

BAB I

PENDAHULUAN

1.1. Latar Belakang

Perkembangan terkini dalam bidang *Artificial Intelligence* (AI) telah melahirkan sebuah sub-bidang, yaitu AI Generatif (GenAI). Teknologi ini menandai pergeseran dari model AI tradisional, yang primernya berfokus pada tugas-tugas analitis seperti prediksi dan klasifikasi berdasarkan data yang ada [1]. Sebaliknya, AI Generatif, khususnya *Large Language Models* (LLM), memiliki kemampuan untuk secara aktif terlibat dalam “*knowledge generation*”. LLM adalah model komputasi canggih yang dilatih pada data teks yang sangat besar, memungkinkannya untuk memahami dan menghasilkan konten baru seperti teks, gambar, kode, dan lainnya yang sangat mirip dengan hasil karya manusia, sebagai respons terhadap input dalam bahasa alami [1].

Kemampuan unik LLM ini telah membuka berbagai peluang aplikasi inovatif dalam sektor pendidikan. Sebuah tinjauan sistematis komprehensif yang dilakukan baru-baru ini telah memetakan beberapa domain utama di mana LLM menunjukkan potensi yang signifikan [2]. Domain-domain tersebut meliputi:

- a. Pembuatan Konten Edukasi Medis: LLM telah terbukti mampu menghasilkan deskripsi yang akurat mengenai pengetahuan anatomi dan memberikan saran medis awal menggunakan pendekatan berbasis kuesioner, yang dapat digunakan sebagai alat bantu belajar bagi mahasiswa kedokteran.
- b. Asisten Pembelajaran Bahasa Asing: Dalam pembelajaran bahasa, LLM dapat berperan sebagai mitra percakapan, penyedia konten yang relevan, asisten penulisan, dan pemberi umpan balik yang instan terhadap tata bahasa dan kosa kata.
- c. Bantuan Penelitian Akademis: Bagi para peneliti dan mahasiswa tingkat lanjut, LLM dapat berfungsi sebagai alat untuk mempercepat proses kompilasi literatur, meringkas artikel ilmiah yang kompleks, dan bahkan membantu dalam analisis data awal.
- d. Evaluasi Ujian dan Penilaian: LLM juga digunakan untuk mengevaluasi kinerja model itu sendiri dalam berbagai jenis ujian standar, seperti ujian

masuk sekolah kedokteran atau hukum, yang menunjukkan kemampuannya dalam penalaran tingkat tinggi.

Akan tetapi, penggunaan LLM yang berdiri sendiri memiliki kelemahan yang fundamental, yakni cenderung memiliki *hallucination* untuk menghasilkan informasi yang tidak akurat [2], [3]. Kekurangan ini tentu saja sangat krusial jika diimplementasikan ke dalam bidang edukasi, yang akan mengakibatkan informasi yang salah dan tidak akurat. Hal ini yang menyebabkan LLM konvensional tidak dapat diandalkan untuk tugas yang sangat memerlukan data faktual. Maka dari itu, RAG merupakan solusi dari masalah tersebut, dimana RAG akan memanfaatkan dokumen yang terverifikasi untuk dijadikan sebagai *ground truth*, sehingga menghasilkan data yang lebih akurat dan faktual berdasarkan dokumen yang diberikan dan mengurangi risiko *hallucination*.

Secara khusus, penelitian ini relevan dengan proyek pengembangan *flashcard* otomatis karena literatur secara eksplisit memvalidasi kemampuan LLM dalam menghasilkan materi belajar yang spesifik. Studi-studi telah mendokumentasikan bahwa model canggih seperti GPT-4 mampu membuat "kuis pilihan ganda dan soal menjodohkan dengan berbagai tingkat kesulitan" berdasarkan materi yang diberikan [2]. Selain itu, alat bantu belajar komersial yang ada di pasaran, seperti Scholarcy, menawarkan fitur pembuatan *flashcard* sebagai salah satu fungsi utamanya untuk membantu siswa dalam proses menghafal pengetahuan [4]. Validasi ini menegaskan bahwa penggunaan LLM untuk menghasilkan *flashcard* bukan hanya sebuah konsep teoritis, tetapi merupakan aplikasi yang praktis dan relevan dalam ekosistem teknologi pendidikan saat ini.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, *Generative AI* terutama LLM memiliki potensi besar dalam konten edukatif, akan tetapi masih memiliki kelemahan yakni *hallucination* atau menghasilkan informasi yang tidak akurat karena keterbatasan data. Untuk mengatasi hal tersebut maka arsitektur RAG diajukan untuk meningkatkan akurasi faktual dengan dokumen yang telah diverifikasi sebagai basis informasinya. Oleh karena itu, untuk menyelesaikan masalah tersebut, dibuat penelitian ini dengan rumusan masalah sebagai berikut:

1. Bagaimana implementasi arsitektur RAG (*Retrieval-Augmented Generation*) untuk menghasilkan *flashcard* yang akurat secara faktual dan meminimalkan risiko *hallucination*?
2. Bagaimana kinerja atau performa dari arsitektur RAG ini (kecepatan dan akurasi)?

1.3. Tujuan Penelitian

Tujuan-tujuan dari penelitian ini dirancang untuk memandu proses pengembangan penelitian, maka tujuan penelitian ini adalah sebagai berikut:

1. Merancang, mengimplementasikan, dan mengevaluasi sebuah sistem berbasis arsitektur *Retrieval-Augmented Generation* (RAG) untuk menghasilkan *flashcard* yang akurat secara faktual dan terbukti mampu meminimalkan risiko *hallucination*.
2. Menganalisis dan mengukur kinerja arsitektur *Retrieval-Augmented Generation* (RAG) berdasarkan metrik kecepatan pemrosesan dan tingkat akurasi.

1.4. Batasan Masalah

Untuk menjaga fokus penelitian agar tidak terlalu meluas dari pembahasan yang dimaksud, penelitian ini membatasi ruang lingkupnya. Oleh karena itu, batasan masalah dalam penelitian ini meliputi:

1. Penelitian ini menggunakan model dan *framework open-source*, dimana alat-alat yang digunakan adalah Google Gemini, LangSmith, dan Streamlit. Sehingga memiliki keterbatasan penggunaan dan kemampuan dibandingkan dengan alat-alat yang berbayar.
2. Sumber dataset yang digunakan adalah dokumen berbasis PDF dengan subjek pembahasan terbatas pada topik biologi pada tingkat sekolah (SD-SMA).

1.5. Kerangka Pemikiran

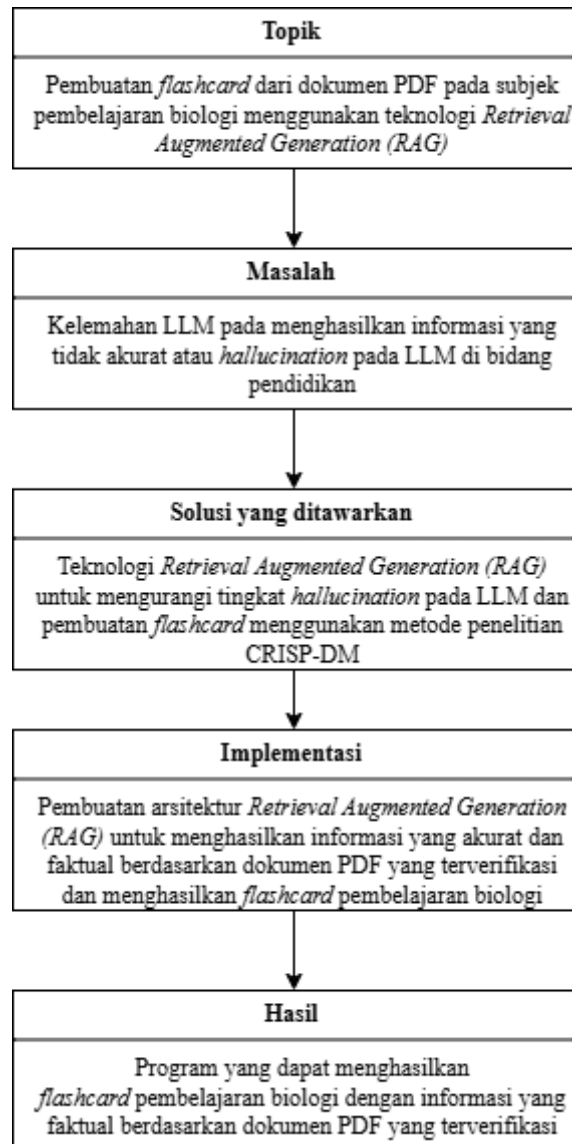
Kerangka pemikiran penelitian ini dimulai dengan identifikasi masalah utama pada *Large Language Models* (LLM) di bidang pendidikan, yaitu kelemahannya dalam menghasilkan informasi yang tidak akurat atau *hallucination* saat membuat materi pembelajaran [3].

Selanjutnya, penelitian ini menemukan bahwa penggunaan teknologi Retrieval-Augmented Generation (RAG) dapat menawarkan solusi yang efektif untuk mengatasi masalah tersebut. Pendekatan ini memanfaatkan kemampuan RAG untuk mengambil informasi yang faktual dan relevan dari dokumen PDF yang telah terverifikasi, serta kemampuan LLM di dalamnya untuk menghasilkan konten *flashcard* yang akurat berdasarkan data yang diperoleh.

Proses ini dimulai dengan tahap *retrieval* yang menganalisis dokumen PDF untuk mengekstrak konteks yang relevan, diikuti dengan penerapan tahap *generation* yang menghasilkan teks *flashcard* berdasarkan informasi yang terverifikasi tersebut. Arsitektur RAG akan diimplementasikan untuk membangun sebuah program yang efektif dalam membuat *flashcard* pembelajaran.

Setelah pengembangan, sistem akan diimplementasikan dan dievaluasi untuk mengukur kemampuannya dalam menghasilkan *flashcard* dengan informasi yang faktual dan sesuai dengan dokumen sumber. Evaluasi ini akan memberikan wawasan tentang efektivitas solusi RAG yang diusulkan dalam mengurangi tingkat *hallucination* dan menciptakan alat bantu belajar yang andal. Berikut adalah bagan kerangka pemikiran untuk penelitian:





Gambar 1.1. Kerangka Pemikiran Penelitian

Berdasarkan Gambar 1.1, penelitian ini bertujuan untuk mengembangkan sebuah program yang dapat secara otomatis membuat *flashcard* pembelajaran biologi dari dokumen PDF. Untuk mengatasi masalah *hallucination* atau informasi tidak akurat yang sering dihasilkan oleh *Large Language Models (LLM)*, penelitian ini menerapkan teknologi *Retrieval-Augmented Generation (RAG)*. Arsitektur RAG memastikan bahwa setiap *flashcard* yang dibuat didasarkan pada konteks yang diambil langsung dari dokumen PDF sumber, sehingga menghasilkan konten pembelajaran yang faktual dan terverifikasi. Hasil akhirnya adalah sebuah aplikasi yang mampu mengubah materi studi menjadi alat bantu belajar yang akurat.

1.6. Sistematika Penulisan

Struktur penulisan laporan penelitian ini dirancang untuk menyajikan kerangka kerja yang sistematis. Adapun sistematika pembahasannya adalah sebagai berikut:

BAB I: PENDAHULUAN

Bab ini menyajikan pengantar penelitian yang meliputi latar belakang, rumusan masalah, tujuan, batasan masalah, kerangka pemikiran, metodologi secara umum, serta sistematika penulisan laporan.

BAB II: TINJAUAN PUSTAKA

Bab ini mengulas berbagai konsep fundamental dan landasan teori yang relevan dengan topik permasalahan. Tinjauan ini menjadi dasar untuk analisis dan pemecahan masalah.

BAB III: METODOLOGI PENELITIAN

Bab ini menjelaskan secara rinci alur kerja penelitian yang diterapkan, mulai dari tahap perencanaan hingga tahap akhir. Di dalamnya mencakup analisis kebutuhan produk serta strategi implementasi aplikasi.

BAB IV: HASIL DAN PEMBAHASAN

Bab ini menyajikan dua bagian utama. Pertama, pemaparan hasil temuan yang diperoleh dari setiap tahapan penelitian sesuai dengan urutan rumusan masalah. Kedua, analisis mendalam terhadap temuan tersebut untuk menjawab pertanyaan penelitian.

BAB V: PENUTUP

Bab ini berisi rangkuman kesimpulan dari keseluruhan penelitian. Selain itu, dipaparkan pula kritik dan saran konstruktif yang diperoleh selama proses pengerjaan tugas akhir sebagai bahan pertimbangan untuk pengembangan penelitian di masa mendatang.