

1. Pendahuluan

Ujaran kebencian adalah ujaran yang ditujukan kepada individu atau kelompok yang tidak mengikuti norma, mengancam, melecehkan, menghina, mempermalukan, menimbulkan kerugian, dan dapat menimbulkan kekacauan sosial [1], [2]. Di Indonesia, ujaran kebencian diatur dalam UU ITE Nomor 11 Tahun 2008, dengan hukuman maksimal 6 tahun. Ujaran kebencian dapat terjadi di mana saja, terutama di media sosial, di mana orang sering mengungkapkan pendapat, kritik, dan masih banyak lagi yang dapat menimbulkan kebencian [3]. Berdasarkan hal tersebut, diperlukan suatu sistem pendeteksi ujaran kebencian di Twitter untuk menciptakan lingkungan regulasi yang baik, aman, dan disiplin. Namun, pesan tweet dibatasi hingga 280 karakter [4]. Dengan panjang karakter yang terbatas, orang menggunakan berbagai cara untuk menyampaikan pesan yang diinginkan, seperti melalui bahasa gaul, singkatan, dan pengurangan huruf dalam kata. Hal ini dapat menyebabkan vocabulary mismatch dan menjadi masalah ketika sistem ingin mengklasifikasikan teks karena sistem dapat menganggap arti yang sama sebagai kata yang berbeda [5]. Untuk mengatasi masalah vocabulary mismatch dapat dikurangi dengan ekspansi fitur. Namun, ada banyak penelitian yang telah dilakukan sebelumnya mengenai deteksi ujaran kebencian. Pengembangan penelitian sistem deteksi ujaran kebencian saat ini menggunakan Deep Learning dan Word Embedding [6]–[12].

Penelitian tentang hybrid deep learning untuk deteksi ujaran kebencian telah dilakukan sebelumnya [9]–[11]. Pada penelitian [9], metode hybrid deep learning, yaitu CNN+GRU, dan dibandingkan dengan metode CNN, GRU, dan Bidirectional Encoder Representation from Transformers (BERT). Dalam studi [10], 2 Metode hybrid yang digunakan yaitu CNN-LSTM dan BiLSTM-CNN dibandingkan dengan metode CNN saja. Dalam penelitian [11], Metode hybrid yang digunakan yaitu CNN+GRU dibandingkan dengan metode CNN dan Support Vector Machine (SVM). Berdasarkan penelitian tersebut meskipun sudah menggunakan hybrid deep learning, namun belum ada yang menggunakan feature expansion untuk mengatasi ketidaksesuaian kosa kata.

Penelitian tentang feature expansion juga telah dilakukan sebelumnya [3], [12]. Pada penelitian [3], metode feature expansion digunakan dengan model Global Vectors (GloVe) untuk mengatasi ketidaksesuaian kosakata dan untuk metode classifier menggunakan algoritma Logistic Regression (LR), Random Forest (RF), dan Artificial Neural Network (ANN). Dalam studi lain [12], metode feature expansion digunakan dengan model Word2Vec dan untuk classifiernya menggunakan SVM dan RF. Namun penelitian-penelitian ini, belum ada yang menggunakan hybrid deep learning untuk metode classifiernya.

Tujuan serta fokus utama pada penelitian ini yaitu pengimplementasian single deep learning yaitu CNN dan GRU, serta hybrid deep learning yaitu CNN-GRU dan GRU-CNN, untuk mendeteksi ujaran kebencian dengan twitter bahasa indonesia dengan menggunakan feature expansion FastText untuk mengatasi vocabulary mismatch. Sejauh yang kami tahu, belum ada penelitian yang membandingkan hybrid deep learning dengan single deep learning yang digabungkan dengan feature expansion untuk mendeteksi ujaran kebencian di Indonesia. Batasan masalah pada penelitian ini yaitu hanya menggunakan data twitter berbahasa indonesia dengan 7 topik yaitu kepolisian, orientasi seksual, politik, covid-19, agama, ras, dan kata-kata khusus.

Jurnal TA ini terdiri dari 5 bagian, yaitu Bagian 1 yang berisi pendahuluan. Bagian II berisi Studi Terkait, yang mencakup pembahasan singkat tentang penelitian yang terkait dengan penelitian kami. Lalu ada metodologi yang digunakan dalam penelitian kami di Bagian III. Kemudian ada hasil dan analisis eksperimen yang dibahas di Bagian IV. Bagian V merupakan bagian terakhir yang berisi kesimpulan.