

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pemahaman terhadap tafsir Al-Qur'an merupakan kebutuhan esensial bagi Muslim awam, namun akses terhadapnya masih menghadapi berbagai hambatan fundamental. Kajian internasional menunjukkan bahwa 82,5% non-penutur Arab "dapat membaca Al-Qur'an namun mengalami kesulitan memahami maknanya" karena faktor bahasa, kultur, dan keterbatasan perangkat interpretatif klasik [1]. Tantangan ini semakin besar ketika dihadapkan pada tafsir klasik seperti *Tafsir Ibn Kathir* yang memiliki volume sangat besar, misalnya edisi Inggris *abridged* yang terdiri atas sepuluh jilid dengan sekitar 650 halaman per jilid [2]. Kompleksitas bahasa Arab klasik, panjang penjelasan, dan minimnya dukungan teknologi modern menyebabkan tingkat aksesibilitas tafsir masih rendah bagi masyarakat umum.

Tren global juga menunjukkan adanya penurunan kemampuan pemahaman bacaan, yang berdampak langsung pada kapasitas individu dalam memproses teks keagamaan yang panjang dan padat. *OECD* melaporkan bahwa kemampuan literasi orang dewasa di banyak negara mengalami stagnasi atau penurunan, dan hampir separuh remaja tidak mencapai tingkat membaca dengan pemahaman yang memadai [3]. Kondisi serupa ditemukan di Indonesia, di mana kemampuan pemahaman bacaan siswa berada pada kategori rendah serta minat baca masih minim [4]. Situasi ini memperkuat urgensi solusi digital yang mampu menjembatani kesenjangan antara kebutuhan memahami tafsir dan kemampuan membaca masyarakat modern.

Dalam beberapa tahun terakhir, teknologi *Retrieval-Augmented Generation (RAG)* dan *semantic retrieval* menunjukkan potensi signifikan untuk meningkatkan kualitas sistem tanya-jawab (Q&A) pada domain ilmu keislaman. Penelitian pada QIAS 2025 menunjukkan bahwa skema *Hybrid RAG* yang mengombinasikan *BM25*, *embedding retrieval*, dan *cross-encoder reranking*, mampu meningkatkan akurasi hingga 25% pada tugas Q&A pengetahuan Islam [5]. Riset lain menyimpulkan bahwa model *Llama-3.2 3B* memiliki kinerja tinggi pada dataset kajian Al-Qur'an [6], sementara benchmark MMTEB mengidentifikasi *multilingual-e5-large-instruct* sebagai model embedding paling unggul untuk *retrieval* multibahasa [7]. Di sisi lain, berbagai studi juga menunjukkan bahwa pendekatan *hybrid retrieval* lebih efektif dibandingkan metode tunggal karena mampu mengatasi keterbatasan *sparse-only* maupun *dense-only retrieval*, serta meningkatkan relevansi dan *faithfulness* jawaban [8], [9].

Melihat adanya kebutuhan pemahaman tafsir yang tinggi, rendahnya aksesibilitas teks tafsir klasik, dan perkembangan teknologi RAG yang semakin matang, implementasi sistem penjawab tafsir berbasis *Hybrid Sparse-Dense Retrieval* menjadi semakin relevan. Integrasi *BM25* dan model embedding *multilingual* memberikan peluang untuk mengatasi hambatan bahasa serta meningkatkan kemampuan pencarian semantik pada dataset tafsir yang panjang dan kompleks. Dengan memanfaatkan *Tafsir Ibn Kathir* versi Inggris sebagai studi kasus, penelitian ini diarahkan untuk menghasilkan sistem tanya-jawab tafsir yang lebih akurat, inklusif, dan adaptif terhadap kebutuhan Muslim awam dalam konteks literasi modern.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang sudah dijelaskan, maka rumusan masalah dalam penelitian ini dapat dirumuskan sebagai berikut:

1. Bagaimana perbandingan kinerja metode retrieval berbasis kata kunci (*Sparse: BM25*), berbasis makna/kontekstual (*Dense: E5 Large Multilingual*), serta kombinasi keduanya (*Hybrid*) dalam menemukan dokumen *Tafsir Ibn Kathir* yang paling relevan?
2. Bagaimana performa sistem RAG dalam menghasilkan jawaban berdasarkan konteks dokumen *Tafsir Ibn Kathir*?
3. Bagaimana performa sistem dari sisi waktu respons (latensi) ketika diimplementasikan sebagai prototipe aplikasi RAG berbasis *Streamlit*?

1.3 Batasan Masalah

Untuk memastikan penelitian tetap fokus, terukur, dan tidak melebar ke aspek yang tidak memengaruhi tujuan utama, beberapa batasan berikut ditetapkan sebagai ruang lingkup sistem yang dikembangkan:

1. Penelitian ini hanya menggunakan *Tafsir Ibn Kathir* versi *abridged*, tanpa memasukkan tafsir lain atau edisi lengkapnya.
2. Sistem dibatasi untuk menjawab penjelasan tafsir, tidak menampilkan ayat Al-Qur'an atau melakukan pencarian ayat.
3. Metode *retrieval* yang digunakan hanya *hybrid BM25 + E5 Large Multilingual*, tanpa penambahan teknik lain seperti *re-ranking* lanjutan atau model *retrieval* alternatif.
4. *Filtering metadata* dibatasi pada surah dan ayat, tanpa atribut tambahan seperti topik, riwayat, atau *asbāh al-nuzūl*.
5. *Input kueri* bersifat *multilingual*, tetapi penelitian tidak mengevaluasi perbedaan performa antar bahasa maupun melakukan analisis linguistik tingkat lanjut.
6. Model generatif yang digunakan adalah *LLaMA 3.2 3B* tanpa *fine-tuning*, sehingga seluruh keluaran bergantung pada kemampuan bawaan model dan instruksi *prompt*.

7. Evaluasi kualitas retrieval dilakukan secara kuantitatif menggunakan metrik matematis. Penelitian ini tidak menggunakan penilaian subjektif manusia, validasi teologis, maupun penilaian pakar tafsir sebagai parameter kebenaran data.
8. Pengembangan antarmuka Streamlit hanya bertujuan untuk memverifikasi alur sistem dan mengukur latensi respons (kecepatan). Oleh karena itu, kualitas atau relevansi jawaban yang muncul pada antarmuka prototipe tidak termasuk dalam cakupan pengujian kualitas sistem, mengingat adanya potensi keterbatasan pada kode prototipe yang tidak memengaruhi performa algoritma retrieval utama.
9. Evaluasi kualitas sistem dibatasi pada metrik retrieval umum dan modul evaluasi yang tidak membutuhkan *ground truth* pihak ketiga. Validasi teologis atau penilaian pakar tafsir tidak dilakukan.
10. Penelitian tidak menilai keakuratan teologi, kesesuaian fikih, atau interpretasi *hermeneutis*; fokusnya pada performa teknis sistem RAG.
11. Pengembangan antarmuka pengguna (*UI/UX*) dan aplikasi *mobile* atau *web* bukan objek utama penelitian. Sistem prototipe Streamlit hanya digunakan untuk mengukur latensi respons, bukan desain tampilan.
12. Penelitian tidak membahas nuansa *sanad*, perbedaan riwayat, atau analisis konteks sejarah; sistem hanya mengambil dan mengolah teks tafsir sebagaimana tercantum dalam korpus.
13. Ruang lingkup sistem tidak mencakup *reasoning* lintas sumber atau penggabungan tafsir lain; seluruh jawaban bersandar pada korpus Ibn Kathir yang digunakan.
14. Optimasi tingkat produksi seperti *scaling*, *distributed retrieval*, atau *caching* tidak menjadi fokus; seluruh eksperimen berjalan pada lingkungan prototipe sesuai sumber daya yang tersedia.

1.4 Tujuan Penelitian

Tujuan penelitian ini adalah untuk menghasilkan sistem penjawab Tafsir Al-Qur'an berbasis *hybrid sparse-dense retrieval* dan RAG yang mampu memberikan jawaban tafsir secara relevan, komprehensif, dan efisien. Secara lebih spesifik, penelitian ini bertujuan untuk:

1. Menganalisis kinerja metode *retrieval* berbasis pencocokan teks (*Sparse: BM25*), pemahaman konteks (*Dense: E5 Large Multilingual*), serta pendekatan campuran (*Hybrid*) untuk menghasilkan dokumen *Tafsir Ibnu Katsir* yang paling relevan.
2. Menilai performa sistem RAG dalam menghasilkan jawaban berdasarkan konteks dokumen tafsir, termasuk konsistensi jawaban dan minimisasi informasi di luar konteks.
3. Mengukur performa sistem dari sisi waktu respons (latensi) ketika diimplementasikan sebagai prototipe aplikasi RAG berbasis *Streamlit*.
4. Menyediakan prototipe sistem RAG sebagai bukti konsep yang dapat digunakan untuk evaluasi teknis lebih lanjut.
5. Memvalidasi parameter-parameter kinerja sistem, seperti akurasi *retrieval*, relevansi jawaban, dan latensi, menggunakan metrik *retrieval* umum dan pendekatan *semantic similarity, token-based (BERTScore)*, serta *context relevancy* tanpa memerlukan *ground truth* pihak ketiga.

Tujuan-tujuan ini diharapkan dapat memberikan gambaran yang terukur mengenai efektivitas pendekatan *hybrid sparse-dense retrieval* dan RAG dalam konteks tafsir Al-Qur'an, sekaligus menjadi dasar pengambilan kesimpulan yang objektif.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan kontribusi yang signifikan, baik dari sudut pandang teknis pengembangan sistem maupun manfaat non-teknis bagi masyarakat luas dan akademisi.

1.5.1 Manfaat Teknis

Secara teknis, penelitian ini memberikan kontribusi nyata dalam pengembangan teknologi informasi, antara lain:

1. **Pengembangan Sistem Penjawab Cerdas:** Menghasilkan sistem penjawab *Tafsir Ibn Kathir* berbasis *hybrid sparse-dense retrieval* dan *Retrieval-Augmented Generation (RAG)*. Sistem ini memungkinkan pengguna untuk mendapatkan penjelasan tafsir secara cepat, relevan, dan komprehensif melalui pendekatan teknologi terkini.
2. **Prototipe dan Referensi Pengembangan:** Menjadi sebuah prototipe sistem RAG yang dapat dijadikan acuan dasar bagi peneliti maupun pengembang perangkat lunak untuk melakukan pengembangan lebih lanjut, seperti optimasi kecepatan sistem (latensi), peningkatan akurasi pencarian, hingga integrasi ke berbagai platform aplikasi lainnya.
3. **Penyediaan Data Empiris Penelitian AI:** Memungkinkan dilakukannya evaluasi teknis terhadap efektivitas kombinasi metode pencarian (*sparse*, *dense*, dan *hybrid*) serta performa *Large Language Model (LLM)* pada domain teks keagamaan. Hal ini memberikan data empiris yang berharga bagi kemajuan bidang *Information Retrieval (IR)* dan Kecerdasan Buatan.

1.5.2 Manfaat Non-Teknis

Selain aspek teknis, penelitian ini juga membawa dampak positif pada aspek edukasi dan sosial, yaitu:

1. **Pengayaan Literatur Akademik:** Memberikan kontribusi pada literatur ilmiah dan metodologi penelitian, khususnya dalam penerapan AI pada studi tafsir Al-Qur'an. Penelitian ini memperkaya khazanah ilmu pengetahuan mengenai bagaimana teknologi dapat mendukung studi agama.
2. **Referensi bagi Institusi Pendidikan:** Menjadi sumber rujukan bagi pengajar, peneliti, atau pengembang sistem edukasi keagamaan dalam memahami cara mengimplementasikan teknologi AI untuk mempermudah akses terhadap materi pembelajaran tafsir yang kredibel.
3. **Peningkatan Literasi Digital Keagamaan:** Mendukung transformasi digital dalam beragama dengan menyediakan medium interaktif. Hal ini diharapkan mampu mendorong masyarakat untuk memahami Al-Qur'an dengan cara yang lebih modern, efisien, dan sesuai dengan perkembangan zaman.

1.6 Sistematika Penulisan

Sistematika penulisan skripsi ini disusun agar pembaca memperoleh gambaran menyeluruh mengenai alur penelitian dan isi tiap bab, sebagai berikut:

BAB I PENDAHULUAN

Berisi latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan. Bab ini memberikan landasan dan arah penelitian sehingga pembaca memahami konteks, ruang lingkup, dan tujuan yang ingin dicapai.

BAB II TINJAUAN PUSTAKA

Memuat literatur dan teori yang relevan dengan penelitian. Fokus utama adalah teori *Information Retrieval* (IR), metode *sparse* dan *dense retrieval*, *hybrid sparse-dense retrieval*, serta *Retrieval-Augmented Generation* (RAG). Bab ini juga membahas studi terdahulu di bidang yang sejenis sebagai dasar analisis dan perbandingan metode yang digunakan.

BAB III METODE PENELITIAN

Membahas secara detail rancangan penelitian, termasuk deskripsi objek penelitian, arsitektur sistem, alur *pipeline retrieval*, metodologi pengembangan prototipe sistem, serta alat dan bahan yang digunakan dalam penelitian. Bab ini menjelaskan langkah-langkah yang dilakukan untuk membangun sistem RAG, sehingga penelitian dapat direplikasi atau dikembangkan lebih lanjut.

BAB IV HASIL DAN PEMBAHASAN

Berisi hasil eksperimen dan analisis terkait performa sistem. Pembahasan mencakup evaluasi kinerja metode retrieval *sparse*, *dense*, dan *hybrid*, performa jawaban sistem RAG berdasarkan konteks dokumen *Tafsir Ibn Kathir*, serta pengukuran latensi prototipe aplikasi berbasis *Streamlit*.

BAB V PENUTUP

Berisi kesimpulan dari seluruh penelitian, saran pengembangan sistem ke depannya, serta implikasi penelitian baik secara teknis maupun non-teknis. Bab ini menyimpulkan capaian penelitian terkait efektivitas *hybrid sparse-dense retrieval* dan RAG dalam penjawab tafsir Al-Qur'an, sekaligus memberikan rekomendasi untuk penelitian atau pengembangan lanjutan.

