

ABSTRAK

Penelitian ini bertujuan mendeteksi indikasi cyberbullying pada komentar pengguna Instagram menggunakan metode Support Vector Machine (SVM). Media sosial Instagram yang banyak digunakan remaja memungkinkan munculnya berbagai bentuk interaksi, termasuk komentar bermuatan negatif yang berpotensi menjadi perundungan daring. Oleh karena itu, diperlukan suatu pendekatan otomatis untuk mengidentifikasi sentimen komentar secara efisien. Data penelitian berupa 1.000 komentar yang dikumpulkan dari Instagram dengan komposisi 537 komentar positif dan 353 komentar negatif. Proses pengolahan data diawali dengan tahap preprocessing teks yang meliputi case folding, tokenizing, stopword removal, dan stemming untuk membersihkan dan menormalkan teks. Selanjutnya dilakukan pembobotan kata menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) guna mengubah teks menjadi representasi numerik yang dapat diproses oleh model klasifikasi. Tahap klasifikasi dilakukan menggunakan algoritma Support Vector Machine untuk memisahkan komentar ke dalam kategori sentimen. Hasil pengujian menunjukkan model memperoleh nilai akurasi sebesar 89,89%, precision 84,40%, recall 98,92%, f1-score 91,08%, dan specificity 80,00%. Nilai evaluasi tersebut menunjukkan bahwa model memiliki kemampuan yang baik dalam mengidentifikasi komentar bermuatan negatif. Dengan demikian, metode SVM dapat digunakan sebagai pendekatan awal dalam membantu proses deteksi cyberbullying pada media sosial, sehingga berpotensi mendukung upaya pemantauan serta pencegahan perundungan daring secara lebih cepat dan sistematis.

Kata kunci: Analisis sentimen, *Cyberbullying*, Instagram, TF-IDF, *Support Vector Machine*

ABSTRACT

This study aims to detect indications of cyberbullying in Instagram user comments using the Support Vector Machine (SVM) method. Instagram, which is widely used by adolescents, enables various forms of interaction, including negative comments that may lead to online harassment. Therefore, an automated approach is required to efficiently identify sentiment within comments. The research dataset consists of 1,000 comments collected from Instagram, comprising 537 positive comments and 353 negative comments. The data processing stage begins with text preprocessing, including case folding, tokenizing, stopword removal, and stemming to clean and normalize the text. Subsequently, word weighting is performed using the Term Frequency–Inverse Document Frequency (TF-IDF) method to transform the text into numerical representation suitable for classification. The classification stage is carried out using the Support Vector Machine algorithm to categorize comments into sentiment classes. The evaluation results show that the model achieved an accuracy of 89.89%, precision of 84.40%, recall of 98.92%, f1-score of 91.08%, and specificity of 80.00%. These performance metrics indicate that the model has good capability in identifying negative comments. Therefore, the SVM method can be utilized as an initial approach to support the detection of cyberbullying on social media, potentially assisting monitoring and prevention efforts in a faster and more systematic manner.

Keywords: *Sentiment Analysis, Cyberbullying, Instagram, TF-IDF, Support Vector Machine*