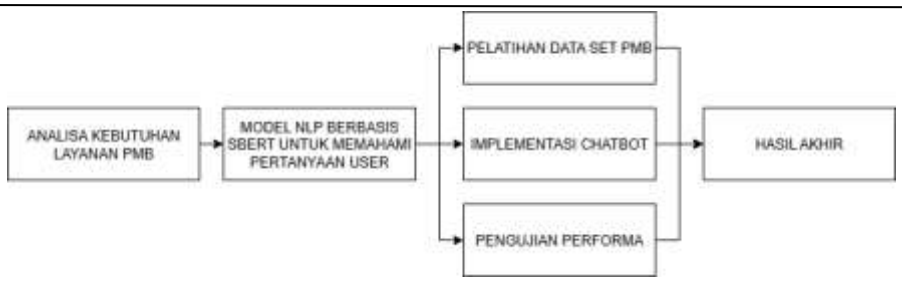
$$F1 = 2 . \frac{Precision\ x\ Recall}{Precision + Recall}$$

Dalam evaluasi *chatbot, SBERT* untuk proses admisi perguruan tinggi Indonesia, F1-score tinggi (misalnya di atas 85%) menunjukkan performa model yang seimbang dalam mengenali *intent* pertanyaan sambil memastikan akurasi respons, yang esensial untuk meningkatkan kepuasan pengguna secara keseluruhan [14].

2.2. Kerangka Berpikir

Alur penelitian ini seperti ditunjukkan pada Gambar 1., dimulai dengan analisis kebutuhan layanan PMB, diikuti dengan pengembangan model NLP berbasis SBERT untuk memahami pertanyaan pengguna. Proses pengembangan meliputi pelatihan model dengan *data set* pertanyaan PMB, implementasi *chatbot*, dan pengujian performa. Hasil akhir berupa *chatbot* yang mampu menjawab pertanyaan dengan akurasi tinggi dan meningkatkan efisiensi layanan.

p-ISSN : 2808-909X e-ISSN : 2798-2505



Gambar 1. Alur Penelitian Chatbot PMB Berbasis SBERT

2.3. Penelitian Terdahulu

Beberapa referensi penelitian yang relevan dengan topik serupa guna mendukung dan memberikan gambaran dan berkaitan dalam penulisan skripsi ini.

Tabel 1. Penelitian Terdahulu [1]

Penulis	Jurnal, Vol, No, Tahun	Permasalahan	Algoritma/ Metode	Hasil/ Akurasi
Alifya Hikmah, Fairuz Azmi, Ratna Astuti Nugrahaeni	Implementasi *Natural Language Processing* Pada *Chatbot* Untuk Layanan Akademik, Vol. 10 No. 1 (2023)	Kurangnya efisiensi dan terbatasnya akses terhadap narasumber resmi yang dapat memberikan informasi akademik secara lengkap dan relevan	(ANN) *Artificial Neural Network.*	Pengujian sistem *chatbot* dengan metode pengujian beta dengan *Usability Testing* mendapatkan hasil sebesar 93% dari total 42 responden dan 6 pertanyaan. Untuk pengujian akurasi pada *chatbot* mendapatkan hasil 100% dari total 54 pertanyaan yang diajukan secara acak. [11]
Dyah Apriliani, Sharfina Febbi Handayani, Irfan Triadi Saputra	Implementasi NLP Dalam Pengembangan Aplikasi *Chatbot* Pada SMK YPE Nusantara Slawi, Vol. 22 No.4 (2023)	Proses penyebaran informasi dan promosi di SMK YPE Nusantara Slawi masih dilakukan secara manual dan belum memanfaatkan teknologi informasi secara optimal, sehingga menyebabkan ketidakefisienan dan keterbatasan akses informasi	(ANN) *Artificial Neural Network.*	Terbangunnya sebuah aplikasi *chatbot* berbasis *website* yang dapat memberikan informasi lengkap seputar SMK YPE Nusantara Slawi secara otomatis dan interaktif, dengan dukungan teknologi NLP untuk meningkatkan efisiensi dalam penyampaian informasi dan promosi sekolah. [12]

		bagi masyarakat dan calon siswa		
Tsaura Rafah Masuzzahra, Khothibul Umam, Hery Mustofa, Maya Rini Handayani	HANA: An AI *Chatbot for Islamic Jurisprudence on Menstruation using SBERT and* TF-IDF, Vol.9 No.3 (2025)	Akses masyarakat, khususnya perempuan, terhadap pengetahuan fikih haid (menstruasi) dalam Islam masih terbatas, dan belum tersedia sistem interaktif yang mampu menjawab pertanyaan-pertanyaan fikih secara akurat dan efisien.	*Sentence -BERT* (SBERT) dengan *siamese network.*	*Chatbot* memberikan skor *cosine similarity* rata-rata sebesar 0.6581 dan tingkat **relevansi semantik berdasarkan validasi ahli mencapai 87%, pengujian *User Acceptance Testing* (UAT)** yang melibatkan 15 responden menunjukkan tingkat kepuasan pengguna sebesar **86,7%.** [13]

3. METODOLOGI PENELITIAN

3.1. Analisa Sistem

Analisa sistem dalam penelitian ini dilakukan untuk memahami kondisi sistem yang ada (sistem berjalan) serta kebutuhan perbaikan yang menjadi dasar pengembangan sistem usulan. Analisis ini berdasarkan observasi langsung terhadap proses layanan informasi PMB yang tersedia saat ini, tanpa melibatkan wawancara nara sumber. Data diperoleh melalui pengamatan terhadap saluran informasi yang digunakan, dengan tujuan mengidentifikasi kelemahan yang ada dan merumuskan solusi yang sesuai untuk program Studi Teknik Informatika Universitas Pamulang.

3.1.1. Analisa Sistem Berjalan

Sistem berjalan untuk layanan informasi PMB di Program Studi Informatika Universitas Pamulang saat ini bergantung pada pendekatan pasif melalui media sosial, media promosi lainnya, dan pos PMB yang disediakan di beberapa titik wilayah kampus. Berdasarkan observasi yang dilakukan terdapat beberapa kelemahan yang teridentifikasi yaitu:

Tabel 2. Data Observasi [2]

No.	Masalah yang teridentifikasi
1.	Informasi yang tersedia melalui media sosial seperti Instagram dan website resmi sering kali tidak diperbarui secara langsung, menyebabkan data yang diberikan kadang - kadang tidak sinkron dengan perkembangan terbaru, seperti jadwal pendaftaran atau persyaratan tambahan.
2.	Media promosi seperti brosur atau poster memiliki keterbatasan akses aksesibilitas, karna hanya efektif bagi calon mahasiswa yang berada di dekat titik distribusi atau kampus

p-ISSN : 2808-909X e-ISSN : 2798-2505

3.	Pos PMB di wilayah kampus memiliki jam operasional terbatas, yang menyulitkan calon mahasiswa di luar jam tersebut untuk mendapatkan informasi

Observasi menunjukkan bahwa kurangnya interaksi langsung dan otomatisasi menyebabkan potensi kebingungan di kalangan calon mahasiswa, terutama pada periode pendaftaran puncak.

3.1.2. Analisa Sistem Usulan

Sistem usulan berupa *chatbot* berbasis *Natural Languange Processing* (NLP) dengan algoritme Sentence-BERT (SBERT) dirancang untuk mengatasi kelemahan sistem berjalan dengan menyediakan solusi otomatis dan cerdas (Gambar 3.1). Berdasarkan kebutuhan yang diidentifikasi dari observasi, sistem ini bertujuan untuk meningkatkan akses informasi PMB dengan respon instan melalui platform digital yang dapat diakses kapan saja. *Chatbot* ini dirancang mampu memahami variasi pertanyaan calon mahasiswa menggunakan analisa semantik SBERT dan terintegrasi dengan *database* resmi PMB untuk memastikan akurasi data, dengan demikian, sistem usulan diharapkan mengurangi ketergantungan pada media pasif dan pos fisik, serta meningkatkan efisiensi layanan informasi.

3.2. Perancangan Sistem

Perancangan sistem merupakan tahap berikutnya setelah analisis kebutuhan, yang bertujuan untuk merancang solusi teknis berdasarkan sistem usulan, proses ini mencakup desain arsitektur sistem, diagram interaksi, dan alur kerja yang akan diimplementasikan dalam pengembangan *chatbot*. Perancangan dilakukan dengan mempertimbangkan kebutuhan fungsional dan non-fungsional yang telah diidentifikasi, serta memastikan kompatibilitas dengan infrastruktur teknologi yang ada di Universitas Pamulang

3.2.1. Perancangan Model *Chatbot*

Arsitektur sistem dirancang menggunakan pendekatan berbasis bot Telegram dengan bahasa pemrograman *Python*. Sistem ini mengintegrasikan API Telegram untuk memungkinkan interaksi pengguna melalui aplikasi Telegram, dengan server yang menangani pemrosesan *Natural Language Processing* (NLP) menggunakan algoritme *Sentence-BERT* (SBERT). Arsitektur terdiri dari tiga modul utama (Tabel 3) yakni:

Tabel 3. Tipe Model Sistem Usulan [3]

Model	Deskripsi Kegunaan
Model Input	Menerima pertanyaan dari calon mahasiswa melalui pesan teks di Telegram, yang diteruskan ke server melalui *webhook* atau *polling* API Telegram.
Modul Pemrosesan NLP	Penelitian ini menggunakan *library sentence-transformers* yang disediakan oleh *Hugging Face* untuk menghasilkan *embedding* kalimat. *Embedding* ini kemudian digunakan dalam perhitungan kesamaan semantik antara pertanyaan pengguna dengan basis data FAQ Penerimaan Mahasiswa Baru (PMB)
Model Output	Modul ini bertugas menghasilkan respons teks yang relevan berdasarkan hasil pemrosesan NLP, kemudian mengirimkannya kembali kepada pengguna melalui API Telegram. Basis data FAQ PMB dikelola dalam bentuk Excel yang dikonversi menjadi JSON,

p-ISSN : 2808-909X e-ISSN : 2798-2505

	sehingga data seperti jadwal, persyaratan, dan biaya kuliah dapat diperbarui dengan mudah oleh admin. Bot mengakses file JSON tersebut untuk memberikan jawaban yang akurat kepada pengguna. Selain itu, server dijalankan secara kontinu agar bot dapat beroperasi 24/7 sesuai dengan kebutuhan layanan PMB.

Tahapan dalam perancangan dan pemodelan *Chatbot* ini dilakukan menggunakan bahasa pemrogram *python* mencakup beberapa tahapan sebagai berikut:

1) Pengumpulan data
   Input data berupa data FAQ yang telah disiapkan diambil dari web dan sumber data FAQ yang dikonfirmasikan dengan pengguna di Lembaga Pemasaran dan Penerimaan Mahasiswa Baru (LPPMB) Universitas Pamulang. Hasil FAQ pada penelitian ini seperti ditunjukkan pada Tabel 4. Data FAQ yang dibuat sebanyak 40 *record* data.

   Tabel 4. FAQ Penerimaan Mahasiswa Baru Program Studi TI [4]

Question	***Answer***
Kapan pendaftaran PMB UNPAM TA 2025/2026 dimulai?	Pendaftaran dibuka mulai 4 Mei 2025 untuk jalur online (gelombang 1).
Apa batas akhir pendaftaran online gelombang 1?	3 Mei 2025.
Bagaimana cara mendaftar secara offline?	Datang ke sekretariat PMB di kampus UNPAM dan isi formulir pendaftaran.
Apa itu SPMB UNPAM?	SPMB adalah Seleksi Penerimaan Mahasiswa Baru ujian seleksi online untuk calon mahasiswa.

Selanjutnya data Penerimaan Mahasiswa Baru (PMB) disimpan dalam *file faq.json*, yang berisi pasangan pertanyaan dan jawaban yang digunakan untuk pelatihan dan pencocokan oleh *SBERT.*

2) Pre-processing Data
   Tahapan selanjutnya dilakukan proses *tokenisasi* dan *embedding* sebagai tahapan *pre-processing* untuk dilanjutkan ke pemodelan SBERT dengan penjelasan detail sebagai berikut:
   a. *Tokenisasi*
      Tokenisasi adalah proses awal dalam pemrosesan teks, di mana pertanyaan dari calon mahasiswa dipecah menjadi unit-unit kecil yang disebut token (Jurafsky & Martin, 2024). Ditunjukkan pada Gambar 3.4 bahwa proses tokenisasi dimulai dengan memisahkan teks berdasarkan spasi dan tanda baca, kemudian mengidentifikasi kata-kata atau frasa yang bermakna (misalnya, "jadwal pendaftaran" menjadi dua token: "jadwal" dan "pendaftaran"). Dalam SBERT, tokenisasi menggunakan WordPiece, sebuah algoritma yang membagi kata-kata kompleks menjadi sub-kata (subword) untuk menangani variasi bahasa, seperti menjadikan "pendaftaran" menjadi "pen", "##daft", dan "##aran", atau bisa

p-ISSN : 2808-909X e-ISSN : 2798-2505

memisahkan dari satu kalimat menjadi beberapa kata seperti "Bagaimana cara mendaftar?" menjadi ["Bagaimana", "Cara", "Mendaftar","?"]

b. Embedding

Embedding adalah proses mengonversi token menjadi representasi numerik (vektor) yang menangkap makna semantik. Dalam SBERT, *embedding* dilakukan menggunakan arsitektur *Transformer* berbasis *bidirectional*, yang memproses teks dari kedua arah (kiri ke kanan dan kanan ke kiri) untuk memahami konteks (Devlin et al., 2019). Setiap token dari tahap *tokenisasi* dimasukkan ke model SBERT, yang menghasilkan vektor berkode tinggi (biasanya 768 dimensi). Vektor ini menggabungkan informasi konteks dari seluruh kalimat, bukan hanya token individu, menggunakan lapisan *attention*. Hasil *embedding* ini digunakan untuk pencocokan dengan data pelatihan dari data *base* PMB. Proses dilanjutkan dengan pencocokan, di mana *embedding* pertanyaan dibandingkan dengan vektor data pelatihan menggunakan *cosine similarity* untuk memilih jawaban paling relevan. Kinerja SBERT divalidasi pada data set STS-B, menunjukkan akurasi tinggi dalam tugas perbandingan semantik.

c. Pemodelan *S-BERT*

Tahapan proses pemodelan *chatbot* berbasis NLP menggunakan algoritme *Sentence-BERT (SBERT)* dikembangkan dengan menggunakan bahasa pemrograman *python* menggunakan *library* terkait seperti pada *Flask, request, jsonify, SentenceTransformer, util, torch, sentence_bleu* dan *SmoothingFunction*. Model dibuat menggunakan *SentenceTransformer* dengan parameter model yang dipilih adalah jenis *sentence-transformers/paraphrase-multilingual-MiniLM-L12-v2.*

d. Evaluasi Model

Untuk memberikan *output* dari model SBERT ini ditentukan sebuah parameter *threshold* sebagai batasan dalam kemampuan memberikan respons menjawab atas pertanyaan yang diterima apakah dapat dilakukan atau tidak berdasarkan *cosine similarity* yakni dengan nilai adalah 0.4 .

3.2.2. *Use Case* Diagram

Use Case Diagram dirancang untuk menggambarkan interaksi antara dua aktor utama, yaitu Admin dan Calon Mahasiswa, dengan sistem *chatbot* berbasis telegram. Diagram pada Tabel 5 menjelaskan fungsi-fungsi utama yang dapat dilakukan oleh masing-masing aktor berdasarkan peran mereka dalam proses Penerimaan Mahasiswa Baru (PMB) di Program Studi Teknik Informatika Universitas Pamulang. Sistem dirancang untuk mengidentifikasi autentikasi tambahan untuk admin.

Tabel 5. Daftar Aktor [5]

Aktor	Fungsi
Calon Mahasiswa	Aktor ini memiliki akses terbatas hanya untuk mengirim pertanyaan terkait PMB, seperti jadwal pendaftaran, persyaratan, dan biaya kuliah. Setelah , mengirim *chat command* ***/ask*** sistem langsung memproses pertanyaan menggunakan algoritme *Sentence-BERT* (SBERT) dan memberikan respons otomatis tanpa memerlukan autentikasi tambahan
Admin	Aktor ini memiliki hak akses lebih luas dan harus melalui autentikasi. Ketika admin memasukkan *chat command* ***/Admin***

p-ISSN : 2808-909X e-ISSN : 2798-2505

	sistem akan meminta input PASSWORD untuk verifikasi. Setelah berhasil, admin dapat mengunggah file Excel yang berisi daftar pertanyaan dan jawaban terbaru untuk memperbarui *database* PMB, memastikan informasi tetap terkini

Sistem dirancang untuk membedakan aktor berdasarkan perintah awal yang dimasukkan. Jika *command* ***/ask*** terdeteksi, sistem mengenali pengguna sebagai calon mahasiswa dan membatasi fungsi hanya untuk pertanyaan PMB, jika *Command* ***/Admin*** terdeteksi, sistem meminta autentikasi ID dan PASSWORD sebelum memberikan akses kepada admin untuk mengunggah file Excel. Mekanisme ini memastikan keamanan data dan pembagian peran yang jelas.

3.2.3. *Activity* Diagram

Activity diagram menggambarkan alur utama sistem *Chatbot* Telegram berbasis FAQ dengan dua peran: Calon Mahasiswa dan Admin. Calon Mahasiswa mengirim pertanyaan terkait PMB melalui antarmuka Telegram; sistem memproses input, mencocokkan dengan database FAQ, dan memberikan jawaban relevan secara otomatis. Ini menunjukkan fungsi *chatbot* sebagai sistem tanya jawab teks berbasis NLP.

Admin bertanggung jawab mengelola data FAQ: melakukan autentikasi, mengunggah file (.json atau .xlsx) berisi pertanyaan-jawaban terbaru, lalu sistem memperbarui database untuk menjaga keakuratan informasi.

Secara keseluruhan, diagram ini memperlihatkan dua alur paralel: interaksi pengguna untuk informasi cepat dan manajemen data oleh admin untuk pembaruan dinamis. Sistem bersifat adaptif, di mana respons *chatbot* bergantung pada data FAQ terkini yang dikelola admin [20].



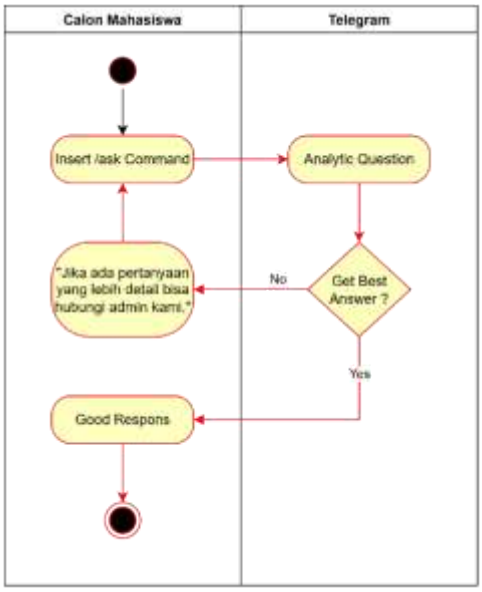
Gambar 2 *Activity* Diagram Admin

Activity diagram pada Gambar 3 menggambarkan alur proses login admin dan pembaruan data FAQ melalui Telegram bot. Proses dimulai saat admin memasukkan perintah /admin, diikuti permintaan kata sandi untuk login. Sistem memvalidasi kata sandi: jika salah, tampil pesan “Password Salah” dan kembali ke tahap login;

p-ISSN : 2808-909X e-ISSN : 2798-2505

Admin kemudian mengunggah *file* Excel berisi data FAQ terbaru. Sistem memproses *file* tersebut, memperbarui *database* FAQ, dan memberikan notifikasi "FAQ Berhasil Diperbarui".

Diagram ini menekankan mekanisme autentikasi aman dan pembaruan data yang hanya dapat dilakukan admin dengan hak akses valid, memastikan keakuratan informasi *chatbot*.



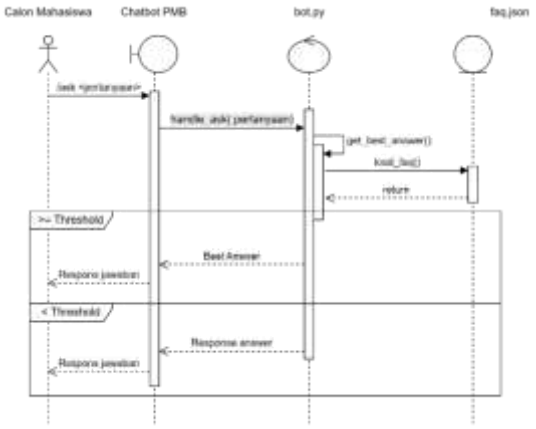
Gambar 7. Activity Diagram Calon Mahasiswa

Activity diagram pada Gambar 7 menggambarkan alur interaksi calon mahasiswa dengan *chatbot* Telegram untuk memperoleh informasi PMB. Proses dimulai saat pengguna mengetik perintah /ask, diikuti analisis pertanyaan oleh sistem. Sistem kemudian mencari jawaban terbaik dengan membandingkan input pengguna terhadap *database* FAQ. Jika ditemukan kecocokan, *chatbot* memberikan respons relevan (good response); jika tidak, tampil pesan alternatif "Jika ada pertanyaan yang lebih detail bisa hubungi admin kami" untuk mengarahkan ke bantuan manual. Proses berakhir setelah *chatbot* memberikan jawaban, memastikan layanan informasi otomatis dan efisien bagi calon mahasiswa.

3.2.4. *Sequance* Diagram

Sequence Diagram menggambarkan urutan interaksi antar objek dalam sistem *chatbot* berbasis Telegram, dengan fokus pada alur komunikasi antara aktor (calon mahasiswa dan admin) serta komponen sistem seperti Telegram API, SBERT Model, dan PMB *Database*. Diagram ini menunjukkan bagaimana pesan atau data ditransmisikan secara berurutan untuk memproses pertanyaan atau memperbarui data, sesuai dengan desain sistem yang diusulkan.

p-ISSN : 2808-909X e-ISSN : 2798-2505



Gambar 4. *Activity* Diagram Calon Mahasiswa

Sequence diagram di atas menggambarkan alur interaksi antara pengguna, bot Telegram, file utama bot.py, dan basis data FAQ (faq.json). Proses dimulai saat pengguna mengirim perintah /ask <pertanyaan> ke bot Telegram, yang diteruskan ke fungsi handle_ask() di bot.py. Fungsi ini memanggil get_best_answer() untuk mencari jawaban terbaik: memuat data FAQ via load_faq(), menganalisis kesamaan semantik menggunakan model SBERT, menghitung *similarity score*, dan membandingkan dengan *threshold*. Jika score ≥ threshold, kembalikan jawaban relevan; jika tidak, berikan respons alternatif (misal "Pertanyaan belum tersedia, hubungi admin"). Hasil akhir dikirim kembali ke bot Telegram untuk ditampilkan kepada pengguna. Diagram ini memperlihatkan mekanisme pencocokan teks otomatis berbasis *similarity* untuk respons cepat dan akurat pada *chatbot* FAQ.

4. HASIL DAN PEMBAHASAN

4.1. Implementasi Model *Chatbot*

Implementasi dilakukan sebagai realisasi dari desain yang telah dirancang pada bab sebelumnya. Sistem *chatbot* berbasis Telegram dikembangkan menggunakan bahasa pemrograman *Python* dengan *library python-telegram-bot* untuk integrasi dengan API Telegram dan *Hugging Face* untuk model Sentence-BERT (SBERT). Server diatur menggunakan *Flask* untuk menangani *webhook* dari Telegram, memastikan komunikasi *real-time* antara pengguna dan sistem. Data Penerimaan Mahasiswa Baru (PMB) disimpan dalam file faq.json, yang berisi pasangan pertanyaan dan jawaban yang digunakan untuk pelatihan dan pencocokan oleh SBERT

4.2. Implementasi *Software*

Implementasi pada aplikasi *chatbot* penerimaan mahasiswa baru terdiri dari:

a. Bahasa pemrograman yang digunakan adalah *python*.
b. *Library* utama yang digunakan adalah:
   - *Python-telegram-bot* (integrasi API Telegram).
   - *Flask (webhook handler*).
   - *Sentence-transformers* (model SBERT).
   - *Transformers* dari *Hugging Face.*
   - Json untuk memproses *dataset* FAQ.
c. Server: *Webhook* HTTPS pada *hosting*.

p-ISSN : 2808-909X e-ISSN : 2798-2505

d. Platform *Chatbot*: Telegram Bot API.
e. *Dataset* PMB: File faq.json yang berisi pertanyaan dan jawaban.

4.3. Implementasi *Hardware*

Adapun rincian *hardware* yang digunakan dalam pengembangan *chatbot* penerimaan mahasiswa baru yaitu:

a. Perangkat utama berupa laptop dengan prosesor Intel *Core* i3.
b. Memori (RAM) berkapasitas 4 GB.
c. Media penyimpanan menggunakan SSD berkapasitas 256 GB.
d. Layar 14 inci dengan resolusi *Full* HD.

4.4. Implementasi Antarmuka *Chatbot*

Antarmuka pengguna (UI) pada sistem ini berbasis platform Telegram, sehingga pengguna dapat langsung bertanya mengenai informasi PMB secara *real-time*. Terdapat dua jenis antarmuka yaitu:

4.4.1. Antarmuka Pengguna Calon Mahasiswa

Antarmuka ini digunakan oleh calon mahasiswa untuk mengajukan pertanyaan terkait Penerimaan Mahasiswa Baru.



Gambar 5. Antarmuka Pengguna Calon Mahasiswa

Keterangan:

- /*start* untuk memulai interaksi dan menampilkan menu utama.
- /*ask* <pertanyaan> untuk mengajukan pertanyaan kepada *chatbot*.

Setelah perintah /ask dikirim, sistem akan menampilkan instruksi agar pengguna memasukkan pertanyaan. *Chatbot* kemudian memproses pertanyaan tersebut menggunakan SBERT dan memberikan jawaban berdasarkan kecocokan semantik dengan data pada *faq.json*.

4.4.2. Antarmuka Admin

Antarmuka ini digunakan oleh admin untuk melakukan pengelolaan data *Frequently Asked Questions* (FAQ) yang digunakan oleh *chatbot*, prosesnya dimulai dengan perintah /admin diikuti *password* untuk *login*, kemudian admin dapat mengunggah file Excel baru untuk memperbarui data FAQ. Contoh alur interaksi admin dengan *chatbot* ditunjukkan pada Gambar 5 berikut.

p-ISSN : 2808-909X e-ISSN : 2798-2505



Gambar 6. Antarmuka Pengguna Admin

Keterangan:

Berdasarkan Gambar 6 admin dapat memperbarui data Penerimaan Mahasiswa Baru (PMB) dengan cara mengunggah file Excel yang berisi daftar pertanyaan dan jawaban terbaru.

Setelah *file excel* diunggah, sistem akan melakukan proses validasi format untuk memastikan bahwa struktur data telah sesuai ketentuan. Apabila file dinyatakan valid, sistem akan mengonversi *file* tersebut menjadi *faq.json* baru dan menggantikan *file* sebelumnya yang digunakan *chatbot*.

4.5. Pengujian Sistem

Pengujian sistem dilakukan untuk memastikan bahwa *chatbot* dapat bekerja sesuai dengan tujuan yang ditetapkan, terutama dalam hal pencocokan pertanyaan pengguna dengan data FAQ secara *semantic*.

4.5.1. Pengujian Metode SBERT

Pengujian ini dilakukan untuk mengetahui kemampuan metode *Sentence-BERT* (SBERT) dalam mengukur tingkat kesamaan antara pertanyaan pengguna dengan data pertanyaan–jawaban yang tersimpan dalam basis data FAQ. Pada proses ini, setiap teks dikonversi menjadi vektor *embedding* berdimensi 768 menggunakan model SBERT. Setelah *embedding* diperoleh, nilai kemiripannya dihitung menggunakan metode *Cosine Similarity*. Adapun rumus *Cosine Similarity* adalah sebagai berikut:

$$cosine_{similarity} = \frac{A.B}{||A||||B||}$$

Contoh Perhitungan Manual

Untuk memberikan gambaran proses perhitungan, berikut contoh pengolahan satu data. Meskipun embedding sebenarnya memiliki 768 dimensi, pada contoh ini penulis tampilkan hanya 5 dimensi pertama untuk memudahkan penjelasan.

Pertanyaan Pengguna:

"Kapan pendaftaran PMB UNPAM TA 2025/2026 dimulai?"

Jawaban Sistem:

"Pendaftaran dibuka mulai 8 September 2025 untuk jalur *online* (gelombang 1)."

***Embedding* Pertanyaan (A):**

p-ISSN : 2808-909X e-ISSN : 2798-2505

A = [0.213,0.512,−0.144,0.337,0.109]

***Embedding* Jawaban FAQ (B):**

B = [0.201,0.498,−0.130,0.355,0.097]

Menghitung *Dot Product* A · B

$$\mathbf{A} \cdot \mathbf{B} = (0.213)(0.201) + (0.512)(0.498) + (-0.144)(-0.130) + (0.337)(0.355) + (0.109)(0.097)$$

$$= 0.0428+0.255+0.0187+0.1194+0.0105$$

$$= \mathbf{0.4464}$$

Menghitung Norma Masing-masing Vektor

Norma vektor A:

$$|| \mathbf{A} || = \sqrt{0.213^2 + 0.512^2 + (-0.144)^2 + 0.337^2 + 0.109^2}$$

$$= \sqrt{0.045 + 0.262 + 0.020 + 0.113 + 0.011}$$

$$= \sqrt{0.451} = \mathbf{0.6715}$$

Norma vektor B:

$$|| \mathbf{B} || = \sqrt{0.201^2 + 0.498^2 + (-0.130)^2 + 0.355^2 + 0.097^2}$$

$$= \sqrt{0.040 + 0.248 + 0.016 + 0.126 + 0.009}$$

$$= \sqrt{0.439} = \mathbf{0.6625}$$

Hitung *Cosine Similarity*

$$\textit{Cosine Similarity} = \frac{0.4464}{(0.6715)(0.6625)} = \frac{0.4464}{0.4448} = \mathbf{0.8822}$$

Nilai 0.8822 menunjukkan tingkat kesamaan yang sangat tinggi antara pertanyaan pengguna dan jawaban FAQ yang diberikan oleh sistem.

Hasil lengkap pengujian akurasi sistem dalam mencocokkan pertanyaan pengguna dengan data FAQ ditunjukkan pada Tabel 6 berikut:

Tabel 6. Hasil Pengujian Metode SBERT [6]

No	Pertanyaan	Jawaban Sistem	Cosine Sim	Relevan
1	Kalau mau tahu akreditasi Unpam, ceknya di mana ya?	Untuk informasi resmi dan sertifikat akreditasi terbaru, Anda dapat mengunjungi laman LPM Unpam di : Cek Disini	0.79	Ya
2	Biaya kuliah S1 di Unpam Serang kira-kira berapa?	Universitas Pamulang (UNPAM) menawarkan biaya kuliah yang terjangkau dengan sistem pembayaran yang fleksibel bagi seluruh mahasiswa. Untuk melihat rincian lengkap biaya kuliah setiap program studi, silakan kunjungi tautan berikut: Lihat Rincian Biaya Kuliah Pascasarjana	0.85	Ya

3	Kalau kuliah D3 atau D4 di Unpam Tangsel habis berapa?	Universitas Pamulang (UNPAM) menawarkan biaya kuliah yang terjangkau dengan sistem pembayaran yang fleksibel bagi seluruh mahasiswa. Untuk melihat rincian lengkap biaya kuliah setiap program studi, silakan kunjungi tautan berikut: Lihat Rincian Biaya Kuliah UNPAM Tangsel	0.83	Ya
4	S2 di Unpam biayanya mahal nggak?	Universitas Pamulang (UNPAM) menawarkan biaya kuliah yang terjangkau dengan sistem pembayaran yang fleksibel bagi seluruh mahasiswa. Untuk melihat rincian lengkap biaya kuliah setiap program studi, silakan kunjungi tautan berikut: Lihat Rincian Biaya Kuliah Pascasarjana	0.74	Ya
5	Pembayaran uang kuliah di Unpam bisa lewat apa aja?	Pembayaran biaya kuliah dapat dilakukan dengan mudah, baik secara offline maupun online. Offline: Melalui Loket Keuangan UNPAM di Kampus Pusat, Viktor, atau Serang. Online: Melalui bank mitra resmi UNPAM, yaitu: • Bank Mandiri Panduan • Bank BTN Panduan • Bank BJB Panduan • Bank BNI Panduan • Bank BPR Universitas Pamulang Kampus II. Pastikan Anda mengikuti petunjuk pembayaran sesuai ketentuan yang berlaku agar proses administrasi berjalan lancar	0.85	Ya

4.5.2. Analisis Hasil Pengujian

Pada tahap ini dilakukan pengujian untuk mengevaluasi kinerja metode *Sentence-BERT (SBERT)* dalam mencocokkan pertanyaan pengguna dengan kumpulan FAQ yang tersedia. Berdasarkan total 40 data pengujian diperoleh hasil:

- 36 data relevan (jawaban sesuai)
- 4 data tidak relevan (jawaban tidak sesuai)

Perhitungan Akurasi Sistem:

$$\text{Akurasi} = \frac{36}{40} \times 100\% = 90\%$$

Dengan demikian, model SBERT yang digunakan memiliki tingkat akurasi 90**%,** menunjukkan performa yang cukup baik dalam pencarian semantik berbasis FAQ.

Analisis Nilai *Cosine Similarity*:

a) Kategori Relevan (36 data)

p-ISSN : 2808-909X e-ISSN : 2798-2505

Berdasarkan hasil tabel pengujian di atas nilai *cosine similarity* untuk data relevan berada pada rentang 0.71 – 1. Beberapa nilai bahkan mendekati 1.0, seperti 0.95 0.97, 0.98. Nilai yang tinggi ini menunjukkan bahwa *SBERT* berhasil memahami makna kalimat meskipun struktur kata tidak sama persis.

b) Kategori Tidak Relevan (4 data)
Berdasarkan analisis hasil sistem *question-answering* berbasis SBERT, ditemukan diskrepansi antara skor *cosine similarity* yang tinggi (0,72–0,78) dengan ketidakrelevanan jawaban. Fenomena ini terutama disebabkan oleh bias kata kunci yang menyesatkan, di mana model memberikan skor tinggi akibat *overlap* leksikal atau semantik parsial, meskipun dokumen tidak menjawab inti pertanyaan. Contohnya, pertanyaan "Lulusan D3 bisa lanjut S1 di Unpam nggak?" memperoleh skor 0,72 karena kemunculan kata "ijazah" dan "SMA/SMK/MA", sedangkan dokumen hanya berisi daftar persyaratan administrasi tanpa menjelaskan kebijakan kelanjutan studi. Secara keseluruhan SBERT terbukti efektif digunakan sebagai mesin pencarian semantik untuk *chatbot* Penerimaan Mahasiswa Baru.

4.5.3. Evaluasi Kinerja Menggunakan *Precision, Recall* dan *F1-Score*

Setelah melakukan pengujian nilai *cosine similarity* dan akurasi relevansi, evaluasi tambahan dilakukan menggunakan metrik *Precision, Recall, dan F1-Score*. Metrik ini digunakan untuk mengukur sejauh mana sistem dapat memberikan jawaban yang relevan terhadap pertanyaan pengguna. Dari total 405 data pengujian, diperoleh:

- *True Positive* (TP) = 36 (jawaban relevan dan sesuai)
- *False Positive* (FP) = 4 (jawaban dianggap relevan tetapi tidak sesuai)
- *False Negative* (FN) = 0
- *True Negative* (TN) = 0

a) Precision

$$Precision = \frac{36}{36+4} = \frac{36}{40} = 0.9\ (\mathbf{90}\%)$$

b) Recall

$$Recall = \frac{36}{36+0} = 1.0\ (\mathbf{100}\%)$$

c) F1-Score

$$F1 = 2 \times \frac{0.9 \times 1.0}{0.9 + 1.0}$$
$$F1 = 0.9473(\mathbf{94.73}\%)$$

Dari hasil evaluasi diperoleh bahwa ***Precision* 90%** menunjukkan bahwa sebagian jawaban yang diberikan *chatbot* sesuai dan relevan. ***Recall* 100%** menandakan bahwa seluruh pertanyaan yang seharusnya relevan berhasil ditemukan oleh sistem. **F1-Score 94.73%** menunjukkan performa keseluruhan sistem berada pada kategori baik.

5. KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan hasil penelitian dan pengembangan yang telah dilakukan, dapat disimpulkan bahwa penelitian ini berhasil merancang dan mengembangkan *chatbot* berbasis *Natural Language Processing* (NLP) dengan algoritme *Sentence-BERT* (SBERT) yang mampu memproses dan memahami pertanyaan terkait layanan Penerimaan Mahasiswa Baru

(PMB) secara semantik dan relevan dengan F1-*Score* 94.73%. Implementasi *chatbot* telah dibuat *prototipe* menggunakan aplikasi telegram dan API server menggunakan *Flask*, mampu menggambarkan penambahan fitur layanan otomatis dan efisien serta mampu memberikan respons cepat dan akurat dalam layanan PMB.

5.2. Saran

Berdasarkan hasil penelitian yang telah diperoleh, penulis menyarankan agar implementasi *chatbot* ke depannya dapat diintegrasikan dengan sistem lain yang telah ada di Universitas Pamulang, seperti fitur *chatting* pada web sistem pendaftaran *online* sehingga dapat memperluas akses dan layanan. Selain itu, diperlukan evaluasi dan peningkatan performa *chatbot* secara rutin, termasuk penambahan dan pembaruan *data set* pertanyaan agar tingkat akurasi tetap terjaga dan terus meningkat. Pengembangan selanjutnya juga dapat dilakukan dengan menambahkan fitur pembelajaran berkelanjutan agar *chatbot* semakin memahami variasi pertanyaan yang lebih kompleks. Selain itu layanan ini dapat dikembangkan bagi PMB program studi lainnya bahkan di luar PMB yang ada di Universitas Pamulang.

DAFTAR PUSTAKA

[1] M. Hamzani Wahid, D. Tejo Kumoro, and Z. Abidin, "Implementasi Chatbot Berbasis Natural Language Processing Pada Web Lppmb Di Uniqhba Implementation of a Natural Language Processing-Based Chatbot on the LPPMB Website at UNIQHBA," *SIJ*, vol. 7, no. 2, pp. 489–496, 2024, doi: https://doi.org/10.37824/sij.v7i2.2024.758.

[2] A. Yudin, I. Maulana, A. Fitriani, D. Kasoni, and F. Aprilyani, "Sistem Penerimaan Mahasiswa Baru Cerdas Berbasis Web," vol. 14, no. 1, Feb. 2025, doi: https://doi.org/10.51998/jsi.v14i01.600.

[3] R. Fitria, R. P. Tulodo, and A. Sofyan, "Pemanfaatan chatbot untuk informasi pendaftaran calon mahasiswa baru di perguruan tinggi," *IT-Explore: Jurnal Penerapan Teknologi Informasi dan Komunikasi*, vol. 4, no. 2, pp. 221–232, Jun. 2025, doi: 10.24246/itexplore.v4i2.2025.pp221-232.

[4] K. Shiva *et al.*, "Natural Language Processing for Customer Service Chatbots: Enhancing Customer Experience," *Original Research Paper International Journal of Intelligent Systems and Applications in Engineering IJISAE*, vol. 2024, no. 22s, pp. 155–164, 2024, [Online]. Available: www.ijisae.org

[5] N. Bhardwaj, "Natural Language Processing in AI Chatbots," Dec. 2024. doi: https://dx.doi.org/10.2139/ssrn.5044217.

[6] W. Brown, G. Wilson, and O. Johnson, "Understanding the Role of Chatbots in Enhancing Customer Service," Aug. 06, 2024. doi: 10.20944/preprints202408.0321.v1.

[7] B. Juarto and A. S. Girsang, "Neural Collaborative with Sentence BERT for News Recommender System," *INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION*, 2021, doi: https://dx.doi.org/10.30630/joiv.5.4.678.

[8] R. M. Holis, P. E. P. Utomo, and B. F. Hutabarat, "Semantic FAQ Chatbot Using SBERT (Sentence-BERT) and Cosine Similarity for Academic Services," *Brilliance: Research of Artificial Intelligence*, vol. 5, no. 2, pp. 915–922, Oct. 2025, doi: 10.47709/brilliance.v5i2.7027.

p-ISSN : 2808-909X e-ISSN : 2798-2505

[9] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pp. 3982–3992, Nov. 2019, doi: https://doi.org/10.18653/v1/D19-1410.

[10] J. Devlin, M.-W. Chang, K. Lee, K. T. Google, and A. I. Language, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, vol. 1, pp. 4171–4186, Jun. 2019, doi: https://doi.org/10.18653/v1/N19-1423.

[11] A. Hikmah, F. Azmi, and R. A. Nugrahaeni, "Implementasi Natural Language Processing Pada Chatbot Untuk Layanan Akademik," *e-Proceeding of Engineering*, vol. 10, pp. 371–382, Feb. 2023.

[12] D. Apriliani, S. Febbi Handayani, I. Triadi Saputra, T. Informatika, and P. Harapan Bersama, "Implementasi Natural Language Processing (NLP) Dalam Pengembangan Aplikasi Chatbot Pada SMK YPE Nusantara Slawi Implementation of Natural Language Processing (NLP) in the Development of a Chatbot Application at SMK YPE Nusantara Slawi," 2023.

[13] T. R. Masuzzahra, K. Umam, H. Mustofa, and M. R. Handayani, "HANA: An AI Chatbot for Islamic Jurisprudence on Menstruation using SBERT and TF-IDF," *Journal of Applied Informatics and Computing (JAIC)*, vol. 9, no. 3, pp. 1013–1024, Jun. 2025, doi: https://doi.org/10.30871/jaic.v9i3.9449.

[14] D. Jurafsky and J. H. Martin, *Speech and Language Processing An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models Third Edition*. 2024. Accessed: Jan. 18, 2026. [Online]. Available: https://web.stanford.edu/~jurafsky/slp3/ed3bookaug20_2024.pdf

[15] S. Bird, E. Klein, and E. Lope, *Natural Language Processing with Python*, 1st ed. O'Reilly Media, Inc, 2009. Accessed: Jan. 19, 2026. [Online]. Available: https://tjzhifei.github.io/resources/NLTK.pdf

[16] R. Giri, L. Yadav, and V. Rakhde, "NLP in Chatbot Customer Service," *INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN SCIENCE ENGINEERING TECHNOLOGY*, vol. 13, no. 5, pp. 7598–7605, 2024, doi: 10.15680/IJIRSET.2024.1305173.

[17] "SentenceTransformers Documentation." Accessed: Jan. 19, 2026. [Online]. Available: https://sbert.net/docs/package_reference/sentence_transformer/SentenceTransformer.html

[18] S. Russell and P. Norvig, *Artificial Intelligence A Modern Approach Fourth Edition Global Edition*, 4th ed. Pearson Education © 2021, 2021. Accessed: Jan. 18, 2026. [Online]. Available: http://lib.ysu.am/disciplines_bk/efdd4d1d4c2087fe1cbe03d9ced67f34.pdf

[19] J. Han, J. Pei, and H. Tong, "Data, easurements, and data preprocessing," in *Data Mining*, 4th ed., 2023, ch. 2, pp. 23–84. doi: https://doi.org/10.1016/B978-0-12-811760-6.00012-6.

[20] M. Adam, M. Wessel, and A. Benlian, "AI-based chatbots in customer service and their effects on user compliance," *Electronic Markets*, vol. 31, no. 2, pp. 427–445, Jun. 2021, doi: 10.1007/s12525-020-00414-7.