

ABSTRAK

PERBANDINGAN METODE *K-NEAREST NEIGHBORS* DAN *SUPPORT VECTOR MACHINE* DALAM ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP LAYANAN PINJAMAN ONLINE

Oleh

MUTIARA WIDDI AYUNINTYAS

Dalam penelitian ini, dilakukan analisis sentimen terhadap opini masyarakat di media sosial Twitter mengenai layanan pinjaman *online* yang ada dan membagi dataset opini ke dalam dua kelas yaitu sentimen positif dan sentimen negatif. Penggunaan layanan pinjaman *online* oleh masyarakat Indonesia beberapa tahun terakhir ini masih menjadi pilihan yang populer dan masih ramai diperbincangkan. Hal tersebut seharusnya menjadi perhatian untuk mengetahui apa yang terjadi dalam masyarakat. Pengklasifikasian dalam penelitian ini dilakukan dengan menggunakan metode *K-Nearest Neighbor* (KNN) dan *Support Vector Machine* (SVM). Selanjutnya, dilakukan juga skema penyeimbangan data pada data *unbalanced* menggunakan metode *oversampling* dan *undersampling* data untuk mendapatkan metode dengan performa terbaik. Hasil klasifikasi menggunakan metode-metode tersebut menghasilkan kinerja yang lebih baik menggunakan algoritma *Support Vector Machine* (SVM) pada dataset *undersampling* dengan nilai *f1-score* sebesar 47,41%. Sementara pada algoritma *K-Nearest Neighbors* (KNN), dihasilkan kinerja yang terbaik dengan nilai *f1-score* sebesar 43,60% pada dataset *undersampling*. Jadi, perbandingan dari hasil kinerja kedua metode tersebut menunjukkan bahwa algoritma *Support Vector Machine* (SVM) menghasilkan performa yang lebih baik dalam melakukan analisis sentimen masyarakat terhadap layanan pinjaman online dibandingkan dengan algoritma *K-Nearest Neighbors* (KNN).

Kata kunci: Analisis Sentimen, *K-Nearest Neighbors*, *Support Vector Machine*, Pinjaman Online, Twitter

ABSTRACT

COMPARISON OF THE K-NEAREST NEIGHBORS AND SUPPORT VECTOR MACHINE METHODS IN THE ANALYSIS OF TWITTER USER SENTIMENT FOR ONLINE LOAN SERVICES

By

MUTIARA WIDDI AYUNINTYAS

In this study, sentiment analysis was carried out on public opinion on social media Twitter regarding existing online loan services and dividing the opinion dataset into two classes, namely positive sentiment and negative sentiment. The use of online loan services by the people of Indonesia in recent years is still a popular choice and is still being widely discussed. This should be a concern to know what is happening in society. Classification in this study was carried out using the K-Nearest Neighbors (KNN) and Support Vector Machine (SVM) methods. Furthermore, a data balancing scheme is also carried out on unbalanced data using the oversampling and undersampling data methods to get the method with the best performance. The classification results using these methods produce better performance using the Support Vector Machine (SVM) algorithm on the undersampled dataset with an f1-score value of 47,41%. Meanwhile, in the K-Nearest Neighbors (KNN) algorithm, the best performance is produced with an f1-score of 43,60% on the undersampled dataset. So, a comparison of the performance results of those two methods shows that the Support Vector Machine (SVM) algorithm produces better performance in conducting an analysis of public sentiment about online loan services compared to the K-Nearest Neighbors (KNN) algorithm.

Keywords: Sentiment Analysis, K-Nearest Neighbors, Support Vector Machine, Online Loans, Twitter