

Pengembangan Aplikasi Chatbot Untuk Layanan Penerimaan Mahasiswa Baru Berbasis Natural Language Processing

Agustin Setiyorini¹, Moh. Arif Ma'ruf Setiawan², Bayu Prasetyo³

^{1,2,3} Fakultas Teknik Universitas Janabadra, Jl. Tentara Rakyat Mataram, Yogyakarta, 55231, Indonesia

Info Artikel

Riwayat Artikel:

Received 2025-07-23

Revised 2025-11-03

Accepted 2025-11-05

Corresponding Author:

Agustin Setiyorini

Email: agustin@janabadra.ac.id



This is an open access article under the [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license.

Abstract – The New Student Admission (PMB) information service at Universitas Janabadra has traditionally relied on manual interactions limited to office hours, often leading to delayed responses and increased workloads for administrative staff. This situation highlights the need for a digital solution capable of providing fast, accurate, and real-time information services to prospective students. This study aims to develop and implement a chatbot system based on Natural Language Processing (NLP) to improve the efficiency of PMB information services. The system was developed using the CodeIgniter 4 framework by integrating the Naive Bayes Classifier algorithm for intent classification and the Levenshtein Distance algorithm for text similarity matching. The training dataset was compiled from frequently asked questions by prospective students and validated by PMB staff. Evaluation results show that the chatbot automatically answered 70% of 500 test questions with an accuracy rate of 92% and an average response time of 0.5 seconds, reducing the administrative workload by up to 30%. A user satisfaction survey involving 100 respondents indicated that 85% were satisfied with the system's speed and accuracy. The findings demonstrate that NLP-based chatbots are effective in enhancing the efficiency and quality of higher education information services. The novelty of this study lies in the integration of the Naive Bayes and Levenshtein Distance algorithms within an Indonesian-language PMB service context, optimized for real-time university information systems.

Keywords: Admission Services; Chatbot; Levenshtein Distance; Natural Language Processing; Naive Bayes.

Abstrak – Layanan informasi Penerimaan Mahasiswa Baru (PMB) di Universitas Janabadra selama ini masih dilakukan secara manual dan terbatas pada jam kerja, sehingga sering menyebabkan keterlambatan respons serta meningkatnya beban kerja staf administrasi. Kondisi ini menimbulkan kebutuhan akan sistem digital yang mampu memberikan layanan informasi secara cepat, akurat, dan real-time kepada calon mahasiswa. Penelitian ini bertujuan untuk mengembangkan dan mengimplementasikan sistem chatbot berbasis Natural Language Processing (NLP) sebagai solusi peningkatan efisiensi layanan PMB. Sistem dikembangkan menggunakan framework CodeIgniter 4 dengan penerapan algoritma Naive Bayes Classifier untuk klasifikasi intent dan Levenshtein Distance untuk pencocokan kemiripan teks. Dataset pelatihan disusun berdasarkan kumpulan pertanyaan umum calon mahasiswa dan diverifikasi oleh staf PMB. Hasil evaluasi menunjukkan chatbot mampu menjawab 70% dari 500 pertanyaan secara otomatis dengan tingkat akurasi 92% dan waktu respons rata-rata 0,5 detik, serta menurunkan beban kerja staf administrasi hingga 30%. Survei terhadap 100 pengguna menunjukkan tingkat kepuasan sebesar 85% terhadap kecepatan dan ketepatan respons sistem. Dengan demikian, chatbot berbasis NLP terbukti efektif dalam meningkatkan efisiensi dan kualitas layanan informasi pendidikan tinggi. Kebaruan penelitian ini terletak pada integrasi algoritma Naive Bayes dan Levenshtein Distance dalam konteks layanan PMB berbahasa Indonesia, yang dioptimalkan untuk sistem informasi universitas secara real-time.

Kata Kunci: Chatbot, Layanan PMB, Levenshtein Distance, Natural Language Processing, Naive Bayes.

I. PENDAHULUAN

Penerimaan Mahasiswa Baru (PMB) merupakan salah satu proses penting dalam operasional perguruan tinggi karena menjadi pintu awal interaksi antara calon mahasiswa dan institusi. Proses ini melibatkan berbagai kegiatan komunikasi yang berkaitan dengan informasi program studi, biaya kuliah, jadwal pendaftaran, hingga prosedur administrasi lainnya. Selama ini, proses layanan informasi PMB di banyak perguruan tinggi, termasuk Universitas Janabadra, masih dilakukan secara manual melalui staf administrasi atau media komunikasi konvensional yang terbatas pada jam kerja. Kondisi ini menyebabkan keterlambatan penyampaian informasi, peningkatan beban kerja staf, serta menurunnya efisiensi pelayanan ketika volume pertanyaan calon mahasiswa meningkat secara signifikan [1].

Masalah utama yang dihadapi adalah keterbatasan waktu layanan dan kurangnya pemanfaatan teknologi informasi untuk mendukung layanan PMB secara *real-time*. Layanan manual membuat calon mahasiswa kesulitan memperoleh informasi di luar jam operasional, sementara staf administrasi mengalami kesulitan dalam

menangani banyak pertanyaan secara bersamaan [2][3]. Situasi ini menunjukkan perlunya inovasi digital yang mampu meningkatkan kecepatan, akurasi, dan ketersediaan informasi sepanjang waktu.

Seiring perkembangan teknologi digital, penerapan *Artificial Intelligence* (AI), khususnya *Natural Language Processing* (NLP), telah membuka peluang baru dalam meningkatkan kualitas layanan pendidikan tinggi. Salah satu implementasi yang menonjol adalah penggunaan *chatbot*, sistem otomatis yang dirancang untuk berinteraksi dengan pengguna melalui teks atau suara guna menyampaikan informasi secara efisien [4][5]. *Chatbot* memungkinkan pengguna memperoleh jawaban secara instan dan dapat beroperasi selama 24 jam, sehingga sangat relevan untuk mendukung layanan PMB [6].

Berbagai penelitian terdahulu telah menunjukkan efektivitas *chatbot* dalam konteks layanan pendidikan. *Chatbot* yang dikembangkan oleh Politeknik TEDC Bandung, misalnya, mampu mencapai akurasi sebesar 90,77% dan tingkat kepuasan pengguna sebesar 89,78% [7]. Sementara itu, Di Universitas Muhammadiyah Sukabumi, sistem UMMIBOT berhasil mencapai akurasi fungsionalitas 83,67% [8]. sedangkan Universitas Muhammadiyah Makassar melaporkan peningkatan efisiensi layanan PMB melalui integrasi teknologi NLP [9].

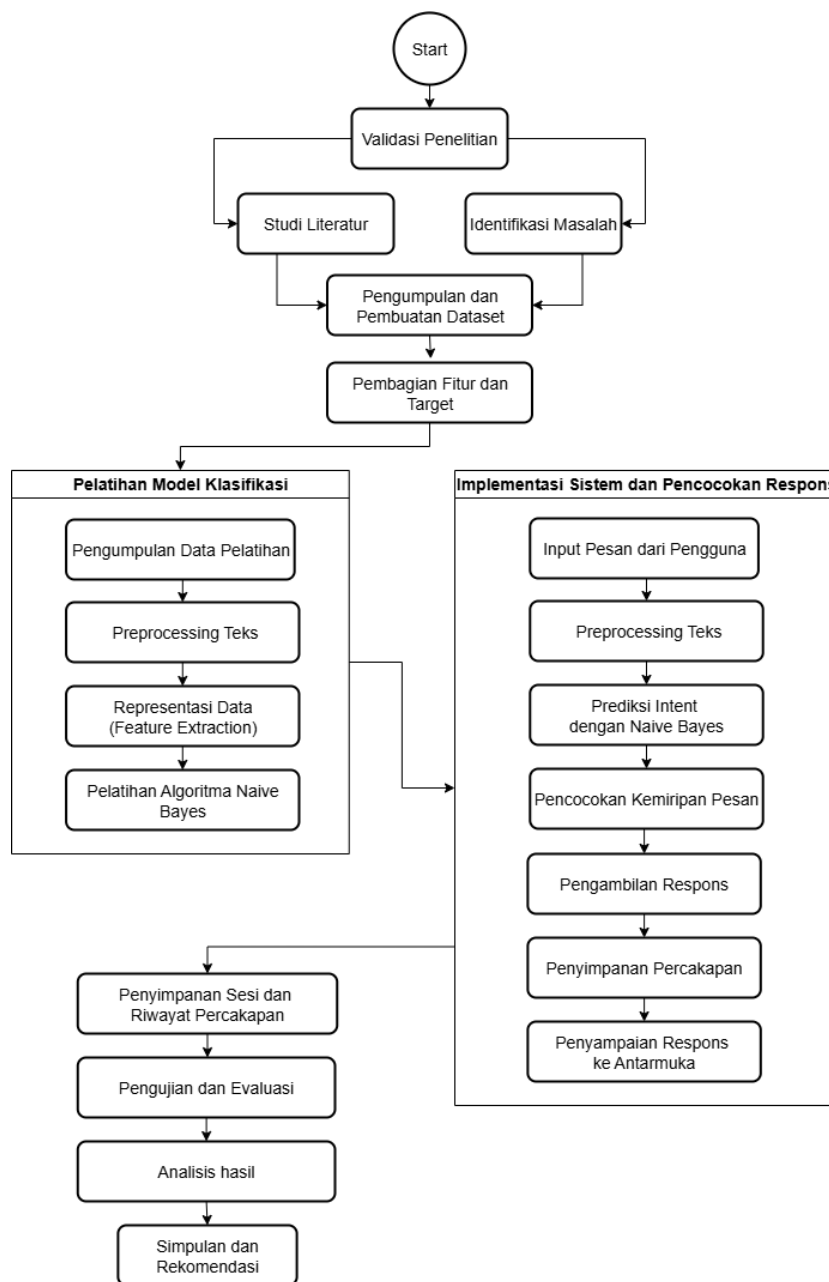
Dari sisi teknis, sejumlah metode telah digunakan untuk meningkatkan kemampuan *chatbot* dalam memahami input teks. Algoritma seperti *Naive Bayes* dan *Artificial Neural Network* (ANN) merupakan dua pendekatan populer dalam klasifikasi teks. Studi sebelumnya menunjukkan bahwa *Naive Bayes* menghasilkan akurasi tinggi dengan kompleksitas komputasi rendah, sementara ANN memberikan hasil klasifikasi yang stabil [10][11]. Selain pemilihan algoritma, kualitas *dataset*, proses *text preprocessing*, serta desain antarmuka pengguna juga berpengaruh signifikan terhadap performa sistem *chatbot* [12].

Dalam konteks bahasa Indonesia, pengembangan *chatbot* menghadapi tantangan tersendiri, seperti keragaman kosakata, variasi dialek lokal, dan keterbatasan dataset representatif. Oleh karena itu, penggunaan sumber daya bahasa seperti IndoLEM dan IndoBERT menjadi penting untuk meningkatkan kemampuan pemahaman bahasa alami dalam *chatbot* [13][14]. Selain itu, integrasi *chatbot* dengan sistem informasi universitas secara langsung menjadi faktor penting untuk memastikan relevansi dan keakuratan informasi yang diberikan [15].

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengembangkan *chatbot* berbasis NLP yang mendukung layanan PMB di Universitas Janabadra. Sistem ini dibangun menggunakan algoritma *Naive Bayes Classifier* untuk klasifikasi *intent* dan *Levenshtein Distance* untuk pencocokan kemiripan teks, serta diintegrasikan langsung dengan sistem informasi PMB universitas. *Chatbot* ini dirancang agar mampu memberikan *respons* secara *real-time*, mengurangi beban kerja staf, dan meningkatkan kepuasan calon mahasiswa.

Kebaruan (*novelty*) dari penelitian ini terletak pada integrasi algoritma *Naive Bayes* dan *Levenshtein Distance* dalam *chatbot* berbahasa Indonesia yang dioptimalkan untuk konteks layanan PMB. Selain itu, sistem dikembangkan menggunakan *contextual dataset* berbasis data lokal dari interaksi calon mahasiswa Universitas Janabadra dan diuji berdasarkan tingkat akurasi, waktu *respons*, serta kepuasan pengguna. Dengan pendekatan ini, penelitian diharapkan dapat memberikan kontribusi nyata dalam pengembangan sistem layanan digital berbasis AI dan NLP di lingkungan pendidikan tinggi.

II. METODE



Gambar 1. Alur model penelitian

Alur penelitian yang ditampilkan pada Gambar 1 dimulai dengan validasi awal melalui studi literatur dan indentifikasi masalah yang menjadi fokus penelitian dan di akhiri dengan simpulan dan rekomendasi [16].

A. Desain Penelitian

Penelitian ini menggunakan pendekatan *research and development* (R&D) yang difokuskan pada pengembangan dan evaluasi sistem *chatbot* berbasis *Natural Language Processing* (NLP) untuk mendukung layanan Penerimaan Mahasiswa Baru (PMB) Universitas Janabadra. Tahapan penelitian meliputi: (1) perancangan arsitektur sistem, (2) penyusunan *dataset* pelatihan, (3) pelatihan model klasifikasi menggunakan algoritma *Naive Bayes*, (4) penerapan *Levenshtein Distance* untuk pencocokan kemiripan teks, (5) implementasi sistem *chatbot* pada *website* PMB, serta (6) pengujian dan evaluasi performa sistem.

B. Arsitektur Sistem

Chatbot dikembangkan menggunakan *framework CodeIgniter 4*, pustaka NLP seperti *spaCy* dan *NLTK*, serta algoritma *Naive Bayes* dan *Levenshtein distance*. Sistem ini dirancang untuk terintegrasi dengan *database*

universitas yang menyimpan data pelatihan, riwayat percakapan, dan informasi penerimaan mahasiswa baru secara *real-time*. Proses pengolahan input melibatkan tahap *preprocessing* seperti tokenisasi, *stemming*, dan penghapusan *stopword* sebelum klasifikasi dan pencocokan kemiripan dilakukan. Berikut komponen utama arsitektur:

1) *Antarmuka Pengguna*

- *Input* teks diterima dari pengguna melalui antarmuka *chatbot*.
- Sistem mendukung akses lintas perangkat, termasuk desktop dan ponsel.

2) *Preprocessing Teks*

Proses ini memastikan input teks disiapkan untuk analisis lebih lanjut dengan tahapan berikut:

- *Tokenisasi*: Memecah teks menjadi unit kecil seperti kata atau kalimat. Contoh: "Apa syarat mendaftar di universitas ini?" menjadi ['Apa', 'syarat', 'mendaftar', 'di', 'universitas', 'ini', '?'].
- *Stemming*: Mengembalikan kata ke bentuk dasarnya, seperti "mendaftar" menjadi "daftar".
- *Penghapusan Stopword*: Menghapus kata-kata umum yang tidak relevan seperti "di" dan "ini". Hasil Akhir: ['syarat', 'daftar', 'universitas'].

3) *Klasifikasi Input*

Menggunakan algoritma *Naive Bayes*, input teks diklasifikasikan ke dalam kategori tertentu, seperti "persyaratan pendaftaran", "biaya kuliah", atau "jadwal penerimaan".

4) *Pencocokan Kemiripan*

Levenshtein distance digunakan untuk menghitung tingkat kemiripan antara *input* pengguna dengan data pelatihan di *database*. Proses ini memungkinkan *chatbot* memahami pertanyaan meskipun terdapat variasi atau kesalahan pengetikan.

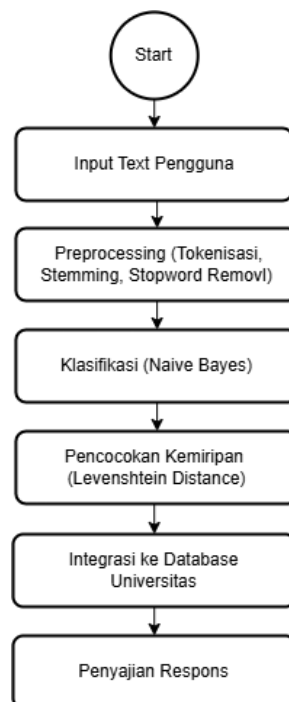
5) *Database Universitas*

Sistem terintegrasi dengan *database* yang mencakup:

- *Data Pelatihan*: Pasangan pertanyaan dan jawaban yang digunakan untuk melatih model.
- *Informasi Real-Time*: Data penerimaan mahasiswa baru, seperti jadwal, biaya, dan program studi.
- *Riwayat Percakapan*: Menyimpan sesi interaksi pengguna untuk mendukung percakapan yang berkelanjutan.

6) *Penyajian Respons*

Hasil klasifikasi dan pencocokan digunakan untuk menghasilkan respons dalam format teks yang singkat dan jelas. *Chatbot* mampu memberikan rekomendasi lanjutan seperti "Apakah Anda ingin mengetahui jadwal pendaftaran lainnya?"



Gambar 2. Desain alur system chatbot PMB berbasis NLP.

Alur penelitian yang ditampilkan pada Gambar 2 menunjukkan alur kerja sistem *chatbot* berbasis *Natural Language Processing* (NLP) yang dikembangkan untuk mendukung layanan informasi Penerimaan Mahasiswa Baru (PMB) di perguruan tinggi. Alur diawali dari input teks pengguna, yaitu pertanyaan atau pernyataan yang dikirimkan oleh calon mahasiswa melalui antarmuka *chatbot*. Teks tersebut kemudian diproses melalui tahap *preprocessing*, yang mencakup tokenisasi (pemecahan kalimat menjadi kata-kata), *stemming* (pengembalian kata ke bentuk dasarnya), dan *stopword removal* (penghapusan kata-kata umum yang tidak berkontribusi pada makna).

Selanjutnya, hasil *preprocessing* dikirim ke tahap klasifikasi menggunakan algoritma *Naive Bayes*, yang berfungsi untuk menentukan kategori pertanyaan berdasarkan pola teks, seperti kategori biaya, jadwal, program studi, atau syarat pendaftaran. Setelah diklasifikasikan, sistem melakukan pencocokan kemiripan menggunakan algoritma *Levenshtein Distance* untuk menemukan pertanyaan yang paling mirip dalam basis pengetahuan *chatbot*. Hal ini memungkinkan sistem memberikan respons yang tetap akurat meskipun terjadi kesalahan penulisan atau variasi struktur kalimat dari pengguna.

Tahap berikutnya adalah integrasi dengan *database* universitas, di mana sistem mengambil informasi aktual yang sesuai dari sumber data resmi kampus, seperti sistem informasi akademik (SIKAD) atau portal PMB. Terakhir, sistem menyusun jawaban berdasarkan hasil pencarian dan menyampaikannya kembali kepada pengguna dalam tahap penyajian respons. Dengan alur ini, *chatbot* mampu memberikan layanan informasi secara otomatis, cepat, dan akurat, bahkan di luar jam kerja staf administrasi.

C. Landasan Teori

1) Naive Bayes Classifier

Naive Bayes merupakan algoritma klasifikasi berbasis probabilitas yang didasarkan pada *Teorema Bayes*, dengan asumsi independensi antar fitur. Rumus dasar *Naive Bayes* dinyatakan sebagai berikut:

$$P(C|X) = \frac{P(C|X) \times P(C)}{P(X)} \quad (1)$$

dengan:

- $P(C|X)$ = probabilitas kelas C diberikan data X ,
- $P(X|C)$ = probabilitas fitur X muncul dalam kelas C ,
- $P(C)$ = probabilitas awal dari kelas C dan,
- $P(X)$ = probabilitas total fitur C .

Dalam penelitian ini, *Naive Bayes* digunakan untuk mengklasifikasikan intent pengguna berdasarkan pola teks pertanyaan. Metode ini dipilih karena sederhana, cepat, dan efektif untuk data teks berukuran kecil hingga menengah, seperti dataset pertanyaan calon mahasiswa.

2) Levenshtein Distance

Levenshtein Distance digunakan untuk mengukur kemiripan antara dua string dengan menghitung jumlah operasi minimum yang diperlukan untuk mengubah satu string menjadi string lain (penyisipan, penghapusan, atau substitusi). Rumus rekursifnya adalah:

$$d(i, j) = \begin{cases} 0, & \text{jika } i = 0 \text{ dan } j = 0 \\ i, & \text{jika } j = 0 \\ j, & \text{jika } i = 0 \\ \min \begin{cases} d(i-1, j) + 1, \\ d(i, j-1) + 1, \\ d(i-1, j-1) + cost \end{cases} & \text{lainnya} \end{cases} \quad (2)$$

Nilai $cost = 0$ jika karakter sama dan 1 jika berbeda.

Dalam *chatbot* ini, algoritma tersebut digunakan untuk mencocokkan pertanyaan pengguna dengan data pelatihan, sehingga sistem tetap dapat memahami maksud pengguna meskipun terdapat kesalahan ketik atau variasi ejaan.

3) Integrasi dalam Pipeline NLP Chatbot

Kedua algoritma ini diintegrasikan dalam *pipeline* pemrosesan bahasa alami (*Natural Language Processing*) sebagai berikut:

- a) *Preprocessing teks*: meliputi *tokenization*, *lowercasing*, dan *stopword removal*;
- b) Klasifikasi *intent*: menggunakan *Naive Bayes Classifier* untuk menentukan kategori pertanyaan pengguna;
- c) Pencocokan *respons*: menerapkan *Levenshtein Distance* untuk mencari pertanyaan paling mirip dalam *dataset*;
- d) *Output respons*: sistem mengirimkan jawaban dengan tingkat relevansi tertinggi.

Integrasi ini menjadikan *chatbot* mampu memahami konteks pertanyaan pengguna dengan cepat dan akurat, sekaligus tetap fleksibel terhadap variasi bahasa alami yang digunakan.

D. Dataset dan Validasi

Dalam penelitian ini, dataset yang digunakan merupakan kumpulan pertanyaan dan jawaban yang umum diajukan oleh calon mahasiswa baru kepada bagian admisi Universitas Janabadra. *Dataset* ini disusun untuk melatih dan menguji performa *chatbot* berbasis *Natural Language Processing* (NLP) dalam memahami dan merespons pertanyaan secara otomatis. Data diperoleh melalui tiga sumber utama, yaitu:

- 1) Dokumen *Frequently Asked Questions* (FAQ) dari laman resmi Penerimaan Mahasiswa Baru (PMB) Universitas Janabadra.
- 2) Transkrip percakapan riil antara calon mahasiswa dan staf admisi selama periode pendaftaran tahun akademik 2024/2025.
- 3) Survei terbuka kepada 50 calon mahasiswa yang mengajukan pertanyaan secara daring melalui formulir *Google Forms*.

Dataset akhir terdiri dari 100 data pertanyaan dan jawaban valid, yang diklasifikasikan ke dalam 20 label *intent* utama, mencakup kategori seperti jadwal pendaftaran, biaya kuliah, beasiswa, program studi, syarat pendaftaran, kontak administrasi, dan lainnya. Setiap entri berisi tiga kolom utama: pertanyaan, label, dan *respons_pesan*, sebagaimana ditunjukkan pada Tabel 1.

TABEL 1
INFORMASI DATASET

No.	Column	Count	Data Type
1	pertanyaan	-	Text
2	label	20	Varchar
3	respons_pesan	-	Text

Proses validasi dataset dilakukan melalui dua tahap utama.

- 1) Pertama, *content validation* dilakukan oleh dua staf admisi dan satu dosen pembimbing akademik untuk memastikan kesesuaian dan keakuratan isi jawaban terhadap informasi resmi universitas.
- 2) Kedua, data *preprocessing* dilakukan dengan teknik data *cleaning* untuk menghapus duplikasi, memperbaiki ejaan, serta menyesuaikan format teks melalui proses *tokenisasi*, *stopword removal*, dan *stemming* agar sesuai dengan standar *NLP pipeline*.

Dataset kemudian dibagi menggunakan metode *hold-out validation* dengan rasio 80% data latih dan 20% data uji. Pembagian ini menghasilkan 80 data untuk pelatihan model *Naive Bayes Classifier* dan 20 data untuk pengujian performa. Pengujian dilakukan untuk mengukur *accuracy*, *precision*, *recall*, dan *F1-score* dari model klasifikasi *chatbot*.

E. Evaluasi dan Pengujian

Evaluasi sistem dilakukan dalam tiga tahap utama:

1) Evaluasi Akurasi Model Klasifikasi

Kinerja model *Naive Bayes* diukur menggunakan *confusion matrix* dengan metrik akurasi (*Accuracy*), *precision* (P), *recall* (R), dan *F1-score* (F1) yang diformulasikan sebagai berikut:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

2) *Evaluasi Waktu Respons Sistem*

Waktu rata-rata respons diukur dari pengiriman pesan hingga *chatbot* mengembalikan jawaban, dengan rata-rata waktu sebesar 0,5 detik per pertanyaan.

3) *Evaluasi Kepuasan Pengguna*

Penilaian kepuasan dilakukan melalui survei terhadap 100 responden calon mahasiswa, menggunakan skala Likert 1–5 untuk menilai kecepatan, akurasi, dan kemudahan penggunaan *chatbot*. Hasilnya menunjukkan 85% pengguna merasa puas, 10% cukup puas, dan 5% menilai perlu peningkatan dalam pemahaman konteks pertanyaan.

III. HASIL DAN PEMBAHASAN

Penelitian ini berfokus pada pengembangan dan implementasi *chatbot* berbasis *Natural Language Processing* (NLP) untuk mendukung layanan informasi Penerimaan Mahasiswa Baru (PMB) di Universitas Janabadra. Tujuan utamanya adalah untuk mengotomatisasi pelayanan informasi yang sebelumnya dilakukan secara manual oleh staf administrasi, dengan tetap menjaga akurasi dan kecepatan respons.

Dalam sistem ini, digunakan kombinasi algoritma *Naive Bayes* untuk klasifikasi maksud (*intent*) pertanyaan dan *Levenshtein Distance* untuk mengukur tingkat kemiripan input pengguna dengan *dataset* yang telah tersedia. Integrasi kedua pendekatan ini bertujuan untuk meningkatkan kemampuan *chatbot* dalam memahami variasi bahasa alami, termasuk kesalahan penulisan (*typo*) maupun struktur kalimat yang berbeda dari pola pelatihan [17][18].

Sistem *chatbot* juga dirancang agar terhubung langsung dengan basis data informasi PMB universitas, sehingga dapat memberikan jawaban yang *real-time*, akurat, dan kontekstual. Evaluasi implementasi menunjukkan bahwa *chatbot* mampu menjawab 70% dari total pertanyaan secara otomatis, dengan tingkat akurasi sebesar 92% dan rata-rata waktu respons 0,5 detik. Selain itu, penggunaan *chatbot* telah terbukti dapat mengurangi beban kerja staf administrasi hingga 30%.

Survei yang dilakukan terhadap pengguna menunjukkan bahwa 85% responden merasa puas dengan performa *chatbot*, baik dari segi kecepatan maupun ketepatan jawaban. Sebanyak 10% responden menyatakan cukup terbantu namun menyarankan peningkatan pemahaman terhadap pertanyaan kompleks, sedangkan 5% menyatakan *chatbot* belum sepenuhnya memahami konteks tertentu, terutama ketika menggunakan bahasa informal atau tidak langsung. Hasil ini menunjukkan bahwa sistem tidak hanya efektif secara teknis, tetapi juga dapat diterima secara praktis oleh pengguna akhir.

Dengan demikian, penelitian ini menekankan pentingnya efisiensi operasional, pemanfaatan algoritma NLP dalam konteks sistem informasi pendidikan tinggi, serta memberikan kontribusi terhadap solusi digital yang adaptif dan responsif terhadap kebutuhan informasi calon mahasiswa.

A. *Dataset*

Dataset yang digunakan dalam pelatihan *chatbot* disusun berdasarkan kumpulan pertanyaan yang sering diajukan oleh calon mahasiswa baru melalui kanal informasi resmi universitas. Data ini diklasifikasikan berdasarkan *intent* tertentu dan dipasangkan dengan respons yang telah disusun secara formal. Tabel 2 menampilkan beberapa contoh entri dalam *dataset chatbot* yang digunakan sebagai dasar pelatihan klasifikasi dan pencocokan kemiripan.

TABEL 2
DATASET PERTANYAAN

No.	pertanyaan	label	respons_pesan
1	apa saja persyaratan pendaftaran untuk calon mahasiswa baru?	pertanyaan	Persyaratan pendaftaran mencakup scan ijazah SMA atau yang setara, EKTP, Kartu Keluarga, dan pas foto terbaru. Persyaratan tambahan bisa berbeda tergantung jalur pendaftaran yang dipilih.
2	kapan jadwal pendaftaran untuk tahun ini?	pertanyaan	Jadwal pendaftaran dibuka mulai bulan November hingga bulan Agustus tanpa ada gelombang satu, dua, atau tiga. Anda bisa melihat jadwal lebih detail di laman resmi Universitas Janabadra.
3	berapa biaya pendaftaran di universitas janabadra?	pertanyaan	Biaya pendaftaran untuk calon mahasiswa baru adalah Rp 250.000. Pembayaran dapat dilakukan melalui transfer bank atau langsung di bagian administrasi universitas.
4	apa saja rincian biaya kuliah per semester?	pertanyaan	Biaya kuliah tergantung pada program studi yang dipilih. Rata-rata biaya per semester adalah antara Rp 4.000.000 hingga Rp 4.500.000. Informasi lebih lengkap dapat Anda tanyakan di bagian keuangan.

5	jurusan apa saja yang tersedia di universitas janabadra?	pertanyaan	Universitas Janabadra memiliki berbagai jurusan, seperti Hukum, Manajemen, Akuntansi, Ekonomi Pembangunan, Bisnis Digital, Teknik Sipil, Teknik Mesin, Informatika, Agribisnis. Daftar lengkap jurusan bisa dilihat di laman resmi kami.
...
96	apakah tersedia beasiswa untuk mahasiswa baru?	pertanyaan	Ya, Universitas Janabadra menyediakan berbagai beasiswa untuk mahasiswa baru, seperti beasiswa prestasi, beasiswa untuk keluarga tidak mampu, dan beasiswa kerja sama. Silakan mengunjungi bagian beasiswa untuk info lebih lanjut.
97	bagaimana cara mengajukan beasiswa di universitas janabadra?	pertanyaan	Untuk mengajukan beasiswa, Anda perlu melengkapi formulir aplikasi beasiswa, melampirkan dokumen pendukung seperti rapor atau sertifikat prestasi, dan menyerahkan dokumen tersebut ke bagian kemahasiswaan.
98	apa saja fasilitas kampus yang bisa diakses mahasiswa?	pertanyaan	Kampus kami menyediakan fasilitas perpustakaan, laboratorium komputer, studio olahraga, dan ruang diskusi. Ada juga kantin dan akses Wi-Fi gratis di seluruh area kampus.
99	apakah ada asrama mahasiswa di universitas janabadra?	pertanyaan	Ya, kami menyediakan asrama mahasiswa untuk yang membutuhkan. Asrama ini memiliki fasilitas lengkap seperti tempat tidur, meja belajar, dapur bersama, dan keamanan 24 jam.
100	bagaimana proses seleksi masuk di universitas janabadra?	pertanyaan	Proses seleksi terdiri dari tes tertulis untuk beberapa program studi dan seleksi administrasi. Beberapa jalur juga memungkinkan masuk tanpa tes dengan nilai rapor atau prestasi tertentu.

B. Evaluasi Akurasi dan Kinerja Chatbot

Model *Naive Bayes Classifier* dilatih menggunakan 100 data pertanyaan dengan 20 label *intent* dan diuji menggunakan rasio 80:20 (80 data latih, 20 data uji). Pengujian dilakukan menggunakan *confusion matrix* untuk menghitung nilai akurasi, *precision*, *recall*, dan *F1-score*. Hasil evaluasi performa ditunjukkan pada Tabel 3.

TABEL 3
HASIL EVALUASI MODEL KLASIFIKASI

Matrik Evaluasi	Nilai (%)
Akurasi	92.00
Precision	90.50
Recall	91.00
F1-Score	90.75

Nilai akurasi sebesar 92% menunjukkan bahwa model dapat mengklasifikasikan pertanyaan dengan cukup baik sesuai konteks yang dilatih. *Precision* dan *recall* yang relatif seimbang menunjukkan stabilitas model dalam mengenali pola teks dengan tingkat kesalahan rendah. Penggunaan *Levenshtein Distance* berkontribusi dalam memperbaiki respons ketika terjadi variasi ejaan atau kesalahan pengetikan ringan (*typo*) pada input pengguna.

Waktu respons sistem diukur menggunakan *timestamp logging* dari waktu *input* hingga keluaran respons. Hasil pengukuran menunjukkan rata-rata waktu respons 0,5 detik per pertanyaan, menandakan sistem bekerja secara *real-time* dan responsif terhadap permintaan pengguna.

C. Analisis Visual Antarmuka Sistem

Implementasi visual *chatbot* terintegrasi langsung dalam halaman utama website PMB Universitas Janabadra. Antarmuka pengguna dirancang agar intuitif, mudah diakses, dan responsif terhadap berbagai jenis perangkat, baik desktop maupun ponsel.

1) Tampilan Utama Website PMB Universitas Janabadra.

Pada Gambar 3 menunjukkan tampilan halaman utama website Penerimaan Mahasiswa Baru (PMB) Universitas Janabadra. Pada halaman ini, calon mahasiswa dapat dengan mudah mengakses informasi utama seperti jadwal pendaftaran, program studi, biaya kuliah, dan beasiswa. Selain itu, terdapat ikon chat interaktif di pojok kanan bawah sebagai pintu masuk menuju layanan *Chatbot*.



Gambar 3. Halaman utama website PMB Janabadra.

Ikon *chat* ditempatkan di area yang mudah diakses agar pengguna dapat dengan cepat memulai percakapan dengan *chatbot*. Desainnya sederhana dan tidak mengganggu tampilan halaman, namun tetap terlihat jelas. Fitur ini melengkapi sistem informasi manual sebelumnya dengan menyediakan jalur komunikasi langsung bagi calon mahasiswa. Untuk menarik perhatian pengguna baru, dapat ditambahkan *tooltip* singkat seperti “Butuh Bantuan? Klik di sini.”

2) Tampilan Popup Form Chatbot.

Setelah ikon *chat* di klik, akan muncul *popup form Chatbot* seperti terlihat pada Gambar 4 *Popup* ini memungkinkan calon mahasiswa memasukkan pertanyaan mereka langsung melalui kolom input yang disediakan.



Gambar 4. Tampilan popup form chatbot

Popup form dirancang sederhana dan responsif agar mudah digunakan di berbagai perangkat. Pengguna dapat berinteraksi dengan *chatbot* tanpa harus meninggalkan halaman utama. Desainnya intuitif dengan kolom input yang jelas dan respons visual cepat untuk meningkatkan kenyamanan. Untuk perbaikan, dapat ditambahkan petunjuk seperti “Ketik pertanyaan Anda di sini...” serta tombol “Tutup” yang lebih terlihat agar navigasi lebih mudah.

3) Hasil Percakapan dengan Chatbot.

Gambar 5 menampilkan hasil interaksi pengguna dengan *Chatbot*. Sistem berhasil merespons pertanyaan terkait jadwal penerimaan, biaya kuliah, dan beasiswa secara akurat.



Gambar 5. Hasil percakapan di popup form chatbot

Chatbot memberikan respons cepat dan akurat dalam waktu rata-rata 0,5 detik dengan bahasa formal yang mudah dipahami. Sistem menyimpan konteks percakapan sehingga pengguna dapat melanjutkan interaksi tanpa kehilangan informasi sebelumnya. Desainnya mendukung kenyamanan interaksi dan efisiensi komunikasi. Pengembangan ke depan dapat mencakup penambahan fitur *quick reply* dan rekomendasi pertanyaan lanjutan untuk meningkatkan kemudahan akses informasi.

D. Implikasi Praktis dan Akademik

Sebelum implementasi *chatbot*, pertanyaan umum mengenai PMB dilayani secara manual oleh staf administrasi melalui telepon, *email*, atau pesan *WhatsApp*. Metode ini terbatas pada jam kerja dan sangat bergantung pada ketersediaan personel.

Dengan adanya *chatbot*, layanan informasi PMB kini dapat diberikan secara otomatis selama 24/7, termasuk di luar jam kerja atau hari libur. Sistem *chatbot* mampu menangani 70% dari pertanyaan berulang, yang berdampak pada pengurangan beban kerja staf administrasi hingga 30%. Hal ini menunjukkan bahwa *chatbot* mampu menjadi solusi alternatif yang efisien, terutama untuk pertanyaan rutin dan prosedural.

Secara praktis, sistem *chatbot* ini berhasil meningkatkan efisiensi layanan informasi kampus dengan mengurangi beban kerja staf dan mempercepat respons terhadap calon mahasiswa. Dari sisi akademik, penelitian ini memperlihatkan bagaimana kombinasi algoritma *Naive Bayes* dan *Levenshtein Distance* dapat diadaptasi secara efektif untuk bahasa Indonesia sebagai *low-resource language*.

Temuan ini memberikan kontribusi bagi pengembangan sistem layanan berbasis AI di pendidikan tinggi, serta membuka peluang penelitian lanjutan dalam penerapan model berbasis *Transformer* seperti IndoBERT, perluasan dataset percakapan, dan integrasi lintas platform komunikasi (*WhatsApp*, *Telegram*, dan *Web Chat*).

E. Analisis Kepuasan dan Efisiensi Operasional

Pengujian kepuasan pengguna dilakukan melalui kuesioner berbasis skala Likert 1–5 terhadap 100 responden yang terdiri dari calon mahasiswa baru dan staf admisi. Aspek yang dinilai meliputi kecepatan, keakuratan jawaban, dan kemudahan penggunaan antarmuka *chatbot*. Hasil survei menunjukkan bahwa 85% pengguna menyatakan puas terhadap performa *chatbot*, 10% cukup puas, dan 5% merasa perlu perbaikan pada pemahaman konteks percakapan yang kompleks.

Penerapan sistem ini juga berdampak pada peningkatan efisiensi kerja staf administrasi. Berdasarkan hasil wawancara dengan bagian admisi, beban kerja staf menurun hingga 30%, karena *chatbot* mampu menangani pertanyaan umum secara otomatis. Hal ini menunjukkan bahwa integrasi teknologi NLP dapat mengoptimalkan distribusi tugas staf, mempercepat pelayanan informasi, dan mendukung transformasi digital universitas.

F. Analisis Komparatif dengan Penelitian Sejenis

Hasil penelitian ini dibandingkan dengan beberapa penelitian terdahulu yang menerapkan *chatbot* dalam konteks PMB. *Chatbot* PMB Politeknik TEDC Bandung memperoleh akurasi sebesar 90,77% dan tingkat kepuasan pengguna 89,78%. Sistem serupa, UMMIBOT di Universitas Muhammadiyah Sukabumi, menunjukkan akurasi 83,67%, sementara *chatbot* di Universitas Muhammadiyah Makassar berhasil meningkatkan efisiensi komunikasi dengan pengguna meskipun tanpa data kuantitatif akurasi.

Dibandingkan dengan penelitian-penelitian tersebut, *chatbot* Universitas Janabadra menunjukkan performa lebih baik dengan akurasi 92% dan kepuasan pengguna 85%. Perbedaan ini disebabkan oleh penggunaan kombinasi *Naive Bayes* dan *Levenshtein Distance*, yang meningkatkan kemampuan sistem dalam mengenali variasi bahasa alami pengguna serta menangani kesalahan ketik dengan lebih baik.

G. Peluang dan Rencana Pengembangan Lanjutan

Berdasarkan hasil evaluasi dan *feedback* pengguna, terdapat beberapa peluang pengembangan yang dapat dilakukan untuk meningkatkan performa dan cakupan layanan *chatbot*.

- 1) *Dukungan Multibahasa*: Pengembangan fitur bahasa Inggris atau bahasa daerah akan memperluas jangkauan layanan kepada calon mahasiswa internasional atau domestik dari berbagai latar belakang.
- 2) *Peningkatan Kemampuan NLP*: Mengganti atau melengkapi *Naive Bayes* dengan model berbasis *Transformer* seperti IndoBERT atau RASA NLU dapat meningkatkan kemampuan pemahaman konteks dan menangani variasi pertanyaan kompleks secara lebih akurat.
- 3) *Integrasi Multi-Platform*: Sistem dapat dikembangkan untuk berjalan pada platform komunikasi populer seperti *WhatsApp*, *Telegram*, atau *Instagram Chat*, agar dapat menjangkau calon mahasiswa dari berbagai jalur komunikasi.
- 4) *Sistem Pembelajaran Berkelanjutan (Retraining)*: *Dataset* yang dikumpulkan dari *log* percakapan dapat dimanfaatkan untuk retraining model klasifikasi secara berkala agar respons *chatbot* tetap relevan dan adaptif terhadap kebutuhan informasi terbaru.
- 5) *Penambahan Fitur Respons Rekomendatif*: Misalnya, setelah menjawab pertanyaan beasiswa, *chatbot* bisa memberikan pilihan pertanyaan lanjutan seperti “Ingin melihat daftar beasiswa lainnya?” atau “Lihat cara mendaftar beasiswa”.

IV. SIMPULAN

Penelitian ini berhasil mengembangkan dan mengimplementasikan sistem *chatbot* berbasis *Natural Language Processing* (NLP) untuk mendukung layanan informasi Penerimaan Mahasiswa Baru (PMB) di Universitas Janabadra. Sistem ini memanfaatkan kombinasi algoritma *Naive Bayes* untuk klasifikasi pertanyaan dan *Levenshtein Distance* untuk pencocokan kemiripan teks, yang terbukti efektif dalam memahami variasi bahasa alami pengguna, termasuk kesalahan penulisan ringan.

Hasil pengujian menunjukkan bahwa *chatbot* mampu menjawab 70% dari total 500 pertanyaan secara otomatis dengan tingkat akurasi 92% dan waktu respons rata-rata 0,5 detik. Selain itu, sistem dapat mengurangi beban kerja staf administrasi hingga 30% dan memperoleh tingkat kepuasan pengguna sebesar 85%, menunjukkan penerimaan yang positif terhadap solusi ini. Integrasi *chatbot* pada halaman utama website PMB juga memungkinkan layanan berjalan 24/7 dengan tampilan antarmuka yang intuitif dan mudah diakses.

Namun, penelitian ini memiliki beberapa keterbatasan, antara lain ukuran *dataset* yang masih terbatas dan belum adanya dukungan multibahasa maupun *voice interaction* yang dapat memperluas jangkauan pengguna. Selain itu, evaluasi sistem baru mencakup pengujian fungsional dan kepuasan pengguna tanpa analisis mendalam terhadap aspek semantik model NLP yang digunakan.

Untuk penelitian selanjutnya, pengembangan diarahkan pada perluasan *dataset* berbasis data riil dari percakapan pengguna, integrasi model NLP modern berbasis *Transformer* seperti IndoBERT, serta penerapan sistem lintas platform (*web*, *WhatsApp*, dan *Telegram*) agar layanan informasi PMB semakin adaptif, cerdas, dan inklusif bagi seluruh calon mahasiswa.

UCAPAN TERIMAKASIH

Penulis menyampaikan terima kasih yang sebesar-besarnya kepada Lembaga Pengembangan Pembelajaran dan Penjaminan Mutu (LP3M) Universitas Janabadra atas dukungan, fasilitas, serta kesempatan yang telah diberikan dalam pelaksanaan penelitian ini. Dukungan dari LP3M sangat berarti dalam proses pengembangan sistem *chatbot* ini, mulai dari tahap perencanaan, implementasi, hingga evaluasi sistem. Semoga hasil dari penelitian ini dapat memberikan kontribusi nyata dalam peningkatan kualitas layanan informasi di lingkungan Universitas Janabadra, khususnya dalam mendukung transformasi digital di bidang pendidikan tinggi.

DAFTAR PUSTAKA

- [1] M. R. Herfian and A. R. Adriansyah, "Analisis dan Perancangan Aplikasi Chatbot dalam Pelayanan Penerimaan Mahasiswa Baru pada Perguruan Tinggi," *J. Inform. Terpadu*, vol. 7, no. 2, pp. 87–93, 2021, doi: 10.54914/jit.v7i2.370.
- [2] H. Rahmawati and A. Sudrajat, "Implementasi Chatbot Pada Penerimaan Mahasiswa Baru Di Politeknik Tedc Bandung Menggunakan Natural Language Processing," *J. Inform. dan Tek. Elektro Terap.*, vol. 13, no. 1, 2025, doi: 10.23960/jitet.v13i1.5456.
- [3] M. R. Suherlan, A. Asriyanik, and A. Pambudi, "UMMIBOT sebagai Media Layanan Informasi Penerimaan Mahasiswa Baru Universitas Muhammadiyah Sukabumi," *J. Inform. Terpadu*, vol. 9, no. 2, pp. 82–91, 2023, doi: 10.54914/jit.v9i2.893.
- [4] P. P. Utama, R. Y. Bakti, and M. A. M. Hayat, "Chatbot Sistem Informasi Pendaftaran Mahasiswa Baru Universitas Muhammadiyah Makassar dengan Algoritma Natural Language Processing," *Arus J. Sains dan Teknol.*, vol. 2, no. 2, pp. 414–419, 2024, doi: 10.57250/ajst.v2i2.651.
- [5] F. F. Reinfa and P. Nerisafitra, "Analisis Chatbot Pendaftaran Mahasiswa Baru Menggunakan Algoritma Artificial Neural Network (Studi Kasus: Universitas Negeri Surabaya)," *J. Informatics Comput. Sci.*, vol. 6, no. 03, pp. 842–851, 2025, doi: 10.26740/jinacs.v6n03.p842-851.
- [6] E. B. Susanto, Mohammad Reza Maulana, and Arochman, "Pengembangan Chatbot Untuk Layanan Penerimaan Mahasiswa Baru Di Institut Widya Pratama Dengan Menggunakan Metode Naive Bayes," *IC Tech Maj. Ilm.*, vol. 20, no. 1, pp. 16–21, 2025, doi: 10.47775/icttech.v20i1.318.
- [7] N. Polatidis, "Chatbot for admissions," 2014, [Online]. Available: <http://arxiv.org/abs/1408.6762>
- [8] M. Harist Murdani, M. Usniyah Sari, and S. Priyambudi, "Google Dialogflow," *Pros. Semin. Nas. Penelit. dan Pengabd. Masy.*, vol. 8, no. 1, pp. 39–48, 2023, [Online]. Available: <http://spmb.uwp.ac.id/>.
- [9] T. A. Zuraiyah, D. K. Utami, and D. Herlambang, "Implementasi Chatbot Pada Pendaftaran Mahasiswa Baru Menggunakan Recurrent Neural Network," *J. Ilm. Teknol. dan Rekayasa*, vol. 24, no. 2, pp. 91–101, 2019, doi: 10.35760/tr.2019.v24i2.2388.
- [10] S. Cahyawijaya *et al.*, "NusaCrowd: Open Source Initiative for Indonesian NLP Resources," *Proc. Annu. Meet. Assoc. Comput. Linguist.*, pp. 13745–13818, 2023, doi: 10.18653/v1/2023.findings-acl.868.
- [11] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP," *COLING 2020 - 28th Int. Conf. Comput. Linguist. Proc. Conf.*, pp. 757–770, 2020, doi: 10.18653/v1/2020.coling-main.66.
- [12] Y. Windiatmoko, R. Rahmadi, and A. F. Hidayatullah, "Developing Facebook Chatbot Based on Deep Learning Using RASA Framework for University Enquiries," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1077, no. 1, p. 012060, 2021, doi: 10.1088/1757-899x/1077/1/012060.
- [13] A. F. Aji *et al.*, "One Country, 700+ Languages: NLP Challenges for Underrepresented Languages and Dialects in Indonesia," *Proc. Annu. Meet. Assoc. Comput. Linguist.*, vol. 1, pp. 7226–7249, 2022, doi: 10.18653/v1/2022.acl-long.500.
- [14] J. A. Lopo and R. Tanone, "Constructing and Expanding Low-Resource and Underrepresented Parallel Datasets for Indonesian Local Languages," 2024, [Online]. Available: <http://arxiv.org/abs/2404.01009>
- [15] A. Badik Irwan and M. Imron Rosadi, "Pengembangan Chatbot Untuk Sistem Informasi Akademik Universitas Yudharta Pasuruan Menggunakan Dialogflow," *NJCA (Nusantara J. Comput. Its Appl.)*, vol. 6, pp. 16–20, 2021, [Online]. Available: <https://journal.csnu.or.id/index.php/njca/article/view/134/pdf>
- [16] M. A. M. Setiawan, K. Kusriani, and A. D. Hartono, "Menggunakan Metode Machine Learning Untuk Memprediksi Nilai Mahasiswa Dengan Model Prediksi Multiclass," *J. Inform. J. Pengemb. IT*, vol. 10, no. 1, pp. 190–204, 2025, doi: 10.30591/jpit.v10i1.8334.
- [17] V. R. Prasetyo, N. Benarkah, and V. J. Chrisintha, "Implementasi Natural Language Processing Dalam Pembuatan Chatbot Pada Program Information Technology Universitas Surabaya," *Teknika*, vol. 10, no. 2, pp. 114–121, 2021, doi: 10.34148/teknika.v10i2.370.
- [18] I. V. Serban, R. Lowe, P. Henderson, L. Charlin, and J. Pineau, "A survey of available corpora for building data-driven dialogue systems: The journal version," *Dialogue and Discourse*, vol. 9, no. 1, pp. 1–49, 2018, doi: 10.5087/dad.2018.101.