

INTISARI

X atau Twitter adalah salah satu sosial media platform interaksi online yang populer. Namun, pengguna juga dapat mempengaruhi hal buruk seseorang untuk melakukan bullying. Data menunjukkan bahwa lebih dari 60% anak-anak dan 40% orang dewasa menjadi target cyberbullying, dengan tren yang diperkirakan akan meningkat. Menurut Cloudwars, 24% korban melaporkan pelecehan diskriminatif di X. Fenomena ini menjadi hal yang penting untuk dikaji bersamaan dengan pengolahan teks, mengingat dampak psikologis yang ditimbulkannya pada korban. Oleh karena itu, penelitian ini bertujuan untuk menerapkan algoritma Support Vector Machine yang mampu mengidentifikasi tweet (komentar) pada sosial media X yang mengandung pesan cyberbullying. Serta mengetahui kinerja metode ADASYN dalam menangani permasalahan kesenjangan kelas. Penelitian menggunakan algoritma Support Vector Machine, dataset terdiri dari 945 komentar yang dilabeli secara manual, didukung oleh ekstraksi fitur TF-IDF dan ADASYN. Hasil menunjukkan bahwa model tanpa ADASYN mencapai akurasi 74.71% dengan F1-score 69.92% menggunakan komposisi data 80:20. Setelah penerapan ADASYN, performa model meningkat dengan akurasi 77.25% dan F1-score 77.18. Peningkatan juga terjadi pada skenario data 70:30, sebelum menggunakan ADASYN mencapai akurasi 74.42% dengan F1-score 70.26, setelah penerapan ADASYN menjadi akurasi 76.26% dan F1-score 76.19%. Peningkatan pada semua metrik dan skenario menunjukkan bahwa ADASYN adalah teknik yang sangat diperlukan untuk tugas klasifikasi, terutama ketika menangani kumpulan data yang tidak seimbang.

Kata kunci: Cyberbullying, SVM, X, ADASYN

ABSTRACT

X or twitter is one of the most popular social media platforms for online interaction. However, users can also influence others to bully. Data shows that more than 60% of children and 40% of adults are targets of cyberbullying, with the trend expected to increase. According to Cloudwars, 24% of victims reported discriminatory harassment on X. This phenomenon is important to study in conjunction with text processing, given the psychological impact it has on victims. Therefore, this research aims to build a system that is able to identify tweets (comments) on X social media that contain cyberbullying messages. As well as knowing the performance of the ADASYN method in dealing with class gap problems. The research uses the Support Vector Machine algorithm, the dataset consists of 945 manually labeled comments, supported by feature extraction TF-IDF and ADASYN. Results showed that the model without ADASYN achieved 74.71% accuracy with an F1-score of 69.92% using 80:20 data composition. After the application of ADASYN, the model performance improved with an accuracy of 77.25% and F1-score of 77.18. The improvement also occurred in the 70:30 data scenario, before using ADASYN, it reached an accuracy of 74.42% with an F1-score of 70.26, after applying ADASYN it became an accuracy of 76.26% and an F1-score of 76.19%. The improvement in all metrics and scenarios shows that ADASYN is an indispensable technique for classification tasks, especially when dealing with imbalanced data sets.

Keywords: Cyberbullying, SVM, X, ADASYN