

ABSTRAK

Penelitian ini dilakukan untuk menganalisis sentimen publik terhadap Anies Baswedan di media sosial X, mengingat pentingnya pemahaman akan pandangan publik dalam konteks politik dan komunikasi massa. Dalam penelitian ini ada dua algoritma *machine learning* yaitu *Support Vector Machine* (SVM) dan *Random Forest* digunakan untuk membangun model klasifikasi sentimen. Data yang digunakan terdiri dari 1050 *tweet* yang dikumpulkan melalui proses *crawling*. Data ini kemudian dibagi menjadi 80% untuk pelatihan model dan 20% untuk pengujian. Setelah melakukan *crawling*, data akan melalui tahap *preprocessing* yang mencakup labeling manual, *cleaning*, *case folding*, *tokenizing*, *stopword removal* dan *stemming* diikuti dengan *matrix generation* menggunakan metode *TF-IDF*. Selanjutnya algoritma *Support Vector Machine* (SVM) dan *Random Forest* diterapkan untuk membangun model klasifikasi sentimen. Hasil penelitian menunjukkan bahwa *Random Forest* unggul dibandingkan *Support Vector Machine* (SVM) dengan mencatatkan rata-rata akurasi sebesar 99.68%, presisi 99.47%, dan *recall* 99.56%, sedangkan *Support Vector Machine* (SVM) mencatatkan rata-rata akurasi 99.36%, presisi 98.95%, dan *recall* 99.09%. Evaluasi model dilakukan menggunakan *K-Fold Cross Validation* yang mengindikasikan bahwa *Random Forest* memiliki performa yang lebih stabil dan konsisten. Kesimpulan dari penelitian ini menunjukkan bahwa metode *Random Forest* lebih efektif dan andal untuk digunakan dalam klasifikasi sentimen dibandingkan dengan *Support Vector Machine* (SVM) terutama dalam hal keandalan dan stabilitas performa.

Kata Kunci: Analisis Sentimen, Anies Baswedan, Media Sosial X, *Support Vector Machine*, *Random Forest*.

ABSTRACT

This study was conducted to analyze public sentiment towards Anies Baswedan on the social media platform X, given the importance of understanding public opinion in the context of politics and mass communication. In this research, two machine learning algorithms, Support Vector Machine (SVM) and Random Forest, were used to build sentiment classification models. The data used consisted of 1050 tweets collected through a crawling process. This data was then split into 80% for model training and 20% for testing. After data collection, the data underwent preprocessing stages, including manual labeling, cleaning, case folding, tokenizing, stopword removal, and stemming, followed by matrix generation using the TF-IDF method. Subsequently, the Support Vector Machine (SVM) and Random Forest algorithms were applied to build the sentiment classification models. The results showed that Random Forest outperformed the Support Vector Machine (SVM), achieving an average accuracy of 99.68%, precision of 99.47%, and recall of 99.56%, while the Support Vector Machine (SVM) recorded an average accuracy of 99.36%, precision of 98.95%, and recall of 99.09%. Model evaluation was conducted using K-Fold Cross Validation, indicating that Random Forest demonstrated more stable and consistent performance. The conclusion of this study indicates that the Random Forest method is more effective and reliable for sentiment classification compared to the Support Vector Machine (SVM), particularly in terms of reliability and performance stability.

Keywords: *Sentiment Analysis, Anies Baswedan, Social Media X, Support Vector Machine, Random Forest.*