

BAB V PENUTUP

5.1 Kesimpulan

Penelitian ini bertujuan untuk merancang dan mengimplementasikan metode *ensemble weighted soft-voting* dalam menggabungkan Logistic Regression dan IndoBERT pada tugas deteksi aspek multi-label dan klasifikasi sentimen per aspek terhadap cuitan Twitter berbahasa Indonesia mengenai Timnas Indonesia pada era kepelatihan Patrick Kluivert. Perancangan *ensemble* dilakukan melalui penggabungan probabilitas keluaran kedua model dasar dengan penentuan bobot dan *threshold optimal* menggunakan data validasi tanpa melakukan pelatihan ulang model. Berdasarkan rangkaian proses pengumpulan data, praproses, pemodelan, serta evaluasi yang telah dilakukan, diperoleh beberapa temuan utama sebagai berikut:

1. Pada tugas deteksi aspek multi-label, *ensemble* memperoleh nilai Macro-F1 sebesar 0,87, lebih tinggi dibandingkan IndoBERT dan Logistic Regression. Peningkatan juga terlihat pada Micro-F1 dan Macro-AUC, dimana *ensemble* memperoleh nilai 0,86 dan 0,95.
2. Pada tugas klasifikasi sentimen per aspek, *ensemble* mencapai Macro-F1 sebesar 0,77 dan Micro-F1 sebesar 0,77, lebih tinggi dibandingkan IndoBERT maupun Logistic Regression.

Pendekatan *ensemble weighted soft-voting* memberikan peningkatan kinerja dibandingkan kedua model tunggal dalam deteksi aspek multi-label dan klasifikasi sentimen per aspek. Meskipun peningkatan yang diperoleh relatif kecil, hasil tersebut menunjukkan bahwa kombinasi kedua model memberikan kontribusi tambahan dalam meningkatkan stabilitas prediksi pada analisis sentimen multi-aspek.

5.2 Saran

Meskipun penerapan metode *ensemble weighted soft-voting* menunjukkan peningkatan performa yang konsisten dibandingkan model tunggal, peningkatan tersebut tidak terlalu signifikan. Oleh karena itu, penelitian selanjutnya dapat mengeksplorasi teknik *ensemble* lainnya, seperti *stacking* atau *blending*, untuk mengetahui apakah kombinasi model yang lebih kompleks dapat memberikan peningkatan kinerja yang lebih signifikan. Penelitian selanjutnya juga dapat mempertimbangkan penggunaan teknik evaluasi dan pendekatan pengujian tambahan untuk memperoleh gambaran performa model yang lebih menyeluruh serta dapat membantu memperkuat kesimpulan terkait efektivitas metode *ensemble*.

Selain itu, perluasan cakupan data dan periode waktu pengambilan data juga dapat dipertimbangkan untuk menguji konsistensi performa model pada kondisi opini publik yang berbeda. Hal ini penting untuk menguji ketahanan pendekatan *ensemble* terhadap variasi distribusi aspek dan sentimen. Penelitian selanjutnya juga dapat mempertimbangkan pemanfaatan model bahasa yang lebih mutakhir atau pendekatan berbasis LLM (*Large Language Model*) untuk mengevaluasi kemungkinan peningkatan pemahaman makna teks dalam tugas analisis sentimen multi-aspek.