

ABSTRAK

Media sosial adalah sebuah sarana komunikasi berbasis daring yang dapat mempermudah penggunaanya dalam berinteraksi antar sesama pengguna tanpa adanya batasan wilayah dan waktu. Indonesia yang memiliki angka pengguna sosial media tertinggi di dunia. Platform sosial media Twitter merupakan salah satu tempat dimana para penggunaanya dapat mencurahkan seluruh isi hati dalam bentuk cuitan. Interaksi yang bebas serta beragam pada Twitter memiliki pengaruh yang cukup besar dalam kondisi psikologis penggunaanya. Cyberbullying atau perundungan secara daring adalah suatu tindakan penghinaan atau menyakiti perasaan orang lain secara sengaja dan berulang-ulang pada sosial media, pesan, atau dengan cara lain. Pada penelitian ini menggunakan metode klasifikasi RoBERTa untuk mendeteksi tweet yang mengandung cyberbullying dengan skor terbaik accuracy 86.9% dan f1-score 77.5%.

Kata kunci : media sosial, twitter, cyberbullying, RoBERTa.