

Sentiment Analysis on Indonesian Tweets About the 2024 Election

Alfan Ramadhan Sembiring^{1)*}, Chandra Kusuma Dewa²⁾

^{1,2)}Department of Informatics, Universitas Islam Indonesia, Yogyakarta, Indonesia
¹22917024@students.uii.ac.id, ²chandra.kusuma@uui.ac.id

Submitted : Jan 16, 2025 | **Accepted** : Jan 21, 2025 | **Published** : Jan 30, 2025

Abstract: This study investigates public sentiment on Indonesian Twitter regarding the 2024 General Election, employing machine learning and deep learning techniques, including Naïve Bayes, Support Vector Machines (SVM), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU). The dataset was collected using a Tweet Harvest method with the keyword "Pemilu" and underwent preprocessing steps such as case folding, removal of symbols and URLs, stop word elimination, and tokenization to ensure data quality. Performance evaluation metrics, including accuracy, precision, recall, and F1-score, were applied to assess the models' effectiveness. Naïve Bayes achieved the highest accuracy of 64%, followed by SVM at 63%, LSTM at 60%, and GRU at 57%. The findings indicate that traditional models like Naïve Bayes and SVM perform effectively on smaller datasets with structured features, while deep learning models excel in capturing complex sequential dependencies. However, deep learning methods exhibited overfitting tendencies, indicating the need for better regularization and optimization techniques. Furthermore, it emphasizes the potential of integrating traditional algorithms with advanced methods to enhance sentiment classification accuracy and generalizability across diverse datasets.

Keywords: deep learning, gru, lstm, machine learning, naive bayes, sentiment analysis, svm

INTRODUCTION

Social media has now become an integral part of our daily routine. Tweets often include only a few words that have practical meaning, and these words are very important in the classification stage (Talaat, 2023). The General Election in 2024 is a very important part of Indonesia's democracy, especially when entering the campaign period where many people express themselves towards the General Election on Social Media such as X or what was known in the past as Twitter.

Social media has grown phenomenally over the past decade (Kar & Debbarma, 2023). For example, Twitter has 321 million active users. On this social media application, people are addicted to expressing their opinions, suggestions, likes, dislikes, and experiences in a very casual way (Abid et al., 2020). Although many social media platforms are used for online discussion and information exchange, Twitter still holds an important position in terms of text data dissemination (Polyzos, Efstathios & Wang, 2022). Along with the evolution of social media networks, the sheer volume of social media text available for sentiment analysis has increased several fold, leading to a robust corpus (Iddrisu et al., 2023)

One of the very important things in analyzing public opinion is to examine their feelings. Here, the concept of sentiment analysis comes into (Pandey & Vishwakarma, 2023). Sentiment Analysis is a branch of Natural Language Processing (NLP) that focuses on analyzing people's views, feelings, and emotions (Talaat, 2023). Sentiment Analysis models play an important role. Many sectors, including political challenges, marketing, public policy, disaster management, and public health, rely on emotion detection (Talaat, 2023). However, its relevance and application in the context of General Elections in Indonesia is still relatively little explored.

One of the Deep Learning methods that is widely used to perform Sentiment Analysis is the RNN method. CNN-based methods are gaining popularity along with the development of the deep learning field. Nevertheless, most CNN techniques currently focus on specific jobs (Alruily, 2023). In general, Recurrent Neural Network (RNN), Long Short Term Memory (LSTM), Gated Recurrent Unit (GRU), Convolutional Neural Network (CNN) and recursive neural networks are often used to encode word sequences to determine the sentiment polarity of a particular aspect, sentence or document (Liu & Shen, 2020).

The enthusiasm of many people in expressing themselves on social media has resulted in Twitter becoming one of the people's destinations to find out about other users' opinions on the 2024 General Election. From the

*name of corresponding author



many tweets available in general, it can be used as a data source for research purposes and classify what topics are widely conveyed on Twitter. Sentiment analysis in social media analysis are used to identify problems described in user-generated content, ideas and opportunities, and important discussions (Mulyani et al., 2024) quickly, effectively, and cheaply, and extrapolate common themes across documents (Melton et al., 2021).

In previous studies, machine learning algorithms were commonly used, while some research focused on one or two deep learning algorithms. However, this study employs a combination of two machine learning algorithms and two deep learning algorithms on a dataset of tweets related to the 2024 election. This approach highlights the complexity of the research and demonstrates a more comprehensive contribution compared to earlier studies.

LITERATURE REVIEW

Research on sentiment analysis has developed rapidly, utilizing various machine learning methods and datasets applied to textual data from diverse contexts, ranging from social media to product reviews. In sentiment analysis on social media, (Fitri et al., 2019) employed Naive Bayes, Decision Tree, and Random Forest algorithms to analyze tweets related to anti-LGBT issues, where Naive Bayes achieved the highest accuracy of 83.43%. (Hulu & Lhaksana, 2019) applied SVM with TF-IDF to tweets discussing the 2019 Indonesian presidential election, achieving a maximum accuracy of 62.88%. Similarly, (Muhajir et al., 2024) used SVM and LDA on tweets with keywords "Bjorkanism" and "Hacker Bjorka," consistently achieving accuracy rates above 85%. Meanwhile, (Damayanti & Lhaksana, 2024) employed SVM with Word2Vec to analyze tweets about the 2024 Indonesian presidential election, achieving the highest accuracy of 90.75%.

Sentiment analysis has also been widely applied to product and app reviews. (Awotunde et al., 2023) used Naive Bayes on business review data from Kaggle and achieved an accuracy of 96.08%. (Idris et al., 2025) analyzed reviews of the Tokopedia app using SVM, Random Forest, and Naive Bayes, with SVM combined with FastText achieving the highest accuracy of 89%. (Akbar et al., 2024) applied Naive Bayes to reviews of the InDriver online motorcycle taxi app, achieving 76% accuracy. (Murfi et al., 2024) evaluated combinations of CNN, LSTM, GRU, and hybrid models on e-commerce reviews from Shopee, Tokopedia, and Lazada. Using the BERT text representation method, the LSTM-CNN model achieved the best accuracy on the Tokopedia dataset, while CNN-LSTM provided the best precision on the Shopee dataset.

Other studies focused on more specific domains and datasets. (Iddrisu et al., 2023) analyzed tweets discussing the impact of the COVID-19 pandemic on air travel, with SVM achieving the highest accuracy of 89%. (Aakash et al., 2024) compared models such as LSTM and GRU on product reviews, with GRU achieving the highest recall rate of 92%. (AL-Smadi et al., 2023) concentrated on Arabic hotel reviews, where the GRU model achieved 91.4% accuracy and an F1-score of 90.86%. (Talaat, 2023) used BiLSTM and BiGRU on tweet datasets from HuggingFace, with the RoBERTa model achieving the best accuracy of 91.72%.

Overall, these studies demonstrate that various machine learning methods, such as Naive Bayes, SVM, LSTM, GRU, and modern approaches like BERT, can achieve high performance in sentiment analysis. However, the final performance heavily depends on the type of data and the context of the research.

Sentiment Analysis

In general, Sentiment Analysis deals with the mining of information related to sentiments or opinions from a group for a specified topic. In addition, sentiments from some applications are received at the document level (Saberi & Saad, 2017). Sentiment Analysis is a multi-step process that includes data capture, extraction, pre-processing, and feature extraction (Talaat, 2023). Naive Bayes has recently been widely used in classification techniques, especially in social media networks such as Twitter, using several methods, including Unigram Naive Bayes, Multinomial Naive Bayes and Maximum Entropy Classification (Samsir et al., 2021). The calculation of probability categories in Naive Bayes is based on the basic principle of Bayes' Theorem. This algorithm calculates the probability of data belonging to a particular category using the following formula:

$$P(X) = \frac{P(C)P(C)}{P(X)} \quad (1)$$

The classification in Support Vector Machines (SVM) is based on the principle of finding the optimal hyperplane that maximally separates data points of different categories. This algorithm calculates the decision boundary using the following formula :

$$f(x) = \text{sgn} \left(\sum_{n=1}^N \alpha_i - y_i K(x'_n, x) + b \right) \quad (2)$$

Natural Language Processing (NLP) has greatly benefited from the rise of Deep Neural Networks (DNN), due to their high performance with less need for engineered features. There are two main DNN architectures:

*name of corresponding author



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) (Yin et al., 2017). Recurrent neural network (RNN) has been widely used in sequence learning problems such as action recognition, scene labeling and language processing, and has achieved impressive results (Li et al., 2018). However, RNNs often suffer from the classic “Vanishing” and “Exploding” gradient problems, resulting in failure to learn long-range dependencies (Rana, 2016).

Long Short-Term Memory (LSTM) was introduced to alleviate the vanishing gradient problem, and has become one of the most popular RNN architectures to date (DiPietro & Hager, 2019). The equation for the Forget Gate can be seen in equation 3.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (3)$$

The Input Gate used to determined the importance of the new information. The equation of input gate can seen in equation 4.

$$i_t = \sigma(x_t * U_i + H_{t-1} * W_i) \quad (4)$$

Output gate control cell state that is used to computer the current hidden state. The equation of output gate pretty similar to the previous two gates.

$$o_t = \sigma(X_t * U_o + H_{t-1} * W_o) \quad (5)$$

The cell state in an LSTM acts as the memory of the network, carrying long-term information across time steps, while the hidden state represents the output of the LSTM at a specific time step, encapsulating both the current input and relevant past information for short-term dependencies. The equation for cell state and hidden state can be seen in equation 6 and 7.

$$C_t = f_t * C_{t-1} + i_t \cdot C_t \quad (6)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (7)$$

The Gated Recurrent Unit (GRU) is a slightly more simplified variation of the LSTM. It combines the Forget gate and input gates into one “update gate” and has an additional “reset gate”. The final model is simpler than the standard LSTM model and is becoming increasingly popular (Rana, 2016). The equation for the update gate (z_t) can be seen in equation 8.

$$z_t^j = \sigma(W_z x_t + u_z h_t - 1)^j \quad (8)$$

The Update gate (z_t^j) controls how much information from the previous unit's activation (h_{t-1}) will be retained or updated with new information from the current input (x_t). The value of z_t^j ranges between 0 and 1, where 0 means “forget completely” and 1 means “keep completely”. A sigmoid activation function (σ) is used to generate values within this range.

The reset gate determines how much information from the previous hidden state should be discarded. Similar to the update gate, the reset gate also uses a sigmoid activation function to produce a value between 0 and 1. The equation for the reset gate (r_t^j) can be seen in equation 9.

$$r_t^j = \sigma(W_r x_t + r h_t - 1)^j \quad (9)$$

Candidate activation is a potential representation of the new hidden state. The hyperbolic tangent (\tanh) activation function is used to generate values between -1 and 1. The equation for candidate activation can be seen in equation 10.

$$h_t^{-j} = \tanh(W x_t + U(r_t \cdot h_t - 1)^j) \quad (10)$$

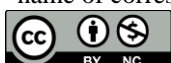
Hidden state is a weighted combination of the previous hidden state ($h_{(t-1)^j}$) and candidate activation ($h_t^{(-j)}$). The weighting is controlled by the update gate (z_t^j). The equation for the hidden state is listed in equation 11.

$$h_t^j = (1 - z_t^j) h_{t-1}^j + z_t^j h_t^{-j} \quad (11)$$

METHOD

This section explains the research methodology used to conduct sentiment analysis from textual data collected from social media. The approach leverages deep learning methods and statistical algorithms to produce a thorough and comprehensive analysis. This section details the research approach from start to finish, including data collection, preprocessing, analysis, and interpretation of the results. The use of Naïve Bayes, SVM, LSTM and

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

GRU for sentiment analysis was chosen due to their ability to handle sequential text data. This research aims to provide deeper insights into public opinion through sentiment analysis of unstructured data. By applying a rigorous methodological approach at each stage of the study, the findings are expected to make a significant contribution to the field of text analysis and social understanding. This chapter will outline the various techniques and tools used, and how each method contributes to the overall goals of the research. The following is a flowchart of the system which can be seen in Figure 1.

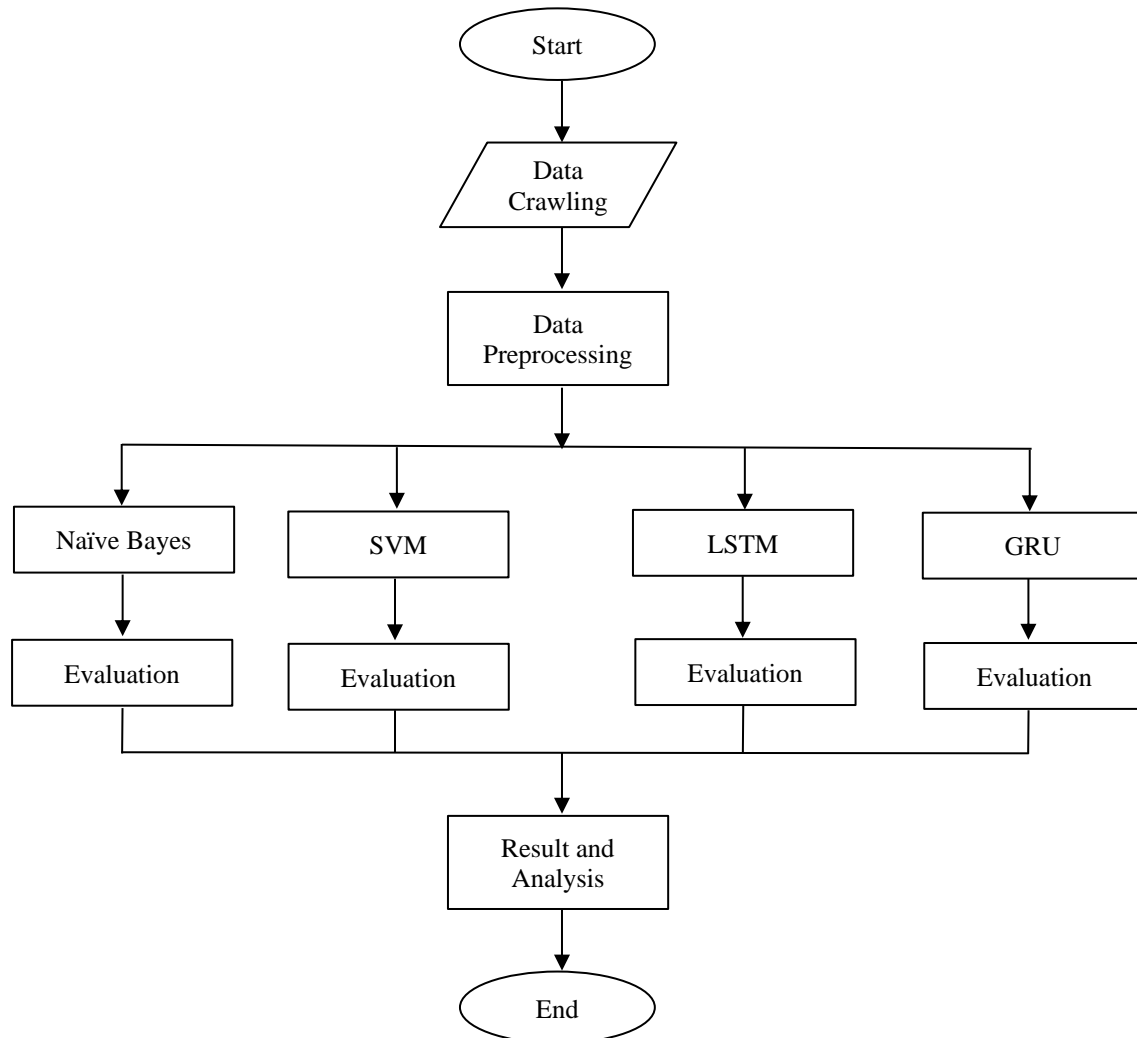


Figure 1. System overview flowchart

Dataset

The data was collected from the social media platform X/Twitter using the Tweet Harvest technique, which is as follows: identifying goals and criteria, twitter access and data harvesting. The detail of the process start with the crawling process was conducted using Python libraries and a Twitter Authorization Token, with the keyword “Pemilu” (Election) specified for the crawling process, along with search dates in accordance with the problem constraints. The collected data was then saved in CSV (Comma-Separated Values) format. The crawling process was performed several times due to Twitter’s limit on the number of tweets that can be accessed. The results of the data crawling can be seen in the following table.

Table 2. Crawling Data

Crawling Data
Kerja Keras PSI Cegah Kekalahan Terulang Seperti Pemilu 2019 - https://t.co/SKg61sB7cH https://t.co/p4PiNJ41aJ #IndonesiaKuatIndonesiaMaju #DiplomasiProgresif02 #02Melanjutkan Sebenarnya gua udah muak bgt dah mantengin huru-hara pemilu ini gua mencoba buat ga peduli tapi GUA GA SUDI DIPIMPIN SAMA PRESIDEN DAN WAKIL MODELAN GITU belum pendukung-pendukungnya juga rusak semua lagi ga ada yang bener

*name of corresponding author



@RcyberProj0 @Gerindra @prabowo @gibrantweet @budimuni98 @handoko2411 @FALDA81 Survei hanya penggiringan opini lucu nya pemilu negeri KONOHA hasil pilpres bisa ditentukan seblm pencoblosan Dalam menghadapi Pemilu Damai 2024 Anggota BhabinKamtibmas Polsek Sukajadi dipimpin Kanit Binmas Iptu Subahagia melaksanakan kegiatan Patroli dialogis dengan menyambangi Ketua PPK Kec. Suakajadi Bpk. Cecep bertempat di Kantor Kec. Sukajadi Kota Bandung. <https://t.co/az2TMeZc0R>
@nongandah @Paltiwest Ingat pemilu 2019 banyak juga pendukung prabowo ditangkap

Preprocessing

To ensure transparency and reproducibility in NLP research, it is important to provide clear documentation of preprocessing decisions and their impact on the analysis. This includes considering the type of text mining conducted and the characteristics of the dataset (Hickman et al., 2022). The following are several stages in preprocessing in this research:

Text Normalization

Text normalization, which includes converting text to lowercase and handling punctuation, is also crucial for ensuring consistency across the text data (MR ADEPU RAJESH & DR TRYAMBAK HIWARKAR, 2023).

Stemming

Stemming are techniques used to reduce words to their base or root form. Stemming involves cutting off prefixed or suffixes, while lemmatization uses linguistic knowledge to derive the base form. These methods are vital for reducing dimensionality and improving the performance of NLP systems (Jabbar et al., 2023).

Stop Words Removal

Stop word removal is essential for optimizing processes in information retrieval and text analytics systems. By filtering out redundant words, these systems can achieve higher accuracy and performance (Kaur & Saini, 2015). The removal of stop words reduces the size of the corpus by 35-45%, which in turn improves the efficiency and accuracy of text mining applications by decreasing time and space complexity (Ladani & Desai, 2020).

Tokenization

Splitting text into individual words or tokens is fundamental for most NLP tasks (Chai, 2023). It enhances the accuracy of part-of-speech tagging and other analyses (Camacho-Collados & Pilehvar, 2017).

Table 3. Text after Preprocessing

Preprocessing Step	Text After Preprocessing
Original Text	Kerja Keras PSI Cegah Kekalahan Terulang Seperti Pemilu 2019 - https://t.co/SKg61sB7cH https://t.co/p4PiNJ41aJ #IndonesiaKuatIndonesiaMaju #DiplomasiProgresif02 #02Melanjutkan
Case Folding	kerja keras psi cegah kekalahan terulang seperti pemilu 2019 - https://t.co/SKg61sB7cH https://t.co/p4PiNJ41aJ #indonesiakuatindonesiamaju #diplomasi progresif02 #02melanjutkan
Removing Symbol, URL, and Numbering	kerja keras psi cegah kekalahan terulang seperti pemilu indonesiakuatindonesiamaju diplomasi progresif lanjut
Stop words removal	kerja keras psi cegah kalah ulang pemilu indonesiakuatindonesiamaju diplomasi progresif lanjut
Tokenization	['kerja', 'keras', 'psi', 'cegah', 'kalah', 'ulang', 'pemilu', 'indonesiakuatindonesiamaju', 'diplomasi progresif', 'lanjut']

Naïve Bayes

Among the assortment of current classification techniques, the Naïve Bayes (NB) classifier has played a prominent role because of its simplicity, tractability and efficiency (Blanquero et al., 2021). Naïve Bayes calculates probability for each sentiment class based on text features represented by TF-IDF. Multinomial Naïve Bayes was used, which is suitable for text data represented as word frequency distributions.

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Support Vector Machine

SVMs are based on convex optimization, ensuring no local minima and efficient solutions. They often yield sparse solutions, meaning only a subset of training data points (support vectors) are used, which is beneficial for handling large datasets (Moguerza & Muñoz, 2006); (Alessia Mammone, 2009). The SVM algorithm works by finding the hyperplane that best separates the classes in the feature space. The implementation in this study used the Radial Basis Function (RBF) kernel and C hyper parameter of 1, which maps data to a higher-dimensional space to make it linearly separable.

Long Short-Term Memory

LSTMs introduce gate functions into the cell structure, which effectively manage long-term dependencies by mitigating issues like **exploding** and vanishing gradients that are common in traditional RNNs (Yu et al., 2019); (Landi et al., 2021). The architecture of LSTM includes forget gates and output activation functions, which are critical components for its performance (Greff et al., 2017). Early stopping was employed to halt training when the validation loss stopped improving for three consecutive epochs, restoring the best model weights.

Gated Recurrent Unit

GRUs are generally more computationally efficient than LSTMs due to their simpler architecture, which includes fewer gates. This efficiency makes them faster in training and suitable for applications with limited computational resources (Li et al., 2018). Early stopping was also employed to train GRU.

RESULT

In this research, we provide our sentiment analysis result using Naïve Bayes, SVM, LSTM, and GRU. The analysis will compare classification metrics, such as accuracy, precision, recall, and F1-score. Confusion Matrix was used for machine learning method, it was particularly useful for identifying the type and sources of errors, making them a stringent evaluation tool. While Accuracy and Loss per epoch was used in deep learning method. Research has identified conditions under which stability is maintained, ensuring that accuracy improves as loss decreases. This research has result as followed:

Table 4. Result Classification of Performance Evaluation Naive Bayes Method

Classification Categories	True Negative	True Neutral	True Positive	Precision	Recall	F1-Score
Prediction Negative	195	65	33	66%	75%	71%
Prediction Neutral	48	183	83	58%	57%	58%
Prediction Positive	15	73	186	68%	62%	65%

Based on the results in Table 4, data with a total of 881 tweets, obtained 195 tweets declared true negative, 183 declared true neutral, and 186 declared true neutral. The average percentage (%) for accuracy is 64%, precision 64%, recall 65%, and F1-Score with 64%. This research is using Multinomial Naïve Bayes, which is can effectively manage multiple classes and heterogeneous features, which is beneficial in diverse datasets (Pavlov et al., 2004).

Table 5. Result Classification of Performance Evaluation Support Vector Machine Method

Classification Categories	True Negative	True Neutral	True Positive	Precision	Recall	F1-Score
Prediction Negative	166	65	25	65%	65%	65%
Prediction Neutral	76	200	90	55%	62%	58%
Prediction Positive	15	90	187	72%	62%	67%

Based on the results in Table 5, data with a total of 881 tweets, obtained 166 tweets declared true negative, 200 declared true neutral, and 187 declared true neutral. The average percentage (%) for accuracy is 63%, precision 64%, recall 63%, and F1-Score with 63%.

*name of corresponding author



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Table 6. Evaluation Of Accuracy Rates Per Epoch

Classification Categories	Epoch (%)									Runtime
	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	
LSTM	0.376	0.411	0.535	0.565	0.583	0.56	0.593	0.563	-	3m.2s
GRU	0.393	0.462	0.512	0.545	0.542	0.56	0.555	0.573	0.573	2m 5s

Table 6 shows the accuracy evaluation of LSTM and GRU models per epoch during training. The training used 10 epochs with early stopping applied if the validation loss did not improve. The LSTM model achieved the highest accuracy of 59.3% at epoch 7 and stopped training after 3 minutes 2 seconds. The GRU model reached its peak accuracy of 57.3% at epoch 9, completing training in a shorter time of 2 minutes 5 seconds.

Table 7 Classification Report for LSTM Model

Classification Categories	Precision	Recall	F1-Score	Support
Negative	64%	70%	67%	128
Neutral	52%	60%	56%	161
Positive	69%	52%	66%	151

The table shows the classification report for an LSTM model. For the Negative category, the precision is 64%, recall is 70%, F1-Score is 67%, and support is 128. For the Neutral category, the precision is 52%, recall is 60%, F1-Score is 56%, and support is 161. For the Positive category, the precision is 69%, recall is 52%, F1-Score is 66%, and support is 151. The average precision is 62%, recall is 61%, and F1-Score is 61%.

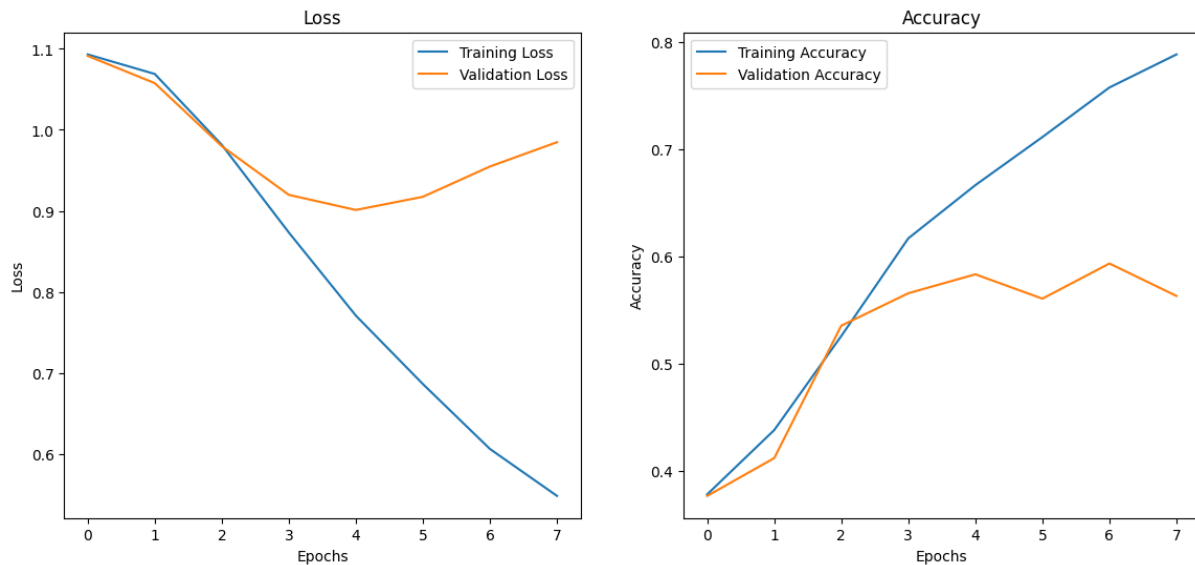


Fig 2. Loss and Accuracy Graph for LSTM model.

From figure 2, the graphs show the model's performance over 7 epochs in terms of loss and accuracy for both training and validation data. The training loss steadily decreases, indicating effective learning, while the validation loss initially decreases but starts to increase after epoch 4, signaling potential overfitting. Similarly, training accuracy improves consistently, reaching approximately 80% by epoch 7, whereas validation accuracy rises until epoch 3 before fluctuating and stabilizing around 60%. This suggests that while the model performs well on the training data, its generalization to unseen data is limited after a certain point.

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Table 8 Classification Report for GRU

Classification Categories	Precision	Recall	F1-Score	Support
Negative	74%	55%	63%	128
Neutral	50%	75%	60%	161
Positive	68%	47%	56%	151

The table shows the classification report for a GRU model. For the Negative category, the precision is 74%, recall is 55%, F1-Score is 63%, and support is 128. For the Neutral category, the precision is 50%, recall is 75%, F1-Score is 60%, and support is 161. For the Positive category, the precision is 68%, recall is 47%, F1-Score is 56%, and support is 151.

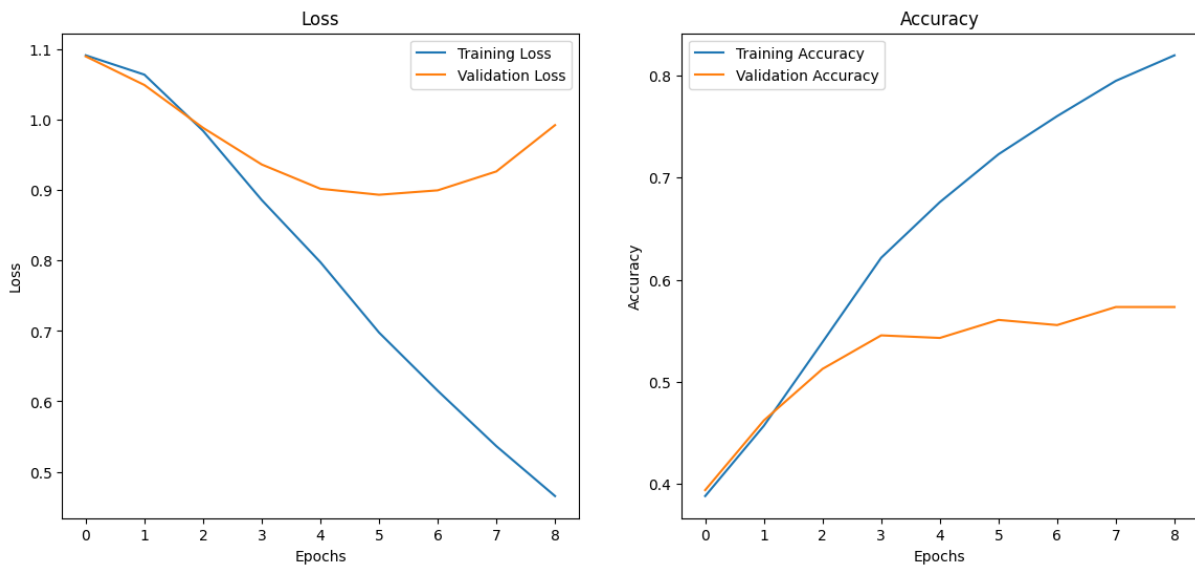


Fig. 3 Loss and Accuracy Graph for GRU model.

The graph showed improved performance over 8 epochs, characterized by a decrease in loss values on both training and validation data, as well as an increase in accuracy on both types of data. However, there is a slight indication of overfitting as shown by the validation loss value being slightly higher than the training loss.

DISCUSSIONS

The findings from Table 4 and Table 5 indicate that Naïve Bayes achieved an accuracy of 64%, slightly outperforming SVM at 63%. Both models demonstrated strengths in handling structured datasets with balanced sentiment distributions. In contrast, the deep learning models struggled with generalization, with LSTM reaching 60% accuracy and GRU at 57%. Despite this, LSTM models often provide higher stability and accuracy in predictions, although GRU models had slightly better prediction speed and accuracy (Chen et al., 2024).

The analysis of validation and training loss trends revealed overfitting in the deep learning models, as observed in Figures 2 and 3. For LSTM, validation loss began increasing after the fourth epoch, even as training accuracy continued to rise. GRU exhibited similar behavior, with validation loss exceeding training loss by the ninth epoch. These trends emphasize the importance of incorporating techniques like dropout (Srivastava et al., 2014), early stopping (Makarova et al., 2021), and hyperparameter settings such as learning rate, batch size, and regularization parameters (L1, L2) significantly influence overfitting and model performance (Xu et al., 2023).

CONCLUSION

This study demonstrates the application of machine learning and deep learning methods in analyzing Twitter sentiment surrounding Indonesia's 2024 General Election. Naïve Bayes and SVM emerged as efficient models for sentiment classification in smaller datasets, while LSTM and GRU exhibited potential for more complex analyses despite overfitting challenges. These findings suggest that combining traditional and advanced models could optimize sentiment analysis performance. Future research should focus on expanding datasets, integrating multilingual analysis, and employing hybrid models to enhance predictive accuracy and generalizability. Insights gained from this research can inform political strategies, policymaking, and public engagement initiatives.

*name of corresponding author

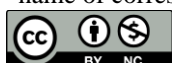


This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

REFERENCES

- Aakash, Gupta, S., & Noliya, A. (2024). URL-Based Sentiment Analysis of Product Reviews Using LSTM and GRU. *Procedia Computer Science*, 235(2023), 1814–1823. <https://doi.org/10.1016/j.procs.2024.04.172>
- Abid, F., Li, C., & Alam, M. (2020). Multi-source social media data sentiment analysis using bidirectional recurrent convolutional neural networks. *Computer Communications*, 157(December 2019), 102–115. <https://doi.org/10.1016/j.comcom.2020.04.002>
- Akbar, M. Z., Ikhsan, M., & Zufria, I. (2024). Sentiment Analysis of the Indriver Online Ojek Application using the Naïve Bayes Classifier Method. *Sinkron*, 8(3), 2037–2048. <https://doi.org/10.33395/sinkron.v8i3.13934>
- AL-Smadi, M., Hammad, M. M., Al-Zboon, S. A., AL-Tawalbeh, S., & Cambria, E. (2023). Gated recurrent unit with multilingual universal sentence encoder for Arabic aspect-based sentiment analysis. *Knowledge-Based Systems*, 261. <https://doi.org/10.1016/j.knosys.2021.107540>
- Alessia Mammone, M. T. and N. C. (2009). SVM: Support Vector Machines. *The Top Ten Algorithms in Data Mining, December*, 37–59. <https://doi.org/10.1201/9781420089653-10>
- Alruily, M. (2023). Sentiment analysis for predicting stress among workers and classification utilizing CNN: Unveiling the mechanism. *Alexandria Engineering Journal*, 81(July), 360–370. <https://doi.org/10.1016/j.aej.2023.09.040>
- Awotunde, J. B., Misra, S., Katta, V., & Adebayo, O. C. (2023). An Ensemble-Based Hotel Reviews System Using Naive Bayes Classifier. *CMES - Computer Modeling in Engineering and Sciences*, 137(1), 131–154. <https://doi.org/10.32604/cmes.2023.026812>
- Blanquero, R., Carrizosa, E., Ramírez-Cobo, P., & Sillero-Denamiel, M. R. (2021). Variable selection for Naïve Bayes classification. *Computers and Operations Research*, 135, 105456. <https://doi.org/10.1016/j.cor.2021.105456>
- Camacho-Collados, J., & Pilehvar, M. T. (2017). On the Role of Text Preprocessing in Neural Network Architectures: An Evaluation Study on Text Categorization and Sentiment Analysis. *ArXiv, abs/1707.0*. <https://doi.org/10.18653/v1/W18-5406>
- Chai, C. P. (2023). Comparison of text preprocessing methods. *Natural Language Engineering*, 29(3), 509–553. <https://doi.org/DOI:10.1017/S1351324922000213>
- Chen, X., Yang, L., Xue, H., Li, L., & Yu, Y. (2024). A Machine Learning Model Based on GRU and LSTM to Predict the Environmental Parameters in a Layer House, Taking CO2 Concentration as an Example. *Sensors*, 24(1). <https://doi.org/10.3390/s24010244>
- Damayanti, L., & Lhaksmana, K. M. (2024). Sentiment Analysis of the 2024 Indonesia Presidential Election on Twitter. *Sinkron*, 8(2), 938–946. <https://doi.org/10.33395/sinkron.v8i2.13379>
- DiPietro, R., & Hager, G. D. (2019). Deep learning: RNNs and LSTM. In *Handbook of Medical Image Computing and Computer Assisted Intervention*. Elsevier Inc. <https://doi.org/10.1016/B978-0-12-816176-0.00026-0>
- Fitri, V. A., Andreswari, R., & Hasibuan, M. A. (2019). Sentiment analysis of social media Twitter with case of Anti-LGBT campaign in Indonesia using Naïve Bayes, decision tree, and random forest algorithm. *Procedia Computer Science*, 161, 765–772. <https://doi.org/10.1016/j.procs.2019.11.181>
- Greff, K., Srivastava, R. K., Koutnik, J., Steunebrink, B. R., & Schmidhuber, J. (2017). LSTM: A Search Space Odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, 28(10), 2222–2232. <https://doi.org/10.1109/TNNLS.2016.2582924>
- Hickman, L., Thapa, S., Tay, L., Cao, M., & Srinivasan, P. (2022). Text Preprocessing for Text Mining in Organizational Research: Review and Recommendations. *Organizational Research Methods*, 25(1), 114–146. <https://doi.org/10.1177/1094428120971683>
- Hulu, A. M. F., & Lhaksmana, K. M. (2019). Analisis Sentimen Politik pada Twitter Menggunakan Metode Support Vector Machine (Studi Kasus : Pilpres 2019). *E-Proceeding of Engineering*, 6(2), 2017–2020.
- Iddrisu, A. M., Mensah, S., Boafo, F., Yeluripati, G. R., & Kudjo, P. (2023). A sentiment analysis framework to classify instances of sarcastic sentiments within the aviation sector. *International Journal of Information Management Data Insights*, 3(2). <https://doi.org/10.1016/j.ijime.2023.100180>
- Idris, M., Rifai, A., & Tania, K. D. (2025). *Sentiment Analysis of Tokopedia App Reviews using Machine Learning and Word Embeddings*. 9(1), 210–219.
- Jabbar, A., Iqbal, S., Tamimy, M. I., Rehman, A., Bahaj, S. A., & Saba, T. (2023). An Analytical Analysis of Text Stemming Methodologies in Information Retrieval and Natural Language Processing Systems. *IEEE Access*, 11, 133681–133702. <https://doi.org/10.1109/ACCESS.2023.3332710>
- Kar, P., & Debbarma, S. (2023). Sentimental analysis & Hate speech detection on English and German text collected from social media platforms using optimal feature extraction and hybrid diagonal gated recurrent neural network. *Engineering Applications of Artificial Intelligence*, 126(January). <https://doi.org/10.1016/j.engappai.2023.107143>
- Kaur, J., & Saini, J. R. (2015). A Natural Language Processing Approach for Identification of Stop Words in Punjabi Language. *International Journal of Data Mining And Emerging Technologies*, 5(2), 114.

*name of corresponding author



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

- <https://doi.org/10.5958/2249-3220.2015.00015.4>
- Ladani, D. J., & Desai, N. P. (2020). Stopword Identification and Removal Techniques on TC and IR applications: A Survey. *2020 6th International Conference on Advanced Computing and Communication Systems, ICACCS 2020*, 466–472. <https://doi.org/10.1109/ICACCS48705.2020.9074166>
- Landi, F., Baraldi, L., Cornia, M., & Cucchiara, R. (2021). Working Memory Connections for LSTM. *Neural Networks*, *144*, 334–341. <https://doi.org/10.1016/j.neunet.2021.08.030>
- Li, S., Li, W., Cook, C., Zhu, C., & Gao, Y. (2018). Independently Recurrent Neural Network (IndRNN): Building A Longer and Deeper RNN. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 5457–5466. <https://doi.org/10.1109/CVPR.2018.00572>
- Liu, N., & Shen, B. (2020). Aspect-based sentiment analysis with gated alternate neural network. *Knowledge-Based Systems*, *188*. <https://doi.org/10.1016/j.knosys.2019.105010>
- Makarova, A., Shen, H., Perrone, V., Klein, A., Faddoul, J., Krause, A., Seeger, M., & Archambeau, C. (2021). Overfitting in Bayesian Optimization: an empirical study and early-stopping solution. *ArXiv, abs/2104.08166*. <https://consensus.app/papers/overfitting-in-bayesian-optimization-an-empirical-study-makarova-shen/ce3cdd176b0e5faab18721b504c56fd8/>
- Melton, C. A., Olusanya, O. A., Ammar, N., & Shaban-Nejad, A. (2021). Public sentiment analysis and topic modeling regarding COVID-19 vaccines on the Reddit social media platform: A call to action for strengthening vaccine confidence. *Journal of Infection and Public Health*, *14*(10), 1505–1512. <https://doi.org/10.1016/j.jiph.2021.08.010>
- Moguerza, J. M., & Muñoz, A. (2006). Support vector machines with applications. *Statistical Science*, *21*(3), 322–336. <https://doi.org/10.1214/088342306000000493>
- MR ADEPU RAJESH, & DR TRYAMBAK HIWARKAR. (2023). Exploring Preprocessing Techniques for Natural LanguageText: A Comprehensive Study Using Python Code. *International Journal of Engineering Technology and Management Sciences*, *7*(5), 390–399. <https://doi.org/10.46647/ijetms.2023.v07i05.047>
- Muhajir, M., Rosadi, D., Mada, U. G., & Regency, S. (2024). *Sentiment Analysis and Topic Modelling of Bjorka Using Support Vector Machine and Latent Dirichlet Allocation*. *5*(1), 57–66.
- Murfi, H., Theresia Gowandi, S., Ardaneswari, G., & Nurrohmah, S. (2024). BERT-based combination of convolutional and recurrent neural network for Indonesian sentiment analysis. *Applied Soft Computing*, *151*(June 2023). <https://doi.org/10.1016/j.asoc.2023.111112>
- Pandey, A., & Vishwakarma, D. K. (2023). Progress, Achievements, and Challenges in Multimodal Sentiment Analysis Using Deep Learning: A Survey. *Applied Soft Computing*, *152*(December 2023), 111206. <https://doi.org/10.1016/j.asoc.2023.111206>
- Pavlov, D., Balasubramanyan, R., Dom, B., Kapur, S., & Parikh, J. (2004). Document preprocessing for Naive Bayes classification and clustering with mixture of multinomials. *KDD-2004 - Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 829–834. <https://doi.org/10.1145/1014052.1016922>
- Polyzos, Efstathios & Wang, F. (2022). Polyzos, Efstathios & Wang, Fang. *Energy Economics*, *114*.
- Rana, R. (2016). *Gated Recurrent Unit (GRU) for Emotion Classification from Noisy Speech*. 1–9. <http://arxiv.org/abs/1612.07778>
- Saberi, B., & Saad, S. (2017). Sentiment analysis or opinion mining: A review. *International Journal on Advanced Science, Engineering and Information Technology*, *7*(5), 1660–1666. <https://doi.org/10.18517/ijaseit.7.5.2137>
- Samsir, Irmayani, D., Edi, F., Harahap, J. M., Jupriaman, Rangkuti, R. K., Ulya, B., & Watrianthos, R. (2021). Naives Bayes Algorithm for Twitter Sentiment Analysis. *Journal of Physics: Conference Series*, *1933*(1). <https://doi.org/10.1088/1742-6596/1933/1/012019>
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.*, *15*, 1929–1958. <https://doi.org/10.5555/2627435.2670313>
- Talaat, A. S. (2023). Sentiment analysis classification system using hybrid BERT models. *Journal of Big Data*, *10*(1). <https://doi.org/10.1186/s40537-023-00781-w>
- Xu, C., Coen-Pirani, P., & Jiang, X. (2023). Empirical Study of Overfitting in Deep Learning for Predicting Breast Cancer Metastasis. *Cancers*, *15*(7). <https://doi.org/10.3390/cancers15071969>
- Yin, W., Kann, K., Yu, M., & Schütze, H. (2017). *Comparative Study of CNN and RNN for Natural Language Processing*. <http://arxiv.org/abs/1702.01923>
- Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. *Neural Computation*, *31*(7), 1235–1270. https://doi.org/10.1162/neco_a_01199