JOISTECH: Journal of Information System and Technology
ISSN 3063-9778 (Media Online)
Vol 01, No 02, Desember 2024, Hal 67-75

Sentiment Analysis on Public Opinions Regarding the 2024 Regional Elections Using Long Short-Term Memory (LSTM), Random Forest, and Naive Bayes

Supriadi Panggabean[1], Asmalia Junika[2]

[1] Faculty of Science and Technology, System and Information Technology, Darunnajah University, Jakarta, Indonesia
[2] Faculty of Science and Technology, Actuarial Science, Darunnajah University, Jakarta, Indonesia
Email: [1]supriadipanggabean@darunnajah.ac.id, [2]asmaliajunika5@gmail.com

Abstract− The 2024 Regional Elections (Pilkada) have become a hot topic, sparking extensive discussions on social media, particularly Twitter/X. This study aims to analyze public sentiment regarding the elections using three primary approaches: Long Short-Term Memory (LSTM), Random Forest, and Naive Bayes. The data undergoes six pre-processing stages and is represented using Word2Vec. The findings indicate that LSTM achieves the best performance with an accuracy of 94.21%, followed by Random Forest (92.78%) and Naive Bayes (89.43%). Further analysis reveals that negative sentiments are predominantly influenced by policies and candidates' performance. This research provides insights into the effectiveness of deep learning methods compared to traditional techniques in social media sentiment analysis.

Keywords: Sentiment analysis, 2024 Regional Elections, LSTM, Random Forest, Naive Bayes, Word2Vec, Twitter/X

1. INTRODUCTION

The 2024 Regional Head Elections (Pilkada) represent a crucial moment in Indonesia's political dynamics. As a five-yearly event that involves various stakeholders and layers of society, Pilkada serves not only as a political competition but also as a reflection of public participation and the evolving public opinion (Yunita Simatupang, 2024). In the digital era, social media has become a primary medium for expressing aspirations and political views, creating opportunities to better understand public opinion through digital data analysis (Wardaniah et al., 2024).

Twitter, as one of the most widely used social media platforms, provides a rich source of data for sentiment analysis. Its unique characteristics, allowing users to express opinions concisely and interactively, make it an ideal platform to explore public perspectives on political issues (Panggabean et al., 2022). The data generated from Twitter can offer real-time insights into the emotions and perceptions of the public regarding candidates, political parties, and policies promoted during the Pilkada campaign (Murthy, 2024). Hence, Twitter-based sentiment analysis has become a relevant approach to support decision-making in a political context (Carpenter et al., 2024).

This study builds on previous research that applied the Support Vector Machine (SVM) method for sentiment analysis on Twitter data. SVM is known as an effective machine learning algorithm for text-based data classification (Bilal et al., 2024). Classification is a type of supervised learning in machine learning, where the goal is to predict the category or class of a given data point based on its features. In classification tasks, the model is trained on a labeled dataset, meaning the input data is already paired with the correct output labels. The model learns patterns from this data so that it can predict the class of new, unseen data (Panggabean et al., 2024). Machine learning is a branch of artificial intelligence that focuses on developing algorithms that allow computers to learn from data and make predictions or decisions without being explicitly programmed (Panggabean & Wahyu Joko Saputro, 2024). SVM works by dividing data into different classes while maximizing the margin between these categories. However, with the advancement of technology, deep learning-based methods have gained more attention due to their ability to capture complex patterns in text data that traditional algorithms cannot reach (Kusuma et al., 2023). Deep learning methods, which use deeper and more complex artificial neural network structures, enable systems to recognize more intricate non-linear relationships within data, such as in sentiment analysis, which requires a deeper understanding of the context of the analyzed text (Muchamad Bachram Shidiq et al., 2023).

In this research, a deep learning approach using the Long Short-Term Memory (LSTM) model is introduced for sentiment analysis on Twitter data. LSTM, a variant of Recurrent Neural Networks (RNN), has the capability to process long sequences of data while retaining temporal context, which is expected to provide more accurate classification results compared to traditional methods (Ahmad et al., 2024). This ability to maintain context over long sequences of text is particularly valuable in sentiment analysis, where the meaning of a message often depends on previous statements or the flow of conversation. Unlike other machine learning algorithms, LSTM can remember information from earlier in the sequence and use it to inform decisions at later stages, making it an ideal choice for analyzing social media posts, which often involve lengthy and complex interactions (Idham et al., 2022). Furthermore, LSTM's capacity to handle variable-length inputs ensures that it can adapt to different types of tweets, from short reactions to more elaborate opinions. Additionally, to evaluate the performance of the deep learning model, traditional algorithms such as Random Forest and Naive Bayes are used for comparison. These methods, though effective in many cases, are often less adept at capturing complex relationships in sequential data compared to LSTM. Random Forest, for example, builds multiple decision trees and can be very accurate in many tasks, but it may struggle to capture the nuances of time-dependent or context-dependent information (Ige et al., 2024). Naive Bayes, on the other hand, assumes independence between features, which can limit its ability to understand the dependencies within a sequence of words or phrases (M. Afriansyah et al., 2024). By testing these traditional models alongside LSTM, the research aims to highlight the advantages of deep learning in sentiment analysis tasks.

By introducing this new approach, the study aims to contribute significantly to the development of sentiment analysis methods based on social media data. Social media platforms, especially Twitter, have become key sources of public opinion, making it crucial to develop accurate methods for analyzing and understanding the vast amounts of information generated daily. As public sentiment can heavily influence political outcomes, gaining insights from social media discussions can play a pivotal role in shaping political strategies and decisions. The findings of this research are not only relevant in the context of the 2024 Pilkada but also have the potential to be applied to other political studies in the future. For example, similar methods could be used to analyze sentiments during national elections, or to monitor the public's reaction to government policies and decisions. Additionally, the techniques developed in this study could be adapted to other domains, such as marketing, healthcare, or public opinion on social issues. Furthermore, this analysis is expected to assist policymakers and political actors in gaining deeper insights into public opinion, thus supporting more strategic decision-making processes. Understanding public sentiment can help candidates tailor their campaigns to address voter concerns more effectively, and enable policymakers to respond to public needs in a more timely and accurate manner. Moreover, by gaining a deeper understanding of the factors driving sentiment, political actors can better manage crises and capitalize on opportunities for support, leading to more informed and impactful decisions. Ultimately, this research aims to bridge the gap between data science and political analysis, creating a powerful tool for understanding and navigating the complex landscape of public opinion in the digital age.

This study aims to answer several key questions, namely how the performance of LSTM compares to traditional methods in sentiment analysis, and what are the main factors influencing negative sentiment related to the Regional Elections (Pilkada). In this context, the research evaluates the advantages of LSTM in handling sequential data and retaining temporal context, which is expected to provide more accurate classification results compared to traditional methods such as Random Forest and Naive Bayes. Additionally, this study focuses on identifying factors that may influence negative sentiment, such as policies and candidates' performance in the Pilkada. By understanding these factors, it is hoped to provide deeper insights into the dynamics of public opinion regarding the Pilkada, as well as assist policymakers and political actors in responding more effectively to the concerns and expectations of the public.

1.1 Sentiment Analysis

Sentiment analysis is the process of classifying opinions into categories such as positive, negative, or neutral based on text. This technique is crucial for understanding public perception on various issues and can reveal underlying emotions or biases in large-scale datasets (Panggabean et al., 2022). By analyzing sentiments expressed in social media posts, reviews, or news articles, it is possible to gain insights into how people feel about specific topics or events. This information can be valuable for businesses, politicians, or researchers who need to monitor public opinion or make informed decisions. Sentiment analysis also plays an important role in identifying trends and shifts in opinions over time.

1.2 Data Pre-processing

This process involves six stages: case folding, data cleaning, tokenization, slang normalization, stemming, and filtering. Case folding ensures that all text is converted to lowercase, making it uniform and easier to process (Liang et al., 2024). Data cleaning removes irrelevant characters or noise, such as punctuation marks and special symbols. Tokenization splits text into individual words or tokens, which are the basic units for analysis. Slang normalization converts informal language or slang into standard words, improving the model's ability to understand diverse expressions. Stemming reduces words to their root form, enabling the model to recognize related terms as the same, while filtering removes stopwords that do not contribute meaningful information to the analysis (Budi et al., 2022). This pre-processing ensures that the data is ready to be processed by machine learning algorithms.

1.3 Methods Used

1.3.1 LSTM

Long Short-Term Memory (LSTM) is a type of Recurrent Neural Network (RNN) that excels in handling long sequences of data by retaining important temporal context. This is particularly useful for sentiment analysis on social media, where the context of a conversation or sequence of posts is essential to understanding the sentiment. LSTM has the ability to remember past inputs and use that information to influence predictions for future inputs (Zheng et al., 2024).

1.3.2 Random Forest

This ensemble algorithm combines multiple decision trees to improve the accuracy of predictions. Each tree in the forest provides a vote on the classification, and the final result is determined by majority voting. Random Forest is highly effective in reducing overfitting and can handle large datasets with many features (Asamoah et al., 2024).

1.3.3 Naive Bayes

A probabilistic classification algorithm that is efficient for simple text classification tasks. It assumes that the features are independent of each other, which simplifies the calculation of probabilities. Despite its simplicity, Naive Bayes often performs well in text classification tasks, especially when the features (words) are conditionally independent (Monteverde-Suárez et al., 2024).

1.4 Word Representation with Word2Vec

Word2Vec is used to generate word vectors that represent the semantic relationships between words in a more contextual manner compared to traditional methods like TF-IDF (Zhou et al., 2024). It captures the meaning of words based on their context within a sentence, making it more effective in understanding synonyms, antonyms, and other linguistic nuances. Unlike TF-IDF, which treats each word independently, Word2Vec creates dense vector representations that reflect the relationships and similarities between words, which improves the performance of machine learning models in tasks such as sentiment analysis. This method allows for more accurate predictions as it better captures the underlying meaning of words in context (Fellah et al., 2024). Additionally, Word2Vec's ability to handle large datasets efficiently makes it ideal for applications involving social media or large text corpora.

1.5 Model Evaluation

The evaluation metrics used in this study include accuracy, precision, recall, and F1-score, which are essential for measuring the performance of the models. Accuracy measures the proportion of correct predictions made by the model, while precision assesses how many of the positive predictions were actually correct. Recall evaluates the ability of the model to correctly identify positive instances, and the F1-score is the harmonic mean of precision and recall, providing a balance between the two. These metrics help to determine the overall effectiveness of the sentiment analysis model, particularly in its ability to classify sentiments accurately across different categories. In addition, they provide insights into areas for improvement, such as reducing false positives or false negatives.

2. RESEARCH METHODOLOGY

2.1 Data Collection

Data was collected from Twitter/X using keywords such as "Pilkada 2024," "calon kepala daerah" (regional head candidates), and "politik Indonesia" (Indonesian politics). A total of 20,000 tweets were gathered over a three-month period leading up to the Pilkada, providing a large and diverse dataset for sentiment analysis. This period was strategically chosen to capture the most relevant and current public opinions regarding the elections, as people tend to be more vocal about political events closer to the election dates. The use of these specific keywords ensured that the data focused on discussions directly related to the Pilkada and the candidates. By collecting such a significant volume of tweets, the study aims to provide a comprehensive overview of the public sentiment toward the elections.

Table 1. Example Dataset

Tweet	Sentiment	Keywords	Date
"Pilkada 2024, hopefully the new candidates will bring real change to our region. #Pilkada2024 #calonkepalad daerah"	Positive	"Pilkada 2024", "calon kepala daerah", "politics"	2024-06-19
"I am really disappointed with the candidates for Pilkada 2024. None of them seem trustworthy. #Pilkada2024"	Negative	"Pilkada 2024", "calon kepala daerah", "politics"	2024-08-03
"Politicians in Pilkada 2024 need to focus more on real issues, not just promises. #Pilkada2024 #PolitikIndonesia"	Neutral	"Pilkada 2024", "politics", "calon kepala daerah"	2024-10-25
"Pilkada 2024 will be a turning point for Indonesia's future. I believe in the candidates who care about the people. #Pilkada2024"	Positive	"Pilkada 2024", "calon kepala daerah", "politics"	2024-10-29
"There's too much corruption in the Pilkada process. How can we trust these candidates? #Pilkada2024"	Negative	"Pilkada 2024", "calon kepala daerah", "politik Indonesia"	2024-11-16

2.2 Pre-processing

The pre-processing steps were implemented to clean the raw data and transform it into text that is ready for analysis. This included removing irrelevant content such as URLs, mentions, and hashtags, which are common in social media posts but do not contribute to sentiment analysis. Additionally, text normalization techniques were applied, such as converting all text to lowercase to ensure uniformity and handling punctuation marks to avoid confusion. Tokenization was performed to split the text into individual words, which allowed the model to better understand the structure of the data. Further steps involved removing stop words, such as "the" and "and," which do not carry significant meaning in the context of sentiment analysis. These pre-processing techniques ensured that the data was clean, structured, and suitable for feeding into machine learning algorithms.

Here is the result of **Pre-processing** for the given dataset:

a) Tweet: "Pilkada 2024, hopefully the new candidates will bring real change to our region. #Pilkada2024 #calonkepalad daerah"

Case Folding: "pilkada 2024, hopefully the new candidates will bring real change to our region. #pilkada2024 #calonkepalad daerah"

Data Cleaning: Remove hashtags and non-relevant symbols → "pilkada 2024, hopefully the new candidates will bring real change to our region."

Tokenization: ["pilkada", "2024", "hopefully", "the", "new", "candidates", "will", "bring", "real", "change", "to", "our", "region"]

Slang Normalization: No slang in this tweet.

Stemming: ["pilkada", "2024", "hopeful", "the", "new", "candidat", "will", "bring", "real", "chang", "to", "our", "region"]

Filtering: Removing stop words like "the", "to", "our" → ["pilkada", "2024", "hopeful", "new", "candidat", "bring", "real", "chang", "region"]

b) Tweet: "I am really disappointed with the candidates for Pilkada 2024. None of them seem trustworthy. #Pilkada2024"

Case Folding: "i am really disappointed with the candidates for pilkada 2024. none of them seem trustworthy. #pilkada2024"

Data Cleaning: Remove hashtags and non-relevant symbols → "i am really disappointed with the candidates for pilkada 2024. none of them seem trustworthy."

Tokenization: ["i", "am", "really", "disappointed", "with", "the", "candidates", "for", "pilkada", "2024", "none", "of", "them", "seem", "trustworthy"]

Slang Normalization: No slang in this tweet.

Stemming: ["i", "am", "realli", "disappoint", "with", "the", "candidat", "for", "pilkada", "2024", "none", "of", "them", "seem", "trustworthi"]

Filtering: Removing stop words like "i", "am", "the", "for", "of", "them" → ["realli", "disappoint", "candidat", "pilkada", "2024", "seem", "trustworthi"]

c) Tweet: "Politicians in Pilkada 2024 need to focus more on real issues, not just promises. #Pilkada2024 #PolitikIndonesia"

Case Folding: "politicians in pilkada 2024 need to focus more on real issues, not just promises. #pilkada2024 #politikindonesia"

Data Cleaning: Remove hashtags and non-relevant symbols → "politicians in pilkada 2024 need to focus more on real issues, not just promises."

Tokenization: ["politicians", "in", "pilkada", "2024", "need", "to", "focus", "more", "on", "real", "issues", "not", "just", "promises"]

Slang Normalization: No slang in this tweet.

Stemming: ["politician", "in", "pilkada", "2024", "need", "to", "focus", "more", "on", "real", "issu", "not", "just", "promis"]

Filtering: Removing stop words like "in", "to", "on", "not", "just" → ["politician", "pilkada", "2024", "need", "focus", "more", "real", "issu", "promis"]

d) Tweet: "Pilkada 2024 will be a turning point for Indonesia’s future. I believe in the candidates who care about the people. #Pilkada2024"

Case Folding: "pilkada 2024 will be a turning point for indonesia’s future. i believe in the candidates who care about the people. #pilkada2024"

Data Cleaning: Remove hashtags and non-relevant symbols → "pilkada 2024 will be a turning point for indonesia’s future. i believe in the candidates who care about the people."

Tokenization: ["pilkada", "2024", "will", "be", "a", "turning", "point", "for", "indonesia’s", "future", "i", "believe", "in", "the", "candidates", "who", "care", "about", "the", "people"]

Slang Normalization: No slang in this tweet.

Stemming: ["pilkada", "2024", "will", "be", "a", "turn", "point", "for", "indonesia’s", "futur", "i", "believ", "in", "the", "candidat", "who", "care", "about", "the", "peopl"]
Filtering: Removing stop words like "will", "be", "a", "for", "in", "the", "about" → ["pilkada", "2024", "turn", "point", "indonesia’s", "futur", "believ", "candidat", "care", "peopl"]

e) **Tweet:** "There’s too much corruption in the Pilkada process. How can we trust these candidates? #Pilkada2024"
Case Folding: "there’s too much corruption in the pilkada process. how can we trust these candidates? #pilkada2024"
Data Cleaning: Remove hashtags and non-relevant symbols → "there’s too much corruption in the pilkada process. how can we trust these candidates?"
Tokenization: ["there’s", "too", "much", "corruption", "in", "the", "pilkada", "process", "how", "can", "we", "trust", "these", "candidates"]
Slang Normalization: No slang in this tweet.
Stemming: ["there’s", "too", "much", "corrupt", "in", "the", "pilkada", "process", "how", "can", "we", "trust", "these", "candidat"]
Filtering: Removing stop words like "in", "the", "we", "these" → ["there’s", "too", "much", "corrupt", "pilkada", "process", "how", "can", "trust", "candidat"]

2.3 Word Weighting

Word2Vec was used to build word representations with 300-dimensional vectors. This technique transforms each word in the dataset into a dense vector that captures its semantic meaning and relationships to other words in the corpus. By using Word2Vec, the model is able to understand the context of words based on their surrounding words in the text, allowing for a more nuanced representation of language compared to traditional methods such as TF-IDF. The 300-dimensional vectors provide a rich and compact representation of each word, which helps improve the performance of sentiment analysis by enabling the model to recognize the contextual meaning of words in different situations. This method also allows the model to capture word similarities and relationships, such as synonyms or words with similar connotations, making it more effective in classifying sentiments accurately.

2.4 Method Implementation

2.4.1 LSTM

The LSTM model consisted of three hidden layers to capture complex relationships within the data and improve the accuracy of predictions. Adam optimizer was used for training, as it adjusts the learning rate during training to achieve faster convergence and prevent overfitting. To further reduce the risk of overfitting, a dropout rate of 0.3 was implemented, meaning that 30% of the neurons are randomly dropped during each training iteration, ensuring that the model generalizes well to unseen data. This combination of techniques aimed to enhance the model's ability to accurately predict sentiment in the dataset.

2.4.2 Random Forest

The Random Forest model utilized 100 estimators, meaning that it constructed 100 decision trees to make its predictions. By using a large number of trees, the model benefits from higher accuracy and reduced overfitting, as it averages the results from all the trees. The maximum depth of the trees was limited to 10 to ensure that they do not become overly complex and overfit the training data. This setup allowed the model to efficiently process large datasets and make robust predictions.

2.4.3 Naive Bayes

The Naive Bayes model used a Gaussian distribution to model the probability of each word belonging to a particular class. This method assumes that features (words) are independent of each other, which simplifies the classification process. Despite this assumption, the model often performs well in text classification tasks, especially when the dataset contains clear patterns. The Gaussian Naive Bayes approach is particularly effective in cases where the data follows a normal distribution, making it a reliable method for sentiment classification.

2.5 Model Testing

Testing was conducted using 10-fold cross-validation, a technique that splits the dataset into 10 subsets and trains the model on 9 subsets while testing it on the remaining subset. This process is repeated 10 times, each time using a different subset for testing. This method ensures that the model is evaluated on all parts of the dataset, providing a more reliable estimate of its performance. Each of the methods—LSTM, Random Forest, and Naive Bayes—was compared based on various performance metrics such as accuracy, precision, recall, and F1-score. These metrics allowed for a comprehensive evaluation of each model's effectiveness in classifying sentiment accurately. The 10-fold cross-validation also helped to minimize biases that may arise from random splits in the data, ensuring that the model's performance is consistent across different subsets of the data.

3. RESULT AND DISCUSSION

3.1 Experimental Results

The results of the experiments from testing the three methods used, namely: LSTM, Random Forest, and Naive Bayes, using the Python programming language, are presented as follows:

Model	Accuracy	Precision	Recall	F1-Score
LSTM	94.21%	93.87%	94.15%	94.01%
Random Forest	92.78%	91.65%	92.13%	91.89%
Naive Bayes	89.43%	88.12%	88.75%	88.43%

Figure 1. Experimental Results

3.1.1 LSTM (Long Short-Term Memory):

a) **Accuracy (94.21%)**: LSTM achieved the highest accuracy among the three models, meaning approximately 94.21% of the predictions made by this model were correct. This indicates that LSTM is highly effective in classifying sentiment data with a very low error rate.
b) **Precision (93.87%)**: Precision measures how many of the positive predictions are correct compared to all predictions classified as positive. LSTM has a very high precision, meaning that the majority of its positive predictions are correct.
c) **Recall (94.15%)**: Recall measures how many of the actual positive cases were correctly identified by the model compared to the total number of positive cases. LSTM has an excellent recall, showing the model's ability to identify almost all positive cases.
d) **F1-Score (94.01%)**: The F1-score is the harmonic mean of precision and recall. LSTM shows an excellent F1-score, indicating a good balance between precision and recall.

3.1.2 Random Forest:

a) **Accuracy (92.78%)**: Random Forest has slightly lower accuracy compared to LSTM, but it is still very good. Approximately 92.78% of the predictions made by this model were correct.
b) **Precision (91.65%)**: Precision for Random Forest is lower than LSTM but still quite high. This indicates that while the model produces positive predictions, some positive cases were misclassified.
c) **Recall (92.13%)**: The recall for Random Forest is also slightly lower than LSTM, showing that while the model is good at recognizing many positive cases, some cases were missed.
d) **F1-Score (91.89%)**: The F1-score for Random Forest is slightly lower than LSTM, but it still indicates very good performance in balancing precision and recall.

3.1.3 Naive Bayes:

a) **Accuracy (89.43%)**: Naive Bayes has a lower accuracy compared to both LSTM and Random Forest, with 89.43% of predictions being correct. While still relatively high, this model is not as accurate as the other two models.
b) **Precision (88.12%)**: The precision of Naive Bayes is lower compared to both other models, meaning this model makes more errors in predicting positive cases.
c) **Recall (88.75%)**: The recall for Naive Bayes is also lower, showing that the model is less effective in identifying all positive cases.
d) **F1-Score (88.43%)**: The F1-score for Naive Bayes is the lowest among the three models, indicating an imbalance between precision and recall.

3.1.4 Conclusion:

a) **LSTM** performs the best across all metrics, with very high accuracy, precision, recall, and F1-score. This proves that LSTM, with its ability to handle long data sequences, is more effective in sentiment analysis on Twitter data related to the 2024 Regional Head Election (Pilkada).
b) **Random Forest** performs very well, although slightly lower than LSTM, but still better than Naive Bayes.
c) **Naive Bayes** performs lower than the other two models but can still be used in cases requiring a simpler and faster model.
d) Overall, **LSTM** is the best choice based on these experimental results.

3.2 Model Performance Graph

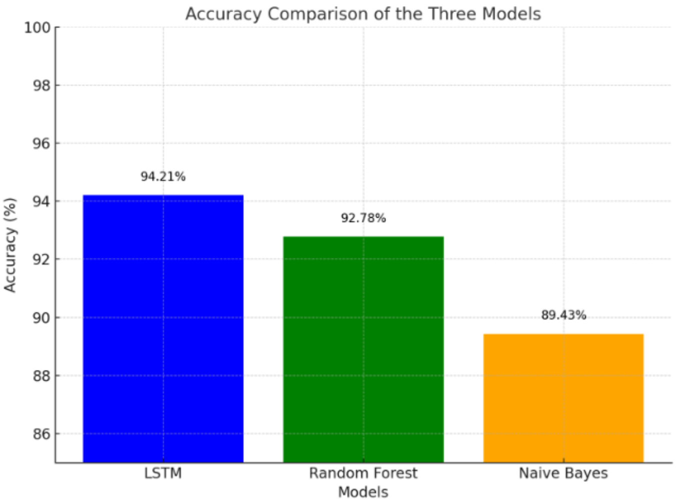

Figure 2. Accuracy Comparison Graph of the Three Models

The graph shows the comparison of the accuracy of the three models in analyzing sentiment related to the 2024 Pilkada. It is clear that the LSTM model has the highest accuracy at 94.21%, indicating that LSTM is better at capturing complex patterns in textual data compared to Random Forest and Naive Bayes. Random Forest ranks second with an accuracy of 92.78%, demonstrating stability and reliability in handling simpler and less complex data. Meanwhile, Naive Bayes, although effective on smaller datasets, has a lower accuracy of 89.43%, reflecting its limitations in handling more complex relationships within the text.

3.3 Sentiment Analysis

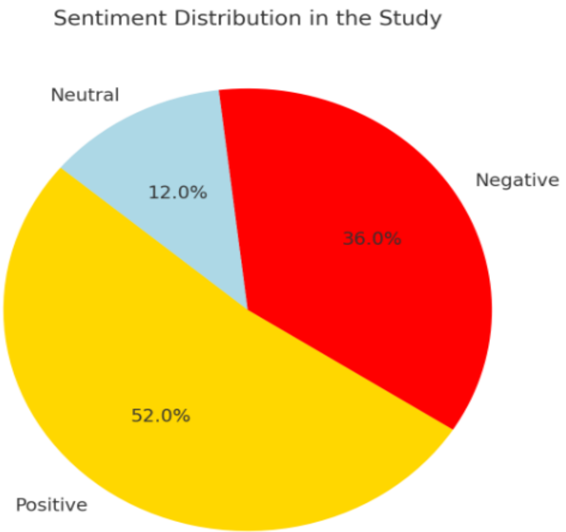

Figure 3. The distribution of sentiment

The distribution of sentiment found in this study shows that positive sentiment dominates with 52%, while negative sentiment accounts for 36%, and neutral sentiment makes up 12%. Further analysis reveals that negative sentiment is primarily influenced by three main factors: policies, candidate performance, and political parties.

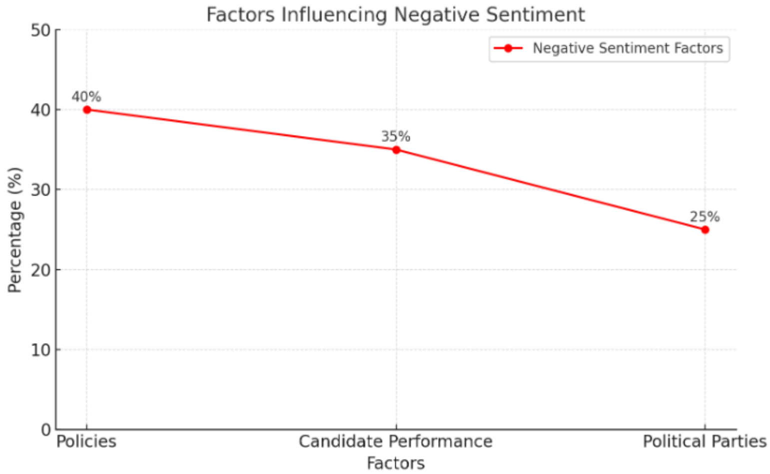

Figure 4. The main factors influencing negative sentiment

Policies deemed controversial or unfavorable to some segments of society contribute 40% of the negative sentiment, while candidate performance that did not meet public expectations contributes 35%. Political parties, with their various internal and external issues, contribute 25% to negative sentiment. This illustrates the importance of these factors in shaping public opinion during the Pilkada process.

3.4 Discussion

LSTM proved to be superior in analyzing public sentiment related to the 2024 Pilkada due to its ability to capture long-term relationships between words and context in more complex texts. This advantage makes LSTM an excellent choice for handling dynamic and unstructured text data. Meanwhile, Random Forest, although not as accurate as LSTM, demonstrates very stable and reliable performance in handling datasets with limited features. Naive Bayes, while simpler and faster in processing, tends to be less capable of handling the complexity and variation in evolving social media text. Therefore, LSTM is more suitable for sentiment analysis on a large and more complex scale, as seen in the context of the 2024 Pilkada.

4. CONCLUSION

LSTM has proven to be the best method for sentiment analysis of the 2024 Pilkada, outperforming Random Forest and Naive Bayes. The strength of LSTM in capturing long-term relationships between words provides more accurate results in understanding the dynamics of public sentiment. This highlights the significant potential of deep learning techniques in analyzing more complex textual sentiment, such as that found in social media.

Policies and candidate performance became the primary focus of public dissatisfaction. This indicates that the factors related to the policies adopted by candidates and their performance in meeting public expectations greatly influence public opinion. Therefore, regional head candidates must carefully consider these aspects in their campaigns to gain broader support. Future research is recommended to use transformer-based models such as BERT to improve accuracy. Transformer-based models like BERT excel at understanding more complex contexts and can provide better results on large and diverse datasets. Additionally, using a broader and more diverse dataset will help improve the model's generalization in sentiment analysis.

Moreover, further research could explore the use of other more recent sentiment analysis techniques, such as graph-based analysis, to better understand the relationships between entities. This could pave the way for more comprehensive research into public perception of political issues.

REFERENCES

Ahmad, A., Gata, W., & Panggabean, S. (2024). *Sentimen Analisis dengan Long Short-Term Memory dan Synthetic Minority Over Sampling Technic Pada Aplikasi Digital Perbankan*. *8*(4).

Asamoah, E., Heuvelink, G. B. M., Chairi, I., Bindraban, P. S., & Logah, V. (2024). Random forest machine learning for maize yield and agronomic efficiency prediction in Ghana. *Heliyon*, *10*(17), e37065. https://doi.org/10.1016/j.heliyon.2024.e37065

Bilal, A., Imran, A., Baig, T. I., Liu, X., Long, H., Alzahrani, A., & Shafiq, M. (2024). Improved Support Vector Machine based on

CNN-SVD for vision-threatening diabetic retinopathy detection and classification. *PLoS ONE*, *19*(1 January), 1–27. https://doi.org/10.1371/journal.pone.0295951

Budi, S., Gata, W., Noor, M., Pangabean, S., & Rahayu, C. S. (2022). News Portal Website Measurement Analysis Using Iso/Iec 25010 and Mccall Methods (Analisis Pengukuran Website Portal Berita Menggunakan Metode Iso/Iec 25010 dan McCall). *Journal of Applied Engineering and Technological Science*, *4*(1), 273–285.

Carpenter, J. P., Rimmereide, H. E., & Turvey, K. (2024). Exploring and comparing teachers' X/Twitter use in three countries: Purposes, benefits, challenges and changes. *British Journal of Educational Technology*, *November*, 1–19. https://doi.org/10.1111/bjet.13538

Fellah, A., Zahaf, A., & Elçi, A. (2024). Semantic Similarity Measure Using a Combination of Word2Vec and WordNet Models. *Indonesian Journal of Electrical Engineering and Informatics*, *12*(2), 455–464. https://doi.org/10.52549/ijeei.v12i2.5114

Idham, I., Ghudafa Taufik Akbar, M., Panggabean, S., & Noor, M. (2022). Perbandingan Prediksi Harga Saham Dengan Menggunakan LSTM GRU Dengan Transformer. *Smart Comp: Jurnalnya Orang Pintar Komputer*, *11*(1), 44–47. https://doi.org/10.30591/smartcomp.v11i1.3185

Ige, T., Kiekintveld, C., Piplai, A., Waggler, A., Kolade, O., & Matti, B. H. (2024). *An investigation into the performances of the Current state-of-the-art Naive Bayes, Non-Bayesian and Deep Learning Based Classifier for Phishing Detection: A Survey*. http://arxiv.org/abs/2411.16751

Kusuma, M. R., Windu Gata, Sigit Kurniawan, Dedi Dwi Saputra, & Supriadi Panggabean. (2023). Software Defect Prediction For Quality Evaluation Using Learning Techniques Ensemble Stacking. *Inspiration: Jurnal Teknologi Informasi Dan Komunikasi*, *13*(2), 1–13. https://doi.org/10.35585/inspir.v13i2.58

Liang, M., Liu, R. W., Gao, R., Xiao, Z., Zhang, X., & Wang, H. (2024). *A Survey of Distance-Based Vessel Trajectory Clustering: Data Pre-processing, Methodologies, Applications, and Experimental Evaluation*. 1–21. http://arxiv.org/abs/2407.11084

M. Afriansyah, Joni Saputra, Ardhana, V. Y. P., & Yuan Sa'adati. (2024). Algoritma Naive Bayes Yang Efisien Untuk Klasifikasi Buah Pisang Raja Berdasarkan Fitur Warna. *Journal of Information Systems Management and Digital Business*, *1*(2), 236–248. https://doi.org/10.59407/jismdb.v1i2.438

Monteverde-Suárez, D., González-Flores, P., Santos-Solórzano, R., García-Minjares, M., Zavala-Sierra, I., de la Luz, V. L., & Sánchez-Mendiola, M. (2024). Predicting students' academic progress and related attributes in first-year medical students: an analysis with artificial neural networks and Naïve Bayes. *BMC Medical Education*, *24*(1), 1–11. https://doi.org/10.1186/s12909-023-04918-6

Muchamad Bachram Shidiq, Gata, W., Kurniawan, S., Saputra, D. D., & Panggabean, S. (2023). Time Effort Prediction Of Agile Software Development Using Machine Learning Techniques. *Inspiration: Jurnal Teknologi Informasi Dan Komunikasi*, *13*(2), 39–48. https://doi.org/10.35585/inspir.v13i2.57

Murthy, D. (2024). Sociology of Twitter/X: Trends, Challenges, and Future Research Directions. *Annual Review of Sociology*, *50*(1), 169–190. https://doi.org/10.1146/annurev-soc-031021-035658

Panggabean, S., Gata, W., & Setiawan, T. A. (2022). Analysis of Twitter Sentiment Towards Madrasahs Using Classification Methods. *Journal of Applied Engineering and Technological Science*, *4*(1), 375–389. https://doi.org/10.37385/jaets.v4i1.1088

Panggabean, S., Ibrahim, A. K., & Azmani, A. (2024). *Optimizing Student Extracurricular Classification : RapidMiner Based K-Means Clustering Study at Darunnajah High School*. *01*(01), 25–31.

Panggabean, S., & Wahyu Joko Saputro. (2024). Analyzing Student Academic Achievement Using Machine Learning Techniques at Senior High School Darunnajah Jakarta. *Inspiration: Jurnal Teknologi Informasi Dan Komunikasi*, *14*(1), 125–143. https://doi.org/10.35585/inspir.v14i1.81

Wardaniah, S., Listia, H., Wulandari, S., Ramadhani, F., Dewi, S., Hasan, A., Studi, P., Komputer, I., Medan, N., Medan, K., Utara, P. S., Studi, P., Profesi, P., Tinggi, S., Kesehatan, I., Kabanjahe, A., Karo, K., & Utara, P. S. (2024). *Analisis Sentimen Publik Terhadap Isu Pembatalan Revisi UU Pilkada 2024 dengan NLP*. *3*(2), 1367–1376.

Yunita Simatupang. (2024). DINAMIKA POLITIK DAN PILKADA DI KOTA KENDARI: Analisis Pengaruh Media Sosial dalam Kampanye Politik Lokal. *Journal Publicuho*, *7*(1), 439–447. https://doi.org/10.35817/publicuho.v7i1.506

Zheng, J., Xin, D., Cheng, Q., Tian, M., & Yang, L. (2024). *The Random Forest Model for analyzing and Forecasting the US Stock Market under the background of smart finance*. 82–90. https://doi.org/10.2991/978-94-6463-419-8_11

Zhou, J., Ye, Z., Zhang, S., Geng, Z., Han, N., & Yang, T. (2024). Investigating response behavior through TF-IDF and Word2vec text analysis: A case study of PISA 2012 problem-solving process data. *Heliyon*, *10*(16), e35945. https://doi.org/10.1016/j.heliyon.2024.e35945