

## **LEMBAR PERNYATAAN**

Dengan ini saya, Maulana Malik Ibrahim Lubis, menyatakan sesungguhnya bahwa Tugas Akhir saya dengan judul Analisis Sentimen Pemilihan Umum Tahun 2024 di Indonesia pada X Menggunakan Random Forest dengan FastText beserta dengan seluruh isinya adalah merupakan hasil karya sendiri, dan saya tidak melakukan penjiplakan yang tidak sesuai dengan etika keilmuan yang berlaku dalam masyarakat keilmuan. Saya siap menanggung resiko/sanksi yang diberikan jika di kemudian hari ditemukan pelanggaran terhadap etika keilmuan dalam buku TA atau jika ada klaim dari pihak lain terhadap keaslian karya,

Bandung, 30 Agustus 2024

Yang Menyatakan

Maulana Malik Ibrahim Lubis

**Analisis Sentimen**  
**Pemilihan Umum Tahun 2024 di Indonesia pada X**  
**Menggunakan Random Forest dengan FastText**  
**Maulana Malik Ibrahim Lubis<sup>1</sup>, NamaPenulis<sup>2</sup>, NamaPenulis<sup>3</sup>**

<sup>1,2,3</sup>Fakultas Informatika, Universitas Telkom, Bandung  
<sup>1</sup>mhs@students.telkomuniversity.ac.id, <sup>2</sup>pembimbing1@telkomuniversity.ac.id,  
<sup>3</sup>pembimbing2@telkomuniversity.ac.id

---

**Abstrak**

Pemilihan Umum (Pemilu) merupakan salah satu momen krusial dalam proses demokrasi yang menentukan arah politik dan kepemimpinan suatu negara. Di Indonesia, Pemilu 2024 telah menjadi topik hangat yang mendominasi perbincangan publik, terutama di media sosial. Media sosial, khususnya platform X, berperan penting dalam mencerminkan dan membentuk opini publik mengenai isu-isu politik. Penelitian ini menyelidiki analisis sentimen terhadap Pemilihan Umum 2024 di Indonesia melalui media sosial X, menggunakan pendekatan Random Forest dan Decision Tree dengan FastText. Pemilu memiliki peran kunci dalam menentukan arah politik dan kepemimpinan negara. Peran media sosial, khususnya X, semakin penting dalam memahami sentimen publik terkait isu-isu politik. Studi sebelumnya menunjukkan efektivitas analisis sentimen menggunakan machine learning, namun tantangan seperti overfitting dan kompleksitas data tetap perlu diatasi. Penelitian ini menggunakan Random Forest untuk menangani kompleksitas data dan FastText untuk representasi kata dalam analisis teks. Metode Decision Tree juga digunakan untuk perbandingan performa. Hasil penelitian menunjukkan hasil terbaik dengan akurasi sebesar 70,60% dan f1-score sebesar 70,09%, menunjukkan efektivitas pendekatan ini dalam analisis sentimen politik di media sosial.

**Kata kunci : analisis sentimen, pemilihan umum, random forest, decision tree, fasttext**

---

**Abstract**

*General elections are one of the crucial moments in the democratic process that determine the political direction and leadership of a country. In Indonesia, the 2024 General Election has become a hot topic that dominates public conversations, especially on social media. Social media, particularly platform X, plays an important role in reflecting and shaping public opinion on political issues. This research investigates the sentiment analysis of the 2024 General Election in Indonesia through social media X, using Random Forest and Decision Tree approaches with FastText. Elections play a key role in determining the political direction and leadership of the country. The role of social media, especially X, is increasingly important in understanding public sentiment related to political issues. Previous studies have shown the effectiveness of sentiment analysis using machine learning, but challenges such as overfitting and data complexity still need to be addressed. This research uses Random Forest to handle data complexity and FastText for word representation in text analysis. Decision Tree method is also used for performance comparison. The results show the best results with an accuracy of 70.60% and f1-score of 70.09%, demonstrating the effectiveness of this approach in analyzing political sentiment on social media.*

**Keywords: sentiment analysis, general election, random forest, decision tree, fasttext**

---

## 1. Pendahuluan

### Latar Belakang

Pemilihan umum di Indonesia merupakan salah satu proses demokratisasi yang penting dalam sistem pemerintahan negara. Dilakukan setiap lima tahun sekali, pemilihan umum menjadi momen penting bagi rakyat Indonesia untuk menentukan arah politik dan kepemimpinan negara. Peran media sosial semakin signifikan dalam proses politik dan penelitian sosial. Di antara platform media sosial, X telah menjadi sumber data yang vital dalam menganalisis sentimen publik, terutama dalam konteks pemilihan umum. Analisis sentimen dapat memberikan wawasan yang berharga tentang pandangan dan sikap masyarakat terhadap berbagai isu politik dan sosial. Penggunaan X sebagai sumber data untuk analisis sentimen tidak terbatas hanya di Indonesia saja, melainkan juga telah dilakukan dalam konteks pemilihan umum di negara-negara lain [1]. Setelah dilaksanakannya pemilihan umum, penting untuk melakukan analisis sentimen untuk memahami dinamika politik dan opini publik yang berkembang. Analisis sentimen tidak hanya membantu dalam mengukur respons masyarakat terhadap keputusan politik, tetapi juga mengidentifikasi pola perilaku pemilih serta mendeteksi isu-isu yang muncul dan memengaruhi opini publik.

Seiring dengan kemajuan teknologi dan metodologi penelitian, penelitian-penelitian sebelumnya telah menghasilkan beragam temuan dan metode. Pada tahun 2021 telah dilakukan penelitian analisis sentimen pada media sosial Twitter dengan topik pemilihan presiden Indonesia tahun 2019 menggunakan *neural network* [1]. Penelitian ini menggunakan lima model machine learning yang menggunakan neural network yaitu LSTM, CNN, CNN+LSTM, GRU+LSTM, dan Bidirectional LSTM. Didapatkan hasil bahwa model Bidirectional LSTM mendapatkan performa yang paling baik dengan nilai akurasi 84.6%.

Namun, dalam proses penelitian, terdapat beberapa tantangan yang perlu diatasi. Salah satunya adalah kemungkinan terjadinya *overfitting* dan kesulitan dalam menangani fitur-fitur tidak linear dalam data. Untuk mengatasi hal ini, metode Random Forest dapat menjadi pilihan yang efektif karena mampu menangani kerumitan dan variabilitas dalam data dengan baik. Beberapa penelitian sebelumnya menunjukkan bahwa metode Random Forest dapat menghasilkan performa yang lebih baik dalam berbagai bidang dibandingkan dengan metode *machine learning* lainnya [2][3][4]. Dalam analisis sentimen, Random Forest juga digunakan dalam penelitian oleh Talha Ahmed Khan et al. pada tahun 2024, yang menghasilkan akurasi sebesar 78,56% [5]. Topik penelitian yang menggunakan metode Random Forest sering dibandingkan dengan Decision Tree. Hal ini karena Random Forest adalah kumpulan dari banyak Decision Tree tunggal. Penggunaan Random Forest biasanya menghasilkan performa yang lebih baik dibandingkan hanya menggunakan satu Decision Tree [6]. Selain itu, penggunaan data teks yang bersifat tidak terstruktur seringkali menjadi masalah tersendiri dalam pengembangan model *machine learning*.

Untuk menghadapi tantangan ini, teknik *word embedding* dapat digunakan untuk mengkodekan kata-kata ke dalam ruang vektor berdimensi rendah yang memungkinkan penangkapan hubungan semantik antar kata, sehingga model dapat memahami makna dan konteks kata [7]. Salah satu teknik *word embedding* yang dapat digunakan adalah FastText. Penelitian sebelumnya menunjukkan bahwa FastText dapat memberikan hasil yang lebih baik dibandingkan dengan teknik lainnya [8]. Dengan demikian, menggunakan Random Forest bersama FastText dalam analisis sentimen pemilihan umum di Indonesia memadukan kekuatan keduanya untuk meningkatkan performa model. Random Forest efektif dalam mengatasi kerumitan dan variabilitas data serta mengurangi *overfitting*, sementara FastText mengatasi data teks tidak terstruktur dan masalah *out-of-vocabulary (OOV)* dengan menangkap hubungan semantik antar kata.

### Topik dan Batasannya

Penelitian ini mengambil topik sentimen analisis di media sosial X pada topik pemilu 2024 di Indonesia. Unggahan yang diambil dari bulan November 2023 sampai dengan April 2024. Unggahan diambil menggunakan kata kunci pemilu dan pemilihan umum serta tagar #Pemilu2024. Selanjutnya untuk melakukan *word embedding* menggunakan FastText dengan korpus dari artikel wikipedia Bahasa Indonesia.

### Tujuan

Penelitian ini bertujuan untuk mengeksplorasi pengaruh penggunaan metode Random Forest dalam meningkatkan performa dalam analisis sentimen terkait pemilihan umum tahun 2024 di Indonesia. Metode Random Forest dipilih karena kemampuannya dalam menangani klasifikasi data yang kompleks dan beragam, yang sering ditemui dalam analisis sentimen terhadap isu politik. Jumlah data yang digunakan pada penelitian ini berjumlah 5000 baris. Penelitian ini juga bertujuan untuk menginvestigasi dampak penggunaan FastText dalam mengelola data teks yang tidak terstruktur dalam analisis sentimen yang sama. FastText dipilih karena kemampuannya dalam memahami konteks bahasa yang kompleks dan kemampuannya dalam menangani teks dalam berbagai bahasa, termasuk bahasa Indonesia. Dengan demikian, penelitian ini diharapkan dapat memberikan wawasan yang berharga dalam meningkatkan akurasi analisis sentimen terkait pemilihan umum tahun 2024 di Indonesia melalui penerapan metode Random Forest dan FastText.

### **Organisasi Tulisan**

Penyusunan tulisan dimulai dengan melakukan tinjauan pustaka yang mencakup analisis sentimen dalam berbagai topik. Kemudian, akan dijelaskan metodologi yang digunakan dalam penelitian ini. Tahap selanjutnya akan dilakukan pembahasan evaluasi hasil penelitian. Terakhir, akan disampaikan kesimpulan dan saran.