

The COVID-19 Tweets Classification Based on Recurrent Neural Network

Arif Laksito ^{a,*}, Nuruddin Wiranda ^b, Shofiyati Nur Karimah ^c, Mardhiya Hayaty ^a

^a Faculty of Computer Science, Universitas Amikom Yogyakarta, Yogyakarta, Indonesia

^b Department of Computer Education, Lambung Mangkurat University, Banjarmasin, Indonesia

^c Graduate School of Advanced Science, Japan Advanced Institute of Science and Technology (JAIST), Nomi, Ishikawa, Japan

Corresponding author: *arif.laksito@amikom.ac.id

Abstract— Due to its extensive use in both public and commercial contexts, sentiment analysis on Twitter has recently received much attention, particularly concerning tweets about COVID-19. Information about COVID-19 has been widely spread over social media, resulting in various views, opinions, and emotions about this pandemic, significantly impacting people's health. It is exceedingly challenging for the authorities to find these rumors on these public platforms manually. This paper proposes a framework for text classification using the RNN model and its updates, such as LSTM, BiLSTM, and GRU. This study aims to determine the best recurrent network model for handling cases of Twitter data classification. We utilized Twitter data relevant to COVID-19 and the lockdown with four classification classes (sad, joy, fear, and anger). In addition, this study aims to prove whether GloVe pre-trained word embedding can increase the accuracy of model predictions. The training and testing datasets were split into 80% and 20%, respectively. Therefore, in this experiment an early stopping technique was used with a limit of 15 epochs and a minimum delta of 0.01, meaning that training will be stopped if there is no improvement of 0.1% accuracy after 15 epochs. We used the f1-score average to measure the accuracy of the classification task results. The test results show that the BiLSTM model with GloVe word embedding yields the best f1-score compared to other models. Moreover, in all model testing, the f1-score value of the 'fear' class displays the highest accuracy compared to other classes.

Keywords—Text classification; recurrent neural network; COVID-19 tweets.

Manuscript received 6 Apr. 2023; revised 5 Nov. 2023; accepted 9 Jan. 2024. Date of publication 29 Feb. 2024.
IJASEIT is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.



I. INTRODUCTION

The global COVID-19 pandemic, which emerged in late 2019, has brought about unprecedented health challenges and given rise to a flood of public discourse and emotional expressions on social media platforms. Twitter, being a prominent medium for real-time information sharing and emotional venting, has become a valuable resource for understanding the diverse spectrum of human reactions to this public health crisis. Detecting rumors and sentiments on such a vast public platform manually is an arduous task, necessitating the deployment of robust automated tools.

In the era of deep learning and natural language processing, the ability to automatically categorize, understand, and extract valuable information from vast volumes of textual data has become an essential endeavor. Among the many techniques developed for this purpose, Recurrent Neural Networks (RNNs) have emerged as powerful tools for text classification.

These networks, inspired by the human brain's ability to process data sequences, are particularly well-suited for tasks involving sequential data, such as language.

In response, this study presents a comprehensive framework for text classification utilizing Recurrent Neural Networks (RNNs) and their advanced variations, including Long Short-Term Memory (LSTM), Bidirectional LSTM (BiLSTM), and Gated Recurrent Unit (GRU) models. Additionally, we also investigated the used of GloVe word embedding on the classification models. The objective is to determine the most effective RNN model for classifying Twitter data relevant to COVID-19. The results of this study are expected to offer the most recent comparison of deep learning models on sentiment analysis.

The remainder of this article is organized as follows. Part II explains earlier studies relevant to the techniques utilized, the application of each approach to RNNs, and analyses of the suggested experimental framework. Part III discusses the

experimental findings and Part IV concludes the study with recommendations for further investigation.

II. MATERIAL AND METHOD

A. Related Works

Several studies related to COVID-19 sentiment analysis have been conducted using RNN models [1]–[3]. Kaur [1] investigated sentiment analysis of 600 tweets about COVID-19 and its implications on mental health. The author assigned positive, negative, and neutral sentiment ratings using Recurrent Neural Network (RNN) and Support Vector Machine (SVM) sentiment categorization. Meanwhile, Chintalapudi [2] performed a sentiment analysis using information collected from Twitter. Phrases of "thank you," "well," and "good" foster a friendly environment among healthcare authorities. However, terms of "trump," "kill," "death," and "die" cause unwarranted anxiety in individuals. These conclusions compel local governments to install fact-checkers on social media to counter misleading propaganda. In the subsequent year, Singh [4] performed a sentiment analysis using publicly accessible Twitter data taken from the Kaggle database. Based on 179,108 tweets about COVID-19, there were 45% of positive, 30% neutral, and 25% negative tweets.

Alabdulkreem used Recurrent Neural Network (RNN) and Glove Word Embedding [5] to study depression predictions in Arab women during the COVID-19 epidemic phase using Tweet data. Based on 10,000 tweets from 200 Saudi Arabian users, it was determined that Arab women initially suffered from depression during the COVID-19 epidemic period. Research on the application of machine learning to categorize the feelings of student feedback was done by Edalati [6]. RNN + Glove is one of the machine learning techniques employed. This model yields the highest accuracy of all other models at 98.29%.

Bangyal [7] examined the use of deep learning to identify COVID-19 fake content. This study employed LSTM, BiLSTM, Gated Recurrent Unit (GRU), RNN, and Convolutional Neural Network (CNN) as deep learning techniques. Various metrics were used to measure performance, including precision, recall, accuracy, and F1-score. Results of the GRU show an accuracy rate of 90%, precision of 0.91, recall of 0.95, and F1-score of 0.93.

Abdelminaam [8] researched on how to improve a smart framework for automatically spotting COVID-19 misleading information on Twitter. The tests' results show accuracy, precision, recall, and F1-score of 81.49%, 82.26%, 81.49%, and 81.35%, respectively. Meanwhile, Raamkumar [9] studied the application of deep learning, including GRU, on COVID-19 social media content to examine how general public perceived physical distance during the pandemic. Furthermore, user-generated information from social media was also characterized using the Health Belief Model (HBM) in order to better understand public health behavior. The findings of this work show that, in the context of physical distance treatments, a deep learning-based text classifier was successful in creating an appropriate categorization of COVID-19 Facebook comments using the HBM construct.

Research on detecting and classifying COVID-19 false information on social media Twitter was conducted by

Pathway [10], which employed several deep learning models, including GRU with Glove word embeddings, to effectively detect COVID-19 false information. The findings demonstrate that the overall deep learning model—which consists of a CNN with an accuracy rate of 93.92%—outperforms the traditional classification strategy. Meanwhile, the accuracy of GRU using the GloVe word embedding is 92.14%. Research on a decision support system that can categorize Twitter sentiments and explain prediction outcomes using XAI approaches was done by Hamed [11]. His study produced the following average accuracy numbers: 86% for the SVM model, 78.4% for RNN, 78.8% for LSTM, 78.6% for GRU, and 79% for Bi-directional RNN.

The Long Short-Term Memory (LSTM) model, Multichannel Convolutional Neural Network (MC-CNN), and K-Nearest Neighbors (KNN) have all been used in research by Alenezi [12] to identify COVID-19 disinformation on Twitter. The LSTM model performs with 95.51% accuracy, 99.99% recall, 97.7% F-measure, and 95.51% precision. Chandra [13] conducted research on sentiment analysis of COVID-19 cases in India using the BERT and LSTM models. The results of this investigation revealed that most of the tweets were optimistic, annoyed and joking, expressing optimism, fear and uncertainty during the emergence of COVID-19 cases in India. Alam [14] studied the sentiment of the COVID-19 vaccination reaction using data from Twitter. LSTM and Bi-directional LSTM (BiLSTM) models were employed. According to Twitter statistics, 33.96% of people had positive reactions to the COVID-19 vaccination, 17.55% had negative reactions, and 48.49% had neutral reactions. The accuracy of the LSTM performance is 90.59%, while the accuracy of the BiLSTM is 90.83%.

Yeasmin [3] studied the COVID-19 pandemic predictions and sentiment analysis of Twitter users. The CNN and LSTM models, in addition to three-dimensional Glove embedding, Word2Vec embedding, and encoding methods for data conversion, are all examples of deep learning. According to the study, most user sentiments are neutral. It was also discovered that the deep learning developed utilizing Word2Vec and the feature extraction method for encoding exceeded Glove's method for embedding. The Word2Vec extraction method combined with the LSTM yields the most significant result. Regarding the COVID-19 epidemic, a case study of lockdown in New York, Miao [15] researched attitude detection and opinion monitoring by combining Glove and LSTM. Precision, recall, F1-score, and accuracy were all attained with performance outcomes from the LSTM and Glove combination of 0.57, 0.56, 0.55, and 0.69. Based on LSTM and CNN, Al-Sarem [16] did a study on detecting COVID-19 rumors on social media. Three static word embedding models, including word2vec, Glove, and FastText, were used in the study. The accuracy, precision, recall, and F1-score performance metrics for the LSTM + Glove are 85.42%, 0.86, 0.85, and 0.85, respectively.

A BiLSTM model study was done by Arbane [17] on the classification of COVID-19 sentiment on social media. The performance outcomes are 83% accurate. During the COVID-19 epidemic, To [18] studied the use of machine learning to spot anti-vaccine tweets. In this study, the performance of the Bi-LSTM and bidirectional encoder representations from transformers (BERT) is compared using traditional machine

learning techniques like the support vector machine (SVM) and naive bayes (NB). The results of the BERT model outperform the Bi-LSTM, SVM, and NB models. Using COVID-19 tweets, Kabir [19] studied the identification, analysis, and visualisation of emotions. The investigation demonstrates how the epidemic causes a rise in negative sentiments. It also demonstrates how individuals have grown more optimistic as a result of the pandemic.

Mengistie [20] studied sentiment analysis using deep learning to analyze public reviews of COVID-19. In his study, a hybrid model made up of CNN and BiLSTM was employed with FastText and GloVe word embeddings. The accuracy of the CNN-BiLSTM model with the FastText pre-trained model was 99.33%, while the GloVe pre-trained model was 97.55%.

To the best of our knowledge, no research has used various types of RNN algorithms and word embedding for COVID-19 classification tweets. This study addressed a gap in the literature by assessing the best RNN models for text classification using GloVe word embedding about COVID-19.

B. Text Classification Based on Recurrent Neural Network

Recurrent Neural Network (RNN) architecture, one of the Artificial Neural Network (ANN) designs, is typically used in text classification models based on deep learning [21]. In this work, the RNN was built using four different architectures: the Simple RNN (Vanilla RNN), GRU, LSTM, and Bi-directional LSTM. Each of these structures has the word embedding approach implemented using a 6B GloVe with a dimension of 50.

1) *Recurrent Neural Network (RNN)*: The recurrent neural network (RNN) allows the processing of sequential input for sequence identification and prediction [21]. Since the input and output in conventional neural networks are independent, the following word in a phrase is frequently predicted to be problematic. With the addition of a hidden layer, which serves as a memory for certain information in a sequence, RNN can address the challenges.

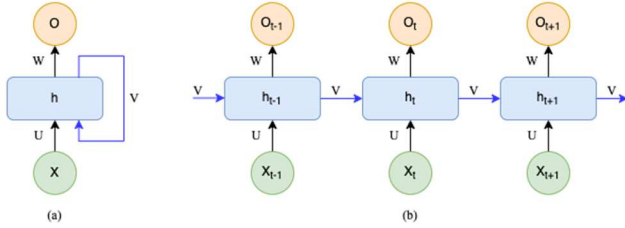


Fig. 1 Architecture of (a) Folded RNN, and (b) Unfolded RNN

Fig. 1 depicts the RNN architecture, where x , o , and h represent the input, output, and hidden states, respectively. Whereas U , V , and W represent the internal parameters.

The input layer consists of a vector series over time t $\{\dots, x_{t-1}, x_t, x_{t+1}, \dots\}$, where $x_t = (x_1, x_2, \dots, x_N)$. The hidden units in the hidden layer, which are determined by a weight matrix U , are connected to the input units in the RNN. The hidden layer has M hidden units, denoted by the notation $h_t = (h_1, h_2, \dots, h_M)$, which are linked to one another throughout the time period via recurrent connections.

The hidden layer defines the state space or memory of the system as $h_t = f_h(O_t)$, where $O_t = Ux_t + Vh_{t-1} + b_h$. The bias vector of the hidden unit is denoted by b_h , while the

activation function for the hidden layer is specified as $f_h()$. The hidden unit is connected to the output layer by a W -weighted connection.

The W units in the output layer are determined as $O_t = f_o(Wh_t + b_o)$, where $f_o()$ is the activation function and b_o is the bias vector in the output layer. The output layer has the following W units: $O_t = (O_1, O_2, \dots, O_p)$. Since the input-target pair is sequential all the time, the above steps are repeated consequently $t = (1, 2, \dots, T)$.

2) *Long Short-Term Memory (LSTM) and Bi-directional LSTM (BiLSTM)*: By utilising the capacity to comprehend sequential relationships, recurrent connections can enhance the efficiency of neural networks [21]. However, the technique used to train the RNN can significantly restrict the memory created by recurrent connections. All the models used thus far had experienced exploding or vanishing gradients during the training process, which prevented the network from learning long-term sequential dependencies in data.

The concept of Long Short-term Memory (LSTM), introduced by [22], is one of the most well-liked and effective ways to reduce the vanishing and exploding gradient effects. This method converts the "sigmoid" or "tanh" hidden unit structure into a memory cell, where gates manage the inputs and outputs. This model resembles a recurrent network (Fig. 1) by using cell memory in the hidden state part, as seen in Fig. 2. A cell memory is constructed of input, forget, output, and cell activation gates.

First, the forget gate determines which data will be removed from the cell state. Then the input gate chooses which data will be updated in the cell state. The cell state could be modified after establishing these two points. Finally, the output gate chooses what network will be produced in its final form.

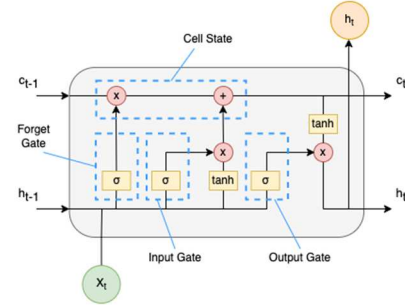


Fig. 2 LSTM Cell Memory

The following equations 1 to 6 define the state of each node in this process.

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

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \quad (2)$$

$$\tilde{C}_t = (\tanh(W_c \cdot [h_{t-1}, x_t]) + b_c) \quad (3)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t, \quad (4)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \quad (5)$$

$$h_t = O_t * \tanh(x_t), \quad (6)$$

where x_t represents the current input, W and b represent the weights and biases, and h_{t-1} represents the hidden state of the preceding layer. Additionally, σ is the sigmoid function, f_t displays the forget gate's output, and c_{t-1} and c_t respectively

show the cell states of the previous and current layers. The output of the output gate and the hidden state are shown to the following layer, respectively, by o_t and h_t .

On the other hand, the Bi-directional LSTM (BiLSTM) model is an improved model that addresses issues with the LSTM standard for classification scenarios [23]. This model employs a forward hidden layer and a backward hidden layer to analyze input data in two directions (from left to right and from right to left), as the architecture seen in Fig. 3.

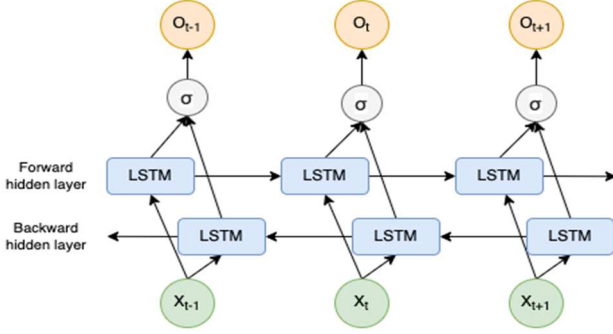


Fig. 3 BiLSTM Architecture

3) *Gate Recurrent Unit (GRU)*: Short-term memory in conventional RNNs can be overcome using one of the models introduced by [24]. Similar to LSTMs, Gated Recurrent Units (GRUs) manage and maintain information via reset gates and update gates. The unit cell of the GRU is seen in Fig. 4.

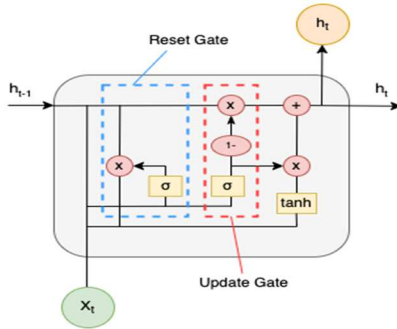


Fig. 4 GRU Cell's Components

The hidden state of GRU is calculated using Equation 7–10 below:

$$r_t = \sigma(W_{xr}x_t + W_{hr}h_{t-1} + b_r), \quad (7)$$

$$z_t = \sigma(W_{xz}x_t + W_{hz}h_{t-1} + b_z), \quad (8)$$

$$\tilde{h}_t = \tanh(W_{xh}x_t + W_{hh}(r_t \odot h_{t-1}) + b_h), \quad (9)$$

$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot \tilde{h}_t \quad (10)$$

where the reset gate, update gate, input vector, and output vector are denoted by the signs r_t , z_t , x_t , and h_t , respectively. W and b display the weights and biases. Meanwhile, the circled dot operator \odot is element-wise multiplication.

C. Method

This study proposes a framework using the RNN model and its updates, such as LSTM, BiLSTM, and GRU. This study aims to determine the best recurrent network model for handling cases of Twitter data classification related to COVID-19 topics and lockdowns. In addition, this study aims

to prove whether using GloVe pre-trained word embedding [25] can increase the accuracy of model predictions.

1) *Experimental Setup*: We employed eight scenarios, as indicated in Table 1, to determine the optimal RNN model and the impact of word embedding. It is believed that sequential input and output models have successfully employed RNN approaches. The RNN model diagram for multi-label classification is shown in Fig. 5, where the green circle, blue box, and yellow circle represent the input, recurrent hidden layer, and target or output, respectively.

TABLE I
MODEL SCENARIOS

No	Word embedding	Model
1		RNN
2	GloVe word embedding	LSTM
3		BiLSTM
4		GRU
5		RNN
6	Without word embedding	LSTM
7		BiLSTM
8		GRU

A keras library running on Python 3.7.10 was employed in this experiment. The experiment was run on macOS machines with Intel dual core i5 specifications with a speed of 2.7GHz with 8GB of memory.

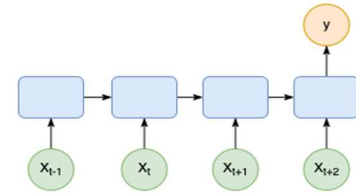


Fig. 5 RNN model for text classification

2) *GloVe Word Embedding*: Word embedding is a sort of word representation that enables representations of words with related meanings. Word embedding aims to transform textual data into vector values. Term Frequency-Inverse Document Frequency (TF-IDF) [26], [27], Latent Semantics Analysis [28], and Neural Network Models like Word2Vec and GloVe [25] are several methods that have been utilized to represent words as vector values. The Global Vector for Word Representation (GloVe) is an extension of the word2vec method to make word vectors more efficient. The GloVe technique combines local context-based learning at word2vec with global statistics from factorization matrices like LSA. The GloVe uses unsupervised learning techniques in which a group of words' meanings are recognized without the need for human verification.

The most used words are gathered as the context in the first phase of GloVe. To create the co-occurrence matrix X , the next step is to scan the corpus of words. P_j^i is the probability of the word j appearing in the context word i , where i is the index of the words that appear frequently and j is the other words in the corpus,

$$P_j^i = P(j|i) = \frac{x_{ij}}{x_i}. \quad (10)$$

The probability ratio of co-occurrence can be calculated by considering the two words i and j and the context word k ,

$$F(w_i, w_j \tilde{w}_k) = \frac{P_{ij}}{P_{jk}}. \quad (11)$$

The loss function J can be determined at the last step using the following equation.

$$J = \sum_{i,j=1}^V f(X_{ij})(W_i^T \tilde{w}_j + b_i + b_j - \log X_{ij})^2. \quad (12)$$

The training aims to minimize the least square error. Each word in this training is assigned to a different vector value.

3) *Dataset*: We utilized Twitter data relevant to COVID-19 and the lockdown from March 23 to July 15, 2020, which we obtained from Kaggle [29] for training and testing. After going through the data cleaning and labeling processes, the text is ready to be employed in experiments.

This dataset has four labels: fear, sadness, anger, and joy. There are 3,090 total datasets, and each label has an associated value of 801 for class fear, 795 for sad, 767 for anger, and 727 for joy.

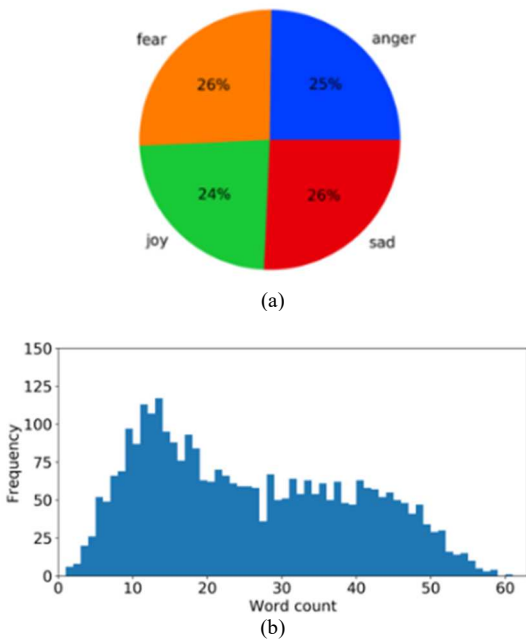


Fig. 6 (a) The composition of the dataset for each class, (b) Distribution of words in the dataset

As shown in Fig. 6(a), the label data composition is balanced, hence no additional data balancing step is necessary. Meanwhile, there are 73,356 unique words in the dataset, with an average word count of 23.73 and a maximum word count of 60. Fig. 6 depicts the overall dataset distribution and composition.

4) *Model Configuration*: The computing process was not considered in the experiment since we employed a small dataset, and it did not significantly impact the model's performance. The model utilized has 64 units, and the softmax function is applied in a dense layer at the final layer to address the text classification problems. Several outputs were fed to a dropout layer (value of 0.5), followed by a global max pooling layer, a dense layer, and four softmax layers on the top of the network.

The training and testing datasets were split into 80% and 20%, respectively. 20% of the training data was further used for validation during the training process. The input consists of text up to 60 characters long and possibly contain symbol t between 1 to 60. The pre-trained GloVe used has a dimension of 100 with a total of 6 million tokens from the data source Wikipedia 2014 and Gigaword 5 [25].

The issue with neural network training is to find out how many epochs to employ. A model that has too many epochs is overfitted, whereas a model with too few epochs is underfitted. Therefore, in this experiment, an early stopping technique was used with a limit of 15 epochs and a minimum delta of 0.01, meaning that training will be stopped if there is no improvement of 0.1% accuracy after 15 epochs. The evaluation was carried out by f1-score. We used the f1-score average as a measure of the accuracy of the classification task results.

III. RESULTS AND DISCUSSION

The results demonstrate that, in the word-embedded COVID-19 twitter dataset, the recurrent network model with memory expansion as in LSTM and GRU outperforms the baseline RNN model in terms of accuracy values. This is in line with research by [30], [31]. However, the accuracy of the dataset without word embedding indicates not much different from that of the RNN model and the other three models.

Furthermore, utilizing word embedding GloVe improves text classification accuracy compared to not using word embedding in all the evaluated models. The results of this study are aligned with that of [5]. Similarly, among the three models, the bidirectional technique for the 2-way training process presents the greatest accuracy across all classes.

The stopping epoch outcomes for each model were evaluated throughout the model training process with a maximum of 100 epochs using the Adam optimizer, a learning rate value of 0.01. We use a similar configuration for each model with an embedding dimension of 100, and the maximum sequence length is 60. The early-stopping approach is displayed in Table 3.

TABLE II
F1-SCORE FROM THE EXPERIMENTS

Class	Without word embedding				GloVe word embedding			
	RNN	LSTM	BiLSTM	GRU	RNN	LSTM	BiLSTM	GRU
Sad	0.53	0.55	0.56	0.57	0.65	0.67	0.67	0.66
Joy	0.57	0.53	0.53	0.57	0.59	0.66	0.70	0.67
Fear	0.69	0.70	0.68	0.72	0.75	0.79	0.82	0.76
Anger	0.70	0.67	0.68	0.66	0.63	0.74	0.78	0.75
Accuracy	0.62	0.61	0.61	0.63	0.66	0.72	0.74	0.71

According to Table 3, the number of epochs of the RNN model with and without GloVe word embedding stopped at

26 and 28, respectively. Meanwhile, the other three models stopped between epochs 16 and 18. This demonstrates that the

LSTM, BiLSTM, and GRU models converged faster than the RNN model.

Moreover, in all model testing, the f1-score value of 'fear' class displays the highest accuracy compared to other classes. The COVID-19 dataset demonstrates that it is simpler to predict fearful sensations, particularly when GloVe word embedding is used.

TABLE III
STOPPING EPOCH FOR MODEL TRAINING

No	Word embedding	Model	Stopping epoch
1		RNN	26
2	GloVe word	LSTM	18
3	embedding	BiLSTM	16
4		GRU	18
5		RNN	28
6	Without word	LSTM	16
7	embedding	BiLSTM	18
8		GRU	17

IV. CONCLUSION

This study used the Twitter dataset to compare the RNN model with other RNN derivative models about COVID-19 and lockdown. In general, predictions without word embedding generate model accuracy that is similar to each RNN model. The application of GloVe word embedding greatly improved the accuracy of all models' predictions. The evaluation shows that the BiLSTM model with GloVe word embedding yields the best accuracy compared to other models. The training produces a model, which is compiled in the H5 file format. Based on our experiment results, it is feasible to implement the training model for sentiment prediction related to COVID-19, such as real-time opinion prediction, social media sensing, and so forth. We believe that using this model for automated opinion prediction in actual data will help the authorities determine future policies.

Despite the promising finding, there is still much to learn about text classification, especially when it comes to different languages and datasets. Future research studies are expected to examine the usage of different word embedding techniques, such as word2vec or fastText. Additionally, verifying the configuration of the RNN architecture's number of layers and other models is a remaining challenge to be resolved by researchers in the scope of text classification.

ACKNOWLEDGMENT

This research and APC were funded by the Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) Universitas Amikom Yogyakarta following the grant number RP-1675234867000.

REFERENCES

[1] H. Kaur, S. U. Ahsaan, B. Alankar, and V. Chang, "A Proposed Sentiment Analysis Deep Learning Algorithm for Analyzing COVID-19 Tweets," *Information Systems Frontiers*, vol. 23, no. 6, pp. 1417–1429, 2021, doi: 10.1007/s10796-021-10135-7.

[2] N. Chintalapudi, G. Battineni, and F. Amenta, "Sentimental analysis of COVID-19 tweets using deep learning models," *Infect Dis Rep*, vol. 13, no. 2, pp. 329–339, 2021, doi: 10.3390/IDR13020032.

[3] N. Yeasmin *et al.*, "Analysis and Prediction of User Sentiment on COVID-19 Pandemic Using Tweets," *Big Data and Cognitive Computing*, vol. 6, no. 2, 2022, doi: 10.3390/bdcc6020065.

[4] C. Singh, T. Imam, S. Wibowo, and S. Grandhi, "A Deep Learning Approach for Sentiment Analysis of COVID-19 Reviews," *Applied Science*, vol. 12, no. 8, 2022, doi: 10.3390/app12083709.

[5] E. Alabdulkreem, "Prediction of depressed Arab women using their tweets," *J Decis Syst*, vol. 30, no. 2–3, pp. 102–117, 2021, doi:10.1080/12460125.2020.1859745.

[6] M. Edalati, A. S. Imran, Z. Kastrati, and S. M. Daudpota, "The Potential of Machine Learning Algorithms for Sentiment Classification of Students' Feedback on MOOC," *Lecture Notes in Networks and Systems*, vol. 296, pp. 11–22, 2022, doi: 10.1007/978-3-030-82199-9_2.

[7] W. H. Bangyal *et al.*, "Detection of Fake News Text Classification on COVID-19 Using Deep Learning Approaches," *Comput Math Methods Med*, vol. 2021, 2021, doi: 10.1155/2021/5514220.

[8] D. S. Abdelminaam, F. H. Ismail, M. Taha, A. Taha, E. H. Houssein, and A. Nabil, "CoAID-DEEP: An Optimized Intelligent Framework for Automated Detecting COVID-19 Misleading Information on Twitter," *IEEE Access*, vol. 9, no. December 2019, pp. 27840–27867, 2021, doi: 10.1109/access.2021.3058066.

[9] A. S. Raamkumar, S. G. Tan, and H. L. Wee, "Use of health belief model-based deep learning classifiers for COVID-19 social media content to examine public perceptions of physical distancing: Model development and case study," *JMIR Public Health Surveill*, vol. 6, no. 3, pp. 1–8, 2020, doi: 10.2196/20493.

[10] P. Pathwar and S. Gill, "Tackling COVID-19 Infodemic Using Deep Learning," *Lecture Notes on Data Engineering and Communications Technologies*, vol. 99, pp. 319–335, 2022, doi: 10.1007/978-981-16-7182-1_26.

[11] S. H. Hamed, H. Elbakry, H. Elghareeb, and S. Elhishi, "Using XAI Techniques to Persuade Text Classifier Results: A Case Study of Covid-19 Tweets," *Indian J Sci Technol*, vol. 15, no. 30, pp. 1484–1494, 2022, doi: 10.17485/ijst/v15i30.397.

[12] M. N. Alenezi and Z. M. Alqenaei, "Machine learning in detecting covid-19 misinformation on twitter," *Future Internet*, vol. 13, no. 10, pp. 1–20, 2021, doi: 10.3390/fi13100244.

[13] R. Chandra and A. Krishna, "COVID-19 sentiment analysis via deep learning during the rise of novel cases," *PLoS One*, vol. 16, no. 8 August, pp. 1–26, 2021, doi: 10.1371/journal.pone.0255615.

[14] K. N. Alam *et al.*, "Deep Learning-Based Sentiment Analysis of COVID-19 Vaccination Responses from Twitter Data," *Comput Math Methods Med*, vol. 2021, 2021, doi: 10.1155/2021/4321131.

[15] L. Miao, M. Last, and M. Litvak, "Tracking social media during the COVID-19 pandemic: The case study of lockdown in New York State," *Expert Syst Appl*, vol. 187, p. 115797, 2022, doi:10.1016/j.eswa.2021.115797.

[16] M. Al-Sarem, A. Alsaedi, F. Saeed, W. Boulila, and O. Ameerbakhsh, "A novel hybrid deep learning model for detecting covid-19-related rumors on social media based on lstm and concatenated parallel cnns," *Applied Sciences (Switzerland)*, vol. 11, no. 17, 2021, doi:10.3390/APP11177940.

[17] 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, no. November 2021, p. 118710, 2023, doi: 10.1016/j.eswa.2022.118710.

[18] Q. G. To *et al.*, "Applying machine learning to identify anti-vaccination tweets during the covid-19 pandemic," *Int J Environ Res Public Health*, vol. 18, no. 8, 2021, doi: 10.3390/ijerph18084069.

[19] M. Y. Kabir and S. Madria, "EMOCOV: Machine learning for emotion detection, analysis and visualization using COVID-19 tweets," *Online Soc Netw Media*, vol. 23, no. September 2020, p. 100135, 2021, doi: 10.1016/j.osnem.2021.100135.

[20] T. T. Mengistie and D. Kumar, "Deep Learning Based Sentiment Analysis on COVID-19 Public Reviews," *3rd International Conference on Artificial Intelligence in Information and Communication, ICAIIC 2021*, no. April, pp. 444–449, 2021, doi: 10.1109/ICAIIIC51459.2021.9415191.

[21] H. Salehinejad, S. Sankar, J. Barfett, E. Colak, and S. Valaee, "Recent Advances in Recurrent Neural Networks," Dec. 2017, doi:10.48550/arxiv.1801.01078.

[22] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.

[23] Y. Imrana, Y. Xiang, L. Ali, and Z. Abdul-Rauf, "A bidirectional LSTM deep learning approach for intrusion detection," *Expert Syst Appl*, vol. 185, p. 115524, Dec. 2021, doi:10.1016/j.eswa.2021.115524.

- [24] K. Cho *et al.*, "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation," *EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*, pp. 1724–1734, Jun. 2014, doi:10.48550/arxiv.1406.1078.
- [25] J. Pennington, R. Socher, and C. D. Manning, "GloVe: Global Vectors for Word Representation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014, pp. 1532–1543. doi: 10.3115/v1/D14-1162.
- [26] K. S. Jones, "A statistical interpretation of term specificity and its application in retrieval," *Journal of Documentation*, vol. 28, no. 1, pp. 11–21, 1972, doi: 10.1108/EB026526/full/pdf.
- [27] S. Robertson, "Understanding inverse document frequency: On theoretical arguments for IDF," *Journal of Documentation*, vol. 60, no. 5, pp. 503–520, 2004, doi: 10.1108/00220410410560582/FULL/PDF.
- [28] Scott Deerwester, Susan T. Dumais, George W. Furnas, and Thomas K. Landauer, "Indexing by latent semantic analysis," *Journal of the American Society for Information Science*, vol. 41, no. 6, pp. 391–407, 1990, doi: 10.1002/aris.1440380105.
- [29] S. Kumar, "Covid 19 Indian Sentiments on covid19 and lockdown," *Dataste of Twiiter sentiment of indians on covid 19*. Accessed: January 07, 2023. [Online]. Available: <https://www.kaggle.com/surajkum1198/twitterdata>
- [30] A. Shewalkar, D. nyavanandi, and S. A. Ludwig, "Performance Evaluation of Deep neural networks Applied to Speech Recognition: Rnn, LSTM and GRU," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 9, no. 4, pp. 235–245, 2019, doi:10.2478/jaiscr-2019-0006.
- [31] R. Ni and H. Cao, "Sentiment Analysis based on GloVe and LSTM-GRU," *Chinese Control Conference, CCC*, vol. 2020-July, pp. 7492–7497, Jul. 2020, doi: 10.23919/CCC50068.2020.9188578.
- [29] S. Kumar, "Covid 19 Indian Sentiments on covid19 and lockdown," *Dataste of Twiiter sentiment of indians on covid 19*. Accessed: January 07, 2023. [Online]. Available: <https://www.kaggle.com/surajkum1198/twitterdata>
- [30] A. Shewalkar, D. nyavanandi, and S. A. Ludwig, "Performance Evaluation of Deep neural networks Applied to Speech Recognition: Rnn, LSTM and GRU," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 9, no. 4, pp. 235–245, 2019, doi:10.2478/jaiscr-2019-0006.
- [31] R. Ni and H. Cao, "Sentiment Analysis based on GloVe and LSTM-GRU," *Chinese Control Conference, CCC*, vol. 2020-July, pp. 7492–7497, Jul. 2020, doi: 10.23919/CCC50068.2020.9188578.