

PERBANDINGAN ALGORITMA NAÏVE BAYES, SVM DAN RANDOM FOREST PADA ANALISIS SENTIMEN PENGUNA X TERHADAP IKN

ABSTRAK

Abstract: Sentiment analysis was explored to understand social media users' opinions towards the Indonesian Capital City (IKN) through the X platform with machine learning and lexicon-based algorithms. This research uses three algorithms: Naïve Bayes, Support Vector Machine (SVM), and Random Forest. The aim of this research is to test and compare the performance of the three algorithms to determine the best in classifying sentiment data from the X platform. The data consists of 10,000 tweets collected using the crawling method with the Python Harvest Library and Node.js, using keywords related to IKN. Based on the algorithm performance test, it was concluded that SVM had the highest performance compared to Naïve Bayes and Random Forest, producing an accuracy of 87%, precision 87%, recall 87%, and f-1 score 87%. This research uses the CRISP-DM Data Mining framework to ensure a structured and systematic approach to the analysis process.

Keywords: sentiment analysis, X, naïve bayes, support vector machine (SVM), random forest, IKN.

Abstrak: Analisis sentimen diesplorasi untuk memahami opini pengguna media sosial terhadap Ibu Kota Nusantara (IKN) melalui platform X dengan algoritma machine learning dan berbasis lexicon-based. Penelitian ini menggunakan tiga algoritma: Naïve Bayes, Support Vector Machine (SVM), dan Random Forest. Tujuan penelitian ini adalah menguji dan membandingkan performa ketiga algoritma tersebut untuk menentukan yang terbaik dalam pengklasifikasian data sentimen dari platform X. Data terdiri dari 10.000 tweet yang dikumpulkan menggunakan metode crawling dengan Library Python Harvest dan Node.js, menggunakan kata kunci terkait IKN. Berdasarkan uji performa algoritma, disimpulkan bahwa SVM memiliki performa tertinggi dibandingkan Naïve Bayes dan Random Forest, dengan menghasilkan accuracy 87%, precision 87%, recall 87%, dan f-1 score 87%. Penelitian ini menggunakan framework Data Mining CRISP-DM untuk memastikan pendekatan terstruktur dan sistematis dalam proses analisis.

Kata Kunci: analisis sentimen, X, naïve bayes, support vector machine (SVM), random forest, IKN.