

## **BAB V**

### **PENUTUP**

#### **5.1 Kesimpulan**

Berdasarkan beberapa penelitian yang dilakukan mengenai analisis sentimen dengan menggunakan algoritma machine learning, dapat disimpulkan bahwa algoritma seperti Random Forest, Support Vector Machine (SVM), dan XGBoost menunjukkan kinerja yang baik, namun dengan hasil yang berbeda tergantung pada konteks dan data yang digunakan. Penelitian oleh (Ramadhan et al., 2019) menggunakan data Twitter untuk menganalisis sentimen terhadap program acara di SCTV. Dalam penelitian tersebut, algoritma Naive Bayes dan Support Vector Machine (SVM) diterapkan. Hasilnya menunjukkan bahwa metode SVM memberikan akurasi terbaik, yaitu sebesar 88,57% untuk semua program acara. Sementara itu penelitian oleh (Daryanti & Widodo, 2024) menganalisis sentimen terhadap maskapai penerbangan Indonesia dengan data ulasan dari X (Twitter) dan TripAdvisor, yang mencakup 6.469 ulasan. Penelitian ini menggunakan algoritma Support Vector Machine (SVM), Naive Bayes, dan Random Forest. Hasilnya menunjukkan bahwa SVM dengan kernel RBF mencapai akurasi 89%, Naive Bayes memperoleh akurasi 79%, dan Random Forest memberikan akurasi tertinggi, yaitu 91%, sehingga dinyatakan sebagai metode paling efektif untuk analisis sentimen dalam konteks ini. Oleh karena itu, metode SVM dan random Forest memiliki potensi besar untuk diaplikasikan dalam menganalisis opini masyarakat terhadap tayangan televisi di platform Twitter. Dalam penelitian (Christanto et al., 2023) melakukan penelitian Analisis Perbandingan Decision Tree, Support Vector Machine, dan Xgboost dalam Mengklasifikasi Review Hotel Trip Advisor menguji tiga algoritma Machine Learning, yaitu Decision Tree Classifier, Support Vector Machine (SVM) dan Xgboost Classifier, dalam melakukan analisis sentimen terhadap review hotel di platform Trip advisor. Hasilnya menunjukkan bahwa Xgboost memiliki tingkat keakuratan (accuracy) yang paling tinggi, mencapai 99%, dibandingkan dengan Decision Tree

(97%) dan Support Vector Machine (98%). Dengan demikian, Xgboost dianggap sebagai algoritma terbaik untuk melakukan analisis sentimen pada review hotel di Trip advisor.

Dalam capstone project yang telah dikerjakan dengan judul Analisis Sentimen Menggunakan Metode Random Forest, SVM, dan XGBoost Pada Tayangan TV Berdasarkan Ulasan Masyarakat di Media Sosial Twitter, Hasil penelitian kami menunjukkan bahwa Algoritma Random Forest menunjukkan hasil accuracy (84%), SVM menunjukkan hasil (86%), dan XGBoost (86%), Oleh karena itu dapat disimpulkan bahwa Metode XGBoost merupakan algoritma yang paling baik. Hal ini disebabkan oleh kemampuannya yang stabil dan akurat dalam memproses data ulasan masyarakat, serta memberikan hasil yang lebih konsisten dibandingkan dengan Random Forest dan SVM. Oleh karena itu, XGBoost dapat dianggap sebagai pilihan algoritma yang optimal untuk analisis sentimen terhadap tayangan televisi berdasarkan ulasan di media sosial Twitter.

## **5.2 Saran**

Berdasarkan berbagai penelitian yang dianalisis, algoritma Random Forest, Support Vector Machine (SVM), dan XGBoost secara umum menunjukkan kinerja yang baik dalam penerapan analisis sentimen, dengan tingkat akurasi yang bervariasi tergantung pada karakteristik data dan konteks analisis yang digunakan. Setiap algoritma memiliki kelebihan masing-masing dalam menangani data teks untuk klasifikasi sentimen. SVM sering kali menonjol dalam mengolah data dengan dimensi tinggi dan bekerja dengan baik pada data yang terpisah secara linier, sedangkan Random Forest dikenal dengan pendekatannya yang berbasis ensemble yang dapat mengurangi risiko overfitting dengan melakukan pengambilan keputusan dari banyak pohon keputusan.

Dalam berbagai studi yang membandingkan ketiga algoritma tersebut, seperti yang terlihat pada penelitian yang menggunakan data ulasan di media sosial, algoritma XGBoost cenderung memberikan hasil

yang lebih unggul dalam hal akurasi dan kestabilan performa. Hal ini terlihat dari hasil capstone project yang dianalisis, di mana XGBoost menunjukkan akurasi tertinggi sebesar 86%, disusul oleh SVM dengan akurasi yang sama, namun dengan konsistensi yang sedikit lebih rendah, dan Random Forest dengan akurasi 84%.

Keunggulan XGBoost dapat dikaitkan dengan kemampuannya dalam menangani data yang kompleks melalui teknik boosting berbasis gradient, yang secara iteratif meningkatkan performa dengan meminimalkan kesalahan prediksi pada setiap tahap pelatihan. Selain itu, algoritma ini dikenal mampu mengatasi data dengan dimensi besar dan memberikan hasil yang lebih stabil dalam berbagai skenario pengujian.

Namun demikian, penting untuk dicatat bahwa pemilihan algoritma yang optimal tidak hanya bergantung pada tingkat akurasi yang dicapai, melainkan juga harus mempertimbangkan karakteristik dataset, kompleksitas data, ukuran dataset, serta kebutuhan spesifik dalam konteks analisis yang dilakukan. Dalam beberapa kasus, algoritma lain seperti SVM atau Random Forest dapat lebih sesuai, terutama pada data dengan dimensi yang lebih rendah atau saat interpretabilitas model menjadi prioritas utama.

Dengan demikian, meskipun XGBoost menunjukkan hasil yang paling unggul dalam studi yang dianalisis, evaluasi menyeluruh terhadap kebutuhan analisis dan karakteristik data sangat disarankan sebelum menentukan algoritma yang paling sesuai untuk digunakan dalam analisis sentimen di masa mendatang.