

PENGEMBANGAN CHATBOT UNTUK LAYANAN PENERIMAAN MAHASISWA BARU DI INSTITUT WIDYA PRATAMA DENGAN MENGGUNAKAN METODE NAIVE BAYES

Eko Budi Susanto¹⁾, Mohammad Reza Maulana²⁾, Arochman³⁾

Fakultas Teknologi Informasi, Program Studi Teknik Informatika, Institut Widya Pratama
ekobudi.s@stmik-wp.ac.id¹⁾, reza@stmik-wp.ac.id²⁾, arochman@stmik-wp.ac.id³⁾

Abstrak

Pada penelitian ini, sebuah chatbot untuk penerimaan mahasiswa baru dikembangkan dengan memanfaatkan algoritma Naïve Bayes. Permasalahan yang ingin diselesaikan adalah tingginya frekuensi pertanyaan berulang dari calon mahasiswa yang harus dilayani oleh staf administrasi, sehingga menimbulkan beban kerja dan keterlambatan dalam memberikan informasi. Dengan chatbot, proses pelayanan informasi dapat diotomatisasi dan dilakukan selama 24 jam. Sumber data yang digunakan adalah dialog antara staf penerimaan dan calon mahasiswa, yang didasarkan pada informasi dari brosur penerimaan mahasiswa baru Institut Widya Pratama (d.h. STMIK Widya Pratama) tahun akademik 2024-2025. Data tersebut disusun dalam format JSON, mencakup tiga elemen: struktur pertanyaan, klasifikasi, dan jawaban. Langkah-langkah pra-pemrosesan data meliputi penormalan teks, pengkodean label, vektorisasi, dan word embedding. Algoritma Naïve Bayes dipilih untuk melatih model karena kemampuannya dalam memprediksi respons dari data diskrit. Hasil pengujian menunjukkan bahwa algoritma ini mencapai tingkat akurasi 88%, melampaui kinerja algoritma K-Neighbors-Classifer. Model yang dihasilkan diimplementasikan dalam aplikasi chatbot berbasis desktop yang dijalankan melalui command prompt, menggunakan bahasa pemrograman Python dan beberapa pustaka seperti pickle, string, dan numpy. Aplikasi ini mampu memberikan jawaban atas pertanyaan pengguna dengan mengakses model yang tersimpan dalam file *.pkl. Hasil penelitian menunjukkan bahwa algoritma Naïve Bayes efektif dalam memprediksi pertanyaan pengguna, dengan akurasi yang memadai untuk aplikasi chatbot penerimaan mahasiswa baru.

Kata kunci: NLP, ChatBot, Naïve Bayes, PMB

1. Pendahuluan (bold 11 pt)

Kemajuan teknologi informasi dan komunikasi (TIK) telah mendorong perubahan besar dalam berbagai sektor kehidupan, termasuk dunia pendidikan. Perguruan tinggi, seperti Institut Widya Pratama, dituntut untuk menyesuaikan diri dengan perkembangan teknologi agar dapat meningkatkan mutu layanan bagi calon mahasiswa (Heryati et al., 2023; Nugraha et al., 2024). Salah satu upaya inovatif yang dapat dilakukan adalah penerapan chatbot sebagai sarana komunikasi interaktif dalam proses penerimaan mahasiswa baru. Chatbot merupakan perangkat lunak yang mampu menirukan percakapan manusia melalui teks atau suara, sehingga mampu menyajikan informasi secara cepat dan praktis (Nurul Puteri et al., 2022), (Diantoni et al., 2024).

Dalam proses penerimaan mahasiswa baru, seringkali terdapat pertanyaan yang berulang terkait program studi, syarat pendaftaran, jadwal seleksi, hingga prosedur pembayaran. Kehadiran chatbot dapat membantu mengurangi beban tugas staf administrasi dan memberikan layanan yang lebih responsif kepada calon mahasiswa. Selain itu, chatbot memiliki keunggulan karena dapat beroperasi selama 24 jam penuh, memungkinkan calon mahasiswa memperoleh informasi kapan saja tanpa terbatas waktu.

Algoritma Naive Bayes dipilih sebagai metode pengembangan chatbot ini karena kemampuannya yang efektif dalam klasifikasi teks. Naive Bayes merupakan metode machine learning yang berlandaskan teorema Bayes dengan asumsi bahwa setiap fitur bersifat independen. Metode ini dikenal andal dalam berbagai penerapan pengolahan bahasa alami

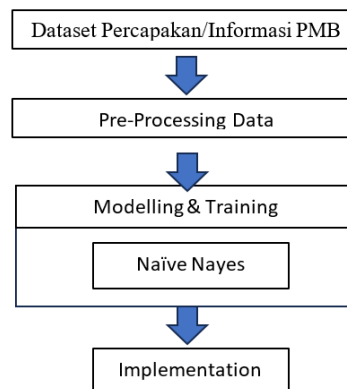
(NLP), termasuk pengembangan chatbot (Aelani & Gustaman, 2021). Kelebihan Naive Bayes meliputi kemudahan penerapan, kecepatan proses, serta ketepatan yang tinggi dalam pengolahan data berbasis teks.

Naive Bayes mampu meramalkan respons yang tepat berdasarkan pola data yang telah dilatih sebelumnya (Diantoni et al., 2024). Dengan penerapan metode ini, diharapkan chatbot yang dikembangkan dapat menyajikan informasi yang relevan dan memenuhi kebutuhan calon mahasiswa. Penelitian yang dilakukan oleh Diantoni dkk menunjukkan bahwa penggunaan metode Naive Bayes menghasilkan performa model dengan akurasi mencapai 88,9%. Keefektifan algoritma Naive Bayes telah terbukti, dan dengan penambahan data serta penyesuaian model, akurasi chatbot dapat meningkat secara signifikan (Diantoni et al., 2024). Oleh karena itu, chatbot ini memiliki potensi besar untuk meningkatkan kualitas interaksi dengan pengguna serta memberikan kontribusi positif dalam penyebaran informasi di lingkungan pendidikan. Susanto, dkk melakukan evaluasi kinerja algoritma naive bayes dengan k-nn pada dataset pmb, hasilnya didapatkan bahwa algoritma naive bayes memiliki performa yang lebih baik dibandingkan dengan k-nn (Budi Susanto et al., 2024)

Penelitian ini bertujuan untuk merancang chatbot yang dapat mendukung proses pelayanan penerimaan mahasiswa baru di Institut Widya Pratama dengan mengadopsi metode Naive Bayes. Diharapkan, penerapan chatbot ini dapat meningkatkan efisiensi serta mutu layanan informasi bagi calon mahasiswa, sehingga menciptakan pengalaman yang lebih baik dalam proses pendaftaran. Permasalahan yang menjadi latar belakang penelitian ini adalah masih manualnya proses layanan informasi PMB yang menyebabkan keterlambatan dalam merespon pertanyaan berulang dari calon mahasiswa, serta keterbatasan waktu layanan oleh staf. Selain itu belum banyaknya studi yang mengimplementasikan chatbot berbasis algoritma Naive Bayes secara spesifik dalam konteks PMB di perguruan tinggi kecil-menengah seperti Institut Widya Pratama juga menjadi pertimbangan dalam penelitian ini. Penelitian ini memberikan kontribusi yaitu pengembangan dan evaluasi chatbot berbasis

Naive Bayes yang mampu memberikan layanan informasi PMB secara otomatis, akurat, dan efisien. Kontribusi ini ditampilkan dalam bagian metode (pengumpulan dan pelatihan model dari dataset PMB), hasil dan pembahasan (evaluasi akurasi dan implementasi sistem), serta kesimpulan (rekomendasi pengembangan lebih lanjut).

2. Metode Penelitian



Gambar 1 Metode Penelitian

Dari Gambar 1 metode penelitian dapat dijabarkan sebagai berikut

2.1 Dataset

Dataset yang digunakan yaitu: percakapan tentang PMB. Langkah pertama adalah mengumpulkan dataset yang terdiri dari percakapan-percakapan yang berkaitan dengan penerimaan mahasiswa baru. Dataset ini bisa mencakup berbagai pertanyaan dan jawaban umum yang diajukan oleh calon mahasiswa, seperti informasi tentang program studi, biaya pendaftaran, prosedur pendaftaran dan yang lainnya

2.2 Pre-Processing Data

Data Cleansing. Setelah dataset dikumpulkan, langkah berikutnya adalah membersihkan data dari elemen-elemen yang tidak relevan atau noise. Proses ini mencakup penghapusan kata-kata yang tidak penting, seperti tanda baca atau stop words, serta penanganan kesalahan penulisan yang dapat mempengaruhi kualitas data.

Vektorisasi. Proses ini mengubah teks yang telah diberi label menjadi vektor angka melalui teknik seperti TF-IDF (Term Frequency-Inverse Document Frequency) atau Bag of Words.

Vektorisasi ini memungkinkan algoritma pembelajaran mesin untuk memahami dan mengolah teks.

Embedding adalah teknik yang digunakan untuk merepresentasikan kata-kata dalam bentuk vektor di ruang berdimensi rendah, yang lebih mudah diproses oleh model. Teknik ini mempertahankan hubungan semantik antar kata sehingga model dapat mengenali konteks yang lebih luas.

2.3 Modeling & Training

Pada tahap ini, model pembelajaran mesin menggunakan algoritma Naive Bayes dikembangkan untuk memprediksi jawaban yang sesuai berdasarkan input dari pengguna. Pemodelan ini bertujuan untuk menciptakan chatbot yang dapat memberikan respons yang relevan dan akurat. Pada tahap ini dataset dibagi menjadi 2 bagian yaitu untuk data training dan data testing. Sebanyak 20% untuk data testing, sisanya 80% untuk data training.

Setelah model dibangun, model tersebut dilatih menggunakan dataset yang telah diproses. Proses pelatihan ini bertujuan untuk mengoptimalkan kemampuan model dalam memahami dan merespons berbagai pertanyaan yang mungkin diajukan oleh calon mahasiswa.

2.4 Implementation

Setelah model dilatih dan diuji, langkah terakhir adalah mengimplementasikan chatbot ke dalam sistem yang akan digunakan oleh pengguna. Implementasi ini mencakup integrasi chatbot ke dalam bentuk aplikasi yang berjalan di command prompt window menggunakan software vscode dengan bahasa pemrograman python, di mana calon mahasiswa dapat mengajukan pertanyaan dan menerima jawaban secara otomatis.

3. Hasil dan Pembahasan

Penelitian ini memanfaatkan dataset yang berisi interaksi antara petugas penerimaan mahasiswa baru dan calon mahasiswa. Jawaban yang diberikan petugas didasarkan pada informasi yang tertera di brosur penerimaan mahasiswa baru Institu Widya Pratma (d.h. STMIK Widya Pratama) tahun akademik 2024-2025. Dataset ini memiliki tiga variabel penting: 'patterns' yang merepresentasikan pola percakapan, 'tag' yang mengkategorikan

pertanyaan, dan 'respon' yang berisi jawaban. Dataset terdiri dari 127 record dan 3 atribut.

Tabel 1. Dataset PMB

No	Field	Keterangan
1	Patterns	Teks yang berisi pertanyaan atau pernyataan calon mahasiswa terkait PMB.
2	Tag	Label yang menandai jenis topik yang dibahas dalam pola pertanyaan
3	Respon	Teks yang berisi jawaban yang diberikan oleh petugas PMB.

Dataset ini mencakup tiga atribut: pola pertanyaan, kategori, dan jawaban (NgodingPython, 2021). Pola pertanyaan berisi teks pertanyaan terkait penerimaan mahasiswa baru, kategori mengklasifikasikan jenis pertanyaan, dan jawaban akan dipilih secara acak oleh algoritma. Dataset disimpan dalam format JSON.

```

1 // @context X
2 data-chat {} intents {} [] intent {} 10
3 "intents": [
4 ...
5
6 "tag": "telepon",
7
8 "patterns": ["nomor telepon", "telepon", "telpon", "telp", "nomor HP", "hand phone", "phone"],
9
10 "responses": ["0800", "hi kak 0800"]
11
12
13
14 "tag": "fakultas",
15
16 "patterns": ["fakultasnya apa", "info fakultas", "fakultasnya apa saja", "ada fakultas apa"],
17
18 "responses": ["Fakultas Ekonomi dan Bisnis, Fakultas Teknologi Informasi", "Ada 2 fakultas kak, Ekonomi dan Bisnis, Teknologi Info
19 Fakultas Ekonomi dan Bisnis memiliki progi Bisnis Digital"]
20
21
22
23 "tag": "progi",
24
25 "patterns": ["Progidnya apa saja", "Program Studi", "Jurusannya", "Progidnya", "Programnya", "Informasi Program", "Informasi Progi"],
26
27 "responses": ["SI Sistem Informasi, SI Teknik Informatika, D3 Manajemen Informatika, D3 Komputerisasi Akuntansi, SI Bisnis Digital"]
28
29
30

```

Gambar 2 Penggatan Isi Dataset *.json

Pre-Processing data mencakup normalisasi teks melalui case folding. Proses ini mengubah huruf kapital menjadi huruf kecil dan membersihkan teks dari tanda baca. Contoh implementasi case folding dapat dilihat di bawah ini.

Tabel 2. Hasil Case Folding & Cleansing

Pertanyaan	Case Folding & Cleansing
Manajemen	manajemen
Informatika itu apa?	informatika itu apa
"Halo, ada yang bisa saya bantu?"	halo ada yang bisa saya bantu

Pada proses case folding dan cleansing melibatkan proses label encoding, sebuah teknik

untuk memungkinkan penyimpanan objek dalam berkas atau basis data.

```
# 1.load data
path = 'data-chat/intents.json'
jp = JSONParser()
jp.parse(path)
df = jp.get_dataframe()

# Load model
with open('model-chatbot-wp.pkl', 'rb') as f:
    model = pickle.load(f)

# interaski dengan bot
print("[INFO] Anda sudah terhubung dengan IWIMA ...")
while True:
    chat = input("Anda --> ")
    res, tag = bot_respon(chat, model, jp)
    print(f"IWIMA --> {res}")
    if tag=="bye" :
        break
```

Gambar 6 Penggalan Kode Program ChatBot

File *.pkl dipanggil dengan perintah “pickle.load(..)”, kemudian model digunakan untuk memprediksi hasil percakapan yang dimasukkan oleh pengguna.

```
[INFO] Anda sudah terhubung dengan IWIMA ...
Anda --> Assalamualaikum
IWIMA --> Wassalamualaikum
Anda --> selamat pagi
IWIMA --> Pagi Juga
Anda --> siapa widya pratama
IWIMA --> Kenalin, aku Widya Pratama!
Anda --> apa saja program studinya
IWIMA --> S1 Sistem Informatasi, S1 Teknik Informatika, D3 Manaje
Anda --> dimana kampusnya
IWIMA --> Ini kak link nya https://pmb.iwima.ac.id
Anda --> baik, terima kasih ya
IWIMA --> Kalau ada masalah, hubungi aku lagi ya
```

Gambar 7 Hasil respon dari chatbot

Hasil evaluasi menunjukkan bahwa akurasi model mencapai 88%, yang merupakan nilai cukup tinggi untuk skenario layanan informasi berbasis teks. Jika dibandingkan dengan penelitian oleh Aelani & Gustaman (2021) yang menggunakan Naïve Bayes dalam konteks serupa, akurasi sistem ini berada dalam rentang yang kompetitif. Selain itu, penelitian oleh Diantoni et al. (2024) juga membuktikan bahwa metode ini unggul dalam kecepatan pelatihan dan ketepatan pada data terbatas. Dalam studi ini, performa Naïve Bayes juga mengungguli algoritma K-Nearest Neighbors (KNN), yang cenderung mengalami penurunan performa pada data dengan dimensi tinggi seperti teks. Hal ini memperkuat posisi Naïve Bayes sebagai

algoritma yang tepat untuk chatbot berbasis teks PMB.

4. Kesimpulan dan Saran

Pada penelitian ini Algoritma Naive Bayes memiliki kemampuan memprediksi pertanyaan yang berupa kalimat yang diinputkan oleh pengguna. Tingkat akurasi algoritma naive bayes sebesar 88% dengan dataset percakapan/pertanyaan dari pengguna. Model yang dilatih yaitu algoritma naïve bayes diimplementasikan pada aplikasi chat bot yang berbasis dekstop dan berjalan pada command prompt window. Proses pengujian alpha aplikasi dapat merespon jawaban dari pertanyaan pengguna.

Untuk pengembangan berikutnya aplikasi ini perlu ditambahkan dataset percakapan atau pertanyaan pengguna. Semakin banyak data yang digunakan untuk melatih model, semakin baik kemampuannya dalam memahami variasi bahasa dan konteks. Meskipun Naive Bayes menunjukkan hasil yang baik, dapat dipertimbangkan untuk mencoba algoritma klasifikasi teks lain yang lebih unggul kinerjanya, seperti Support Vector Machines (SVM), Recurrent Neural Networks (RNN), atau Transformers (misalnya, BERT).

DAFTAR PUSTAKA

- Aelani, K., & Gustaman, G. (2021). *Chatbot for Information Service of New Student Admission Using Multinomial Naïve Bayes Classification and TF-IDF Weighting*.
- Budi Susanto, E., Reza Maulana, M., Agus Setiawan, T., & Ilyas, A. (2024). Evaluasi Komprehensif Algoritma Naive Bayes Dan Knn Dalam Pengembangan Chatbot Cerdas Untuk Layanan Informasi PMB. *Jurnal Teknologi Informasi*, 8(2). <https://stmik-wp.ac.id/>
- Diantoni, C., Mufidah, R., Triana Informatika, H., Karawang, S., Ronggo Waluyo, J. H., Timur, T., Karawang, J., & Barat, I. (2024). Membangun Chatbot untuk Informasi Magang dan Studi Independen Kampus Merdeka Dengan Algoritma Naive Bayes.

- Jurnal Mahasiswa Teknik Informatika*, 8(2).
<https://kampusmerdeka.kemdikbud.go.id/>
- Heryati, D., Zulkifli, I., Fajri, R. M., & Kom, S. (2023). Aplikasi Chatbot untuk Penerimaan Mahasiswa Baru Universitas Indo Global Mandiri Menggunakan Deep Learning. *The Journal of Intelligent Networks and IoT Global*, 1(1).
- NgodingPython. (2021, December 4). [LIVE] Weekend Project: Membuat Chatbot Sederhana.
<https://www.youtube.com/watch?v=HihKQ8F0k4c&t=929s>.
- Nugraha, M., Sakinah, L., Setiawan, R. A., & Mulyani, H. (2024). Rancang Bangun Sistem Informasi Penerimaan Mahasiswa Baru Berbasis Web dengan Menggunakan Framework Laravel. *Jurnal Informatika Dan Teknik Elektro Terapan*, 12(2).
<https://doi.org/10.23960/jitet.v12i2.4179>
- Nurul Puteri, A., Tamrin, F., Rahman Nasir, K., Widya Anggraeni, D., & Arafah, M. (2022). Aplikasi Chatbot untuk Layanan Informasi Penerimaan Mahasiswa Baru. *Prosiding Seminar Nasional Teknik Elektro Dan Informatika (SNTEI) 2022-Teknik Informatika*, 168–174.