

ABSTRAK

Di Indonesia, ada beberapa perusahaan penyedia layanan internet, salah satunya adalah PT Telekomunikasi Seluler yang mengoperasikan layanan jaringan internet seluler melalui Telkomsel. Dikarenakan banyak yang menggunakan operator seluler Telkomsel maka umumnya pengguna dari operator tersebut akan menyampaikan opini mereka terhadap kualitas layanan jaringan internet yang diberikan. Opini tersebut dapat merepresentasikan tentang kepuasan pengguna pada layanan jasa tersebut. Kepuasan pengguna seringkali digunakan oleh sebuah perusahaan sebagai bahan penilaian dan evaluasi dari layanan mereka. Untuk mengukur kepuasan pengguna tersebut, salah satu cara yang dapat diterapkan yaitu analisis sentimen dengan cara melakukan klasifikasi terhadap opini pengguna dan mengelompokkan kedalam kelas positif, negatif, atau netral. Dalam penelitian analisis sentimen ini menggunakan model algoritma *machine learning* yaitu *K-Nearest Neighbour*, *Support Vector Machine*, dan *Ensemble KNN-SVM* yang berbasis *majority vote* dan berbasis *average*. Dalam penelitian ini data yang diambil berasal dari Twitter dengan rentang waktu 7 Juli 2020 hingga 31 Desember 2022 dengan total jumlah data sebesar 30004 data dan diambil sampel yang diberi label sebesar 3900 data. Dari penggunaan data sampel tersebut, nilai akurasi pada model KNN pada $K=15$ memberikan hasil akurasi sebesar 83.21%, model SVM pada $C=100$ memberikan hasil akurasi sebesar 84.33%, model *Ensemble KNN-SVM Majority Vote* atau *Hard Vote* memberikan hasil akurasi sebesar 83.26%, dan model *Ensemble KNN-SVM Average* atau *Soft Vote* memberikan hasil akurasi sebesar 84.79%. Selain itu keempat model tersebut melakukan prediksi sentimen terhadap data yang belum dilabel dan keempat model tersebut memprediksi mayoritas sentimennya yaitu negatif. Sehingga dapat disimpulkan bahwa opini masyarakat terhadap kualitas layanan jaringan internet telkomsel adalah negatif. Secara keseluruhan, penggunaan model klasifikasi KNN, SVM, dan *Ensemble KNN-SVM* dalam melakukan analisis sentimen dapat dikatakan baik dan mampu untuk memprediksi sentimen pada sebuah data yang belum berlabel dan yang berlabel.

Kata Kunci: *K-Nearest Neighbour*, *Support Vector Machine*, Analisis Sentimen, *Ensemble Learning*, Jaringan Internet, Telkomsel

ABSTRACT

In Indonesia, there are several internet service provider companies, one of which is PT Telekomunikasi Cellular which operates cellular internet network services through Telkomsel. Because many use Telkomsel cellular operators, generally users of these operators will express their opinions on the quality of internet network services provided. The opinion can represent user satisfaction with the service. User satisfaction is often used by a company as an assessment and evaluation of their services. To measure user satisfaction, one way that can be applied is sentiment analysis by classifying user opinions and grouping them into positive, negative, or neutral classes. This sentiment analysis research uses machine learning algorithm models, namely K-Nearest Neighbor, Support Vector Machine, and KNN-SVM Ensemble which are majority vote-based and average-based. In this study, the data taken came from Twitter with a period of July 7, 2020, to December 31, 2022, with a total amount of data of 30004 data and a labeled sample of 3900 data was taken. From the use of the sample data, the accuracy value of the KNN model at $K = 15$ gave an accuracy result of 83.21%, the SVM model at $C = 100$ gave an accuracy result of 84.33%, the KNN-SVM Majority Vote or Hard Vote Ensemble model gave an accuracy result of 83.26%, and the KNN-SVM Average or Soft Vote Ensemble model gave an accuracy result of 84.79%. In addition, the four models predict sentiment against unlabeled data and all four models predict the majority of sentiment is negative. So it can be concluded that public opinion on the quality of Telkomsel's internet network services is negative. Overall, the use of KNN, SVM, and Ensemble KNN-SVM classification models in conducting sentiment analysis can be said to be good and able to predict sentiment on unlabeled and labeled data.

Keywords: *K-Nearest Neighbor, Support Vector Machine, Sentiment Analysis, Ensemble Learning, Internet Network, Telkomsel*