
1. Pendahuluan

Latar Belakang

Dalam era digital yang semakin berkembang, penggunaan media sosial telah menjadi bagian dari kehidupan sehari-hari masyarakat. Salah satu media sosial yang sangat populer adalah Twitter. Baru-baru ini Twitter telah berganti nama menjadi 'X'. 'X' merupakan media sosial yang memiliki sifat anonim [1]. Sifat anonim platform 'X' ini mempermudah orang untuk melakukan cyberbullying karena pengguna dari berbagai belahan dunia dapat secara bebas menyampaikan pendapat dan ekspresi mereka [2].

Cyberbullying merupakan masalah serius yang muncul di dunia maya, *Cyberbullying* dapat berupa ujaran kebencian, *trolling*, konten dewasa, rasisme, pelecehan, atau kata-kata kotor [3]. Dampak dari *cyberbullying* sering kali jauh lebih buruk dibandingkan dengan penindasan fisik atau verbal yang terjadi di dunia nyata, karena penyebaran informasi yang cepat dan mudah diakses dapat memperburuk trauma yang dialami korban [4].

Selain itu, adanya pelaku dengan identitas yang tersembunyi membuat proses pelaporan dan penindakan menjadi lebih sulit. Dengan demikian, deteksi dan pencegahan *cyberbullying* di media sosial, khususnya di platform seperti 'X', menjadi sangat penting untuk mengurangi dampak negatif ini. Hal ini mendorong perlunya pengembangan metode yang efektif dalam mendeteksi dan mengklasifikasikan tindakan *cyberbullying*. Deteksi yang lebih akurat tidak hanya berguna untuk mengidentifikasi pelaku, tetapi juga untuk melindungi korban dan mencegah terjadinya perundungan lebih lanjut.

Penelitian ini menggunakan pendekatan hybrid deep learning CNN-LSTM untuk mendeteksi *cyberbullying* karena berbagai penelitian sebelumnya telah membuktikan bahwa metode ini lebih efektif dibandingkan tanpa hybrid. Model *Convolutional Neural Network* (CNN) terbukti efektif dalam mengekstraksi fitur lokal dari teks, serta membantu mengidentifikasi pola dan struktur yang signifikan dalam kalimat atau dokumen [5]. Sementara itu, model *Long Short-Term Memory* (LSTM) dapat secara efektif menganalisis konteks teks yang memiliki kompleksitas dan panjang yang tinggi [6].

Dengan mengkombinasikan CNN dan LSTM, model ini dapat mengidentifikasi karakteristik teks penting dan memahami hubungan kontekstual dalam data sekuensial. Metode ini telah terbukti lebih efektif daripada metode tunggal dalam mendeteksi *cyberbullying* [7].

Topik dan Batasannya

Penelitian ini berfokus pada pengembangan metode deteksi *cyberbullying* menggunakan pendekatan hybrid deep learning CNN-LSTM serta menganalisis pengaruh fitur semantik dalam meningkatkan akurasi deteksi *cyberbullying* pada metode tersebut.

Tidak hanya itu, kombinasi antara CNN dan LSTM dengan ekstraksi fitur menjadi kunci penting dalam menghasilkan performa yang baik [8]. Ekstraksi fitur merupakan tahap krusial dalam proses *machine learning* yang bertujuan untuk mengidentifikasi dan memilih fitur-fitur yang paling relevan dari data yang kompleks [8].

Selain itu, penelitian ini juga mengeksplorasi pengaruh fitur semantik, yaitu IndoBERT, pada deteksi *cyberbullying*. IndoBERT adalah model bahasa berbasis BERT yang telah dilatih khusus untuk bahasa Indonesia. Pemanfaatan IndoBERT sebagai fitur semantik dapat membantu model memahami konteks teks secara lebih mendalam dan relevan dalam bahasa Indonesia, yang diharapkan dapat meningkatkan performa deteksi *cyberbullying*. Dengan mengombinasikan fitur semantik IndoBERT dengan ekstraksi fitur menggunakan TF-IDF dan ekspansi fitur menggunakan FastText, penelitian ini bertujuan untuk menggali lebih dalam mengenai kontribusi fitur-fitur ini dalam mendeteksi *cyberbullying* secara lebih efektif. Penelitian ini memiliki beberapa batasan sebagai berikut:

1. Dataset yang digunakan terdiri dari pesan-pesan yang berkaitan dengan *cyberbullying* di platform 'X'.
2. Data yang digunakan merupakan pesan-pesan yang berbahasa Indonesia.
3. Metode yang akan diterapkan mencakup penggunaan *Convolutional Neural Network* (CNN) dan *Long Short-Term Memory* (LSTM).
4. Teknik ekstraksi fitur yang akan digunakan mencakup TF-IDF, sedangkan untuk ekspansi fitur akan menggunakan FastText.

Tujuan

Tujuan dari penelitian ini adalah membangun model Hybrid Deep Learning mendeteksi *cyberbullying* di platform 'X' dengan menggunakan pendekatan *hybrid deep learning* yang menggabungkan metode *Convolutional Neural Network* (CNN) dan *Long Short-Term Memory* (LSTM). Penelitian ini juga bertujuan untuk menganalisis pengaruh fitur semantik, yaitu IndoBERT, dalam meningkatkan akurasi deteksi *cyberbullying* pada model *hybrid deep learning* CNN-LSTM. Selain itu, penelitian ini mengintegrasikan teknik ekstraksi fitur TF-IDF dan ekspansi fitur FastText guna meningkatkan performa model dalam mendeteksi tindakan *cyberbullying*. Dengan mengeksplorasi pengaruh fitur semantik pada deteksi

cyberbullying menggunakan model *hybrid deep learning* CNN-LSTM, penelitian ini dapat memberikan kontribusi ilmiah baru, mengingat topik ini belum banyak dibahas dalam penelitian sebelumnya.