Indonesian Journal of Electrical and Electronics Engineering (INAJEEE), Vol 7, No 2, 2024, page 76-82

Prediction of Transformer Aging Loss in 150kV Waru Substation Using GRU-LSTM Method Based on Temperature and Load

Ahmad Mustafa[1*], Lusia Rakhmawati[2]
[1,2,3,4] Electrical Engineering Department, Universitas Negeri Surabaya
A5 Building Ketintang Campus, Surabaya 60231, Indonesia
[1*]ahmad.20001@mhs.unesa.ac.id,
[2] lusiarakhmawati@unesa.ac.id

Abstract – The increasing demand for electricity in Indonesia highlights transformers' crucial role in the electrical system. This study utilizes GRU (Gated Recurrent Unit) and LSTM (Long Short-Term Memory) in a deep learning framework to predict aging losses of transformer unit 5 at the 150 kV Waru Substation. The aim is to enhance grid reliability, and efficiency, and prevent disruptions like power outages. Conducted at the 150 kV Waru Substation, the research focuses on transformer loading and temperature data. Data preprocessing involves normalizing load, oil temperature, and winding temperature data. The model architecture combines GRU and LSTM to capture short-term and long-term patterns in time series data. Training employs the Adam optimizer with customized learning rates, and performance evaluation uses metrics such as MSE, RMSE, MAPE, and MAE. Results indicate the GRU-LSTM model trained with a batch size of 64 and 75 epochs achieves superior performance: MSE of 0.0000129008474202634, RMSE of 0.00359177496793207, MAPE of 0.3943965%, and MAE of 0.00000832911556912471. This model forecasts transformer 5's aging loss over the next 30 days with an average daily deterioration rate of 0.001378178 pu/day, peaking at 0.0030481415 pu.

Keywords: Aging Loss Prediction, GRU, LSTM, Transformer

I. INTRODUCTION

Electricity has become a crucial component for the future of the country. The demand for electricity in Indonesia continues to increase in line with the growth of industry and population [1]. PT. PLN (Persero), through the 150 kV Waru Substation, plays a role in meeting the electrical energy needs for industry and residential areas in Surabaya and Sidoarjo. One of the vital pieces of equipment in the operation of the substation is the transformer. The 150 kV Waru Substation has five power transformer units with varying workloads. A transformer is an electrical device that has a magnetic circuit and windings with two or more coils [2]. A transformer converts electrical power (current and voltage) from an alternating current (AC) system to different current and voltage levels with the same frequency through the process of electromagnetic induction [3]. According to the IEC (International Electrotechnical Commission), a transformer will have a normal lifespan of 30 years with a continuous load of 100% and a hotspot temperature of 98°C. When a transformer is loaded to its maximum capacity (100%) of its total power capacity, it will have a lifespan of 20 years at an ambient temperature of 20°C [4].

However, based on field data from the 150 kV Waru Substation, one of the transformer units has been in operation since 1998, which results in a longer operational time compared to other transformers, potentially causing disturbances. These disturbances occur due to several factors, such as short circuit faults, increased load, chemical mechanisms, insulation degradation, and other elements. These disturbances lead to a reduction in the transformer's lifespan, ultimately affecting its performance within the operational limit [5]. If Transformer 5 experiences a disturbance, it will impact several 20 kV feeders supplied by Transformer 5.

Maintenance of transformers at the 150 kV Waru Substation has been carried out by PT. PLN (Persero) conducts routine testing every two years using insulation resistance testing and load testing methods to evaluate the condition and performance of the transformers. The maintenance approach in use is proving less efficient due to various factors affecting the transformer system, including higher loads, insulation degradation, and other elements that introduce disruptions and diminish the transformer's longevity. Moreover, this method cannot accurately predict the transformer's remaining lifespan. Over time, transformers will experience a decline in efficiency, thus limiting their long-term performance.

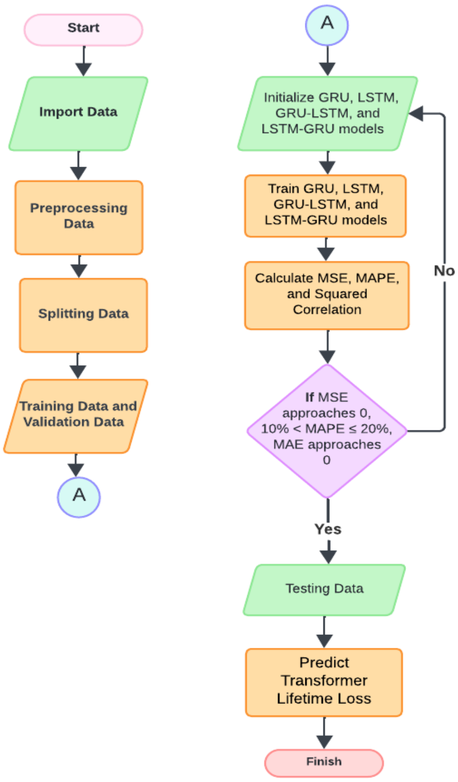
Therefore, the weaknesses of the periodic testing maintenance method necessitate a more effective solution by leveraging current technological advancements. One such technological application is the use of machine learning. Machine learning, a subset of artificial intelligence, is centered on creating algorithms and techniques that allow computers to learn from data and make predictions or decisions accordingly [6]. Several machine learning methods have been previously employed to predict transformer lifespan reduction, such as Decision Tree, Random Forest, and Linear Regression [7]. The

Linear Regression method conducted by R.S Juwita and L. Liliana [8] for forecasting the lifespan reduction of a 60 MVA transformer used transformer load data. However, the linear regression machine learning method is less suitable for predicting transformer lifespan reduction because the relationship between input variables such as temperature, load, and operating time with the remaining life of the transformer is often non-linear [9]. Linear regression tends to assume a simple linear relationship between these variables, whereas the factors influencing transformer lifespan can be much more complex and non-linear.

Instead, deep learning methods can provide a more effective solution. With deep learning, we can leverage its capability to model complex, non-linear relationships that may occur in the influence of variables on transformer lifespan reduction [10]. The use of Deep Learning methods for predicting transformer lifespan reduction has been previously explored. Research by A. Novian [11], utilized Long Short Term Memory (LSTM), to estimate transformer lifespan reduction over one year using temperature and load as parameters. Additionally, Worku Abebe Degife and Bor-Shen Lin [12], discovered that the GRU model surpassed traditional machine learning models such as MLP and LSTM in evaluation metrics, including mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R2).In the study, it was concluded that GRU is capable of capturing complex and non-linear relationships among factors, thus achieving high prediction accuracy. Furthermore, in the research by W. Yang et al. [13] they utilized the GRU method for forecasting electric load during wildfire seasons, where the GRU model outperformed the LSTM model.

Furthermore, the GRU method was also used in the study by M. Abumohsen et al. [14], researchers utilized the GRU method to forecast electric load using data from the SCADA system at the Tubas Electricity Company in Palestine. Their comparison of three deep learning algorithms (LSTM, GRU, and RNN) revealed that the GRU model outperformed the others in terms of accuracy and error reduction, demonstrating its effectiveness in load prediction. GRU has simpler computational complexity and requires less memory compared to LSTM. Moreover, GRU and LSTM each have their own advantages in sequential data processing. GRU is more efficient in information processing due to its simpler structure, whereas LSTM is effective in handling long-term dependencies in information [15].

Therefore, this study proposes the GRU-LSTM method for predicting the lifespan reduction of Transformer Unit 5 at the 150 kV Waru Substation. Utilizing GRU-LSTM is expected to achieve better performance in enhancing prediction accuracy and overall model performance compared to using each method independently. This study also compares four models: GRU, LSTM, GRU-LSTM, and LSTM-GRU. Additionally, the proposed GRU-LSTM method can contribute to maintaining the reliability and efficiency of the electrical system and preventing disruptions such as power outages and other electrical system failures.



Algorithm 1. Pseudocode GRU-LSTM Model

```
1. Initialize the Sequential model:
   model = Sequential()

2. Add the GRU layer:
   - Number of units: 64
   - Return only the output for the last
timestep (return_sequences=False)
   model.add(GRU(64,
return_sequences=False))

3. Add the LSTM layer:
   - Number of units: 64
   - Input shape: (None, 3)
   - Return the full sequence of outputs
(return_sequences=True)
   model.add(LSTM(64, input_shape=(None,
3), return_sequences=True))

4. Add the Dense layer:
   - Number of neurons: 1
   model.add(Dense(1))

5. Set the learning rate for the optimizer:
   - Learning rate: 0.001
   optimizer = Adam(learning_rate=0.001)

6. Compile the model with the optimizer and
loss function:
   - Optimizer: Adam with the custom
learning rate
   - Loss function: Mean Squared Error
(MSE)
   model.compile(optimizer=optimizer,
loss='mse')
```

Figure 1. Flowchart and listing program of GRU-LSTM Modeling

II. METHODS

Datasets

The data collection was carried out at the 150 kV Waru Substation, which includes daily inspection data and daily transformer load recordings that are recorded every day. Transformer load data, winding temperature, and oil temperature are the input data. Meanwhile, the target data is the transformer aging loss. Subsequently, the data was normalized and divided into training data, testing data, and validation data. The purpose of this data division is to ensure that the GRU-LSTM model created is capable of making accurate predictions on unseen data. In this study, the transformer load and temperature are considered constant every hour throughout a day, and the forecasting is done for daily periods within each month, specifically predicting lifespan reduction in January 2024.

GRU-LSTM Model

The proposed model in this study uses deep learning with the GRU-LSTM model, as well as LSTM-GRU, GRU, and LSTM as comparison algorithms. The steps to integrate the model are shown in Figure 1.

1. Data Import
   In this section, daily temperature and load data, formatted into CSV, will be imported using the pandas library.
2. Data Preprocessing
   The data is normalized using min-max scaling, a simple technique to rescale data within a predefined range. Additionally, at this stage, input data values are transformed to range between 0 and 1 to facilitate deep learning model compatibility.
3. Data Splitting
   In this research, the dataset is segmented into two main parts: the training dataset and the testing dataset. The training dataset is employed to train the model and determine optimal weights, while the testing dataset is used to assess the model's performance. Typically, the split ratio between training and testing datasets ranges from 80% for training and 20% for testing purposes.
4. Model Initialization
   Initialization of GRU, LSTM, GRU-LSTM, and LSTM-GRU models, including hyperparameter tuning such as neurons, verbose, epochs, and batch size. The architecture of the GRU for this research is illustrated in Figure 2, while the LSTM architecture is depicted in Figure 3.
5. Training Data
   After initializing the GRU and LSTM models, the training of the GRU, LSTM, GRU-LSTM, and LSTM-GRU models is performed. During model training, the data used comprises both training and validation data.
6. Model Evaluation
   Model performance is evaluated using Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE), and Mean Absolute Error (MAE). If the results are deemed poor or unsatisfactory, the model initialization for GRU-LSTM, LSTM-GRU, GRU, and LSTM is revisited, and hyperparameters are readjusted. If the results are satisfactory, the process proceeds to forecasting using test data with the developed model.

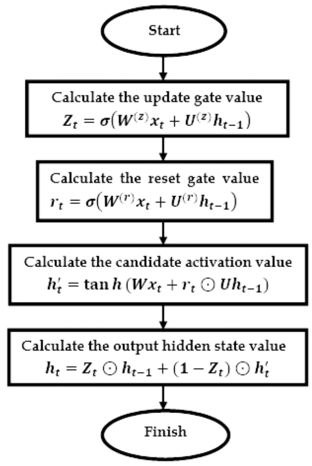


Figure 2. Flowchart of GRU Architecture

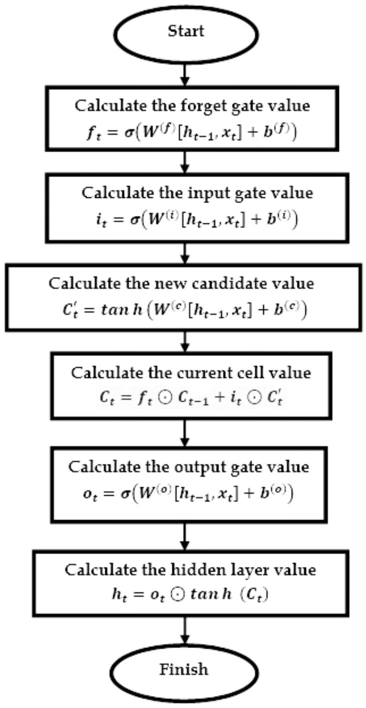


Figure 3. Flowchart of LSTM Architecture

7. Model Testing
   The testing of the models involves feeding the preprocessed and divided test data into the previously constructed GRU-LSTM, LSTM-GRU, GRU, and LSTM models to evaluate their performance. The testing results provide predictions of transformer lifespan reduction based on the input data, which includes transformer temperature and load.

III. RESULT AND DISCUSSION

This research involved varying hyperparameters like the optimizer (ADAM and RMSprop) and adjusting Learning Rates to 0.001 and 0.01. It also explored different batch sizes (32, 64, 128) and epochs (50, 75, 100). In various studies, the ADAM optimizer has been found to perform optimally under specific hyperparameter settings [16]-[17]. Optimizers such as Adam, which adaptively adjust the learning rate during training, are often used due to their higher stability and efficiency [18]. Additionally, Based on the research conducted by Job Dip Das et al., it has been found that utilizing hyperparameter Batch Sizes (BS) between 16 and 64 yields superior performance for GRU-LSTM models [19]. On the other hand, hyperparameter settings for the Learning Rate (LR) should ideally start with a smaller value, such as 0.001 or 0.01. A Learning Rate that is too large can cause the model to not converge or to overfit [20]. A lower Root Mean Squared Error (RMSE) indicates better prediction accuracy, reflecting reduced prediction errors and improved model precision in estimating actual values [21]. Besides, Research by Eliana Vivas et al.[22], suggests that a good MAPE (Mean Absolute Percentage Error) value is typically lower. A lower MAPE indicates that the model's predictions are more precise and closely match the actual values, reflecting higher accuracy in forecasting or estimation tasks. Next, a comparison of the training and testing results of the four models including GRU, LSTM, GRU-LSTM, and LSTM-GRU will be presented.

Table 1. Comparison of Model Training Results

Hyperparameter and Evaluation	**Models**			
	GRU	**LSTM**	**GRU-LSTM**	**LSTM-GRU**
Optimizer	Adam	Adam	Adam	Adam
Learning Rate	0,01	0,01	0,01	0,01
Batch Size	32	32	64	64
Epoch	100	100	75	100
MSE	0,001026	0,0000226	0,000046	0,000045
RMSE	0,032024	0,015035	0,006772	0,006728
MAE	0,000075	0,000061	0,000014	0,000013
MAPE (%)	4,071297	3,637880	0,743127	0,667350

Table 2. Comparison of Model Testing Results

Hyperparameter and Evaluation	**Models**			
	GRU	**LSTM**	**GRU-LSTM**	**LSTM-GRU**
Optimizer	Adam	Adam	Adam	Adam
Learning Rate	0,01	0,01	0,01	0,01
Batch Size	32	32	64	64
Epoch	100	100	75	100
MSE	0,001209	0,000503	0,000013	0,000013
RMSE	0,034766	0,022419	0,003592	0,003582
MAE	0,000054	0,000055	0,000008	0,000009
MAPE (%)	2,780356	2,884454	0,394397	0,466057

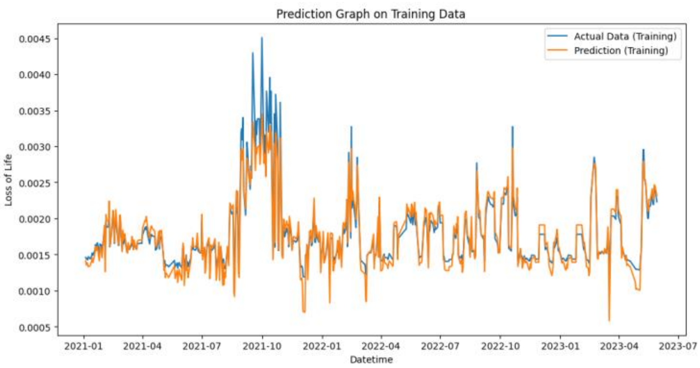


Figure 4. Graph of Training for GRU Model

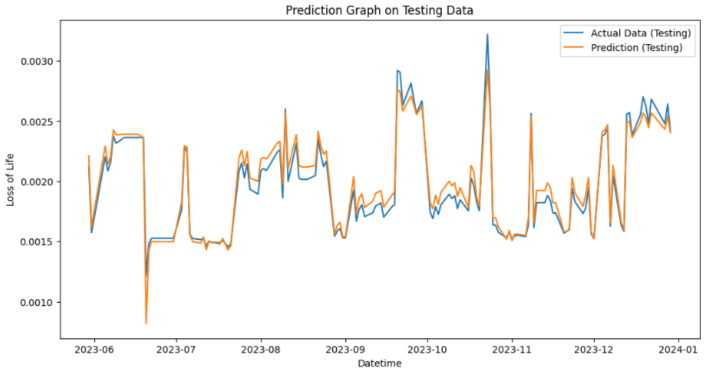


Figure 5. Graph of Testing for GRU Model

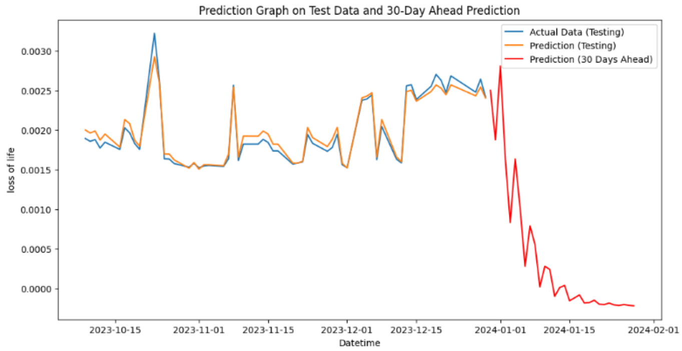


Figure 6. Graph of 30-Day Prediction using GRU Model

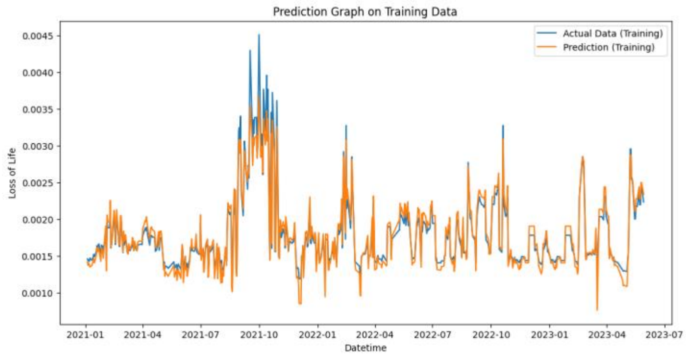


Figure 7. Graph of Training for LSTM Model

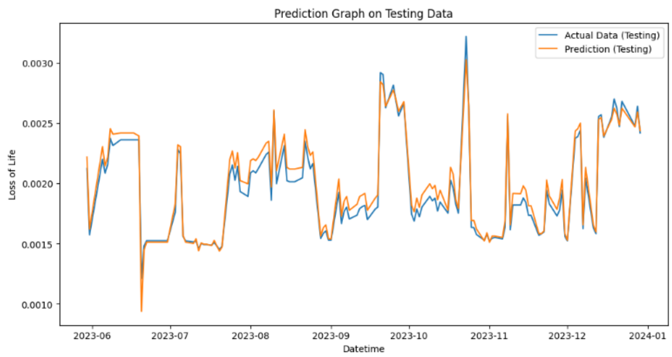


Figure 8. Graph of Testing for LSTM Model

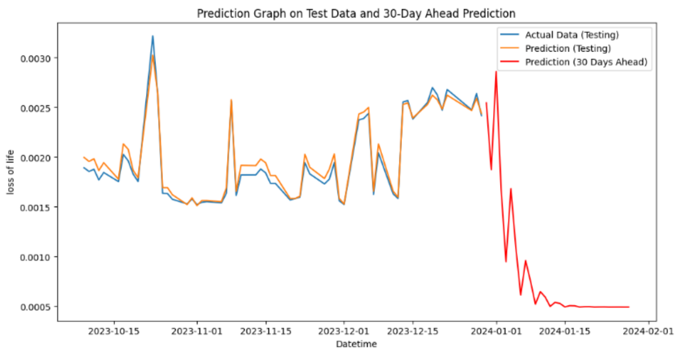


Figure 9. Graph of 30-Day Prediction using LSTM Model

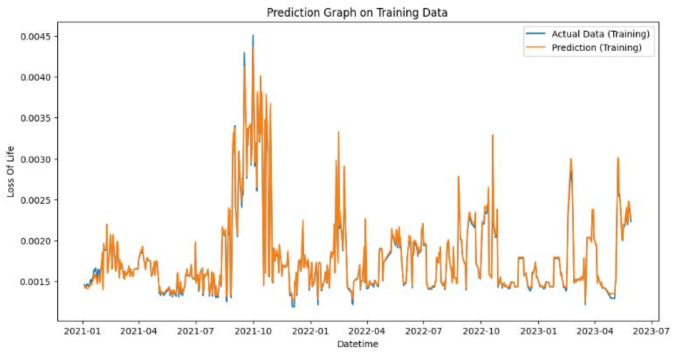


Figure 10. Graph of Training for GRU-LSTM Model

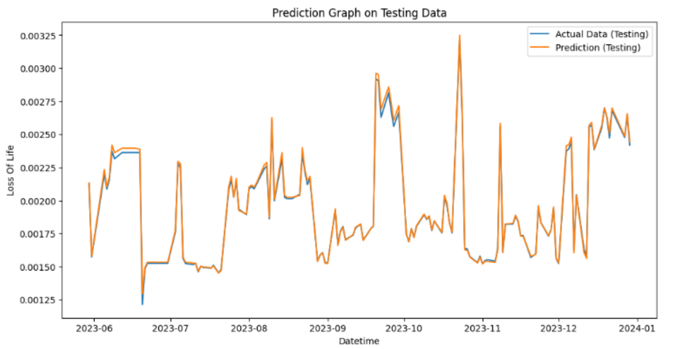


Figure 11. Graph of Testing for GRU-LSTM Model

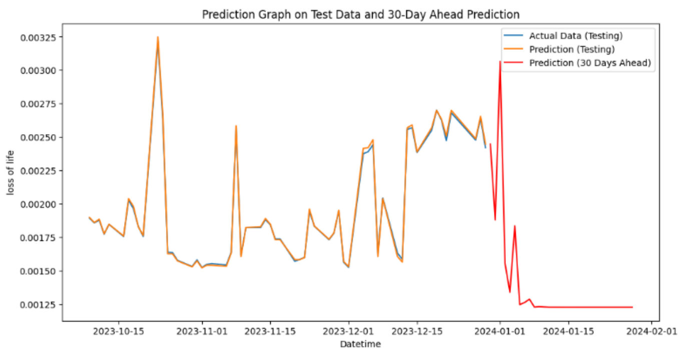


Figure 12. Graph of 30-Day Prediction using GRU-LSTM Model

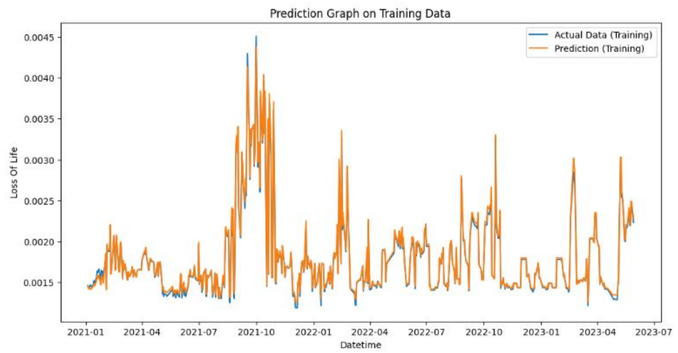


Figure 13. Graph of Training for LSTM-GRU Model

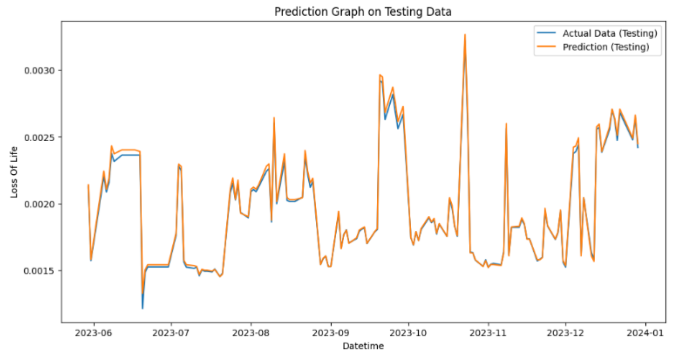


Figure 14. Graph of Testing for LSTM-GRU Model

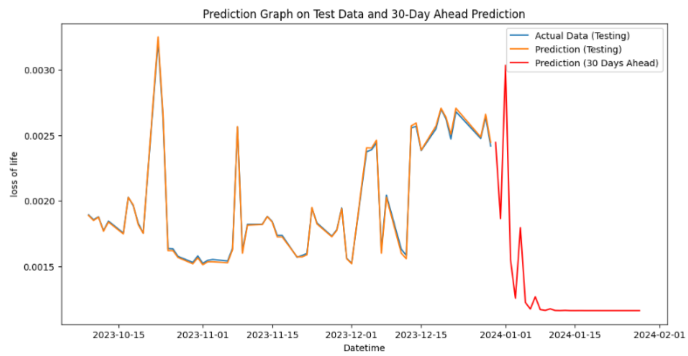


Figure 15. Graph of 30-Day Prediction using LSTM-GRU Model

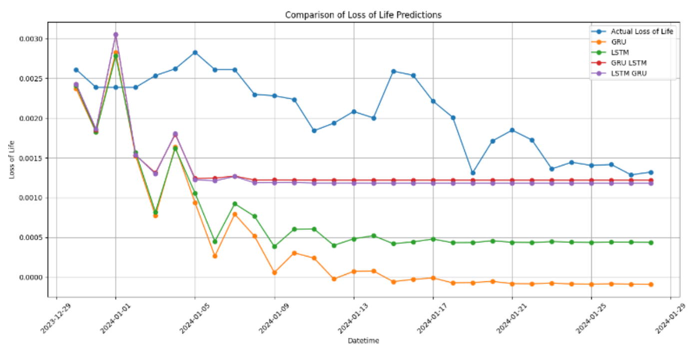


Figure 16. Comparison of 4 Models

Based on **Table 1** and **Table** 2, the LSTM-GRU model shows the best performance during training with the lowest MSE, RMSE, MAE, and MAPE values, whereas, during testing, the GRU-LSTM model performs best with the lowest MSE, RMSE, MAE, and MAPE values. Additionally, based on **Figure 6** the graph shows actual testing data and predictions for the next 30 days for GRU models. The red line (30-day prediction) shows a sharp decline immediately after the prediction period starts around January 1, 2024. The average daily transformer aging loss predicted for the next 30 days is approximately 0.000442297 pu/day. The highest predicted daily aging loss is 0.002826909 pu on January 1, 2024. Next is based on **Figure 9** the graph shows actual testing data and predictions for the next 30 days for LSTM models. Sharp decline in daily transformer aging loss, which then flattens to 0.0005 pu/day after January 15, 2024. The predicted average daily transformer aging loss for the next 30 days is approximately 0.000781039 pu/day. The highest predicted daily aging loss is 0.0027794456 pu on January 1, 2024. Additionally, The Graph of 30-Day Prediction using the GRU-LSTM Model is shown in **Figure 12,** in the graph, it is evident that the prediction (red line) shows a spike around January 1st, and this value is the highest throughout the 30-day prediction period. After January 1st, the graph shows a decrease and stabilization until the end of the prediction period (around January 10th), with values approaching 0.0012 pu/day. From the graph, the highest spike on January 1st appears to be around 0.003 pu, which is consistent with the description. After this spike, the values stabilize around 0.0012 pu/day. **Figure 15**, The prediction pattern for the next 30 days using the LSTM-GRU model is almost identical to the GRU-LSTM model. The predicted average daily transformer aging loss for the next 30 days by the LSTM-GRU model is approximately 0.001351554 pu/day, with the highest predicted daily aging loss being 0.0030481871 pu on January 1, 2024. Additionally, **Figure 16** shows a comparison of transformer aging loss predictions over a specific period using four different methods: GRU, LSTM, GRU-LSTM, and LSTM-GRU. GRU-LSTM and LSTM-GRU demonstrate higher stability with consistent aging loss values after the initial period. GRU-LSTM achieved an MSE of 0.000849473, an RMSE of 0.029145723, a MAPE of 32.52%, and an MAE of 0.000726526.

IV. CONCLUSION

In this research, various experiments were carried out to fine-tune hyperparameters such as the choice of optimizer (ADAM optimizer, RMSprop) and adjusting Learning Rates to 0.001 and 0.01. Additionally, different combinations of batch sizes (32, 64, 128) and epochs (50, 75, 100) were tested. The study also compared the performance of four models: GRU, LSTM, GRU-LSTM, and LSTM-GRU. The best results were obtained using the GRU-LSTM model with Adam Optimizer, batch size of 64, and 75 epochs. The GRU-LSTM model successfully predicted the aging of Unit 5 transformer at the 150 kV Waru Substation for the next 30 days with an average rate of 0.001378178 pu/day. The highest predicted aging rate was 0.0030481415 pu on January 1, 2024.

In future research, data collection can be conducted in more detail by recording hourly loads to increase the dataset size, enabling models to predict further ahead. Additionally, modifying the model architecture for comparison with previously tested architectures is also recommended.

ACKNOWLEDGMENT

Thank you to Dr. Lusia Rakhmawati, S.T., M.T., my supervisor, for guiding this research to completion. Also, I extend my gratitude to the 150 kV Waru Substation for allowing me to collect data necessary for this research.

REFERENCES

[1] D. Hartono, S. H. Hastuti, A. A. Balya, and W. Pramono, "Modern energy consumption in Indonesia: Assessment for accessibility and affordability," *Energy Sustain. Dev.*, vol. 57, pp. 57–68, 2020.

[2] L. I. Sakhno, O. I. Sakhno, and M. S. Krylov, "Equivalent Circuit of the Transformer with Two Primary and One Secondary Windings," *Proc. 2020 IEEE Conf. Russ. Young Res. Electr. Electron. Eng. EIConRus 2020*, vol. 11, no. 2, pp. 1297–1300, 2020.

[3] SPLN, *Buku Pedoman Pemeliharaan Transformator Tenaga*. Jakarta, 2014.

[4] F. Azhar, Y. Rahmawati, and I. Fadlika, "Linear Trend Application on the Estimated Age of Distribution Transformer Based on the Load Growth and Environmental Temperature," *Frontier Energy System and Power Engineering*, vol. 1, no. 2. p. 7, 2019.

[5] K. Ibrahim, R. M. Sharkawy, H. K. Temraz, and M. M. A. Salama, "Reliability calculations based on an enhanced transformer life expectancy model," *Ain Shams Eng. J.*, vol. 13, no. 4, p. 101661, 2022.

[6] S. Raschka, J. Patterson, and C. Nolet, "Machine learning in python: Main developments and technology trends in data science, machine learning, and artificial intelligence," *Inf.*, vol. 11, no. 4, 2020.

[7] A. Alqudsi and A. El-Hag, "Application of machine learning in transformer health index prediction," *Energies*, vol. 12, no. 14, pp. 1–13, 2019.

[8] D. A. Trianggana and P. Pendekatan, "Metode Regresi Linear," *J. Media Infotama*, 2020.

[9] G. Turabee *et al.*, "The Role of Neural Networks in Predicting the Thermal Life of Electrical Machines," *IEEE Access*, vol. 8, pp. 40283–40297, 2020.

[10] P. R. de O. da Costa, A. Akçay, Y. Zhang, and U. Kaymak, "Remaining useful lifetime prediction via deep domain adaptation," *Reliab. Eng. Syst. Saf.*, vol. 195, no. August 2019, p. 106682, 2020.

[11] A. Novian and U. T. Kartini, "Peramalan Loss Of Life Tranformator Berdasarkan Loading dan Temperature Menggunakan LSTM di Gardu Induk 150KV Buduran," *J. Tek. Elektro*, vol. 12, no. 1, pp. 81–86, 2023.

[12] W. A. Degife and B. S. Lin, "Deep-Learning-Powered GRU Model for Flight Ticket Fare Forecasting," *Appl. Sci.*, vol. 13, no. 10, 2023.

[13] W. Yang, S. N. Sparrow, and D. C. H. Wallom, "A generalised multi-factor deep learning electricity load forecasting model for wildfire-prone areas," 2023.

[14] M. Abumohsen, A. Y. Owda, and M. Owda, "Electrical Load Forecasting Using LSTM, GRU, and RNN Algorithms," *Energies*, vol. 16, no. 5, pp. 1–31, 2023.

[15] Y. Touzani and K. Douzi, "An LSTM and GRU based trading strategy adapted to the Moroccan market," *J. Big Data*, vol. 8, no. 1, 2021.

[16] S. Y. Sen and N. Ozkurt, "Convolutional Neural Network Hyperparameter Tuning with Adam Optimizer for ECG Classification," *Proc. - 2020 Innov. Intell. Syst. Appl. Conf. ASYU 2020*, no. 978, pp. 0–5, 2020.

[17] A. A. Pierre, S. A. Akim, A. K. Semenyo, and B. Babiga, "Peak Electrical Energy Consumption Prediction by ARIMA, LSTM, GRU, ARIMA-LSTM and ARIMA-GRU Approaches," *Energies*, vol. 16, no. 12, 2023.

[18] D. Yi, J. Ahn, and S. Ji, "An effective optimization method for machine learning based on ADAM," *Appl. Sci.*, vol. 10, no. 3, 2020.

[19] J. Dip Das, R. K. Thulasiram, C. Henry, and A. Thavaneswaran, "Encoder–Decoder Based LSTM and GRU Architectures for Stocks and Cryptocurrency Prediction," *J. Risk Financ. Manag.*, vol. 17, no. 5, 2024.

[20] Y. Jin *et al.*, "Autolrs: Automatic Learning-Rate Schedule By Bayesian Optimization on the Fly," *ICLR 2021 - 9th Int. Conf. Learn. Represent.*, pp. 1–22, 2021.

[21] H. Tian, P. Wang, K. Tansey, J. Zhang, S. Zhang, and H. Li, "An LSTM neural network for improving wheat yield estimates by integrating remote sensing data and meteorological data in the Guanzhong Plain, PR China," *Agric. For. Meteorol.*, vol. 310, no. 17.

[22] E. Vivas, H. Allende-Cid, and R. Salas, "A systematic review of statistical and machine learning methods for electrical power forecasting with reported mape score," *Entropy*, vol. 22, no. 12, pp. 1–24, 2020.