Twitter Social Media-Based Sentiment Analysis Using Bi-LSTM Method With Genetic Algorithms Optimization

Girindra Syukran Prahasto [1], Erwin Budi Setiawan [2*]

*erwinbudisetiawan@telkomuniversity.ac.id
[1,2]School of Computing, Telkom University
Bandung

Abstract-Advances in information technology, particularly social media platforms such as Twitter, can be used to explore public sentiment around the much-discussed 2024 Indonesian Presidential Election. Using sentiment analysis as part of text mining, we focus on distinguishing positive and negative polarity using Natural Language Processing (NLP) techniques with to detect the accuracy of tweet polarity regarding the 2024 Indonesian Presidential Election. Specifically, we implement the Bidirectional Long Short-term Memory (Bi-LSTM) method, an enhanced version of LSTM, for sentiment analysis. The text is preprocessed, TF-IDF is used for word importance weighting, and Word2Vec is used for efficient learning of high-quality words. To optimize the accuracy of the model, we used Genetic Algorithm (GA), a heuristic approach rooted in the principles of genetics and natural selection. GA operates on a chromosome-based population, aligned with Darwinian evolutionary concepts. This research aims to compare the accuracy of the Bi-LSTM model with various feature extraction methods, including TF-IDF and Word2Vec, in measuring the polarity of election-related tweets. This research highlights the comparison and improvement of the accuracy of each scenario in the built model. The accuracy score results in this research was 83%, where the accuracy score increases from the baseline by 7.98%.

Keywords: Sentiment analysis; Bi-LSTM; TF-IDF, Word2Vec, Genetic Algorithm

Article info: *submitted January 12, 2024, revised April 7, 2025, accepted June 15, 2025*

1. Introduction

Information technology is an advancement in the field of information that facilitates the implementation of daily tasks, both in obtaining and disseminating information and disseminating information [1]. Social media is a form of advancement in information technology, and Twitter is one of the most popular platforms among the public. Nowadays, there are many social media, an example of which is Twitter, which is a convenient text-based social networking and micro-blogging application. Twitter users can access and express opinions on various topics that are being discussed around the world [2]. Many people use Twitter to voice their opinions on the 2024 Indonesian Presidential Election.

Sentiment analysis is a subset of text mining that explores the opinions, sentiments, evaluations, behaviors, and emotions of individuals towards an entity, such as a product, service, organization, individual, problem, topic, event, and its attributes [3]. The process of sentiment analysis involves determining the polarity of text in a document or set of words, by identifying whether it is positive, negative, or neutral [4]. In this research, the polarity that will be used is positive and negative.

Specifically, in this research, the method in Natural Language Processing (NLP) used is Bi-LSTM as a baseline to perform sentiment analysis. Bidirectional Long Short-term Memory (Bi-LSTM) is a developed form of LSTM method that is able to produce output from both directions, both from the left and right, at each stage of the process [5].

The extraction used is TF-IDF to give weight to the words. TF-IDF is a weighting technique that integrates Term Frequency and Document Frequency. The score is given based on the importance of the word from the frequency of its occurrence in the text set [6]. Feature extraction TF-IDF weights signify greater representativeness, leading to their retention, while lower-weighted words are considered less representative and are omitted. A judicious choice of word features not only accelerates information retrieval but also maintains model performance. Feature extraction involves assigning scores to fundamental words, and subsequently transforming these results into vectors that serve as representations for each word [7]. Word2Vec is used to insert words from the TF-IDF extraction results. The goal of Word2Vec is to insert high quality words that can be learned efficiently, especially in a larger corpus of text used for training [8].

Research from Hai Zhou [6], the base Bi-LSTM model demonstrated an impressive accuracy of 99.8% in classifying THUCNews article text, comprising 20,000 data points across 10 categories. However, upon employing TF-IDF feature extraction, the accuracy experienced a marginal decrease of 0.8% to reach

99%. Similarly, when applied to Taobao review text with an equivalent dataset, the base Bi-LSTM model achieved an accuracy score of 91.6%, but with TF-IDF feature extraction, this accuracy declined by 5.2% to 86.7%. The research effectively underscored the performance contrast between the base Bi-LSTM and Bi-LSTM + TF-IDF models.

Experimental outcomes revealed significant reductions in the number of parameters and training time when TF-IDF was utilized for data processing. In lengthy texts, parameters decreased by approximately 80%, and computation time was reduced to only 10% of the original, albeit with a minor 1% reduction in model effectiveness (accuracy, F1). Conversely, in short texts, the parameter decrease was around 25%, accompanied by a substantial 90% reduction in training time, leading to a 2-5% decrease in model effectiveness. These findings highlight the distinct focus of TF-IDF on capturing local information, while LSTM excels in capturing global information. The intriguing observation is that TF-IDF enhances text feature extraction and optimizes the constructed model.

In this research, Genetic Algorithm (GA) will be used to find the best accuracy score of the built model. GA is a method with a heuristic approach that evolved from the principles of genetics and the process of natural selection from Darwin's theory of evolution [9]. GA is the optimal choice because it applies Darwinian evolutionary principles, works on a population represented as chromosomes, and is evaluated to achieve the most optimal value [10]. GA optimization can increase accuracy scores based on research from Alsaleh [11] who used the CNN-GA model where GA as optimization to find the best fitness value on the SNAD dataset. The baseline accuracy score was 84.32%, but by applying GA optimization, the accuracy score increased by 4.39% to 88.71%.

The main contribution of this research is to find the accuracy score of the Bi-LSTM model to detect the accuracy of tweet polarity regarding the 2024 Indonesian Presidential Election. To date, to the best of the author's knowledge, there has been no research that analyzes sentiment with a model like the one in this study. This research will implement four main scenarios, with the first scenario being the baseline. There is a comparison between the Bi-LSTM model as the baseline, Bi-LSTM with TF-IDF feature extraction, and Word2Vec feature expansion. In the end, the built model will be applied GA optimization for Bi-LSTM to find the best accuracy.

2. Related Work

Arbane et al conducted a classification of text sentiment uploaded by Twitter and Reddit users against the COVID-19 pandemic [5]. The methods used for this research are bidirectional LSTM and embedding matrix to detect polarity sentiment from Twitter as much as and Reddit against the COVID-19 pandemic with four scenarios. The result of his research is that the Bi-LSTM model in the fourth scenario gets the best accuracy of 97.52% with three polarity labels. The dataset used in the fourth scenario is Reddit comments totaling 451,554 comments.

There is also research from Kartika Puerwandar et al [12] collecting tweets using Support Vector Machine (SVM) algorithm

trained to detect tweets related to real-time weather. TF-IDF feature extraction is applied to the model to convert a collection of text into a numeric vector. The data converted into numeric vectors is then applied to parameterize the train and test data. By using a linear kernel, the SVM algorithm obtained an accuracy of 85%.

Farrikh Alzami et al [13] examined sentiment analysis using several feature extractions such as Word Bags, TF-IDF, Word2Vector combined with several machine learning models such as Random Forest, SVM, KNN and Naïve Bayes. The dataset used is the case of Amazon food reviews totaling 568,454 reviews. The best accuracy score of TF-IDF is found in the SVM model with an accuracy score of 87.35%, Word2Vec is found in the Random Forest model with an accuracy score of 77.8%, and TF-IDF & Word2Vec is found in the Random Forest model with an accurcy score of 74.6%.

Rizki Tri Setiawan used the LSTM method and Genetic Algorithm Optimization to analyze the sentiment of Bank BCA stock price [14]. In this research, GA is an important optimization in getting the accuracy value of the model that has been built. GA emulates the specification process observed in natural evolution, where optimization involves selection, crossover, and mutation to iteratively evolve a population of potential solutions and search for optimal or near-optimal solutions to a problem. The accuracy score of the LSTM model is 43%, GloVe-LSTM scores 71%, and GloVe-LSTM-GA scores 87%.

3. Methods

An overview of this research will be presented in the flowchart Figure 1. The earliest stage starts with crawling data from Twitter with keywords related to the 2024 Indonesian Presidential Election. The results of crawling the data are given a polarity or label in the form of positive or negative. After being given polarity or labeling is the preprocessing stage so that the data is ready to proceed to the next stage. The data splitting stage is applied to produce train data and test data. The data is then applied to the Bi-LSTM model, TF-IDF feature extraction, Word2Vec feature expansion, GA optimization, and finally performance evaluation.

1. Crawling Data

The dataset used in this research is tweets on social media Twitter related to the 2024 Indonesian Presidential Election. Some keywords related to the topic are "Capres", "Calon presiden", "Anies Baswedan", "Prabowo Subianto", and "Ganjar Pranowo", as shown in Table 1. We obtained the data using Twitter Application Program Interface (API) and tweet-harvest library with Python programming language. The result of the data that has been crawled reaches 37,391 tweets which are stored in a file with Comma Separated Value (CSV) format.

Table 1. Keyword and quantity of dataset

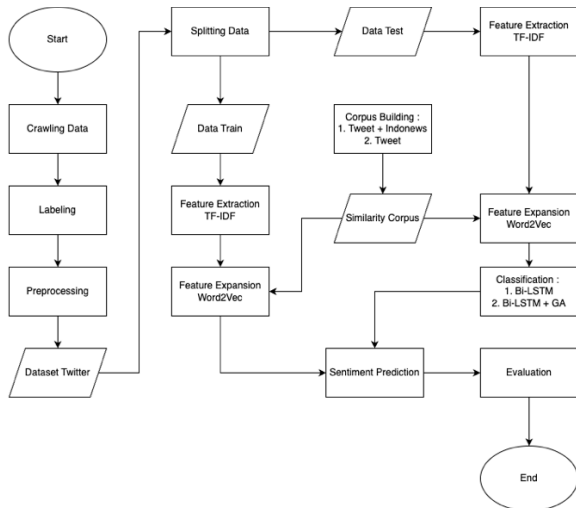| Keyword	Quantity
Capres	7,296
Calon Presiden	6,972
Anies Baswedan	10,434
Prabowo Subianto	4,662
Ganjar Pranowo	8,027

Figure 1. Model of Bi-LSTM for sentiment analysis flowchart

2. Labeling Data

In sentiment analysis, data labeling is required to assign a polarity to the text in the form of a positive or negative label. Discrete-valued polarity prediction is referred to as a classification problem. It aims to determine the polarity of a text [15].

This research involves three correspondents to determine the polarity of the text in the form of labels "1" for positive and "-1" for negative. For the result of text polarity, the dominant label from the three correspondents will be selected.

Table 2. Quantity of each polarity

Polarity	Quantity
Positive	21,866
Negative	15,525

3. Preprocessing

Author names need to be written in full without title and academic degree. If there are more than one author, one of them will be the corresponding author, whose name is written followed by an asterisk. If there are more than one author and they are from different affiliations, names are followed by a number that resembles each affiliation. One number is for one affiliation. The number is written in superscript, so is the asterisk for the corresponding author.

Information for affiliation includes name of department (or division), name of institution, and city. Name of country may be added if necessary. An email address is written for the corresponding author.

Data that has been collected previously needs to be preprocessed before being applied to the model. The purpose of preprocessing itself is to improve the quality of documents that will be applied to the model built [13]. There are several stages that are divided into this technique, starting from data cleaning, case folding, tokenizing, data normalization, stop word removal, and stemming. The following is an explanation of the stages in Preprocessing:

A. Data Cleaning

Each text in the dataset needs to be cleaned of some unnecessary elements, such as punctuations, emojis, address links, usernames, links, and hashtags.

B. Case Folding

Each word in the text needs to be case folded, where capital letters are converted to lower case.

C. Tokenizing

The separation of each word into a single word is applied to tokenizing.

D. Data Normalization

Normalization strives to standardize words with similar meanings that may vary. This procedure is implemented to address the heightened ambiguity associated with non-standard words.

E. Stop Word Removal

In this stage, a Python library called PySastrawi is used to assist in removing stop words. This process involves reducing the dimensionality of the data by eliminating words that do not contribute to the sentiment, such as personal pronouns, conjunctions, and prepositions.

F. Stemming

This stage converts words with affixes into basic words only. The affixes to be removed are prefixes or initial affixes, including "me-", "ber-", "pe-", "per-", "te-", "ter-", and "ke-". In addition to prefixes, the affixes that are removed are suffixes or final affixes, including "-I", "-an", and "-nya". PySastrawi Python library was again used to apply stemming to the textWriting citations and list of reference

4. Bidirectional Long Short-Term Memory (Bi-LSTM)

Bidirectional Long Short-Term Memory (Bi-LSTM) created by Schuster and Paliwal [16] is a method in NLP for learning document matrices and modeling sentiment analysis, in this context in Indonesian. LSTM itself is an extension of the standard Recurrent Neural Network (RNN) that is more effective and scalable. The main goal of LSTM is to avoid the missing gradient problem that RNN face, where gradients propagated back through the network decay or grow exponentially over time [17], [18] . RNN are not so effective in understanding a context behind inputs with long-term dependencies. LSTM with its complex dynamics consisting of several so-called memory blocks, can overcome this problem by effectively outputting and making predictions based on time series data. However, there is one drawback of the LSTM itself, which is that it cannot consider the entire context in the text because it only moves in one direction, i.e. only forward. This leads to low classification performance as the next words will be dismissed. Therefore, Bi-LSTM emerged to overcome the low classification performance by moving forward and backward and connecting the two hidden layers of LSTM to the output layer. [19], [20]. The architecture of the BI-LSTM model can be seen in the Figure 2.

5. Feature Expansion TF-IDF

Data that has previously been preprocessed will then be converted into a vector representation called TF-IDF. Term

Frequency - Inverse Document Frequency (TF-IDF) is a word frequency scoring system that highlights words from documents with greater interest, i.e. words that are less common across documents. [12]. TF-IDF combines both Term Frequency and Document Frequency methods. Term Frequency is a weighting concept to find the frequency of term appearance in a document. A document will have different lengths, so there is a possibility that a word appears in a long document compared to a short document. Document Frequency is the number of documents or texts in which a term appears. When the frequency is small, the weight value will be small.

In specialized documents, Term Frequency (TF) is used to measure how significant a word is by looking at how often it appears in the document. TF can be considered as the result of Bag of Words (BoW). In the TF-IDF method, the second element is the Inverse Document Frequency (IDF). In IDF, a word is considered important in a document if it does not appear frequently in other documents.
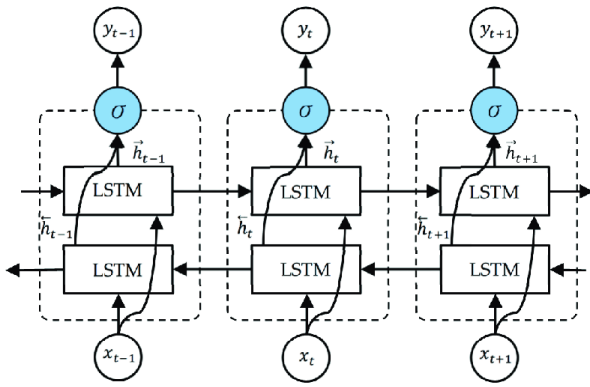

Figure 2. Architecture of Bi-LSTM model

6. Feature Expansion TF-IDF

Data that has previously been preprocessed will then be converted into a vector representation called TF-IDF. Term Frequency - Inverse Document Frequency (TF-IDF) is a word frequency scoring system that highlights words from documents with greater interest, i.e. words that are less common across documents. [12]. TF-IDF combines both Term Frequency and Document Frequency methods. Term Frequency is a weighting concept to find the frequency of term appearance in a document. A document will have different lengths, so there is a possibility that a word appears in a long document compared to a short document. Document Frequency is the number of documents or texts in which a term appears. When the frequency is small, the weight value will be small.

In specialized documents, Term Frequency (TF) is used to measure how significant a word is by looking at how often it appears in the document. TF can be considered as the result of Bag of Words (BoW). In the TF-IDF method, the second element is the Inverse Document Frequency (IDF). In IDF, a word is considered important in a document if it does not appear frequently in other documents.

7. Feature Expansion Word2Vec

Word embedding is a distributed representation that includes properties of words in a vector of real numbers that capture syntactic features and semantic relations of words. Word embedding is an attribute or feature learning technique in Natural Language Processing (NLP) where words that have similar, similar contexts, or relatively similar semantics, are mathematically grouped into a vector space. One of the models that can generate word embedding with long dimensions is Word2Vec. Word2Vec is one of the continuous learning tools that generate word embedding. Depending on the number of words predicted, there are two types of Word2Vec models, namely the Continuous Bag-of-Words (CBOW) model and the Skip-Gram model [21].

Word2Vector essentially positions words in the feature space, where their placement is dictated by their semantic meaning. In other words, words with similar meanings are clustered together, and the distance between two words reflects their semantic similarity [22]. The computational approach employs cosine similarity, which can be expressed as follows:

$$similarity = \cos\theta = \frac{words_1.words_2}{||words_1||.||words_2||} \qquad (1)$$

8. Genetic Algorithm Optimization

The Genetic Algorithm (GA) stands out as one of the pioneering population-based stochastic algorithms introduced in history. Alongside other Evolutionary Algorithms (EAs), GA operates through key mechanisms, namely selection, crossover, and mutation. This algorithm has been briefly outlined in this chapter, and its application to various case studies has been explored to assess its performance. Despite its advantages, genetic algorithms face a drawback in that selecting incorrect parameters can diminish the accuracy of the generated outcomes [23].

Natural selection aims to preserve the best individuals by replicating them, ensuring their survival in subsequent iterations. The crossover operator facilitates the creation of new individuals by combining genetic material from two parents. The most common method for selecting parents is the roulette wheel technique. The mutation operator, on the other hand, replaces the worst individuals with new ones, and the extent of replacement is determined by the mutation rate parameter [24].

9. Performance Measurement

Evaluating or measuring performance serves the purpose of determining the accuracy of a created algorithmic model. In performance measurement, several criteria are evaluated, including accuracy, precision, F1 Score, and recall. It employs confusion matrix for performance measurement. A confusion matrix is a tool used to analyze how well a classification model identifies tuples of different data. Within the confusion matrix, various terms exist, such as True Positive, indicating tuples labeled correctly as positive by the classifier; True Negative, denoting tuples labeled correctly as negative by the classifier; False Positive, signifying tuples labeled incorrectly as negative by the classifier, and False Negative, indicating tuples labeled incorrectly as positive by the classifier [25].
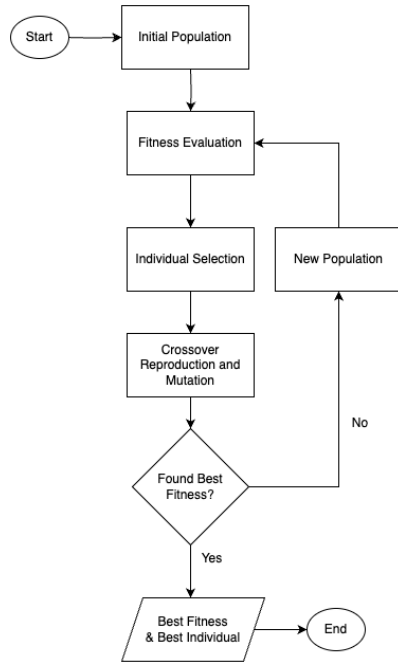
Figure 3 Architecture of Genetic Algorithm

Figure 4. Confussion matrix

Based on the confusion matrix previously described, there are several equations that can be calculated to get the performance value, these equations are as follows:

A. Accuracy

Accuracy is a measurement of what proportion of the total prediction is true of the entire data, the formula is as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \qquad (2)$$

B. Precision

Precision is a measurement of how precise the results of a model are, the formula is as follows:

$$Precision = \frac{TP}{TP + FP} \qquad (3)$$

C. Recall

Recall is a measurement of how complete a model is, the formula is as follows:

$$Recall = \frac{TP}{TP + FN} \qquad (4)$$

D. F1 Score

F-measure is the harmonic mean of precision and recall, the formula is as follows:

$$F1\ Score = \frac{2 * Precision * Recall}{Precision + Recall} \qquad (5)$$

4. Result Of Experiment

There were several experiments conducted for this research. Specifically, there are several scenarios conducted to make comparisons of accuracy scores and F1-Score, then find the best accuracy. Each scenario went through five runs, then the average of the accuracy score and F1-Score was calculated.

1. First Scenario

This scenario aims to find the best baseline based on the accuracy achieved. Different data splitting is applied to the Bi-LSTM model with several hyperparameters which can be seen in Table 3. The ratio of splitting data to be applied starts from 70:30, 80:20, and 90:10.

Table 3. Parameters in Bi-LSTM model

Parameter	Value
Embedding size	128
Bi-LSTM first hidden layer	128
Bi-LSTM second hidden layer	64
Dropout	0.2
Recurrent dropout	0.2
Batch size	128
Activation	sigmoid
Loss function	Binary cross-entropy
Optimizer	rmsprop
Number of epochs	10
Patience (Early Stopping)	3
Minimum delta	0.0001

The results of the accuracy score and F1 score are in the following table:

Table 4. Result of first scenario

Ratio	Accuracy(%)	F1-Score(%)
70:30	72.24	72.09
80:20	74.22	73.73
90:10	**75.02**	**74.64**

The 90:10 ratio is the best data splitting ratio where the accuracy score reaches 75.02% and F1-Score 74.64%. This ratio is the best result from the first scenario and will be used in the next scenarios.

2. Second Scenario (Bi-LSTM + TF-IDF)

In the second scenario, there is an additional feature, namely TF-IDF feature extraction. With TF-IDF, the embedding size parameter in Table 3 will not be used. The parameter whose value will be compared and searched for the best accuracy and F1-Score is the max-features parameter. Max-features that will be compared are 3000, 4500, and 6000. The results of some of these experiments are shown in the following table:

Table 5. Result of second scenario

Max-features	Accuracy(%)	F1-Score(%)
3000	81.46 (+ 6.44)	80.94 (+ 6.3)
4500	**81.72 (+ 6.7)**	**81.12 (+ 6.48)**
6000	81.57 (+ 6.55)	81 (+ 6.36)

Max-features 4500 is the best value in Bi-LSTM + TF-IDF with an accuracy value of 81.72% and F1-Score 81.12% with an accuracy increase of 6.7% from the baseline. The value of 4500 in max-features will be used in the next scenario.

3. Third Scenario (Bi-LSTM + TF-IDF + Word2Vec)

In the third scenario, Word2Vec was added as a feature expansion. This scenario is divided into four times to find the best accuracy score and F1-Score. The parameters that will be explored in this scenario are top-n and corpus that has been previously trained. The corpus build for Word2Vec has been trained before, the first corpus is tweet only, and the second corpus is tweet + Indonews corpus. The results of several experiments are shown in the following table:

Table 6. Result of third scenario

Corpus	Top-n	Accuracy(%)	F1-Score(%)
Tweet only	1	82.04 (+ 7.01)	81.52 (+ 6.88)
	5	80.44 (+ 5.42)	79.88 (+ 5.24)
Tweet + Indonews	1	**82.41 (+ 7.39)**	**81.89 (+ 7.25)**
	5	81.94 (+ 6.92)	81.43 (+ 6.79)

Based on these results, the corpus that will be used for the next scenario is the tweet corpus + Indonews with top-1 similarity. The accuracy score reached 82.41% with an increase of 7.39% from the baseline.

4. Fourth Scenario (Bi-LSTM + TF-IDF + Word2Vec + GA)

In the fourth scenario, there is additional GA optimization to find the best fitness value, where the best fitness value means the best accuracy score. First, the best units value and dropout value are searched, then the best individual of activation is searched.

In the first experiment, there are three individual units listed, namely 64, 128, and 256. The second individual is a dropout with a range of 0.1 to 0.5 randomly. The individual is then applied to find the best activation which in this case the individuals are "sigmoid", "softmax", and "relu" activation. Each individual will have five genes by cycling the two individuals. Other parameters used in GA are found in the Table 7 and Table 8.

Table 7. Evolutionary GA parameters

Parameter	Value
Cxpb	0.5
Mutpb	0.2
Ngen	10

Table 8. Genetic GA parameters

Parameter	Value
Mate	Indpb = 0.5
Mutate	Indpb = 0.2
Select	Tournsize = 3

Table 9. Result of fourth scenario

Best Individual	Best Accuracy (%)
Units = 256 ropout = 0.158 Activation = sigmoid	**83 (+ 7.98)**

The result of this scenario is that the best individual is obtained by unit value 256, dropout value 0.158, and activation sigmoid. The highest accuracy score reaches 83% with an increase of 7.98% from the baseline.

5. Conclusion

This research to find the accuracy score of the Bi-LSTM model to detect the accuracy of tweet polarity regarding the 2024 Indonesian Presidential Election. The dataset used was obtained from Twitter with the keywords "Capres", "Calon Presiden", "Anies Baswedan", "Prabowo Subianto", and "Ganjar Pranowo" for a total of 37,391 tweets. The polarity of each tweet has a total of 21,866 positive tweets and 15,525 negative tweets. There are four scenarios that have been described in the previous section of Bi-LSTM classification with TF-IDF extraction features, Word2Vec expansion features, and Genetic Algorithm Optimization. The highest accuracy score result is in the fourth scenario with 83% accuracy, where the accuracy score increases from the baseline by 7.98%.

Reference

[1] W. Setyowati, R. Widayanti, and D. Supriyanti, "Implementation of E-Business Information System in Indonesia : Prospects and Challenges," *International Journal of Cyber and IT Service Management*, vol. 1, no. 2, pp. 180–188, Oct. 2021, doi: 10.34306/ijcitsm.v1i2.49.

[2] A. D. Dubey, "Twitter Sentiment Analysis during COVID19 Outbreak," *SSRN Electronic Journal*, 2020, doi: 10.2139/ssrn.3572023.

[3] M. Birjali, M. Kasri, and A. Beni-Hssane, "A comprehensive survey on sentiment analysis: Approaches, challenges and trends," *Knowl Based Syst*, vol. 226, p. 107134, Aug. 2021, doi: 10.1016/j.knosys.2021.107134.

[4] D. H. Abd, A. R. Abbas, and A. T. Sadiq, "Analyzing sentiment system to specify polarity by lexicon-based," *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 1, pp. 283–289, Feb. 2021, doi: 10.11591/eei.v10i1.2471.

[5] M. Arbane, R. Benlamri, Y. Brik, and A. D. Alahmar, "Social media-based COVID-19 sentiment classification model using Bi-LSTM," *Expert Syst Appl*, vol. 212, p. 118710, Feb. 2023, doi: 10.1016/j.eswa.2022.118710.

[6] H. Zhou, "Research of Text Classification Based on TF-IDF and CNN-LSTM," *J Phys Conf Ser*, vol. 2171, no. 1, p. 012021, Jan. 2022, doi: 10.1088/1742-6596/2171/1/012021.

[7] Y. Zhang, Y. Zhou, and J. Yao, "Feature Extraction with TF-IDF and Game-Theoretic Shadowed Sets," 2020, pp. 722–733. doi: 10.1007/978-3-030-50146-4_53.

[8] B. Jang, M. Kim, G. Harerimana, S. Kang, and J. W. Kim, "Bi-LSTM Model to Increase Accuracy in Text Classification: Combining Word2vec CNN and Attention Mechanism," *Applied Sciences*, vol. 10, no. 17, p. 5841, Aug. 2020, doi: 10.3390/app10175841.

[9] Gunawan D, Riana D, Ardiansyah D, Akbar F, and Alfarizi S, "Komparasi Algoritma Support Vector Machine Dan Naïve Bayes Dengan Algoritma Genetika Pada Analisis Sentimen Calon Gubernur Jabar 2018-2023," *Jurnal Teknik Komputer AMIK BSI*, vol. 6, no. 1, pp. 121–129, Jan. 2020.

[10] L. K. Karnatapu, S. P. Annavarapu, and U. V. Nanduri, "Multi-Objective Reservoir Operating Strategies by Genetic Algorithm and Nonlinear Programming (GA–NLP) Hybrid Approach," *Journal of The Institution of Engineers (India): Series A*, vol. 101, no. 1, pp. 105–115, Mar. 2020, doi: 10.1007/s40030-019-00419-2.

[11] D. Alsaleh and S. Larabi-Marie-Sainte, "Arabic Text Classification Using Convolutional Neural Network and Genetic Algorithms," *IEEE Access*, vol. 9, pp. 91670–91685, 2021, doi: 10.1109/ACCESS.2021.3091376.

[12] K. Purwandari, R. B. Perdana, J. W. C. Sigalingging, R. Rahutomo, and B. Pardamean, "Automatic Smart Crawling on Twitter for Weather Information in Indonesia," *Procedia Comput Sci*, vol. 227, pp. 795–804, 2023, doi: 10.1016/j.procs.2023.10.585.

[13] F. Alzami, E. D. Udayanti, D. P. Prabowo, and R. A. Megantara, "Document Preprocessing with TF-IDF to Improve the Polarity Classification Performance of Unstructured Sentiment Analysis," *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, pp. 235–242, Aug. 2020, doi: 10.22219/kinetik.v5i3.1066.

[14] R. T. Setiawan and E. B. Setiawan, "Sentiment Analysis of BBCA Stock Price on Twitter Data Using LSTM and Genetic Algorithm Optimization," *sinkron*, vol. 8, no. 4, pp. 2479–2489, Oct. 2023, doi: 10.33395/sinkron.v8i4.12825.

[15] W. Muhammad, M. Mushtaq, K. N. Junejo, and M. Y. Khan, "SENTIMENT ANALYSIS OF PRODUCT REVIEWS IN THE ABSENCE OF LABELLED DATA USING SUPERVISED LEARNING APPROACHES," *Malaysian Journal of Computer Science*, vol. 33, no. 2, pp. 118–132, Apr. 2020, doi: 10.22452/mjcs.vol33no2.3.

[16] M. J. Hamayel and A. Y. Owda, "A Novel Cryptocurrency Price Prediction Model Using GRU, LSTM and bi-LSTM Machine Learning Algorithms," *AI*, vol. 2, no. 4, pp. 477–496, Oct. 2021, doi: 10.3390/ai2040030.

[17] O. Alharbi, "A Deep Learning Approach Combining CNN and Bi-LSTM with SVM Classifier for Arabic Sentiment Analysis," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 6, 2021, doi: 10.14569/IJACSA.2021.0120618.

[18] A. Purwarianti and I. A. P. A. Crisdayanti, "Improving Bi-LSTM Performance for Indonesian Sentiment Analysis Using Paragraph Vector," in *2019 International Conference of Advanced Informatics: Concepts, Theory and Applications (ICAICTA)*, IEEE, Sep. 2019, pp. 1–5. doi: 10.1109/ICAICTA.2019.8904199.

[19] I. Lasri, A. Riadsolh, and M. ElBelkacemi, "Self-Attention-Based Bi-LSTM Model for Sentiment Analysis on Tweets about Distance Learning in Higher Education," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 18, no. 12, pp. 119–141, Jun. 2023, doi: 10.3991/ijet.v18i12.38071.

[20] Y.-H. Li, L. N. Harfiya, K. Purwandari, and Y.-D. Lin, "Real-Time Cuffless Continuous Blood Pressure Estimation Using Deep Learning Model," *Sensors*, vol. 20, no. 19, p. 5606, Sep. 2020, doi: 10.3390/s20195606.

[21] Alvi Rahmy Royyan and Erwin Budi Setiawan, "Feature Expansion Word2Vec for Sentiment Analysis of Public Policy in Twitter," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 1, pp. 78–84, Feb. 2022, doi: 10.29207/resti.v6i1.3525.

[22] I. A. Asqolani and E. B. Setiawan, "Hybrid Deep Learning Approach and Word2Vec Feature Expansion for Cyberbullying Detection on Indonesian Twitter," *Ingénierie des systèmes d information*, vol. 28, no. 4, pp. 887–895, Aug. 2023, doi: 10.18280/isi.280410.

[23] 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, p. 012017, Nov. 2020, doi: 10.1088/1742-6596/1641/1/012017.

[24] A. Rafdi, H. Mawengkang, and S. Efendi, "Sentiment Analysis Using Naive Bayes Algorithm with Feature Selection Particle Swarm Optimization (PSO) and Genetic Algorithm," *International Journal of Advances in Data and Information Systems*, vol. 2, no. 2, Nov. 2021, doi: 10.25008/ijadis.v2i2.1224.

[25] S. H. Rukmawan, F. R. Aszhari, Z. Rustam, and J. Pandelaki, "Cerebral Infarction Classification Using the K-Nearest Neighbor and Naive Bayes Classifier," *J Phys Conf Ser*, vol. 1752, no. 1, p. 012045, Feb. 2021, doi: 10.1088/1742-6596/1752/1/012045.