



Hybrid Deep Learning with GloVe and Genetic Algorithm for Sentiment Analysis on X: 2024 Election

Mahrnunissa Azmima Fitria^{1*}, Erwin Budi Setiawan²

^{1,2}School of Computing, Faculty of Informatics, Telkom University, Indonesia

Abstract.

Purpose: This research analyzes sentiment on the 2024 Indonesian Presidential Election using data from X, employing a hybrid CNN-GRU model optimized with a Genetic Algorithm (GA) to improve accuracy and efficiency. It also explores GloVe feature expansion for enhanced sentiment classification, aiming for deeper insights into public opinion through advanced deep learning and optimization techniques.

Methods: This research employs a deep learning approach that integrates Convolutional Neural Network (CNN) and Gated Recurrent Unit (GRU) models, Term Frequency-Inverse Document Frequency (TF-IDF), Global Vectors (GloVe), and GA. The dataset comprises 62,955 Indonesian tweets focusing on the 2024 General Election using various keywords.

Result: The results indicated that the Genetic Algorithm significantly improved model accuracy. The CNN-GRU + GA model achieved 84.72% accuracy for the Top 10 ranking, a 1.94% increase from the base model. In comparison, the GRU-CNN + GA model achieved 84.69% accuracy for the Top 5 ranking, a 2.76% increase from the base model, demonstrating enhanced performance with GA across configurations.

Novelty: This research uses a hybrid CNN-GRU model to introduce a novel sentiment analysis approach for the 2024 Indonesian Presidential Election. The model enhances accuracy by combining CNN's spatial feature extraction with GRU's temporal context capture and GloVe's word semantics. Genetic Algorithm optimization further refines performance. Comprehensive pre-processing ensures high-quality data, and focusing on election-specific keywords adds relevance. This study advances sentiment analysis through its innovative hybrid model, feature expansion, and optimization techniques.

Keywords: Sentiment analysis, Election, X, CNN, Gated recurrent unit, GloVe, Genetic algorithm, Hybrid deep learning

Received June 2024 / **Revised** August 2024 / **Accepted** August 2024

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).



INTRODUCTION

Sentiment analysis is the process of determining sentiment and classifying the polarity of text in a document or sentence to be categorized as positive, negative, or neutral. This technique is often used in social media, especially X, to understand the public's perception of an entity such as a particular service, product, individual, or topic [1]. X provides a real-time platform where users can express their opinions dynamically, making it a rich data source for sentiment analysis [2], [3]. In the context of the 2024 General Election in Indonesia, sentiment analysis via X can provide insights into people's opinions towards Presidential candidates and the evolving political dynamics [4].

In recent years, numerous studies have been conducted to improve the accuracy and efficiency of sentiment analysis models. Deep learning methods such as GRU (Gated Recurrent Unit) and CNN (Convolutional Neural Network) have shown high accuracy in sentiment analysis [5]–[8]. For instance, research by Kiran Baktha and his colleagues demonstrated that the GRU model achieved the highest accuracy in sentiment analysis of Amazon product reviews [9]. GRU is particularly effective in handling sequential data, such as text, due to its ability to retain information over time without the vanishing gradient problem often faced by traditional RNN models.

*Corresponding author.

Email addresses: mahrnunissa@student.telkomuniversity.ac.id (Fitria)*, erwinbudisetiawan@telkomuniversity.ac.id (Setiawan)

DOI: [10.15294/sji.v11i3.8467](https://doi.org/10.15294/sji.v11i3.8467)

Similarly, the Convolutional Neural Network (CNN) has proven to be effective in sentiment analysis. A study by Aldiansyah and Priyo demonstrated that the CNN model achieved 88.21% accuracy in sentiment analysis of public opinion towards Smartfren 4G network services [10]. CNN's ability to perform feature extraction from text data by utilizing convolution layers to detect important patterns and features makes it a valuable tool for sentiment classification.

Despite these advancements, there is still a need for more comprehensive approaches that can further improve the performance of sentiment analysis models. For example, the use of GloVe (Global Vectors for Word Representation) feature expansion has been shown to enhance sentiment analysis performance. GloVe is a word embedding model that uses global statistical information from the entire document to generate word representations in vector form. Research by Sani Kamış and his colleagues found that the use of GloVe improved sentiment analysis performance by 5%-7% compared to Word2Vec [11]. The vector representation generated by GloVe is richer in context, making it more effective in understanding and classifying text sentiment.

In addition, research by Severyn and Moschitti explored the effectiveness of CNN in sentiment analysis, but they focused primarily on sentence-level classification without integrating sequential data processing, which limits the model's ability to capture the context of the entire document [12]. On the other hand, Zhang et al. proposed a hybrid approach combining CNN with LSTM (Long Short-Term Memory) networks for text classification, highlighting the potential of hybrid models but leaving room for optimization in terms of feature selection and model efficiency. These studies underline the importance of exploring hybrid models but also point out that the combination of CNN and GRU, as well as the integration of GloVe and Genetic Algorithm, remains underexplored [13].

Optimization techniques such as Genetic Algorithms (GA) can significantly improve the accuracy of sentiment analysis models by optimizing feature selection and model parameters, thereby enhancing overall performance. Research by Riska Aryanti and her colleagues demonstrated that the GA-based Support Vector Machine algorithm improved the average accuracy value and AUC in public transportation sentiment analysis [14]. Additionally, Loussaief & Abdelkrim provided a comprehensive study on the integration of Genetic Algorithms with deep learning models, specifically highlighting how GA can be employed to optimize hyperparameters and feature selection to enhance classification performance. Their findings suggest that GA is particularly effective in improving model efficiency and accuracy in complex classification tasks. However, while their research offers valuable insights, it was primarily conducted on traditional datasets, leaving the application of GA in more dynamic and real-time environments, such as social media sentiment analysis, which was relatively unexplored [15]. This gap presents an opportunity for further exploration, particularly in the context of real-time sentiment analysis during politically charged events like elections.

However, the existing research primarily focuses on individual techniques such as CNN, GRU, or GloVe, without fully exploring the potential of combining these methods in a hybrid model. Moreover, the majority of studies do not address the specific challenges posed by analyzing sentiment in real-time political contexts, such as the 2024 Indonesian Presidential Election, where the public's opinions can rapidly shift due to emerging political events.

This research focuses on proposing a novel approach that integrates a hybrid CNN-GRU model with GloVe feature expansion and Genetic Algorithm optimization for sentiment analysis related to the 2024 Election on X. This hybrid approach leverages the strengths of CNN in terms of feature extraction and GRU in handling sequential data, which is further expected to improve the accuracy and robustness of sentiment analysis. Additionally, by incorporating Genetic Algorithm optimization, this research aims to develop a more efficient and accurate model for classifying the sentiment of public opinion towards the 2024 Election, thus providing valuable insights for political stakeholders.

The importance of sentiment analysis in the context of the 2024 Indonesian Presidential Election lies in its ability to capture and analyze the public's evolving perceptions and attitudes towards the candidates. This is crucial for understanding the dynamics of voter behavior and for informing campaign strategies. The insights generated from this analysis can help political parties and candidates to tailor their messages and strategies in real-time, potentially influencing the outcome of the election. Therefore, this study not only

contributes to the field of sentiment analysis but also holds practical significance in the context of political decision-making during elections.

METHODS

This research employed the application of a hybrid Convolutional Neural Network (CNN) and Gated Recurrent Unit (GRU) model optimized using a Genetic Algorithm for sentiment analysis of tweets related to the 2024 General Election in Indonesia [16]. Data was collected through X crawling and then processed through pre-processing stages such as data cleaning, normalization, stopwords removal, stemming, and tokenization [17]. Feature expansion was performed using GloVe to improve word representation. The CNN-GRU hybrid model was used to classify sentiment as positive, negative, or neutral, with Genetic Algorithm optimization to improve model accuracy. Performance evaluation employed the Confusion Matrix to calculate accuracy, precision, and recall values. The steps conducted in this study are shown in Figure 1.

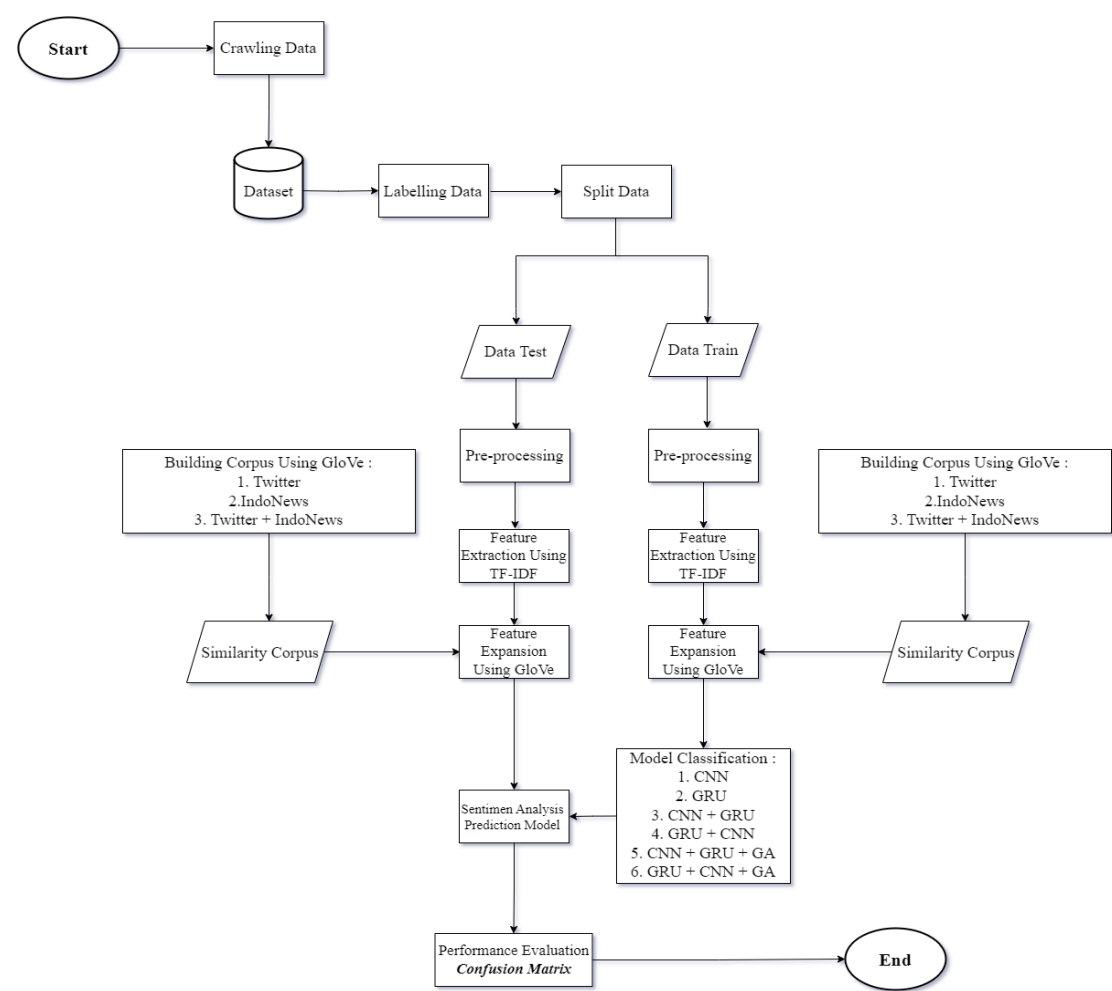


Figure 1. Research flow

Data crawling

The data was retrieved from X using a web scraping technique with the help of the X API. Crawling was conducted with keywords related to the 2024 Indonesian Presidential Election to collect relevant tweets. The following keywords were selected to capture a diverse range of discussions surrounding the election #Pilpres2024, #AniesMuhaimin, #PrabowoGibran, #PolitikIndonesia, #GanjarMahfud, and #DebatPilpres. These keywords ensured a comprehensive analysis of public sentiment by reflecting various facets of the electoral process and voter opinions. A total of 62,955 tweets were collected within a certain period before the election. Table 1 shows the amount of data collected for each keyword used in the data crawling process.

Table 1. Quantity of crawled data

Keyword	Tweet Quantity
Pilpres 2024	7.326
AniesMuhaimin	10.341
PrabowoGibran	25.420
PolitikIndonesia	4.013
GanjarMahfud	9.778
DebatPilpres	6.077

Data labeling

The pre-processed data is then manually labeled to determine the sentiment of the tweet as positive, negative, or neutral. The quantity of specific criteria for determining sentiment tweets is outlined in Table 2.

Table 2. Quantity of labeled data

Label	Quantity
Positive	21.114
Neutral	20.644
Negative	21.197

Data pre-processing

The data crawling process produces unstructured data, so it often contains noise i.e. data that is not relevant in the classification process. Therefore, data pre-processing is carried out to reduce the level of noise [18]. Five stages of data preprocessing were applied in this study, namely [19]:

- Data Cleaning**
This process involves removing irrelevant elements from the text, such as symbols, punctuation marks, numbers, URLs, and missing data. The goal is to ensure the data is cleaner and ready for analysis.
- Case Folding**
This step converts all letters in the text to lowercase. It is done to eliminate differences between uppercase and lowercase letters that could affect the analysis.
- StopWord**
This process removes words considered insignificant for analysis, such as conjunctions. By removing these words, the model can focus more on meaningful words.
- Stemming**
This step involves removing affixes from words, such as prefixes, suffixes, or infixes. This helps to simplify words to their base form, making them easier to analyze.
- Tokenization**
The final step breaks the text into separate words, usually separated by spaces. This process is essential for further analysis, such as creating text vectors.

Table 3. Example of data preprocessing

Preprocessing	Text
Original Tweet	<i>Inilah pendukung DUNGU dari PASLON 2, apa RG itu CAPRES? Klo dungu jangan ditunjukan ke publik https://twitter.com/SpentaMainyu01/status/1731859356772077715</i>
Data Cleaning	<i>Inilah pendukung DUNGU dari PASLON apa RG itu CAPRES Klo dungu jangan ditunjukan ke publik</i>
Case Folding	<i>inilah pendukung dungu dari paslon apa rg itu capres klo dungu jangan ditunjukan ke publik</i>
StopWord	<i>pendukung dungu paslon rg capres dungu ditunjukan publik</i>
Stemming	<i>dukung dungu paslon rg capres dungu tunjuk publik</i>
Tokenization	<i>["dukung", "dungu", "paslon", "rg", "capres", "dungu", "tunjuk", "publik"]</i>

Term frequency-inverse document frequency (TF-IDF)

Feature extraction affects classification precision. This research used the TF-IDF method, which is widely used. TF-IDF was a statistical method for measuring the importance of words in a document set. The tweet data in this study was assigned a TF-IDF value to determine the significance of each word. A high TF-IDF value is influenced by the high frequency of the word in the document and the low occurrence of the word in other documents. Term Frequency (TF) is the number of occurrences of word i in data j , divided by the total occurrences of all words in data [16], [19], [20]–[22]. The formula for calculating TF is [23]:

$$tf_{ij} = \frac{f_{d(i)}}{f_{d(j)}} \quad (2.1)$$

Inversed Document Frequency (IDF) aims to reduce the weight of terms contained in all documents. The following is the formula used to calculate the IDF value [23]:

$$idf(t, D) = \log \log (-)^N_{df(i)} \quad (2.2)$$

Global vector (GloVe)

GloVe is a word embedding model that represents words in vector form. Its advantage lies in its ability to use global statistical information from the entire document, thus capturing the meaning of words based on the statistical distribution in the text corpus. GloVe produces word representations that account for the relationships between words globally, making it suitable for natural language processing and tasks such as text classification [24].

Building corpus

The corpus is used to create a top-ranked dataset that shows similarity based on its rank [25]. The process begins by collecting data from X and news articles from various news sources in Indonesia, followed by data pre-processing. This top-ranked corpus is then used to expand the features. In this study, three corpus expansions were conducted: Tweet, IndoNews, and Tweet+IndoNews, with the aim of extracting additional features to enhance the machine learning model's performance. The quantity of corpus constructed using GloVe is presented in Table 4.

Table 4. Quantity of labeled data

Corpus Similarity	Quantity
Tweet	62.955
IndoNews	126.673
Tweet+IndoNews	189.628

Convolutional neural network (CNN)

CNN is a deep learning model that operates convolution and integrates multiple processing layers inspired by the biological nervous system. This modeling method was used in this research. CNNs have become a popular tool in deep learning, especially in the image-processing community. The basic structure of CNN involves a convolution layer, a pooling layer, and a fully connected layer (see Figure 2). With feature extraction, text can also be processed for classification using this method. CNN architecture consists of several layers, such as a subsampling layer, convolutional layer, loss layer, and fully connected layer [26]–[29].

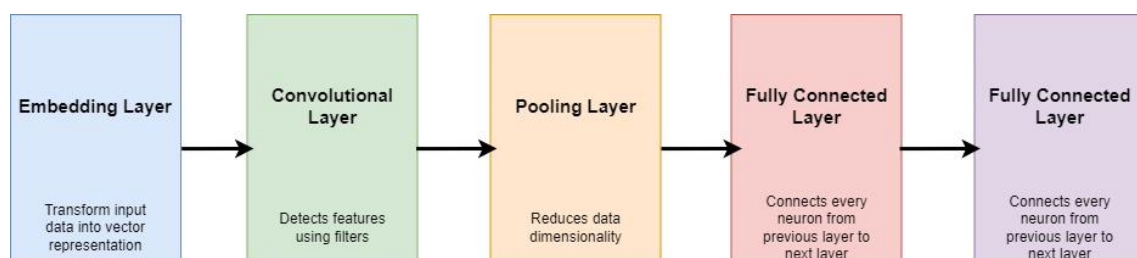


Figure 2. Model convolutional neural network [29]

Gated Recurrent Unit (GRU)

GRU is a Recurrent Neural Network (RNN) algorithm that is simpler and more efficient than Long Short-Term Memory (LSTM). GRU was developed to overcome the vanishing gradient problem and improve RNN training by using a reset gate and an update gate instead of an output gate and a forget gate. The GRU model involves three main steps: reset gate calculation, candidate hidden state calculation, and final hidden state calculation (see Figure 3). The reset gate determines the information that is forgotten or remembered, while the update gate determines the information that is passed on to the future. The candidate hidden state is calculated using a hyperbolic tangent activation function. GRU is proven to be faster and more efficient in memory usage than LSTM [17].

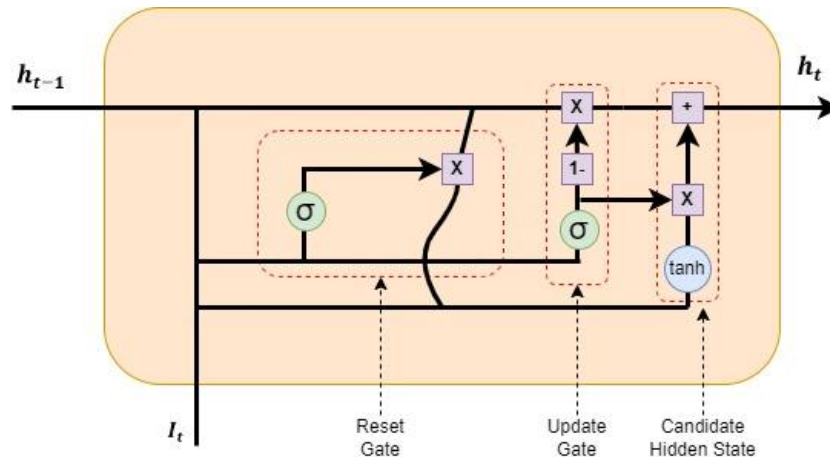


Figure 3. Model gated recurrent unit [17]

Hybrid deep learning

Hybrid Deep Learning is a new model that combines different deep learning [30]

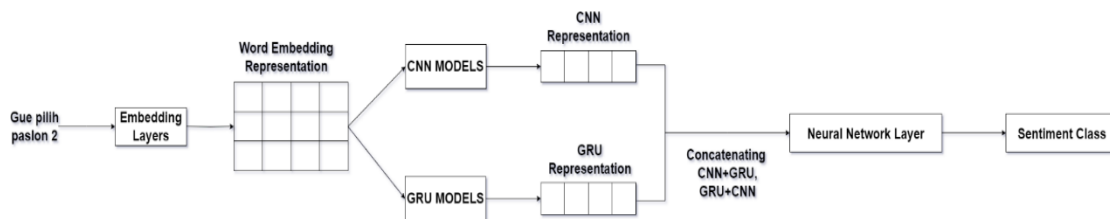


Figure 4. CNN and GRU

In this research, a combination of two classification algorithms, known as a hybrid model, was performed to achieve a higher accuracy level [30]. This approach is inspired by the results of previous research by Li-Xia Luo in the context of sentiment analysis. In her research, the use of a combined CNN and GRU model with LDA feature extraction has been shown to achieve the best accuracy score (see Figure 4). Therefore, this research applied a similar approach by combining CNN and GRU algorithms in order to achieve optimal accuracy.

Model optimization with genetic algorithm (GA)

GA is used to optimize the hyperparameters of the CNN-GRU hybrid model. This optimization process includes:

- Population Initialization**
Creating an initial population with various combinations of hyperparameters.
- Selection**
Selecting the best individuals based on the model's performance on the training data.
- Crossover and Mutation**
Generates new individuals with better hyperparameter combinations.
- Evaluation**
Measures model performance based on accuracy, precision, recall, and F1-score.

Model training and evaluation

The data was divided into training data and test data with a proportion of 80:20. The model was trained using the training data and evaluated using the test data. The evaluation metrics used include accuracy, precision, recall, and F1-score.

RESULTS AND DISCUSSIONS

This research evaluates the CNN-GRU hybrid model optimized with a Genetic Algorithm (GA) for sentiment analysis of tweets related to the 2024 General Election in Indonesia. Several experimental scenarios were conducted to measure the performance of the model with and without additional features, as well as with various preprocessing techniques. The results show that the use of the CNN-GRU hybrid model with GA optimization and GloVe feature expansion significantly improves the accuracy of sentiment analysis compared to individual models.

Table 5. First scenario: base model (CNN and GRU without feature extraction)

Split Ratio	Accuracy CNN	Accuracy GRU
90:10	83.01%	82.41%
80:20	83.10%	82.38%
70:30	81.98%	81.61%
60:40	82.25%	80.92%

This study evaluates the performance of a CNN-GRU hybrid model optimized with a Genetic Algorithm (GA) for sentiment analysis of tweets related to the 2024 General Election in Indonesia. In the first scenario, the basic CNN and GRU models without feature extraction show that CNN is slightly superior to GRU. The highest accuracy was achieved by CNN with 83.10% at a split ratio of 80:20, while GRU achieved the highest accuracy of 82.41% at a split ratio of 90:10. See Table 5. These results show that CNN has a better ability to handle raw data than GRU, with especially in the context of sentiment analysis on X.

Table 6. Second scenario: use of TF-IDF

Max Feature	Accuracy CNN	Accuracy GRU
2000	82.72% (-0,46)	82.56% (+0,18)
5000	83.35% (+0,30)	82.61% (+0,24)
10000	83.28% (+0,21)	82.92% (+0,61)
15000	83.56% (+0,55)	82.49% (+0,09)
20000	83.33% (+0,27)	82.44% (+0,03)

In the second scenario, the use of TF-IDF method for feature extraction shows an increase in accuracy in both CNN and GRU. CNN achieved the highest accuracy of 83.56% at Max Feature 15000, while GRU achieved the highest accuracy of 82.92% at Max Feature 10000. See Table 6. This improvement showed that the number of features extracted had a significant impact on model performance. CNN showed higher sensitivity to increasing the number of features than GRU, which remained stable but slightly lower.

Table 7. Third scenario: N-gram usage

N-Gram	Accuracy CNN	Accuracy GRU
Unigram (1, 1)	83.56% (+0,55)	82.92% (+0,61)
Bigram (2, 2)	78.10% (-6,01)	82.07% (-0,41)
Trigram (3, 3)	71.69% (-13,7)	82.18% (-0,27)
Unigram + Bigram (1, 2)	83.37% (+0,32)	81.98% (-0,52)
Bigram + Trigram (2, 3)	78.22% (-5,87)	81.68% (-0,88)
Unigram + Bigram + Trigram (1, 3)	83.30% (+0,24)	81.28% (-1,37)

The third scenario evaluated the use of N-Grams in feature extraction. The results showed that the Unigram combination gave the best results for CNN with 83.56% accuracy, while GRU showed the best performance on Unigram with 82.92% accuracy. See Table 7. The combination of various N-Grams generally improved the accuracy of CNN models more significantly than GRU.

Table 8. Fourth scenario: similarity corpus (CNN)

Rank	Tweet	News	Tweet + News
TOP 1	83.76% (+0,79)	82.97% (-0,15)	82.79% (-0,37)
TOP 5	83.70% (+0,72)	83.12% (+0,02)	82.76% (-0,40)
TOP 10	82.86% (-0,28)	82.81% (-0,34)	83.89% (+0,95)
TOP 15	82.30% (-0,96)	82.93% (-0,20)	82.42% (-0,81)

Table 9. Fourth scenario: similarity corpus (GRU)

Rank	Tweet	News	Tweet + News
TOP 1	82.68% (+0,32)	81.78% (-0,76)	82.71% (+0,36)
TOP 5	82.54% (+0,15)	82.03% (-0,46)	82.49% (+0,09)
TOP 10	82.97% (+0,67)	82.95% (+0,65)	81.64% (-0,93)
TOP 15	82.27% (-0,16)	82.34% (-0,08)	82.37% (-0,04)

In the fourth scenario, the similarity corpus feature was used to evaluate the performance of the model on tweet data, news, and a combination of both. The results showed that the CNN and GRU methods achieved the highest accuracy on Tweet data, ranking TOP 1 with 83.76% accuracy for CNN and TOP 10 with 82.97% accuracy for GRU. See Table 8 and 9. This indicated that Tweet data provided better results. This suggests that tweets as a data source have higher quality and relevance in sentiment analysis related to elections than other data sources. News data and tweet+news combinations tend to yield lower accuracy, indicating that focusing on one consistent data source is more effective.

Table 10. Fifth scenario: hybrid CNN-GRU

Rank	Tweet	News	Tweet + News
TOP 1	81.33% (+0,13)	82.23% (+1,24)	81.94% (+0,88)
TOP 5	81.38% (+0,19)	80.84% (-0,46)	81.26% (+0,04)
TOP 10	82.29% (+1,31)	81.37% (+0,18)	81.29% (+0,08)
TOP 15	81.61% (+0,48)	81.73% (+0,62)	81.53% (+0,38)

Table 11. Fifth scenario: hybrid GRU-CNN

Rank	Tweet	News	Tweet + News
TOP 1	81.69% (+0,66)	81.86% (+0,87)	82.49% (+1,65)
TOP 5	81.98% (+1,02)	82.00% (+1,04)	81.64% (-0,93)
TOP 10	82.29% (+1,40)	82.16% (+1,24)	81.23% (-1,43)
TOP 15	82.25% (+1,35)	81.30% (+0,18)	82.32% (-0,10)

The fifth scenario evaluated the performance of the CNN-GRU and GRU-CNN hybrid models. The results showed that the GRU-CNN hybrid model provided higher accuracy than CNN-GRU, especially on the tweet and news data (TOP 1 ranking), with accuracy reaching 82.49% (Table 10 and 11). This indicated that the GRU-CNN model sequence was more effective in handling more complex data. This better performance could be due to GRU ability to capture long-term dependencies before CNN further extracts the features.

Table 12. Sixth scenario: hybrid CNN-GRU + GA

Rank	Tweet		News		Tweet News	+
TOP 1	82.31% (0,95)	(-)	81.29% (2,17)	(-)	82.70% (0,48)	(-)
TOP 5	83.33% (+0,27)		83.93% (+0,99)		83.88% (+0,93)	
TOP 10	84.72% (+1,94)		82.96% (0,16)	(-)	82.82% (0,33)	(-)
TOP 15	84.30% (+1,44)		83.88% (+0,93)		83.45% (+0,42)	

Table 13. Sixth scenario: hybrid GRU-CNN + GA

Rank	Tweet	Berita	Tweet + Berita
TOP 1	82.32% (-0,10)	81.89% (-0,63)	82.25% (-0,19)
TOP 5	84.69% (+2,76)	82.91% (+0,60)	82.76% (+0,37)
TOP 10	83.59% (+1,40)	84.22% (+2,19)	83.96% (+1,88)
TOP 15	83.49% (+1,31)	83.89% (-1,79)	83.78% (+1,66)

To improve the accuracy of the CNN-GRU and GRU-CNN models, the sixth scenario incorporated Genetic Algorithm optimization into the hybrid model. The results demonstrated that the use of GA significantly enhanced the model's accuracy compared to the version without optimization. With GA optimization, the CNN-GRU model achieved a higher prediction accuracy when using Tweet data with the TOP 10 keywords, increasing from 82.29% to 84.72%, and the GRU-CNN model showed a 2.76% improvement over the model without GA in TOP 5 keywords with Tweet data (Table 12). Notably, the accuracy achieved with GA optimization surpassed the highest accuracy obtained by the models without GA, highlighting the effectiveness of GA in identifying the optimal parameters that enhance the model's performance in sentiment classification.

The overall results show that the combination of deep learning methods and feature extraction techniques can provide significant improvements in sentiment analysis. The use of TF-IDF, N-Gram, and similarity corpus helps in improving the accuracy of the model. While the CNN-GRU hybrid model performed well, the GRU-CNN sequence proved more effective in some cases. The addition of the genetic algorithm also contributed to the improvement of accuracy, indicating further potential for optimization.

The results from this study underscore the effectiveness of incorporating Genetic Algorithm (GA) optimization into hybrid deep learning models. To build on these findings, future research could explore several areas. First, examining the application of GA optimization in other advanced models could provide insights into its comparative effectiveness. Second, investigating the impact of contextual features, such as user demographics or tweet location, on sentiment analysis could improve model accuracy. Additionally, developing real-time sentiment monitoring systems could facilitate dynamic tracking of public opinion during key events. These approaches could further improve the accuracy and applicability of sentiment analysis models.

Comparison between each scenario

The study evaluated the significance of accuracy changes in each scenario using the Z-value and P-value for statistical testing. A 95% confidence level was applied for the Z-value. Changes were deemed highly significant if Z-Value > 1.96 and P-Value < 0.01, significant if Z-Value > 1.96 and P-Value < 0.05, and insignificant otherwise. Table 13 presents the results of the statistical significance testing for all scenarios. According to the table, the changes in the scenarios S1→S2, S2→S3, S3→S4, and S4→S5 did not show statistically significant improvements as their Z-Values and P-Values did not meet the criteria for significance (Z-Value > 1.96 and P-Value < 0.05). However, the change from scenario S1→S6 showed a significant improvement with a Z-Value of 5.90 and a P-Value of 3.44. This demonstrates that the progression from the baseline scenario S1 to the final scenario S6 can significantly enhance the performance. The other scenarios did not exhibit statistically significant changes based on the given confidence level and criteria. See Table 14.

Table 14. Result of statistical significance tests in various scenarios

	S1→S2	S2→S3	S3→S4	S4→S5	S5→S6	S1→S6
Z-Value	19.40	nan	1.357	0.044	0.975	5.909
P-Value	0.000	nan	0.174	0.964	0.329	3.440
Significant?	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE

This research shows the importance of using hybrid models and feature optimization in social media sentiment analysis. By focusing on relevant data and appropriate feature extraction techniques, models can be more effective in capturing sentiment and providing more accurate insights. In the context of the 2024 General Election, this approach can help in understanding public opinion and evolving political dynamics, providing a useful tool for researchers and policymakers.

CONCLUSION

Expansion optimized by a Genetic Algorithm for sentiment analysis on X related to the 2024 Indonesian Presidential Election. The results show that the use of a Genetic Algorithm significantly improves the accuracy of the sentiment analysis model. The CNN-GRU + GA hybrid model achieved the highest accuracy of 84.72% for the TOP 10 ranking, and the GRU-CNN + GA hybrid model achieved the maximum accuracy of 84.69% for the TOP 5 ranking. Compared to the study that applies a CNN-GRU model without GA to analyze the sentiment of Bank BCA shares, this study demonstrated that the GloVe feature expansion enhances the GRU model's accuracy to 74.56% from 74.24% with TF-IDF, although it decreased the CNN model's accuracy by 73.9%. The combined CNN-GRU model showed an improvement, outperforming individual CNN and GRU models with a marginal increase in accuracy (+1.88 for CNN and +1.86 for GRU with GloVe and TF-IDF). The substantial improvement in accuracy in the second study (approximately 10% higher than the first study) underscored the effectiveness of the Genetic Algorithm in enhancing model performance. This suggests that incorporating optimization techniques like GA can significantly boost the accuracy and robustness of sentiment analysis models. For future work, it is suggested to explore other deep learning models to compare with the CNN-GRU approach. Also, it can be expanded to the investigation into how contextual features, like user demographics or tweet location, impact on sentiment analysis. Additionally, future research can develop real-time sentiment monitoring systems to dynamically track public opinion during key events.

REFERENCES

- [1] W. A. Prabowo and C. Wiguna, "Sistem Informasi UMKM Bengkel Berbasis Web Menggunakan Metode SCRUM," *J. Media Inform. Budidarma*, vol. 5, no. 1, p. 149, 2021, doi: 10.30865/mib.v5i1.2604.
- [2] S. Riyadi, N. Gita Mahardika, C. Damarjati, and S. Ramli, "Modified Convolutional Neural Network for Sentiment Classification: A Case Study on The Indonesian Electoral Commission," *Sci. J. Informatics*, vol. 11, no. 2, p. 493, 2024, doi: 10.15294/sji.v11i2.4929.
- [3] J. Jumanto, M. A. Muslim, Y. Dasril, and T. Mustaqim, "Accuracy of Malaysia Public Response to Economic Factors During the Covid-19 Pandemic Using Vader and Random Forest," *J. Inf. Syst. Explor. Res.*, vol. 1, no. 1, pp. 49–70, 2022, doi: 10.52465/joiser.v1i1.104.
- [4] and U. A. J. Khatib Sulaiman, D. Setiyawati, N. Cahyono, "Analisa Sentimen Pengguna Sosial Media Twitter Terhadap Perokok di Indonesia," *Indones. J. Comput. Sci. Attrib.*, vol. 12, no. 1, p. 262, 2023.
- [5] M. Alruily, "Sentiment analysis for predicting stress among workers and classification utilizing CNN: Unveiling the mechanism," *Alexandria Eng. J.*, vol. 81, no. August, pp. 360–370, 2023, doi: 10.1016/j.aej.2023.09.040.
- [6] S. Liao, J. Wang, R. Yu, K. Sato, and Z. Cheng, "CNN for situations understanding based on sentiment analysis of twitter data," *Procedia Comput. Sci.*, vol. 111, no. 2015, pp. 376–381, 2017, doi: 10.1016/j.procs.2017.06.037.
- [7] M. Usama *et al.*, "Removal notice to 'Equipping recurrent neural network with CNN-style attention mechanisms for sentiment analysis of network reviews' (Equipping recurrent neural network with CNN-style attention mechanisms for sentiment analysis of network reviews (2019) 148 (98–106), (S0140366419305456), (10.1016/j.comcom.2019.11.009))," *Comput. Commun.*, vol. 148, p. 250, 2019, doi: 10.1016/j.comcom.2019.11.009.
- [8] Z. Ahmed and J. Wang, "A fine-grained deep learning model using embedded-CNN with BiLSTM

- for exploiting product sentiments,” *Alexandria Eng. J.*, vol. 65, pp. 731–747, 2023, doi: 10.1016/j.aej.2022.10.037.
- [9] I. of E. and E. Engineers, D. of E. and C. E. Adhiparasakthi Engineering College, and M. S. Institute of Electrical and Electronics Engineers, “Proceedings of the 2017 IEEE International Conference on Communication and Signal Processing (ICCSPP),” 2017.
 - [10] D. of I. Universitas Diponegoro, Faculty of Science and Mathematics, I. S. Institute of Electrical and Electronics Engineers, I. S. B. Universitas Diponegoro, and I. of E. and E. Engineers, “ICICoS 2019: The 3rd International Conference on Informatics and Computational Sciences: Proceedings: October 29th-30th, 2019, Semarang, Central Java, Indonesia,” 2019.
 - [11] S. Kaniş and D. Goularas, “Evaluation of Deep Learning Techniques in Sentiment Analysis from Twitter Data,” *Proc. - 2019 Int. Conf. Deep Learn. Mach. Learn. Emerg. Appl. Deep. 2019*, pp. 12–17, 2019, doi: 10.1109/Deep-ML.2019.00011.
 - [12] A. Severyn and A. Moschitti, “Twitter Sentiment Analysis with deep convolutional neural networks,” *SIGIR 2015 - Proc. 38th Int. ACM SIGIR Conf. Res. Dev. Inf. Retr.*, pp. 959–962, 2015, doi: 10.1145/2766462.2767830.
 - [13] Y. Zhang and B. Wallace, “A Sensitivity Analysis of (and Practitioners’ Guide to) Convolutional Neural Networks for Sentence Classification,” 2015, [Online]. Available: <http://arxiv.org/abs/1510.03820>
 - [14] R. Aryanti, A. Saryoko, A. Junaidi, S. Marlina, Wahyudin, and L. Nurmalia, “Comparing Classification Algorithm with Genetic Algorithm in Public Transport Analysis,” *J. Phys. Conf. Ser.*, vol. 1641, no. 1, 2020, doi: 10.1088/1742-6596/1641/1/012017.
 - [15] S. Loussaief and A. Abdelkrim, “Convolutional Neural Network hyper-parameters optimization based on Genetic Algorithms,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 10, pp. 252–266, 2018, doi: 10.14569/IJACSA.2018.091031.
 - [16] M. F. Avidiansyah and E. B. Setiawan, “The Influence of Sentiment on the Movement of Bank BCA (BBKA) Shares with the CNN-GRU Classification Model and Feature Expansion with GloVe,” *2023 Int. Conf. Data Sci. Its Appl. ICoDSA 2023*, no. August 2023, pp. 105–110, 2023, doi: 10.1109/ICoDSA58501.2023.10277052.
 - [17] K. U. Wijaya and E. B. Setiawan, “Hate Speech Detection Using Convolutional Neural Network and Gated Recurrent Unit with FastText Feature Expansion on Twitter,” *J. Ilm. Tek. Elektro Komput. dan Inform.*, vol. 9, no. 3, pp. 619–631, 2023, [Online]. Available: 10.26555/jiteki.v9i3.26532
 - [18] R. Alshalan and H. Al-Khalifa, “A deep learning approach for automatic hate speech detection in the saudi twittersphere,” *Appl. Sci.*, vol. 10, no. 23, pp. 1–16, 2020, doi: 10.3390/app10238614.
 - [19] K. U. WIJAYA, “Deteksi Ujaran Kebencian Menggunakan CNN+GRU dengan Ekspansi Fitur FastText di Twitter,” Universitas Telkom, 2023. [Online]. Available: <https://repositori.telkomuniversity.ac.id/pustaka/203982/deteksi-ujaran-kebencian-menggunakan-cnn-gru-dengan-ekspansi-fitur-fasttext-di-twitter-dalam-bentuk-pengganti-sidang-artikel-jurnal.html>
 - [20] M. D. D. SREYA, “Penggunaan Metode GloVe untuk Ekspansi Fitur pada Analisis Sentimen Twitter dengan Naïve Bayes dan Support Vector Machine,” Universitas Telkom, 2022. [Online]. Available: <https://repositori.telkomuniversity.ac.id/pustaka/178395/penggunaan-metode-glove-untuk-ekspansi-fitur-pada-analisis-sentimen-twitter-dengan-na-ve-bayes-dan-support-vector-machine.html>
 - [21] S. Akuma, T. Lubem, and I. T. Adom, “Comparing Bag of Words and TF-IDF with different models for hate speech detection from live tweets,” *Int. J. Inf. Technol.*, vol. 14, no. 7, pp. 3629–3635, 2022, doi: 10.1007/s41870-022-01096-4.
 - [22] S. W. Kim and J. M. Gil, “Research paper classification systems based on TF-IDF and LDA schemes,” *Human-centric Comput. Inf. Sci.*, vol. 9, no. 1, 2019, doi: 10.1186/s13673-019-0192-7.
 - [23] M. Nurjannah and I. Fitri Astuti, “PENERAPAN ALGORITMA TERM FREQUENCY-INVERSE DOCUMENT FREQUENCY (TF-IDF) UNTUK TEXT MINING Mahasiswa S1 Program Studi Ilmu Komputer FMIPA Universitas Mulawarman Dosen Program Studi Ilmu Komputer FMIPA Universitas Mulawarman,” *J. Inform. Mulawarman*, vol. 8, no. 3, pp. 110–113, 2013.
 - [24] A. Nurdin, B. Anggo Seno Aji, A. Bustamin, and Z. Abidin, “Perbandingan Kinerja Word Embedding Word2Vec, Glove, Dan Fasttext Pada Klasifikasi Teks,” *J. Tekno Kompak*, vol. 14, no. 2, p. 74, 2020, doi: 10.33365/jtk.v14i2.732.
 - [25] I. Gde Bagus Janardana Abasan and E. B. Setiawan, “Empowering hate speech detection: leveraging linguistic richness and deep learning,” *Bull. Electr. Eng. Informatics*, vol. 13, no. 2, pp.

- 1371–1382, 2024, doi: 10.11591/eei.v13i2.6938.
- [26] H. K. Farid, E. B. Setiawan, and I. Kurniawan, “Selection for Hoax News Detection on Twitter using Convolutional Neural Network,” *Indones. J. Comput.*, vol. 5, no. December 2020, pp. 23–36, 2020, doi: 10.34818/indojc.2021.5.3.506.
 - [27] M. Zulqarnain, R. Ghazali, Y. M. M. Hassim, and M. Rehan, “A comparative review on deep learning models for text classification,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 19, no. 1, pp. 325–335, 2020, doi: 10.11591/ijeecs.v19.i1.pp325-335.
 - [28] W. Meng, Y. Wei, P. Liu, Z. Zhu, and H. Yin, “Aspect Based Sentiment Analysis with Feature Enhanced Attention CNN-BiLSTM,” *IEEE Access*, vol. 7, pp. 167240–167249, 2019, doi: 10.1109/ACCESS.2019.2952888.
 - [29] A. Mulyanto, E. Susanti, F. Rossi, W. Wajiran, and R. I. Borman, “Penerapan Convolutional Neural Network (CNN) pada Pengenalan Aksara Lampung Berbasis Optical Character Recognition (OCR),” *J. Edukasi dan Penelit. Inform.*, vol. 7, no. 1, p. 52, 2021, doi: 10.26418/jp.v7i1.44133.
 - [30] M. U. Salur and I. Aydin, “A Novel Hybrid Deep Learning Model for Sentiment Classification,” *IEEE Access*, vol. 8, pp. 58080–58093, 2020, doi: 10.1109/ACCESS.2020.2982538.