

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan mengenai implementasi ensemble learning untuk analisis sentimen pada aplikasi Reku, dapat disimpulkan beberapa hal sebagai berikut:

1. Berdasarkan perbandingan hasil pengukuran metrik seperti akurasi, presisi, recall, dan F1-score, model ensemble learning dengan metode Weighted Soft Voting menunjukkan performa yang lebih unggul dibandingkan dengan model tunggal (CNN, SVM, dan IndoBERT) dan Soft Voting. Pencapaian akurasi tertinggi sebesar 91% pada model Weighted Soft Voting, yang lebih tinggi dibandingkan dengan CNN (83%), SVM (84%), IndoBERT (89%), Soft Voting (89%). Model Weighted Soft Voting tidak hanya unggul dalam hal akurasi, tetapi juga menunjukkan prediksi yang lebih seimbang terhadap kedua kelas sentimen, yaitu positif dan negatif.
2. Kombinasi model yang paling efektif dalam menghasilkan performa optimal adalah dengan pendekatan Weighted Soft Voting karena mampu mengintegrasikan kekuatan dari model deep learning (CNN), machine learning klasik (SVM), dan transformer pretrained (IndoBERT). Sistem akhir dari penelitian ini juga diimplementasikan dalam bentuk aplikasi web berbasis Streamlit, yang memungkinkan pengguna untuk melakukan analisis sentimen terhadap ulasan baru secara interaktif. Validasi terhadap data terbaru menunjukkan bahwa model Weighted Soft Voting tetap mampu mempertahankan akurasi tinggi sebesar 90%, yang menandakan stabilitas performa dalam penggunaan nyata.
3. Hasil analisis sentimen terhadap data ulasan aplikasi Reku menunjukkan bahwa sentimen positif merupakan kategori yang dominan. Namun demikian, ulasan negatif tetap ditemukan dan menjadi masukan penting untuk pengembangan fitur serta peningkatan kualitas layanan di masa mendatang.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, beberapa saran yang dapat diberikan untuk pengembangan lebih lanjut adalah sebagai berikut:

1. Penelitian berikutnya disarankan untuk menambahkan kategori netral agar klasifikasi sentimen menjadi lebih representatif terhadap ekspresi pengguna yang tidak sepenuhnya positif maupun negatif. Selain itu, model yang dikembangkan sebaiknya diuji pada dataset dari aplikasi lain agar kemampuan generalisasi sistem dapat dievaluasi secara lebih menyeluruh.
2. Pengembangan sistem web dapat ditingkatkan dengan menambahkan fitur visualisasi distribusi sentimen, penjelasan hasil prediksi (*explainability*), serta mendukung lebih banyak format input. Aspek efisiensi seperti waktu inferensi dan ukuran model juga dapat menjadi fokus evaluasi agar sistem lebih ringan dan cepat diakses.
3. Menyempurnakan proses validasi dengan melibatkan lebih banyak annotator atau metode semi-otomatis lain. Selain itu, model sebaiknya dikembangkan agar lebih adaptif terhadap penggunaan bahasa yang bervariasi, termasuk bahasa gaul, singkatan, atau campuran bahasa Indonesia-Inggris yang umum ditemukan dalam ulasan pengguna.