

BAB V

PENUTUP

5.1. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, maka dapat ditarik kesimpulan yang menjawab rumusan masalah sebagai berikut.

1. Model *fine-tuning* IndoBERT mampu melakukan klasifikasi sentimen berbasis aspek pada ulasan aplikasi JConnect Mobile dengan performa yang baik. Proses pelatihan model dilakukan menggunakan metode *5-Fold Cross Validation* dengan kombinasi *hyperparameter* berupa *epoch* 5 dan 10, *batch size* 16 dan 32, serta *dropout* 0,2, 0,4, dan 0,6 untuk memperoleh konfigurasi terbaik pada setiap aspek. Berdasarkan hasil *5-Fold Cross Validation*, model *fine-tuning* IndoBERT secara konsisten menghasilkan performa yang lebih baik dibandingkan *fine-tuning* mBERT pada seluruh aspek. Pada aspek tampilan, *fine-tuning* IndoBERT memperoleh *mean F1-Score* sebesar 0,86, lebih tinggi dibandingkan *fine-tuning* mBERT yang memperoleh 0,80. Pada aspek fitur dan performa, *fine-tuning* IndoBERT memperoleh *mean F1-Score* sebesar 0,83, sedangkan *fine-tuning* mBERT memperoleh 0,79. Sementara itu, pada aspek layanan, *fine-tuning* IndoBERT memperoleh *mean F1-Score* sebesar 0,74, lebih tinggi dibandingkan mBERT yang memperoleh 0,67. Hasil tersebut menunjukkan bahwa *fine-tuning* IndoBERT lebih mampu memahami karakteristik ulasan berbahasa Indonesia dibandingkan mBERT. Model IndoBERT dengan konfigurasi *hyperparameter* terbaik selanjutnya dievaluasi menggunakan data uji untuk mengukur kemampuan generalisasi terhadap data yang belum pernah digunakan selama proses pelatihan. Berdasarkan hasil evaluasi, model memperoleh *Macro F1-Score* sebesar 0,80 pada aspek tampilan, 0,83 pada aspek fitur dan performa, serta 0,81 pada aspek layanan. Hasil tersebut menunjukkan bahwa model mampu mempertahankan performa yang baik dalam melakukan klasifikasi sentimen berbasis aspek pada data baru.
2. Model dengan performa terbaik berhasil diimplementasikan ke dalam sistem berbasis *web*. Sistem menyediakan fitur utama, yaitu analisis ulasan tunggal, analisis massal, dan *dashboard* dokumentasi *dataset*. Pada fitur analisis

tunggal, sistem dapat menerima *input* ulasan dan menampilkan prediksi sentimen beserta *confidence score* untuk setiap aspek. Pada fitur analisis massal, sistem mampu memproses file CSV dan menghasilkan visualisasi hasil analisis sentimen. Selain itu, fitur *dashboard* dokumentasi *dataset* menyajikan hasil prediksi seluruh ulasan JConnect Mobile periode 2021–2025 dalam bentuk visualisasi yang informatif, sehingga memudahkan pengguna dalam memahami distribusi sentimen pada setiap aspek.

5.2. Saran

Berdasarkan keseluruhan hasil penelitian, beberapa saran disampaikan sebagai bahan pertimbangan dalam pengembangan penelitian selanjutnya, yaitu sebagai berikut.

1. Pada aspek tampilan tidak ditemukan ulasan dengan sentimen netral sehingga model tidak dapat dilatih untuk mengenali maupun memprediksi sentimen netral pada aspek tampilan. Oleh karena itu, penelitian selanjutnya disarankan menggunakan *dataset* yang lebih beragam dengan distribusi kelas sentimen yang lebih seimbang, sehingga seluruh kelas sentimen dapat dipelajari oleh model dan menghasilkan kemampuan klasifikasi yang lebih baik.
2. Melakukan penentuan dan identifikasi aspek yang divalidasi oleh ahli di bidang perbankan dan ahli evaluasi kualitas perangkat lunak untuk memastikan aspek yang digunakan relevan, komprehensif, dan sesuai dengan konteks aplikasi *mobile banking*.
3. Membandingkan performa IndoBERT dengan model *transfer learning* lain yang terbaru, seperti XLM-RoBERTa dan *Large Language Model* (LLM), untuk mengetahui model dengan performa terbaik dalam analisis sentimen bahasa Indonesia.
4. Mengembangkan sistem analisis berbasis *web* agar dapat analisis secara *real-time* dengan mengintegrasikan langsung dengan API Google Play Store, App Store, dan media sosial agar proses pengambilan data ulasan dapat dilakukan secara otomatis.