

1. Pendahuluan

Perkembangan internet membuat pencarian suatu informasi menjadi lebih mudah melalui media sosial dibandingkan dari koran, majalah, buku, atau sumber informasi fisik lainnya. Melalui media sosial, masyarakat dapat bebas berbagi informasi yang mereka miliki dengan cepat dan mudah. Kemudahan dalam berbagi informasi ini tidak selalu berdampak baik kepada masyarakat, salah satu dampak negatifnya adalah tersebarnya informasi atau berita palsu yang biasa disebut dengan *Hoax*. *Hoax* adalah informasi sesat dan berbahaya yang memiliki tujuan untuk mempengaruhi opini dan persepsi pembaca [1]. *Hoax* sengaja ditulis untuk menyesatkan pembaca agar mengandalkannya informasi palsu yang membuat informasi tersebut sulit untuk dideteksi kebenarannya berdasarkan konten berita [2].

Kesulitan dalam mendeteksi *hoax* adalah suatu persoalan serius yang apabila tidak ditangani dengan benar dapat berdampak besar terhadap berbagai hal. Deteksi *hoax* sangatlah berguna untuk mengetahui kebenaran dari suatu informasi.

Untuk mengetahui kebenaran dari suatu informasi, beberapa penelitian telah dilakukan dengan cara membangun sistem yang berfungsi untuk mendeteksi *hoax* di media sosial. Salah satu pendekatan untuk sistem deteksi *hoax* yang dapat digunakan yaitu pendekatan *Deep Learning*. Terdapat beberapa metode pada pendekatan *deep learning*, diantaranya *Long Short-Term Memory* (LSTM), *Bidirectional Long Short-Term Memory* (Bi-LSTM), *Gated Recurrent Unit* (GRU), *Bidirectional Gated Recurrent Unit* (Bi-GRU), dan *1 Dimensional-Convolutional Neural Network* (1D-CNN).

Metode LSTM telah dilakukan pada penelitian [3] dengan hasil akurasi 85.50% dan metode GRU dengan hasil akurasi 91.00%. Penelitian [4] dilakukan menggunakan metode Bi-GRU dan memberikan hasil akurasi 81.75%. Penelitian [5] melakukan pengujian dengan berbagai pendekatan *machine learning* pada tiga dataset berbeda yang besar dan beragam. Penelitian tersebut menerapkan *Global Vectors* (GloVe) sebagai fitur untuk mendapatkan representasi vektor dari kata-kata, dan mendapatkan akurasi tertinggi pada model *deep learning* sebesar 95.00% dengan model Bi-LSTM menggunakan gabungan dua korpus GloVe. Pada penelitian [6], metode Bi-LSTM menghasilkan tingkat akurasi 96.20% dan metode 1D-CNN dengan akurasi tertinggi yaitu 97.90%. Berdasarkan hasil penelitian tersebut, metode 1D-CNN memiliki akurasi tertinggi diikuti dengan metode Bi-LSTM.

Pada penelitian ini, terdapat tiga metode dengan pendekatan *deep learning* yang digunakan, yaitu Bi-LSTM, 1D-CNN dan *Hybrid* Bi-LSTM-1D-CNN yang merupakan gabungan dari dua metode sebelumnya. Dalam membangun model dari ketiga metode tersebut, digunakan TF-IDF sebagai ekstraksi fitur dan GloVe sebagai ekspansi fitur untuk pembangunan korpus. Data yang digunakan pada penelitian ini adalah data *tweet* masyarakat Indonesia pada media sosial Twitter.

Topik dan Batasannya

Berdasarkan penjelasan pada latar belakang, perumusan masalah pada penelitian ini yaitu bagaimana mengimplementasikan metode Bi-LSTM, 1D-CNN, dan *Hybrid* Bi-LSTM-1D-CNN untuk mendeteksi *hoax*, bagaimana menerapkan ekstraksi fitur untuk mendeteksi *hoax*, bagaimana menerapkan ekspansi fitur untuk mendeteksi *hoax*, bagaimana mendapatkan model dengan performansi terbaik dalam mendeteksi *hoax*.

Batasan masalah pada penelitian ini adalah sistem deteksi *hoax* yang dibangun menggunakan metode Bi-LSTM, 1D-CNN, *Hybrid* Bi-LSTM-1D-CNN, dan *tweet* masyarakat Indonesia mengenai berita Tragedi Kanjuruhan dan Kasus Ferdy Sambo yang digunakan sebagai dataset penelitian.

Tujuan

Berdasarkan penjelasan pada rumusan masalah, tujuan dari penelitian ini yaitu mengembangkan sistem deteksi *hoax* dengan metode Bi-LSTM, 1D-CNN, dan *Hybrid* Bi-LSTM-1D-CNN, mengimplementasikan fitur ekstraksi untuk deteksi *hoax*, mengimplementasikan fitur ekspansi untuk deteksi *hoax*, dan membandingkan masing-masing sistem untuk mendapatkan model dengan performansi terbaik dalam mendeteksi *hoax*.

Organisasi Tulisan

Jurnal TA ini terdiri dari 5 bagian, yaitu Bab 1 berisi pendahuluan dan penjelasan latar belakang, Bab 2 berisi penjelasan studi terkait, Bab 3 berisi penjelasan rancangan sistem dan metode penelitian, Bab 4 berisi penjelasan hasil pengujian dan analisis hasil pengujian, dan Bab 5 berisi kesimpulan dari penelitian yang telah dilakukan.