

ABSTRAK

Ujaran kebencian adalah ujaran yang mengandung kebencian, menyerang, dan berapi-api yang dimaksudkan untuk menginspirasi orang lain untuk melakukan kekerasan atau menyakiti orang atau kelompok lain. Dalam penyampaiannya, ujaran kebencian juga sering disertai dengan bahasa kasar. Bahasa kasar adalah ungkapan yang mengandung frasa atau kata-kata kasar dan dikomunikasikan secara lisan atau tertulis kepada lawan bicara, yang berdampak mempercepat terjadinya konflik sosial jika disertai dengan ujaran kebencian.

Pada penelitian ini akan menggunakan metode *deep learning Attention Based Long Short Term Memory* (LSTM) menggunakan *word embedding Fasttext*. LSTM dipilih karena dapat menangani masalah *vanishing gradient* yang tidak bisa ditangani oleh metode *Recurrent Neural Network* (RNN). *Word embedding fasttext* dipilih karena dapat mengatasi permasalahan *out of vocabulary* atau dengan kata lain dapat memberikan representasi kata yang tidak muncul dalam data latih. Sedangkan *Attention Based Mechanism* ditambahkan untuk menangani masalah saat input memiliki *sequence* yang sangat panjang.

Dari hasil pengujian, tingkat performa dari model *Attention Based LSTM* menggunakan *word embedding fasttext* yang telah dibangun memiliki nilai rata-rata *Accuracy* sebesar 71,703% *Precision* sebesar 70,815%, *Recall* sebesar 71,701% dan *F1-Score* sebesar 71,255%. Model dapat mengklasifikasikan teks dengan cukup baik, akan tetapi masih model masih memiliki kekurangan pada saat mengklasifikasi beberapa teks terutama pada kelas ujaran kebencian. Hal ini disebabkan karena adanya kata-kata yang bersinggungan antara kelas ujaran kebencian dengan kelas netral dan kelas bahasa kasar. Selain itu, dataset yang tidak seimbang (*unbalanced dataset*) juga menjadi faktor model masih melakukan beberapa kesalahan dalam mengklasifikasikan teks.

Kata Kunci: Ujaran Kebencian, Bahasa Kasar, Twitter, Klasifikasi Teks, *Deep Learning*, *Word Embedding*, *Fasttext*, *Long Short Term Memory*, *Attention Based Mechanism*