

# Buzzer Account Detection in Political Hate Tweets Using IndoBERT and Ensemble Learning: Case Study of the Indonesian Presidential Election 2024

Fizio Ramadhan Herman<sup>1\*</sup>, Ade Romadhony<sup>2</sup>

<sup>1,2</sup>*School of Computing, Telkom University,  
Bandung, Indonesia*

\*[fizioramadhan@gmail.com](mailto:fizioramadhan@gmail.com)

## Abstract

The Indonesian Presidential Election of 2024 has seen a widespread use of social media such as Twitter for political campaigning and discussion. However, this has also enabled the spread of hate speech from buzzer accounts that are created to influence public opinions. This study implements a machine learning approach to classify buzzer accounts that are spreading hate speeches during the presidential election period. By utilizing IndoBERT for hate speech classification and a traditional machine learning model to classify buzzer accounts. This study analyzes 62,341 tweets for hate speech classification and 961 accounts for buzzer account classification. Our implementation of IndoBERT achieved a strong performance with 91.12% of precision and recall, and 91.19% accuracy and F1-score in hate speech classification. While for buzzer account classification, we compared Decision Tree, Random Forest, and XGBoost, with Decision Tree achieving the highest performance of 64% precision, recall, accuracy, and F1-Score. Our results demonstrate the effectiveness of combining deep learning for hate speech classification with traditional machine learning for buzzer account classification, contributing to the development of more effective content filtering for election discourse on social media.

**Keywords:** buzzer detection, ensemble learning, IndoBERT, presidential election, sentiment analysis, social media

## I. INTRODUCTION

The 2024 Indonesian Presidential Election has captured the attention of Indonesians for these past few months. Since late November 2023, there has been a growing discussion regarding this topic, especially on the internet, due to the political party campaigns have started. Social media such as Twitter and Facebook, play a significant role in politics [1] by providing spaces for individuals to share views and engage in political discussions [2]. As the campaign season intensifies, the use of social media by Indonesian political parties to promote an advertise candidates is one example that demonstrates the role that social media occupies in politics. However, the open nature of Twitter can both encourage informed political discourse and create opportunities for the dissemination of hate speech, which could potentially shape public opinions about presidential candidates.

Increasing amount of political hate speech on Twitter has emerged as a notable concern. Based on a study conducted by Mozafari et. al [3], hate speech is commonly defined as any form of communication that involves the criticism of an individual or a group based on some characteristics such as gender, nationality, religion, race, etc. It is important to consider the potential consequences of this issue, as it could lead to division within communities, incite violence, and undermine the integrity of fair elections. In order to tackle this problem, the proposed research seeks to develop a robust deep learning approach to detect and identify buzzer accounts that are specifically created or used to spread political hate speech during the 2024 Indonesian Presidential Election.

While previous studies have explored the role of Twitter in political polarization [2] the influence of e-campaigns on voters [4] [5] and the detection of malicious social bots [6], there is a need for a comprehensive study focused on the detection of buzzer accounts actively engaged in spreading political hate speech during the Indonesian Presidential Election. Previous research has also demonstrated significant progress in buzzer detection. Suciati et al. [1] evaluated multiple ensemble learning algorithms for buzzer detection, achieving 62.3% accuracy and 61.3% precision using AdaBoost, while Ibrahim et al. [8] developed an automatic buzzer detection system combined with sentiment analysis to predict election results with a mean absolute error of 0.61%. However, while these studies have broadened our understanding of buzzer detection and political sentiment analysis separately, these leaves a gap in the integrated study of hate speech-spreading buzzer accounts. A detailed review of related studies is provided in Section 2.

This research aims to bridge this gap by leveraging ensemble learning for buzzer detection and deep learning approaches, specifically IndoBERT, for hate speech classification. Through collaboration with *NoLimit Indonesia*, we analyze a dataset of 62,341 tweets and 961 accounts to evaluate our proposed approach in identifying buzzer accounts that spread hate speeches during election periods. The study focuses on developing a framework that integrates ensemble learning and IndoBERT, assessing its effectiveness on a real Twitter dataset collected during the election period, and establishing quantitative benchmarks for identifying hate speech-spreading buzzer accounts. Additionally, the dataset used in this research is expected to serve as a resource for future studies on hate speech classification and buzzer account detection.

## II. LITERATURE REVIEW

This section will analyze the recent research carried out in the field of buzzer account detection and hate speech classification. We will study this body of work by splitting it into several components to conduct a comprehensive review.

### A. Buzzer Account

1) *Definition of Buzzer Account*: According to a study conducted by Rismi Juliadi in 2017 [9], a "buzzer" on Twitter is anticipated to spread information that will have an impact on their followers. The widespread adoption of buzzers can be understood through the lens of the Diffusion of Innovation Theory by Everett Rogers [10]. This theory explains that communicators or messengers who get information from the mass media have the

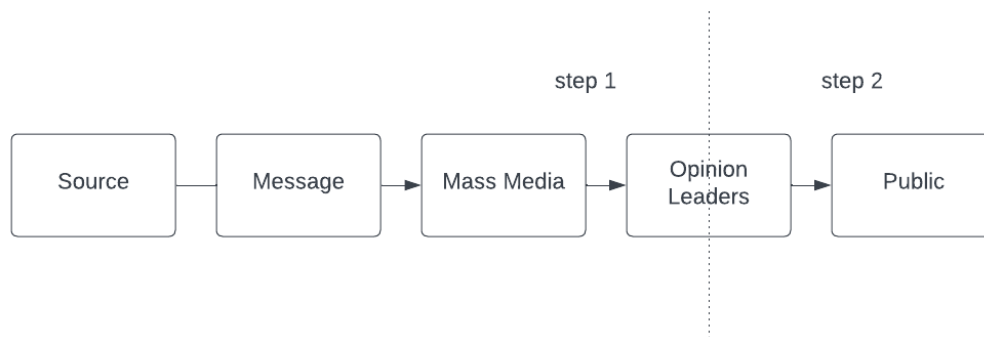


Fig. 1. Two Step Flow Model (Katz & Lazarsfeld, 1955)

ability to influence those around them. This process of influence is particularly well-illustrated through the following model in Fig. 1.

As described in the two-step flow model, the communication process is divided into two stages. The initial source communicates messages through mass media. Then, the message is received by the opinion leader. During the second stage, the opinion leader relays messages from the mass media to the public. These messages may already include the opinion leader's interpretations and responses [11].

2) *Characteristics of Buzzer Account*: Within the social media landscape, two primary categories of buzzer accounts exist: automated and human-operated [8]. Automated buzzers, essentially computer programs, are programmed to react on specific triggers. These programs can be designed for broadcast spamming, reaching a wide audience, or context-aware spamming, targeting specific groups [6]. As stated by Ibrahim in 2015 [8], an automated buzzer might automatically re-tweet posts containing keywords like "pemilu" (election). Human-operated buzzers, in contrast, are individuals who are paid or driven by strong allegiance to a particular candidate. These accounts frequently post content and actively engage with related tweets, potentially influencing public opinion. Notably, local media reports suggest a possible link between candidates and paid social media users who promote specific campaigns [9]. This highlights the potential for manipulation within the online political sphere.

### B. *Political Hate Speech*

In today's political landscape, social media is commonly used to express opinions and for campaigning. Platforms like Twitter and Facebook can be utilized to communicate one's views in favor of a political party or political candidate [12]. However, while the creation of online communities is not inherently a significant issue, it can also become concerning when politically active individuals utilize these platforms to spread radicalism [2]. This radicalization can be facilitated by the widespread use of provocative language, as defined by Seigner et al. [13], which includes aggressive expressions, disobedience towards regulations, swearing, and antagonistic behavior.

### C. *Related Works: Buzzer Account Detection and Hate Speech Classification*

Several studies have explored methods for detecting buzzer accounts on Twitter. Ibrahim et al. [8] conducted research on buzzer detection for the 2014 Indonesian Presidential Election. They employed a machine learning approach to develop a computational model for classifying Twitter users as buzzers or non-buzzers. Their method achieved an accuracy of 86% in detecting buzzer accounts. Suciati et al. [1] investigated buzzer detection for the 2019 Indonesian Presidential Election. They examined the performance of four machine learning algorithms: AdaBoost (AB), Gradient Boosting (GB), Extreme Gradient Boosting (XGB), and Histogram-based Gradient Boosting (HGB). Their results showed that AdaBoost achieved the best accuracy 62.3% and precision 61.3% with 25 features, while XGB attained the highest recall 67.9%.

Hate speech classification is another important task related to social media analysis, particularly in the context of elections. Geni et al. [19] conducted sentiment analysis on Indonesian tweets related to the 2024 election using various models, including IndoBERT, which achieved the highest accuracy of 83.5%. This method can be used for tasks like hate speech classification. While significant progress has been made in both buzzer detection and hate speech classification independently, there remains a critical gap in addressing the intersection of these challenges specifically, the detection of buzzer accounts that spread hate speech during election periods.

## III. RESEARCH METHOD

### A. *System Overview Description*

To develop a model that can identify whether a Twitter account that posts hate speech is a buzzer or non-buzzer, a dataset is required. Through collaboration with NoLimit Indonesia, we gathered a dataset of 62,341 tweets along with their corresponding account features. The system implementation begins by using IndoBERT to classify these tweets for hate speech content. If a tweet is not identified as hate speech, it is disregarded. For tweets classified as containing hate speech, we then examine the frequency of the account's appearance in the dataset. Based on the characteristic behavior of buzzer accounts, which involve high posting frequencies, we

set a threshold of 10 appearances in the dataset to differentiate potential buzzer accounts from regular users. If an account appears more than 10 times in the gathered dataset, we proceed to classify it using ensemble learning methods (Decision Tree, Random Forest, and XGBoost) to determine whether it is a buzzer or non-buzzer account. If the account appears less than 10 times, it is ignored from further analysis. The final output consists of a classification result indicating whether each qualifying account is categorized as a buzzer or non-buzzer based on their posting patterns (Fig. 2).

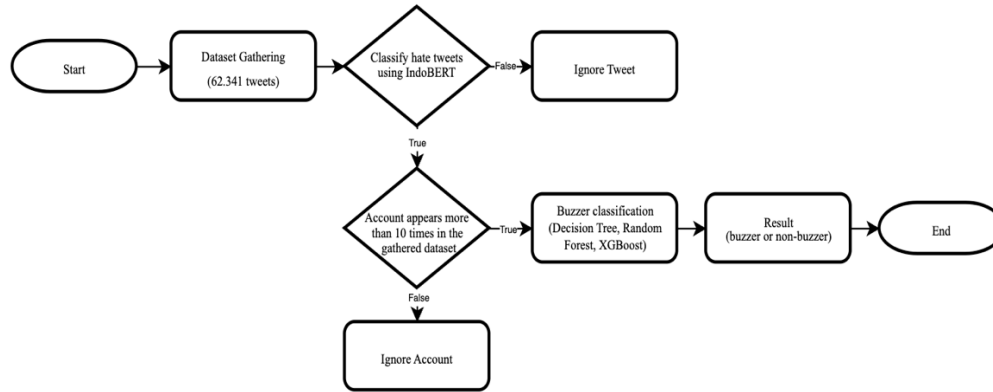


Fig. 2. System Overview

**B. Dataset Preparation**

1) *Aimed Dataset*: The dataset aimed for this research composed of social media posts on Twitter where the post includes hashtags that have the potential to incite hate speech related to the 2024 Indonesian Presidential Election, along with accompanying account features. Here in Fig. 3 are some cases for preferred tweets for the dataset. For the account features, we have selected nine specific features based on similar research conducted by Suciati *et. al* [1]. These features will be presented in Table 2.

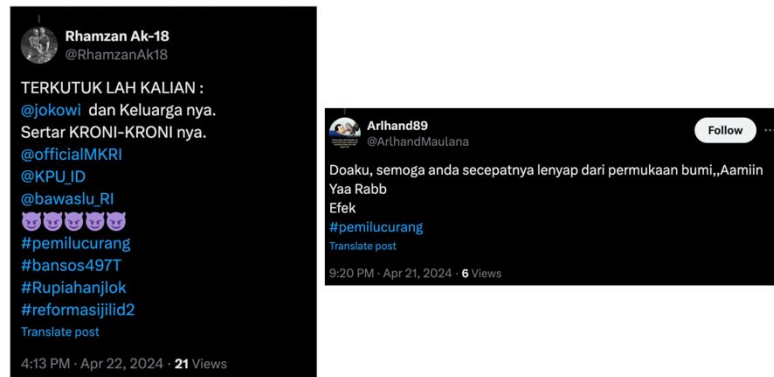


Fig. 3. Hate Speeches on Twitter

Based on the research conducted by Suciati *et. al* [1], there were originally 64 features before feature selection was performed. They used mutual information (MI) to avoid the curse of dimensionality. Among these features, only 11 were utilized due to possessing the highest MI scores. However, for our research, we are constrained to utilizing only 9 of these features, as obtaining the remaining data is unfeasible.

2) *Data Collection Method*: The data for this study were collected through a collaboration with *NoLimit Indonesia*, a company that provides social media analysis as a service. *NoLimit Indonesia* supplied all the required data, such as tweet text, metadata and account features of a user. Their team handle the data gathering process, allowing us to focus on the analysis and modelling work.

TABLE I  
 DESCRIPTION OF ACCOUNT FEATURES

No	Features Name	Types of Data	Description
1	from_username	string	Unique Twitter handle/username of the account (e.g., @username)
2	followers	integer	Total number of current followers for the account
3	account_age	Integer	Account age in years (calculated as current date minus account creation date)
4	post_frequency_per_year	float	Average number of posts per year
5	post_frequency_per_month	float	Average number of posts per month
6	post_frequency_per_day	float	Average number of posts per day
7	content	text	Full text content of the tweet/post, including any hashtags and mentions
8	final_sentiment	boolean	Binary classification of content (0: non-hate speech, 1: hate speech)
9	is_buzzer	boolean	Binary classification of account type (0: non-buzzer, 1: buzzer account)

3) *Data Labeling*: For data labeling, *NoLimit Indonesia* provided a dataset of 62,341 tweets that were already labeled with sentiments (hate/non-hate). For buzzer account labeling, we can determine whether an account is a buzzer based on the definition and characteristics of buzzers as explain in Section 2, where buzzers are identified through several key characteristics:

- a. *Information Spreading Patterns*: Based on the Two-Step Flow Model and Diffusion of Innovation Theory [10], buzzer accounts act as opinion leaders who actively spread and amplify messages to influence their followers.
- b. *Account Type*: Following Ibrahim's research [8], buzzer accounts can be categorized into two automated buzzers and human-operated buzzers. Automated buzzer are a programmed accounts that automatically react to specific triggers, such as retweeting posts containing keywords like “pemilu” (election) or any other specific keywords.
- c. *Behavioral Patterns*: We also identify buzzer accounts by examining the frequency and consistency of tweet posting, pattern of message, and connection to organized political campaign or one of specific presidential candidates.

4) *Data Preprocessing*: Data preprocessing conducted for this research dataset includes several steps, namely:

- a. *Text Cleaning*: involves removing any irrelevant characters or symbols from the text data to ensure consistency and improve the quality of the analysis. This typically includes removing punctuation marks, special characters, mentions/urls, leading/trailing whitespace and converting all letters to lowercase.
- b. *Tokenization*: the process of breaking down a text into smaller units called tokens. These tokens can be words, phrases, or even individual characters. Tokenization facilitates further analysis by providing a structured representation of the text, enabling algorithms to process and understand the underlying content more efficiently.
- c. *Slang Words Removal*: This process involves identifying and replacing slang words in the text data with their corresponding standard Indonesian translations using a slang word dictionary. The dictionary contains a collection of informal, non-standard words and phrases commonly used in casual conversations, especially among youth and in online communication. For example, "ak" is a slang word for "saya"(I), "bkn" means "bukan"(not), and "jgn" is short for "jangan"(don't). By mapping these slang words to their standard equivalents, we can standardize the text data, reduce the variability introduced by informal language usage, and improve the performance of subsequent natural language processing tasks.

### C. Proposed Method

1) *IndoBERT for Hate Speech Classification*: IndoBERT is a transformer-based pretrained language model designed specifically for the Indonesian language. It is built on the BERT architecture, which utilizes bidirectional encoding to capture contextual information from both directions in a sentence. For hate speech

classification, IndoBERT can be fine-tuned by adding a classification head, enabling it to learn task-specific patterns from labeled data. IndoBERT is chosen for hate speech classification because its language-specific design ensures optimal understanding of Indonesian linguistic nuances, and its bidirectional encoding captures complex contextual information necessary for identifying hate speech.

2) *Ensemble Learning for Buzzer Detection*: Ensemble learning combines predictions from multiple base models to improve overall classification performance. This method is particularly effective for buzzer detection, as it leverages diverse classifiers to handle various data patterns. Common techniques include bagging (e.g., Random Forest) and stacking, which combines outputs from base models through a meta-classifier. Ensemble learning is chosen for buzzer detection due to its abilities in handling noisy and diverse data.

#### D. Evaluation Method

The evaluation of the hate speech detection and buzzer account classification task involves assessing the models' performance using a variety of evaluation metrics tailored to the specific objectives of the research. The evaluation metric is used in two distinct stages: the training stage, which involves the learning process, and testing stage. During the testing phase, an assessment metric was utilized to measure the effectiveness of the classifier when applied to new, unseen data [18]. There are four usual evaluation metrics that are often used in a classification task, which are precision, recall, F1 score, and accuracy.

1) *Precision*: Precision is used to measure the positive patterns that are correctly predicted from the total predicted patterns in a positive class. The formula for calculating precision is shown in Equation 1. In this formula,  $tp$  stands for true positive and  $fp$  stands for false positive. The precision formula calculates the ratio of true positive predictions ( $tp$ ) to the total number of positive predictions ( $tp + fp$ ). A higher precision value indicates a lower number of false positives.

$$Precision = \frac{tp}{tp + fp} \quad (1)$$

2) *Accuracy*: Accuracy is used as an evaluation metric in classification tasks because it provides a straightforward measure of a model's overall performance. This metric is particularly useful when the classes are balanced and the cost of false positives and false negatives is similar. The formula for calculating accuracy is shown in Equation 2. In this accuracy formula  $tp$  stands for true positive,  $tn$  stands for true negative,  $fp$  stands for false positive, and  $fn$  stands for false negative. The accuracy formula calculates the ratio of correctly predicted instances (both true positives and true negatives) to the total number of instances. A higher accuracy value indicates a higher number of correct predictions (both positives and negatives) relative to the total number of predictions made.

$$Accuracy = \frac{tp + tn}{tp + fp + tn + fn} \quad (2)$$

3) *Recal*: Recall is used to measure the fraction of positive patterns that are correctly classified. The formula for calculating recall is shown in Equation 3. In the recall formula,  $tp$  stands for true positive and  $fn$  stands for false negative. The recall formula calculates the ratio of true positive predictions to the total number of actual positive instances, which is the sum of true positives and false negatives. A higher recall value indicates that the model is effective at identifying most of the positive instances in the dataset.

$$Recall = \frac{tp}{tp + fn} \quad (3)$$

4) *F1 Score*: The F1 score is used as an evaluation metric in classification tasks because it provides a balanced measure of a model's performance by combining precision and recall into a single metric. The formula for calculating F1 score is shown in Equation 4. This metric is particularly useful when there is an imbalance between the number of positive and negative samples, or when the costs of false positives and false negatives are significantly different..

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

## IV. RESULTS AND DISCUSSION

### A. Dataset Characteristics

Our analysis encompassed two primary datasets. The first dataset consisted of 62341 tweets for hate speech detection, with 1249 (66.17%) classified as hate speech and 21094 (33.83%) as non-hate speech. For buzzer account detection, after applying our filtering criteria of minimum 10 tweets per account, we analyzed 961 accounts, comprising 580 (60.35%) buzzer accounts and 381 (39.65%) non-buzzer accounts. Fig. 4 presents the detailed distribution of both datasets, showing the class proportions for hate speech and buzzer account detection.

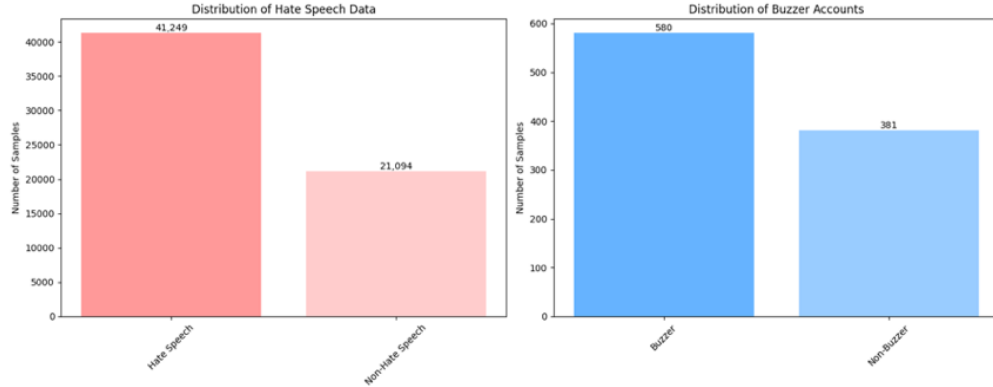


Fig. 4. Dataset Distribution

### B. Experimental Setup

For hate speech classification, we implemented the IndoBERT model. Additionally, we employed three machine learning models for buzzer account detection: Decision Tree, Random Forest, and XGBoost. Table 2 presents the complete hyperparameter configurations for each model.

TABLE II  
 MODEL HYPERPARAMETER CONFIGURATION

Model	Parameter	Value
IndoBERT	Training Batch Size	32
	Evaluation Batch Size	64
	Number of Epochs	4
	Learning Rate	0.00002
	Weight Decay	0.02
	Warmup Steps	500
	Warmup Ratio	0.15
	Gradient Accumulation Steps	1
Decision Tree	Max Depth	5
	Min Samples Split	15
	Min Samples Leaf	4
Random Forest	Number of Estimators	200
	Max Depth	10
	Min Samples Split	15
	Min Samples Leaf	2
XGBoost	Number of Estimators	100
	Learning Rate	0.01
	Max Depth	5
	Subsample	0.9
	Gamma	0.2
	Min Child Weight	3

C. Hate Speech Classification Result

The IndoBERT model demonstrated a robust performance with 91.12% precision, recall, accuracy, and F1-score. Analysis of the confusion matrix reveals strong discrimination ability, with 7,881 true negatives and 3,490 true positives, while maintaining relatively low false positives (446) and false negatives (652). These metrics demonstrate the model's effectiveness in detecting hate speech tweets. The balanced score across all metrics indicates consistent performance. With only 1,098 misclassifications out of 12,469 total samples, this suggest that IndoBERT effectively learned the linguistic pattern for hate speech detection in Indonesian political discourse. These results are visualized in Fig. 5 and Fig. 6 which illustrate the model's performance metrics and classification distribution respectively.

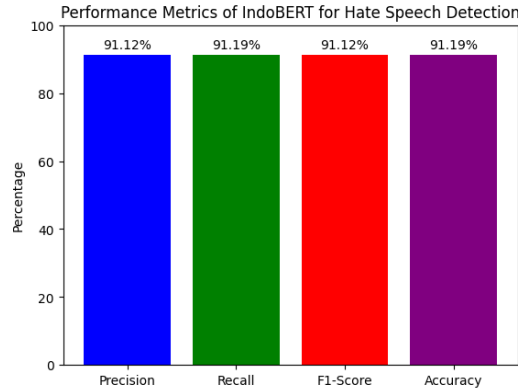


Fig. 5. Performance Metrics (Precision, Recall, F1-Score, and Accuracy) of IndoBERT

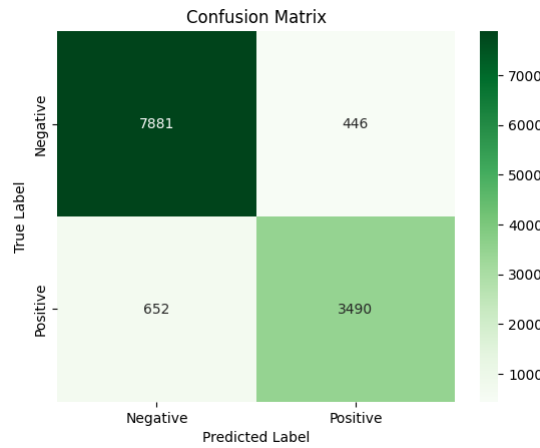


Fig. 6. Confusion Matrix of IndoBERT

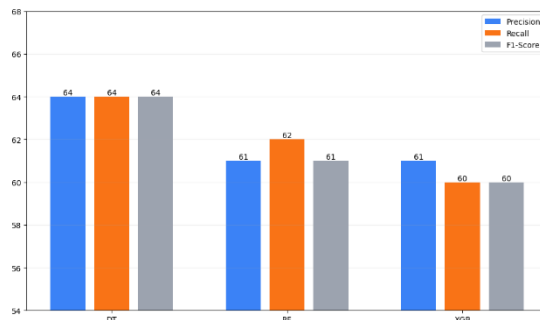


Fig. 7. Model Performances Comparison

#### *D. Buzzer Account Classification Result*

For buzzer account detection, we evaluated three machine learning models: Decision Tree (DT), Random Forest (RF), and XGBoost (XGB). The Decision Tree classifier achieved the highest performance with 64% precision, recall, and F1-Score. Random Forest and XGBoost performed slightly lower, achieving 61% and 60% F1-scores respectively. These results reflect the inherent complexity of buzzer account detection, where sophisticated account behaviors can make classification challenging. The performance metrics are visualized in Fig. 7.

#### *E. Discussion*

IndoBERT demonstrated excellent performance in hate speech classification, achieving 91.12% accuracy with balanced precision and recall scores. This indicates the model's strong capability in understanding Indonesian political discourse and distinguishing hate speech patterns. However, there is still room for improvement through several approaches:

1. Expanding the dataset to include more diverse forms of political hate speech.
2. Incorporating contextual features to better understand implicit forms of hate speech.

Meanwhile, the buzzer account detection achieved moderate performance (64% accuracy with Decision Tree), highlighting the need for a more comprehensive dataset. The current dataset of 961 accounts, while providing a foundation for classification, appears insufficient to capture the full complexity of buzzer behavior patterns. Future work should focus on:

1. Collecting a larger, more diverse dataset of buzzer accounts.
2. Developing more sophisticated feature engineering techniques to capture evolving buzzer behaviors.
3. Exploring other machine learning methods to better handle the variety of buzzer account strategies.

## V. CONCLUSION

This research demonstrates the significant presence of buzzer accounts spreading hate speech on Twitter during the 2024 Indonesian Presidential Election, highlighting the urgent need for effective detection mechanisms before implementing countermeasures. Our analysis reveals that out of 62,341 tweets analyzed, approximately 66.17% contained hate speech, indicating the substantial scale of coordinated negative messaging in political discourse.

By leveraging advanced natural language processing and machine learning techniques, we have developed a comprehensive framework for creating healthier social media environments during election periods. Our implementation of IndoBERT for hate speech classification achieved robust performance with 91.12% precision and F1-score, and 91.19% accuracy and recall, demonstrating its effectiveness in identifying harmful content in Indonesian political discourse. For buzzer account detection, the Decision Tree classifier emerged as the most effective among traditional machine learning models, achieving 64% across all metrics (precision, recall, F1-score, and accuracy).

These results validate the potential of combining deep learning approaches for sentiment analysis with traditional machine learning for behavior classification in addressing social media manipulation. While the performance metrics indicate promising capabilities, they also highlight areas for future improvement, particularly in buzzer account detection through expanded datasets and enhanced feature engineering.

## DATA AND COMPUTER PROGRAM AVAILABILITY

Data and program used in this paper can be accessed in the following site: [github.com](https://github.com)

## REFERENCES

- [1] A. Suciati, A. Wibisono and P. Mursanto, "Twitter Buzzer Detection for Indonesian Presidential Election," in *2019 3rd International Conference on Informatics and Computational Sciences (ICICoS)*, 2019.
- [2] M. Conover, J. Ratkiewicz, M. Francisco, B. Goncalves, F. Menczer and A. Flammini, "Political Polarization on Twitter," in *Proceedings of the International AAAI Conference on Web and Social Media*, 2021.
- [3] M. Mozafari, R. Farahbakhsh and N. Crespi, "A BERT-Based Transfer Learning Approach for Hate Speech Detection in Online Social Media," in *Complex Networks and Their Applications VIII. COMPLEX NETWORKS 2019*, Springer, Cham, 2020.
- [4] I. Satria, A. Kurnia and Y. Nurhadryani, "Influence of presidential candidates e-campaign towards voters in 2014 presidential election in Bogor City," in *2014 International Conference on Advanced Computer Science and Information System*, 2014.
- [5] S. Kiyohara, "A Study on How Technological Innovation Affected the 2008 U.S. Presidential Election: Young Voters' Participation and Obama's Victory," in *2009 Ninth Annual International Symposium on Applications and the Internet*, 2009.
- [6] M. Latah, "Detection of malicious social bots: A survey and a refined taxonomy," *Expert Systems with Applications*, vol. 151, p. 11383, 2020.
- [7] S. B. Kotsiantis, I. Zaharakis and P. Pintelas, "Supervised machine learning: A review of classification techniques," *Emerging artificial intelligence applications in computer engineering*, vol. 160, pp. 3-24, 2007.
- [8] M. Ibrahim, O. Abdillah, A. F. Wicaksono and M. Adriani, "Buzzer Detection and Sentiment Analysis for Predicting Presidential Election Results in a Twitter Nation," in *2015 IEEE International Conference on Data Mining Workshop (ICDMW)*, 2015.
- [9] R. Juliadi, "The construction of Buzzer identity on social media (A descriptive study of Buzzer identity in Twitter)," in *Proceedings of the 3rd International Conference on Transformation in Communications 2017 (IcoTiC 2017)*, 2017.
- [10] E. M. Rogers, *Diffusion of Innovations*, Free Press, 2003.
- [11] S. Choi, "The Two-Step Flow of Communication in Twitter-Based Public Forums," *Social Science Computer Review*, vol. 33, pp. 696-711, 2015.
- [12] T. A. a. B. I. Arafat, R. Mahendra and D. A. Salehah, "Demographic Analysis of Candidates Supporter in Twitter During Indonesian Presidential Election 2019," in *2020 International Conference on ICT for Smart Society (ICISS)*, 2020.
- [13] B. D. C. Seigner, H. Milanov, E. Lundmark and D. A. Shepherd, "Tweeting like Elon? Provocative language, new-venture status, and audience engagement on social media," *Journal of Business Venturing*, vol. 38, 2023.

- [14] H. S. Iqbal, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Computer Science*, pp. 1-6, 2021.
- [15] Z. M., Ehsan-Ul-Haq, R. S., J. K., S. and Hussain, "Improving Hate Speech Detection of Urdu Tweets Using Sentiment Analysis," *IEEE Access*, vol. 9, pp. 84296-84305, 2021.
- [16] E. del Valle and L. de la Fuente, "Sentiment analysis methods for politics and hate speech contents in Spanish language: a systematic review," *IEEE Latin America Transactions*, vol. 21, pp. 408-418, 2023.
- [17] T. J. M. and A. S., "Buzzer Detection on Twitter Using Modified Eigenvector Centrality," in *2018 5th International Conference on Data and Software Engineering (ICoDSE)*, 2018.
- [18] M. Hossin and S. M.N, "A Review on Evaluation Metrics for Data Classification Evaluations," *International Journal of Data Mining & Knowledge Management Process*, vol. 5, pp. 1-11, 2015.