

KOMPARASI ALGORITMA *NAÏVE BAYES* DAN *SUPPORT VECTOR MACHINE* (SVM) PADA ANALISIS SENTIMEN IMPLEMENTASI PROGRAM KARTU INDONESIA PINTAR KULIAH

Oleh

Ni Putu Ana Rainita, NIM 2115091075

Jurusan Teknik Informatika

Program Studi Sistem Informasi

ABSTRAK

Pemilihan model klasifikasi yang tepat sangat penting dalam analisis sentimen berbasis teks, terutama ketika distribusi data tidak seimbang. Kondisi ini dapat memengaruhi akurasi dan performa model dalam mengklasifikasikan sentimen. *Naive Bayes* dan *Support Vector Machine* (SVM) merupakan dua algoritma yang sering digunakan dalam klasifikasi teks, namun perbandingan performa keduanya pada data tidak seimbang masih perlu dikaji lebih lanjut. Penelitian ini bertujuan untuk membandingkan performa kedua algoritma dalam mengklasifikasikan sentimen masyarakat terhadap implementasi Program Kartu Indonesia Pintar (KIP) Kuliah. Data yang digunakan dalam penelitian ini terdiri dari 1.644 tweet yang dikumpulkan dari media sosial, dengan distribusi sentimen yang didominasi oleh opini negatif. Untuk mengatasi ketidakseimbangan distribusi kelas pada data, digunakan metode *Synthetic Minority Oversampling Technique* (SMOTE) guna memperbanyak data pada kelas minoritas. Hasil evaluasi menunjukkan bahwa sebelum penerapan SMOTE, algoritma SVM memiliki akurasi sebesar 92%, presisi 91%, dan *recall* 77%. Sementara itu, algoritma *Naive Bayes* mencatatkan akurasi 79%, presisi 68%, dan *recall* 78%. Setelah dilakukan *oversampling* dengan SMOTE, performa kedua model meningkat signifikan. SVM mencapai nilai akurasi, presisi, dan *recall* sebesar 99%, sedangkan *Naive Bayes* memperoleh 95% pada ketiga metrik tersebut. Meskipun SVM menunjukkan akurasi lebih tinggi, hasil penelitian ini menunjukkan bahwa *Naive Bayes* memiliki performa yang lebih stabil baik sebelum maupun sesudah proses penyeimbangan data dilakukan.

Kata kunci: KIP-K, Analisis Sentimen, *Naive Bayes*, SVM, SMOTE

**COMPARISON OF NAÏVE BAYES AND SUPPORT VECTOR MACHINE
(SVM) ALGORITHMS IN SENTIMENT ANALYSIS OF IMPLEMENTATION
KARTU INDONESIA PINTAR KULIAH PROGRAM**

Bye:

Ni Putu Ana Rainita, NIM 2115091075

Informatics Engineering

Information Systems Study Program

ABSTRACT

The selection of an appropriate classification model is crucial in text-based sentiment analysis, especially when the data distribution is imbalanced. This condition can affect the accuracy and performance of the model in classifying sentiment. Naive Bayes and Support Vector Machine (SVM) are two algorithms commonly used in text classification, but their performance on imbalanced data still requires further investigation. This study aims to compare the performance of both algorithms in classifying public sentiment toward the implementation of the Kartu Indonesia Pintar (KIP) Kuliah Program. The dataset used in this study consists of 1,644 tweets collected from social media, with sentiment distribution dominated by negative opinions. To address the class imbalance issue, the Synthetic Minority Oversampling Technique (SMOTE) was applied to increase the number of data in the minority class. Evaluation results show that before applying SMOTE, the SVM algorithm achieved an accuracy of 92%, precision of 91%, and recall of 77%. Meanwhile, the Naive Bayes algorithm recorded an accuracy of 79%, precision of 68%, and recall of 78%. After applying SMOTE, the performance of both models improved significantly. SVM reached 99% in accuracy, precision, and recall, while Naive Bayes achieved 95% across all three metrics. Although SVM demonstrated higher accuracy, the results of this study indicate that Naive Bayes showed more stable performance both before and after the data balancing process.

Keywords: KIP-K, Sentimen Analysis, Naïve Bayes, SVM, SMOTE