

ABSTRAK

Perkembangan dan pertumbuhan media sosial ditandai dengan peningkatan jumlah pengguna, khususnya pengguna Twitter. Twitter merupakan *platform* media sosial di mana pengguna dapat menulis pendapat, mengakses berita terbaru, dan menyampaikan kritik serta mengizinkan pengguna lain untuk berkomentar. Komentar pada twitter disebut tweet. Pengguna dapat saling terhubung untuk membicarakan tentang yang sedang menjadi topik pembicaraan utama. Semakin banyak pengguna, semakin beragam pula komentar yang ditulis seseorang dalam sebuah *postingan*. Tanpa disadari, beberapa pesan dan komentar mengandung unsur *bullying* atau kekerasan yang dapat berdampak negatif pada diri sendiri dan orang lain. Oleh karena itu, pesan dan komentar di Twitter harus dianalisis menggunakan analisis sentimen. Dalam penelitian ini membahas analisis sentimen menggunakan algoritma *Naive Bayes* dan *K-Nearest Neighbor*. Informasi yang diperoleh dibagi menjadi dua kategori, yaitu kategori positif dan negatif. Selain itu, data melewati fase pra-pemrosesan, yaitu *cleaning*, *casefolding*, *tokenizing*, *filtering stopwords* dan TF-IDF. Pada penelitian ini data dibagi menjadi 60% data latih dan 40% data uji. Hasil akurasi menggunakan metode *confusion matrix* dengan algoritma *Naive Bayes* mencapai akurasi 81%, sedangkan algoritma *K-Nearest Neighbor* mencapai akurasi 77%.

Kata Kunci: Analisis Sentimen, *Bullying*, Twitter, *Naive Bayes*, *K-Nearest Neighbor*.



ABSTRACT

The development and growth of social media is marked by an increase in the number of users, especially Twitter users. Twitter is a social media platform where users can write opinions, access the latest news, and provide criticism and allow other users to comment. Comments on Twitter are called tweets. Users can connect with each other to talk about what is currently the main topic of conversation. The more users, the more diverse comments someone writes in a post. Without realizing it, some messages and comments contain elements of bullying or violence which can have a negative impact on oneself and others. Therefore, messages and comments on Twitter must be analyzed using sentiment analysis. This study discusses sentiment analysis using the Naive Bayes and K-Nearest Neighbor algorithms. The information obtained is divided into two categories, namely positive and negative categories. In addition, the data goes through pre-processing phases, namely cleaning, case folding, tokenizing, filtering stopwords and TF-IDF. In this study, the data was divided into 60% training data and 40% test data. The accuracy results using the confusion matrix method with the Naive Bayes algorithm reach 81% accuracy, while the K-Nearest Neighbor algorithm achieves 77% accuracy.

Keywords: Sentiment Analysis, Bullying, Twitter, Naïve Bayes, K-Nearest Neighbor.

