

BAB V

PENUTUP

5.1. Kesimpulan

Berdasarkan hasil skripsi yang telah dipaparkan, skripsi ini mengembangkan 48 skenario berbasis algoritma *Support Vector Machine* (SVM) dan *Random Forest* (RF). Pengembangan skenario dilakukan dengan menggunakan pembagian dataset Holdout pada rasio 80:20 dan 70:30, teknik pembobotan kata berbasis *Term Weighting* (TF-IDF dan TF-ABS), serta teknik *resampling* *Random Oversampling* (ROS) dan *Random Undersampling* (RUS). Dari 48 skenario yang diimplementasikan, disimpulkan bahwa algoritma *Support Vector Machine* (SVM) menunjukkan kinerja yang lebih unggul dibandingkan *Random Forest* (RF) dalam klasifikasi sentimen dengan deteksi sarkasme pada komentar video YouTube terkait PILKADA 2024.

Pada klasifikasi sentimen, algoritma SVM dengan teknik pembobotan kata TF-IDF, metode penyeimbangan data menggunakan *Random Oversampling* (ROS), dan pembagian data dengan rasio 80:20 menghasilkan akurasi terbaik, yaitu 0.87 pada data train dan 0.80 pada data test. Model ini juga menunjukkan stabilitas kinerja yang tinggi tanpa indikasi *overfitting* maupun *underfitting*. Sementara itu, pada klasifikasi sarkasme, algoritma SVM dengan teknik pembobotan kata TF-IDF, tanpa penerapan teknik *resampling*, dan pembagian data yang sama, memberikan hasil terbaik dengan akurasi sebesar 0.80 pada data train dan 0.73 pada data test, meskipun terdapat variasi kinerja dalam beberapa skenario.

Di sisi lain, algoritma *Random Forest* (RF) dengan teknik pembobotan kata TF-ABS, metode penyeimbangan data menggunakan *Random Oversampling* (ROS), dan pembagian data dengan rasio 70:30 menghasilkan akurasi 0.73 untuk klasifikasi sentimen. Sedangkan pada klasifikasi sarkasme, algoritma RF dengan teknik pembobotan kata TF-IDF, metode penyeimbangan data yang sama, dan pembagian data dengan rasio 70:30 mencapai akurasi sebesar 0.69. Hasil ini mengindikasikan bahwa meskipun SVM lebih unggul dalam mendeteksi sarkasme, *Random Forest* tetap menunjukkan potensi pada beberapa skenario klasifikasi.

Secara keseluruhan, algoritma SVM terbukti memiliki kinerja yang lebih baik dibandingkan RF dalam menangkap pola pada klasifikasi sentimen dengan deteksi sarkasme, menjadikannya pilihan yang lebih efektif untuk analisis pada komentar video YouTube terkait PILKADA 2024.

5.2. Saran Pengembangan

Untuk penelitian selanjutnya, beberapa saran dapat dipertimbangkan. Pertama, eksplorasi lebih lanjut terhadap langkah *preprocessing* tambahan, seperti normalisasi teks, *tokenizing*, dll, dapat dilakukan untuk meningkatkan kualitas data sebelum proses klasifikasi. Penerapan langkah-langkah ini diharapkan mampu meningkatkan akurasi model dengan mengurangi variasi kata yang tidak perlu.

Selanjutnya, penggunaan algoritma lain, seperti *deep learning* (LSTM, GRU, CNN, dll.), dapat menjadi alternatif potensial dalam meningkatkan performa model. Algoritma ini memiliki kemampuan yang lebih baik dalam memahami pola data yang kompleks, terutama dalam mendeteksi sarkasme yang sering kali memiliki struktur bahasa yang tidak langsung.

Kemudian, peningkatan jumlah dataset juga perlu dipertimbangkan agar model dapat mempelajari pola yang lebih beragam dan menghasilkan prediksi yang lebih akurat. Dengan jumlah data yang lebih besar, model diharapkan mampu mengenali berbagai variasi komentar yang lebih luas, sehingga meningkatkan generalisasi terhadap data baru.

Dengan mempertimbangkan saran-saran ini, diharapkan penelitian selanjutnya dapat lebih mengoptimalkan proses klasifikasi sentimen dengan deteksi sarkasme serta meningkatkan keandalan model dalam menganalisis opini di media sosial.