

ALGORITMA NAÏVE BAYES PADA CHATBOT WHATSAPP UNTUK LAYANAN INFORMASI PPDB DI MTS ATTAQWA 04 BEKASI

Ade Yusuf Ardiyansyah^{1*}, Nadya Safitri²

^{1,2}Program Studi Teknik Informatika, Fakultas Informatika dan Desain, Universitas Bina Insani
Jl. Siliwangi No. 6, Sepanjang Jaya, Rawalumbu, Kota Bekasi, Jawa Barat 17114

Keywords: Chatbot; Naïve Bayes; Natural Language Processing; PPDB; WhatsApp.

Correspondent Email:

adeyusufardiyansyah01@gmail.com

Abstrak. Penerimaan Peserta Didik Baru (PPDB) di MTs Attaqwa 04 Kabupaten Bekasi saat ini masih dilayani secara manual melalui WhatsApp oleh sekitar sepuluh panitia yang merangkap mengajar, dengan intensitas pertanyaan 8–10 pesan per hari dan dapat meningkat dua kali lipat pada periode tertentu. Kondisi ini menyebabkan keterlambatan respons dan ketidaktersediaan layanan 24 jam. Penelitian ini bertujuan merancang dan mengimplementasikan chatbot WhatsApp berbasis Natural Language Processing (NLP) dengan algoritma Multinomial Naïve Bayes untuk klasifikasi intent pertanyaan calon pendaftar. Sistem dikembangkan menggunakan metode Extreme Programming (XP) dengan integrasi WhatsApp HTTP API (WAHA), backend Python-Flask, ekstraksi fitur Term Frequency–Inverse Document Frequency (TF-IDF), dan pipeline praproses yang mencakup case folding, cleansing, normalisasi berbasis aturan, serta fuzzy correction. Model dilatih menggunakan kurang lebih 1.000 data intent dan diuji pada 223 data uji holdout. Hasil pengujian menunjukkan akurasi klasifikasi sebesar 85,65%, sedangkan pengujian fungsional terhadap 14 skenario seluruhnya berhasil. Pengujian User Acceptance Testing (UAT) terhadap 28 responden menghasilkan skor rata-rata 4,18–4,43 pada skala Likert 1–5, dengan kepuasan keseluruhan 4,39. Sistem dinilai mampu memberikan layanan informasi PPDB secara otomatis, cepat, dan tersedia 24 jam.



Copyright © 2026 The Author(s),
Published by Jurnal Informatika dan
Teknik Elektro Terapan. This is an
open-access article distributed under
the terms of the Creative Commons
Attribution-NonCommercial 4.0
International License (CC BY-NC
4.0).

Abstract. The New Student Admission (PPDB) process at MTs Attaqwa 04 Bekasi Regency is currently handled manually through WhatsApp by about ten committee members who also serve as teachers, receiving 8–10 daily queries that may double during certain periods. This condition leads to delayed responses and the unavailability of 24-hour service. This research aims to design and implement a WhatsApp chatbot based on Natural Language Processing (NLP) using the Multinomial Naïve Bayes algorithm for classifying applicant question intents. The system was developed using the Extreme Programming (XP) method, integrated with the WhatsApp HTTP API (WAHA), a Python-Flask backend, Term Frequency–Inverse Document Frequency (TF-IDF) feature extraction, and a preprocessing pipeline including case folding, cleansing, rule-based normalization, and fuzzy correction. The model was trained on approximately 1,000 intent data and tested on 223 holdout samples. Test results showed a classification accuracy of 85.65%, while functional testing of 14 scenarios was entirely successful. The User Acceptance Testing (UAT) with 28 respondents yielded average scores between 4.18–4.43 on a 1–5 Likert scale, with overall satisfaction of

4.39. The system effectively delivers automated, fast, and 24-hour PPDB information services.

1. PENDAHULUAN

Penerimaan Peserta Didik Baru (PPDB) merupakan tahapan awal pelayanan sekolah kepada masyarakat sekaligus bagian penting dari manajemen pendidikan. Proses ini tidak hanya berkaitan dengan penerimaan siswa, tetapi juga mencerminkan kualitas layanan dan akses informasi yang diberikan lembaga pendidikan kepada calon peserta didik beserta orang tua [1]. Ketersediaan informasi yang mudah dipahami menjadi faktor penting agar proses pendaftaran dapat berjalan lancar. Dalam perkembangannya, chatbot telah menjadi solusi populer untuk menyediakan informasi secara cepat dan otomatis pada layanan PPDB karena mampu meningkatkan efisiensi, mempercepat respons, dan menyajikan informasi yang akurat [2].

MTs Attaqwa 04 merupakan salah satu lembaga pendidikan menengah berbasis Islam di Kabupaten Bekasi yang berada di bawah naungan Yayasan Attaqwa. Setiap tahun sekolah ini menyelenggarakan PPDB untuk menerima calon siswa baru. Berdasarkan hasil observasi dan wawancara, seluruh pertanyaan calon siswa maupun orang tua masih dilayani secara manual melalui aplikasi WhatsApp. Penggunaan WhatsApp sebagai media komunikasi sejalan dengan temuan [3] yang menyatakan bahwa WhatsApp telah menjadi sarana komunikasi daring utama lembaga pendidikan dalam menyampaikan informasi PPDB.

Dalam praktiknya, panitia PPDB yang berjumlah sekitar sepuluh orang harus mengetik balasan satu per satu, sementara mereka juga memiliki tugas utama mengajar dan menjalankan kegiatan administrasi. Intensitas pertanyaan mencapai 8–10 pesan per hari dan dapat meningkat dua kali lipat pada periode tertentu seperti bulan Ramadan, dengan total sekitar 500 calon pendaftar dalam satu periode PPDB. Kondisi tersebut menyebabkan keterlambatan respons, kesalahan pengetikan (typo) yang berpotensi menimbulkan kesalahpahaman informasi, serta hilangnya momentum calon pendaftar untuk segera mendaftar.

Permasalahan serupa juga ditemukan pada [4] dan [5], yang menegaskan bahwa institusi pendidikan sering menghadapi keterbatasan staf serta waktu operasional yang tidak mampu melayani 24 jam. Implementasi chatbot berbasis Natural Language Processing (NLP) menjadi alternatif yang krusial karena mampu memahami maksud (intent) pertanyaan dan memberikan respons otomatis tanpa terikat jam kerja, sekaligus meringankan beban administratif panitia.

Beberapa penelitian terdahulu menunjukkan bahwa algoritma Naïve Bayes banyak digunakan dalam pengembangan chatbot karena kemampuannya melakukan klasifikasi teks. Penelitian [6] melaporkan tingkat akurasi 88% pada chatbot PMB Institut Widya Pratama, sedangkan [7] memperoleh akurasi 88,9% pada chatbot informasi magang dengan dataset terbatas. Namun, sebagian besar penelitian tersebut diterapkan pada institusi pendidikan umum seperti SMK dan perguruan tinggi, sementara implementasi chatbot berbasis WhatsApp untuk layanan informasi PPDB pada lingkungan madrasah dengan memanfaatkan dataset percakapan nyata masih sangat terbatas. Hal inilah yang menjadi celah penelitian (research gap) yang diangkat pada studi ini.

Berdasarkan kondisi tersebut, penelitian ini bertujuan untuk merancang dan mengimplementasikan sistem chatbot WhatsApp berbasis NLP dengan algoritma Multinomial Naïve Bayes sebagai layanan informasi PPDB di MTs Attaqwa 04. Sistem diharapkan mampu memberikan respons otomatis terhadap pertanyaan yang sering diajukan calon pendaftar, mengurangi keterlambatan respons, serta menyediakan layanan 24 jam tanpa bergantung pada ketersediaan petugas.

2. TINJAUAN PUSTAKA

2.1. Penerimaan Peserta Didik Baru (PPDB)

PPDB merupakan tahap awal manajemen pelayanan sekolah kepada masyarakat dan menjadi sarana utama penyampaian informasi kepada calon peserta didik sejak tahap

pendaftaran [1]. Pelaksanaan PPDB mengacu pada regulasi pendidikan nasional, dan penerapan PPDB berbasis online bertujuan meningkatkan transparansi, keadilan, serta pemerataan akses pendidikan [8].

2.2. Layanan Informasi

Layanan informasi merupakan kegiatan pemberian bantuan agar individu memperoleh pengetahuan dan pemahaman yang diperlukan dalam mengambil keputusan [9]. Layanan ini dapat disampaikan melalui berbagai media, termasuk media sosial, sehingga lebih mudah diakses [10].

2.3. Chatbot

Chatbot adalah asisten virtual interaktif yang dirancang untuk membantu pengguna dengan menjawab berbagai pertanyaan berdasarkan basis pengetahuan yang telah diintegrasikan [11]. Chatbot juga merupakan agen cerdas yang mampu meniru kemampuan manusia dalam melakukan percakapan dengan pengguna [12].

2.4. WhatsApp dan WAHA

WhatsApp merupakan platform perpesanan instan yang memfasilitasi komunikasi lintas waktu dan ruang dan menjadi salah satu aplikasi paling populer di Indonesia [13]. WAHA (WhatsApp HTTP API) adalah tools open-source yang dapat digunakan untuk mengirim dan menerima pesan WhatsApp melalui API serta mendukung otomatisasi dan integrasi dengan berbagai sistem [14]. WAHA berfungsi sebagai gateway yang memungkinkan integrasi sistem dengan aplikasi WhatsApp melalui protokol HTTP, sehingga memudahkan pengiriman dan penerimaan pesan secara otomatis [15].

2.5. Natural Language Processing (NLP)

NLP merupakan teknik pemrosesan bahasa alami yang digunakan untuk menganalisis teks. Dalam pengembangannya, NLP berperan penting dalam meningkatkan kinerja dan akurasi model klasifikasi melalui serangkaian prosedur prapemrosesan data seperti pembersihan teks, penormalan huruf, penghapusan kata umum, hingga pengubahan kata ke bentuk dasarnya [16]. Tahap prapemrosesan teks menjadi langkah awal yang penting sebelum dilakukan proses pengolahan lebih lanjut [17].

2.6. Naïve Bayes

Naïve Bayes merupakan metode pengklasifikasian statistik yang digunakan untuk memprediksi probabilitas keanggotaan suatu kelas. Metode ini bekerja dengan menghitung sekumpulan nilai probabilitas melalui penjumlahan frekuensi dan kombinasi nilai dari dataset [16]. Algoritma ini didasarkan pada Teorema Bayes dan digunakan untuk memprediksi peluang berdasarkan pengalaman sebelumnya [18]. Persamaan dasar Teorema Bayes ditunjukkan pada Persamaan (1).

$$P(C_i|X) = P(X|C_i) \cdot P(C_i) / P(X) \quad (1)$$

Keterangan: X adalah data dengan kelas yang belum diketahui; C_i adalah hipotesis data merupakan kelas spesifik; $P(C_i|X)$ adalah probabilitas posterior; $P(C_i)$ adalah probabilitas prior; $P(X|C_i)$ adalah likelihood; dan $P(X)$ adalah probabilitas X. Naïve Bayes memiliki keunggulan kecepatan klasifikasi karena hanya melibatkan perhitungan probabilitas berdasarkan parameter statistik hasil pelatihan [19]. Sebagai perbandingan, metode berbasis deep learning seperti IndoBERT membutuhkan sumber daya komputasi yang lebih tinggi dan waktu pemrosesan yang lebih lama [20], sehingga kurang optimal pada lingkungan VPS dengan sumber daya terbatas.

2.7. JavaScript dan Python

JavaScript merupakan bahasa pemrograman web yang bersifat client-side dan banyak digunakan untuk membangun tampilan aplikasi web yang dinamis [21][22]. Python merupakan bahasa pemrograman tingkat tinggi yang berorientasi objek dan banyak digunakan untuk pengolahan data karena fleksibilitas dan kemudahan implementasinya, terutama dalam pemrosesan data dan visualisasi [23][24].

2.8. Extreme Programming (XP)

Extreme Programming (XP) merupakan metodologi pengembangan perangkat lunak yang menekankan kualitas perangkat lunak serta kemampuan beradaptasi terhadap perubahan kebutuhan pengguna [25]. Tahapan utama XP terdiri dari planning, designing, coding, dan testing yang dilakukan secara iteratif sehingga memungkinkan pengembangan sistem secara bertahap dan adaptif [26].

3. METODE PENELITIAN

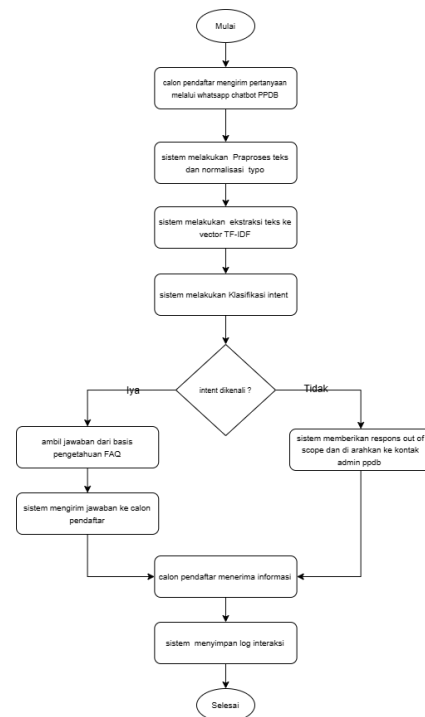
Penelitian ini dirancang sebagai penelitian rekayasa perangkat lunak menggunakan pendekatan kualitatif-kuantitatif. Pendekatan kualitatif diperoleh dari hasil observasi dan wawancara dengan staf panitia PPDB MTs Attaqwa 04, sedangkan pendekatan kuantitatif diperoleh dari hasil evaluasi performa model klasifikasi dan pengujian penerimaan pengguna.

3.1. Teknik Pengumpulan Data

Pengumpulan data dilakukan melalui tiga teknik. Pertama, observasi langsung di MTs Attaqwa 04 untuk memetakan alur pelayanan informasi yang masih dilakukan secara manual melalui WhatsApp. Kedua, wawancara dengan Bapak Imam Dharmadi selaku staf panitia PPDB untuk mengidentifikasi kebutuhan basis pengetahuan dan kendala komunikasi yang terjadi. Ketiga, studi pustaka terhadap jurnal ilmiah, dokumen operasional SOP PPDB, serta data riwayat percakapan WhatsApp panitia.

3.2. Model AI yang Digunakan

Model AI yang dibangun pada penelitian ini menggabungkan pendekatan NLP dengan algoritma Multinomial Naïve Bayes sebagai classifier intent. Pipeline pemrosesan terdiri dari empat tahap utama: (1) text preprocessing meliputi case folding, cleansing karakter, normalisasi berbasis aturan, dan fuzzy correction; (2) feature extraction menggunakan TF-IDF untuk mengubah teks menjadi vektor numerik berbobot; (3) text classification menggunakan Multinomial Naïve Bayes dengan Laplace smoothing; serta (4) pemilihan jawaban dari basis pengetahuan FAQ berdasarkan intent hasil prediksi atau respons out of scope apabila confidence di bawah ambang. Alur menyeluruh sistem usulan ditunjukkan pada Gambar 2.



Gambar 2. Flowchart Sistem Chatbot PPDB yang Diusulkan

Sumber: Hasil Penelitian (2026)

Pembobotan fitur TF-IDF dihitung melalui Persamaan (2) sampai (4) berikut.

$$TF(t,d) = f(t,d) / \sum f(t',d) \quad (2)$$

$$IDF(t) = \log(N / (1 + Df(t))) + 1 \quad (3)$$

$$TF-IDF(t,d) = TF(t,d) \times IDF(t) \quad (4)$$

Pada algoritma Multinomial Naïve Bayes, probabilitas kata pada kelas dihitung dengan Laplace smoothing sebagaimana ditunjukkan pada Persamaan (5).

$$P(w_i|c) = (N_{w_i,c} + 1) / (N_c + |V|) \quad (5)$$

Keterangan: $N_{w_i,c}$ adalah jumlah kemunculan kata w_i pada kelas c ; N_c adalah total token pada kelas c ; dan $|V|$ adalah jumlah kosakata unik.

3.3. Model Pengembangan Sistem

Model pengembangan sistem yang digunakan adalah Extreme Programming (XP) [25][26] dengan enam tahapan: planning, design, coding, testing, release, dan iteration. Tahap planning melakukan identifikasi kondisi layanan informasi PPDB dan kebutuhan

fungsional-nonfungsional sistem. Tahap design menyusun arsitektur dan flowchart sistem chatbot. Tahap coding mengimplementasikan sistem menggunakan Python-Flask sebagai backend dengan integrasi WhatsApp melalui WAHA API. Tahap testing dilakukan melalui alpha testing internal dan beta testing dengan pihak sekolah. Tahap release men-deploy sistem ke server VPS. Tahap iteration melakukan evaluasi dan penyempurnaan berdasarkan hasil pengujian dan umpan balik pengguna. Alur tahapan tersebut diilustrasikan pada Gambar 1.



Gambar 1. Tahapan Extreme Programming dalam Pengembangan Chatbot PPDB

Sumber: Hasil Penelitian (2026)

3.4. Dataset

Dataset disusun dari kebutuhan informasi PPDB yang paling sering ditanyakan calon pendaftar, seperti jadwal, syarat dokumen, biaya, fasilitas, dan kontak panitia. Materi dikumpulkan dari wawancara dan evaluasi pertanyaan riil melalui kanal WhatsApp PPDB. Dataset terdiri atas dua komponen yang disimpan dalam format CSV: `train_intents.csv` berisi sekitar 1.000 contoh pertanyaan dengan label intent, dan `kb_faq.csv` berisi pasangan intent dan jawaban resmi yang akan dikirim ke pengguna.

3.5. Teknik Pengujian

Pengujian sistem dilakukan melalui tiga skema. Pertama, evaluasi model menggunakan data uji holdout dengan metrik akurasi. Kedua, pengujian fungsional menggunakan metode alpha testing dengan skenario-skenario inti layanan PPDB. Ketiga, User Acceptance Testing (UAT) menggunakan kuesioner skala Likert 1–5 yang dianalisis dengan rumus rata-rata pada Persamaan (6) [31].

$$\bar{x} = \sum (f_i \cdot s_i) / n \quad (6)$$

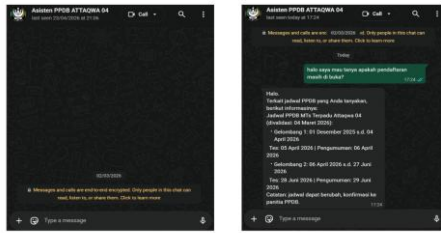
Keterangan: $\bar{x}$ adalah nilai rata-rata pernyataan; f_i adalah frekuensi responden yang memilih skor ke- i ; s_i adalah nilai skor ke- i ; dan n adalah jumlah total responden.

4. HASIL DAN PEMBAHASAN

4.1. Implementasi Sistem

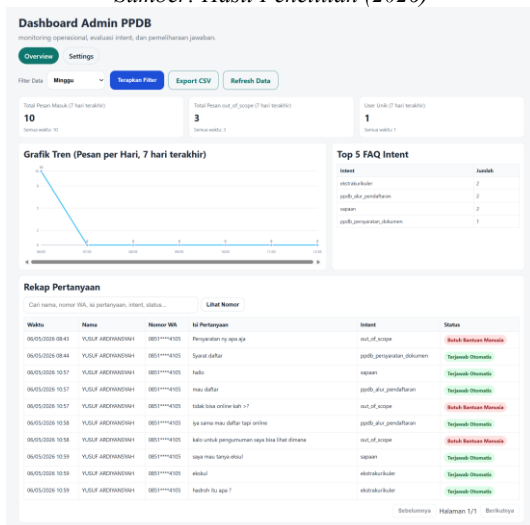
Sistem chatbot diimplementasikan dalam bentuk arsitektur tiga lapis. Lapis pertama adalah WhatsApp sebagai antarmuka pengguna, terhubung ke WAHA yang berperan sebagai gateway. WAHA dikonfigurasi sebagai webhook yang meneruskan pesan masuk ke backend. Lapis kedua adalah backend Python berbasis Flask yang menerima webhook, mengekstrak metadata pesan (ID pengirim, isi pesan, timestamp, metadata sesi), kemudian mengalirkan teks ke pipeline pemrosesan NLP. Lapis ketiga adalah mesin AI yang melakukan pra-proses, ekstraksi fitur TF-IDF, klasifikasi intent menggunakan model Multinomial Naïve Bayes yang telah dilatih dan disimpan sebagai file `naive_bayes_pipeline.joblib`, serta penentuan jawaban dari `kb_faq.csv`.

Antarmuka pengguna sistem ditampilkan pada Gambar 3, di mana pengguna mengirim pertanyaan dalam bahasa natural melalui WhatsApp dan menerima respons otomatis dari chatbot tanpa perlu format perintah khusus. Sementara itu, sistem juga dilengkapi dashboard admin berbasis web yang ditunjukkan pada Gambar 4. Dashboard menampilkan ringkasan total pesan masuk, jumlah pesan out of scope, pengguna unik, grafik tren pesan harian, top FAQ intent, serta rekap percakapan dalam tabel terstruktur lengkap dengan fitur filter periode, pencarian, dan export CSV.



Gambar 3. Tampilan Antarmuka Chatbot pada WhatsApp

Sumber: Hasil Penelitian (2026)



Gambar 4. Tampilan Dashboard Admin PPDB

Sumber: Hasil Penelitian (2026)

4.2. Hasil Praproses Teks

Tabel 1 menunjukkan ilustrasi transformasi bertahap dari input mentah pengguna menjadi bentuk teks akhir yang siap diproses oleh model.

Tabel 1. Hasil Transformasi Praproses Teks

No	Tahap	Hasil
1	Input asli	Syarat nya apa min??
2	Case folding	syarat nya apa min??
3	Cleansing	syarat nya apa min
4	Normalisasi berbasis aturan	syaratnya apa admin
5	Fuzzy correction	syaratnya apa admin
6	Output praproses	syaratnya apa admin

Sumber: Hasil Penelitian (2026)

Pada Tabel 1, input mentah “Syarat nya apa min??” yang mengandung kapitalisasi tidak baku, tanda baca berlebih, kata terpotong, dan singkatan tidak baku, secara bertahap disamakan menjadi bentuk normal “syaratnya

apa admin?”. Hal ini menunjukkan bahwa kombinasi normalisasi berbasis aturan dan fuzzy correction efektif menangani variasi bahasa informal yang lazim ditemukan pada percakapan WhatsApp.

4.3. Evaluasi Awal Model

Evaluasi awal dilakukan menggunakan data uji holdout yang dipisahkan dari data latih. Tabel 2 menunjukkan rekap hasil evaluasi.

Tabel 2. Rekap Evaluasi Awal Model

Kategori Pengujian	Jumlah Sampel	Prediksi i Benar	Akurasi
Data uji holdout	223	191	85,65%

Sumber: Hasil Penelitian (2026)

Berdasarkan Tabel 2, model chatbot berhasil mengklasifikasikan 191 dari 223 sampel uji dengan benar, sehingga diperoleh akurasi sebesar 85,65%. Tingkat akurasi ini relatif kompetitif jika dibandingkan dengan studi terdahulu pada domain serupa, antara lain [6] yang memperoleh akurasi 88% dan [7] yang memperoleh akurasi 88,9%. Perbedaan tersebut dapat dijelaskan oleh karakteristik dataset penelitian ini yang berasal dari riwayat percakapan WhatsApp nyata sehingga memiliki variasi bahasa, singkatan, dan typo yang lebih tinggi dibandingkan dataset berbasis brosur ataupun dokumen formal. Hasil ini menunjukkan bahwa kombinasi TF-IDF dan Multinomial Naïve Bayes mampu bekerja stabil pada lingkungan operasional nyata dengan data terbatas.

4.4. Hasil Pengujian Fungsional

Pengujian fungsional dilakukan menggunakan metode alpha testing terhadap 14 skenario inti layanan PPDB, mencakup jadwal pendaftaran, biaya, persyaratan dokumen, alur pendaftaran, program unggulan, ekstrakurikuler, kegiatan siswa, MBG, beasiswa, respons out of scope, serta pengujian follow-up konteks. Seluruh skenario menghasilkan output yang sesuai dengan ekspektasi (status Berhasil). Hasil ini menunjukkan sistem mampu mengenali intent dengan benar, memberikan jawaban relevan dari basis pengetahuan, mengembalikan respons out of scope ketika pertanyaan berada di luar cakupan, serta mempertahankan konteks percakapan pada pertanyaan lanjutan.

4.5. Hasil Pengujian Pengguna (UAT)

UAT dilakukan terhadap 28 responden dari total populasi 119 calon pendaftar PPDB tahun ajaran 2026/2027 menggunakan kuesioner skala Likert 1–5. Pengujian difokuskan pada lima aspek, yaitu fungsionalitas sistem, kinerja sistem, pengalaman pengguna, efisiensi dan produktivitas, serta performa chatbot NLP. Tabel 3 menyajikan rekap hasil UAT pada sepuluh pernyataan pengujian.

Tabel 3. Rekap Hasil UAT 28 Responden

No	Pernyataan	Rata-rata	Kategori
P1	Respons otomatis tanpa keterlambatan	4,18	Setuju
P2	Informasi akurat dan sesuai kebutuhan	4,43	Sangat Setuju
P3	Kejelasan informasi menu dan persyaratan	4,36	Sangat Setuju
P4	Mudah digunakan tanpa pengetahuan teknis	4,18	Setuju
P5	Memahami variasi kalimat pertanyaan	4,18	Setuju
P6	Percakapan berkesinambungan	4,21	Sangat Setuju
P7	Informasi lokasi sekolah sesuai dan jelas	4,18	Setuju
P8	Eskalasi ke admin pada kasus yang tepat	4,21	Sangat Setuju
P9	Bot membantu memperoleh informasi cepat	4,29	Sangat Setuju
P10	Kepuasan keseluruhan terhadap sistem	4,39	Sangat Setuju

Sumber: Hasil Penelitian (2026)

Berdasarkan Tabel 3, seluruh pernyataan memperoleh skor rata-rata di rentang 4,18–4,43 yang termasuk dalam kategori “Setuju” hingga “Sangat Setuju”. Skor tertinggi diperoleh pada pernyataan keakuratan informasi (P2 = 4,43), diikuti kepuasan keseluruhan (P10 = 4,39) dan kejelasan informasi menu (P3 = 4,36). Skor terendah, yaitu pada aspek kinerja kecepatan respons (P1), kemudahan penggunaan (P4), pemahaman variasi kalimat (P5), dan kejelasan informasi lokasi (P7), seluruhnya tetap berada pada kategori “Setuju” dengan nilai 4,18. Hasil ini mengonfirmasi bahwa sistem dapat diterima dengan baik oleh pengguna sasaran.

4.6. Pembahasan

Hasil penelitian ini menjawab dua rumusan masalah utama. Pertama, terkait kemampuan chatbot memberikan respons cepat: skor P1 sebesar 4,18 menunjukkan responden menilai sistem mampu memberikan respons otomatis

tanpa keterlambatan yang berarti, didukung oleh hasil uji fungsional yang menunjukkan 100% skenario berjalan sesuai ekspektasi. Kedua, terkait ketersediaan layanan 24 jam: arsitektur sistem yang menempatkan model klasifikasi pada backend Flask dengan deployment di VPS memungkinkan layanan berjalan tanpa bergantung pada keberadaan petugas, sehingga pertanyaan yang masuk di luar jam kerja tetap mendapatkan respons otomatis.

Dibandingkan dengan penelitian terdahulu, kontribusi utama studi ini ada pada tiga aspek. Pertama, dari sisi platform, sebagian besar penelitian sebelumnya seperti [27], [28], [29], dan [30] mengimplementasikan chatbot pada platform website atau Telegram. Penelitian ini menggunakan WhatsApp yang sudah familiar bagi masyarakat sehingga menjangkau pengguna lebih luas, sejalan dengan [3]. Kedua, dari sisi dataset, penelitian ini menggunakan dataset percakapan WhatsApp nyata, bukan turunan dari brosur seperti pada [6]. Ketiga, dari sisi tahan typo, kombinasi normalisasi berbasis aturan dan fuzzy correction yang diintegrasikan ke pipeline NLP membantu sistem tetap mengenali maksud pertanyaan meskipun pengguna mengetik dengan kesalahan ringan aspek yang masih menjadi keterbatasan pada [28].

Beberapa keterbatasan tetap perlu dicatat. Sistem hanya melayani pertanyaan dalam Bahasa Indonesia berbasis teks, belum mendukung pesan suara, gambar, atau dokumen. Selain itu, sistem berfokus pada jalur pendaftaran reguler dan tidak menangani kasus khusus seperti siswa pindahan dengan dokumen non-standar yang tetap memerlukan penanganan langsung oleh panitia.

5. KESIMPULAN

Berdasarkan hasil penelitian, perancangan, implementasi, dan pengujian sistem chatbot WhatsApp berbasis NLP dengan algoritma Multinomial Naïve Bayes pada layanan informasi PPDB MTs Attaqwa 04 Kabupaten Bekasi, dapat disimpulkan hal-hal sebagai berikut:

1. Sistem chatbot WhatsApp berbasis NLP berhasil diimplementasikan dan mampu memberikan respons otomatis terhadap pertanyaan calon pendaftar melalui integrasi WAHA dan algoritma Multinomial Naïve

Bayes. Evaluasi model menghasilkan akurasi 85,65% pada data uji holdout, sementara pengujian fungsional terhadap 14 skenario seluruhnya berhasil. Pernyataan P1 pada UAT memperoleh skor rata-rata 4,18 yang berarti responden setuju sistem mampu memberi respons tanpa keterlambatan yang berarti.

2. Sistem dapat memberikan layanan informasi PPDB 24 jam melalui media WhatsApp, sehingga lebih mudah diakses calon pendaftar maupun orang tua. Hal ini didukung skor kepuasan keseluruhan (P10) sebesar 4,39 yang termasuk kategori sangat setuju, serta skor keakuratan informasi (P2) sebesar 4,43.

3. Penerapan kombinasi praproses berbasis aturan dan fuzzy correction terbukti efektif menangani variasi bahasa informal khas WhatsApp seperti typo dan singkatan, yang sebelumnya menjadi keterbatasan beberapa penelitian terdahulu.

Pengembangan lanjutan dapat diarahkan pada penambahan jumlah dataset intent untuk meningkatkan kemampuan pemahaman sistem, perluasan dukungan media non-teks (gambar dan dokumen), serta pengayaan fitur analitik pada dashboard admin seperti monitoring performa real-time dan statistik pertanyaan pengguna yang lebih lengkap.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada MTs Attaqwa 04 Kabupaten Bekasi yang telah memberikan kesempatan untuk pelaksanaan observasi dan penelitian, kepada Ibu Nadya Safitri, S.T., M.Kom. selaku dosen pembimbing, serta kepada seluruh civitas akademika Universitas Bina Insani yang telah mendukung penyelesaian penelitian ini.

DAFTAR PUSTAKA

- [1] A. S. Atmaja, A. S. Sidabalok, M. Raihan, F. A. Putra, dan N. Ifkah, "Perancangan Sistem Informasi Penerimaan Peserta Didik Baru (PPDB) Berbasis Web," vol. 2, hlm. 515–523, 2024.
- [2] A. Farizi, I. N. Farida, dan D. Swanjaya, "Penerapan Chatbot Berbasis Natural Language Processing untuk Layanan Informasi PPDB di SMK PGRI 1 Kediri," vol. 9, hlm. 1044–1053, 2025.
- [3] I. Amirulloh, M. W. Pertiwi, dan T. Wibisono, "Rancang Bangun Chatbot WhatsApp Menggunakan Node.js dan Model Natural Language Processing untuk Layanan PPDB SMK YPC Tasikmalaya," Jurnal Informatika dan Teknik Elektro Terapan (JITET), vol. 12, 2024.
- [4] Mubarak, "Implementasi Natural Language Processing dalam Perancangan Aplikasi Chatbot pada FIKTI UMSU," vol. 8, hlm. 11992–12001, 2024.
- [5] D. M. Alfiansyah, L. Setiyani, dan D. F. Wati, "Pengembangan Chatbot Berbasis Web untuk Layanan Informasi di Horizon," vol. 7, 2025.
- [6] E. B. Susanto dan M. R. Maulana, "Pengembangan Chatbot untuk Layanan Penerimaan Mahasiswa Baru di Institut Widya Pratama dengan Menggunakan Metode Naive Bayes," hlm. 16–21, 2025.
- [7] C. Diantoni, R. Mufidah, dan H. Triana, "Membangun Chatbot untuk Informasi Magang dan Studi Independen Kampus Merdeka dengan Algoritma Naive Bayes," vol. 8, 2024.
- [8] A. V. Uamira, "Proses Pembuatan Peraturan Daerah Terkait Kebijakan Pendidikan di Sumatera Barat: PPDB," vol. 3, 2025.
- [9] S. Qomaria, M. T. Arifin, dan A. Djonu, "Pemberian Layanan Informasi untuk Mengurangi Perilaku Membolos pada Siswa Kelas X SMK Negeri 1 Maumere," vol. 14, 2022.
- [10] F. C. Wulur dan D. Mulyanti, "Analisis Pengaruh Pemanfaatan Media Sosial di Pemerintah," vol. 2, hlm. 37–45, 2023.
- [11] S. Hariyanto, I. Fenriana, D. Surya, D. Putra, dan D. Lasut, "Perancangan Virtual Assistant Chatbot Berbasis Website sebagai Alat Promosi dan Dukungan Pemasaran," Rubinstein: Jurnal Multidisiplin, vol. 2, 2023.
- [12] R. Parina, "Aplikasi Chatbot Sebagai Media Pembelajaran Interaktif SD N 17 Kota Bengkulu Berbasis Android," vol. 18, hlm. 121–127, 2022.
- [13] I. S. Juniar dan K. Nasionalita, "Motif Penggunaan Media pada Pengguna Fitur WhatsApp Story (Studi pada Generasi Millenial di Kota Bandung)," vol. 10, hlm. 451–456, 2022.
- [14] Roziqbahtiar, "Install WAHA: WhatsApp API Open-Source Tanpa Biaya," roziqbahtiar.com, 2025.
- [15] M. H. Firmansyah, "Membangun Self WhatsApp Gateway: Panduan Integrasi WAHA (Docker) ke Laravel," Medium, 2026.
- [16] M. A. Prasetya, M. Wulandari, dan S. A. Nikmah, "Implementasi NLP (Natural Language Processing) Dasar pada Analisis Sentiment Review Spotify," vol. 3, hlm. 145–153, 2024.
- [17] M. Hafid, N. Firmansyah, dan S. N. Budiman, "Sistem Rekomendasi Konversi Mata Kuliah

- Berbasis NLP Menggunakan Metode BERT di Universitas Islam Balitar," vol. 4, hlm. 674–682, 2025.
- [18] I. M. A. Arya, I. M. G. Dwija, dan I. G. Aris, "Perbandingan Algoritma Naive Bayes Berbasis Feature Selection Gain Ratio dengan Naive Bayes Konvensional dalam Prediksi Komplikasi Hipertensi," JTIM: Jurnal Teknologi Informasi dan Multimedia, vol. 6, hlm. 37–49, 2024.
- [19] N. Adila, "Analisis Komparatif Algoritma K-Nearest Neighbors dan Naive Bayes untuk Klasifikasi Target Pemasaran pada Media Sosial," vol. 5, hlm. 32–39, 2026.
- [20] S. Perbandingan, K. T. Dan, dan I. Untuk, "Studi Perbandingan Kinerja TF-IDF dan IndoBERT untuk Rekomendasi Resep Berdasarkan Ketersediaan Bahan Makanan Berbasis Website," vol. 9, hlm. 1–8, 2025.
- [21] I. P. Sari, A. Jannah, A. M. Meuraxa, A. Syahfitri, dan R. Omar, "Perancangan Sistem Informasi Penginputan Database Mahasiswa Berbasis Web," 2022.
- [22] P. D. Nusantara, S. B. Nauli, R. Sibarani, dan F. Zuli, "Pelatihan Pemrograman Web JavaScript Dasar pada Komunitas Sel Gereja Kasih Karunia Indonesia (GEKARI)," vol. 5, 2023.
- [23] S. Junaidi, M. Devegi, dan H. Kurniawan, "Pelatihan Pengolahan dan Visualisasi Data Penduduk Menggunakan Python," vol. 4, hlm. 151–162, 2023.
- [24] D. E. Apriani, G. B. Siburian, R. Situmorang, S. A. Simamora, dan T. A. Hutapea, "Menyelesaikan Permasalahan Supremum dan Infimum Suatu Himpunan dengan Menggunakan Python," vol. 3, 2025.
- [25] A. Fikriyadi, D. Purnomo, H. Ulya, A. Makassari, dan A. Saifudin, "Metode Extreme Programming untuk Memberikan Informasi Penjualan PT. Megalestari Epack Sentosaraya," vol. 1, hlm. 625–632, 2023.
- [26] M. R. Musyaffa, B. Prasetyo, dan M. A. Komara, "Pengembangan Aplikasi Al-Qur'an Digital Berbasis Mobile dengan Integrasi Chatbot Tanya Jawab Islami Menggunakan Metodologi Extreme Programming," vol. 9, hlm. 8350–8356, 2025.
- [27] A. Setiyorini, "Pengembangan Aplikasi Chatbot untuk Layanan Penerimaan Mahasiswa Baru Berbasis Natural Language Processing," Jurnal Informatika: Jurnal Pengembangan IT, vol. 10, hlm. 1115–1126, 2025.
- [28] R. R. Rafi, A. Setiawan, dan A. A. Rosadi, "Chatbot Penerimaan Peserta Didik Baru di SMK Intensif Baitussalam Menggunakan NLP," vol. 4, hlm. 130–135, 2025.
- [29] M. E. Fauzi, "PublicTalk: Sistem Chatbot Pintar Berbasis Natural Language Processing untuk Layanan Pemerintahan Digital," vol. 3, hlm. 426–433, 2025.
- [30] J. Khatib dan S. Dalam, "Penggunaan Teknologi Natural Language Processing dalam Sistem Chatbot untuk Peningkatan Layanan Informasi Administrasi Publik," Indonesian Journal of Computer Science, vol. 13, hlm. 1227–1242, 2024.
- [31] N. Hartono dan A. A. Muin, "Penggunaan User Acceptance Testing (UAT) pada Pengujian Sistem Informasi Pengelolaan Keuangan dan Inventaris Barang," 2025.