ABSTRAK

YouTube adalah platform berbagi konten media yang memfasilitasi pengguna untuk menonton, menyimpan, dan menyebarkan beragam jenis video. Sejumlah tanggapan dipengaruhi oleh informasi yang disampaikan oleh penyebar konten dan mencerminkan apakah isi dari video tersebut dianggap positif atau negatif. Penerapan data mining dalam analisis sentimen menjadi pendekatan yang signifikan untuk memahami makna dari setiap komentar. Pengklasifikasian komentar sebagai positif atau negatif sangat berguna bagi pengguna YouTube. Klasifikasi ini bertujuan untuk melakukan analisis sentimen pengguna terhadap informasi dari video youtube melalui komentar dengan menggunakan metode *hybrid Naïve Bayes - Support Vector Machine*. Dataset yang dianalisis berasal dari kanal Youtube CNN Indonesia dengan 5 video bertema "The Politician" yang dipilih berdasarkan jumlah komentar teratas. Metode klasifikasi *Multinominal Naïve Bayes* digunakan untuk mencari probabilitas dari setiap kata dengan menggunakan metode *likelihood ratio*. Metode klasifikasi *Support Vector Machine* digunakan untuk memprediksi sentimen dengan meminimalkan *loss function*. Hasil penelitian dari 5 dataset menunjukan rata-rata nilai nilai *Precision* sebesar 79.8%, *Recall* sebesar 92.6%, *F1 Score* sebesar 85.2%, dan Akurasi sebesar 80.4%.

Kata kunci: komentar youtube, analisis sentimen, Naïve Bayes Support Vector Machine.

ABSTRACT

YouTube is a media content-sharing platform that facilitates users to watch, save, and distribute various types of videos. The perception of whether the content of the videos is considered positive or negative is influenced by the information conveyed by the content creators. The application of data mining in sentiment analysis has become a significant approach to understanding the meaning behind each comment. Classifying comments as positive or negative is highly beneficial for YouTube users. The classification in this study aims to analyze the sentiment of YouTube users towards the information in shared videos through the comment section, using a hybrid Naïve Bayes - Support Vector Machine method. The analyzed dataset consists of 5 videos with the theme "The Politician" from the CNN Indonesia YouTube channel, selected based on the highest number of comments. The Multinomial Naïve Bayes classification method is used to determine the probability of each word using the likelihood ratio method. The Support Vector Machine classification method is employed to predict sentiment by minimizing the loss function. The research results from the 5 datasets show an average Precision value of 79.8%, Recall of 92.6%, F1 Score of 85.2%, and Accuracy of 80.4%.

***Keywords**: YouTube comments, sentiment analysis, Naïve Bayes Support Vector Machine.*