

ABSTRAK

Penelitian ini dilatarbelakangi oleh kebutuhan akan sistem informasi kampus yang mampu memberikan jawaban relevan meskipun informasi akademik dan berita kampus terus diperbarui setiap semester. Pendekatan *Retrieval-Augmented Generation* (RAG) dapat meningkatkan kualitas jawaban dengan memanfaatkan dokumen sebagai sumber informasi, namun penggunaan metode *retrieval* tunggal masih memiliki keterbatasan dalam menangani karakteristik kueri yang beragam. Penelitian ini bertujuan mengimplementasikan *Adaptive Hybrid Retrieval* yang menggabungkan BM25, FAISS, *Query-aware Parameter Selection*, dan *Time-Based Weighting* pada *chatbot* informasi kampus Universitas Amikom Purwokerto. Penelitian menggunakan data pengumuman akademik, berita kampus, dan informasi umum yang diproses menjadi korpus *retrieval*, kemudian dievaluasi menggunakan 30 kueri yang terdiri atas kategori leksikal, semantik, dan temporal. Kinerja sistem diukur menggunakan metrik *Precision*, *Recall*, *Mean Reciprocal Rank* (MRR), dan *normalized Discounted Cumulative Gain* (nDCG). Hasil penelitian menunjukkan bahwa *Adaptive Hybrid Retrieval* memberikan performa terbaik dibandingkan BM25, FAISS, dan *Hybrid Baseline* dengan nilai *Precision* sebesar 0,41, *Recall* sebesar 0,54, *MRR* sebesar 0,83, dan *nDCG* sebesar 0,64. Hasil tersebut menunjukkan bahwa penggabungan pencarian leksikal, pencarian semantik, serta pembobotan adaptif berdasarkan karakteristik kueri dan keterbaruan dokumen mampu meningkatkan kualitas *retrieval* secara lebih konsisten pada domain informasi kampus. Dengan demikian, *Adaptive Hybrid Retrieval* dapat diterapkan sebagai mekanisme *retrieval* yang efektif untuk mendukung *chatbot* layanan informasi kampus berbasis *Retrieval-Augmented Generation*.

Kata kunci: *Hybrid Retrieval*, BM25, FAISS, *Retrieval-Augmented Generation*, *Chatbot*

ABSTRACT

The increasing volume of campus information and the diversity of user queries require an information retrieval system capable of providing accurate and relevant responses. This study aims to implement an Adaptive Hybrid Retrieval approach by integrating BM25, FAISS, Query-aware Parameter Selection, and Time-Based Weighting into a Retrieval-Augmented Generation (RAG)-based campus information chatbot at Universitas Amikom Purwokerto. The dataset consists of academic announcements, campus news, and general campus information that were processed into a retrieval corpus. The system was evaluated using 30 test queries representing lexical, semantic, and temporal query categories. Retrieval performance was measured using Precision, Recall, Mean Reciprocal Rank (MRR), and normalized Discounted Cumulative Gain (nDCG). The experimental results show that Adaptive Hybrid Retrieval achieved the best overall performance compared with BM25, FAISS, and the Hybrid Baseline, obtaining a Precision of 0.41, Recall of 0.54, MRR of 0.83, and nDCG of 0.64. These findings indicate that combining lexical retrieval, semantic retrieval, and adaptive weighting based on query characteristics and document freshness improves retrieval quality more consistently for campus information services. Therefore, Adaptive Hybrid Retrieval can serve as an effective retrieval component for Retrieval-Augmented Generation-based campus information chatbots.

Keywords: Hybrid Retrieval, BM25, FAISS, Retrieval-Augmented Generation, Chatbot