

BAB 5. PENUTUP

5.1. Kesimpulan

Berdasarkan penelitian implementasi sistem tanya jawab Al-Qur'an berbasis *mobile* menggunakan pendekatan *Retrieval-Augmented Generation* (RAG), pencarian semantik (*Semantic Search*), dan evaluasi kualitas jawaban pada dataset resmi Kemenag RI, dapat disimpulkan bahwa sistem telah berhasil dirancang, diimplementasikan, dan dievaluasi sesuai tujuan penelitian. Secara umum, penelitian ini menjawab dua rumusan masalah utama, yaitu perancangan dan implementasi arsitektur sistem tanya jawab berbasis RAG, serta pengukuran tingkat keandalan dan akurasi semantik sistem melalui evaluasi otomatis.

a. Rumusan Masalah 1 Perancangan dan Implementasi Arsitektur

Sistem berhasil dirancang dan diimplementasikan sebagai aplikasi *mobile* Flutter yang terhubung ke *backend* FastAPI berbasis RAG, dengan komponen inti berupa *semantic retrieval Guardrail* anti-halusinasi dan dataset Al-Qur'an Kemenag. Arsitektur sistem lebih menonjolkan akurasi, *Grounding*, dan validitas teologis daripada optimasi latensi atau skalabilitas.

b. Rumusan Masalah 2 (Keandalan dan Akurasi Semantik)

Keandalan sistem terbukti melalui pengujian teknis (*Unit*), pengujian *black-box*, serta evaluasi kualitas jawaban pada Bab 4, yang menunjukkan jawaban relevan, ter-*Grounding* pada sumber resmi Kemenag RI, dan layak digunakan sebagai media bantu tanya jawab Al-Qur'an.

Dengan demikian sistem terbukti layak dan dapat digunakan sebagai media tanya jawab Al-Qur'an yang andal dan minim halusinasi.

5.2. Saran

Berdasarkan hasil pengembangan dan evaluasi yang telah dilakukan, terdapat beberapa saran untuk pengembangan penelitian selanjutnya agar sistem dapat menjadi lebih matang dan bermanfaat.

a. Perluasan Basis Pengetahuan

Penelitian selanjutnya disarankan menambah cakupan sumber atau memperkaya variasi corpus dataset yang digunakan seperti b-b tafsir lain, terutama untuk pertanyaan tematik yang lebih luas. Penambahan dataset yang lebih beragam akan membantu sistem menangani konteks yang lebih kompleks tanpa kehilangan *Grounding* pada sumber resmi.

b. Peningkatan Ketepatan *Retrieval*

Mekanisme *Semantic Search* dan pengaturan ambang *SIMILARITY_THRESHOLD* masih dapat diuji lebih lanjut untuk memperoleh keseimbangan yang lebih baik antara penolakan adaptif dan cakupan jawaban. Penelitian berikutnya dapat melakukan uji sensitivitas terhadap ambang skor dan strategi *retrieval* agar hasil jawaban semakin stabil pada berbagai jenis *intent*.

c. Penyempurnaan Evaluasi Kualitas Jawaban

Evaluasi pada penelitian ini sudah mencakup *Black-box Testing*, RAGAS, PSSUQ, dan validasi pakar. Namun, penelitian berikutnya dapat menambahkan jumlah dataset evaluasi atau memperluas jumlah responden agar hasil pengukuran lebih representatif. Pengujian tambahan juga dapat dilakukan terhadap skenario pertanyaan yang lebih beragam dan lebih menantang.

d. Peningkatan Pengalaman Pengguna

Frontend dapat dikembangkan lebih lanjut agar interaksi pengguna menjadi lebih nyaman, misalnya dengan penyajian referensi yang lebih interaktif, pengelolaan riwayat yang lebih rinci, atau dukungan pencarian ulang berdasarkan percakapan sebelumnya. Penyempurnaan ini akan membantu pengguna memahami jawaban sistem secara lebih cepat dan lebih jelas.

e. Penguatan Monitoring dan Regression Test

Karena sistem melibatkan banyak komponen seperti *Intent Routing*, SSE, *Guardrail*, dan sinkronisasi Firestore, penelitian berikutnya disarankan menambahkan pengujian regresi yang lebih otomatis dan konsisten pada setiap perubahan fitur. Langkah ini penting agar stabilitas sistem tetap terjaga ketika dilakukan pengembangan lanjutan.

f. Pengembangan Dukungan Validasi Keislaman yang Lebih Luas

Validasi pakar pada penelitian ini sudah menunjukkan hasil yang baik, namun ke depan sistem dapat diuji oleh lebih banyak validator atau dengan rubrik penilaian yang lebih rinci. Hal ini akan memperkuat aspek teologis dan meningkatkan kepercayaan pengguna terhadap hasil jawaban sistem.

g. Penguatan Batasan Aplikasi dan Skalabilitas

Penelitian lanjutan perlu memfokuskan roadmap skalabilitas karena sistem saat ini masih berada pada level prototipe akademik. Penguatan yang disarankan meliputi autoscaling *service*, cache untuk *query* berulang, serta *pipeline* CI/CD yang konsisten agar sistem siap dipakai pada beban pengguna yang lebih besar.

h. Penguatan *Backup Plan* LLM

Strategi *fallback* saat ini masih berpusat pada *Provider cloud* dan *Safe refusal response*. Penelitian berikutnya dapat mengkaji *fallback* hybrid, misalnya *Provider* sekunder yang benar-benar independen, *caching* jawaban untuk *query* berulang, atau model lokal untuk mode darurat. Namun, implementasi lokal perlu dihitung cermat karena menuntut spesifikasi perangkat, optimasi inferensi, dan evaluasi ulang groundedness pada *domain* tafsir.

Dengan demikian, pengembangan lanjutan diharapkan tidak hanya meningkatkan kualitas teknis sistem, tetapi juga memperkuat manfaatnya sebagai media bantu pemahaman Al-Qur'an yang kredibel, aman, dan sesuai dengan kebutuhan pengguna.