

ABSTRAK

Kelompok Bimbingan Ibadah Haji (KBIH) berperan penting dalam memberikan layanan informasi kepada jemaah, namun layanan konsultasi manual yang terbatas pada jam operasional tertentu belum mampu menjawab kebutuhan informasi jemaah yang bersifat kontekstual, khususnya pada aspek *fiqh* haji. Pemanfaatan *chatbot* berbasis *Artificial Intelligence* (AI) berpotensi menjadi solusi, namun model generatif rentan mengalami halusinasi, menghasilkan informasi yang terdengar meyakinkan namun tidak didukung sumber valid yang berisiko tinggi pada domain keagamaan. Penelitian ini bertujuan mengembangkan dan mengevaluasi *chatbot* layanan haji berbasis IndoT5-Base yang dioptimalkan dengan teknik *Low-Rank Adaptation* (LoRA) dan diintegrasikan dengan mekanisme *Retrieval-Augmented Generation* (RAG) bertingkat sebagai strategi mitigasi halusinasi AI. Metodologi penelitian mengadaptasi kerangka kerja *Cross-Industry Standard Process for Data Mining* (CRISP-DM), meliputi tahap pengumpulan data dari FAQ KBIH Miftahul Ulum dan Buku Tuntunan Manasik Haji dan Umrah Kemenag RI 2024, augmentasi data, *fine-tuning* model dengan LoRA, implementasi mekanisme *retrieval* bertingkat dengan empat mode respons (*retrieval* langsung, *generate* berkonteks, *abstain*, dan *retrieval* terkunci), serta evaluasi menggunakan metrik ROUGE-L, BLEU-4, BERTScore, *response time*, dan validasi manual oleh pengurus KBIH. Hasil pengujian terhadap 6.306 data menunjukkan model *Hybrid* v1 memperoleh ROUGE-L 0,517 dan BERTScore-F1 0,833, sementara pendekatan *Hybrid* v2 (*retrieval* bertingkat) menghasilkan BLEU-4 tertinggi (0,401) dan waktu respons tercepat (0,15 detik). Validasi manual terhadap 20 sampel jawaban menghasilkan skor ketepatan faktual rata-rata 4,54 dan kelengkapan 4,14 dari skala 5, dengan mekanisme *retrieval* bertingkat terbukti mengeliminasi risiko fabrikasi pada 48,7% pertanyaan yang dijawab langsung dari basis pengetahuan terverifikasi. Meskipun target halusinasi belum sepenuhnya tercapai yaitu dengan potensi menyesatkan sebesar 15%, hasil penelitian menunjukkan bahwa pendekatan *retrieval* bertingkat efektif membatasi risiko halusinasi pada kategori pertanyaan berisiko tinggi.

Kata Kunci: *Chatbot*, Haji, IndoT5, LoRA, *Retrieval-Augmented Generation*, Halusinasi AI

ABSTRACT

KBIH organizations play an important role in providing information services to pilgrims. However, manual consultation services limited to certain operational hours have been unable to meet pilgrims' need for contextual information, particularly regarding fiqh al-hajj (Islamic jurisprudence on the pilgrimage). The use of Artificial Intelligence (AI)-based chatbots has the potential to serve as a solution; however, generative models are prone to hallucination, producing information that sounds convincing but is not supported by valid sources, a high-risk issue in the religious domain. This study aims to develop and evaluate a hajj service chatbot based on IndoT5-Base, optimized using Low-Rank Adaptation and integrated with a tiered Retrieval-Augmented Generation mechanism as a strategy for mitigating AI hallucination. The research methodology adapts the CRISP-DM framework, encompassing the stages of data collection from the FAQ of KBIH Miftahul Ulum and the 2024 Indonesian Ministry of Religious Affairs' Guidebook for Hajj and Umrah Rituals (Buku Tuntunan Manasik Haji dan Umrah), data augmentation, model fine-tuning with LoRA, implementation of a tiered retrieval mechanism with four response modes (direct retrieval, context-based generation, abstention, and locked retrieval), and evaluation using ROUGE-L, BLEU-4, BERTScore, response time metrics, and manual validation by KBIH administrators. Testing on 6,306 data points showed that the Hybrid v1 model achieved a ROUGE-L score of 0.517 and a BERTScore-F1 of 0.833, while the Hybrid v2 approach (tiered retrieval) produced the highest BLEU-4 score (0.401) and with the fastest response time (0.15 seconds). Manual validation of 20 sample answers yielded an average factual accuracy score of 4.54 and a completeness score of 4.14 on a 5-point scale, with the tiered retrieval mechanism proven to eliminate fabrication risk in 48.7% of questions answered directly from the verified knowledge base. Although the hallucination-reduction target was not fully achieved, with a residual misleading-response potential of 15%, the results demonstrate that the tiered retrieval approach effectively limits hallucination risk in high-risk question categories.

Keywords: Chatbot, Hajj, IndoT5, LoRA, Retrieval-Augmented Generation, AI Hallucination