
Deteksi Teks yang Mengandung Keinginan Bunuh Diri pada Media Sosial X Menggunakan Metode Hybrid Deep Learning CNN-BiLSTM

Nurlailiyah Salsabilah Valentina¹, Erwin Budi Setiawan²

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

⁴Divisi Digital Service PT Telekomunikasi Indonesia

¹salsabilavalenn@students.telkomuniversity.ac.id,

²erwinbudisetiawan@telkomuniversity.ac.id

Abstrak

Bunuh diri adalah salah satu penyebab kematian terbesar di dunia. Diantara banyak individu yang mengalami tekanan mental, sebagian memilih mengekspresikan perasaan mereka melalui media sosial seperti X. Unggahan-unggahan tersebut sering terdapat tanda-tanda adanya keinginan untuk mengakhiri hidup, yang dapat menjadi peringatan dini jika terdeteksi dengan tepat. Penelitian ini bertujuan mengembangkan sistem yang mampu mendeteksi teks dengan indikasi keinginan bunuh diri. Sistem ini dibangun dengan menggabungkan model deep learning CNN dan BiLSTM, menggunakan TF-IDF sebagai metode ekstraksi fitur dan FastText sebagai metode ekspansi fitur dengan menggunakan corpus similarity sebanyak 64.045 data. Serangkaian skenario dilakukan pada model yang dibangun menggunakan 49.022 data dalam bahasa Inggris yang dikumpulkan dari platform X. Hasil dari pengujian menunjukkan bahwa model hybrid mendapatkan nilai akurasi terbaik. Model hybrid yang mendapatkan nilai akurasi terbaik tersebut adalah CNN-BiLSTM dengan nilai akurasi sebesar 79,21% yang mana mendapatkan kenaikan akurasi dari model baseline sebesar 6,9%. Hasil tersebut menunjukkan bahwa model hybrid dengan kombinasi ekstraksi fitur TF-IDF dan ekspansi fitur FastText mampu mendeteksi indikasi keinginan bunuh diri pada unggahan di sosial media X.

Kata kunci: deteksi bunuh diri, hybrid deep learning, TF-IDF, FastText, X

Abstract

Suicide is one of the leading causes of death worldwide. Among the many individuals experiencing mental distress, some choose to express their feelings through social media platforms like X. These posts often contain signs of a desire to end one's life, which can serve as an early warning if detected accurately. This study aims to develop a system capable of detecting texts with indications of suicidal intent. The system is built by combining CNN and BiLSTM deep learning models, using TF-IDF for feature extraction and FastText for feature expansion, leveraging a corpus of 64,045 data samples. A series of scenarios were conducted on the model, which was built using 49,022 English-language data collected from the X platform. The results of the testing show that the hybrid model achieved the best accuracy score. The hybrid model, CNN-BiLSTM, achieved an accuracy of 79.21%, which represents an improvement of 6.9% compared to the baseline model. These results indicate that the hybrid model, with its combination of TF-IDF feature extraction and FastText feature expansion, is effective in detecting signs of suicidal intent in posts on the X social media platform.

Keywords: suicide detection, hybrid deep learning, TF-IDF, FastText, X

1. Pendahuluan

Latar Belakang

Bunuh diri merupakan salah satu masalah kesehatan masyarakat yang serius, dengan angka kematian mencapai sekitar 0,7 juta orang meninggal setiap tahun, terutama pada kalangan muda dan paruh baya [1]. Tindakan ini dipengaruhi oleh beberapa faktor yang dikelompokkan menjadi tiga kategori yaitu faktor kesehatan, faktor lingkungan, dan faktor yang terkait dengan riwayat pribadi [2]. Faktor tersebut akan memicu adanya perasaan yang tidak dapat diterima yang kemudian memunculkan ide untuk melukai diri sendiri hingga tindakan bunuh diri [3]. Ide melakukan tindakan bunuh diri dapat ditangani oleh ahli medis [4]. Akan tetapi, karena berbagai alasan seperti stigma sosial terhadap masalah mental, sebagian individu menghindari pengobatan medis dan memilih untuk mengekspresikan perasaan mereka melalui *platform* media sosial [5], [6]. Salah satu *platform* media sosial yang digunakan untuk mengekspresikan perasaan adalah Twitter, yang sekarang disebut dengan X.

X adalah *platform* media sosial yang paling umum digunakan untuk membagikan kata-kata, foto, video, dan tautan untuk saling berinteraksi [7]. Beberapa individu memanfaatkan anonimitas yang tersedia untuk berbagi kekhawatiran dan ketegangan yang mereka alami, termasuk mengekspresikan keinginan atau niat bunuh diri, mencari nasihat mengenai bunuh diri, bahkan percakapan yang mendorong tindakan bunuh diri [8]. Penggunaan X sebagai sarana menyampaikan perasaan berpotensi memiliki risiko penularan ide bunuh diri kepada individu lain [9]. Oleh karena itu, data pengguna di X dapat digunakan untuk mendeteksi tanda-

tanda adanya indikasi keinginan bunuh diri dan melakukan pencegahan terhadap tindakan tersebut.

Dalam mengatasi masalah tersebut, telah dilakukan berbagai penelitian terkait deteksi bunuh diri, beberapa diantaranya menggunakan model gabungan *deep learning Convolutional Neural Network* (CNN) dan *Long-Short Term Memory* (LSTM), yang menghasilkan akurasi sebesar 90,3% [10]. Penelitian lainnya melakukan pengujian menggunakan model *hybrid Convolutional Neural Network* (CNN) dan *Bidirectional Long-Short Term* (BiLSTM) yang mendapatkan hasil akurasi sebesar 94,29% [11]. Selain itu, penelitian lainnya yang juga menggunakan model *hybrid CNN-BiLSTM* mendapatkan nilai akurasi sebesar 95,% [12].

Penelitian ini menggunakan model *hybrid deep learning Convolutional Neural Network* (CNN) dan *Bidirectional Long-Short Term Memory* (BiLSTM) untuk mendeteksi teks yang mengandung indikasi bunuh diri. CNN digunakan untuk mengklasifikasikan dan menganalisis teks, sementara BiLSTM digunakan untuk mempertimbangkan konteks menyeluruh dengan memproses informasi dari dua arah [13]. Penggabungan kedua model tersebut dapat meningkatkan akurasi prediksi [14] dalam mendeteksi adanya indikasi keinginan bunuh diri.

Topik dan Batasannya

Penelitian ini berfokus pada deteksi teks yang mengandung indikasi keinginan bunuh diri menggunakan *hybrid deep learning*, yaitu CNN dan BiLSTM. Dengan menambahkan *Term Frequency-Inverse Document Frequency* (TF-IDF) dan *word embedding* FastText, model dapat mengubah masukan yang diberikan dan meningkatkan akurasi prediksi dibandingkan dengan model *baseline* [15]. Ekstraksi fitur TF-IDF memberikan pembobotan dalam mengidentifikasi kata-kata penting dalam teks [16]. Sedangkan ekspansi fitur FastText yang mengukur kesamaan kata [17].

Penelitian ini memiliki beberapa batasan sebagai berikut:

1. Dataset yang digunakan berupa teks dari X yang berkaitan dengan bunuh diri.
2. Data yang digunakan merupakan teks berbahasa Inggris.
3. Menggunakan model CNN dan BiLSTM.
4. Menggunakan ekstraksi fitur TF-IDF dan ekspansi fitur FastText.

Tujuan

Tujuan dari penelitian ini adalah membangun model *hybrid deep learning* untuk mendeteksi indikasi keinginan bunuh diri di *platform* X dengan menggunakan model CNN dan BiLSTM. Penelitian ini juga bertujuan untuk menggabungkan ekstraksi fitur TF-IDF dan ekspansi fitur FastText untuk meningkatkan pemahaman semantik dan meningkatkan nilai akurasi model dalam mendeteksi indikasi keinginan bunuh diri dalam teks.

2. Studi Terkait

Penelitian terkait deteksi indikasi keinginan bunuh diri diantaranya menerapkan *hybrid deep learning*, seperti yang dilakukan oleh Renjith Shini, dkk [10]. Dengan menggunakan Word2Vec sebagai *word embedding* pada input model LSTM-Attention-CNN dan juga pengoptimalan parameter, diperoleh nilai akurasi sebesar 90,3%. Dikarenakan dataset tidak seimbang, evaluasi model dengan nilai akurasi belum tentu tepat. Sehingga dilakukan evaluasi menggunakan ukuran yang akurat seperti presisi mendapat nilai 91,6%, recall mendapat nilai sebesar 93,7%, dan F1-score mendapat nilai sebesar 92,6%.

Penelitian lainnya dilakukan oleh Mohaiminul Islam Bhuiyan, dkk [11] yang melakukan pengujian *hybrid deep learning* (CNN) dan (BiLSTM) dengan menggunakan *word embedding* Word2Vec sebagai ekspansi fiturnya. Penelitian ini melakukan *fine tuning* pada berbagai komponen model untuk meningkatkan kinerja model. Salah satu penyesuaian yang dilakukan adalah teknik regulasi terutama pada L2 *regularization* dan layer *dropout*. Selain itu, digunakan *early stopping* untuk menghindari terjadinya *overfitting*. Dari pengujian yang dilakukan, didapat bahwa model *hybrid* yang dilakukan *fine tuning* mendapatkan peningkatan akurasi dibandingkan model awal yaitu sebesar 92,81%.

Penelitian lainnya terkait deteksi keinginan bunuh diri dilakukan oleh Aldhyani, dkk [12]. Penelitian ini melakukan pengujian pada *hybrid deep learning* CNN dan BiLSTM yang digabungkan dengan ekstraksi fitur TF-IDF dan Word2Vec, serta membandingkannya dengan model XGBoost. Model *hybrid* CNN-BiLSTM mampu menangkap pola linguistik yang kompleks dalam mendeteksi bunuh diri. Diketahui bahwa, model *hybrid deep learning* CNN-BiLSTM mendapatkan hasil akurasi terbaik sebesar 95% dan lebih baik dibandingkan dengan model XGBoost yang mendapatkan akurasi terbaik sebesar 91,5%.

Berdasarkan penelitian sebelumnya yang berkaitan dengan deteksi indikasi bunuh diri, dapat dikatakan bahwa model gabungan dari *deep learning* memiliki nilai akurasi yang lebih baik.

3. Sistem yang Dibangun

Dalam membangun sistem untuk mendeteksi indikasi keinginan bunuh diri pada *platform* X, penelitian ini menggunakan model *hybrid deep learning* yang dapat dilihat pada Figure 1.

Tahapan yang diusulkan meliputi: data *crawling*, data *labeling*, data *pre-processing*, data *split* menjadi data train dan data test, ekstraksi fitur TF-IDF, ekspansi fitur FastText, kemudian mengklasifikasikan data menggunakan empat model yaitu: CNN, BiLSTM, CNN-BiLSTM, BiLSTM-CNN. Terakhir, mengevaluasi performa model yang telah dibangun.