

Implementasi Question Answering Berbasis Chatbot Telegram Pada Tafsir Al-Jalalain Menggunakan Langchain dan LLM

Febrian Rizki Adi Sutiyo, Nazruddin Safaat Harahap*, Surya Agustian, Reski Mai Candra

Fakultas Sains dan Teknologi, Teknik Informatika, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

Email: ¹12050116997@students.uin-suska.ac.id, ^{2,*}nazruddin.safaat@uin-suska.ac.id, ³surya.agustian@uin-suska.ac.id,

⁴reski.candra@uin-suska.ac.id

Email Penulis Korespondensi: nazruddin.safaat@uin-suska.ac.id

Abstrak—Perkembangan teknologi menjadi sangat penting untuk pengambilan informasi yang efisien, akurat dan cepat. Tafsir Al-Jalalain menjadi salah satu Tafsir Al-Qur'an yang terkenal, dan digunakan sebagai sumber pedoman hidup umat muslim. Untuk mendapatkan informasi tentang tafsir bisa melalui media informasi seperti internet atau pada ahli Tafsir Al-Qur'an. Namun untuk mendapatkan informasi tersebut butuh waktu yang banyak untuk melakukan penyaringan terhadap informasi yang efisien, akurat dan cepat. Dari permasalahan tersebut dibutuhkan sistem yang mampu menjawab pertanyaan manusia secara akurat, efektif dan cepat. Pada penelitian ini tentang implemetansi *Question Answering* berbasis *Chatbot* Telegram menggunakan *langchain* dan *LLM* diharapkan mampu menjadi solusi untuk memberikan informasi Tafsir Al-Jalalain yang akurat, efektif dan cepat. Sistem *Question Answering* akan melakukan pembelajaran pada data Tafsir Al-Jalalain dengan menggunakan model Bahasa yaitu *Large Language Model* sehingga diharapkan mampu memberikan informasi yang akurat, efektif dan cepat. Hasil evaluasi dari penelitian dengan melakukan penyebaran kusioner kepada mahasiswa jurusan Ilmu Al-Qur'an dan Tafsir UIN SUSKA Riau sebanyak tujuh responden mendapatkan persentase sebesar 84.29%.

Kata Kunci: Tafsir Al-Jalalain; Question Answering; Chatbot; Langchain; LLM

Abstract—Technological developments are very important for efficient, accurate and fast information retrieval. Tafsir Al-Jalalain is one of the famous Tafsir Al-Qur'an, and is used as a source of life guidance for muslims. To get information about tafsir, you can go through information media such as the internet or from experts in Tafsir Al-Qur'an. However, to get information it takes a lot of time to filter the information efficiently, accurately and quickly. This problem requires a system that is able to answer human questions accurately, effectively and quickly. In this research, it is hoped that the implementation of telegram Chatbot-based Question Answering using Langchain and LLM will be a solution for providing information on Tafsir Al-Jalalain that is accurate, effective and fast. The Question Answering system will carry out learning on the Tafsir Al-Jalalain data using a language model, namely the Large Language Model, so that it is expected to be able to provide accurate, effective and fast information. The evaluation results of the research by distributing questionnaires to students majoring in Al-Qur'an and Tafsir Science at UIN SUSKA Riau, as many as seven respondents, obtained a percentage of 84.29%.

Keywords: Tafsir Al-Jalalain; Question Answering; Chatbot; Langchain; LLM

1. PENDAHULUAN

Perkembangan teknologi informasi dan telekomunikasi saat ini berkembang sangat pesat yang dapat disalurkan dengan mudah dan cepat [1]. Teknologi informasi dan telekomunikasi menjadi salah satu pengaruh besar bagi dunia saat ini [2]. Teknologi informasi dan telekomunikasi memberikan kemudahan kepada Masyarakat dalam mendapatkan dan memahami informasi, seperti dari sumber internet [3]. Internet memiliki kumpulan data informasi yang sangat besar, untuk mendapatkan informasi yang akurat, sangat penting memiliki pengetahuan dan waktu yang cukup untuk mendapatkan informasi yang ingin dicari, karena harus membaca dan melakukan proses penyaringan terhadap data informasi yang di inginkan.

Tafsir Al-Jalalain adalah salah satu kitab Tafsir Al-Qur'an yang bersumber kuat dari AL-Qur'an, kitab suci Al-Qur'an terdiri dari 30 Juz sehingga memiliki data yang sangat banyak [4]. Informasi mengenai Tafsir Al-Jalalain dapat diakses melalui teknologi dengan cara menjelajahi internet, maupun dengan melalui sarana klasik seperti konsultasi dengan ahli Tafsir Al-Qur'an. ketika ingin mendapatkan informasi yang akurat, dibutuhkan waktu yang banyak dikarenakan harus membaca dan melakukan proses penyaringan informasi terlebih dahulu sehingga baru mendapatkan informasi yang sesuai. Karena masalah tersebut, saat ini banyak pengembang sistem membangun sebuah sistem tanya jawab yang memungkinkan pengguna untuk bertanya mengenai kebutuhan informasi dengan Bahasa manusia, kemudian sistem akan mengembalikan informasi dalam bentuk teks kalimat sebagai jawabannya [5]. Sistem tanya jawab (*Question Answering*) dapat memberikan pengalaman pengguna dengan lebih baik dan membantu memberikan informasi yang relevan [6].

Question Answering merupakan salah satu sistem yang bisa menjadi solusi terhadap banyaknya Masyarakat yang membutuhkan informasi yang bisa di akses dengan cepat dan akurat [5]. Pada penelitian yang dilakukan oleh Made Agus Putra Subali dan Puritan Wijaya tentang Sistem *Question Answering* untuk Bahasa Bali menggunakan Metode *Rule-Based* dan *String Similarity*, penulis pada penelitian ini menggunakan kalimat tanya biasa dalam Bahasa bali. Data yang digunakan merupakan lima puluh dokumen dari Bahasa bali. Tahap pengujian menggunakan dua puluh kalimat tanya. Hasil dari penelitian ini memperoleh jawaban berdasarkan *rules* untuk setiap pertanyaan dan memperoleh akurasi sebesar 40% [7]. Penelitian lainnya tentang *Question Answering System* dalam bentuk Chatbot pada Platform Line untuk Mata Kuliah Pelajaran Sejarah SMA/MA dengan menggunakan Algoritma *Levenshten Distance*, penelitian ini bertujuan untuk membantu siswa SMA/MA dalam mencari informasi tentang mata Pelajaran Sejarah karena pemanfaatan buku pada mata

Pelajaran tersebut atau yang lainnya tidak begitu efektif dan efisien dalam pencarian informasinya. Penelitian ini menggunakan algoritma *levensthen distance* untuk mencocokkan kata kunci berdasarkan inputan pengguna dan kata kunci yang ada pada *knowledge base*. Hasil dari penelitian tersebut memperoleh akurasi sebesar 100% dengan pengujian menggunakan pertanyaan yang berasal dari *knowledge base*. Terdapat juga pertanyaan yang langsung ditanyakan oleh siswa SMA/MA dengan skor sebesar 72% [8].

Chatbot merupakan sebuah sistem yang dirancang untuk simulasi percakapan antara 2 pihak, yaitu sistem dan manusia. *Chatbot* sudah berdiri sejak tahun 1945 dan sampai saat ini terus berkembang menggunakan kecerdasan buatan. Salah satu penerapan *Chatbot* yang menggunakan kecerdasan buatan ialah *ChatGPT* yang dikembangkan *openAI* dan bermanfaat untuk membantu memahami respon tepat yang diinginkan manusia. *OpenAI* merupakan sebuah organisasi yang bertujuan mengembangkan teknologi kecerdasan buatan untuk memudahkan manusia dalam melakukan pekerjaan sehari-hari [9]. Terdapat penelitian terkait tentang Implementasi *Chatbot* Pada Pendaftaran Mahasiswa Baru Menggunakan *Recurrent Neural Network* dengan tujuan membangun *Chatbot* telegram otomatis yang dapat berkomunikasi dengan pengguna mengenai informasi pendaftaran mahasiswa baru di Universitas Pakuan menggunakan *Recurrent Neural Network (RNN)* untuk klasifikasi teks. Hasil pengujian pada penelitian ini dengan menggunakan data sebanyak 251 kalimat pertanyaan tentang pendaftaran mahasiswa baru di Universitas Pakuan. Hasil dari penelitian, *Chatbot* dapat menjawab pertanyaan dengan akurasi sebesar 88%, *presisi* sebesar 95% dan *recall* sebesar 92% [10]. Pada penelitian lainnya tentang *Open Artificial Intelligence Analysis Using Chat GPT Integrated with Telegram Bot* yang bertujuan untuk mengembangkan *ChatGPT* versi *Chatbot* telegram menggunakan algoritma *GPT-3*. Sehingga *ChatGPT* dapat diakses langsung melalui *Chatbot* Telegram, dan tidak perlu mengakses melalui website *ChatGPT* yang membutuhkan waktu lebih lama, namun penelitian ini hanya menganalisis perbandingan *ChatGPT* berbasis web, dan juga *ChatGPT* berbasis *Chatbot* telegram [9]. Pada penelitian yang dilakukan oleh R. Qardafil dan A. Sujarwo tentang Optimasi Kualitas *Chatbot* Jala Tech Melalui Implementasi *Auto-GPT* dalam aplikasi Jali yang berfungsi meningkatkan kinerja, akurasi, serta kepuasan pengguna. Penelitian ini menggunakan metode analisis komparatif antara *Auto-GPT* dan *ChatGPT* yang sudah ada. Hasilnya diharapkan bisa memberikan wawasan dan rekomendasi yang relevan untuk meningkatkan kualitas aplikasi Jali [6].

Large Language Model (LLM) adalah salah satu jenis model Bahasa kecerdasan buatan yang termasuk ke dalam *Natural Language Processing (NLP)*. *LLM* dapat membuat teks dengan kemampuan seperti layaknya manusia. *LLM* ditandai dengan banyak parameter dan dilatih pada Kumpulan teks yang besar, serta memiliki kemampuan untuk menghasilkan output yang relevan kontekstual dan gramatikal [11]. Sedangkan *langchain* merupakan kerangka kerja yang digunakan dalam pengembangan aplikasi dengan *LLM*. *Langchain* memungkinkan pengembang untuk membangun aplikasi menggunakan *LLM* untuk meningkatkan penyesuaian, akurat, dan relevansi dari model [12]. Penelitian yang pernah dilakukan oleh Oughuzan Topsakal dan T.Cetin Akinci (2023) tentang *Creating Large Language Model Applications Utilizing Langchain: A Primer on Developing LLM Apps Fast*. Penelitian ini memiliki peran besar dalam dunia pengembangan aplikasi dengan *Large Language Model (LLM)*. Penelitian ini juga menggunakan *Langchain* sebagai kerangka kerja yang memungkinkan pengembang untuk memanfaatkan *LLM* dengan baik, efektif dan efisien. *Langchain* juga dapat meningkatkan kualitas dan relevansi *output* yang dihasilkan. Melalui penelitian ini Topsakal dan Akinci menunjukkan bahwa *LLM* tidak hanya sebuah alat yang kuat untuk memproses teks, namun juga dapat diterapkan secara luas dalam berbagai konteks, termasuk penulisan kode dan tugas-tugas lainnya dalam bidang kecerdasan buatan [11].

Flask merupakan *microframework* yang berjalan pada Bahasa pemrograman *python* dengan tujuan untuk membantu para pengembang dalam membangun website dengan terstruktur [13]. Terdapat penelitian tentang Implementasi *FLASK* pada Sistem Penentuan Minimal Order untuk Tiap Item Barang di *Distribution Center* pada PT XYZ berbasis Website. Penelitian ini bertujuan untuk membangun sistem minimal order berbasis website dengan menerapkan Bahasa pemrograman *Python* dan *Framework FLASK*. Penelitian ini menggunakan metode *waterfall* sebagai tahap penelitiannya yang dimulai dari *System Analyst, Planning and Design System, Implementation* dan *Testing*. Sehingga mendapatkan hasil sebuah sistem yang berfungsi untuk mengatur minimal order tiap barang [14].

Berdasarkan penelitian sebelumnya yang telah diuraikan diatas, belum terdapat penelitian yang membahas secara menyeluruh tentang penerapan *Langchain* dan *LLM* dalam implementasi Sistem *Question Answering*, sehingga peneliti melakukan penelitian tentang Implementasi *Question Answering* berbasis *Chatbot* Telegram pada data Tafsir Al-Jalalain. Dalam melakukan penelitian ini, penulis menggunakan *Langchain*. *Langchain* dapat membantu dalam menerapkan *LLM* pada aplikasi *Natural Language Processing (NLP)* seperti *Question Answering*. *Langchain* digunakan sebagai kerangka kerja untuk membantu Sistem *Question Answering* yang menerapkan *Large Language Model (LLM)* dari *OpenAI* sebagai penyedia *API* dan model Bahasa. Tujuan dari penelitian ini diharapkan dapat membantu dalam mencari informasi tentang Tafsir Al-Jalalain secara akurat, efektif dan cepat.

2. METODOLOGI PENELITIAN

2.1 Tahap Penelitian

Tahap penelitian adalah serangkaian proses yang dilakukan dalam melakukan penelitian. Pada penelitian ini tahap pertama yang dilakukan yaitu mengumpulkan studi literatur pada penelitian terdahulu [15]. Peneliti juga mengumpulkan data Tafsir Al-Jalalain yang digunakan sebagai basis data untuk menghasilkan informasi dari sistem *Question Answering* [16]. Kemudian dilanjutkan dengan Analisa kebutuhan data dan Analisa model Bahasa, kemudian tahap selanjutnya

perancangan Sistem *Question Answering*, kemudian implementasi dari hasil analisa dan perancangan, lalu pada tahap terakhir melakukan evaluasi apakah sistem *Question Answering* sudah bekerja dengan baik dan membantu pengguna dalam mencari informasi yang diinginkan.



Gambar 1. Alur Tahapan Penelitian

2.1.1 Studi Lileratur

Studi literatur dilakukan peneliti sebagai tahap pertama dalam pengumpulan data dengan melakukan pencarian, pembelajaran serta pemahaman dari berbagai sumber lileratur seperti karya ilmiah, jurnal laporan penelitian, dan website yang terkait dengan judul peneliti gunakan.

2.1.2 Pengumpulan Data Tafsir Al-Jalalain

Peneliti melakukan pengumpulan data Tafsir Al-Jalalain dan Al-Qur'an dari sumber *github open source* yaitu <https://github.com/renomureza/quran-API-id>. Data Tafsir Al-Jalalain di masukkan kedalam *database MYSQL*.

2.1.3 Analisa Kebutuhan Data

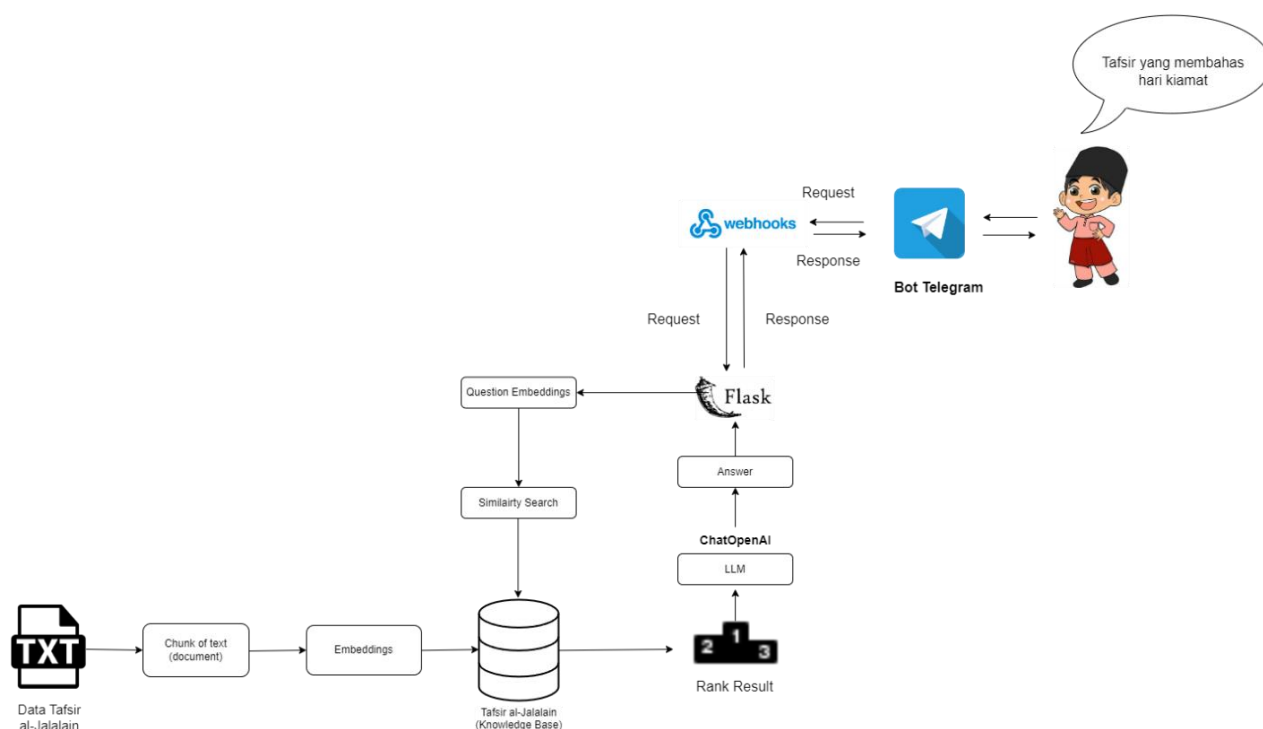
Analisa kebutuhan data digunakan untuk menganalisis kebutuhan data yang akan digunakan untuk penelitian. Peneliti mengumpulkan dan melakukan pemilihan data Tafsir Al-Jalalain yang terdiri dari nomor ayat, nama surah, urutan surah, teks arab, dan teks Tafsir Al-jalalain yang kemudian di *ekspor* dalam format file txt lengkap pada Al-Qur'an Juz 30.

2.1.4 Analisa Model Bahasa

Analisa model Bahasa adalah proses menganalisa model Bahasa yang akan digunakan untuk mengimplentasikan *Question Answering* yang dapat berfungsi dengan baik. Peneliti menggunakan *Large Language Model (LLM)* untuk melatih model pada Kumpulan data Tafsir Al-jalalain yang besar dan akan menghasilkan teks baru secara otomatis. *LLM* yang digunakan yaitu model *GPT-3.5 Turbo* dari *OPENAI*. Peneliti juga menggunakan sebuah framework yaitu *Langchain* untuk membantu meningkatkan penyesuaian keakuratan dan relevansi dari model. *Langchain* membantu dalam megolah data Tafsir Al-Jalalain pada proses *embeddings* data agar dipahami oleh model [17], serta menyimpan data Tafsir Al-Jalalain dalam sebuah *vectore store* yang menjadi *knowledge base* dari penelitian [18]. *Langchain* membantu dalam memproses pertanyaan dari pengguna dengan beberapa proses didalamnya sehingga dapat memberikan sebuah jawaban dari pertanyaan tersebut.

2.1.5 Perancangan Sistem Question Answering

Perancangan Sistem *Question Answering* adalah proses menggambarkan seluruh tahap perancangan Sistem *Question Answering* yang dimulai dari memproses data hingga *output Chatbot Telegram Question Answering*.



Gambar 2. Arsitektur Sistem Question Answering

Gambar 2 merupakan arsitektur yang menggambarkan perancangan Sistem Question Answering secara menyeluruh yang dimulai dari persediaan data Tafsir Al-Jalalain hingga respon dari pengguna.

a. Data Tafsir Al-Jalalain lengkap 30 Juz dengan format file txt

Tabel 1. Data Tafsir Al-Jalalain

id	Gabungan (Nama Surah, Urutan Surah, Urutan Ayat, Teks Al-Qu'ran)	Tafsir Al-Jalalain
[1]	Al-Fatihah Surah Ke-1 : Ayat ke 1 Bunyi ayat : بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ	Tafsir Al-Jalalain : (Dengan nama Allah Yang Maha Pemurah lagi Maha Penyayang)
[2]	Al-Fatihah Surah Ke-1 : Ayat ke 2 Bunyi ayat : الْحَمْدُ لِلَّهِ رَبِّ الْعَالَمِينَ	Tafsir Al-Jalalain : (Segala puji bagi Allah) Lafal ayat ini merupakan kalimat berita,.....
[..]		
[114]	An-Nas Surah Ke 114: Ayat ke 6 Bunyi ayat: مِنَ الْجِنَّةِ وَالنَّاسِ	Tafsir Al-Jalalain : (Dari jin dan manusia") lafal ayat ini menjelaskan pengertian setan yang menggoda itu,.....

Tabel 1 menunjukkan sampel data dari Tafsir Al-Jalalain yang sudah melakukan proses analisis data sehingga menghasilkan struktur data yang diantaranya terdapat id, gabungan (nama surah, urutan surah, urutan ayat, teks Al-Qur'an).

b. *Chunk of Text*

Chunk of Text adalah proses memotong dokumen menjadi bagian bagian yang kecil, sehingga dapat digunakan oleh model Bahasa. Peneliti melakukan proses *Chunk of Text* menggunakan *re.split* yang bekerja dengan cara memotong data berdasarkan pola yang sudah dibuat [19]. Peneliti membuat pola *chunk of text* dengan memotong berdasarkan ayat pada Al-Qur'an.

```
[Document(page_content='Al-Ikhlas Surah Ke-112 : Ayat ke 1 Bunyi ayat : اللَّهُ لَا إِلَهَ إِلَّا هُوَ الْحَمْدُ لِلَّهِ رَبِّ الْعَالَمِينَ', metadata={'source': 'Tafsir Al-Jalalain Surah Al-Ikhlas Ayat Ke 1'}), Document(page_content='Al-Ikhlas Surah Ke-112 : Ayat ke 2 Bunyi ayat : الْحَمْدُ لِلَّهِ رَبِّ الْعَالَمِينَ', metadata={'source': 'Tafsir Al-Jalalain Surah Al-Ikhlas Ayat Ke 2'})]
```

Gambar 3. Contoh *Chunk of Text*

Gambar 3 menunjukkan contoh dari proses *Chunk of Text* yang menghasilkan potongan data berdasarkan ayat-ayat pada Al-Qur'an.

c. *Embeddings*

Embeddings adalah proses mengubah teks yang sudah melalui proses *Chunk of Text* menjadi bentuk angka agar dapat dipahami oleh model Bahasa. *Embeddings* dilakukan dengan menggunakan *Api Key* dari *Open AI* yang kemudian di simpan pada *chromadb* sebagai knowledge base.

d. Knowledge Base

Knowledge Base merupakan tempat penyimpanan dokumen dan informasi pada sebuah sistem *database* dengan tujuan untuk memudahkan seorang *developer* [18]. peneliti menggunakan *Chromadb* yang merupakan *database* vektor *open source*, sehingga dapat dipakai secara gratis dan berlisensi dibawah *Apache 2.0* [20]. *Chromadb* digunakan peneliti untuk menyimpan Hasil data dari proses *Chunk of Text* dan *embeddings*.

e. Similarity Search

Similarity Search berfungsi untuk mengukur Tingkat kemipiran teks pertanyaan dengan data yang ada pada *Knowledge Base*. Proses *similarity search* yaitu dengan mengukur Tingkat kemiripan dari pertanyaan pengguna dengan *kwoledge base* Tafsir Al-Jalalain dengan.

f. Rank Result

Proses *rank result* yaitu mengambil tingkat kemiripan teks dari proses *similarity search* dengan default sebanyak empat sumber yang paling mirip seperti yang ditunjukkan pada gambar 4.

```
"Referensi": "Al-Baqarah Surah Ke-2 : Ayat ke 176 Bunyi ayat : وَالَّذِينَ آمَنُوا وَعَمِلُوا الصَّالِحَاتِ فِي الْآخِرَةِ أُولَئِكَ يُرْجَوْنَ الْغَوْثِ الْأَعْلَىٰ وَاللَّهُ يَرْضَاهُ غَرْثًا قَبِيضًا Tafsir Al-Jalalain : (Demikian itu), yakni apa-apa yang telah disebutkan seperti menelan api dan seterusnya (disebabkan oleh karena) (Allah telah menurunkan Alkitab dengan sebenarnya) berkaitan dengan menurunkan, maka mereka berselisih padanya, mereka beriman pada sebagian dan kafir pada sebagian dengan jalan menyembunyikannya. (Dan sesungguhnya orang-orang yang berselisih tentang Alkitab) yakni orang-orang Yahudi dan ada pula yang mengatakan bahwa mereka itu adalah orang-orang musyrik, yaitu tentang Alquran, sebagian mengatakannya sebagai syair, yang lain sihir dan sebagiannya lagi sebagai tenung (berada dalam penyimpangan yang jauh) dari kebenaran.\n",
"sumber": "AL-Baqarah Ayat Ke 176"
},
{
  "Referensi": "Al-Insyiqaq Surah Ke-84 : Ayat ke 7 Bunyi ayat : فَأَمَّا مَنْ أُوْتِيَ الْكِتَابَ فَكَانَ مُعَذِّبًا فَأَبَىٰ تَوْبَهُ وَكَانَ مِن دُونِ الْحَقِّ وَمِنَ الْأَحْزَابِ فَأُولَٰئِكَ أَصْحَابُ الْأَنْفَارِ الَّتِي فِيهَا يَكْتُمُونَ الَّذِينَ كَفَرُوا يَسْتَرْفِعُونَ أَصْوَابَهُمْ وَيَقُولُ لَوْنُنَا لَوْنُ الْغَايَةِ وَمَن يُضْلِلِ اللَّهُ فَمَا لَهُ هَادٍ فَذَٰلِكَ يُزَكِّي اللَّهُ طَائِفًا مِّنْهُمْ لِيُظْهِرَهُمْ لِنَبِيِّهِ وَأُولَٰئِكَ فِي سَعِيدٍ طَافٍ Tafsir Al-Jalalain : (Adapun orang yang diberikan kitabnya) yakni kitab catatan amalnya (dari sebelah kanannya) dia adalah orang yang beriman.\n",
"sumber": "Tafsir Al-Jalalain Surah Al-Insyiqaq Ayat Ke 7"
},
{
  "Referensi": "Al-Baqarah Surah Ke-2 : Ayat ke 146 Bunyi ayat : وَإِذْ قَرَأْتَ الْقُرْآنَ فَاسْمِعْ يَٰعِزُّونَ أَصْوَابَهُمْ وَاسْمِعُوا الْوَعْدَ الَّذِي لَكُمْ وَكُونُوا لَنَا مَدِينًا مِّمَّنْ آمَنَ اللَّهُ إِنَّهُمْ يَخِفُّونَنَا وَلِآلِهِمُ الْوَحْيُ الْكَافِرُ وَالْكَافِرُ أَجْمَلٌ أَجْمَلِينَ Tafsir Al-Jalalain : (Orang-orang yang Kami beri Alkitab mengenalnya) Muhammad (sebagaimana mereka mengenal anak-anak mereka sendiri) karena disebutkan ciri-cirinya dalam kitab-kitab suci mereka. Kata Ibnu Salam, "Sesungguhnya ketika aku melihatnya, maka aku pun segera mengenalnya, sebagaimana aku mengenal putraku sendiri, bahkan lebih kuat lagi mengenal Muhammad." (Dan sesungguhnya sebagian di antara mereka menyembunyikan kebenaran) maksudnya ciri-cirinya itu (padahal mereka mengetahui) keadaanmu dan siapa kamu yang sebenarnya.\n",
"sumber": "AL-Baqarah Ayat Ke 146"
},
{
  "Referensi": "Abasa Surah Ke-80 : Ayat ke 13 Bunyi ayat : وَكَذَٰلِكَ نَجْزِي الْمُحْسِنِينَ Tafsir Al-Jalalain : (Di dalam kitab-kitab) menjadi Khabar yang kedua, karena sesungguhnya ia dan yang sebelumnya berkedudukan sebagai jumlah Mu'taridhah atau kalimat sisipan (yang dimulikan) di sisi Allah.\n",
"sumber": "Tafsir Al-Jalalain Surah Abasa Ayat Ke 13"
},
{
  "Referensi": "Gafir Surah Ke-40 : Ayat ke 2 Bunyi ayat : وَلَقَدْ أَنزَلْنَا إِلَيْكَ الْكِتَابَ فَاتَّبِعْهُ وَذَكَرْ لَّكَ الْوَعْدَ الَّذِي لَكَ وَالْكَرِيمُ Tafsir Al-Jalalain : (Diturunkan Kitab ini) yakni Alquran, menjadi Mubtada (dari Allah) Khabar Mubtada (Yang Maha Perkasa) di dalam kerajaan-Nya (lagi Maha Mengetahui) tentang makhluk-Nya.\n",
"sumber": "Gafir Surah Ke-40 Ayat Ke 2"
}
```

Gambar 4 Contoh Rank Result

g. LLM (Large Language Model)

Selanjutnya rank result akan masuk ke proses *LLM* menggunakan model *gpt-3.5-turbo* yang akan mengolah *rank result* dan menghasilkan sebuah jawaban.

h. Flask

Flask digunakan pada sisi *Backend* Sistem *Question Answering* sebagai penyedia *API* yang akan digunakan oleh *Chatbot Telegram*.

i. Webhook

Webhook adalah salah satu konsep *API*, *Webhook* juga disebut *callback* yang digunakan oleh aplikasi untuk menyediakan aplikasi lain dengan informasi secara langsung [21]. Pada penelitian ini *Webhook* digunakan sebagai jembatan untuk menghubungkan *Chatbot* dengan *server*

j. Chatbot Telegram

Chatbot telegram digunakan sebagai tampilan yang akan berinteraksi langsung dengan pengguna. Peneliti membuat *Chatbot telegram* dari *BotFather* yang sudah disediakan oleh pihak Telegram.

2.1.6 Implementasi Basis Data

Implementasi Basis Data menggunakan Bahasa pemrograman *python*. Keuntungan menggunakan *python* karena menyediakan banyak *library* yang bisa dipakai untuk mengimplementasikan sistem *Question Answering*, salah satunya yaitu *Langchain*. *Langchain* merupakan kerangka kerja untuk *Large Language Model (LLM)* dapat merespon konteks yang diberikan dengan baik. Data Tafsir Al-Jalalain akan melalui proses pemotongan data (*Chunk of Text*) dan proses *Embeddings* yang kemudian disimpan pada *Knowledge Base*, selanjutnya peneliti menggunakan *Flask* untuk membuat *service API* yang digunakan sebagai *Backend* dari Sistem *Question Answering*.

2.1.7 Implementasi Tampilan

Implementasi tampilan menggunakan fitur Telegram yaitu *Chatbot* sebagai tampilan antarmuka sistem *Question Answering* yang dapat digunakan langsung sebagai penghubung antara sistem dan pengguna. Keuntungan menggunakan *Chatbot* telegram sendiri yaitu mudah dalam membuatnya dan dapat diakses melalui aplikasi telegram.

2.1.8 Evaluasi Question Answering

Eveluasi adalah tahap akhir untuk menguji serta mengevaluasi sistem kepada pengguna dan sebagai bukti kepada pengguna bahwa sistem dapat berfungsi sesuai dengan kebutuhan dan pemahaman pengguna. Evaluasi dilakukan dengan cara memberikan kuisioner kepada responden mahasiswa Jurusan Ilmu Al-Qur'an dan Tafsir UIN Suska Riau sebanyak 7 mahasiswa dengan memberikan 10 pertanyaan yang akan ditanyakan kepada Chatbot Question Answering Tafsir Al-Jalalain, kemudian para responden tersebut menilai apakah jawaban yang dihasilkan sudah akurat atau belum.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pengumpulan dan Analisa Data

Peneliti melakukan pengumpulan data Tafsir Al-Jalalain dari sumber github <https://github.com/renomureza/quran-API-id> lengkap 30 Juz Al-Quran. Data tersebut dimasukkan ke dalam database MYSQL, kemudian di Analisa untuk memilih atribut dari data Tafsir Al-jalalain. Hasil analisis data memiliki atribut yang diantaranya yaitu nomor ayat, nama surah, urutan surah, teks arab, dan teks Tafsir Al-jalalain yang kemudian di *ekspor* kedalam format file txt secara lengkap 30 juz, hasil data berikut yang akan digunakan oleh peneliti.

3.2 Hasil Perancangan

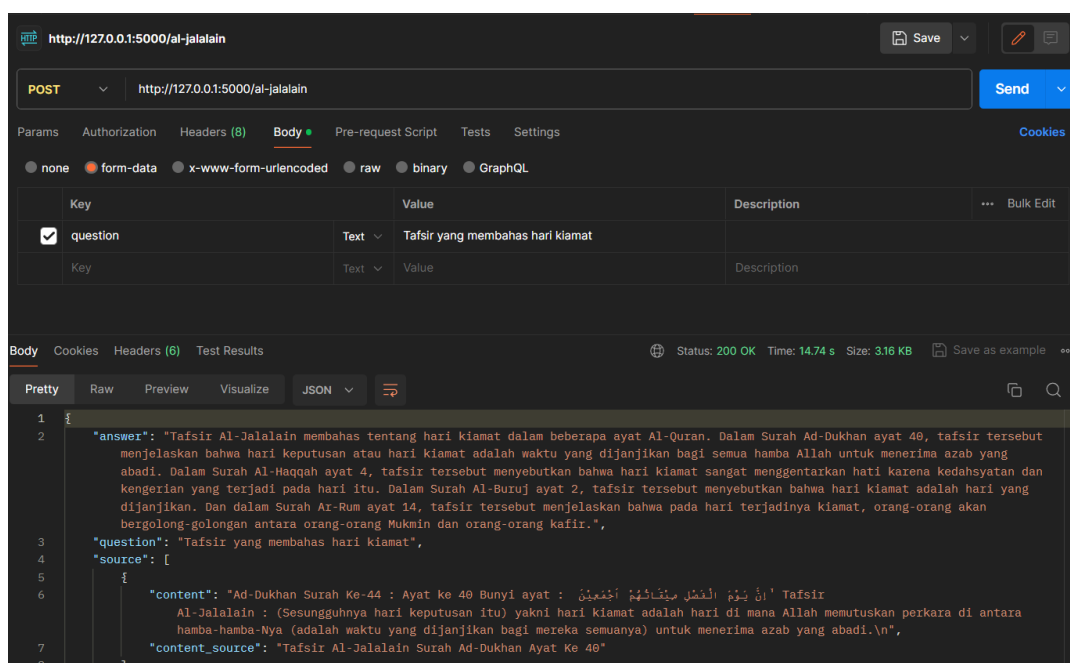
Pada proses ini peneliti menggunakan model Bahasa dari *OpenAI* yaitu “*gpt-3.5-turbo*” dan juga *framework Langchain* yang digunakan untuk proses mengembangkan sistem *Question Answering* yang dimulai dari pemotongan data, kemudian proses *embeddings* yang nantinya bisa dipahami oleh model Bahasa dan akan melakukan proses penyimpanan pada *chroma db* sebagai *knowledge base* serta berakhir pada proses mengubungkan sistem *Question Answering* pada *Chatbot* Telegram yang ditunjukkan pada Gambar 2. Hasil perancangan sistem *Question Answering* selanjutnya akan masuk ke proses implementasi.

3.3 Hasil Implementasi

Implementasi adalah tahap yang dilakukan peneliti setelah melakukan serangkaian proses yang dimulai dari studi literatur dan mengumpulkan data Tafsir Al-Jalalain, menganalisa data Tafsir Al-Jalalain dan menganalisa model yang digunakan dalam penelitian, kemudian merancang arsitektur sistem *Question Answering* lalu bisa di implementasi oleh peneliti. Implementasi dilakukan dengan membagi dua tahap yaitu impelementasi pada basis data dan juga tampilan.

3.3.1 Hasil Implementasi Basis Data

Berikut hasil implementasi Basis data berupa *service API*:



Gambar 5 Contoh Output Service API

Gambar 5 menunjukkan contoh hasil *service API* dari sistem *Question Answering*.

3.1.2 Hasil Implementasi Tampilan

Berikut merupakan hasil implementasi dari Sistem *Question Answering* berbasis *Chatbot Telegram* yang menggunakan *service API* dari *Flask*.

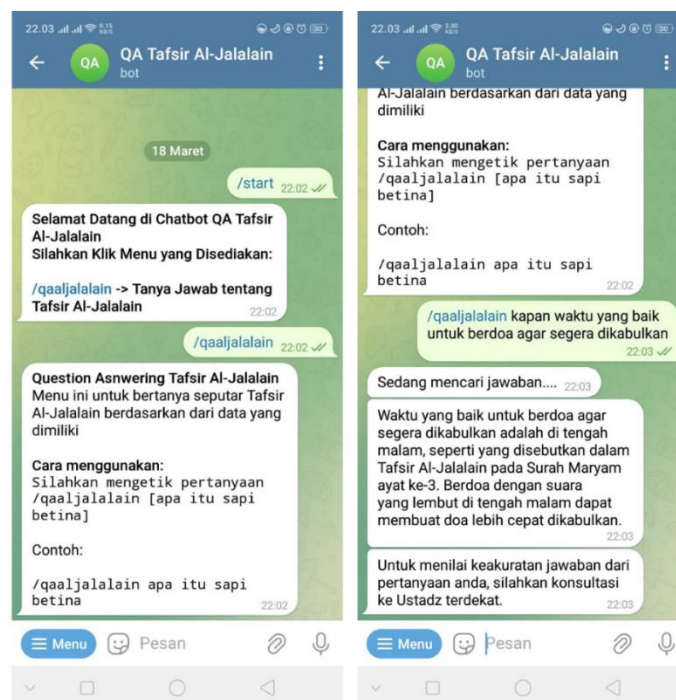
a. Tampilan Awal *Chatbot Telegram*



Gambar 6. Tampilan Awal *Chatbot Telegram*

Tampilan pada gambar 6 merupakan tampilan awal untuk memulai percakapan dengan *Chatbot Telegram*, pengguna dapat memulai dengan memasukkan perintah `/start` untuk melihat panduan penggunaan Sistem *Question Answering*.

b. Tampilan Menu *Question Answering*



Gambar 7. Tampilan Menu *Question Answering*

Pada gambar 7 menunjukkan tampilan antarmuka untuk pengguna dapat bertanya kepada *Chatbot Telegram* dengan perintah /qaaljalalain “pertanyaan”. Sehingga nanti menu *Question Answering* ini akan mengolah pertanyaan dari pengguna dan akan mengembalikan jawaban yang diharapkan sesuai dengan apa yang ditanyakan oleh pengguna.

3.4 Hasil Evaluasi Question Answering

Evaluasi dilakukan dengan menggunakan sampel 10 pertanyaan yang akan diuji oleh Mahasiswa Jurusan Ilmu Al-Qur'an dan Tafsir UIN SUSKA Riau. 10 sampel pertanyaan akan dibandingkan antara jawaban dari Sistem *Question Answering* dengan pemahaman jawaban dari Mahasiswa/I Jurusan Ilmu Al-Qur'an dan Tafsir UIN SUSKA Riau. Pada evaluasi terdapat lima jawaban yang dapat dipilih oleh responden, diantaranya sangat setuju (SS), setuju (ST), ragu-ragu (RG), Tidak Setuju (TS), sangat tidak setuju (STS) dengan skala penilaian 5-1. Perhitungan Hasil evaluasi menggunakan rumus *User Acceptance Test (UAT)* untuk melihat apakah sistem dapat berfungsi sesuai dengan kebutuhan dan pemahaman pengguna [22].

Tabel 2. Hasil Evaluasi *Question Answering*

No	Pertanyaan	SS	ST	RG	TS	STS
1	Mengapa wajib seorang muslim menunaikan shalat 5 waktu	3	4	0	0	0
2	Ketika seseorang telah meninggal, apa yang terjadi saat berada di alam kubur	0	5	2	0	0
3	Apakah kitab taurat merupakan petunjuk bagi bani israel	2	5	0	0	0
4	Apakah boleh menikahi anak Perempuan dari saudara laki-laki ayah	0	4	3	0	0
5	Hukum membatalkan puasa karena sakit	4	3	0	0	0
6	Apa maksud shadaqah dalam Q.S At-Taubah	2	5	0	0	0
7	Kapan waktu yang baik untuk berdoa agar segera dikabulkan	4	3	0	0	0
8	Kenapa seorang muslim Perempuan wajib untuk memakai jilbab dan pakaian yang tertutup	3	4	0	0	0
9	Pada waktu shalat jum'at, apakah boleh melakukan jual beli	4	3	0	0	0
10	Siapa saja yang boleh menerima zakat	2	2	2	1	0
Total		24	38	7	1	0

Hasil perhitungan pada total jawaban responden sebagai berikut

Sangat Setuju (5) : 24 * 5 : 120
 Setuju (4) : 38 * 4 : 152
 Ragu – ragu (3) : 7 * 3 : 21
 Tidak setuju (2) : 1 * 2 : 2
 Sangat Tidak Setuju (1) : 0 * 1 : 0
Total Skor : 295

Kemudian dilakukan perhitungan nilai X (skor tertinggi) sebagai berikut:

X = skor tertinggi * (jumlah pertanyaan * jumlah responden)
 X = 5 * (10*7)
 = 350

Setelah itu dilakukan perhitungan hasil persentase menggunakan persamaan rumus *User Acceptance Test (UAT)* sebagai berikut:

$$\begin{aligned}
 \text{Persentase} &= \frac{\text{Total Skor}}{X} * 100\% \\
 &= \frac{295}{350} * 100\% \\
 &= 84.29\%
 \end{aligned}$$

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, dimulai dari tahap studi literatur, pengumpulan data Tafsir Al-Jalalain, analisis dan perancangan, implementasi, serta evaluasi. Dapat disimpulkan bahwa penulis telah berhasil mengimplementasikan *Question Answering* berbasis *Chatbot Telegram* menggunakan *Langchain* dan *LLM* dengan serangkaian proses yang telah dilakukan. Dari hasil evaluasi Sistem *Question Answering* sudah berfungsi dengan baik dan mendapatkan persentase sebesar 84.29% dari evaluasi responden yang merupakan mahasiswa jurusan Ilmu Al-Qur'an dan Tafsir UIN SUSKA Riau dengan menguji sistem *Question Answering* menggunakan 10 sampel pertanyaan. Namun masih terdapat kekurangan pada Sistem *Question Answering* yaitu Sistem belum mampu secara akurat menghasilkan jawaban yang sesuai dari pertanyaan pengguna dan juga beberapa pertanyaan pengguna tidak bisa dijawab oleh Sistem *Question Answering*, karena Sistem *Question Answering* dilatih dari model Bahasa menggunakan model *GPT-3.5-Turbo*. Kekurangan pada penelitian ini dapat di idenfitikasi kembali untuk pengembangan penelitian selanjutnya. Terdapat saran untuk penelitian selanjutnya yang bisa dikembangkan lebih lanjut yaitu mengidentifikasi agar pertanyaan dari pengguna

bisa di olah oleh sistem *Question Answering* dengan model Bahasa serta bisa memperbaiki pertanyaan dari pengguna, sehingga dapat memungkinkan untuk selalu menghasilkan jawaban yang akurat.

REFERENCES

- [1] D. Apriliani, S. F. Handayani, T. N. Anugrahaeni, A. Miftahudin, L. Nurarifiah, and I. T. Saputra, "Aplikasi Question Answer Sebagai Media Pembelajaran Interaktif Untuk Mata Pelajaran Akuntansi," *JMM (Jurnal Masy. Mandiri)*, vol. 7, no. 2, p. 2003, 2023, doi: 10.31764/jmm.v7i2.13867.
- [2] A. Y. Chandra, D. Kurniawan, and R. Musa, "Perancangan Chatbot Menggunakan Dialogflow Natural Language Processing (Studi Kasus: Sistem Pemesanan pada Coffee Shop)," *J. Media Inform. Budidarma*, vol. 4, no. 1, p. 208, 2020, doi: 10.30865/mib.v4i1.1505.
- [3] S. Husada, "Rancang Bangun Dialog Manager Menggunakan Semantic Interpretation Pada Smiqa (Swamedikasi Interactive Question Answering) System," *J. Teknol. Pint.*, vol. 2, no. 8, pp. 1–10, 2022, [Online]. Available: <http://teknologipintar.org/index.php/teknologipintar/article/view/219>
- [4] H. Nur, "METODE PENGAJARAN PONDOK PESANTREN : STUDI KASUS KH. YAHYA MASDUQI DALAM KAJIAN KITAB TAFSIR AL-JALALAIN YANG DITERAPKAN DI PONDOK PESANTREN MIFTAHUL MUTA'ALLIMIN BABAKAN CIWARINGIN CIREBON," vol. 4, pp. 27–39, 2023.
- [5] R. A. Yunmar and I. W. W. Wisesa, "Pengembangan Mobile-Based Question Answering System Mobile-Based Question Answering System Development With Ontology Based Knowledge," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 4, pp. 693–700, 2020, doi: 10.25126/jtiik.202072255.
- [6] R. Qardafil, A. Sujarwo, and M. It, "Optimasi Kualitas Chatbot Jala Tech Melalui Implementasi Auto-GPT," vol. 12, no. 2, 2023.
- [7] M. A. P. Subali and P. Wijaya, "Sistem Question Answering untuk Bahasa Bali menggunakan Metode Rule-Based dan String Similarity," *Techno.Com*, vol. 20, no. 2, pp. 300–308, 2021, doi: 10.33633/tc.v20i2.4390.
- [8] A. Setiawan, O. N. Pratiwi, and ..., "Question Answering System Dalam Bentuk Chatbot Pada Platform Line Untuk Mata Pelajaran Sejarah Sma/ma Dengan Menggunakan Algoritma Levenshtein ...," *eProceedings ...*, vol. 8, no. 5, pp. 9794–9802, 2021, [Online]. Available: <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/15756%0Ahttps://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/15756/15469>
- [9] G. F. Avisyah, I. J. Putra, and S. S. Hidayat, "Open Artificial Intelligence Analysis using ChatGPT Integrated with Telegram Bot," *J. ELTIKOM*, vol. 7, no. 1, pp. 60–66, 2023, doi: 10.31961/eltikom.v7i1.724.
- [10] T. A. Zuraiyah *et al.*, "IMPLEMENTASI CHATBOT PADA PENDAFTARAN MAHASISWA BARU MENGGUNAKAN RECURRENT NEURAL NETWORK," *J. Ilm. Teknol. dan Rekayasa*, vol. 24, no. 2, pp. 91–101, 2019.
- [11] O. Topsakal and T. C. Akinci, "Creating Large Language Model Applications Utilizing LangChain: A Primer on Developing LLM Apps Fast," *Int. Conf. Appl. Eng. Nat. Sci.*, vol. 1, no. 1, pp. 1050–1056, 2023, doi: 10.59287/icaens.1127.
- [12] Langchain, "Introduction Langchain," *python.langchain.com*, 2023. <https://python.langchain.com/docs/community> (accessed Nov. 07, 2023).
- [13] H. L. Walingkas and P. O. N. Saian, "Penerapan Framework Flask pada Pembangunan Sistem Informasi Pemasok Barang," *J. JTik (Jurnal Teknol. Inf. dan Komunikasi)*, vol. 7, no. 2, pp. 227–234, 2023, doi: 10.35870/jtik.v7i2.729.
- [14] D. F. Ningtyas and N. Setiyawati, "Implementasi Flask Framework pada Pembangunan Aplikasi Purchasing Approval Request," vol. 1, no. 1, pp. 19–34, 2021, doi: 10.25008/janitra.v1i1.120.
- [15] A. Rachman, I. Mardhiyah, and M. Jannah, "Implementasi Chatbot FAQ pada Aplikasi Monev Kinerja Direktorat Jenderal Anggaran Menggunakan Framework Rasa Open Source," *Media Online*, vol. 4, no. 1, pp. 62–72, 2023, doi: 10.30865/klik.v4i1.1020.
- [16] E. Febriansyah, E. Nirmala, F. I. Komputer, T. Informatika, U. Pamulang, and K. T. Selatan, "PERANCANGAN SISTEM INFORMASI JUAL BELI PROPERTI MENGGUNAKAN CHAT BOT TELEGRAM YANG TERINTEGRASI DENGAN WEB MENGGUNAKAN METODE PROTOTYPE," vol. 1, no. 2, pp. 279–284, 2023.
- [17] F. Castillo, "Embeddings: Meaning, Examples and How To Compute," *arize.com*, 2023. <https://arize.com/blog-course/embeddings-meaning-examples-and-how-to-compute/> (accessed Nov. 08, 2023).
- [18] A. C. Dela Cruz, "What is a Knowledge Base: Definition, Types, Features, and Examples," *klutch.app*, 2022. <https://klutch.app/blog/knowledge-base-definition-types-features-examples/> (accessed Nov. 07, 2023).
- [19] Vishal, "Python Regex Split String Using re.split()," *pynative.com*, 2021. https://pynative.com/python-regex-split/#llc_comments (accessed Feb. 08, 2024).
- [20] Langchain, "Chroma," *python.langchain.com*, 2023. <https://python.langchain.com/docs/integrations/vectorstores/chroma> (accessed Feb. 08, 2024).
- [21] T. Istiana, G. S. B. Dharmawan, B. Prakoso, P. J. Komunikasi, and C. Author, "Pengembangan Sistem Diseminasi Prakiraan Cuaca Menggunakan Aplikasi Bot Telegram dengan Metode Webhook," vol. 12, no. April, 2020.
- [22] F. F. Nursaid, A. Hendra Brata, and A. P. Kharisma, "Pengembangan Sistem Informasi Pengelolaan Persediaan Barang Dengan ReactJS Dan React Native Menggunakan Prototype (Studi Kasus : Toko Uda Fajri)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 4, no. 1, pp. 46–55, 2020.