

BAB V PENUTUP

5.1 Kesimpulan

Penelitian ini menganalisis sentimen pada komentar YouTube tentang Kabinet Merah Putih 2024 menggunakan algoritma SVM dan XGBoost. Proses preprocessing mencakup text cleaning, case folding, normalisasi, tokenizing, stopwords removal, dan stemming untuk memastikan data siap untuk pemodelan. Data kemudian dibagi menggunakan K-Fold Cross Validation dan Split Data 80:20, menghasilkan variasi hasil pada kedua metode tersebut. Evaluasi dengan metrik akurasi, precision, recall, dan F1-score menunjukkan bahwa SVM dengan kernel RBF memiliki performa terbaik dibandingkan model lainnya.

Berdasarkan hasil penelitian, dilakukan hyperparameter tuning menggunakan GridSearchCV untuk menemukan kombinasi parameter terbaik guna meningkatkan performa model. Pada SVM dengan kernel RBF, tuning dilakukan untuk menemukan kombinasi C dan γ yang optimal, dan hasilnya menunjukkan bahwa $C = 2$ dan $\gamma = 1$ memberikan akurasi tertinggi sebesar 88,02% pada K-Fold Cross Validation, serta 86,15% pada Split Data 80:20. Selain itu, model SVM dengan kernel Linear, Sigmoid, dan Polynomial juga dituning menggunakan GridSearchCV untuk memperoleh parameter terbaik masing-masing. Model XGBoost turut dioptimalkan dengan GridSearchCV, dengan hasil terbaik pada kombinasi $\text{max_depth} = 5$ dan $\text{n_estimators} = 150$, yang menghasilkan akurasi 84,74% pada K-Fold dan 84,23% pada Split Data 80:20.

Secara keseluruhan, algoritma SVM, terutama dengan kernel RBF, lebih unggul dalam menganalisis sentimen publik terhadap Kabinet Merah Putih 2024 dibandingkan dengan XGBoost, baik pada metode validasi K-Fold Cross Validation maupun Split Data 80:20. Oleh karena itu, SVM dengan kernel RBF dapat menjadi pilihan yang lebih efektif dalam aplikasi analisis sentimen pada komentar publik di media sosial, khususnya untuk topik-topik politik yang memiliki dinamika tinggi.

5.2 Saran

Dalam penelitian ini, meskipun penggunaan algoritma SVM dengan kernel RBF menunjukkan hasil terbaik, ada beberapa area yang dapat diperbaiki dan dikembangkan lebih lanjut. Pertama, penelitian ini hanya menggunakan dua algoritma, yaitu SVM dan XGBoost, sementara algoritma lain seperti Random Forest, Naive Bayes, atau deep learning (seperti BERT dan LSTM) dapat memberikan hasil yang lebih bervariasi dan mungkin lebih optimal untuk dataset ini. Penelitian selanjutnya dapat mengeksplorasi penerapan ensemble learning atau kombinasi model untuk meningkatkan akurasi dan generalisasi. Selain itu, teknik preprocessing yang diterapkan bisa ditingkatkan dengan menggunakan lemmatization, yang lebih terstruktur daripada stemming, serta mengembangkan fitur analisis sentimen berbasis aspek atau analisis emosi untuk mendapatkan wawasan yang lebih dalam mengenai opini publik.

Selanjutnya, eksperimen dengan rasio pembagian data yang berbeda (70:30 atau 90:10) dan penerapan 10-Fold Cross Validation dapat memberikan estimasi performa yang lebih kuat. Selain itu, penelitian ini juga bisa diperluas dengan mengumpulkan komentar dari lebih banyak video YouTube atau platform lain yang relevan, seperti Twitter atau Facebook, untuk mendapatkan data yang lebih representatif. Dengan langkah-langkah tersebut, penelitian di masa depan dapat memperoleh hasil yang lebih optimal dalam bidang analisis sentimen politik di media sosial.