

ABSTRAK

Penggunaan Media Sosial memberikan beberapa manfaat bagi manusia untuk berinteraksi dalam masyarakat. Namun terkadang penyebaran informasi yang dilakukan oleh individu atau kelompok tidak dapat dibenarkan dan menimbulkan penyalahgunaan di media sosial. Penelitian ini tentang pendeteksian ujaran kebencian di media sosial khususnya kolom komentar Instagram dan juga Twitter dengan menggunakan salah satu arsitektur Recurrent Neural Network yaitu Bidirectional LSTM. Recurrent Neural Network adalah jenis jaringan saraf yang berisi loop, yang memungkinkan informasi disimpan di dalam jaringan. Penelitian ini bertujuan untuk mengidentifikasi ujaran kebencian di media sosial khususnya platform Instagram dan Twitter dengan beberapa aspek seperti Abusive, Individual, dan Other. Hasil klasifikasi Hate Speech menggunakan LSTM dan fasttext mencapai 81% untuk skor f1 dan berdasarkan aspek ujaran kebencian, aspek Abusive mencapai hasil lebih tinggi yaitu 84% pada skor f1.

Kata kunci: *Ujaran kebencian, Social Media, Instagram, Twitter, Fasttext, BiLSTM*

