

BAB V

KESIMPULAN DAN SARAN

5.1 KESIMPULAN

Berdasarkan hasil penelitian mengenai analisis sentimen ulasan aplikasi SIGNAL–Samsat Digital Nasional menggunakan metode IndoBERT, dapat disimpulkan beberapa hal sebagai berikut:

1. Performa Model pada Dataset Utama

Penelitian ini menunjukkan bahwa model IndoBERT mampu memberikan performa yang sangat baik dalam melakukan klasifikasi sentimen terhadap 10.000 ulasan aplikasi SIGNAL dari Google Play Store. Setelah dilakukan fine-tuning pada model *indobenchmark/indobert-base-p2* dengan konfigurasi terbaik pada *learning rate* $2e-5$, model menghasilkan akurasi sebesar 99,5% dengan nilai precision, recall, dan F1-score yang tinggi dan konsisten pada seluruh kategori sentimen (positif, netral, dan negatif). Hasil analisis menunjukkan bahwa mayoritas ulasan pengguna bersifat netral (43,5%) dan positif (41,6%), sedangkan hanya 14,9% yang negatif. Visualisasi wordcloud mendukung hasil ini, memperlihatkan kata-kata seperti "pajak", "bayar", "bantu", dan "terima kasih" yang mendominasi ulasan positif, meskipun beberapa kata seperti "tidak", "gagal", dan "proses" menunjukkan adanya keluhan pengguna yang masih perlu diperhatikan oleh pengembang aplikasi.

2. Hasil Pengujian Model terhadap Dataset Eksternal

Pengujian pada tiga dataset eksternal menunjukkan bahwa kemampuan generalisasi model IndoBERT bervariasi. Pada dataset Mobil Listrik, akurasi mencapai 70,3%, namun sentimen negatif dan positif sering salah diklasifikasikan sebagai netral. Dataset Review Threads memberikan akurasi lebih rendah, 57,4%, terutama pada sentimen negatif. Sementara itu, dataset Review Dana menunjukkan hasil terbaik dengan akurasi 82,6%, walau kelas negatif tetap sulit dikenali. Hasil ini mengindikasikan bahwa model bekerja baik pada domain yang mirip dengan data latih, tetapi performanya menurun pada domain berbeda dan cenderung bias terhadap kelas netral.

5.2 SARAN

1. Eksplorasi Model dan Parameter Lain – Penelitian selanjutnya dapat mencoba model lain seperti IndoBERTweet atau IndoRoBERTa, serta melakukan pengujian dengan variasi jumlah *epoch*, dan *batch size* untuk mendapatkan performa yang lebih optimal.
2. Eksplorasi terhadap pendekatan aspect-based sentiment analysis (ABSA) dapat memberikan pemahaman yang lebih mendalam terhadap aspek layanan atau fitur yang paling banyak mendapat ulasan negatif atau positif.
3. Pemanfaatan Metadata Ulasan – Menambahkan data seperti rating bintang, tanggal ulasan, dan versi aplikasi dapat membantu menganalisis perubahan sentimen dari waktu ke waktu.

