

ABSTRAK

Peningkatan penggunaan produk herbal di Indonesia diikuti oleh kebutuhan masyarakat terhadap informasi yang cepat, relevan, dan bersumber jelas dalam menentukan produk yang sesuai dengan keluhan kesehatan. Penelitian ini bertujuan untuk merancang, mengimplementasikan, dan mengevaluasi HerbaAI, yaitu *chatbot* berbasis web untuk konsultasi dan rekomendasi produk herbal HNI yang mengintegrasikan model LLaMA 3.1 8B Instant dengan pendekatan *Retrieval-Augmented Generation* (RAG). Pengembangan sistem menerapkan metode *Rapid Application Development* yang terdiri atas tahapan *Requirements Planning*, *User Design*, *Construction*, dan *Cutover*. Basis pengetahuan dibangun dari 306 dokumen penyakit dan produk HNI melalui proses praproses, *chunking*, pembentukan *embedding* menggunakan *intfloat/multilingual-e5-large-instruct*, serta penyimpanan pada ChromaDB. Sistem dikembangkan menggunakan Next.js, FastAPI, LangChain, SQLite, dan Groq API. Evaluasi dilakukan melalui *black box testing*, ROUGE, dan RAGAS terhadap *golden dataset* yang memuat 20 pertanyaan. Sistem RAG memperoleh nilai ROUGE-1 sebesar 0,4998, ROUGE-2 sebesar 0,3533, dan ROUGE-L sebesar 0,4437, lebih tinggi dibandingkan sistem non-RAG. Evaluasi RAGAS menghasilkan nilai *faithfulness* sebesar 0,8440, *answer relevancy* sebesar 0,9518, *context precision* sebesar 0,9211, dan *context recall* sebesar 0,7504. Hasil penelitian menunjukkan bahwa HerbaAI mampu menghasilkan jawaban yang relevan dan didukung oleh konteks, meskipun cakupan konteks hasil *retrieval* masih perlu ditingkatkan.

Kata kunci: *Chatbot*, Herbal, LLaMA 3.1, *Retrieval-Augmented Generation*, RAGAS

ABSTRACT

The increasing use of herbal products in Indonesia has been accompanied by a growing public need for information that is rapid, relevant, and supported by clearly identified sources to assist in selecting products appropriate to specific health concerns. This study aims to design, implement, and evaluate HerbaAI, a web-based chatbot for consultation and recommendation of HNI herbal products that integrates the LLaMA 3.1 8B Instant model with a Retrieval-Augmented Generation (RAG) approach. The system was developed using the Rapid Application Development method, consisting of the Requirements Planning, User Design, Construction, and Cutover stages. The knowledge base was constructed from 306 documents on diseases and HNI products through preprocessing, chunking, embedding generation using intfloat/multilingual-e5-large-instruct, and storage in ChromaDB. The system was developed using Next.js, FastAPI, LangChain, SQLite, and the Groq API. Evaluation was conducted through black-box testing, ROUGE, and RAGAS using a golden dataset containing 20 questions. The RAG-based system achieved ROUGE-1, ROUGE-2, and ROUGE-L scores of 0.4998, 0.3533, and 0.4437, respectively, outperforming the non-RAG system. The RAGAS evaluation yielded scores of 0.8440 for faithfulness, 0.9518 for answer relevancy, 0.9211 for context precision, and 0.7504 for context recall. The results indicate that HerbaAI is capable of generating relevant and context-supported responses, although the coverage of retrieved context still requires improvement.

Keywords: Chatbot, Herbal Medicine, LLaMA 3.1, Retrieval-Augmented Generation, RAGAS