

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan pada hasil penelitian, model komputasi penilaian skor literasi membaca individu dengan proses *data collection*, *cleansing data*, transformasi data menjadi *dictionary*, tokenisasi, *model FLAN-T5 constructions*, *fine tuning model FLAN-T5*, dan *testing finetuned model* telah berhasil dibuat. Penelitian ini berhasil mengimplementasikan model komputasi menggunakan *FLAN-T5*, yang telah diuji dan menunjukkan kinerja yang memadai. Pendekatan menggunakan prompt yang konsisten dan terstruktur, termasuk pertanyaan, jawaban, dan instruksi penilaian, membantu model dalam memberikan penilaian skor literasi membaca individu yang cenderung akurat.

Mengenai analisa akurasi model dalam penilaian skor uraian literasi membaca individu adalah bahwa penggunaan *FLAN-T5* menunjukkan efektivitasnya dalam mengidentifikasi jawaban pada berbagai tingkat skor literasi, dengan tingkat akurasi diatas 70% di setiap kategori skor dan akurasi 74% secara keseluruhan. Meskipun terdapat beberapa kesalahan, terutama dalam menilai jawaban yang seharusnya mendapatkan skor lebih tinggi atau lebih rendah, model ini secara keseluruhan mampu memberikan analisa dalam penilaian skor uraian literasi membaca individu, terutama dalam mengenali uraian di setiap skor nya. Analisa akurasi kinerja model di penelitian ini menunjukkan perolehan 90% pada kombinasi parameter tertentu pada eksperimen *cross validation*, menunjukkan bahwa pemilihan kombinasi parameter yang tepat dapat mempengaruhi akurasi terhadap kinerja model. *Error rate* juga mengalami penurunan pada eksperimen *cross validation*, menunjukkan peningkatan kemampuan model dalam penilaian dipengaruhi oleh pemilihan kombinasi parameter yang tepat. Waktu komputasi meningkat seiring bertambahnya *epoch* dalam menghasilkan *fine-tuned* model, namun ini merupakan hal yang normal karena semakin besar *epoch* maka iterasi yang dilakukan pun semakin banyak.

5.2 Saran

Dalam pelaksanaan penelitian ini, penulis menyadari bahwa masih terdapat sejumlah kekurangan yang perlu diperhatikan. Adapun saran yang dapat diberikan untuk penelitian selanjutnya adalah sebagai berikut.

1. Mengembangkan dataset dengan variasi yang lebih besar untuk masing-masing kelas untuk dapat meningkatkan kemampuannya dalam membedakan antara tingkat skor yang berbeda.
2. Menambahkan tahap uji model dengan berbagai variasi prompt untuk mengidentifikasi yang dapat mengoptimalkan penilaian oleh model dalam berbagai konteks.
3. *Fine tuning* model dengan data yang memberikan penekanan pada perbedaan kontekstual untuk meningkatkan kemampuan model dalam memahami nuansa jawaban.
4. Menerapkan penelitian dengan fokus pada populasi yang berbeda untuk memastikan generalisasi dan efektivitas model dalam berbagai situasi dan kelompok pengguna.
5. Pastikan penggunaan sumber daya komputasi yang memadai untuk memungkinkan eksperimen dengan kompleksitas yang lebih tinggi dan khususnya *fine tuning* model dengan *epoch* yang lebih besar.