

BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penerapan LLM mampu menghasilkan teks pembaruan artikel Wikipedia yang relevan berdasarkan informasi baru dari artikel berita yang diberikan. Proses ini dimulai dengan pengumpulan data berupa pasangan revisi artikel Wikipedia dan tautan referensi berita, yang kemudian dikombinasikan dalam bentuk *prompt* untuk diberikan sebagai masukan kepada model LLM. Eksperimen dilakukan terhadap tiga model, yaitu Phi-3-mini sebagai *baseline*, serta Gemma 3 dalam dua varian (1B dan 4B) sebagai model utama. Hasil evaluasi menggunakan metrik ROUGE menunjukkan bahwa model Gemma 3 4B mampu memberikan keluaran dengan kualitas tertinggi. Dua kombinasi model dan strategi *prompting* terbaik ditunjukkan oleh Gemma 3 4B sebelum *finetuning* dengan strategi *system prompt* berbahasa Inggris dan Gemma 3 4B setelah *finetuning* dengan strategi *simple prompt*. Selain itu, proses *fine-tuning* terhadap model Gemma 3 1B dan 4B menggunakan data pelatihan yang disusun terbukti meningkatkan kualitas keluaran dibanding model *pretrained*-nya. Hal ini menunjukkan bahwa adaptasi LLM terhadap domain Wikipedia dan gaya penulisannya penting agar dapat menghasilkan keluaran yang lebih sesuai. Beberapa hasil keluaran model juga telah berhasil untuk disunting secara manual ke dalam artikel Wikipedia, sehingga menunjukkan potensi nyata penerapan LLM sebagai alat bantu penyunting dalam praktik pembaruan artikel biografi Wikipedia.

Dengan demikian, dapat disimpulkan bahwa LLM, terutama model Gemma 3, dapat diterapkan dalam pembaruan artikel biografi Wikipedia dengan sumber berita dengan hasil yang optimal apabila disertai perancangan *prompt* yang tepat dan pelatihan lanjutan (*fine-tuning*) pada data yang relevan.

5.2 Saran

Berdasarkan hasil penelitian ini, terdapat beberapa saran yang dapat dipertimbangkan untuk pengembangan dan penelitian lanjutan:

1. Penelitian ini masih terbatas pada jumlah data dan domain biografi tokoh Indonesia. Penelitian selanjutnya disarankan untuk memperluas cakupan data, baik dari segi jumlah revisi maupun jenis artikel, agar model dapat dilatih dan diuji secara lebih komprehensif.
2. Hasil menunjukkan bahwa variasi strategi *prompting* berdampak signifikan terhadap performa model. Oleh karena itu, eksplorasi lebih lanjut terhadap jenis-jenis *prompt* dapat menjadi arah pengembangan yang potensial.
3. Untuk pemanfaatan praktis, hasil dari model dapat diintegrasikan sebagai sistem pendukung penyuntingan semi-otomatis di Wikipedia. Dengan sistem validasi manusia, pendekatan ini berpotensi mempercepat proses pembaruan informasi sambil tetap menjaga akurasi dan kualitas konten.