

ABSTRAK

Layanan informasi akademik di Program Studi Magister Teknologi Informasi Universitas Malikussaleh masih bersifat statis dan tersebar di berbagai halaman website, sehingga mahasiswa kesulitan mengakses data secara cepat dan tepat. Penelitian ini bertujuan mengembangkan *chatbot* akademik berbasis *Large Language Model* (LLM) dengan metode *hybrid retrieval* yang menggabungkan data tidak terstruktur dari dokumen akademik dan data terstruktur dari basis data relasional, serta mengimplementasikan *Token-Oriented Object Notation* (TOON) untuk efisiensi token. Model LLM yang digunakan adalah *meta-llama/llama-4-scout-17b-16e-instruct* melalui platform Groq, sedangkan model *embedding* yang digunakan adalah *paraphrase-multilingual-MiniLM-L12-v2*. Metode yang diterapkan meliputi pemecahan dokumen menjadi *chunk* menggunakan *semantic/structural chunking*, pembuatan *embedding*, penyimpanan ke ChromaDB, serta pencarian dokumen relevan menggunakan *hybrid search* berbasis *cosine similarity* dan *keyword matching*. Untuk data terstruktur, sistem menggunakan *Text-to-SQL* dan mengonversi hasil *query* ke format TOON. Kedua jalur diintegrasikan melalui *Intent Router* berbasis *prompt-based classification*. Evaluasi kinerja *chatbot* dilakukan menggunakan metrik ROUGE-1, ROUGE-2, dan ROUGE-L untuk jalur data tidak terstruktur, serta *Execution Accuracy* dan *Token Reduction Rate* untuk jalur data terstruktur. Hasil evaluasi menunjukkan *chatbot* mampu memberikan jawaban relevan sesuai dokumen dan basis data. Pada jalur RAG, sistem memperoleh rata-rata F1-Score sebesar 92,63% (ROUGE-1), 89,59% (ROUGE-2), dan 92,54% (ROUGE-L). Pada jalur data terstruktur, sistem memperoleh *Execution Accuracy* 100%, sedangkan TOON berhasil menghemat token rata-rata 31,48% dibandingkan JSON. Secara keseluruhan, *chatbot* akademik berbasis LLM dengan *hybrid retrieval* dan TOON dinilai efektif membantu mahasiswa memperoleh informasi akademik secara cepat, tepat, kontekstual, dan efisien.

Kata Kunci: *Chatbot Akademik, Hybrid Retrieval, Large Language Model, RAG, TOON*

ABSTRACT

Academic information services at the Master of Information Technology Study Program, Malikussaleh University, remain static and scattered across various website pages, making it difficult for students to access data quickly and accurately. This research aims to develop an academic chatbot based on Large Language Model (LLM) with hybrid retrieval method that combines unstructured data from academic documents and structured data from relational databases, as well as implementing Token-Oriented Object Notation (TOON) to token efficiency. The LLM model used is meta-llama/llama-4-scout-17b-16e-instruct through the Groq platform, while the embedding model used is paraphrase-multilingual-MiniLM-L12-v2. The methods applied include splitting documents into chunks using semantic/structural chunking, generating embeddings, storing them in ChromaDB, and retrieving relevant documents using hybrid search based on cosine similarity and keyword matching. For structured data, the system uses Text-to-SQL and converts query results to TOON format. Both pipelines are integrated through an Intent Router based on prompt-based classification. Chatbot performance evaluation was conducted using ROUGE-1, ROUGE-2, and ROUGE-L metrics for unstructured data, as well as Execution Accuracy and Token Reduction Rate for structured data. The evaluation results show that the chatbot is able to provide relevant answers according to documents and database data. In the RAG pipeline, the system achieved average F1-Score of 92.63% (ROUGE-1), 89.59% (ROUGE-2), and 92.54% (ROUGE-L). In the structured data pipeline, the system achieved Execution Accuracy of 100%, while TOON successfully saved an average of 31.48% tokens compared to JSON. Overall, the LLM-based academic chatbot with hybrid retrieval and TOON is considered effective in helping students obtain academic information quickly, accurately, contextually, and efficiently.

Keywords: *Academic Chatbot, Hybrid Retrieval, Large Language Model, RAG, TOON*