Indonesian Journal of Artificial Intelligence and Data Mining (IJAIDM)
Vol 7, No.1, March 2024, pp. 141 – 147
p-ISSN: 2614-3372 | e-ISSN: 2614-6150

Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) Methods to Forecast Daily Turnover at BM Motor Ngawi

[1]Larasati, [2]Siti Saadah, [3]Prasti Eko Yunanto
[1,2,3]Faculty of Informatics, Telkom University
Email: [1]larasatii@student.telkomuniversity.ac.id,
[2]sitisaadah@telkomuniversity.ac.id, [3]gppras@telkomuniversity.ac.id

Article Info

Article history:
Received Nov 10th, 2023
Revised Feb 25th, 2024
Accepted Mar 5th, 2024

Keyword:
BM Motor Ngawi
Daily Turnover
Forecasting
LSTM
RNN

ABSTRACT

The number of motorcycles on the report of Indonesian BPS statistics from the Indonesian State Police between 2019 to 2021 by its type has increased annually. Routine motorcycle checks, services, and maintenance are essential to keep a motorcycle in good condition and more durable; therefore, buying spare parts is enlarged in line with the growth of public motorcycle ownership. The necessity of buying spare parts increases with the growth of public motorcycle ownership. Numerous stores in Ngawi offer motorcycle spare parts and check services for routine motorcycle maintenance. One of these stores is BM Motor. To develop an effective product-selling strategy, it is essential to forecast the daily turnover of the shop. To achieve this, the present research aims to analyze the daily turnover using Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM). These methods were applied to a time-series dataset, allowing for an in-depth examination of the patterns and trends in the shop's turnover. The research compares several hyperparameter tunings and scenarios to optimize the models that forecast daily turnover data at the store. The outcomes presented that the LSTM model achieved a lesser MAE score of 0.087, while the RNN model scored 0.092. These findings proved that the LSTM model achieved lower MAE than the RNN model, it means LSTM is more accurate than the RNN model.

Corresponding Author:
Larasati,
School of Computing, Telkom University, Bandung
Jl. Telekomunikasi No. 1, Terusan Buah Batu Bandung 40257, West Java, Indonesia
Email: larasatii@student.telkomuniversity.ac.id
DOI: http://dx.doi.org/10.24014/ijaidm.v7i1.27643

1. INTRODUCTION

The increasing number of motorcycle reports in Indonesian BPS statistics between 2019 to 2021 with a growth percent number of 102.68%. That is a concerning trend that highlights the need for routine motorcycle checks, services, and maintenance. Keeping a motorcycle in good condition and ensuring its durability is of utmost importance. As the number of motorcycles owned by the public continues to grow, acquiring spare parts becomes crucial. With the expansion of public motorcycle ownership, there is a greater demand for spare parts to ensure these vehicles remain in optimal operating condition. In Ngawi, numerous stores offer motorcycle check services and sell spare parts. One of these stores is BM Motor, located in Widodaren since 2003, which provides both motorcycle service and spare parts sales. To adequately prepare for upcoming difficulties and possibilities, it is crucial to develop a highly efficient product sales plan. In order to tackle this issue, the current study centers on a thorough analysis of the daily turnover data prediction for BM Motor. This study utilizes sophisticated analytical approaches, namely Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) methods, to uncover intrinsic patterns and trends in the financial performance of the business over a period of time.

In recent years, many studies have been conducted about forecasting and prediction in any cases. RNN and LSTM models were commonly used, the research had been conducted by Wiranda e. al. [1] predicted pharmaceutical industry product-selling at PT. Metiska Farma's product-selling strategy demands with Mean Absolute Percentage Error (MAPE) score of 12%. Sugiyarto et al. [2] conducted experiments with LSTM-RNN techniques to predict palm oil production. Their approach involved 2 LSTM layers, and they achieved a MAPE score of 2.71% and 2.98% for the training and testing data, respectively. Apple and Microsoft stock price prediction by Hastomo et al. [3] use LSTM to create forecasts with stock movement trends that tend to go down. RNN methods are also used by Sagheer et al. [4] to forecast time-series data of petroleum production with 1890 epoch numbers, generating the best Root Mean Squared Error (RMSE) score of 0.233. Selle et al. [5] used both RNN and LSTM methods to make a prediction comparison of the use of electricity energy with the RMSE average score of 51.05. Moreover, Rafi et al. [6] are using the LSTM Network as one of the research methods for the prediction of short electrical load in Bangladesh power with a Mean Average Error score of 324.7693.

Previous research has shown that LSTM and RNN models are well-suited for accurately processing sequence and time-series data. This research utilizes two models that calculate the performance and accuracy of the model with hyperparameter tuning to forecast the daily turnover of the store. The objective is to improve the shop's capacity to forecast and adjust to changes in demand by utilizing advanced methods of forecasting. This will ultimately contribute to the establishment of a strong and flexible product-selling strategy for BM Motor in the near future.

2. RESEARCH METHOD

2.1. Overview

In this research, a flowchart design, as illustrated in Figure 1, is being developed and implemented. The initial stage, after collecting the dataset, involves preprocessing the data. To facilitate model evaluation, the dataset is divided into two sections: data train and data testing during the data splitting process. The data training phase utilizes an RNN model and an LSTM model. Once the training is completed, model evaluations are conducted to ascertain the most optimal models, with the Mean Absolute Error (MAE) serving as the evaluation metric. The last step is forecasting data, which is done using the most effective model found during the evaluating model phase.

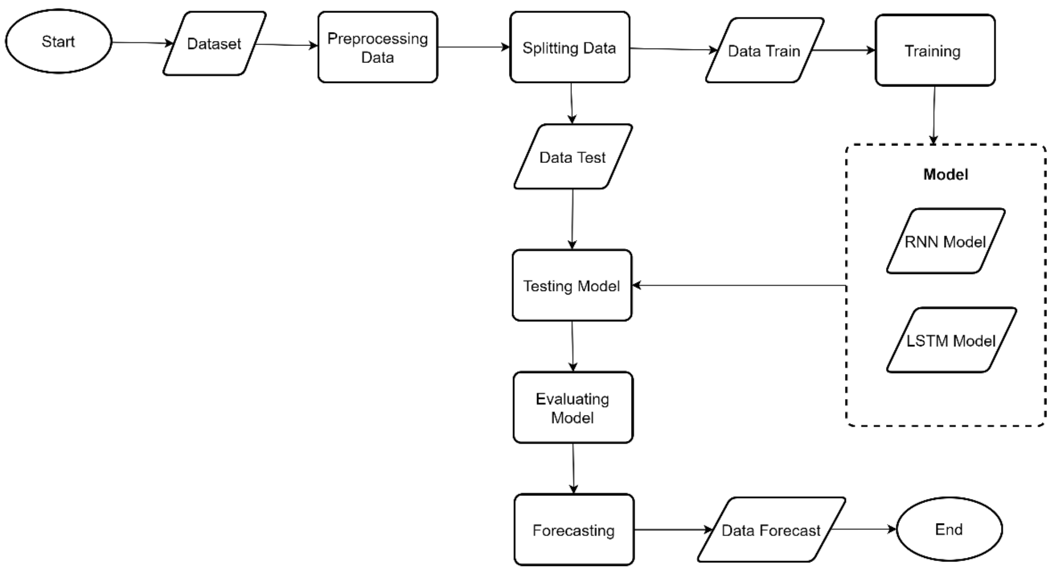


Figure 1. Flowchart Design of Forecasting Experiment

After the dataset has been preprocessed, the next step in the flowchart design is to split the data into data train and data testing. This process involves dividing the dataset into subsets to be used for training and testing purposes. The data training subset is then utilized to train both RNN and LSTM models. These models are trained using various techniques to capture the sequential nature of the data and make accurate predictions. Once the models have been trained, it is essential to evaluate their performances to identify the best models. The evaluation is conducted using MAE as the evaluation score, which counts the errors between predicted and actual values. Lastly, after the evaluation process, the final step in the flowchart design is data forecasting, where the trained models are used to make predictions on new or unseen data instances.

The research uses a daily sales time-series dataset of motorcycle spare parts from BM Motor store, Ngawi. The dataset consists of 2,312 recorded data points including turnover, margin, turnover total, and margin total, as shown in Table 1. This study focuses on turnover profit data and processes univariate data as there is only one type of data used in the experiment.

Table 1. Sample Data of BM Motor

Date	Turnover	Margin	Turnover Total	Margin Total
2016-01-01	388,000	69,000	38,000	69,000
2016-01-02	833,000	156,000	1,221,000	225,000
2016-01-03	504,000	76,000	1,725,000	301,000
2016-01-04	1,575,000	366,000	3,300,000	667,000
2016-01-05	1,121,000	251,000	4,421,000	918,000

To improve data quality, data preprocessing using Min-Max Scaler normalization has been conducted. Then, the data was divided into data train and test. The next step involves training the model with the available data by fitting it with suitable parameters to elevate the performance of the models. Once the training phase is complete, the model's effectiveness is evaluated by testing various scenarios with hyperparameter tuning and determining the best models using MAE as the evaluation score.

The last step is to create a data forecast in the range of periodical times in the forecasting process. Time series forecasting is the process of forecasting upcoming values established on an understanding of past data distributions to involve the future system manner prediction. Time series data is a sequence of data that is composed of specific prevalence, such as either daily or annually from intervals a period that apiece value may have some relationship that is represented by a series of data points listed in chronological order and exist in a wide variety of application areas.

2.2. Recurrent Neural Network

RNN is a method for processing time-series and sequential data that is commonly used within the field of neural networks [7]. The RNN possesses significant strength to utilize the data within a relatively lengthy sequence. RNN method can maintain and capture long-term dependencies as they achieve similar functions for each part in the sequence, with the output of computations intricately tied to all the computations that took place before it [8]. The RNN model is provided in the following equations:

$$h_q = tanh(W_{hx}x_q + W_{hh}h_{q-1} + b_h), \tag{1}$$

$$o_q = W_{oh}h_q + b_o, \tag{2}$$

where W_{hx} ,W_{hh}, and W_{oh} are input-to-hidden, hidden-to-hidden, and hidden-to-output weight matrices, respectively. Bias vectors are denoted with b_h and b_o.

2.3. Long Short-Term Memory

LSTM represents a deviation of RNN [9], commonly referred to as a "gated cell," which extends the RNN model to enable the learning of input with long-term dependencies of each data [10] while minimizing the loss of essential features [11]. Improvement from RNN, this method has the benefit of being able to collect activities over time by cell state [12]. The LSTM model is equipped with three gates that function to regulate its cell state. Firstly, the input gate takes charge of selecting relevant input. Secondly, the forget gate is crucial in eliminating extraneous information. Lastly, the output gate controls the outgoing information flow. These gates work seamlessly to execute the LSTM model's execution [13]. The equations for the LSTM model can be found in Equations (3) - (7):

$$i_i = \sigma(W_i \cdot h_{i-1} + V_i \cdot x_i + b_i), \tag{3}$$

$$o_j = \sigma(W_o \cdot h_{j-1} + V_o \cdot x_j + b_o), \tag{4}$$

$$f_j = \sigma\left(W_f \cdot h_{j-1} + V_f \cdot x_j + b_f\right), \tag{5}$$

$$\tilde{C_j} = \tanh\,(W_C \cdot h_{j-1} + V_C \cdot x_j + b_C\,), \tag{6}$$

$$C_i = i_i \odot \tilde{C_i} + f_i \odot C_{i-1}\,, \tag{7}$$

where i_j, o_j, and f_j represent input gate, output gate, and forget gate. W_i, W_o, W_f, and W_c represent matrices weight, and bias vectors denoted by b_i, b_o, and b_f. Multiplication and sigmoid function are denoted by $\odot$ and σ symbols. Presented in this section is Figure 2, providing a visual representation of the LSTM architecture.

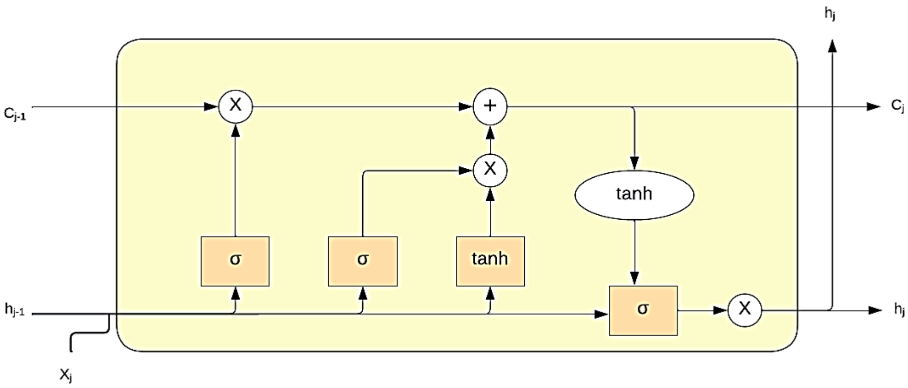


Figure 2. LSTM Cell Architecture [4]

2.3. Mean Absolute Error

MAE is used to measure the closeness between predicted and measured values [14] by calculating the absolute error average [15]. A lower MAE value indicates a more precise model, aiming to get as close to 0 as possible for even greater precision. The given Equation (8) represents the MAE where $\hat{y}_j$ refers to actual data and y_j represents the average of prediction error data.

$$\text{MAE} = \frac{1}{n}\sum_{j=1}^{n}\left|\hat{y}_j - y_j\right| \tag{8}$$

3. RESULTS AND ANALYSIS

The experiments consist of two scenarios to evaluate the effectiveness of the models. Hyperparameter tuning was performed in the first scenario to evaluate the model's parameters such as epoch numbers, data splitting ratio, and dropout number. The objective of this scenario is to identify the best model that can accurately forecast data and provide valuable insights. The measurement of the first scenario used MAE. The second scenario used the setting parameters from the first scenario to forecast with several periods including 6 months, 1 year, 2 years, and 5 years.

Table 2. Hyperparameter Tuning Values

Hyperparameter	Value
Data Split Ratio	70:30, 75:25, 80:20, 85:15, 90:10
Epoch	800; 1000; 1200; 1500
Dropout	0.1; 0.15; 0.2; 0.25

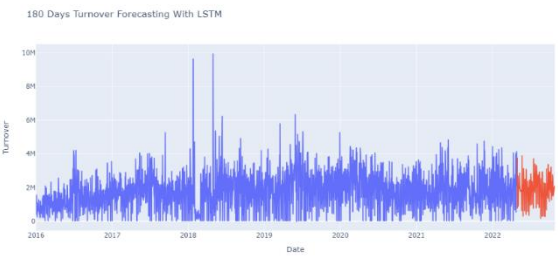
The result of the first scenario is shown in Table 3. This table consists of four main columns: dropout, epoch, and splitting data ratio (including LSTM and RNN models). Each value in the table represents the combination of parameters in Table 2. The best parameter tuning of the LSTM model is in setting the parameter of data splitting in ratio 90:10, drop out number is 0.15, with 800 epoch number yields MAE score result of 0.08753. Besides, the best parameter tuning of the RNN model is in 0.09233 MAE score with the setting data splitting in ratio 90:10, drop out number 0.15, and epoch 1200. Both results prove that 0.15 dropout number and 90:10 data splitting ratio are the best parameters obtained from the first scenario. Afterward, the experiment show the lowest performances of RNN and LSTM model with MAE scores are 0.11191 and 0.10406, where use 85:15 ratio data splitting, 0.1 dropout number, epoch number 1,200 for RNN and 1,000 for LSTM.

Table 3. Hyperparameter Tuning Results

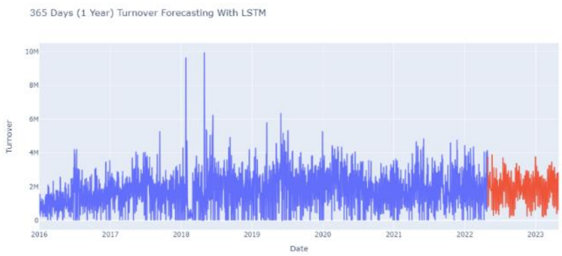
Dropout	Epoch	Data Training: Data Testing Ratio									
		LSTM					RNN				
		70:30	75:25	80:20	85:15	90:10	70:30	75:25	80:20	85:15	90:10
0.10	800	0.09392	0.09588	0.09944	0.09755	0.09252	0.10699	0.10649	0.09485	0.09942	0.10229
	1,000	0.09428	0.09533	0.09710	0.10406	0.09633	0.09940	0.10778	0.09428	0.10187	0.09491
	1,200	0.09479	0.10031	0.09505	0.09718	0.09205	0.09968	0.09749	0.10346	0.10326	0.09734
	1,500	0.09241	0.09952	0.09587	0.09676	0.08840	0.09845	0.09647	0.09667	0.11191	0.09427
0.15	800	0.09169	0.09413	0.09918	0.10015	0.08753	0.10699	0.10649	0.09485	0.09942	0.10229

Dropout	Epoch	Data Training: Data Testing Ratio									
		LSTM					RNN				
		70:30	75:25	80:20	85:15	90:10	70:30	75:25	80:20	85:15	90:10
	1,000	0.09255	0.09120	0.09383	0.09413	0.09547	0.10116	0.10260	0.10399	0.10217	0.10634
	1,200	0.09772	0.09816	0.09631	0.09479	0.09154	0.09791	0.09890	0.10112	0.09593	0.09233
	1,500	0.09241	0.09441	0.09320	0.09976	0.09626	0.10338	0.09971	0.09796	0.10603	0.09427
0.20	800	0.09443	0.09816	0.10379	0.10211	0.08946	0.09511	0.09701	0.10069	0.09558	0.09381
	1,000	0.09657	0.09564	0.09725	0.09823	0.09188	0.10020	0.09762	0.10519	0.09711	0.10450
	1,200	0.09489	0.10011	0.09362	0.09398	0.09848	0.10298	0.09951	0.10349	0.09858	0.09620
	1,500	0.09365	0.09914	0.09677	0.09762	0.09410	0.09783	0.10282	0.10031	0.10220	0.09753
0.25	800	0.09421	0.08909	0.09398	0.10154	0.10057	0.09617	0.09546	0.09519	0.10382	0.10259
	1,000	0.09465	0.09734	0.09912	0.09419	0.09973	0.09882	0.09477	0.09722	0.10566	0.10196
	1,200	0.09489	0.10011	0.09362	0.09398	0.09848	0.09847	0.09962	0.09595	0.10550	0.10741
	1,500	0.09365	0.09914	0.09677	0.09762	0.09410	0.09485	0.10303	0.09719	0.09924	0.10420
0.30	800	0.09459	0.09312	0.09120	0.09766	0.09465	0.09263	0.09605	0.09783	0.09653	0.09526
	1,000	0.09523	0.09206	0.09501	0.09656	0.09284	0.09700	0.09524	0.10481	0.09673	0.10520
	1,200	0.09674	0.09206	0.09889	0.10063	0.09705	0.10000	0.09979	0.10223	0.09468	0.10360
	1,500	0.09654	0.09300	0.09428	0.09631	0.10334	0.09881	0.10090	0.09931	0.10015	0.09294

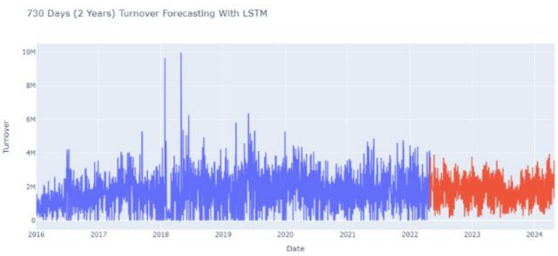
The data for the second scenario has been derived from the best results of the first scenario. Figures 3 and 4 show the result of the second scenario, when RNN and LSTM forecast in various time periods. The blue color on the graph represents actual data, while the red color on the graph represents the results of predictions. Figure 3 shows that the LSTM model can forecast future data and depict upward and downward trends over a predetermined time period.



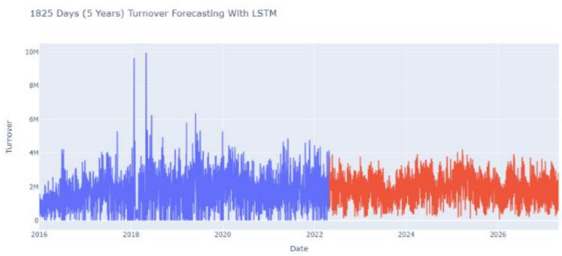
(a)



(b)



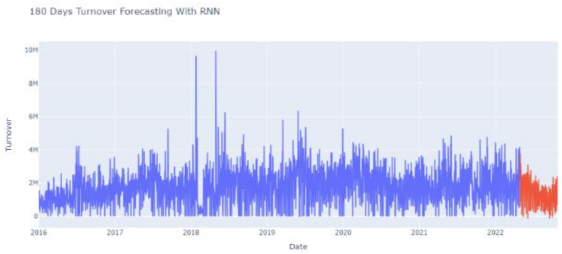
(c)



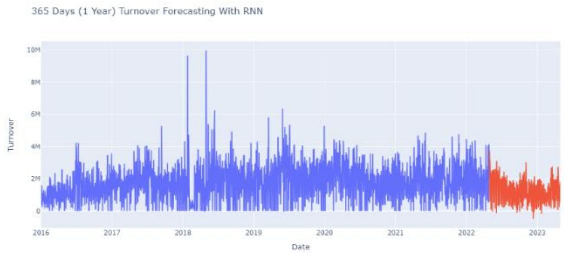
(d)

Figure 3. LSTM forecasting results in (a) 180 days; (b) 1 year; (c) 2 years; and (d) 5 years turnover forecasting with LSTM.

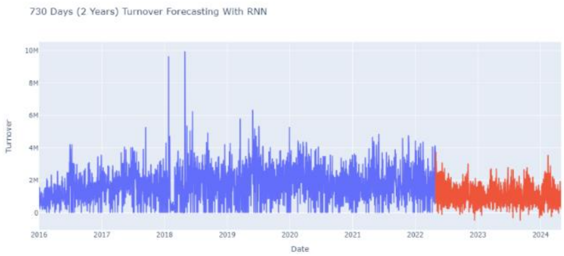
The results show that RNN models consistently exhibit a considerable degree of forecasting changes within a specific period, as shown in Figure 4. The image consists of actual data contained in the blue graph color and predicted data represented in the red graph color. The results of this scenario show good and stable predictive capabilities in the movement of upward and downward trends in the data over a predetermined period. This proves that RNN is a model that can be used as a model to make predictions regarding changes in certain trends over a specified period.



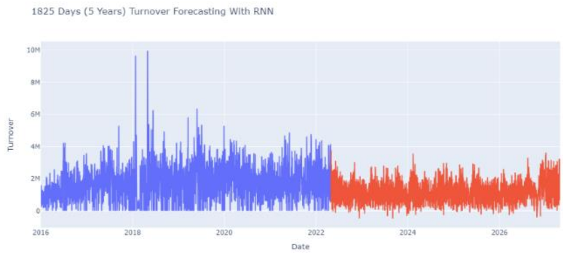
(a)



(b)



(c)



(d)

Figure 4. RNN Forecasting Results (a) 180 days; (b) 1 year; (c) 2 years; and (d) 5 years turnover forecasting with RNN.

4. CONCLUSION

The goal of the study is to evaluate and compare the effectiveness of LSTM and RNN methods to forecast the daily turnover of BM Motor Ngawi. The study encompassed optimizing hyperparameters and running many tests, which indicated that the LSTM model exceeded the RNN model. The LSTM model, trained with a 90:10 data splitting ratio, a 0.15 dropout rate, and 800 epochs, exhibited superior performance compared to the RNN model. Despite both using similar parameters, the LSTM model had a lower MAE score of 0.08753 compared to the RNN model's MAE score of 0.09233, proving that the LSTM model is more accurate in making decisions. The findings of the study suggest that using LSTM instead of RNN can help the store make better sales strategies. In conclusion, the research supports the effectiveness of the LSTM model in predicting the daily turnover of BM Motor Ngawi.

REFERENCES

[1] Wiranda L, Sadikin M. Penerapan Long Short Term Memory Pada Data Time Series untuk Memprediksi Penjualan Produk PT. Metiska Farma. vol. 8. n.d. https://doi.org/https://doi.org/10.23887/janapati.v8i3.19139.

[2] Sugiyarto AW, Abadi AM. Prediction of Indonesian Palm Oil Production Using Long Short-Term Memory Recurrent Neural Network (LSTM-RNN). 2019 1st International Conference on Artificial Intelligence and Data Sciences (AiDAS), IEEE; 2019, p. 53–7. https://doi.org/10.1109/AiDAS47888.2019.8970735.

[3] Widi Hastomo, Sutarno, Sudjiran. Analisis Risiko Investasi dan Prediksi Saham Menggunakan Algortime Machine Learning. Jurnal Ilmiah Komputasi 2022;21:453–62. https://doi.org/10.32409/jikstik.21.3.3104.

[4] Sagheer A, Kotb M. Time series forecasting of petroleum production using deep LSTM recurrent networks. Neurocomputing 2019;323:203–13. https://doi.org/10.1016/j.neucom.2018.09.082.

[5] Selle N, Yudistira N, Dewi C. Perbandingan Prediksi Penggunaan Listrik dengan Menggunakan Metode Long Short Term Memory (LSTM) dan Recurrent Neural Network (RNN). Jurnal Teknologi Informasi Dan Ilmu Komputer 2022;9:155. https://doi.org/10.25126/jtiik.2022915585.

[6] Rafi SH, Nahid-Al-Masood, Deeba SR, Hossain E. A Short-Term Load Forecasting Method Using Integrated CNN and LSTM Network. IEEE Access 2021;9:32436–48. https://doi.org/10.1109/ACCESS.2021.3060654.

[7] DiPietro R, Hager GD. Deep learning: RNNs and LSTM. Handbook of Medical Image Computing and Computer Assisted Intervention, Elsevier; 2020, p. 503–19. https://doi.org/10.1016/B978-0-12-816176-0.00026-0.

[8] Zhu R, Tu X, Xiangji Huang J. Deep learning on information retrieval and its applications. Deep Learning for Data Analytics, Elsevier; 2020, p. 125–53. https://doi.org/10.1016/B978-0-12-819764-6.00008-9.

[9] Sagheer A, Kotb M. Time series forecasting of petroleum production using deep LSTM recurrent networks. Neurocomputing 2019;323:203–13. https://doi.org/10.1016/j.neucom.2018.09.082.

[10] Siami-Namini S, Tavakoli N, Namin AS. The Performance of LSTM and BiLSTM in Forecasting Time Series. 2019 IEEE International Conference on Big Data (Big Data), IEEE; 2019, p. 3285–92. https://doi.org/10.1109/BigData47090.2019.9005997.

[11] Kim S, Kang M. Financial series prediction using Attention LSTM 2019.

[12] Abdel-Nasser M, Mahmoud K. Accurate photovoltaic power forecasting models using deep LSTM-RNN. Neural Comput Appl 2019;31:2727–40. https://doi.org/10.1007/s00521-017-3225-z.

[13] Abbasimehr H, Paki R. Improving time series forecasting using LSTM and attention models. J Ambient Intell Humaniz Comput 2022;13:673–91. https://doi.org/10.1007/s12652-020-02761-x.

[14] Shah AA, Ahmed K, Han X, Saleem A. A Novel Prediction Error-Based Power Forecasting Scheme for Real PV System Using PVUSA Model: A Grey Box-Based Neural Network Approach. IEEE Access 2021;9:87196–206. https://doi.org/10.1109/ACCESS.2021.3088906.

[15] Tian C, Ma J, Zhang C, Zhan P. A deep neural network model for short-term load forecast based on long short-term memory network and convolutional neural network. Energies (Basel) 2018;11. https://doi.org/10.3390/en11123493.

BIBLIOGRAPHY OF AUTHORS

Larasati is currently in her final year of undergraduate studies, majoring in the School of Computing at Telkom University. She enrolled in the Kampus Merdeka program, which has sparked her interest in deep learning and data forecasting and she has been provided with unique opportunities to enhance her knowledge and skills in these specialized fields. As she approaches the culmination of her undergraduate education, She is poised to contribute significantly to the advancement of deep learning and data forecasting.

Siti Saadah is a highly accomplished individual, currently a lecturer and researcher at Telkom University. Her educational background includes a Bachelor's degree in Information Engineering (2004-2009) and a Master's degree in Informatics from Telkom Institute of Technology (2010-2012) with expertise in data mining, artificial intelligence, machine learning, and financial-economic computing, making her an invaluable asset to the institution. Her dedication and passion for her work are evident through her main interests, which revolve around data-based presentations and leading teams. She strives to maximize individuals' potential and actively contributes to society. Saadah's notable background and extensive education make her exceptional in her field.

Prasti Eko Yunanto, a lecturer and researcher in Intelligent Systems, possesses notable teaching experience and expertise in the field. His educational background includes a Bachelor's degree in Information Engineering from the Telkom Institute of Technology (2007-2012) and a Master's degree in Informatics from Telkom University (2013-2015). With his strong academic foundation, he actively engages in teaching a range of subjects in the computer science field, including Algorithm and Programming, Data Structure, Language Theory and Automata, and Artificial Intelligence. Through his comprehensive knowledge and dedication, he enriches the educational experience of students and equips them with valuable skills to excel in their respective careers.