

Abstrak

Media sosial adalah salah satu alat komunikasi yang mendukung pengguna untuk berinteraksi secara sosial dengan menggunakan teknologi. Salah satu platform media sosial yang populer adalah Twitter. Namun, Twitter telah dianggap oleh *virtual police* sebagai salah satu sumber utama penyebaran ujaran kebencian pada media sosial. Dalam penelitian tugas akhir ini, penulis melakukan studi tentang deteksi ujaran kebencian pada *tweet* yang ada pada Twitter Indonesia. Metode yang digunakan dalam penelitian ini adalah klasifikasi multi-label dengan menerapkan metode LSTM dan BiLSTM. Dataset yang digunakan sebanyak 13.169 data *tweet* dan terdapat 12 label kelas ujaran kebencian dari dataset. Hasil penelitian menunjukkan bahwa metode LSTM dan BiLSTM memiliki kinerja yang baik dalam mengklasifikasikan data teks dengan uji coba sebanyak 10 kali dengan rata-rata nilai akurasi seluruh kelas pada LSTM sebesar 78,67% dan BiLSTM sebesar 80,25%. Berdasarkan akurasi yang didapatkan, BiLSTM memiliki akurasi lebih tinggi dibandingkan dengan LSTM sehingga dapat disimpulkan bahwa BiLSTM lebih unggul dibandingkan LSTM dalam mengidentifikasi ujaran kebencian pada media sosial.

Kata kunci: twitter, ujaran kebencian, media sosial, LSTM, BiLSTM