

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penelitian yang dilakukan, dapat disimpulkan bahwa algoritma *Naive Bayes Classifier (NBC)* memiliki performa yang lebih unggul dibandingkan *K-Nearest Neighbor (KNN)* dalam melakukan analisis sentimen terhadap data media sosial X pada Pilkada 2024. Dari total 1.203 tweet yang dikumpulkan, sebanyak 1.187 data berhasil digunakan setelah melalui proses *preprocessing* yang ketat. Model NBC mampu mencapai akurasi sebesar 76,89% dengan hasil *precision* dan *recall* tertinggi pada kategori sentimen positif, sementara KNN hanya mencapai akurasi 74,79% dan menunjukkan kelemahan dalam membedakan sentimen netral dan negatif. Keunggulan NBC terletak pada efisiensinya dalam menangani data teks berdimensi tinggi serta kemampuannya dalam menghasilkan prediksi yang konsisten. Meskipun kedua model mampu mengklasifikasikan data baru dengan tepat, perbedaan kinerja menunjukkan bahwa pemilihan algoritma sangat memengaruhi akurasi dan ketahanan model dalam konteks data dunia nyata.

5.2 Saran

Namun demikian, hasil penelitian ini juga mengungkap beberapa tantangan, seperti ketidakseimbangan distribusi kelas sentimen yang berpotensi memengaruhi hasil klasifikasi. Oleh karena itu, disarankan untuk menerapkan teknik penanganan *class imbalance* seperti *SMOTE* atau *undersampling* agar distribusi data menjadi lebih seimbang. Selain itu, penggunaan algoritma lanjutan seperti *Support Vector Machine*, *Random Forest*, atau metode *Deep Learning* seperti LSTM dan BERT dapat dieksplorasi untuk meningkatkan akurasi dan menangkap konteks sentimen secara lebih mendalam. Strategi evaluasi juga perlu ditingkatkan dengan menerapkan *cross-validation* yang lebih luas, seperti *5-fold* atau *10-fold*, guna mengurangi risiko *overfitting*. Dalam hal *feature engineering*, penggunaan *n-gram*, *word embedding*, dan analisis sintaksis seperti *POS tagging* dapat menjadi pendekatan yang lebih optimal. Penelitian lanjutan juga dianjurkan untuk

melakukan analisis topik dan isu utama yang berkembang selama Pilkada, serta mengembangkan sistem analisis sentimen yang mampu bekerja secara *real-time* dan berbasis lokasi untuk memberikan pemetaan opini publik yang lebih komprehensif.

