KnE Social Sciences
ICESRE
6th International Conference on Education and Social Science Research
Volume 2024

Research Article

# Supporting Learning Information System through Knowledge Management Optimization using Long Short-Term Memory Method

**Ria Rizki Amelia[1*] and Doddy Teguh Yuwono [2]**

[1]Information Technology, AMIK PGRI Kebumen, 54381, Indonesia
[2]Bussiness and Information Faculty, Muhammadiyah University of Palangkaraya, 73111 Indonesia

**ORCID**
Ria Rizki Amelia: https://orcid.org/0009-0005-2748-5575

**Abstract.**
Effective information and knowledge management is vital in many areas, including higher education. The use of artificial intelligence (AI) technology, especially the long short-term memory (LSTM) information system performance patterns in the educational world. This article explores the application of LSTM to optimize knowledge management in colleges, focusing on the prediction of information systems performance. The proposed methods include text classification steps, with measures such as data collection, data pre-processing, word representation, classification, and evaluation. The test results showed that the LSTM model managed to classify reviews labeled positive, neutral, and negative with an accuracy of 33.33%. However, the success of the model was limited by the size of the data set and the pre-processing involved. This research recommends further development with the addition of experimental data, proper preprocessing adjustments, and better hyperparameter identification to improve the accuracy of the prediction results.

**Keywords:** information management, artificial intelegence, LSTM, text classification, knowledge management, accurate prediction

Corresponding Author: Ria Rizki Amelia; email: riarizki@amikpgrikebumen.ac.id

**Published** 12 March 2024

**Publishing services provided by Knowledge E**



Selection and Peer-review under the responsibility of the ICESRE Conference Committee.



## 1. Introduction

In a modern context, effective management of information and knowledge plays a central role in various fields [1] Some of these areas, higher education institutions are centers of learning and research, which rely heavily on information systems to support academic, administrative, and managerial processes. As the system becomes more complex and critical, decision-making based on performance data becomes crucial [2]. To meet these challenges, integration of artificial intelligence (AI) technology emerged as a promising solution [3], through the implementation of LSTM algorithm [4]

**How to cite this article:** Ria Rizki Amelia and Doddy Teguh Yuwono , (2024), “Supporting Learning Information System through Knowledge Management Optimization using Long Short-Term Memory Method ” in 6th *International Conference on Education and Social Science Research*, KnE Social Sciences, pages 361–371. DOI 10.18502/kss.v9i6.15285

The application of LSTM to the prediction and analysis of sequential data has proven to be effective in a range of fields ranging from stock market analysis [5-6] to weather forecasts [4], [7] and natural language processing [8]. In the context of optimizing knowledge management in colleges, the potential of LSTM lies in its ability to predict and identify performance patterns of school information systems [10-11]. It provides valuable information for management, supports decision-making to improve service quality, operational efficiency, and the ability to respond to emerging challenges [12].

However, integrating this kind of technology into the academic world is not without a challenge. Developing and implementing LSTM models requires a deep understanding of data, careful pre-processing, and smooth integration with existing knowledge management systems [13]. In addition, data security and privacy issues must be carefully considered, bearing in mind that the performance estimation process may involve sensitive information [14].

This article aims to examine in depth the concept of knowledge management optimization in colleges through the application of LSTM algorithms. This research will examine how LSTM can be used to predict campus information system performance, with a focus on model development, performance evaluation, and integration of predictions into performance management systems. The expected results of this research provide valuable information for educational institutions that are trying to improve the management of information systems on campus as well as facilitate decision-making processes.

# 2. Method

## 2.1. Text Classification

In this study, to conduct a classification of text requires involving several processes, namely: Data Collection, Data Preprocessing, Word Representation, Classification, and Evaluation/Testing.

The first step in text classification is Data Collection. Data is a collection of review text and is obtained from the campus site. The next step is Data Preprocecing/Cleaning. Datasets [2]. The review text data used still contains some noise. Therefore, the data needs to be processed first. Process Cleaning of Datasets includes:

1. Case Folding, is the process of turning all the capital letters into lower cases.

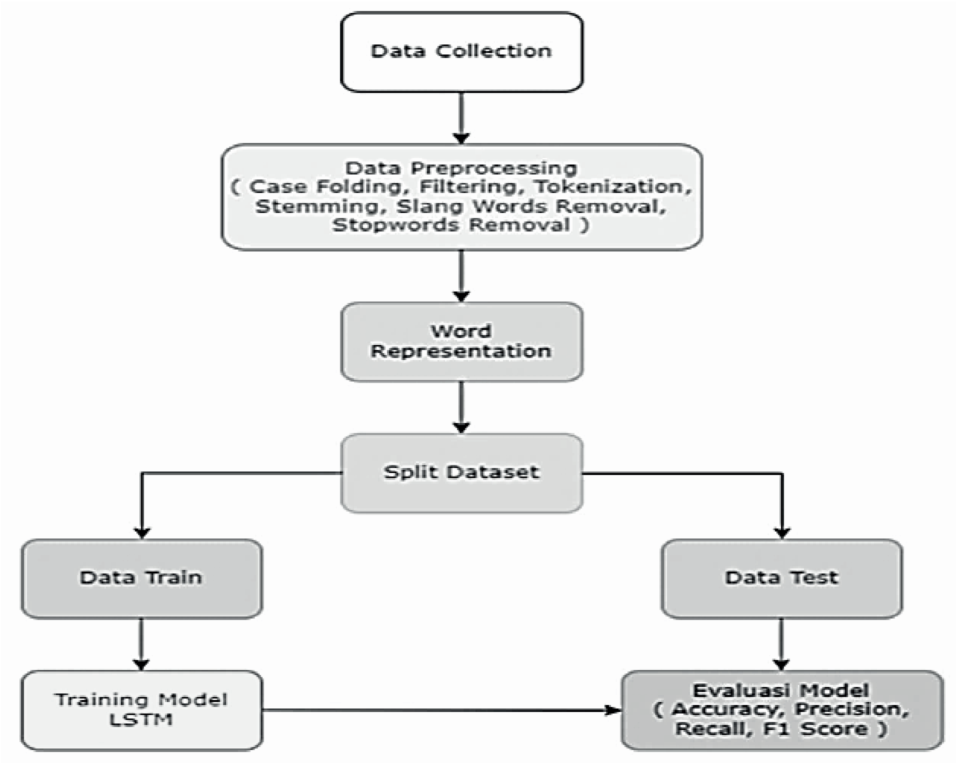


**Figure** 1: Proses Klasifikasi Teks.

2. Filtering, is the process of removing some unnecessary special characters, such as reading marks, numbers, and some other symbols that are not needed in the training process.

3. Tokenization, is a process of cutting sentences into words per word, it is done by cutting a sentence with the reference of each space.

4. Stemming, is the process of removing the substance of a word so that it becomes a basic word.

5. Slang Words Removal, is an process of deleting every vocabulary gaul (slang).

6. Stopwords Removal is a procedure to remove words that do not have meaning or important information in the training process.

After performing the Cleaning of Dataset process, the next process is Word Representation. The process aims to convert each word into a real-number vector by applying a hard-based Trainable Embedding Layer. The trainable embedding layer initiates random weight values and is updated during the model training process using a backpropagation algorithm. The data that has been converted has a dimension, which is 1143 × 20.[7]

After performing the Word Representation process, the next process is Split-datasets or splitting datasets. Datasets are divided into 2, i.e. 80% data train and 20% data test. To classify text on a review, the algorithm used is LSTM.

Long Short Term Memory Networks – commonly referred to simply as “LSTMs” – are one of the variations of Recurrent Neural Networks with extended memory capabilities that are explicitly suitable for dealing with long-term dependencies. The LSTM network was proposed by German researchers Sepp Hochreiter and Juergen Schmid-huber in 1997 as a solution to the problem of missing gradients. The overall LSTM architecture is shown in Figure 2.

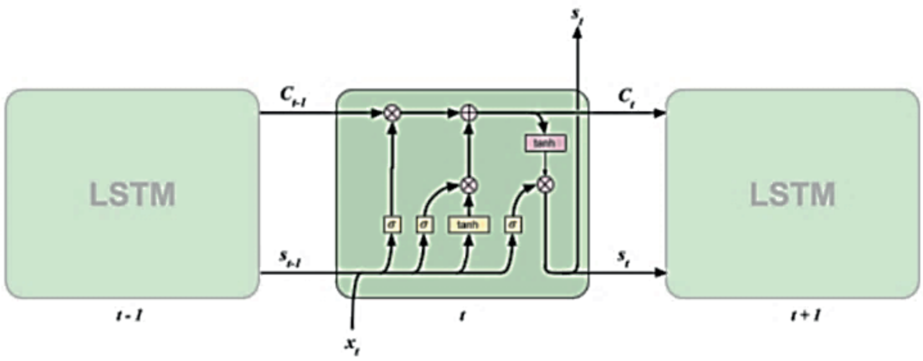


**Figure** 2: Gate of LSTM.

LSTM contains information from the context in the cells that are awake security. Cells control the data to be written, saved, read, and deleted using forget gate, input and output implemented with the element multiplication by the sigmoid [8]. Forget gate studies the weight that controls the rate of decomposition of values stored in memory cells. For example, when input gate and output gate are dead, and forget gate does not cause decay, the memory cellins its value over time causing the error gradient to remain constant during backpropagation [11]. This allows models to remember information for longer periods of time. Mathematically each step can be described as follows:

$$ft = \sigma(Wf[ht\text{-}1,\ xt] + bf)$$

**Figure** 3: Step one.

In the first step, the forget gate layer decides which features will be removed from the cell status by looking at ht–1 and the new xt input.

In the second step, deciding the information stored in the cell status is done in two steps (14). The output gate layer that is the sigmoid layer sets the values to be updated. Then the tan ⊠ layer produces a new candidate value vector $\tilde{C}t$.

$$it = \sigma(Wi[ht\text{-}1, xt] + bi)$$

$$ct = \tanh(Wc[ht-1, xt] + bc)$$

**Figure** 4: Step two.

The old $Ct-1$ cell status is updated to the new $Ct$ cell summarizes the output of the forget gate layer function $ft$ and $it \times \tilde{C}t$[10], [11].

$$ct = ft \times ct-1 + it \times ct$$

**Figure** 5: The old $Ct-1$ cell status

The output is determined in two steps – First, the sigmoid layer decides which parts of the cell are to be removed. The product of the new $Ct$ cell status through the tan ⊠ and the output of the Sigmoid gate produces the selectively determined part $ht$.[4]

Setting up and optimizing hyperparameters is a difficult and experimental task. Costly LSTM model training in terms of memory and computing power [14].

$$ot = \sigma(Wo[ht-1, xt] + bo)$$

$$ht = ot \times \tanh(ct)$$

**Figure** 6: Two steps output.

## 2.2. Evaluation Metrics

Evaluation metrics is a parameter used to measure the quality of a model or machine learning algorithm (13), (14). The evaluation metric used in this evaluation is, Accuracy, Precission, Recall, and also F1-Score. Each of these evaluation metric is formulated as follows:

Accuration is the comparison value of True Positive (TP) and True Negative (TN) predictions with the total amount of data.

Precision is the value of a True Positive (TP) prediction compared to the quantity of predicted positive data.[5]

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

**Figure** 7: Accuration is the comparison value of True Positive (TP) and True Negative (TN).

$$recision = \frac{\mathrm{TP}}{TP + FP}$$

**Figure** 8: Precision is the value of a True Positive (TP).

Recall is the comparison value of a True Positive (TP) prediction to a lot of positive true data.[5]

$$Recall = \frac{\mathrm{TP}}{TP + FN}$$

**Figure** 9: Recall is the comparison value of a True Positive (TP).

There is a difference between precision and recall, where precision has a False Positive variable (FP) whereas recall has the False Negative variable (FN) [5].Next, F1-Score is the comparative value of the weighted precision and recall averages.

$$F1\ Score == \frac{2\mathrm{TP}}{2TP + FN + FP}$$

**Figure** 10: difference between precision and recall.

## 3. Result and Discussion

### 3.1. Dataset

The data set used in this research is a dummy data set from reviews that are on campus websites. The data is a collection of reviews that have been labelled and cleared from the personal information belonging to the user. The existing reviews consist of three different labels, namely Positive, Neutral and Negative.

Here is one example of a Positive Review, "Online rating reviews are very neat and easy to access, all the latest academic information is in handheld.". Next for one is a neutral review, "Campus information system allows students to follow campus events and activities. However, there are some users who feel that this information may be limited or not very relevant. The following is one example of Negative Review, "Functions in campus information systems are often slow in responding to user actions, making the process slow."

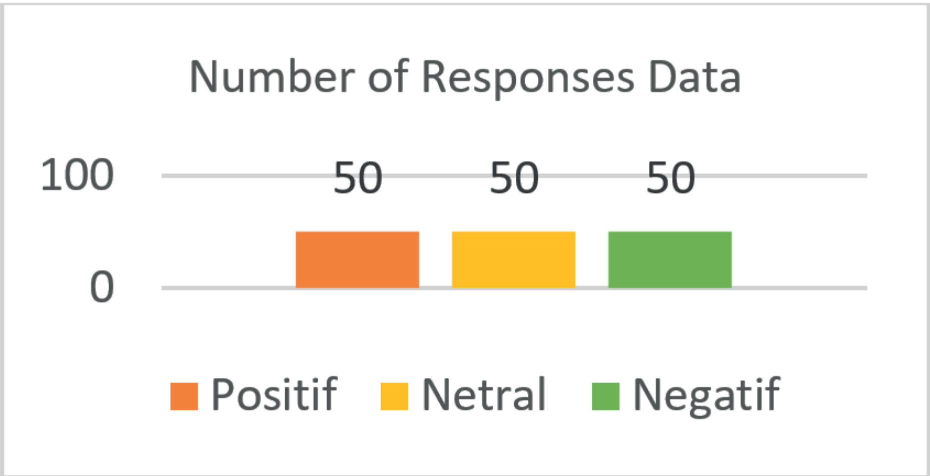


**Figure** 11: Distribution of Responses.

Based on the Figure 12, reviews with positive labels have as many as 50 data; reviews with neutral labels 50 data and reviews with negative labels 50. The total of the total data tagged reviews is 150 which means of each review has a percentage weight of 33.33%.

## 4. Discussion

In this study, the researchers classified reviews into three classes, namely positive labels, neutral labels and negative labels. The algorithm used for the study is LSTM. Figure 13 shows the architecture of the model used. The training model uses the Sequential model with the LSTM architecture and the Dense layer. This model has the first layer of the

Embedding Layer to convert words to vectors, followed by the 64-units LST M layer and the Dens layer with the softmax activation function to generate sentiment probabilities. The model is trained with training data using the fit method. The training process involves several epoch and batch size 32. The test data is also used as validation data during the training process.

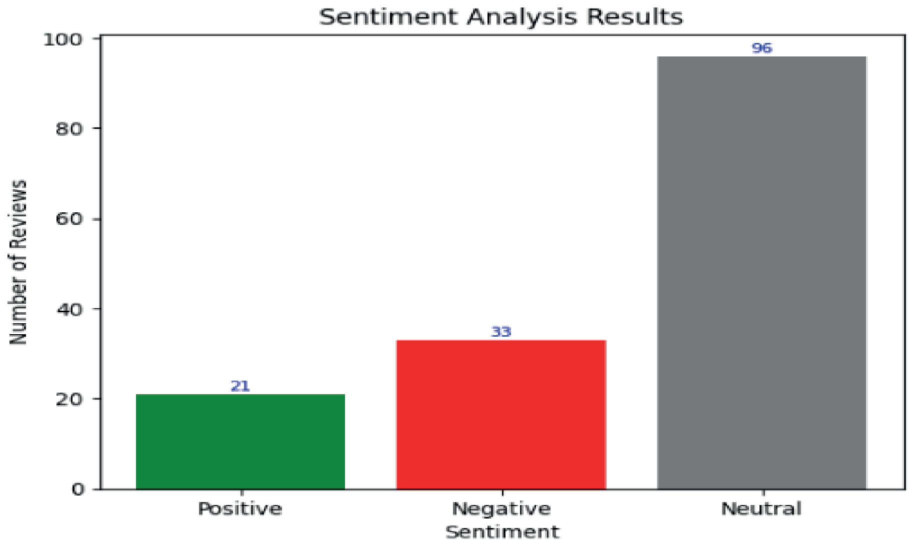


**Figure** 12: Sentiment Analysis Results.

Based on the results of Sentiment Analysis, reviews with positive labels have data of 21, reviews with neutral labels 96 and reviews with negative labels 33. The total of the labeled review data is 150. In the test session of the LSTM model architecture with the hyperparameter set, the following graphical results in Figure 14 were obtained:

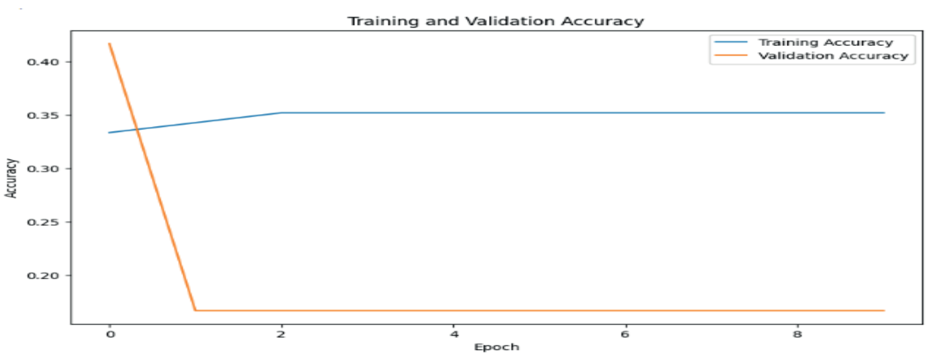


**Figure** 13: Training and Validation Accuracy.

1. Graphic Sentiment Actual (Subplot 1):

   This graph shows the actual sentiment of each tested entity. The color of the stone may be shown in blue (skyblue). From this graph, we can see the distribution of

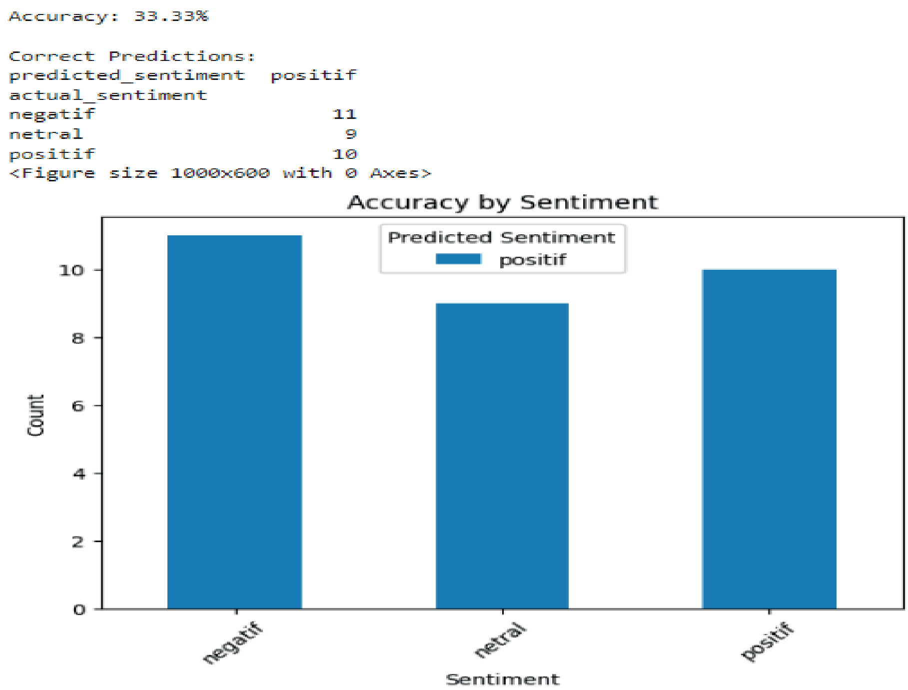


**Figure** 14: Accuracy by Sentiment.

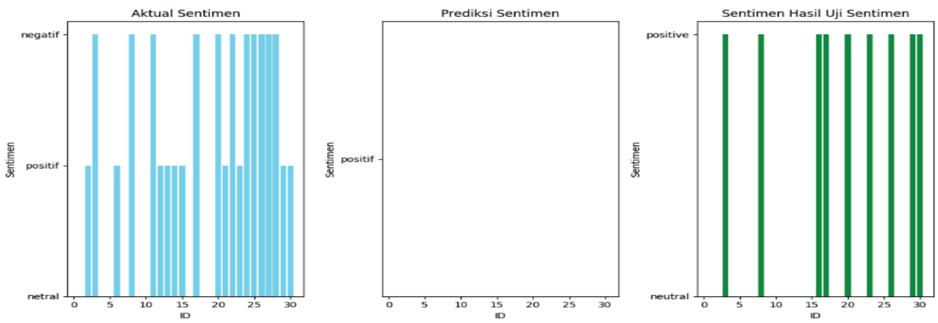


**Figure** 15: Actual, Prediction, Result Sentiment.

actual sentiment from the tested data, whether the majority is positive, negative, or neutral

2. Graphic Sentiment Prediction (Subplot 2):

   This graph shows the sentiment prediction of a trained model. However, in this case, it does not show the Sentiment Prediction of the training model.

3. Graphic Results of Sentiment Test (Subplot 3):

This graph shows a sentiment test result that may be a comparison between actual sentiment and the prediction result. In this graph, we can see where the model successfully predicted correctly shown by the green color.

DOI 10.18502/kss.v9i6.15285

## 5. Conclusion

Based on the results of the research conducted classification of the text on the Indonesian language reviews, the LSTM method produces an accuracy value of 33.33%. Cause of still low accurate results obtained due to the minimum amount of data tested (Dataset) other than the same amount of equal causes the prediction is not performed. The results obtained are some things that need to be recommended for the further development of this research. The researchers suggest that adding test data is crucial in addition to the Data Preprocessing step plays a crucial role in the text classification process, in which the methodology performed must adjust to the characteristics of the data set. In addition, the researchers hope that the development of this study can apply the appropriate hyperparameter setting so that better accuracy results can be obtained.

## References

[1] Putra TI, Suprapto S, Bukhori AF. Model Klasifikasi Berbasis Multiclass Classification dengan Kombinasi Indobert Embedding dan Long Short-Term Memory untuk Tweet Berbahasa Indonesia. Jurnal Ilmu Siber dan Teknologi Digital. 2022 Nov 11;1(1):1-28.

[2] Saputra AW, Wibawa AP, Pujianto U, Utama AB, Nafalski A. LSTM-based Multivariate Time-Series Analysis: A Case of Journal Visitors Forecasting. ILKOM Jurnal Ilmiah. 2022 Apr;14(1):57–62.

[3] Pratama ED. Implementasi Model Long-Short Term Memory (LSTM) pada Klasifikasi Teks Data SMS Spam Berbahasa Indonesia [JMLCI]. The Journal on Machine Learning and Computational Intelligence. 2022 Jul;1(2).

[4] Mahjoub S, Chrifi-Alaoui L, Marhic B, Delahoche L. Predicting Energy Consumption Using LSTM, Multi-Layer GRU and Drop-GRU Neural Networks. Sensors (Basel). 2022 May;22(11):4062.

[5] Selvin S, Vinayakumar R, Gopalakrishnan EA, Menon VK, Soman KP. Stock price prediction using LSTM, RNN and CNN-sliding window model. In2017 international conference on advances in computing, communications and informatics (icacci) 2017 Sep 13 (pp. 1643-1647). IEEE.

[6] Wiranda L, Sadikin M. Penerapan Long Short Term Memory Pada Data Time Series Untuk Memprediksi Penjualan Produk Pt. Metiska Farma. Jurnal Nasional Pendidikan Teknik Informatika: JANAPATI. 2019;8(3):184-96. Zhang Z, Qin H, Yao L, Lu J, Cheng L. Interval prediction method based on Long-Short Term Memory

DOI 10.18502/kss.v9i6.15285 

networks for system integrated of hydro, wind and solar power. Energy Procedia. 2019 Feb;158:6176–82.

[7] Sudriani Y, Ridwansyah IA, Rustini H. Long short term memory (LSTM) recurrent neural network (RNN) for discharge level prediction and forecast in Cimandiri river, Indonesia. InIOP Conference Series: Earth and Environmental Science 2019 Jul 29 (Vol. 299, p. 012037). IOP Publishing.

[8] Masri F, Saepudin D, Adytia D. Forecasting of Sea Level Time Series using Deep Learning RNN, LSTM, and BiLSTM, Case Study in Jakarta Bay, Indonesia. e-Proceeding Eng. 2020;7(2):8544-51.

[9] Wisyaldin MK, Luciana GM, Pariaman H. Pendekatan Long Short-Term Memory untuk Memprediksi Kondisi Motor 10 kV pada PLTU Batubara. J. Kilat. 2020;9(2):311–8.

[10] Afika R, Suprih W, Atikah DA, Fadlan BH. Next word prediction using lstm. Journal of Information Technology and Its Utilization. 2022 Jun;5(1):10–3.

[11] Zhang R. LSTM-based Stock Prediction Modeling and Analysis. In2022 7th International Conference on Financial Innovation and Economic Development (ICFIED 2022) 2022 Mar 26 (pp. 2537-2542). Atlantis Press.

[12] Rizki M, Basuki S, Azhar Y. Implementasi Deep Learning Menggunakan Arsitektur Long Short Term Memory Untuk Prediksi Curah Hujan Kota Malang. Repositor. 2020;2(3):331–8.

[13] Elsworth S, Güttel S. Time series forecasting using LSTM networks: A symbolic approach. arXiv preprint arXiv:2003.05672. 2020 Mar 12.

[14] Cahuantzi R, Chen X, Güttel S. A comparison of LSTM and GRU networks for learning symbolic sequences. InScience and Information Conference 2023 Jul 13 (pp. 771-785). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-37963-5_53.