

## Implementasi Natural Language Processing berbasis Machine Learning untuk Membangun Aplikasi Chatbot Sebagai Media Informasi Perpustakaan

Yudistira Bagus Pratama<sup>1✉</sup>, Ulyma Tahara Arisandi<sup>2</sup>  
<sup>1,2</sup>Universitas Muhammadiyah Bangka Belitung, Indonesia

✉Corresponding Author: yudistira.bagus@unmuhbabel.ac.id

### ABSTRAK

Perkembangan Artificial Intelligence (AI), khususnya dalam bidang Machine Learning (ML) dan Natural Language Processing (NLP), telah mendorong inovasi dalam pengembangan sistem layanan informasi berbasis chatbot. Penelitian ini bertujuan untuk mengembangkan aplikasi chatbot sebagai media layanan informasi perpustakaan di SMA Negeri 1 Kelapa, Bangka Belitung, yang sebelumnya masih menggunakan sistem manual. Metode penelitian yang digunakan adalah kualitatif deskriptif dengan pendekatan pengembangan perangkat lunak menggunakan model waterfall, yang meliputi tahap analisis kebutuhan, desain, implementasi, pengujian, dan pemeliharaan. Model yang digunakan dalam penelitian ini adalah Bidirectional Encoder Representations from Transformers (BERT) untuk klasifikasi intent pengguna. Data yang digunakan meliputi data katalog buku serta data pertanyaan pengguna yang dikategorikan ke dalam beberapa intent. Hasil pelatihan model menunjukkan performa yang baik dengan akurasi sebesar 91% pada data uji dan nilai F1-score sebesar 0,90. Implementasi chatbot berbasis web menggunakan framework Flask menunjukkan bahwa sistem mampu merespons pertanyaan pengguna secara otomatis, akurat, dan kontekstual. Dengan demikian, chatbot yang dikembangkan dapat meningkatkan efisiensi layanan perpustakaan serta mempermudah akses informasi bagi siswa tanpa batasan waktu dan tempat.

**Kata kunci** : chatbot, machine learning, natural language processing, BERT, perpustakaan

### A. Pendahuluan

Artificial Intelligence (AI) awalnya lebih banyak berkembang sebagai bidang teoretis, namun seiring meningkatnya digitalisasi, AI mulai memberikan dampak nyata dalam berbagai aspek kehidupan, termasuk interaksi manusia dan komputer [1]. Perkembangan teknologi komputer yang pesat telah mendorong pemanfaatannya di berbagai sektor, khususnya dalam layanan informasi. Hal ini menuntut peningkatan kualitas sumber daya manusia agar mampu beradaptasi dengan teknologi yang semakin kompleks [2].

Salah satu implementasi AI yang berkembang pesat adalah chatbot. Chatbot merupakan agen percakapan virtual yang mampu meniru komunikasi manusia melalui teks atau suara. Dengan dukungan Machine Learning (ML) dan Natural Language Processing (NLP), chatbot dapat memahami bahasa manusia secara lebih alami serta terus belajar dari interaksi pengguna [3]. Teknologi ini memungkinkan chatbot untuk mengidentifikasi maksud pengguna, seperti pencarian informasi ketersediaan buku, sehingga dapat memberikan respon yang relevan dan cepat [4].

Dalam dunia pendidikan, pemanfaatan teknologi menjadi kebutuhan penting untuk meningkatkan kualitas layanan dan proses pembelajaran [5]. Chatbot berbasis ML dan NLP memiliki potensi besar untuk mendukung layanan informasi yang lebih efektif, modern, dan mudah diakses. Selain membantu penyedia layanan, chatbot juga memberikan kemudahan

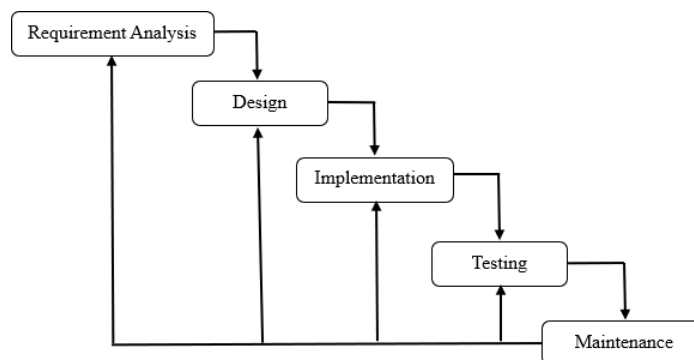
bagi pengguna dalam memperoleh informasi secara efisien tanpa batasan waktu dan tempat [6].

Penelitian ini dilakukan di perpustakaan SMA Negeri 1 Kelapa, Bangka Belitung, yang masih menggunakan sistem pelayanan manual. Kondisi tersebut mengharuskan siswa datang langsung ke perpustakaan untuk memperoleh informasi. Oleh karena itu, pengembangan chatbot menjadi solusi inovatif untuk meningkatkan efisiensi layanan. Penerapan chatbot di lingkungan pendidikan menengah juga menjadi nilai kebaruan penelitian ini.

Dengan adanya chatbot, diharapkan proses pelayanan perpustakaan menjadi lebih optimal, membantu tugas staf, serta mempermudah siswa dalam mengakses informasi. Selain itu, penelitian ini diharapkan dapat mendorong transformasi digital di sekolah dan meningkatkan kualitas layanan pendidikan secara keseluruhan.

## B. Metode

Jenis penelitian ini merupakan penelitian kualitatif deskriptif, dimana data yang dikumpulkan berupa kata-kata, teks, bukan angka-angka [7]. Metode kualitatif dapat digunakan dalam eksplorasi kebutuhan pengguna dan pemahaman kebutuhan terkait dengan masalah yang bisa diselesaikan oleh chatbot [8]. Penelitian ini memakai metode kualitatif karena dalam konteks pengembangan chatbot perpustakaan, metode kualitatif dapat membantu memahami bagaimana pengguna (seperti siswa dan staf perpustakaan) dapat berinteraksi dengan chatbot dan apa yang mereka butuhkan dari aplikasi tersebut. Pada tahap pengembangannya metode waterfall akan digunakan juga pada penelitian ini. Metode waterfall ini merupakan salah satu teknik yang sering digunakann pada berbagai pengembangan perangkat lunak [9]. Tahapan dalam model ini dimulai dari tahap perencanaan hingga tahap pengelolaan (maintenance) dan dilakukan secara bertahap. Pada tahapannya metode waterfall memiliki 5 tahapan, yaitu analisis kebutuhan (requirement analysis), desain (design), implementasi (implementation), (pengujian) testing, dan pemeliharaan (maintenance) [10].



Gambar 1. Tahapan Metode Waterfall.

### a. Requirement analysis

Pada tahap ini, pengumpulan informasi dilakukan untuk menganalisis kebutuhan yang dibutuhkan oleh sistem. Pengambilan data seperti data buku, pertanyaan yang sering di ajukan siswa merupakan data yang dibutuhkan dalam pengembangan *chatbot* untuk perpustakaan serta bahasa pemrograman yang digunakan untuk pengkodean nantinya.

**b. Design**

Setelah keperluan dikumpulkan, pada bagian ini berfokus pada merancang arsitektur sistem dan antarmuka pengguna atau tampilan *chatbot*. Desain sistem mencakup beberapa komponen aplikasi yang akan bekerja satu sama lain dari gabungan desain sistem,

**c. Implementation**

Pada fase ini dilakukan pemrosesan pengkodean di mana tadinya desain sistem yang telah dibuat diubah ke dalam bentuk kode program (Aceng Abdul Wahid, 2020). Bahasa pemrograman *python* akan digunakan pada pengembangan aplikasi *chatbot* ini. Penggunaan bahasa pemrograman seperti *python* ini digunakan karena banyak memiliki berbagai pustaka dan *framework* yang dapat digunakan untuk pengembangan aplikasi chatbot ini. Fase ini dilakukan agar *chatbot* bisa bekerja secara otomatis dengan menggunakan algoritma ML dan NLP. ML digunakan pada tahap ini supaya chatbot dapat belajar terus menerus dari pertanyaan pertanyaan yang sering di ajukan oleh pengguna. Algoritma yang digunakan pada penelitian ini menggunakan *supervised learning*, karena *supervised learning* bekerja dengan data berlabel, di mana setiap input (teks pengguna) memiliki label (intent atau kategori). Kemudian NLP yang akan memproses bahasa atau pertanyaan tadi ke dalam bentuk bahasa manusia yang lebih alami supaya bisa lebih mudah di pahami oleh pengguna nantinya.

**d. Testing**

Setelah setiap komponen sistem dikembangkan, langkah selanjutnya menguji *chatbot* untuk memastikan bahwa sistem *chatbot* berjalan sesuai seperti yang diinginkan. Pengujian ini dilakukan untuk menemukan bug atau kesalahan dalam pengkodean dan memastikan bahwa sistem *chatbot* berjalan dengan baik dalam berbagai kondisi.

**e. Maintenance**

Tahap ini adalah tahap dimana *chatbot* sesudah diimplementasikan, di mana *chatbot* akan dipantau secara teratur dan diperbaiki jika ditemukan bug atau kesalahan pada kode. Pemeliharaan juga mencakup peningkatan pada fitur dan penyesuaian *chatbot* berdasarkan umpan balik pengguna maupun perubahan kebutuhan.

**C. Hasil dan Pembahasan**

**a. Data**

Dalam pengembangan chatbot perpustakaan SMA N 1 Kelapa, teknik pengambilan data dilakukan melalui beberapa metode yang bertujuan untuk memastikan bahwa chatbot dapat memberikan informasi yang akurat dan relevan kepada pengguna. Beberapa data yang dikumpulkan seperti, data buku perpustakaan, beberapa pertanyaan yang sering ditanyakan oleh siswa, dan teks yang digunakan untuk melatih model natural language processing (NLP). Daftar buku perpustakaan, yang diambil langsung dari katalog perpustakaan sekolah, adalah informasi pertama yang digunakan. Detail penting seperti judul buku, penulis, penerbit, dan jumlah halaman disertakan dalam katalog ini. Untuk memberikan informasi kepada siswa yang menginginkannya tentang ketersediaan buku, data ini digunakan untuk meningkatkan basis pengetahuan chatbot.

Tabel 1. Sampel data yang diambil

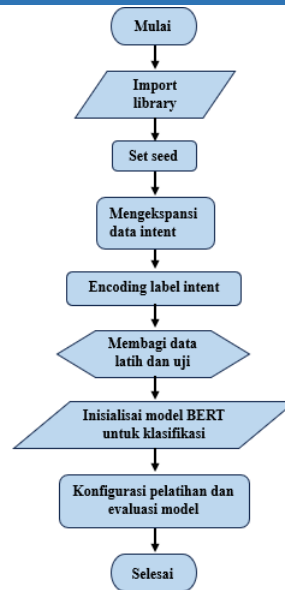
Judul Buku	Penulis	Penerbit	Tahun Terbit	Jumlah Halaman
Kamis Yang Manis	John steinbeck	Basabasi	2019	400hlm
Daun Terakhir	O.henry	Diva Press	2020	156hlm
Cinta Dalam Mimpi	Muyassarotul Hafidzoh	Diva Press	2020	228hlm
Cawan Emas	John steinbeck	Basabasi	2022	288hlm
Monyet Bercerita	Aris kurniawan	Basabasi	2019	192hlm
Trik lolos spmb	Jahen f. rezki	Baduose media	2008	79hlm
Pintar psikotes	Seruni gerri nirwana	Dunia cerdas	2016	340hlm
Buku pintar menghitung pajak	Yustinus prastowo	Raih asa sukses	2016	380hlm
Green economy	Andreas lako	erlangga	2014	156hlm
Dealova	Dyan nuranindya	Gramedia	2024	312hlm
Harus bisa	Dino patti djalal	Red & white publications	2003	437hlm
Mendesain pembelajaran	Trianto, M.pd	Kencana prenatal media	2010	376hlm

Data buku ini kemudian diolah menjadi format pertanyaan dan jawaban dalam intents yang digunakan sebagai dasar untuk melatih chatbot agar mampu merespons pertanyaan pengguna mengenai ketersediaan koleksi buku di perpustakaan secara lebih akurat dan kontekstual. Selain itu, data intent dan respons dikumpulkan melalui wawancara dengan petugas perpustakaan serta observasi terhadap pertanyaan yang sering diajukan oleh siswa.

Pertanyaan-pertanyaan ini kemudian dikategorikan ke dalam intent, seperti informasi tentang buku. Masing-masing intent memiliki daftar patterns (pola pertanyaan) dan responses (jawaban yang diberikan chatbot) yang disusun dalam file intents.csv. Untuk memungkinkan chatbot mengidentifikasi pola kata dan memahami hubungan antar kata dalam konteks tertentu, model ini dilatih menggunakan teks yang dikumpulkan dari berbagai sumber. Dengan teknik ini, chatbot tidak hanya mengandalkan pencocokan kata kunci, tetapi juga dapat memahami makna dari kalimat yang lebih kompleks.

## b. Pelatihan model BERT

Pelatihan model BERT (Bidirectional Encoder Representasion From Transformers) untuk mengenali maksud atau intent dari pertanyaan pengguna merupakan Salah satu bagian penting dalam pengembangan chatbot. Adapun alur pelatihan model bert ini yang di gambarkan dalam bentuk flowchart yang bisa di lihat pada gambar 2 di bawah ini:



Gambar 2. Flowchart Pelatihan Model Bert

Langkah pertama adalah mengimpor seluruh library yang dibutuhkan. Modul random, numpy, dan torch digunakan untuk operasi numerik dasar serta pengaturan seed agar hasil eksperimen konsisten atau reproducible. Pandas dimanfaatkan untuk membaca data dari file CSV, seperti memisahkan dan menata ulang pola kalimat. Library sklearn menyediakan alat untuk mengubah label teks menjadi angka (dengan LabelEncoder), membagi data secara terstratifikasi, dan mengevaluasi performa model menggunakan metrik seperti akurasi dan F1-score. Paket transformers dari Hugging Face digunakan untuk memuat tokenizer dan model BERT yang sudah dilatih sebelumnya, untuk memahami bahasa alami. Objek Dataset dari library dataset membantu mengubah data hasil tokenisasi ke format yang bisa langsung digunakan oleh Trainer. Terakhir, joblib digunakan untuk menyimpan encoder label ke file, agar dapat digunakan kembali saat proses inferensi nanti.

Penetapan seed secara eksplisit di berbagai pustaka memastikan bahwa hasil pelatihan akan sama setiap kali dijalankan. Hal ini penting untuk penelitian agar validasi dan pengujian model tetap konsisten. Lalu menggunakan pustaka Pandas untuk membaca file CSV bernama intents.csv. File ini berisi data intent chatbot dalam bentuk table. Hasil pembacaan file disimpan ke dalam variabel df, yang merupakan objek DataFrame dari Pandas dan berfungsi seperti tabel dalam memori. Namun karena setiap baris pattern berisi banyak kalimat yang dipisahkan dengan tanda|, maka proses ekspansi dilakukan dengan .split() dan .explode() untuk memastikan satu baris hanya mengandung satu pattern. Hal ini meningkatkan jumlah data latih dan memperkaya variasi kalimat untuk tiap intent, sehingga model BERT dapat belajar lebih akurat.

Setelah data pattern dipecah, label intent (tag) dalam bentuk teks diubah menjadi label numerik. Proses encoding ini dilakukan dengan LabelEncoder, yang menghasilkan output numerik dari setiap label kategorikal. Encoder ini disimpan sebagai file .pkl agar bisa digunakan saat proses inferensi pada chatbot.

Dataset yang awalnya belum dilatih kemudian dibagi menjadi 80% data pelatihan, karena model menggunakan data ini untuk belajar mengenali pola intent. Sedangkan 20% data pengujian data ini tidak pernah dilihat oleh model saat dilatih karena digunakan untuk mengukur performa model secara objektif. Teknik stratify memastikan bahwa proporsi jumlah intent di kedua subset tetap seimbang, mencegah bias selama pelatihan.

```

pattern,label,tag
Saya sedang mencari buku,1,cek_buku
Buku ... terbit tahun berapa?,2,info_buku
Nomor ISBN buku ... berapa?,2,info_buku
Bagaimana cara meminjam buku?,0,aturan_peminjaman
Halo,9,salam
Terima kasih infonya,13,terima_kasih
Bisa tunjukkan semua buku,10,semua_buku
Buku ... karya siapa?,6,nama_penulis
Letak perpustakaan,5,lokasi_perpustakaan
Buku ... ada nggak,1,cek_buku
Ada daftar buku cerita pendek,4,jenis_buku
Tampilkan koleksi buku,10,semua_buku Col 3: tag
Apakah buku ... bisa saya pinjam,8,pinjam_buku
Ada jenis buku komik,4,jenis_buku
bolehkah ribut di perpustakaan,12,tatib
Saya mau tahu jam operasional perpustakaan,3,jam_buka
Apa saja isi rak buku,10,semua_buku
Siapa pengarang buku ...?,6,nama_penulis
Tampilkan buku novel,4,jenis_buku
Perpustakaan terletak di mana,5,lokasi_perpustakaan
Arah menuju perpustakaan,5,lokasi_perpustakaan
Saya ingin meminjam buku berjudul ...,8,pinjam_buku
Buku ... masih bisa dipinjam,8,pinjam_buku
Ada buku dari ...,7,penulis_buku
Makasih,13,terima_kasih
Ada buku komik terbaru,4,jenis_buku

```

Gambar 3. data latih

```

pattern,label,tag
Jam buka dan tutup perpustakaan,3,jam_buka
Apakah boleh minum di ruang baca,12,tatib
Saya mau cari buku ...,1,cek_buku
Daftar buku karya ... Apa,7,penulis_buku
Ada buku yang saya cari,1,cek_buku
Yang nulis buku ... Siapa,6,nama_penulis
Aku minjam buku ... dong,8,pinjam_buku
Saya ingin tahu tentang buku ...,2,info_buku
Perpustakaan sekolah ini di mana ya,5,lokasi_perpustakaan
Ada buku berjudul ... di perpustakaan,1,cek_buku
Ada buku pelajaran apa saja,4,jenis_buku
Buku ... itu siapa yang tulis,6,nama_penulis
Apakah perpustakaan punya buku ...,1,cek_buku
"Halo, boleh tanya sesuatu",9,salam
Siapa penerbit buku ...?,2,info_buku
Coba tunjukkan seluruh koleksi buku,10,semua_buku
Perlihatkan buku belajar,4,jenis_buku
Boleh ribut atau ngobrol di perpustakaan,12,tatib
apa syarat menjadi anggota perpustakaan,11,syarat_anggota
Perlihatkan semua buku,10,semua_buku
Bagaimana cara meminjam buku ...,8,pinjam_buku
Bisa info jam layanan perpustakaan,3,jam_buka
Bisa kasih daftar buku tulisan ...,7,penulis_buku
Berapa lama waktu pinjam buku?,0,aturan_peminjaman
Saya ingin pinjam buku novel,4,jenis_buku
"Hai bot, bisa bantu",9,salam

```

Gambar 4. data uji

BERT hanya dapat menerima input dalam bentuk token numerik. Maka dari itu, tokenizer digunakan untuk memotong teks ke dalam token sesuai kebutuhan model, serta memberikan padding agar panjang input seragam. Tokenisasi dilakukan pada data pelatihan dan pengujian. Model klasifikasi BERT dimuat dari model pra-latih "bert-base-uncased", dan jumlah label disesuaikan dengan jumlah kelas intent. Model ini akan belajar mengklasifikasikan input teks ke dalam salah satu label intent. Parameter pelatihan didefinisikan dalam TrainingArguments. Di antaranya adalah jumlah epoch (10), ukuran batch, strategi evaluasi tiap epoch, dan seed untuk konsistensi. Model tidak disimpan tiap epoch karena hanya ingin menyimpan versi akhir.

### c. Evaluasi model

Setelah proses pelatihan model chatbot berbasis machine learning selesai dilakukan, tahap selanjutnya melakukan evaluasi terhadap performa model chatbot yang menggunakan dua jenis data, yaitu data latih dan data uji. Evaluasi ini menggunakan metrik standar klasifikasi, yaitu precision, recall, dan f1-score yang disediakan oleh pustaka sklearn. Model yang digunakan dalam penelitian ini adalah BERT, sedangkan data yang akan di evaluasi yaitu data uji yang berasal dari test\_data dan data latih berasal dari file train\_data.csv yang telah diperbaiki dan diseimbangkan sebelumnya. Evaluasi dilakukan untuk mengukur sejauh mana kemampuan model dalam mengklasifikasikan intent secara akurat. Hasil evaluasi dapat dilihat pada gambar 5 dan 6 di bawah ini:

	precision	recall	f1-score	support
aturan_peminjaman	1.00	1.00	1.00	2
cek_buku	1.00	0.83	0.91	6
info_buku	1.00	0.75	0.86	4
jam_buka	1.00	0.75	0.86	4
jenis_buku	0.75	0.75	0.75	4
lokasi_perpus	0.80	1.00	0.89	4
nama_penulis	1.00	1.00	1.00	4
penulis_buku	0.80	1.00	0.89	4
pinjam_buku	1.00	1.00	1.00	4
salam	0.75	1.00	0.86	3
semua_buku	0.80	1.00	0.89	4
syarat_anggota	1.00	1.00	1.00	3
tatib	1.00	1.00	1.00	3
terima_kasih	1.00	0.75	0.86	4
accuracy			0.91	53
macro avg	0.92	0.92	0.91	53
weighted avg	0.92	0.91	0.91	53

Gambar 5. Evaluasi data uji

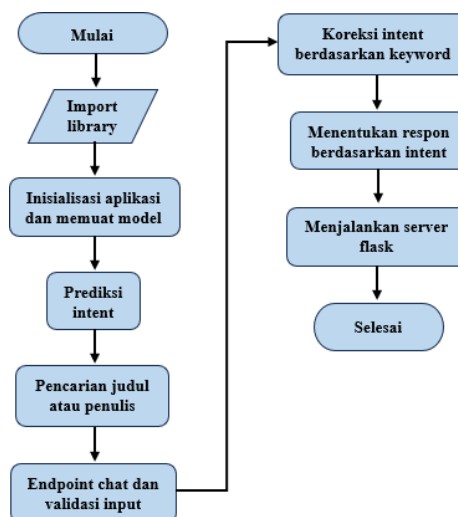
	precision	recall	f1-score	support
aturan_peminjaman	1.00	1.00	1.00	8
cek_buku	1.00	1.00	1.00	23
info_buku	1.00	1.00	1.00	16
jam_buka	1.00	1.00	1.00	13
jenis_buku	1.00	1.00	1.00	16
lokasi_perpus	1.00	1.00	1.00	17
nama_penulis	1.00	1.00	1.00	16
penulis_buku	1.00	1.00	1.00	16
pinjam_buku	1.00	1.00	1.00	16
salam	1.00	1.00	1.00	13
semua_buku	1.00	1.00	1.00	16
syarat_anggota	1.00	1.00	1.00	12
tatib	1.00	1.00	1.00	12
terima_kasih	1.00	1.00	1.00	17
accuracy			1.00	211
macro avg	1.00	1.00	1.00	211
weighted avg	1.00	1.00	1.00	211

Gambar 6. Evaluasi data uji

Hasil evaluasi dataset menunjukkan bahwa model memiliki kemampuan generalisasi yang cukup baik. Pada evaluasi terhadap data latih, model mencapai performa sempurna dengan nilai presisi, recall, dan f1-score sebesar 100% untuk hampir semua kelas, serta akurasi total 100%. Hal ini menandakan bahwa model mampu mengenali pola dalam data latih dengan sangat baik. Namun, saat diuji dengan data uji yaitu data yang tidak pernah dilihat sebelumnya, terjadi sedikit penurunan performa. Meskipun begitu, nilai akurasi tetap tinggi, yakni sebesar 91%, dengan rata-rata f1-score juga sebesar 91%. Penurunan ini tergolong normal dan menunjukkan bahwa model tidak mengalami overfitting secara signifikan.

#### d. Pengembangan Chatbot

Berikut adalah implementasi program utama chatbot dalam file chatbot.py, yang dibangun menggunakan bahasa pemrograman Python dengan framework Flask dan pustaka machine learning. Program ini bertanggung jawab untuk memuat data intent, melatih model klasifikasi, dan melayani permintaan pengguna melalui endpoint API. Adapun alur pengembangan chatbot ini yang di gambarkan dalam bentuk flowchart yang bisa di lihat pada gambar 7 di bawah ini:



Gambar 7. Flowchart Pengembangan Chatbot

Pengembangan chatbot berbasis web dalam penelitian ini diawali dengan proses impor berbagai pustaka (library) yang diperlukan untuk mendukung keseluruhan sistem. Framework Flask digunakan sebagai fondasi utama dalam membangun server API yang memungkinkan komunikasi antara backend dan frontend. Untuk mengatasi keterbatasan akses lintas domain, khususnya pada integrasi dengan aplikasi eksternal seperti Android, digunakan modul Flask-CORS. Selain itu, pustaka pandas dimanfaatkan untuk mengelola dan membaca data dalam format CSV yang berisi informasi koleksi buku perpustakaan. Dalam aspek pemodelan kecerdasan buatan, torch digunakan sebagai kerangka kerja utama berbasis PyTorch untuk menjalankan model deep learning. Sementara itu, model klasifikasi intent berbasis BERT diakses melalui pustaka transformers, dan joblib digunakan untuk memuat label encoder yang telah dilatih sebelumnya.

Tahap berikutnya adalah inisialisasi aplikasi serta pemuatan model yang telah dikembangkan. Aplikasi Flask diinisialisasi sebagai server utama, kemudian dilanjutkan dengan pemanggilan model pemrosesan bahasa alami menggunakan spaCy untuk mendukung ekstraksi entitas dari teks. Dataset buku dimuat dari file CSV ke dalam struktur dataframe menggunakan pandas, sehingga memudahkan proses pencarian dan manipulasi data. Model BERT yang telah dilatih sebelumnya dimuat dari direktori penyimpanan, bersama dengan tokenizer yang berfungsi mengubah teks input menjadi representasi numerik yang dapat diproses oleh model. Label encoder juga dimuat untuk mengonversi hasil prediksi model yang berbentuk numerik menjadi label intent yang dapat dipahami secara semantik.

Proses prediksi intent dilakukan melalui sebuah fungsi khusus yang menerima input teks dari pengguna. Teks tersebut terlebih dahulu ditokenisasi menjadi tensor menggunakan tokenizer, kemudian diproses oleh model BERT untuk menghasilkan nilai logits. Nilai ini selanjutnya dikonversi menjadi probabilitas menggunakan fungsi softmax, dan intent dengan probabilitas tertinggi dipilih sebagai hasil prediksi. Untuk meningkatkan keandalan sistem, diterapkan ambang batas (threshold) tertentu; apabila nilai kepercayaan (confidence) berada di bawah batas tersebut, maka sistem akan mengklasifikasikan input sebagai "unknown". Pendekatan ini penting untuk menangani input yang ambigu atau berada di luar cakupan data pelatihan model.

Selain prediksi berbasis model, sistem juga dilengkapi dengan mekanisme pencarian manual terhadap judul dan penulis buku. Fungsi ini bekerja dengan mengubah input pengguna menjadi huruf kecil, kemudian mencocokkannya dengan data dalam dataframe melalui proses iterasi. Apabila ditemukan kecocokan antara input dengan judul atau nama penulis, sistem akan mengembalikan informasi lengkap terkait buku tersebut, seperti penerbit, tahun terbit, dan jumlah halaman. Jika tidak ditemukan kecocokan, maka fungsi akan mengembalikan nilai kosong sebagai indikasi bahwa data tidak tersedia.

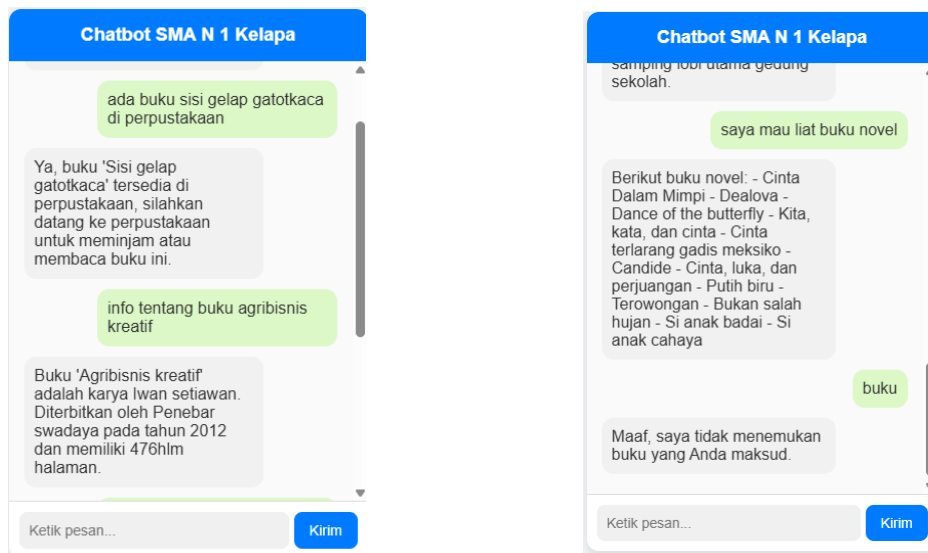
Dalam implementasinya, chatbot menyediakan endpoint API pada rute "/chat" yang hanya menerima metode POST. Input dari pengguna dikirim dalam format JSON dan diekstraksi pada bagian message. Sebelum diproses lebih lanjut, sistem melakukan validasi panjang input untuk memastikan bahwa teks yang diberikan memiliki makna yang cukup. Jika panjang input kurang dari tiga karakter, sistem akan memberikan respons bahwa pertanyaan terlalu singkat dan meminta pengguna untuk memberikan penjelasan yang lebih lengkap.

Untuk meningkatkan akurasi klasifikasi intent, sistem mengombinasikan pendekatan berbasis machine learning dengan aturan berbasis kata kunci (rule-based). Misalnya, jika input mengandung kata "buku" dan kata-kata tertentu seperti "penulis" atau "info", maka intent akan diarahkan menjadi "info\_buku". Demikian pula, jika terdapat kata "pinjam",

maka intent akan disesuaikan menjadi “pinjam\_buku”. Pendekatan hibrida ini terbukti efektif dalam mengurangi kesalahan prediksi yang mungkin terjadi pada model.

Tahap selanjutnya adalah penentuan respons chatbot berdasarkan intent yang telah diperoleh. Untuk intent terkait pengecekan ketersediaan buku, sistem akan memberikan konfirmasi apakah buku tersebut tersedia atau tidak. Sementara itu, untuk intent yang berkaitan dengan informasi buku, sistem akan menyajikan detail lengkap mengenai buku, termasuk penulis, penerbit, tahun terbit, dan jumlah halaman. Jika informasi tidak ditemukan, sistem akan memberikan respons yang sesuai agar pengguna dapat memperjelas pertanyaannya.

Akhirnya, aplikasi dijalankan menggunakan server Flask dalam mode debug. Mode ini memungkinkan pengembang untuk memantau kesalahan secara langsung serta melakukan pembaruan kode secara otomatis tanpa perlu me-restart server secara manual. Fitur ini sangat mendukung proses pengembangan dan pengujian sistem chatbot secara iteratif.



Gambar 8. Hasil Pengembangan

Pada gambar 8 yang menunjukkan, chatbot telah berjalan dengan baik dan mampu merespon input pengguna. Tampilan chatbot didesain dengan antarmuka yang sederhana dan mudah digunakan. Bagian atas chatbot menampilkan header berwarna biru dengan tulisan "Chatbot SMA N 1 Kelapa", yang memberi identitas visual kepada pengguna. Pengguna dapat memasukkan pertanyaan dan mengirimkannya ke chatbot menggunakan kotak input teks dan tombol berwarna biru "Kirim" di bagian bawah.

Berdasarkan hasil pengujian sistem chatbot, yang ditampilkan pada gambar, menunjukkan bahwa chatbot telah mampu memberikan jawaban yang cukup akurat dan sesuai dengan konteks untuk berbagai pertanyaan pengguna. Ketika pengguna bertanya tentang ketersediaan buku "Sisi Gelap Gatotkaca" pada pertemuan awal, chatbot menjawab dengan tepat bahwa buku tersebut tersedia dan menyarankan pengguna untuk mengunjungi perpustakaan. Tanggapan ini menunjukkan bahwa sistem dapat secara akurat mengidentifikasi objek judul buku dan menentukan maksud pertanyaan (maksud cek\_buku).

Selain itu, tanggapan chatbot yang lebih menyeluruh terhadap pertanyaan kedua tentang informasi buku "Agribisnis kreatif" chatbot dapat merespon dengan memberikan informasi tentang buku tersebut. Namun, saat pengguna mengetikkan "buku", chatbot menampilkan pesan "Maaf, saya tidak menemukan buku yang anda maksud". Hal ini menunjukkan bahwa

sistem belum memiliki informasi yang cukup terkait topik yang dimaksud oleh pengguna. Fitur ini merupakan bagian dari fallback response, yang memungkinkan chatbot memberikan tanggapan ketika tidak menemukan jawaban yang sesuai dalam database.

#### D. Simpulan

Penelitian ini bertujuan untuk melakukan pengembangan aplikasi chatbot untuk layanan perpustakaan dengan memanfaatkan algoritma machine learning dan natural language processing. BERT digunakan pada penelitian ini karena cocok untuk menjawab pertanyaan dan klasifikasi teks. Berdasarkan hasil perancangan, pengembangan serta pengujian dari aplikasi chatbot untuk layanan informasi perpustakaan, dengan menerapkan teknologi Machine Learning dan Natural Language Processing terbukti mampu dalam membangun sistem interaktif yang mampu merespons pertanyaan pengguna secara otomatis. Dengan memanfaatkan model BERT sebagai bagian dari proses pelatihan data intents, chatbot ini dapat mengenali maksud pengguna dan memberikan respon yang sesuai. Dari hasil evaluasi model klasifikasi intent chatbot menggunakan algoritma BERT, diperoleh akurasi keseluruhan sebesar 91% dengan nilai macro average F1-score sebesar 0.90, yang menunjukkan bahwa model mampu mengenali dan mengklasifikasikan intent pengguna dengan baik. Dengan demikian, dapat disimpulkan bahwa tujuan penelitian telah tercapai, baik dalam hal pengembangan aplikasi chatbot maupun penerapan teknologi machine learning dan NLP. Chatbot sudah bisa merespon pertanyaan pengguna yang di ajukan oleh pengguna ke chatbot.

#### Daftar Pustaka

- [1] Y. B. Pratama and N. P. Dalimunthe, "Implementasi Teknik Computer Vision Untuk Deteksi Viridiplantae Pada Lahan Pasca Tambang," vol. 3, no. 1, pp. 64–72, 2022, doi: 10.47065/bulletincsr.v3i1.193.
- [2] Y. B. Pratama, N. P. Dalimunthe, and M. Sukma, "Prediksi Spasial Kerapatan Vegetasi Perkotaan dengan Pendekatan Algoritma Time Series Untuk Mendukung Pertumbuhan Ekonomi Hijau," vol. 6, no. 2, pp. 1495–1507, 2025, doi: 10.47065/josh.v6i2.6251.
- [3] M. Mustaqim, A. Gunawan, Y. B. Pratama, and I. Zaliman, "Pengembangan Chatbot Layanan Publik Menggunakan Machine Learning Dan Natural Language Processing," *J. Inf. Technol. Soc.*, vol. 1, no. 1, pp. 1–4, 2023, doi: 10.35438/jits.v1i1.16.
- [4] Yuniarni Yuniarni, Yudistira Bagus Pratama, and Arvi Pramudyantoro, "Pengembangan Virtual Assistant menggunakan Teknologi NLP dengan Metode Algoritma Machine Learning untuk Layanan Informasi Akademik di SMA Negeri 1 Parittiga Berbasis Web," *Mars J. Tek. Mesin, Ind. Elektro Dan Ilmu Komput.*, vol. 3, no. 5, pp. 205–219, 2025, doi: 10.61132/mars.v3i5.1152.
- [5] Julita and P. Dheni Purnasari, "Pemanfaatan Teknologi Sebagai Media Pembelajaran Dalam Pendidikan Era Digital," *J. Educ. Learn. Innov.*, vol. 2, no. 2, pp. 227–239, 2022, doi: 10.46229/elia.v2i2.460.
- [6] N. Rokhman, P. A. Maulan, and N. A. Wirahuda, "Analisis Penilaian Esai Secara Otomatis Menggunakan Natural Language Processing (NLP) dan Cosine Similarity," *Go Infotech J. Ilm. STMIK AUB*, vol. 31, no. 1, pp. 41–52, 2025.
- [7] S. W. Purwanza *et al.*, *Metodologi Penelitian Kuantitatif, Kualitatif, dan Kombinasi*, no. March. 2022.

- [8] A. Adil, *Metode Penelitian Kuantitatif dan Kualitatif: Teori dan Praktik*, no. January. 2023.
- [9] A. A. Maulana, A. Susanto, and D. P. Kusumaningrum, “Rancang Bangun Web Scraping Pada Marketplace di Indonesia,” *JOINS (Journal Inf. Syst.*, vol. 4, no. 1, pp. 41–53, 2019, doi: 10.33633/joins.v4i1.2544.
- [10] W. Feraldi, S. Andriyanto, and Y. A. Rindri, “Rancang Bangun Sistem Informasi Digital Fundraising LAZISMU Universitas Muhammadiyah Bangka Belitung,” *JSITIK J. Sist. Inf. dan Teknol. Inf. Komput.*, vol. 1, no. 1, pp. 1–16, 2022, doi: 10.53624/jsitik.v1i1.151.