

Sentiment Analysis of the 2024 Indonesian Presidential Election using *Fasttext* Feature Expansion and *Recurrent Neural Network* (RNN) with *Genetic Algorithm* Optimization

Inggit Restu Illahi ¹, Erwin Budi Setiawan ^{2*}

^{1,2}*School Of Computing, Telkom University*

Jl. Telekomunikasi No.1 Terusan Buah Batu, Bandung, Jawa Barat, Indonesia, 40257

*

erwinbudisetiawan@telkomuniversity.ac.id

Abstract

Social media is a place to express opinions or feelings, both positive and negative. One of them is to express opinions or feelings about a topic that is currently being discussed. The number of opinions or sentiments related to a topic can be challenging to assess if it leans towards positivity or negativity. Therefore, Sentiment analysis is essential for examining the viewpoints or sentiments on the topic. This study tested 37,391 Twitter user comments on the 2024 Indonesian presidential election. This research employs the RNN methodology, TF-IDF feature extraction, and FastText feature expansion utilizing an IndoNews corpus of as much as 142,545 data and using Genetic Algorithm optimization. The outcomes of this study yielded the highest accuracy when combining TF-IDF feature extraction with max 7000 features, FastText feature expansion on top 5 features, and implementing Genetic Algorithm optimization with a value of 82.72%, accuracy increased by 3.4% from the baseline.

Keywords: FastText Genetic Algorithm, RNN, Sentiment analysis, TF-IDF

I. INTRODUCTION

The use of social media is almost inseparable from people's lives. In this context, social media facilitates interaction between individuals in cyberspace [1]. A favored social media platform among Indonesians is Twitter, which is utilized for sharing personal experiences and gaining diverse perspectives on an issue through user interactions [2]. In this context, the 2024 Indonesian presidential election is one of the topics that attracts the attention of social media users. As a hotly-discussed issue, the presidential election attracted the attention of many people, and Twitter became the main channel for expressing opinions and sentiments related to the topic. However, due to the large number of users, opinions or sentiments towards a topic can vary and it is difficult to assess whether they are positive or negative. Therefore, sentiment analysis becomes very important to understand and evaluate users' views and feelings towards the Indonesian presidential election. By involving sentiment analysis, this research can provide in-depth insights into people's various viewpoints regarding the presidential election, helping to identify trends and opinion patterns that may emerge on social media platforms such as Twitter.

Sentiment analysis is a branch of research that examines human opinions and sentiments on a topic, where the opinions or sentiments produce positive, negative, and neutral values [3][4]. Sentiment analysis research uses deep learning as a processing method that allows computers to understand complex concepts by dissecting them

into simpler components. The deep learning method that is often used is Recurrent Neural Network (RNN) [5], [3], [6], [7]. RNN is a form of Artificial Neural Network (ANN) architecture designed to analyze sentiment based on its ability to process sequential data, especially text data, thus effectively capturing temporal dependencies and patterns in human opinions and sentiments [3]. RNNs have recurrent units that allow them to consider previous contexts, making them effective for Natural Language Processing (NLP) [7]. Several parameters need to be configured in RNN, namely, the number of hidden layers, and the number of nodes. Errors in defining the parameters and their values can result in the use of more time and resources [8].

One alternative that can be used to find the optimal parameters is to use Genetic Algorithm (GA). GA is effective in improving model performance [9][10][11][12]. In addition, the selection of Genetic Algorithm as the optimal parameter search method was chosen because GA is proven effective in improving model performance, can converge quickly and is easy to implement [11]. The author also uses FastText feature expansion in this research. FastText feature expansion is used because it can improve accuracy [13].

This approach is based on the skip-gram model, where each word in the sentence is formed as an n-gram character [14]. Research conducted by Mikolov [15] says that FastText has the potential to produce better word elimination in handling unrecognized words. So, this research will use the RNN method with the expansion of FastText features using Genetic Algorithm optimization to analyze the sentiment of user comments on Twitter.

The main contribution of this research is to find the best accuracy value of the RNN model using TF-IDF feature extraction, Genetic Algorithm optimization, and FastText feature expansion on conversations about the 2024 Indonesian presidential election observed on Twitter. Based on the author's knowledge, no research has been conducted on this research, and this can increase the accuracy value of sentiment analysis. This research applies several methods such as RNN, TF-IDF feature extraction, FastText feature expansion, and Genetic Algorithm optimization. This report includes the following sections, with Section 2 providing related studies. Section 3 presents the system built in the study. Section 4 presents the results and analysis, and section 5 provides the conclusion of this research.

II. LITERATURE REVIEW

Sentiment analysis has been done using various models, one of which is deep learning. Many studies have focused on analyzing social media data, especially those related to certain events or topics. The number of people using social media platforms such as Twitter continues to increase, exceeding 500 million tweets generated by users every day [16]. This attracts the attention of academics, so many studies have been conducted to obtain important information about events that occur on social media such as Twitter.

Research by S.C. Nistor et al [5], in building a sentiment analysis system on Twitter using Recurrent Neural Network (RNN), resulted in an accuracy value of 80.74%. In the study, a Twitter dataset of 1,578,627 tweets was used. This dataset is considered representative because it not only includes textual content consisting of words and numbers, but also includes punctuation marks that form emoticons, hashtags, and emojis. This research uses a deep learning approach with RNN as the main model. However, there is no further information regarding the steps taken to overcome overfitting.

Furthermore, research by Dimuthu Lakmal, et al [17] entitled "Word Embedding Evaluation for Sinhala". In their study, the researchers evaluated various word embedding techniques, including FastText, Word2vec, and Glove, specifically for the Sinhala language. In this evaluation process, the authors used both intrinsic and extrinsic evaluation methods. Of the two, FastText outperformed Word2vec and Glove in terms of overall accuracy, especially when using 300 vector dimensions. This research includes the evaluation of three different types of word embedding, providing a better understanding of the performance of each method for Sinhala. Although FastText with 300 vector dimensions gave the best results, it is important to consider the trade-off between accuracy and computational requirements.

In a study titled "Aspect-Level Sentiment Analysis of Beauty Product Reviews Using Chi-Square and Naive Bayes" [18], the researchers used n-gram and TF-IDF methods to improve the accuracy of the model. The performance results of this study include accuracy of 80.18%, recall of 72.49%, and F1 value of 74.73%. There are shortcomings in this study, namely, the researchers did not translate the review data. By achieving an accuracy of 80.18%, recall of 72.49%, and F1 score of 74.73%, this study provides evidence that the combination of n-gram and TF-IDF methods is effective in handling sentiment analysis tasks at the aspect level.

This research will refer to the above results. In addition, this research will contribute specifically to the development of sentiment analysis regarding the 2024 Indonesian Presidential election by applying the RNN, TF-IDF, Fasttext, and Genetic Algorithm models. Through these applications, it is expected that the research can overcome or improve the imperfections identified in previous research and can improve the accuracy of sentiment analysis of opinions regarding the 2024 Indonesian Presidential election.

III. RESEARCH METHOD

A. System Design

Figure 1 illustrates a flowchart providing a comprehensive depiction of the system implemented in the present research paper. In this research, the authors built a sentiment analysis system using the Recurrent Neural Networks (RNN) method, TF-IDF feature extraction, Fasttext feature expansion and Genetic Algorithm optimization. This research begins with the process of collecting data from twitter social media to analyze the sentiment of the data taken. Then manually labeling the data that has been taken. In the context of this sentiment analysis, labels in the form of sentiment classification are positive and negative. Next is the preprocessing stage to clean and process the initial data.

The numerical representation of the text is obtained through the TF-IDF scheme, which measures the importance of each word in a document against the entire dataset. Next, data from IndoNews is taken to enrich the model in the feature expansion stage. After that, the data is divided into two parts: train data to train the model and test data for model performance evaluation.

The sentiment model uses RNN to understand sequential relationships in text data, enabling sentiment analysis at the sentence or paragraph level. Genetic Algorithm (GA) is applied to automatically optimize the parameters in the RNN model. Next, the trained model is used to predict sentiment from previously unseen data. Finally, the final stage represents the completion of the system flow, marking the completion of the sentiment analysis process.

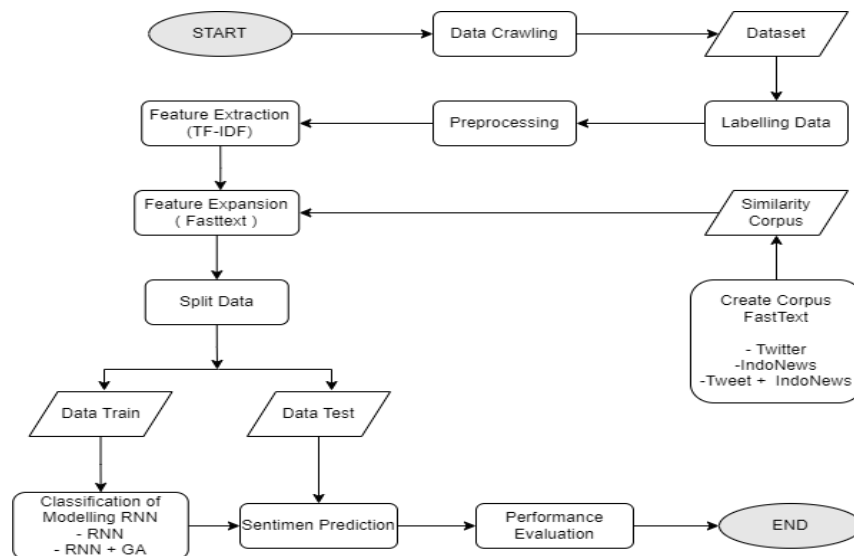


Fig. 1. Flowchart System Sentiment Analysis

B. Crawling Data

Data is obtained through crawling a dataset, a process that entails establishing a connection between the Twitter API and Python programs to retrieve CSV files with “capres”, “calon presiden”, “anies baswedan”, “prabowo subianto” dan “ganjar pranowo” keywords on Table I.

TABLE I
Keyword Data Crawling

Keyword	Ammount	Ratio (%)
Anies Baswedan	10,434	27.90
Ganjar Pranowo	8,027	21.47
Capres	7,296	19.51
Calon Presiden	6,972	18.65
Prabowo Subianto	4,662	12.47
Total	37,391	100

C. Labelling Data

After completing the data collection process and obtaining the dataset, the next step is data labeling. The labeling process is done manually, where each dataset is assigned a sentiment label based on positive (1) and negative (-1) categories. To enhance the credibility of the research, special measures were taken to ensure the validity of the labels. The labeling team consisted of individuals who already knew and recognized the nuances of sentiment. In addition, the dataset was given to 3 raters for checking. This entire process was designed to reduce the possibility of labeling errors and ensure that the labels assigned accurately reflect the true sentiment of the data in the context of this study. The total amount of data and sample data can be seen in Table II and Table III.

TABLE II
Number of Sentiment Labels

Label	Amount	Ratio (%)
Positif	21,866	58.48
Negative	15,525	41.52
Total Data	37,391	100

TABLE III
Example of Dataset

Label	Data
Positif	“Wah pak ganjar pro bgt bkin kue nya krna sedari kecil sring mmbantu ibunya mmbuat kue. Suami idaman bgt”
Negative	“1.000% yakin calon presiden yang suka sama bokep akan tenggelam”

D. Preprocessing Data

One of the most important steps in preparing data for categorization is preprocessing. Data preprocessing is a critical step in data processing that aims to clean, simplify, and prepare raw data to be processed more effectively in analysis or modeling. This process is very important to ensure that the data used in this research is of good quality. In the context of sentiment analysis, data preprocessing aims to optimize text data to suit the needs of the model to be used. In this research, preprocessing is divided into several stages, namely Data Cleansing, Case Folding, Tokenizing, Convert Slang Words, Stopword Removal, and Stemming. This process helps improve the accuracy and performance of the sentiment analysis model in this study.

1) *Data Cleansing*: The goal is to remove punctuation, numbers, symbols, hashtags, URLs, and emoticons from the data set. It aims to simplify the text and remove elements that are not needed in sentiment analysis.

2) *Case Folding*: Aims to convert text from uppercase and lowercase letters to lowercase. This is done so that the model can recognize words with different cases as the same word, thus reducing complexity and improving consistency in text analysis.

3) *Tokenizing*: Aims to convert the text in the document into tokens or words. By tokenizing, the text is broken down into the smallest units that can be processed, making the next steps of analysis easier.

4) *Convert Slang Words*: Aims to convert slang and nonstandard words into standardized words. This helps improve the consistency and accuracy of sentiment analysis, as words often used in social media contexts are often informal.

5) *Stopword Removal*: The process of removing words that are considered unnecessary, such as conjunctions. Removing stopwords helps focus on words that have a greater impact on expressing sentiment.

6) *Stemming*: The aim is to convert a compound word into a base word. Through the implementation of stemming, the model can reduce word variations to their basic form. This enhancement enables the model to better comprehend the core meaning of the text, unaffected by the morphological variations of words. The table below displays the data that has been cleaned after going through the preprocessing process.

TABLE IV
 Results of the dataset after preprocessing

Label	Data before preprocessing	Data after preprocessing
Positif	“Wah pak ganjar pro bgt bkin kue nya krna sedari kecil sring mmbantu ibunya mmbuat kue. Suami idaman bgt”	“ganjar profesional bikin kue suami idam”
Negative	“1.000% yakin calon presiden yang suka sama bokep akan tenggelam”	“calon presiden suka bokep tenggelam”

E. *Feature Extraction TF-IDF*

TF-IDF is an algorithm that computes the weight of each term and assesses the significance of a term within a document. The frequency with which a word appears in a document is known as term frequency (TF). When assigning weight to a term, inverse document frequency (IDF) is utilized to determine the term's frequency of recurrence throughout the document. [41].

TF (Term Frequency) is a metric that gauges the frequency of a particular term within a document, computed as (1).

$$TF = \frac{\text{The number of words in the document}}{\text{Total number of words in the document}} \tag{1}$$

IDF (Inverse Document Frequency) assesses the commonality or rarity of a word across a collection of documents utilizing (2).

$$IDF = \log \frac{\text{Total number of documents}}{\text{Number of documents containing words}} \tag{2}$$

The TF-IDF score computation is demonstrated in (3). It suggests that words having a high frequency within

a given document (TF) and infrequent occurrence over the entire document corpus (IDF) will yield a high TF-IDF score.

$$TF-IDF = TF \times IDF \tag{3}$$

F. Feature Expansion FastText

Feature expansion is an approach to expand the features used to evaluate and understand sentiment from text or data. The features in sentiment analysis include words to identify and assess the sentiment contained in the text. Feature expansion is performed after the feature extraction process is completed. In the context of feature expansion, the initial feature extraction process has been performed to obtain the basic features relevant to sentiment analysis. However, sometimes those features may not be enough to accurately describe the sentiment contained in the text or data. Therefore, by using feature expansion techniques, researchers can expand existing features by adding new features that can improve sentiment understanding. One of the methods used in this research is feature expansion using FastText. FastText is an embedding model based on linear techniques developed by the AI research team at Facebook [19]. This model uses subword vectors through a skipgram model with n-gram characters [20]. In the context of feature expansion, FastText is used to generate a corpus that will be used to expand existing features. The corpus generated with FastText can highlight the semantic similarity between words in the text. FastText has advantages in foreign language processing as it can simplify language processing and reduce dependency on data preprocessing steps [15].

The FastText model contains a corpus obtained from Twitter user comments on the 2024 presidential election in Indonesia and the IndoNews dataset. IndoNews covers Indonesian news, including Liputan 6, Detik.com, CNN Indonesia, Kompas, Republika, and Tempo with a total of 142,545 articles [16]. By expanding the features, sentiment analysis can become more accurate and comprehensive because the features used include richer and more relevant information. This can help in understanding the sentiment contained in the text or data better.

G. RNN (Recurrent Neural Network)

In this research, the use of RNN in neural networks aims to process and handle time-ordered data. Speech recognition, image captioning, text categorization, and other categories have all been successfully applied using RNNs [17]. Figure 2 show the architecture of RNN model.

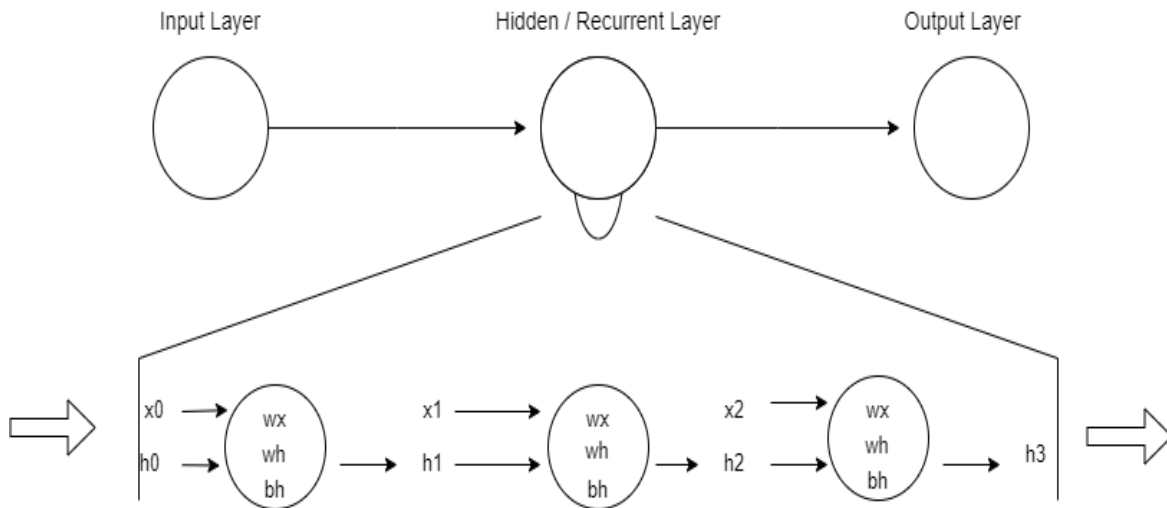


Fig. 2. The architecture of the RNN model

H. GA (Genetic Algorithm)

According to Gunal [23], the Genetic Algorithm is one of the feature selection methods in optimization that takes inspiration from the process of natural evolution. The optimization process involves several key components, each crucial for the algorithm’s success. To begin with, the design of the objective function is pivotal. The objective function serves as the metric for evaluating the performance of potential solutions represented by chromosomes. It quantifies how well a particular subset of features contributes to solving the computational problem at hand.

Chromosomes, in the context of GA, act as the encoded representations of potential solutions. Each gene within a chromosome signifies the presence or absence of a specific feature. Careful consideration must be given to designing chromosomes that effectively capture the relevant features of the problem domain, ensuring the genetic information is appropriately represented for the algorithm to make informed decisions. Recombination, an integral stage in GA, involves combining genetic information from two parent chromosomes to create new offspring. The specific recombination method employed influences the exploration-exploitation trade-off during the search for optimal solutions. Mutations introduce variation by randomly altering genes within chromosomes, preventing premature convergence and promoting diversity within the population.

Survivor selection determines which individuals persist to form the next generation based on their fitness values. The choice of survivor selection strategy impacts the algorithm’s convergence speed and its ability to escape local optima. Hyperparameters, such as mutation rates, crossover probabilities, and population size, play a crucial role in shaping the behavior of the genetic algorithm. Fine-tuning these hyperparameters is essential to achieving optimal performance and ensuring convergence towards the desired subset of features. Iteration continues until the population converges towards the optimal solution, allowing GA to be an efficient method for finding a subset of features that contribute maximally to solving computational problems [24]. The stages of the genetic algorithm optimization process can be seen in the Figure 3.

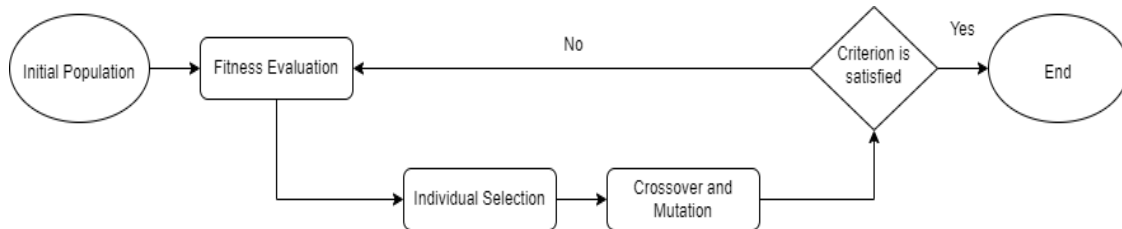


Fig. 3. Flowchart Genetic Algorithm

I. Performance Evaluation

Assessments should be conducted to determine how well the developed system is doing, using accuracy, precision, recall, and F-score values. A confusion matrix, which shows the categorization in terms of actual and expected values, is used to measure performance.

TABLE V
 Confusion Matrix

Predicted Values	Actual Values	
	Positive	Negative
Positive	TP	FP
Negative	FN	TN

The table of system performance contains four terms. A positive forecast whose true state is accurate is called a true positive (TP). When the true state is wrong, a false positive forecast is known as a false positive (FP). A negative prediction in which the true condition is untrue is called a false negative (FN). When the true state is untrue, a true negative prediction, or TN, is made [25]. The following formulas can be used to calculate the performance value:

a. *Accuracy* : The percentage of accurately anticipated data to the entire amount of data is known as accuracy.

$$accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (4)$$

b. *Precision* : The ratio of accurately anticipated positive values to all instances projected as positive is known as precision.

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

c. *Recall* : The ratio of data with accurate positive predictions to all available positive data is known as recall.

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

d. *F1-Score* : F1-Score is a performance metric that includes both recall and precision values.

$$F1 - Score = 2 \times \frac{precision \times recall}{precision+recall} = \frac{2TP}{2TP+FP+FN} \quad (7)$$

IV. RESULTS AND DISCUSSION

This research is divided into four scenarios that use classification models with RNN architecture, dividing the research into four scenarios to comprehensively assess performance. In scenario I, testing to obtain baseline data to gain an initial understanding of the performance of the RNN classification model in sentiment analysis. In scenario II, the research continued by incorporating TF-IDF feature extraction into the testing process. The purpose of this scenario is to see if the addition of TF-IDF feature extraction can improve the performance of the RNN classification model in sentiment analysis. Based on the results obtained in scenarios I and II, scenario III introduces more advanced testing by combining TF-IDF feature extraction with FastText feature expansion. The purpose of this scenario is to evaluate whether the combination of the two techniques can provide a better performance improvement in sentiment analysis. In the last scenario, scenario IV, this research takes a unique step by incorporating a Genetic Algorithm into the model optimization process. The purpose of this scenario is to identify the best individual with the highest fitness value, with the hope of significantly improving the performance of the RNN classification model in sentiment analysis. By dividing the research into four scenarios, the main contribution of the authors is to comprehensively evaluate and compare various techniques and features used in sentiment analysis. This can provide a deeper insight into the factors that contribute to the performance of RNN classification models and help in the development of better methods for sentiment analysis on text.

A. Evaluation Results

In the first scenario, the Recurrent Neural Network (RNN) classification was used to find the base model. This process was repeated five times to ensure consistent results. The total amount of data used was 37,391 comments about the president. The ratio of sizes for training and test data was adjusted to 90:10, 80:20, and 70:30 in the experiment. to understand its effect on model performance. the highest accuracy is reaching 79.32%, and an F1 score of 78.67% when utilizing a test size ratio of 90:10. More detail about the results of the first scenario are shown in Table VI.

TABLE VI
Results of Scenario I

Test Size	Tweet	
	Accuracy (%)	F1-Score (%)
0.1	79.32	78.67
0.2	78.32	77.53
0.3	78.96	78.47

In scenario II, the model from scenario I is integrated into the TF-IDF feature extraction process using the split data. This study involves the comparison of accuracy results for the maximum number of features in TF-IDF, namely 1000, 2000, 5000, 7000, and 8000. The results of this scenario are presented in Table VII.

TABLE VII
Results of Scenario II

Max Feature	Tweet	
	Accuracy (%)	F1-Score (%)
1,000	80.75 (+1.43)	80.13 (+1.46)
2,000	82.14 (+2.82)	81.58 (+2.91)
5,000	82.31 (+2.99)	81.71 (+3.04)
7,000	82.35 (+3.03)	81.82 (+3.15)
8,000	82.30 (+2.98)	81.73 (+3.06)

The second scenario involves an RNN model with a TF-IDF feature extraction process with the resulting TF-IDF matrix containing numerical weights representing how important each word is. The results show that by increasing the maximum number of features in TF-IDF, the performance of the model also improves. The highest accuracy is achieved at 7000 features with a value of 82.35%, indicating that increasing the complexity of the features can make a positive contribution to the performance of the model. According to the authors, when using a maximum of 8000 features, can lead to overfitting, where the model can over understand the details of the training data and the performance on test data decreases. The selection of 7000 features may keep the complexity level from being too high.

For scenario III, the TF-IDF representation results enter the feature expansion process which is implemented using FastText with the corpus of Tweet, IndoNews, and Tweet + IndoNews. The similarity of each corpus is determined based on the top 1, 5, 10, and 15 similarity values. The results of the third scenario are shown in Table VIII below.

TABLE VIII
Results of Scenario III

Top (n)	Tweet		IndoNews		Tweet + Indonews	
	Accuracy (%)	F1-Score (%)	Accuracy (%)	F1-Score (%)	Accuracy (%)	F1-Score (%)
1	82.19 (+2.87)	81.53 (+2.86)	82.16 (+2.84)	81.63 (+2.96)	82.30 (+2.98)	81.70 (+3.03)
5	82.38 (+3.06)	81.77 (+3.10)	82.36 (+3.04)	81.75 (+3.08)	82.43 (+3.11)	81.89 (+3.22)
10	82.43 (+3.11)	81.82 (+3.15)	82.34 (+3.02)	81.74 (+3.07)	82.25 (+2.93)	81.69 (+3.02)
15	82.28 (+2.96)	81.70 (+3.03)	82.11 (+2.79)	81.49 (+2.82)	82.28 (+2.96)	81.72 (+3.05)

In scenario III, experiments were conducted to evaluate feature expansion using FastText with various corpora, namely Tweet, IndoNews, and a combination of both. The results show that accuracy and F1-Score improvements occur by adding information from IndoNews. The highest achievement was achieved in the Top 5 approach with the combined Tweet+IndoNews corpus.

Combining multiple feature expansions such as TF-IDF and FastText in improving the performance of sentiment analysis models has significant benefits. The use of TF-IDF helps extract influential key words, while FastText provides a vector representation. This combination allows the model to understand the meaning of words in context and improves the ability to capture text complexity. While exploring an unexplored combination, it has the potential to provide innovation in sentiment analysis approaches.

In this scenario IV, the RNN model that has been trained will be optimized using the Genetic Algorithm. In Genetic Algorithm optimization, there are steps such as Fitness Evaluation, Individual Selection, Crossover and Mutation, and fitness evaluation. The table below provides details about the parameters used in the Genetic Algorithm. The results of scenario IV testing and the parameters used can be seen in the tables below.

TABLE IX
Genetic GA Parameter

Genetic Parameter	Value
Mate	Indpb = 0.5
Mutate	Mu = 0, sigma = 1, indpb = 0.2
Select	Tournsize = 3

TABLE X
Evolutionary GA Parameter

Genetic Parameter	Value
cxpb	0.5
mutpb	0.2
ngen	10

TABLE XI
Result of GA

Best Individual	Best Accuracy (%)
Units = 255, dropout = 0.122	82.72 (+3.40)

Scenario IV involves optimizing the RNN model using a genetic algorithm. In the genetic algorithm, the fitness function is responsible for evaluating the quality of an individual by training an RNN model based on the individual's parameter configuration and measuring the model's accuracy on validation data. The fitness function provides a numerical value that represents the quality of the solution, where higher values indicate a better solution. The individual is a representation of a potential solution and consists of the parameters number of units in the RNN layer and dropout rate. The results show that through genetic algorithm optimization, the accuracy of the model can be improved. The best individual with unit parameters of 255 and drop out of 0.1222 gives an accuracy (best fitness) of 82.72%, an increase of 3.40% from the baseline. Accuracy only increases by 3.40% because, this happens due to several factors, namely the selection of parameter values used in GA, the limited number of generations used in GA, and the population size being too small which can limit genetic variation and reduce the ability of GA to find better solutions.

B. Discussion

A classification model based on Recurrent Neural Network (RNN) incorporates TF-IDF feature extraction, FastText feature expansion, and Genetic Algorithm (GA) optimization have been shown to significantly improve accuracy. This research using four test scenarios which is four tests were employed to ascertain the accuracy and F1 values for each scenario. Table XII shows the comparison results of the four test scenarios.

TABLE XII
 Comparison of Test Scenario Results

Scenario	Accuracy (%)
Baseline	79.32
Baseline + TF-IDF	82.35 (+3.03)
Baseline + TF-IDF + FastText	82.43 (+3.11)
Baseline + TF-IDF + FastText + GA	82.72 (+3.40)

Scenario III achieved an accuracy rate of 82.43% and an F1-Score of 81.89% while using the Tweet+IndoNews corpus. This represents a 3.11% improvement in accuracy and a 3.22% improvement in F1-Score compared to the baseline. The accuracy experienced a 3.40% improvement after the integration of genetic algorithm optimization in Scenario IV.

V. CONCLUSION

In this research, sentiment analysis was conducted utilizing a Recurrent Neural Network (RNN) with the application of TF-IDF feature extraction, FastText feature expansion, and Genetic Algorithm optimization. The data used are Twitter user comments on the 2024 Indonesian presidential election. The amount of data used was 37.391 consisting of positive and negative comments taken from Twitter social media. In addition, the author uses 142.545 IndoNews data in the feature expansion process to achieve optimal results. Model testing was conducted based on four scenarios combining the Recurrent Neural Network (RNN) classification model, TF-IDF feature extraction, Fasttext feature expansion, and Genetic Algorithm optimization. The TF-IDF feature extraction method can significantly improve the accuracy by using max 7000 features with an accuracy of 82.35%. The application of FastText feature expansion also improved the accuracy results quite well on the top 5 with an accuracy of 82.43%. The best results were obtained after applying Genetic Algorithm optimization with an accuracy value of 82.72% increase of 3.40% from the baseline. Suggestions for further research include increasing the amount of data tested and increasing the methods applied to achieve maximum results.

REFERENCES

- [1] Yanti, P. G., Zabadi, F., & Rahman, F. (2020). The Effect Of Using Social Media Towards Students' Reading Comprehension. *RETORIKA: Jurnal Bahasa, Sastra, dan Pengajarannya*, 13.
- [2] Saad, S. E., & Yang, J. (2019). Twitter sentiment analysis based on ordinal regression. *IEEE Access*, 7, 163677-163685.
- [3] Sanusi, R., Astuti, F. D., & Buryadi, I. Y. (2021). Analisis sentimen pada twitter terhadap program kartu pra kerja dengan recurrent neural network. *JIKO (Jurnal Informatika dan Komputer)*, 5(2), 89-99.
- [4] A. Mittal and S. Patidar, "Sentiment analysis on twitter data: A survei," *ACM Int. Conf. Proceeding Ser.*, pp. 91-95, 2019, doi: 10.1145/3348445.3348466.
- [5] Nistor, S. C., Moca, M., Moldovan, D., Oprean, D. B., & Nistor, R. L. (2021). Building a Twitter sentiment analysis system with recurrent neural networks. *Sensors*, 21(7), 2266.
- [6] Patel, A., & Tiwari, A. K. (2019, February). Sentiment analysis by using recurrent neural network. In *Proceedings of 2nd International Conference on Advanced Computing and Software Engineering (ICACSE)*.
- [7] Cahyadi, R., Damayanti, A., & Aryadani, D. (2020). Recurrent neural network (rnn) dengan long short term memory (lstm) untuk analisis sentimen data instagram. *JIKO (Jurnal Informatika dan Komputer)*, 5(1), 1-9.
- [8] Alibrahim, H., & Ludwig, S. A. (2021, June). Hyperparameter optimization: Comparing genetic algorithm against grid search and bayesian optimization. In *2021 IEEE Congress on Evolutionary Computation (CEC)* (pp. 1551-1559). IEEE.

- [9] Chen, Y., Dong, Z., Wang, Y., Su, J., Han, Z., Zhou, D., ... & Bao, Y. (2021). Short-term wind speed predicting framework based on EEMD-GA-LSTM method under large scaled wind history. *Energy Conversion and Management*, 227, 113559.
- [10] Kara, A. (2021). Multi-step influenza outbreak forecasting using deep LSTM network and genetic algorithm. *Expert Systems with Applications*, 180, 115153.
- [11] Shahid, F., Zameer, A., & Muneeb, M. (2021). A novel genetic LSTM model for wind power forecast. *Energy*, 223, 120069.
- [12] Xu, L., Hou, L., Zhu, Z., Li, Y., Liu, J., Lei, T., & Wu, X. (2021). Mid-term prediction of electrical energy consumption for crude oil pipelines using a hybrid algorithm of support vector machine and genetic algorithm. *Energy*, 222, 119955.
- [13] Nugroho, N. A., & Setiawan, E. B. (2021). Implementation Word2Vec for Feature Expansion in Twitter Sentiment Analysis. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 5(5), 837-842.
- [14] Nevěřilová, Z. (2022). Compressed FastText Models for Czech Tagger. *RASLAN 2022 Recent Advances in Slavonic Natural Language Processing*, 79.
- [15] Fan, Y., Xie, X., Cai, Y., Chen, J., Ma, X., Li, X., ... & Guo, J. (2022). Pre-training methods in information retrieval. *Foundations and Trends® in Information Retrieval*, 16(3), 178-317.
- [16] David Sayce, "The Number of tweets per day in 2020 | David Sayce." <https://www.dsayce.com/social-media/tweets-day/> (accessed Dec 31, 2020).
- [17] Lakmal, D., Ranathunga, S., Peramuna, S., & Herath, I. (2020, May). Word embedding evaluation for sinhala. In *Proceedings of the Twelfth Language Resources and Evaluation Conference* (pp. 1874-1881).
- [18] Novitasari, F., & Purbolaksono, M. D. (2021). Analysis Sentiment Aspect Level on Beauty Product Reviews Using Chi-Square and Naïve Bayes. *Journal of Data Science and Its Applications*, 4(1), 18-30.
- [19] Xu, J., & Du, Q. (2020). TextTricker: Loss-based and gradient-based adversarial attacks on text classification models. *Engineering Applications of Artificial Intelligence*, 92, 103641.
- [20] Lopez, W., Merlino, J., & Rodriguez-Bocca, P. (2020). Learning semantic information from Internet Domain Names using word embeddings. *Engineering Applications of Artificial Intelligence*, 94, 103823. [21] M. Mohandes, A. Hussain, P. Khan, and H. Nuha, "Non-Vital Privacy Compromise for Improved Services in Pilgrimage Seasons," in *2019 IEEE RFID Conference*, 2019.
- [21] Chakraborty, K., Bhattacharyya, S., & Bag, R. (2020). A survey of sentiment analysis from social media data. *IEEE Transactions on Computational Social Systems*, 7(2), 450-464.
- [22] Goularas, D., & Kamis, S. (2019, August). Evaluation of deep learning techniques in sentiment analysis from twitter data. In *2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML)* (pp. 12-17). IEEE. [23] M. Mohandes, A. Hussain, P. Khan, and H. Nuha, "Non-Vital Privacy Compromise for Improved Services in Pilgrimage Seasons," in *2019 IEEE RFID Conference*, 2019.
- [23] Deng, X., Li, Y., Weng, J., & Zhang, J. (2019). Feature selection for text classification: A review. *Multimedia Tools and Applications*, 78, 3797-3816.
- [24] Gen, M., & Lin, L. (2023). Genetic algorithms and their applications. In *Springer handbook of engineering statistics* (pp. 635-674). London: Springer London.
- [25] Heydarian, M., Doyle, T. E., & Samavi, R. (2022). MLCM: Multi-label confusion matrix. *IEEE Access*, 10, 19083-19095.