

ABSTRAK

Metode *Naive Bayes* memiliki performansi yang baik khususnya dalam penerapannya untuk mengklasifikasikan sentimen. Dataset untuk analisa sentimen pengguna Twitter terhadap ketua umum parpol di Indonesia memiliki ketidakseimbangan data yang dapat mengakibatkan berkurangnya performa sistem dalam melakukan klasifikasi. Pada penelitian ini digunakan Algoritma *Adaptive Boosting* dengan tujuan meningkatkan kinerja Algoritma *Naive Bayes* pada data tak seimbang. Penelitian dilakukan dengan memanfaatkan modul *Twitterscraper* pada Python untuk pengumpulan data dan aplikasi Rapidminer untuk melakukan proses klasifikasi. Dataset akan diklasifikasi dengan menggunakan Algoritma *Naive Bayes* dengan 3 rasio perbandingan data yakni 7:3, 8:2 serta 9:1, hasil akurasi terendah dan tertinggi pada masing-masing dataset akan ditambahkan dengan Algoritma *Adaptive Boosting*. Hasil penelitian menunjukkan bahwa setelah dilakukan penambahan Algoritma *Adaptive Boosting* pada analisis sentimen menggunakan Algoritma *Naive Bayes* pada data tak seimbang terjadi peningkatan pada performa sistem dibandingkan dengan sebelum ditambahkan Algoritma *Adaptive Boosting*. Pada dataset dengan tingkat perbedaan data tertinggi yakni dataset Muhaimin Iskandar dengan rasio data 9:1, rata-rata peningkatan performa sistem mencapai 8,70% dengan rata-rata akurasi, presisi dan *recall* sebelum ditambahkan Algoritma *Adaptive Boosting* sebesar 74,61% AUC 0,680 dibandingkan dengan rata-rata akurasi, presisi dan *recall* setelah ditambahkan Algoritma *Adaptive Boosting* sebesar 79,17% AUC 0,789. Hasil ini juga merupakan rata-rata peningkatan performa tertinggi dari keseluruhan hasil pengujian.

Kata Kunci : analisa sentimen *twitter*, *naive bayes*, *adaptive boosting*, ketua umum parpol, ketidakseimbangan data

ABSTRACT

The Naive Bayes method has good performance in its implementation to classify sentiments. The dataset for analyzing twitter users' sentiments towards the chairman of political parties in Indonesia has an imbalance of data and it can reduce the system performance in conducting classification. In this study, the Adaptive Boosting Algorithm is used to increase the performance of the Naive Bayes Algorithm on unbalanced data. The study was conducted by utilizing the Twitterscraper module in Python for data collection and the Rapidminer application to carry out the classification process. The dataset will be classified using the Naive Bayes Algorithm with 3 data comparison ratios namely 7: 3, 8: 2 and 9: 1. The lowest and highest accuracy results in each dataset will be added with Algoritma Adaptive Boosting. The results shows that there is increase in system performance after the addition of the Adaptive Boosting Algorithm in sentiment analysis using the Naive Bayes Algorithm on unbalanced dat. In the dataset with the highest level of data difference of Muhaimin Iskandar dataset with a data ratio of 9: 1, the average increase in system performance reached 8.70% with an average accuracy, precision and recall before adding the Adaptive Boosting Algorithm of 74.61% AUC 0.680 compared to the average accuracy, precision and recall after adding the Adaptive Boosting Algorithm of 79.17% AUC 0.789. This result is also the highest average increase in overall performance of the test results.

Keywords: *Twitter sentiment analysis, Naive Bayes, Adaptive Boosting, Chairman of political parties, data imbalance.*

