

BAB V

SIMPULAN DAN SARAN

5.1 Simpulan

Pengolahan data, terutama dalam skala besar (big data), menjadi semakin krusial dalam era digital saat ini. SQL, sebagai bahasa pemrograman utama untuk mengelola data, memainkan peran penting dalam menggali informasi berharga. Namun, kompleksitas SQL seringkali menjadi hambatan bagi pengguna non-teknis. Natural Language Processing (NLP), khususnya model bahasa besar seperti Llama 2, menawarkan solusi dengan kemampuannya menerjemahkan perintah bahasa alami menjadi query SQL yang valid. Penelitian ini menguji potensi fine-tuning model Llama 2 menggunakan metode QLoRA untuk meningkatkan kemampuannya dalam memahami dan menjawab pertanyaan terkait SQL.

Fine-tuning model LLM merupakan sebuah metode yang dapat dilakukan untuk melatih model tersebut untuk menjalankan *task* yang spesifik berdasarkan *dataset* yang digunakan. Terdapat beberapa metode *fine-tuning* yang dapat digunakan untuk memberikan hasil yang maksimal dengan *resources* yang terbatas. Penelitian ini dilakukan proses *fine-tuning* menggunakan metode QLoRA dimana metode ini berfokus pada kuantisasi metode sebelumnya yaitu LoRA dengan memanfaatkan beberapa *attention layer* yang menjadi target *training* tanpa perlu menggunakan keseluruhan parameter pada model Llama-2 yang digunakan.

Pada penelitian ini, dilakukan sebuah *fine-tuning* dengan menggunakan 10ribu baris data dan menggunakan 1 *epoch* untuk pelatihannya sehingga model akan dilatih sebanyak 1 putaran menggunakan 10ribu data tersebut. Proses *training* berjalan selama 3 jam 10 menit dengan total 2500 steps. Selanjutnya dilakukan uji coba menggunakan *test dataset* dan model berhasil menjawab 4 dari 5 pertanyaan dengan kesamaan pada *test dataset*. Dilakukan juga perhitungan terhadap nilai Akurasi, *precision*, *recall*, *F1-Score*, dan *ROUGE Score*. Pada perhitungan ini, didapat nilai akurasi sebesar 80% secara kualitatif dan 0.90 secara kuantitatif, 0.99 pada *Precision*, 0.89 pada *Recall*, 0.94 pada *F1-Score*, 0.97 pada ROUGE-1 dan

0.96 pada ROUGE-L. Model Llama-2 yang sudah di *fine-tune* ini juga memberikan hasil yang lebih baik terhadap *test dataset* daripada model Llama-2 yang awal sebelum dilakukan *fine-tuning*.

Implementasi *fine-tuning* terhadap model Llama-2 dalam pembuatan model text-to-SQL berhasil dilakukan dengan menggunakan dataset sql-create-context yang berisikan pertanyaan dan jawaban mengenai text-to-SQL. Proses *fine-tuning* dilakukan pada model Llama-2-7b-chat-hf yang memiliki total 7 miliar parameter. Hasil performa dari model Llama-2 yang telah di-*fine-tune* menunjukkan bahwa model mampu mentranslasikan pertanyaan dalam bahasa Inggris menjadi query SQL dengan tingkat akurasi yang memuaskan. Model ini juga berhasil diintegrasikan ke dalam sebuah sistem chatbot yang memudahkan pengguna dalam mengakses informasi mengenai database.

Penelitian ini memiliki beberapa keterbatasan yang perlu diperhatikan. Proses *fine-tuning* model Llama-2 membutuhkan sumber daya komputasi yang besar, mengingat model dengan 7 miliar parameter ini memerlukan perangkat keras dengan kapasitas tinggi untuk memastikan proses pelatihan berjalan dengan baik dan efisien. Selain itu, model yang dihasilkan memiliki kelemahan dalam menjawab pertanyaan yang menggunakan fungsi agregat tertentu, seperti MIN, yang menunjukkan bahwa model belum sepenuhnya optimal dalam memahami dan memproses semua jenis query SQL. Jumlah dataset yang digunakan untuk proses *fine-tuning* juga masih terbatas, dengan hanya 10.000 baris data, yang mungkin belum cukup untuk mencakup seluruh variasi dan kompleksitas query SQL yang mungkin ditemui dalam penggunaan dunia nyata. Penambahan jumlah dan variasi dataset dapat membantu meningkatkan kemampuan model dalam memahami berbagai jenis query SQL.

5.2 Saran

Berdasarkan proses penelitian yang telah dilakukan, penulis dapat menyampaikan beberapa saran yang dapat diperhatikan sebagai panduan bagi penelitian berikutnya karena keterbatasan *resources* serta waktu yang dimiliki, antara lain sebagai berikut.

1. Menggunakan *dataset* yang lebih sesuai dengan *real-life usage* dengan jumlah *record* data yang lebih banyak.
2. Model Llama-2 memiliki parameter yang sangat besar jumlahnya dan keterbatasan *resources* untuk melatih model secara keseluruhan dapat menjadi salah satu faktor untuk efektivitas *fine-tuning* model tersebut. Maka dari itu dapat ditemukan metode *fine-tuning* yang lebih efisien dan *resources* yang lebih mumpuni
3. Disediakan sebuah sarana untuk proses *fine-tuning* dengan *resources* yang lebih baik sehingga mampu melakukan proses *training* dengan *dataset* serta konfigurasi yang lebih baik lagi.

UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA