

## **BAB V**

### **PENUTUP**

#### **5.1 Kesimpulan**

Penelitian ini berhasil mengimplementasikan chatbot layanan informasi aturan akademik UPNVJT berbasis *Retrieval-Augmented Generation* (RAG) dengan mengintegrasikan *Large Language Model* Gemma 3 berukuran 4B parameter melalui Ollama serta basis pengetahuan yang berasal dari dokumen Peraturan Rektor Nomor 25 Tahun 2023 dan Peraturan Rektor Nomor 06 Tahun 2025. Proses prapemrosesan teks yang mencakup ekstraksi dokumen PDF, pembersihan teks, normalisasi struktur dokumen, dan proses *chunking* berhasil membentuk basis pengetahuan yang terstruktur untuk mendukung mekanisme *retrieval*. Penyimpanan potongan dokumen ke dalam ChromaDB dengan representasi *embedding* mendukung proses pencarian informasi yang relevan berdasarkan pertanyaan pengguna. Implementasi arsitektur tersebut menghasilkan sistem chatbot yang mampu memberikan jawaban kontekstual berdasarkan informasi yang tersedia pada dokumen akademik.

Berdasarkan evaluasi terhadap enam skenario pengujian menggunakan metrik RAGAS, konfigurasi terbaik diperoleh pada skenario dengan parameter *chunk size* 1024, *chunk overlap* 400, dan *top-k* 10. Konfigurasi tersebut memperoleh nilai *Faithfulness* sebesar 0.7950, *Answer Relevancy* sebesar 0.7868, *Context Precision* sebesar 0.8093, dan *Context Recall* sebesar 0.8958, yang menunjukkan kualitas jawaban yang baik dari aspek akurasi, relevansi, presisi, dan kelengkapan konteks. Evaluasi *retrieval* juga menunjukkan bahwa konfigurasi *chunk* yang lebih besar dengan *overlap* yang tinggi memberikan performa yang lebih baik dalam menemukan informasi yang relevan dari dokumen akademik. Dengan demikian, chatbot berbasis RAG yang dikembangkan dinilai mampu mendukung penyediaan layanan informasi aturan akademik secara lebih akurat, relevan, dan responsif.

## 5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat dipertimbangkan untuk pengembangan sistem pada penelitian selanjutnya, antara lain :

1. Sistem chatbot pada penelitian ini masih menggunakan pendekatan *stateless*, sehingga setiap pertanyaan diproses secara independen tanpa mempertimbangkan riwayat percakapan sebelumnya. Penelitian selanjutnya disarankan mengembangkan mekanisme *multi-turn conversation* atau *conversational memory* agar chatbot mampu memahami konteks percakapan secara lebih natural dan menghasilkan respons yang lebih kontekstual bagi pengguna.
2. Penggunaan model LLM dengan jumlah parameter yang lebih besar atau model yang telah dilatih khusus pada korpus bahasa Indonesia diharapkan dapat meningkatkan kualitas jawaban sistem. Peningkatan ini terutama ditargetkan pada metrik *Faithfulness* dan *Answer Relevancy* yang masih memiliki ruang perbaikan berdasarkan hasil evaluasi pada penelitian ini.
3. Perluasan basis pengetahuan dengan menambahkan dokumen akademik lain seperti SOP, Surat Keputusan, dan panduan program studi dapat memperluas cakupan informasi yang mampu dijawab oleh sistem. Penambahan sumber dokumen ini diharapkan mampu meningkatkan kelengkapan sistem dalam menjawab berbagai kebutuhan informasi akademik civitas akademika UPNVJT secara lebih menyeluruh.
4. Eksplorasi pendekatan *fine-tuning* model *embedding* pada domain dokumen regulasi berbahasa Indonesia perlu dipertimbangkan untuk meningkatkan akurasi pencarian semantik. Model *embedding* yang dilatih pada korpus domain spesifik diharapkan menghasilkan representasi vektor yang lebih tepat sehingga proses *retrieval* dokumen relevan menjadi lebih presisi.