

BAB I PENDAHULUAN

1.1 Latar Belakang

Pengembangan sistem penjawab pertanyaan (QA) berbasis dokumen telah mendapatkan perhatian signifikan dalam beberapa tahun terakhir. Sistem ini bertujuan untuk memberikan jawaban akurat dan relevan berdasarkan informasi dalam dokumen yang tersedia [1]. Tidak seperti sistem pengambilan informasi tradisional, sistem QA berbasis dokumen memberikan jawaban langsung dengan memahami konteks dan semantik yang mendasari dokumen dan pertanyaan [2].

Proses QA berbasis dokumen melibatkan beberapa langkah: pengambilan dokumen, pengambilan potongan teks, dan ekstraksi jawaban [3]. Penggunaan embedding vektor dan teknik pencarian semantik adalah komponen kunci dalam sistem ini [4]. Embedding vektor adalah representasi numerik yang menangkap makna semantik dan hubungan antara kata, frasa, dan dokumen, yang biasanya dihasilkan menggunakan model berbasis jaringan saraf, seperti BERT [5].

Pencarian semantik mengukur kesamaan antara representasi vektor dari pertanyaan dan dokumen atau potongan teks, memungkinkan pengambilan informasi yang lebih akurat [6]. Model bahasa besar (LLM) seperti GPT-3 dan GPT-4 telah menunjukkan kemampuan luar biasa dalam memahami konteks dan memberikan respons yang akurat [7], [8].

API OpenAI memberikan akses ke model bahasa mutakhir ini, memungkinkan pengembang untuk mengintegrasikan kemampuannya ke dalam sistem QA berbasis dokumen [9], [10]. Dengan memanfaatkan embedding vektor dan teknik pencarian semantik, sistem QA dapat menghasilkan jawaban yang lebih akurat dan kontekstual [11].

LangChain adalah kerangka kerja open-source yang menyederhanakan pengembangan aplikasi yang melibatkan model bahasa besar dan komponen NLP lainnya [12]. Ini menyediakan arsitektur modular yang memungkinkan pengembang untuk mengaitkan berbagai komponen sesuai kebutuhan spesifik [13].

Sejalan dengan ini, Sreeram dan Sai (2023) mendemonstrasikan efektivitas LangChain dalam mengembangkan sistem query yang dapat mengekstrak informasi dari dokumen PDF dengan akurat, menunjukkan fleksibilitas framework ini dalam menangani berbagai jenis data [14].

LangChain dipilih sebagai framework utama untuk pengembangan chatbot ini karena memiliki beberapa keunggulan. Framework ini menyediakan arsitektur modular yang memudahkan integrasi berbagai komponen, mendukung berbagai model bahasa dari OpenAI, dan menawarkan abstraksi tingkat tinggi yang menyederhanakan pengembangan aplikasi berbasis LLM. Selain itu, LangChain unggul dalam mengelola alur kerja kompleks dan mempertahankan konteks dalam interaksi multi-turn, fitur penting untuk chatbot tanya jawab. Topsakal dan Akinci (2023) memperkuat argumen ini dalam penelitian mereka, yang memberikan panduan komprehensif tentang bagaimana LangChain dapat digunakan untuk mengembangkan aplikasi LLM dengan cepat dan efisien, memanfaatkan komponen-komponen seperti prompts, models, chains, dan agents [15].

Dalam bab-bab berikutnya, kita akan membahas detail implementasi dari sistem QA berbasis dokumen kami, memanfaatkan kekuatan LangChain, embedding vektor, dan API OpenAI untuk memberikan jawaban yang akurat dan kontekstual terhadap pertanyaan pengguna.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan, maka dapat dirumuskan suatu masalah yang akan diselesaikan pada penelitian ini yaitu:

1. Saat koleksi dokumen menjadi lebih besar, mencari dokumen atau segmen teks yang paling relevan untuk pertanyaan tertentu menjadi semakin sulit. Metode tradisional berbasis kata kunci seringkali kesulitan menangkap relevansi semantik. Dibutuhkan teknik yang ditingkatkan untuk pengambilan dokumen secara semantik.
2. Meskipun model bahasa besar unggul dalam pertanyaan fakta sederhana, mereka sering kesulitan dengan pertanyaan yang lebih rumit yang

memerlukan penalaran bertahap, menyelesaikan ambiguitas, atau menggabungkan pengetahuan eksternal. Kerangka kerja yang lebih baik diperlukan untuk memungkinkan kemampuan-kemampuan canggih semacam itu.

3. Banyak pertanyaan membutuhkan penggabungan informasi dari beberapa dokumen atau segmen teks. Menggabungkan dan merespon konteks yang tersebar ini secara efektif merupakan tantangan besar yang masih dihadapi oleh model-model saat ini.

1.3 Batasan Masalah

Pada Skripsi ini memiliki beberapa batasan antara lain:

1. Model bahasa besar bisa menunjukkan bias dalam keluarannya berdasarkan data yang mereka latih. Mengurangi bias ini dan memastikan keluaran yang adil dan etis adalah tantangan yang terus berlangsung yang memerlukan pemantauan dan strategi koreksi yang hati-hati.
2. Model OpenAI yang digunakan merupakan model free tier.
3. Penelitian ini akan menggunakan data dokumen dalam format PDF, TXT, atau DOCX. Data non-teks, seperti gambar atau audio, tidak akan dimasukkan dalam penelitian ini.
4. Saat koleksi dokumen menjadi sangat besar, indeksasi, pengambilan, dan pemrosesan yang efisien dapat menimbulkan tantangan skala komputasi yang memerlukan arsitektur terdistribusi yang canggih.
5. Sistem yang diusulkan sangat bergantung pada layanan eksternal seperti OpenAI API dan basis data vektor. Setiap perubahan, gangguan, atau pembaruan harga pada layanan-layanan ini dapat mempengaruhi kinerja, biaya, atau ketersediaan sistem.

1.4 Tujuan Penelitian

Dengan latar belakang, rumusan masalah, dan batasan yang diuraikan di atas, penelitian ini bertujuan untuk mencapai tujuan-tujuan utama berikut:

1. Tujuan utama adalah membangun sistem pertanyaan-jawaban yang dapat

memberikan jawaban yang tepat dan relevan untuk pertanyaan pengguna berdasarkan isi dokumen yang diunggah.

2. Sistem ini akan memanfaatkan dan mengintegrasikan kemajuan terbaru dalam pemrosesan bahasa alami, termasuk model bahasa yang kuat dari OpenAI, teknik penyematan semantik, dan kerangka kerja sumber terbuka seperti LangChain. Tetap berada di garis depan akan memastikan bahwa sistem mendapat manfaat dari kemajuan cepat yang terjadi di bidang-bidang ini.
3. Sistem akan mendukung penerimaan dan pemrosesan dokumen dalam berbagai format seperti file PDF, DOCX, dan TXT. Ini akan memungkinkan pengguna untuk dengan mudah menggabungkan sumber informasi yang beragam ke dalam sistem pertanyaan-jawaban.
4. Melalui desain modular LangChain, arsitektur sistem akan sangat termodularisasi dan dapat diperluas. Hal ini akan memfasilitasi integrasi yang lancar dari komponen-komponen baru, eksperimen yang mudah dengan berbagai teknik, dan kemampuan untuk meningkatkan sistem secara iteratif dari waktu ke waktu.

1.5 Manfaat Penelitian

Mengembangkan sistem pertanyaan-jawaban dokumen yang canggih dengan memanfaatkan model bahasa terbaru, teknik penyematan, dan kerangka kerja sumber terbuka dapat memberikan berbagai manfaat potensial:

1. Dengan memungkinkan pengguna untuk dengan mudah menanyakan dan mengambil informasi relevan dari repositori dokumen besar, sistem bertindak sebagai pengganda pengetahuan. Ini meruntuhkan hambatan untuk mengakses wawasan yang terkubur di berbagai sumber, memberdayakan pengambilan keputusan yang lebih baik.
2. Mencari informasi spesifik secara manual melalui dokumen-dokumen yang banyak memakan waktu dan biaya. Sistem QA yang akurat secara dramatis mempercepat proses ini, menghemat usaha dan sumber daya organisasi

secara signifikan.

3. Alih-alih menggantikan ahli manusia, sistem QA dapat bertindak sebagai alat augmentation - dengan cepat menyediakan informasi relevan kepada manusia untuk membantu mereka membuat keputusan yang lebih terinformasi dan meningkatkan produktivitas mereka.
4. Dengan menggabungkan teknologi terkini dan mengeksplorasi metode-metode baru seperti agregasi konteks dan interpretabilitas, karya ini dapat memajukan bidang pemrosesan bahasa alami dan sistem QA melalui wawasan dan teknik baru.

1.6 Sistematika Penulisan

Sistematika dalam penulisan skripsi adalah sebagai berikut:

BAB I: Pendahuluan

Bab ini berisi tentang penjelasan latar belakang, batasan masalah, tujuan penelitian, manfaat penelitian, metode penelitian, dan sistematika penulisan.

BAB II: Tinjauan Pustaka

Tinjauan pustaka menjelajahi studi terkait dan teknologi yang ada terkait pemrosesan dokumen, pertanyaan dan jawaban, pengembangan *chatbot*, dan integrasi Langchain, dan OpenAI API.

BAB III: Metode Penelitian

Bab metode penelitian memaparkan desain penelitian, arsitektur sistem, dan pendekatan langkah demi langkah yang digunakan untuk mengembangkan *chatbot* pertanyaan dan jawaban berbasis dokumen.

BAB IV: Arsitektur, Pengujian dan Evaluasi Chatbot

Bab ini membahas arsitektur sistem yang dikembangkan, termasuk modul pemrosesan dokumen, teknik NLP, integrasi Langchain dan OpenAI API Key, pengujian dan evaluasi juga disajikan.

BAB V: Penutup

Bab ini menyajikan kesimpulan dan saran dari penelitian tentang chatbot penjawab pertanyaan menggunakan LangChain dan OpenAI API. Sistem ini mempermudah akses informasi dari dokumen besar, menghemat waktu pencarian manual, dan meningkatkan produktivitas. Disarankan untuk terus mengembangkan teknologi dan metode baru, serta meneliti pemrosesan pertanyaan multi-langkah untuk jawaban yang lebih akurat.

