

BAB V

PENUTUP

5.1 Kesimpulan

Penelitian ini berhasil menjawab rumusan masalah terkait kinerja algoritma klasifikasi sentimen (*SVM*, *Logistic Regression*, dan *Random Forest*) yang dioptimasi menggunakan *Grid Search* untuk klasifikasi sentimen komentar *YouTube* berbahasa Indonesia. Sebelum optimasi, model *SVM* menunjukkan akurasi tertinggi sebesar 90,97%, diikuti oleh *Random Forest* (86,72%) dan *Logistic Regression* (87,33%). Setelah optimasi, *SVM* tetap memperoleh akurasi terbaik (92,55%) dengan kombinasi parameter terbaiknya adalah '*C*': 10, '*max_iter*': 1000, '*tol*': 0.0001, diikuti oleh *Logistic Regression* (92,34%) dan *Random Forest* (88,84%), yang menunjukkan bahwa optimasi *Grid Search* berdampak positif pada performa model. Selain itu, teknik ekstraksi fitur dengan *TF-IDF*, *oversampling* menggunakan *SMOTE*, dan tahapan preprocessing (*case folding*, *text cleaning*, *slang words normalization*, *stopword removal*, *tokenizing*, dan *stemming*) terbukti signifikan dalam meningkatkan kinerja model. Penelitian ini juga berhasil mencapai tujuan untuk menganalisis dan membandingkan performa algoritma klasifikasi sentiment sebelum dan sesudah dilakukan optimasi.

5.2 Saran

Untuk penelitian lanjutan, disarankan untuk mengeksplorasi penggunaan model-model *machine learning* yang lebih canggih, seperti model berbasis *deep learning* atau *transformer*, yang dapat menangkap konteks lebih kompleks dalam analisis sentimen. Selain itu, penelitian berikutnya bisa memperkaya kamus pelabelan sentimen agar lebih sesuai dengan bahasa Indonesia, serta mengevaluasi pelabelan sentimen dengan metode manual untuk perbandingan. Penelitian lanjutan juga dapat mencakup analisis pada dataset yang lebih besar dan beragam, serta membandingkan hasil penelitian ini dengan aplikasi klasifikasi sentimen di platform media sosial lainnya. Penambahan teknik seperti *cross-validation* atau *ensemble learning* juga dapat dijadikan pertimbangan untuk meningkatkan akurasi model.