
1. Pendahuluan

Latar Belakang

Bunuh diri menelan korban hampir 800.000 jiwa setiap tahunnya dan menjadi penyebab kematian tertinggi kedua di kalangan generasi muda. Tingkat bunuh diri global mencapai 10,5 per 100.000 penduduk, dan diprediksi bahwa pada tahun 2020, satu orang akan meninggal akibat bunuh diri setiap 20 detik. Sebagian besar (sekitar 79%) kasus bunuh diri terjadi di negara-negara berpenghasilan rendah dan menengah, di mana akses terhadap sumber daya untuk mendeteksi dan menangani masalah ini masih sangat terbatas dan kurang memadai [1].

Deteksi dini bunuh diri memegang peranan penting dalam upaya pencegahan, terutama di kalangan individu yang berisiko tinggi [2]. Proses deteksi ini melibatkan identifikasi tanda- tanda peringatan seperti perubahan perilaku drastis, perasaan putus asa, pernyataan ingin mengakhiri hidup, atau kecenderungan menyakiti diri sendiri [3]. Program deteksi dini di berbagai negara fokus pada pelatihan tenaga kesehatan, pendidik, dan komunitas untuk mengenali gejala awal dan memberikan intervensi yang tepat. Selain itu, teknologi seperti aplikasi kesehatan mental dan layanan bantuan daring turut membantu memantau kondisi psikologis seseorang. Dengan deteksi dini yang efektif, individu yang memiliki kecenderungan bunuh diri dapat segera menerima dukungan, konseling, atau perawatan yang dibutuhkan untuk mencegah terjadinya tindakan fatal [1].

Penelitian ini menggunakan pendekatan *hybrid deep learning* CNN-LSTM untuk model yang digunakan adalah penerapan Convolutional Neural Network (CNN) dan Long Short-Term Memory (LSTM) dengan ekspansi fitur Word2Vec [4]. Metode ini efektif untuk mengolah data teks dari media sosial, seperti Twitter, yang sering kali berisi ungkapan singkat dan emosional. Word2Vec merepresentasikan kata sebagai vektor numerik berdasarkan konteksnya, memungkinkan model mengenali hubungan antar kata, termasuk sinonim dan ekspresi emosional. CNN membantu mengidentifikasi pola lokal seperti frasa dengan makna negatif atau emosional [5], sementara LSTM unggul dalam memahami urutan kata dan konteks yang lebih luas, termasuk emosi tersembunyi dalam struktur kalimat kompleks [6]. Dengan kombinasi CNN dan LSTM serta dukungan Word2Vec, sistem ini mampu mendeteksi tanda- tanda ide bunuh diri secara lebih akurat dalam mendeteksi ide bunuh diri [7].

Meskipun teknologi seperti CNN, LSTM, dan Word2Vec menjanjikan deteksi dini ide bunuh diri, keterbatasan dataset dan bias anotasi manual masih menjadi kendala. Ukuran dataset yang kecil membatasi kemampuan model mengenali berbagai variasi ekspresi emosional, sementara bias dari anotator dapat menyebabkan model salah menafsirkan konteks. Akibatnya, akurasi deteksi menjadi tidak konsisten, mengurangi efektivitas sistem ini dalam penerapan nyata [4].

Topik

Penelitian ini berfokus pada pengembangan metode deteksi potensi bunuh diri di media sosial Twitter menggunakan pendekatan *hybrid deep learning* CNN-LSTM, serta menganalisis pengaruh ekspansi fitur berbasis Word2vec dalam meningkatkan akurasi deteksi teks yang mengandung potensi *suicidal ideation*.

Arsitektur CNN-LSTM dipilih karena kemampuannya dalam memadukan keunggulan CNN dalam menangkap pola lokal seperti frasa bermakna, dengan kekuatan LSTM dalam memahami urutan dan konteks temporal kata dalam teks. Kombinasi ini terbukti efektif dalam berbagai tugas klasifikasi teks, khususnya yang melibatkan emosi atau nuansa psikologis yang kompleks [9].

Penelitian ini turut mengeksplorasi pengaruh fitur semantik Word2vec dalam mendeteksi potensi bunuh diri pada unggahan Twitter. Word2vec memetakan kata ke dalam vektor berdimensi tinggi berdasarkan konteks, sehingga mampu menangkap makna implisit dan emosional. Fitur ini dikombinasikan dengan ekstraksi TF-IDF untuk menyeimbangkan relevansi kata secara statistik dan semantik. Dengan integrasi ke dalam arsitektur CNN-LSTM, kombinasi ini diharapkan dapat meningkatkan akurasi dan sensitivitas deteksi secara kontekstual [10].

Penelitian ini memiliki beberapa batasan sebagai berikut:

1. Dataset yang digunakan terdiri dari unggahan (*tweet*) yang berkaitan dengan isu atau potensi bunuh diri di platform Twitter.
2. Data yang digunakan merupakan pesan-pesan yang berbahasa Inggris.
3. Teknik ekstraksi fitur yang digunakan mencakup TF-IDF

Tujuan

Tujuan dari penelitian ini adalah untuk membangun model *Hybrid deep learning* dalam mendeteksi potensi Bunuh Diri di platform Twitter, dengan menggunakan pendekatan yang menggabungkan metode *Convolutional Neural Network* (CNN) dan *Long Short-Term Memory* (LSTM). Penelitian ini juga bertujuan untuk menganalisis pengaruh fitur semantik Word2vec dalam meningkatkan akurasi deteksi potensi bunuh diri pada model CNN-LSTM. Selain itu, penelitian ini mengintegrasikan teknik ekstraksi fitur TF-IDF dan ekspansi fitur Word2vec guna meningkatkan performa model dalam menangkap konteks semantik dan nuansa emosional pada teks yang berisiko *suicidal*. Untuk meningkatkan performa pelatihan model, penelitian ini juga mengeksplorasi penggunaan optimasi dengan berbagai algoritma optimisasi, yaitu Adam, Nadam, dan Adamax, untuk mengidentifikasi *optimizer* yang memberikan hasil terbaik dalam hal akurasi dan konvergensi model. Dengan mengeksplorasi pengaruh kombinasi fitur semantik dan algoritma optimasi pada deteksi potensi bunuh diri menggunakan model *hybrid deep learning* CNN-LSTM, penelitian ini diharapkan dapat memberikan kontribusi ilmiah baru, mengingat topik ini belum banyak dibahas dalam penelitian sebelumnya.

2. Studi Terkait

Dalam kajian pustaka ini, terdapat beberapa penelitian yang relevan dengan masalah deteksi dini niat bunuh diri di media sosial. Penelitian pertama oleh Tadesse et al. (2020) mengidentifikasi tantangan dalam mendeteksi ide bunuh diri akibat variasi bahasa dan konteks sosial media. Mereka mengembangkan metode berbasis *deep learning* menggunakan CNN dan LSTM untuk meningkatkan akurasi deteksi. Kelebihan dari penelitian ini adalah penggabungan kekuatan CNN untuk ekstraksi fitur dan LSTM untuk menangkap ketergantungan jangka panjang, meskipun terdapat kekurangan terkait ukuran dataset yang terbatas [8].

Selanjutnya, Tam et al. (2022) mengevaluasi efektivitas model ANN, termasuk CNN, LSTM, dan BERT, dalam menganalisis konten Twitter yang berisiko bunuh diri. Penelitian ini menunjukkan bahwa BERT memiliki kinerja terbaik dalam hal presisi dan stabilitas. Namun, penelitian ini juga memiliki keterbatasan, seperti dataset yang hanya mencakup *tweet* berbahasa Inggris [11].

Penelitian oleh Lin et al. (2024) mengusulkan model RoBERTa-CNN yang menggabungkan kekuatan RoBERTa dalam memahami konteks dan CNN dalam mengekstrak fitur lokal. Hasil penelitian ini menunjukkan akurasi tinggi (98%) dan AUC tinggi (97,5%), tetapi memerlukan daya komputasi yang besar dan terbatas pada data teks [12].

Penelitian mengenai deteksi ide bunuh diri telah banyak dilakukan dengan pendekatan teknologi kecerdasan buatan dan pembelajaran mesin. Ji et al. (2021) melakukan tinjauan menyeluruh terhadap metode deteksi ide bunuh diri menggunakan analisis fitur dan teknik pembelajaran mendalam seperti *Convolutional Neural Networks* (CNN) dan *Long Short-Term Memory* (LSTM) pada konten dari media sosial seperti Twitter. Studi tersebut menekankan efektivitas penggunaan word embedding seperti Word2Vec untuk menangkap hubungan antar kata dalam teks singkat dan emosional. Namun, keterbatasan dataset menjadi tantangan signifikan [13].

Aldhyani et al. (2022) melakukan studi mengenai deteksi ide bunuh diri dengan menggunakan metode *deep learning* dan machine learning, khususnya model *hybrid* CNN-BiLSTM dan algoritma XGBoost, pada data Reddit yang berisi postingan dari platform *SuicideWatch*. Penelitian ini menekankan pentingnya penggunaan Word2Vec untuk merepresentasikan kata dalam bentuk vektor, yang memungkinkan model mengenali hubungan antar kata dalam teks singkat dan emosional. Namun, keterbatasan dataset masih menjadi tantangan utama, terutama dalam hal akurasi model dan kemampuannya untuk mengenali variasi bahasa yang lebih beragam di media sosial [14].

Fokus utama dari penelitian ini adalah untuk mengembangkan dan mengevaluasi metode deteksi dini ide bunuh diri pada media sosial Twitter dengan menggabungkan model CNN-LSTM dan ekspansi fitur menggunakan Word2Vec. Penelitian ini bertujuan untuk meningkatkan akurasi dan sensitivitas deteksi ide bunuh diri dengan memperkuat representasi semantik kata-kata dalam *tweet*, yang sering kali bersifat singkat dan emosional. Dengan menggunakan Word2Vec, penelitian ini berharap dapat memperkaya representasi teks, memungkinkan model untuk menangkap hubungan semantik antar kata secara lebih efektif, dan meningkatkan kemampuan model dalam memahami konteks yang kompleks.

3. Sistem yang dibangun

Sistem yang dibangun untuk mendeteksi potensi bunuh diri pada platform Twitter menggunakan model *hybrid deep learning*, ekstraksi fitur TF-IDF dan ekspansi fitur semantik dengan Word2vec dapat dilihat pada Gambar 1.

Langkah-langkah sistem deteksi bunuh diri meliputi data *crawling*, data labelling, pre-processing data, ekstraksi fitur TF-IDF, ekspansi fitur Word2vec, dan splitting data menjadi data uji dan data latih. Serta klasifikasi data dengan empat model: CNN, LSTM, CNN-LSTM, CNN-LSTM. Terakhir, kinerja sistem yang dibangun akan dievaluasi.