

Analisis Sentimen Pemilihan Presiden Indonesia 2024 dengan *Ensemble Learning* untuk Mengurangi Disinformasi di Media Sosial

Jihan Nabilah Hakim¹, Dr. Yuliant Sibaroni, S.Si., M.T.², Sri Suryani Prasetyowati, S.Si., M.Si.³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹nabilahakim@students.telkomuniversity.ac.id, ²yuliant@telkomuniversity.ac.id,

³srisuryani@telkomuniversity.ac.id

Abstrak

Disinformasi di media sosial merupakan masalah yang sering terjadi, terutama dalam konteks politik terkait Pemilihan Presiden 2024. Tujuan dari penelitian analisis sentimen ini adalah untuk mengurangi disinformasi melalui penerapan metode Ensemble Learning dengan menggunakan 4 classifier yaitu SVM, Random Forest, Gradient Boosting, dan Stacked Ensemble, dengan memanfaatkan TF-IDF dan cross validation. Hasil dari penelitian tersebut menunjukkan bahwa SVM, Random Forest, Gradient Boosting Classifier, dan Stacked Ensemble mampu mengklasifikasikan disinformasi dengan akurasi yang signifikan. SVM mencapai akurasi 86.44%, Random Forest mencapai akurasi 87.65%, dan Gradient Boosting Classifier mencapai akurasi 88.20%. Stacked Ensemble mencapai akurasi yang sama dengan Gradient Boosting Classifier. Ensemble Learning melalui Stacked Ensemble terbukti memberikan hasil yang menjanjikan dalam upaya mengurangi disinformasi di media sosial. Hasil penelitian ini memberikan kontribusi penting dalam pengembangan penggunaan analisis sentimen untuk mengatasi masalah disinformasi di media sosial.

Kata kunci: analisis sentimen, ensemble learning, SVM, random forest, gradient boosting classifier, stacked ensemble

Abstract

Disinformation on social media is a frequent problem, especially in the political context related to the 2024 Presidential Election. The purpose of this sentiment analysis research is to reduce disinformation through the application of the Ensemble Learning method using 4 classifiers namely SVM, Random Forest, Gradient Boosting, and Stacked Ensemble, by utilizing TF-IDF and cross validation. The results of the research show that SVM, Random Forest, Gradient Boosting Classifier, and Stacked Ensemble are able to classify disinformation with significant accuracy. SVM achieved 86.44% accuracy, Random Forest achieved 87.65% accuracy, and Gradient Boosting Classifier achieved 88.20% accuracy. Stacked Ensemble achieves the same accuracy as GBC. Ensemble Learning through stacked ensemble has proven to provide promising results in efforts to reduce disinformation on social media. The results of this research make an important contribution in developing the use of sentiment analysis to address the problem of disinformation in social media

Keywords: sentiment analysis, ensemble learning, SVM, random forest, gradient boosting classifier, stacked ensemble
