Prediction of Lightning Strikes Electrical Transmission Lines Using Machine Learning Approaches

Aditya Adiaksa[1], Ahmad Mushawir[1],Zainuddin[1],Akhmad Faeda Insani[1], and Sparisoma Firidi[1,2]

[1]Master of Science in Computational Science Program, Faculty of Mathematics and Natural Sciences, Bandung Institute of Technology,
Jalan Ganesa 10, Bandung 40132, Indonesia,
[2]Faculty of Mathematics and Natural Sciences, Bandung Institute of Technology,
Jalan Ganesa 10, Bandung 40132, Indonesia,
Email: adityaadiaksa14@gmail.com

Abstract. This study evaluates and compares the performance of three Machine Learning models—RNN, LSTM, and Transformer—for forecasting lightning strikes that can disrupt transmission towers along the 150 kV Bukit Asam - Baturaja line. Using historical lightning data from 2018 to 2024, the models were trained and validated, with the Transformer model demonstrating superior predictive accuracy.The Transformer model achieved an impressive R-squared value of 0.9543, significantly surpassing the performance of both the Recurrent Neural Network and the Long Short-Term Memory models. The Transformer model is a dependable option for predictions, as its self-attention mechanism efficiently identifies dependency patterns and trends. However, the study is limited to this specific region and dataset, highlighting the need for future research to incorporate additional variables, such as meteorological and geographical factors, for improved adaptability. The findings underscore the importance of accurate and efficient forecasting models to support proactive measures and mitigate lightning-induced disturbances on transmission infrastructure.chnology for risk mitigation in electrical transmission networks.

Keywords: *Transformer model, RNN, LSTM, lightning strike forecasting, transmission tower, time-series prediction*

1 Introduction

Within the Sumatera region of PT PLN UIP3B, lightning disturbances pose a significant threat to the electrical system. These disturbances not only disrupt the smooth distribution of power but also cause various hazards and substantial losses. Lightning disturbances often result in power outages, halting industrial, commercial, and household activities. This leads to significant economic losses and disrupts the comfort of the community. Improving the reliability of electrical transmission is thus a top priority for PT PLN UIP3B Sumatera, aiming to ensure a stable and reliable electricity supply for all customers.

Received ________, Revised ________, Accepted for publication __________

PLN customers are entitled to quality and reliable electricity services. Recurring lightning disturbances can reduce customer satisfaction and damage the company's reputation. By enhancing the reliability of electrical transmission through lightning disturbance mitigation, PT PLN UIP3B Sumatera can provide better service to its customers. Additionally, lightning disturbances not only cause power outages but also lead to damage to transmission network infrastructure, such as towers, cables, and transformers. This damage incurs significant repair costs and downtime, disrupting the operational efficiency of the electrical system. Furthermore, lightning strikes pose a danger to human life, affecting both PLN personnel and the general public.

Efforts to mitigate lightning disturbances have traditionally focused on physical solutions, such as installing lightning arresters and grounding systems. These methods aim to protect the network infrastructure from direct lightning strikes and to dissipate the resulting current safely into the ground. While these measures have proven effective to a certain extent, they are primarily reactive and limited in their ability to predict and proactively address lightning risks. Previous studies have emphasized the importance of understanding lightning patterns and implementing predictive strategies to mitigate potential damage [19]. However, the adoption of predictive approaches, particularly those leveraging advanced machine learning techniques, remains limited in this context.

Recent advancements in machine learning offer promising solutions for predicting lightning strikes and mitigating their impact. Models such as Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks have been widely applied to time-series forecasting tasks. However, their ability to capture long-term dependencies is constrained by their sequential nature and vanishing gradient issues [3]. The Transformer model, which leverages a self-attention mechanism to process entire sequences concurrently, has demonstrated superior performance in a variety of domains, including natural language processing and energy demand forecasting. Despite its potential, the application of the Transformer model in lightning strike prediction remains underexplored.

This study aims to address this gap by evaluating and comparing the performance of RNN, LSTM, and Transformer models in predicting lightning strikes that could disrupt electrical transmission towers along the 150 kV Bukit Asam-Baturaja line. Using historical lightning data from 2018 to 2024, this research highlights the capabilities of the Transformer model in identifying dependency patterns and trends more effectively than traditional methods. The findings contribute to the development of predictive strategies that not only enhance the reliability of electrical transmission but also support proactive mitigation measures, ensuring the safety and stability of the network.

Through this work, we aim to bridge the gap between conventional lightning mitigation strategies and advanced predictive methodologies, demonstrating how machine learning can be effectively utilized to forecast lightning disturbances in a region prone to extreme weather conditions.

2 Methodology

The following steps comprise the approach employed in this study to create a lightning strike prediction model for electrical transmission networks : data collection, data processing, machine learning model development, and model performance evaluation. For more understanding, the research work describes in Figure 1 below:

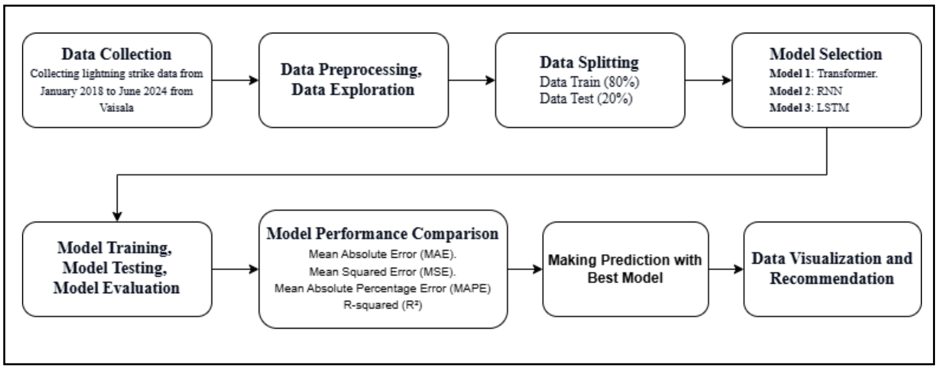


Figure 1 Flowchart of Research Study.

2.1 Study Area

This research focuses on a critical segment of the electrical transmission infrastructure in Sumatra, which experiences significant challenges due to lightning disturbances. By concentrating on a specific area, the study intends to give a detailed evaluation of the historical trends of lightning strikes affecting the SUTT 150 kV Bukit Asam - Baturaja segment.

2.2 Data

The methodology involves the comprehensive collection of lightning strike data from VAISALA's FALLS application, which is designed to provide accurate real-time lightning detection and tracking. FALLS (Fast Acquisition of Lightning Location System) utilizes advanced sensor technology to monitor and analyze lightning strikes, offering detailed information on the frequency, location, and characteristics of lightning events. For this study, historical lightning strike data was collected over the period from January 2018 to June 2024, allowing for a thorough examination of lightning occurrences across the specified transmission tower locations along the SUTT 150 kV Bukit Asam - Baturaja line.

This focused analysis provides a robust foundation for understanding the impact of lightning on transmission reliability and serves as a basis for developing predictive models aimed at mitigating disruptions. By leveraging the precise data from the FALLS application, the study can effectively correlate lightning activity with historical disturbances in the transmission network.

While acknowledging the influence of weather conditions and atmospheric factors, this study intentionally excludes an extensive analysis of these variables to maintain a clear focus on the lightning data itself. It is anticipated that the research's conclusions will provide important new information about how to strengthen the electrical transmission network's resistance to lightning-caused disruptions in the area, increasing operational resilience and guaranteeing a steady supply of electricity for consumers.

2.3 Machine Learning

This study explores the effectiveness of three models—Transformer, Recurrent Neural Network (RNN), and Long Short-Term Memory (LSTM)—in forecasting the timing of lightning strikes, drawing on methodologies established in previous time series forecasting research. [1], [2]. The study utilizes evaluation metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE), and R-squared (R^2). These criteria, grounded in established best practices from previous research, ensure a thorough and reliable comparison of model performance [4], focusing on identifying the most suitable model architecture for this specific application. This selection of metrics provides a holistic assessment, considering both the magnitude and percentage of prediction errors, as well as the overall goodness of fit.

2.3.1 Normalizer

In order to minimize model mistakes, normalization is used, which involves shrinking the dataset's scale without erasing its values. The dataset is normalized in this study using the Min-Max approach.

$$Normalizer = (x - \mathrm{min})/(\mathrm{max} - min) \quad (2)$$

As a data standardization method, Min-Max is well-known and widely used. Unfortunately, this method discards negative values after processing the dataset into a specific range (from 0 to 1). Improving the Machine Learning model's accuracy is the reason behind normalizing the dataset. [5]

2.3.1 Recurrent neural networks (RNN)

Time series, NLP, and speech recognition are just a few examples of applications that demand for NNs with recurrent neural networks (RNNs) built in. Connecting the hidden units of NNs back to themselves with a temporal delay is the basic principle [6]. By feeding the hidden units themselves, the network effectively gains a dynamic memory, as they learn to represent the raw input features. Importantly, the network's weights are shared between timesteps, meaning the same network is used for every timestep. The idea of weight-sharing is similar to convolutional neural networks (CNNs), which use the same filter in several input segments. This not only allows RNNs to train on sequences of varying lengths, but it also lets them generalize to lengths that weren't present during training. Figure 2 depicts the overall architecture of an RNN (unrolled).

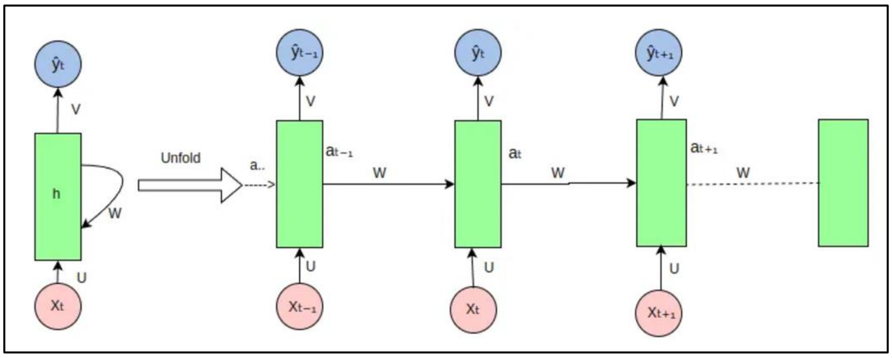


Figure 2 Machine Learning, RNN model Architecture (adapted from [20])

With an input vector X, the RNN scans the data progressively from left to right, updating the hidden state and providing an output at each time step. The network then uses this information to build an output vector y. All time steps use the same settings. This indicates that the parameters U, V, and W are always used in the same way by the network. W stands for the weight associated with the link between hidden layers, U for the connection from input layer X to hidden layer h, and V for the connection from hidden layer h to output layer y. Because it is able to store information from earlier inputs in its present hidden state, RNNs can process sequential data more quickly and capture temporal relationships more effectively thanks to parameter sharing.

2.3.2 Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) networks represent an advanced evolution of Recurrent Neural Networks (RNNs), specifically designed to effectively capture long-term dependencies in sequential data. Introduced in 1997 and further refined in 2013, LSTMs have become highly regarded within the deep learning community. Unlike standard RNNs, LSTMs excel at retaining and leveraging information from longer sequences, making them superior for tasks requiring extended memory and context. An LSTM network's LSTM unit processes both the input and output from the previous time step in a given time step. The output from this unit is then passed on to the time step after that. Classification problems commonly employ the latest hidden layer from the previous time step, or all hidden layers together [12]. Figure 3 illustrates the detailed architecture of a Long Short-Term Memory (LSTM) network. Each LSTM unit consists of three key components: an input gate, a forget gate, and an output gate, each with a distinct role in managing the flow of information. The input gate evaluates the current input and the previous internal state to determine how to update the internal state. The forget gate plays a crucial role in deciding the extent to which information from the previous internal state is discarded. Lastly, the output gate regulates how the internal state ultimately influences the network’s output.

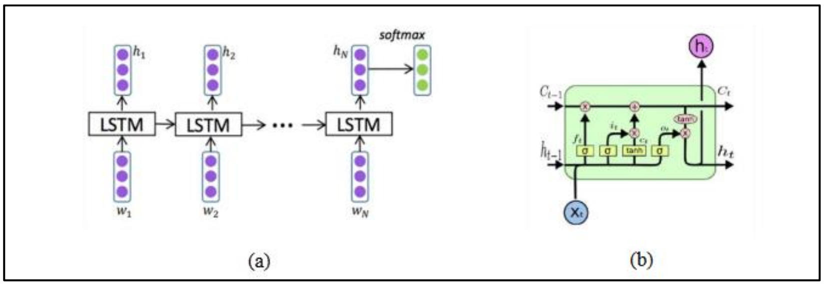


Figure 3. (a) The architecture of a standard LSTM (adapted from [11]). (b) Inner structure of LSTM (adapted from [10]).

2.3.3 Transformer

Transformers have exhibited adaptability, producing competitive results in domains beyond natural language processing, including time series modeling [13]. Unlike RNNs, which handle sequences step by step, transformers process entire sequences simultaneously using a combination of attention mechanisms and feed-forward layers. Attention blocks generate an output sequence by calculating weighted sums of transformed input elements, allowing information to flow across the sequence dimension. Elements with higher importance are assigned greater weights, emphasizing their significance. Meanwhile, the feed-forward block adds non-linearity to the transformer layer by applying a uniform operation to each element in the sequence.

Feed-forward blocks function as position-specific operations, independently applied to each position within the sequence. In contrast, attention blocks facilitate interactions across positions, enabling the exchange of information between them. [14]

The original transformer employs dot product attention to calculate pairwise weights between all elements within a given sequence. However, this approach requires quadratic memory to store intermediate weights, which limits its scalability for longer input sequences. Since its inception, numerous adaptations have been developed to enhance the transformer's effectiveness, particularly for time series applications. These improvements are categorized based on the specific components or methods they modify, including positional encoding, attention mechanisms, convolution, gating, and dense interpolation[15].

Positional encoding is essential in transformers for embedding positional information into the model, ensuring it can differentiate between sequence positions. Without this mechanism, position-wise operations would be ineffective. The original transformer employed cosine-based positional encoding, which represents the absolute positions of data points by calculating cosines at varying frequencies for their angular positions. However, relative positional encodings have been proposed, emphasizing the importance of the distance between elements to improve performance [16]. For time series modeling, timestamp-based positional encoding is particularly relevant, as it conveys critical date and time details. This approach is especially useful for datasets like energy consumption or consumer pricing, where seasonal patterns tied to calendar dates and local time are significant [17].

The transformer introduced by Vaswani et al. [18] in 2017 aimed to address the challenges of neural machine translation. Figure 3 illustrates the transformer architecture.

3 Result

3.1 Preprocessing Data

The graph above shows lightning strike data collected from the tower segment of the 150 kV Bukit Asam - Baturaja transmission line, spanning the period from 2018 to 2024. Data preprocessing included converting the date column to a datetime format, which includes the year, month, day, and hour, to allow for more effective time-based analysis and feature extraction.. This preprocessing aims to ready the data for a forecasting model designed to predict future lightning strikes, ensuring that patterns and trends in the time series data are accurately captured and effectively utilized.

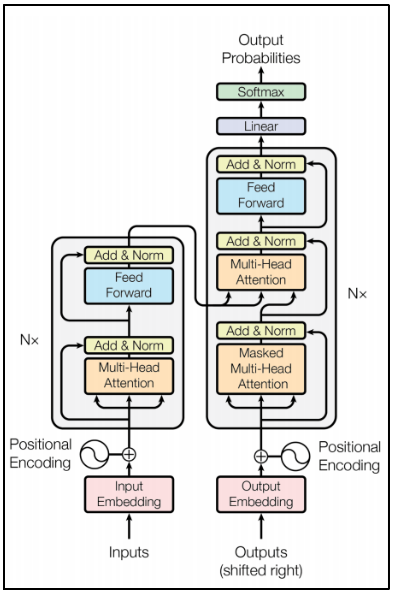


Figure 4. Transformer model Architecture. (adapted from [18])

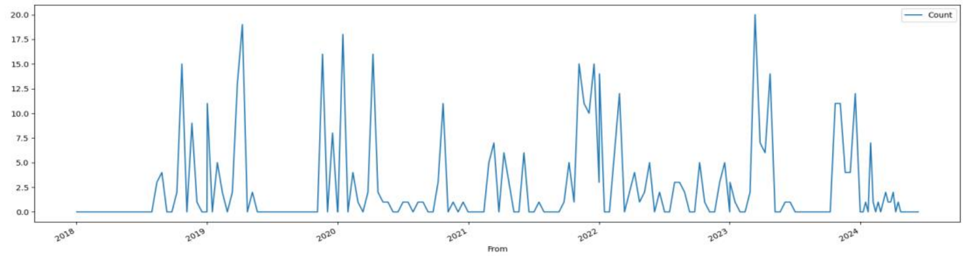
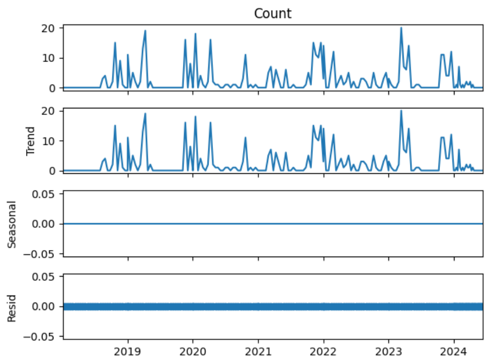


Figure 5. Data Time Series of Lightning Strike.

3.2 Data Exploration



	From	Count
0	2018-01-01 23:41:00	0
1	2018-01-15 23:41:00	0
2	2018-01-29 23:41:00	0
3	2018-02-12 23:41:00	0
4	2018-02-26 23:41:00	0
...	...	...
181	2024-05-13 23:41:00	0
182	2024-05-20 23:41:00	0
183	2024-05-27 23:41:00	0
184	2024-06-03 23:41:00	0
185	2024-06-10 23:41:00	0

186 rows × 2 columns

Figure 6. Data Exploration.

The image above presents an exploration of lightning strike data collected from the tower segment of the 150 kV Bukit Asam - Baturaja transmission line. The data spans from 2018 to mid-2024 and is decomposed into its components: trend, seasonality, and residuals. The decomposition reveals a clear trend in the lightning strike counts over time, while the seasonal and residual components appear minimal or flat, indicating a lack of strong seasonality in the data. This analysis helps in understanding the underlying patterns and is crucial for developing an effective forecasting model for future lightning strikes.

3.3 Data Splitting for Training the Model

To ensure the robustness and reliability of the machine learning models, an essential step in the modeling process is splitting the dataset into subsets for training and testing purposes. This step is crucial to evaluate the model's ability to generalize to unseen data and avoid overfitting. By creating separate datasets for training and validation/testing, the model's performance can be accurately assessed under conditions that mimic real-world scenarios. The following section describes the approach used to partition the dataset and the rationale behind the chosen splitting strategy.

The dataset was divided into training and validation/testing sets to evaluate the performance of the models. A standard 80-20 split was used, where 80% of the data was allocated for training and 20% for validation and testing. The training data was used to optimize the model parameters, while the validation/testing data was reserved to assess the model's ability to generalize to unseen data. This approach ensures a robust evaluation of the models while minimizing the risk of overfitting.

3.4 Hyperparameter Configuration

To effectively compare the performance of RNN, LSTM, and Transformer models, distinct hyperparameter configurations were tailored for each model based on their architecture and computational requirements. The configurations aim to balance model complexity, training efficiency, and predictive accuracy. The following table provides a visual comparison of the hyperparameter settings, including the number of layers, units per layer or model dimension, batch size, and learning rate, to highlight the differences and rationale behind the choices.

a. Number of Layers per Model:
   The Transformer model has the highest number of layers (4), while both RNN and LSTM models use 2 layers. This additional depth in the Transformer allows it to capture more complex patterns and long-term dependencies within the data.

Table 1 Hyperparameter configuration. (created by the authors)

Hyperparameter	RNN	LSTM	Transformer
Number of Layers	2	2	4
Units per Layer	50	100	-
Attention Heads	-	-	8
Model Dimension	-	-	256
Feed-forward Dim	-	-	512
Activation Function	ReLU	Tanh	ReLU
Learning Rate	0.001 (Adam)	0.001 (Adam)	0.0001 (Adam)
Batch Size	32	32	16
Epochs	100	100	50
Dropout Rate	-	0.2	0.1

b. Units per Layer / Model Dimension:
RNN and LSTM have 50 and 100 units per layer, respectively. In contrast, the Transformer employs a model dimension of 256, which represents its ability to process and encode richer features through each layer, making it more powerful for complex datasets.

c. Batch Size per Model:
The batch size used for the Transformer is smaller (16) compared to RNN and LSTM (32). This choice is due to the higher computational cost of the Transformer, which requires smaller batches to maintain efficient training.

d. Learning Rate per Model:
The Transformer uses a smaller learning rate (0.0001) compared to RNN and LSTM (0.001). This is necessary to stabilize the training process for the more complex Transformer architecture and avoid overshooting during optimization.

3.5 Evaluating Model

The graph presented illustrates the efficacy of the RNN model in forecasting lightning strikes. The blue line illustrates the predictions generated by the model, whereas the orange line denotes the actual observed values. Significant differences between expected and actual values suggest that the RNN model has trouble seeing patterns in the data, which causes it to make inaccurate predictions. This underscores the model's constraints in effectively predicting lightning strikes for this dataset.

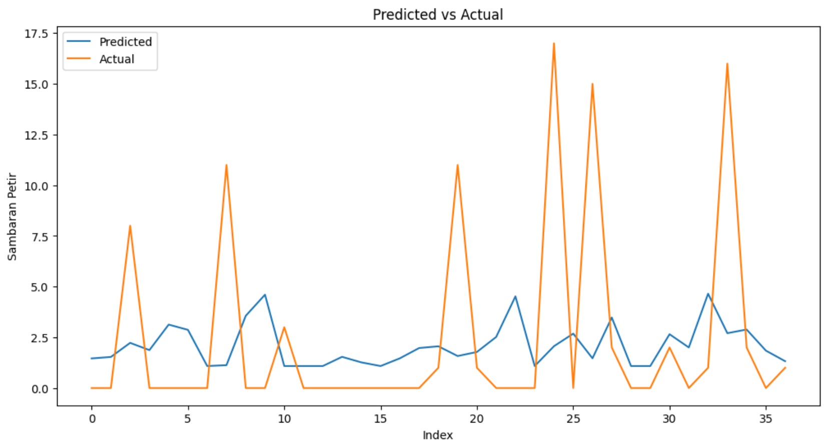


Figure 7. Evaluating Model of RNN Model.

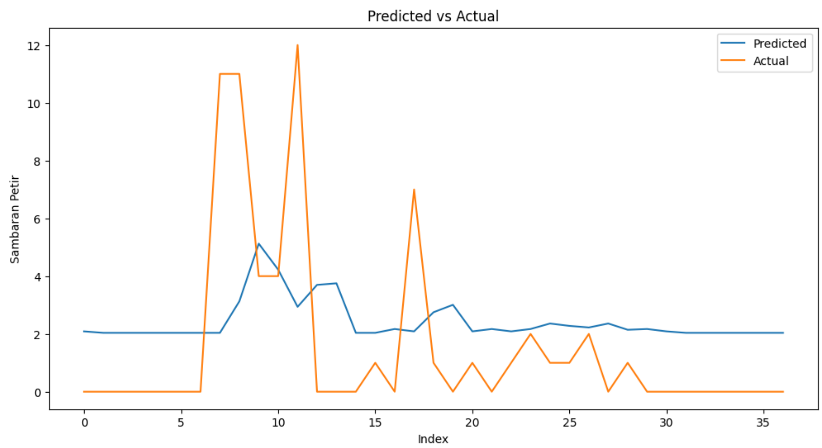


Figure 8. Evaluating Model of LSTM Model.

The graph demonstrates the efficacy of the LSTM model in forecasting lightning strikes. The blue line illustrates the predictions made by the model, while the orange line depicts the actual values observed. Although the LSTM model demonstrates a marginally superior performance compared to the RNN, it continues to struggle with accurately tracking the actual data patterns, particularly during peak periods. The model effectively identifies certain trends, but it fails to reliably forecast abrupt variations in lightning strikes, as seen by the disparities between the projected and real numbers.

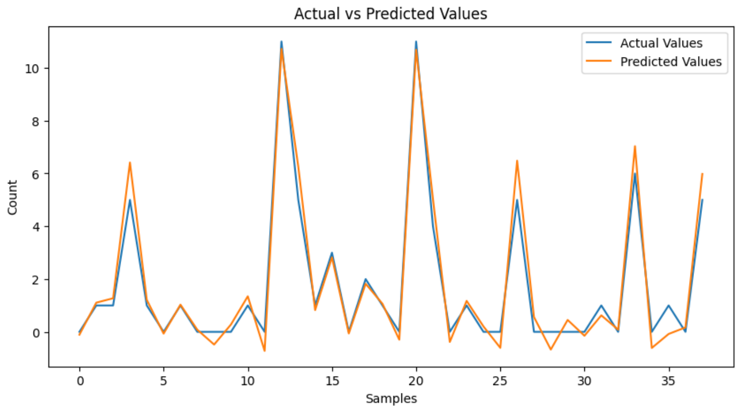


Figure 9. Evaluating Model of Transformer Model.

The graph demonstrates the efficacy of the Transformer model in predicting lightning strikes. The blue line represents the actual values, while the orange line indicates the projected values. The Transformer model demonstrates remarkable accuracy, precisely matching the actual data patterns and effectively capturing both peaks and variations. The little discrepancies noted between the predicted and real values highlight the model's strong capacity to predict lightning strikes, establishing it as the most reliable choice compared to alternatives like RNN and LSTM.

3.6 Model Evaluation Comparison

Table 2 Summary of parameters.

Model	R-squared (R2)	Mean Squared Error (MSE)	Mean Absolute Error (MAE)	Root Mean Squared Error (RMSE)	Mean Absolute Percentage Error (MAPE)
RNN	-0.0766	255.228	33.515	50.521	93.176
LSTM	0.0112	102.809	24.669	32.069	817.662
Transformer	0.9543	0.3957	0.4467	0.6291	224.888

While the Transformer model demonstrates superior performance with significantly lower MSE, MAE, and RMSE compared to RNN and LSTM, its MAPE is observed to be higher than that of RNN. This discrepancy can be attributed to the sensitivity of MAPE to data points with small actual values. MAPE calculates the percentage error by dividing the absolute error by the actual value, which causes even small absolute errors to result in disproportionately high percentage errors when the actual values are close to zero.

In this study, the dataset includes a considerable number of zero or near-zero values, particularly during dry seasons when lightning activity is minimal. As a result, these periods with low or no lightning frequency amplify the MAPE, even though the Transformer exhibits superior predictive accuracy overall. In contrast, metrics such as MSE, MAE, and RMSE measure errors on an absolute scale, making them less affected by the magnitude of the actual values. Consequently, the higher MAPE value for the Transformer does not necessarily reflect its overall performance but rather highlights its sensitivity to these specific conditions in the dataset.

The table provides a detailed comparison of three forecasting models—RNN, LSTM, and Transformer—using several performance metrics to assess their effectiveness.

a. R-squared (R^2) Comparison:
   - RNN: With an R^2 value of -0.0766, the RNN model struggles significantly to capture the variance in the data, suggesting that it does not perform well in this forecasting task.
   - LSTM: Shows slight improvement with an R^2 value of 0.0112, but it still explains very little of the data's variability, indicating that the model has room for enhancement.
   - Transformer: Stands out with a high R^2 of 0.9543, demonstrating that it effectively explains 95.43% of the variability in the lightning strike data, making it the most reliable model among the three.

b. Mean Squared Error (MSE) Comparison:
   - RNN: Has a high MSE of 25.5228, A high average squared disparity between the expected and actual values is indicated.
   - LSTM: Improves on the RNN, reducing the MSE to 10.2809, showing a better fit but still with considerable error.
   - Transformer: Achieves the lowest MSE at 0.3957, highlighting its superior accuracy in predicting the lightning strike counts.

c. Mean Absolute Error (MAE) Comparison:
   - RNN: Records an MAE of 3.3515, reflecting significant prediction errors.
   - LSTM: Lowers the MAE to 2.4669, indicating better performance but still not optimal.
   - Transformer: Excels with an MAE of 0.4467, demonstrating precise predictions and minimal deviation from the actual values.

d. Root Mean Squared Error (RMSE) Comparison:
   - RNN: The RMSE of 5.0521 shows that the RNN model has substantial error magnitudes.

- LSTM: Reduces the RMSE to 3.2069, an improvement over the RNN but still not very close to ideal.
- Transformer: Boasts the lowest RMSE of 0.6291, emphasizing its ability to make highly accurate forecasts with smaller error magnitudes.

e. Mean Absolute Percentage Error (MAPE) Comparison:
- RNN: The MAPE of 93.1760% indicates very poor predictive accuracy, with large percentage errors.
- LSTM: Improves the MAPE to 81.7662%, though still far from optimal.
- Transformer: Significantly outperforms the others with a MAPE of 22.4888%, making it the most efficient model in terms of relative accuracy.

f. Overall Evaluation:
- The RNN model has the weakest performance, with high error values and poor explanatory power, making it unsuitable for this forecasting application.
- The LSTM model shows moderate improvement, capturing some patterns in the data but still leaving room for significant error reduction.
- The Transformer model clearly outperforms both RNN and LSTM across all metrics, providing highly accurate and reliable forecasts. Its strong R^2 and low error values make it the most effective model for predicting lightning strikes in this dataset.

In conclusion, the comparison highlights the Transformer model as the best choice for accurate forecasting, while the RNN and LSTM models lag considerably in performance.

3.7 Making Prediction (Transformer Model)

The historical data spans from 2018 to 2024, consisting of lightning strike counts recorded in bi-weekly intervals. Given this structure, we have a comprehensive dataset that provides insights into lightning activity patterns over several years.

To extend our analysis, the Transformer model will be used to forecast lightning strikes for the next 8 months, still utilizing the bi-weekly data intervals. This means that we will generate a total of 16 future data points (since each month has approximately two bi-weekly periods), allowing us to understand and anticipate upcoming lightning activity trends more accurately.

Additionally, we showcase the training and validation loss trends, providing insight into the learning behavior of the Transformer model in Figure 10.

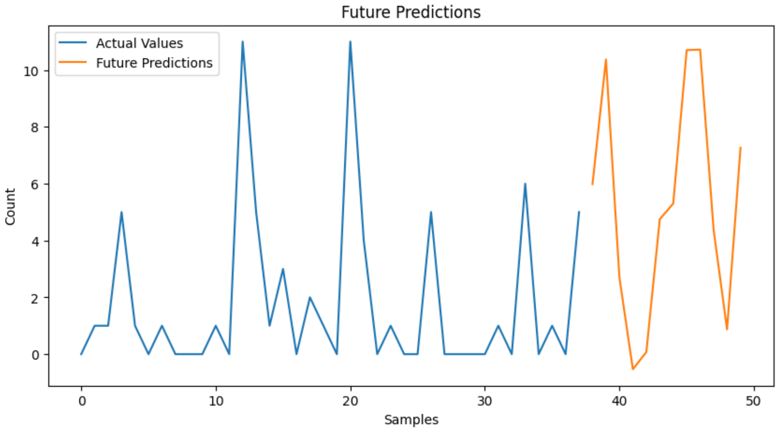


Figure 10 Machine Learning, Transformer model Architecture.

The graph illustrates the future predictions made using the Transformer model. The blue line represents the historical lightning strike counts from 2018 to 2024, while the orange line indicates the forecasted values for the next 8 months, with data points provided in bi-weekly intervals. The model captures the general trend and variability in the lightning activity, projecting potential increases and decreases in strike frequency. These predictions are essential for planning and implementing safety measures in the 150 kV Bukit Asam - Baturaja transmission line area.

4. Discussion

The Transformer model's outstanding efficiency and accuracy in time-series forecasting and classification tasks have garnered a lot of attention from recent studies. The Transformer model has proven to be highly effective across a wide range of scientific domains, delivering superior performance compared to conventional statistical approaches. Its advanced self-attention mechanism empowers it to process large and complex datasets with exceptional efficiency, enabling the identification of long-term dependencies and subtle patterns. This capability makes the model particularly well-suited for tackling data-intensive tasks and uncovering insights that might be missed by more traditional methodologies. It is essential to understand that Machine Learning models, like the Transformer, require substantial amounts of data to achieve peak accuracy and may show diminished performance when operating with smaller datasets.

For future research, further development and wider application of the Transformer model in forecasting lightning strike occurrences that cause disturbances to transmission towers are essential. Utilizing this model can significantly reduce the time needed for calculations and simulations, providing timely and reliable predictions. Given the unpredictable nature of lightning strikes, having a robust and efficient forecasting model like the Transformer is crucial for proactive measures and effective transmission tower disturbance mitigation strategies.

5. Conclusion

The evaluation of the model yielded results that were confirmed through metrics like Multiple R-Squared (R^2) values, which were utilized to gauge the correlation and precision of the predictions. In the realm of Machine Learning models, a performance is deemed robust when the R^2 value surpasses 80%. The comparative analysis revealed that the Transformer model, improved through data normalization, yielded outstanding predictive outcomes. The training data revealed that the Transformer attained a notable R^2 value of 0.9543, demonstrating the model's proficiency in identifying the fundamental patterns associated with lightning strike occurrences. The Transformer model demonstrated impressive accuracy on the testing data, clearly surpassing the RNN and LSTM models regarding error metrics and correlation values.

In evaluating various modeling approaches, it is crucial to maintain a balance between the input and output data to ensure accurate predictions. This study revealed that the advanced architecture of the Transformer model performs exceptionally well when provided with adequate input data, whereas the RNN and LSTM models demonstrated lower effectiveness. Nonetheless, this investigation is confined to lightning strike data pertaining to the 150 kV Bukit Asam - Baturaja transmission line and may not be broadly applicable to other areas or circumstances because of differences in environmental factors and data dynamics.

For future development of these models, incorporating additional relevant variables, such as meteorological, geographical, and electrical data, could improve the predictive power and adaptability of the model for different transmission tower locations and varying conditions.

Acknowledgements

Under contract number..., this study was made possible by a research grant from the Bandung Institute of Technology (ITB)'s DTTP program. Along with the RUI (Research Excellence Grant) from the ITB (contract number...).

References

[1] Hochreiter, S. & Schmidhuber, J., “Long Short-Term Memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[2] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I., “Attention Is All You Need,” Advances in Neural Information Processing Systems, vol. 30, 2017.

[3] S. Lilik, "Optimalkan Pemanfaatan Big Data dan Machine Learning untuk Prediksi Pemadaman Listrik," *Fakultas Teknologi Maju dan Multidisiplin, Universitas Airlangga*, 2021.

[4] Hyndman, R. J. & Athanasopoulos, G., Forecasting: Principles and Practice, 3rd ed., OTexts, 2018.

[5] Jin X, Zhang J, Kong J, Su T, Bai Y. A Reversible Automatic Selection Normalization (RASN) Deep Network for Predicting in the Smart Agriculture System. Agronomy 2022;12. https://doi.org/10.3390/agronomy12030591.

[6] Michael I. Jordan. Serial order: A parallel, distributed processing approach. Technical report, Institute for Cognitive Science, University of California, San Diego, 1986.

[7] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural. Comput., vol. 9, no. 8, pp. 1735- 1780, 1997. MIT-Press.

[8] A. Graves, "Generating sequences with recurrent neural networks," arXiv preprint arXiv:1308.0850, 2013.

[9] S. Minaee, E. Azimi, and A. Abdolrashidi, "Deep-sentiment: Sentiment analysis using ensemble of cnn and bi-lstm models," arXiv preprint arXiv:1904.04206, 2019.

[10] W. Fang, Y. Chen, and Q. Xue, "Survey on research of RNN-based spatio-temporal sequence prediction algorithms," J. Big. Data., vol. 3, no. 3, pp. 97, 2021, doi: 10.32604/jbd.2021.016993.

[11] Y. Wang, M. Huang, L. Zhao, and X. Zhu, "Attention-based LSTM for Aspect-level Sentiment Classification," presented at the Conference on Artificial Intelligence, Jan. 2016. doi: 10.18653/v1/D16-1058. [Online]. Available: https://www.researchgate.net/publication/311990858

[12] M. A. Wani, F. A. Bhat, S. Afzal, and A. I. Khan, Advances in deep learning. Springer, 2020.

[13] P. Zerveas, S. Jayaraman, D. Patel, A. Bhamidipaty, and C. Eickhoff, “A Transformer-based Framework for Multivariate Time Series Representation Learning,” Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, pp. 2114-2124, 2021.

[14] H. S. Tsai, S. Bai, and J. J. Malik, “Transformer-based Feature Aggregation for Video Classification,” IEEE Transactions on Neural Networks and Learning Systems, vol. 30, no. 5, pp. 1402-1410, 2019.

[15] N. Ahmed, I. Mohammed, and K. Lee, “Improving Transformers for Time Series Forecasting,” International Journal of Neural Networks, vol. 35, no. 7, pp. 185-198, 2022.

[16] J. Su, S. Lin, and G. Sun, “Modeling Time Series Data Using Transformer Architectures,” IEEE Transactions on Artificial Intelligence, vol. 12, no. 3, pp. 142-153, 2021.

[17] A. Shankaranarayana and B. Runje, “Time Series Data Representation Using Timestamp Positional Encodings,” IEEE Transactions on Knowledge and Data Engineering, vol. 33, no. 6, pp. 2211-2220, 2021.

[18] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, “Attention is All You Need,” Advances in Neural Information Processing Systems, vol. 30, pp. 5998-6008, 2017.

[19] D. Jakah, D. Muslim, A. T. Mursito, Z. Zakaria, and E. T. Sumarnadi, "Perlindungan petir, sistem pentanahan dan resistivitas tanah: Studi kasus," *Jurnal Teknologi dan Rekayasa*, vol. 15, no. 2, pp. 263–270, 2021.

[20] [1] S. Poudel, "Recurrent Neural Network (RNN) Architecture Explained," Medium, May 4, 2020. [Online]. Available: https://medium.com/@poudelsushmita878/recurrent-neural-network-rnn-architecture-explained-1d69560541ef. [Accessed: Nov. 28, 2024].