

Pendeteksian berita palsu menggunakan RoBERTa dengan Optimalisasi Word Embedding

Adisaputra Nur Arminta¹, Yuliant Sibaroni²

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

¹adisaputraa@student.telkomuniversity.ac.id, ²yuliantsibaroni@telkomuniversity.ac.id

Abstrak

Penyebaran berita palsu (*hoax*) telah menjadi permasalahan serius yang mempengaruhi opini publik dan menciptakan polarisasi di masyarakat. Penelitian ini bertujuan untuk mendeteksi berita palsu menggunakan model *RoBERTa* yang dioptimalkan dengan tiga teknik *word embedding*. *Word embedding* yang digunakan adalah *RoBERTa*, *Word2Vec*, dan *GloVe*. Dataset yang digunakan adalah "Indonesian fact and hoax political news" yang diambil dari Kaggle. Dataset ini memerlukan tahap pre-processing untuk membersihkan ketidakkonsistenan data, seperti mengubah singkatan menjadi kata lengkap dan menghapus tanda baca. Selanjutnya, dilakukan representasi teks menggunakan tiga metode word embedding yaitu *Word2Vec*, *GloVe*, dan *RoBERTa*. Proses pelatihan model dilakukan dengan validasi silang K-Fold untuk meningkatkan generalisasi model. Hasil penelitian menunjukkan bahwa embedding *RoBERTa* mencapai akurasi terbaik 96%, sedangkan word embedding *Word2Vec* mendapatkan akurasi 94%. Word Embedding *Glove* menunjukkan performa paling rendah dengan akurasi 51%. Penelitian ini membuktikan bahwa pemilihan teknik word embedding yang tidak tepat untuk model *RoBERTa* dapat mengurangi akurasi dan efektivitas model dalam mendeteksi berita palsu. Diharapkan bahwa temuan dalam penelitian ini dapat memberikan kontribusi terhadap peningkatan sistem deteksi berita palsu di masa mendatang.

Kata kunci: *hoax*, *RoBERTa*, *GloVe*, *Word2Vec*

Abstract

The dissemination of misinformation (hoaxes) has emerged as a significant issue, shaping public perception and contributing to societal division. This research aims to detect fake news using the *RoBERTa* model which is optimized with three word embedding techniques. The word embeddings used are *RoBERTa*, *Word2Vec*, and *GloVe*. The dataset used is "Indonesian fact and hoax political news" taken from Kaggle. This dataset requires a pre-processing stage to clean up data inconsistencies, such as changing abbreviations into complete words and removing punctuation. Next, text representation is carried out using three word embedding methods: *Word2Vec*, *GloVe*, and *RoBERTa*. The training phase of the model utilizes the K-Fold cross-validation approach to enhance its ability to generalize across different data distributions. The research results show that *RoBERTa* embedding achieves the best accuracy of 96%, while *Word2Vec* word embedding achieves 94% accuracy. Word Embedding *Glove* shows the lowest performance with 51% accuracy. This study demonstrates that an unsuitable word embedding method for the *RoBERTa* model can negatively impact both accuracy and efficiency in identifying fake news. The findings from this research are expected to support advancements in future fake news detection systems.

Keywords: *hoax*, *RoBERTa*, *GloVe*, *Word2Vec*

1. Pendahuluan

1.1 Latar Belakang

Hoax berasal dari *hocus* yaitu untuk menipu sering kali muncul pada topik yang sedang hangat dibicarakan. Tujuannya adalah untuk membujuk atau memanipulasi orang dan kemudian melakukan tindakan yang telah ditetapkan sebelumnya, biasanya menggunakan ancaman atau membuat mereka mempercayai hal-hal yang tidak nyata [1]. *Hoax* adalah informasi palsu yang beredar di berbagai platform dan saluran komunikasi, contohnya seperti media sosial dan situs web di internet.

Dengan berkembangnya teknologi digital, pengguna internet dapat dengan mudah membagikan informasi dan berinteraksi sesama pengguna, namun di sisi lain, informasi tersebut juga bisa disalahgunakan untuk menyebarkan berita palsu yang dapat memberikan dampak negatif kepada masyarakat [2]. Oleh karena itu, penyebaran hoaks di dunia maya perlu segera ditangani, karena dapat memicu ketegangan sosial, meningkatkan rasa permusuhan, dan bahkan menyebabkan konflik antar kelompok [3]. Dengan demikian, pengembangan sistem deteksi hoaks otomatis yang mampu mengidentifikasi serta menangani berita palsu secara cepat menjadi hal yang krusial sebelum informasi tersebut menyebar lebih luas.

Dalam upaya mengatasi penyebaran *hoax*, penggunaan salah satu metode yang menjanjikan untuk mengklasifikasi berita *hoax* adalah menggunakan *transformer* [4]. Salah satu metode *transformer* yaitu adalah *RoBERTa* (*Robustly Optimized BERT Approach*), *RoBERTa* adalah modifikasi dari *BERT* yang sederhana namun efektif [18]. Modifikasi tersebut meliputi melatih model lebih lama dengan *batch* yang lebih besar, menggunakan lebih banyak data, menghapus objektif prediksi kalimat berikutnya, melatih pada urutan yang lebih panjang, dan secara dinamis mengubah pola *masking* yang diterapkan pada data latih. Melalui modifikasi ini, *RoBERTa* dapat menghasilkan hasil terbaik pada berbagai tugas pemrosesan bahasa seperti *GLUE*, *RACE*, dan *SQuAD* [18].

Karena dalam beberapa penelitian sebelumnya telah dilakukan upaya untuk mendeteksi *hoax* menggunakan RoBERTa, Word2vec, dan GloVe. Salah satunya adalah Penelitian yang dilakukan oleh Vladislav Kolev et al. menggunakan model RoBERTa untuk mendeteksi berita palsu berdasarkan analisis emosi dalam judul berita. Model ini mencapai akurasi 88% pada Kaggle Fake and Real News Dataset [7]. Di sisi lain, penelitian oleh Bankar S memanfaatkan pendekatan *hybrid* dengan menggabungkan Word2Vec dan model LSTM untuk menganalisis hubungan semantik antar kata dalam teks berita palsu [8]. Lebih lanjut, penelitian yang dilakukan oleh Ryan Adipradana et al. menggunakan *embedding GloVe* dalam berbagai model jaringan saraf seperti LSTM dan GRU untuk mendeteksi berita palsu dalam bahasa Indonesia. Hasilnya menunjukkan bahwa *embedding GloVe* mampu memberikan performa yang baik dengan akurasi mencapai lebih dari 81% pada tugas klasifikasi berita [9]. Penelitian oleh Shafna et al. menunjukkan bahwa model *RoBERTa* mampu mendeteksi berita palsu dengan akurasi 83,3% pada dataset berbahasa Indonesia [10].

Maka dari itu, penulis mengusulkan penggunaan metode RoBERTa karena kemampuannya dalam menangkap konteks linguistik yang kompleks melalui pelatihan yang lebih intensif dan penggunaan data yang lebih besar [15]. Selain itu, untuk memperkuat pemahaman konteks dalam kalimat, penulis mengintegrasikan teknik Word Embedding seperti Word2Vec dan GloVe. Word2Vec dipilih karena kemampuannya dalam memetakan hubungan semantik antar kata berdasarkan distribusi kata dalam korpus [6], sementara GloVe dipilih karena kemampuannya dalam menangkap hubungan global antar kata melalui faktorisasi matriks ko-okurensi [21]. Integrasi ini diharapkan dapat meningkatkan performa model dalam mengidentifikasi berita hoax dan penerapan integrasi ini diharapkan menjadi solusi yang relevan dan penting dalam mendeteksi hoax, khususnya di era media sosial yang mendorong percepatan penyebaran informasi palsu.

1.2 Topik dan Batasannya

Dalam penelitian ini, penulis merancang dan mengembangkan sebuah sistem untuk melakukan deteksi berita hoax berdasarkan data berita di Indonesia Menggunakan model *RoBERTa* dan *Word Embedding*. Penelitian ini memiliki beberapa Batasan, yakni: (1) Data yang digunakan berasal dari *dataset kaggle* yaitu “Indonesia Fact and Hoax Political News”, dalam dataset tersebut data - data pada teks sudah diklasifikasikan ke dalam dua kategori yaitu *Hoax* dan *Non-Hoax* ;(2) Data yang digunakan dalam penelitian ini adalah sebanyak 5000 dari total keseluruhan *dataset* tersebut.

1.3 Tujuan

Penelitian ini memiliki tujuan untuk mengembangkan sistem yang dapat mengenali berita hoaks dengan menggunakan model deep learning RoBERTa serta optimalisasi teknik word embedding. Selain itu, penelitian ini juga mengeksplorasi bagaimana penggunaan berbagai jenis word embedding, seperti *Word2Vec* dan *GloVe*, dapat memengaruhi kinerja model RoBERTa. Penilaian kinerja sistem akan dilakukan dengan menerapkan metrik evaluasi yang sesuai, seperti akurasi, presisi, recall, dan F1-score, guna menganalisis efektivitas masing-masing metode word embedding yang diterapkan.

2. Studi Literatur

2.1 Penelitian Terkait

Sudah ada beberapa penelitian sebelumnya yang berfokus pada deteksi *hoax* di Indonesia dengan menggunakan berbagai metode pendekatan pemrosesan bahasa alami. Seperti penelitian yang dilakukan oleh Muliono et al. yang berjudul “Hoax Classification in Imbalanced Datasets Based on Indonesian News Title using RoBERTa” [19], menunjukkan bahwa penggunaan model RoBERTa untuk mengklasifikasikan berita hoaks berdasarkan judulnya saja menghasilkan akurasi yang tinggi, yaitu mencapai 99,52%. Crisanadenta et al. juga telah meneliti sebuah model deteksi hoax di twitter menggunakan metode *feed-forward* dan *back-propagation neural networks* [11]. Dengan menggunakan dua metode pembobotan fitur yaitu TF-IDF dan Word2Vec. Hasil penelitian menunjukkan bahwa model *neural networks* dengan TF-IDF mencapai akurasi tertinggi sebesar 78,76% [11].

Selanjutnya, penelitian yang dilakukan oleh Irena et al. menerapkan metode *Decision Tree* untuk mengidentifikasi berita palsu (*hoax*) di Twitter. Hasil penelitian menunjukkan bahwa penggunaan fitur n-gram dapat meningkatkan akurasi dalam mengklasifikasikan *hoax*. Peneliti menunjukkan bahwa penggunaan kombinasi unigram + bigram dan unigram + bigram + trigram memberikan nilai akurasi yang tinggi. Nilai akurasi tertinggi yang dicapai dalam penelitian tersebut adalah 72.91% dengan menggunakan metode *Decision Tree* C4.5, TF-IDF *weighting*, dan *Information Gain feature selection*, dengan rasio data latih 90% dan data uji 10% (90:10) serta menggunakan 5000 fitur unigram [12]. Adrian dan Prasetyowati juga melakukan penelitian yang dimana mereka membandingkan teknik *Word Embedding* yaitu Word2Vec dan GloVe. Hasil penelitian tersebut menunjukkan bahwa LSTM dengan Word2Vec memiliki kinerja yang lebih baik dibanding LSTM dengan GloVe dalam mendeteksi berita hoax dalam bahasa Indonesia. LSTM dengan Word2Vec mencapai nilai akurasi rata-rata sebesar 95%, sementara LSTM dengan GloVe hanya mencapai nilai akurasi rata rata sebesar 90% [13].