MethodsX Riswanda Safira Indah Ghani

by MOHAMMAD GHANI

Submission date: 15-Jun-2024 12:38AM (UTC+0200)
Submission ID: 2402669169
File name: MethodsX_Riswanda_Safira_Indah_Ghani.pdf (1.67M)
Word count: 5806
Character count: 29605

Article title

Comparative analysis of Long Short-Term Memory and Extreme Gradient Boosting methods to predict the ocean wave in Tuban Regency for fisherman safety

Authors

Riswanda Ayu Dhiya'ulhaq, Anisya Safira, Indah Fahmiyah, Mohammad Ghani*

Affiliations

Data Science Technology, Faculty of Advanced Technology and Multidiscipline, Universitas Airlangga, Surabaya 60115, Indonesia

Corresponding author's email address

*mohammad.ghani@ftmm.unair.ac.id

Keywords

Long Short-Term Memory (LSTM), Extreme Gradient Boosting (XGBoost), Wave Prediction

Related research article

None

For a published article

None

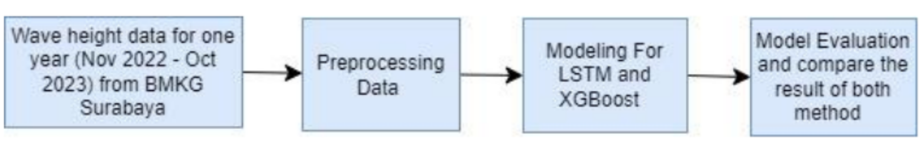
Abstract

The fishing industry has a large role in the Indonesian economy, with potential profits in 2020 of around US$ 1.338 billion. Tuban Regency is one of the regions in East Java that contributes to the fisheries sector. Fisheries relate to the work of fishermen. Accidents in shipping are still a major concern. One of the natural factors that influence shipping accidents is the height of the waves. Fisherman safety regulations have been established by the Ministry of Maritime Affairs and Fisheries and the Meteorology, Climatology and Geophysics Agency. Apart from regulations, the results of wave height predictions using the Long Short-Term Memory (LSTM) and Extreme Gradient Boosting (XGBoost) methods can help fishermen determine shipping departures, thereby reducing the risk of accidents. In this study, the Grid Search hyperparameter tuning process was used for both methods which were carried out on four location coordinates. Based on the analysis results, LSTM is superior in predicting wave height at the first location (RMSE 0.045; MAE 0.029; MAPE 8.671%), second location (RMSE 0.051; MAE 0.035; MAPE 10.64%), and third location (RMSE 0.044; MAE 0.027; MAPE 7.773%). XGBoost provided superior results in the fourth location (RMSE 0.040; MAE 0.025; MAPE 7.286%).

- Hyperparameter tuning with gridsearch is used in LSTM and XGBoost to obtain optimal accuracy
- LSTM outperforms XGBoost in three locations, while XGBoost outperforms in the fourth location.
- Advanced prediction techniques such as LSTM and XGBoost improve fishermen's safety by providing accurate wave height estimates, thereby reducing the possibility of shipping accidents.

Graphical abstract

In general, the stages of this research include secondary data collection, preprocessing data tailored to the characteristics of the data, then modeling data, and making evaluations including comparing the results of the two methods.



Specification table

Subject area	Engineering
More specific subject area	Ocean Engineering
Name of your method	Long Short-Term Memory and Extreme Gradient Boosting
Name and reference of origin method	None
Resource availability	https://github.com/riswandayu/predictoceanwave

Background

Fishing is inherently dangerous, characterized by its 3D characteristics—dangerous, dirty and, difficult (ILO, 2016). Among the natural elements that influence maritime activities, wave height is the most prominent thing. The Meteorology, Climatology and Geophysics Agency (BMKG) provides important safety advice, recommending that fishermen navigate when wave heights do not exceed 1.25 meter (Asmaul Husnah et al., 2023). Estimated wave height plays an important role in determining the safe departure of fishing vessels. Various methodologies, including stochastic techniques and deep learning, are commonly used for this purpose. Stochastic methods such as Autoregressive Integrated Moving Average (ARIMA) are used to predict wave patterns, especially in Jakarta Bay. The ARIMA experiment involving eight combinations produced an optimized Root Mean Square Error (RMSE) value for the ARIMA model (2,2,2) of 0.0106, which shows its ability to predict waves accurately up to 24 hours in advance (Adytia et al., 2019). Another interesting method is Random Forest (RF), which is widely used in operational wave forecasting in the Atlantic Ocean. Campos et al (2021) highlighted the decline in RF performance as forecast time increases, indicating its limitations in long-term predictions.

Long Short-Term Memory (LSTM) networks, which are a modification of Recurrent Neural Networks (RNN), are also gaining traction due to their ability to accurately process and predict time series data. The study by Meng et al (2022) shows the high accuracy of LSTM, especially with 100 input variables, producing minimum Mean Absolute Error (MAE) and RMSE values of 0.015 and 0.019 respectively. The findings of Pramesti et al (2022) in the Bali Strait further strengthen the effectiveness of LSTM in predicting current components with low Mean Absolute Percentage Error (MAPE) values. XGBoost, an Extreme Gradient Boosting algorithm, has emerged as a widely adopted tool for predicting wave height. Research by Hu et al (2021) shows the superior performance of Erie. In addition, the application of XGBoost in predicting wave height on flat beaches produces good Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) values (Tarwidi et al., 2023). The integration of the LSTM and XGBoost methodologies in predicting sea wave height in the Tuban Regency is expected to reduce accidents and losses for fishermen. These efforts are in line with the Sustainable Development Goals (SDGs), especially point 13, which underlines the importance of addressing climate change. By improving maritime safety and utilizing advanced forecasting techniques, these efforts contribute not only to protecting the lives of fishermen but also to addressing global climate challenges.

Method details

Data Source

The data source used in this research is secondary data obtained from Perak Meteorological Station Surabaya II. Wave height data is used for one year, starting from November 2022 to October 2023, with the water area taken from the coordinates -6.21 °S;111.80°E to -6.31 °S; 112.27 °E as in **Fig. 1**. The determination of the location point is based on the activities of fishermen and also the fishery products that are more abundant when the waters are 20 to 30 miles from the coast (Eni, 2017). There are four observed locations, the first location with coordinates -6.21°S;111.80°E, the second location with coordinates -6.22°S;112.04°E, the blue line shows the third location with coordinates -6.26°S;112.20°E, and the green line shows the fourth location with coordinates -6.31°S;112.27°E. This data includes time-related information containing data per day with time every 6 hours starting at 00.00, 06.00, 12.00, and 18.00, longitude values, latitude values, and High Significant Wave (HSW) parameters with a total of 5,348 lines.

Figure 1. Coordinate Map

Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) was introduced by Hochreiter and Schmidhuber in 1997 (Hochreiter and Schmidhuber, 1997). LSTM is one part of deep learning which is a subtype of Recurrent Neural Networks (RNN). LSTM is a modification of RNN that has the advantage of solving the missing gradient problem (Prakash et al., 2021). LSTM can learn and handle very long time dependencies, with a minimum time lag of more than 1,000 time steps (Staudemeyer & Morris, 2019). The LSTM has four interacting layers (Ghosh et al., 2019) where the LSTM architecture is shown in (Marco, 2022). The gates in an LSTM cell can be explained as follows:

$$i_t = \sigma(x_t U^i + h_{t-1} W^i) \tag{1}$$

$$f_t = \sigma(x_t U^f + h_{t-1} - W^f) \tag{2}$$

$$o_t = \sigma(x_t U^o + h_{t-1} - W^o) \tag{3}$$

$$\widetilde{C_t} = \tanh\,(x_t U^g + h_{t-1} W^g) \tag{4}$$

$$C_t = \sigma(f_t * C_{t-1} + i_t * \widetilde{C_t}) \tag{5}$$

$$h_t = \tanh(C_t) * o_t \tag{6}$$

where i_t is the input gate, f_t is the forget gate, o_t is the output gate, W is the connection between the current and previous nodes, U contains the weights of the inputs to the *hidden layer*, $\widetilde{C_t}$ is candidate *hidden layer*, and C is the memory unit.

LSTM Hyperparameter Tuning

The LSTM hyperparameter tuning method uses Grid Search, which is defined as a method of testing a given hyperparameter combination on a grid configuration (Belete and Huchaiah, 2022). The grid search method will search for all possibilities by preparing a grid, which is then evaluated to get the best value among all grids (Anggoro and Mukti, 2021). **Table 1** presents the parameters used in LSTM. In this research, the Tanh activation function and ADAM optimizer are used. The calculation of the Tanh gradient is simple and converges faster than sigmoid (Vijayaprabakaran and Sathiyamurthy, 2022). Adam is an adaptive learning rate technique that lowers the individual learning rate for various parameters (Ismanto and Effendi, 2023).

Table 1. LSTM Hyperparameter Tuning

Parameters	
Hidden Layer	4
Optimizer	Adam
Neuron	30;50
Activation Function	Tanh
Learning Rate	0,001; 0,01; 0,1
Batch Size	32;64
Epoch	50; 100

Extreme Gradient Boosting (XGBoost)

Extreme Gradient Boosting (XGBoost) is a machine learning algorithm introduced by Chen and Guestrin in 2016 (Chen and Guestrin, 2016). XGBoost is one of the boosting methods. Boosting is an ensemble learning technique that combines a set of simpler and weaker sets to improve model accuracy (Malik et al., 2020). XGBoost minimizes regularized losses, and use of L2 regularization aims to obtain a more general model. L2 regularization helps prevent overfitting, adding L2 regularization in the loss function results in more controllable model complexity (Murty & Kiran Kumar, 2019). In general, the L2 regularization equation is as follows:

$$Cost\ Function = loss + \frac{\lambda}{2m} * \sum \|w\|^2 \qquad (7)$$

where λ is the L2 regularization parameter whose value can be optimized for better results, m the number of trees and w is the weight vector or model parameters.

XGBoost Hyperparameter Tuning

The XGBoost hyperparameter tuning method uses Grid Search. Boosting reduces bias and increases variance by increasing the complexity of weak models. By using hyperparameters, overfitting can be prevented (Tarwidi et al., 2023). Table 2 describes the hyperparameters used in XGBoost.

Table 2. XGBoost Hyperparameter Tuning

Parameters	
colsample_bytree	0,5; 0,7; 1
learning_rate	0,05; 0,15; 0,3
max_depth	3; 5; 6
n_estimators	50; 100; 150
reg_lambda	0; 0,1; 0,5; 1
subsample	0,6; 1

Effectively, these hyperparameters affect the tree building and Gain calculation (λ and γ) or the data and feature selection process for each iteration such as subsample and colsample bytree, by adjusting them it can maintain a balance between model complexity and predictive ability, and improve performance by reducing overfitting (Ding, 2022).

Prediction Flowchart with LSTM and XGboost

The research work is divided into two parts, the prediction part with LSTM and XGBoost described in **Fig. 2**.

1. Input Data.
2. Data preprocessing, filtering data based on location and normalizing data. The method used in the normalization process is min-max scaler by converting the actual value into a value with an interval range (de Amorim et al., 2023). The following is the normalization formula.

   $$X' = \frac{x - x_{min}}{x_{max} - x_{min}} \qquad (8)$$

   where X' is the normalized value, x is the actual data value to be normalized, x_{min} is the minimum value of the actual data, and x_{max} is the maximum value of the actual data.

3. Data division with a percentage of 80% for training data and 20% for testing data.
4. LSTM modeling for each location using hyperparameter tuning with grid search and using early stopping to prevent overfitting or underfitting (Jabbar dan Khan, 2015).
5. Modeling XGBoost for each location using hyperparameter tuning with grid search. The best parameter tuning results at each location will be applied in building a prediction model for each location. In grid search, negative square error scoring is used.
6. Model evaluation to see the performance of the model using RMSE, MAE, and MAPE. Here is the evaluation formula (Chang et al., 2018).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad (9)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{10}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\% \tag{11}$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value, dan n is the number of predictions.

7. Denormalization, to return the predicted value to its original value. Here is the formula for denormalization (Putri Udiani et al., 2020).

$$d = y'(\max(y) - \min(y)) + \min(y) \tag{12}$$

d is the denormalized value, y' is the normalized value, max is the minimum value of the actual data, and min is the maximum value of the actual data.

8. Model Interpretation, Interpretation is related to the visualization of data presentation.

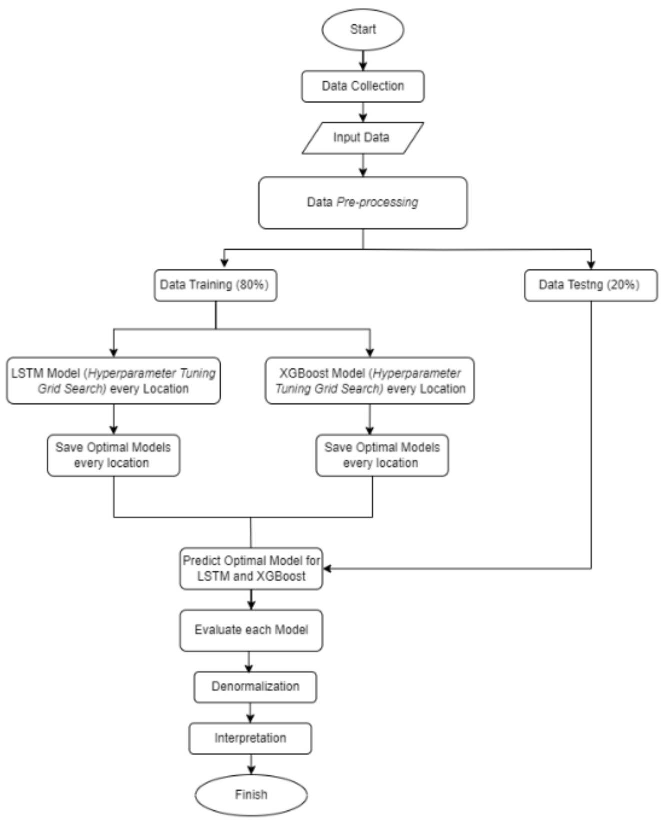


Figure 2. Flowchart

Method validation

Long Short-Term Memory (LSTM)

1st Location (-6.21°S;111.80°E)

Based on the tuning results using grid search for the first location, the negative mean square error value of -0.00352 or RMSE value of 0.05939 is the best parameter using batch size 32; epoch 50; activation tanh; learning rate 0.01; number of neurons 50; optimizer Adam. The best parameters are then used to make wave height predictions, obtained

MAPE evaluation results of 8.671%; MAE 0.029; and RMSE 0.045 for the first location. The prediction of wave height at the first location (**Fig. 3**) shows that between the actual value and the predicted value has a relatively small difference, the difference does not look significant.

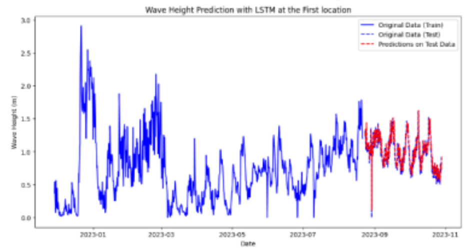


Figure 3. Comparison of Actual and Predicted Values Location 1

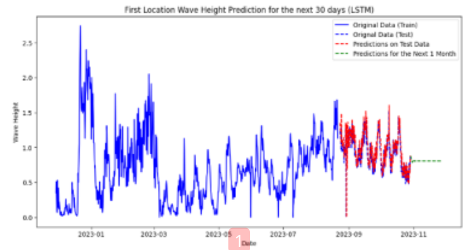


Figure 4. Next 30-Day Wave Height Prediction

Based on the evaluated model, a prediction of the wave height for the next 30 days was made using the historical data in the test data. The prediction is done to predict the wave height at the first location every 6 hours, which is presented in **Fig. 4**. The green line shows the pattern of the predicted wave height for the next 30 days, which tends to decrease when compared to the previous period.

2nd Location (-6.22°S;112.04°E)

The tuning results using grid search for the second location obtained a negative mean square error value of -0.00290 or RMSE value of 0.05385 with parameters batch size 64; epoch 100; activation tanh; learning rate 0.001; number of neurons 50; optimizer Adam. The best parameters are then used to make wave height predictions, obtained MAPE evaluation results of 10.79%; MAE 0.035; and RMSE 0.051 for the second location. The prediction of wave height at the second location (**Fig. 5**) shows that the prediction plot (red line) has shifted slightly to the right but there is no significant difference between the actual value and the predicted value, it can be seen from the pattern between the two lines that have almost the same pattern and height.

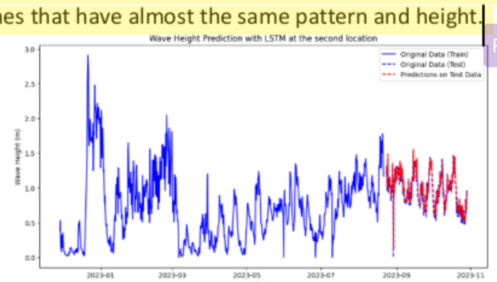


Figure 5. Comparison of Actual and Predicted Values Location 2

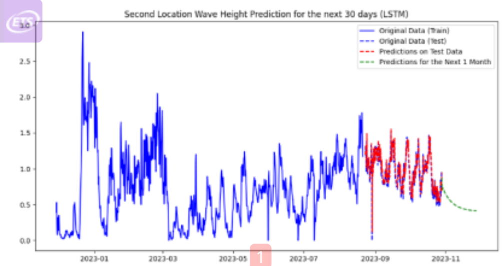


Figure 6. Next 30-Day Wave Height Prediction

Based on the evaluated model, the wave height prediction for the next 30 days was then carried out using the historical data in the test data. The prediction is done to predict the wave height at the first location every 6 hours, which is presented in **Fig. 6**. The green line shows the pattern of the wave height prediction for the next 30 days, which tends to decrease when compared to the previous period.

3rd Location (-6.26°S;112.20°E)

The tuning results using grid search for the third location obtained a negative mean square error value of -0.00214 or RMSE value of 0.04633 with the best parameters using batch size 32; epoch 50; activation tanh; learning rate 0.001; number of neurons 50; optimizer Adam. The best parameters are then used to make wave height predictions, obtained MAPE evaluation results of 7.773%; MAE 0.027; and RMSE 0.044 for the third location. Wave height prediction at the third location (**Fig. 7**) shows that the prediction plot (red line) at the third location tends to have a lower value when compared to the actual data.

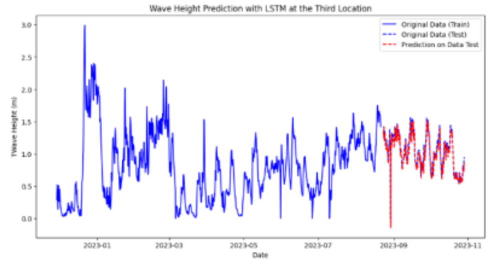


Figure 7. Comparison of Actual and Predicted Values Location 3

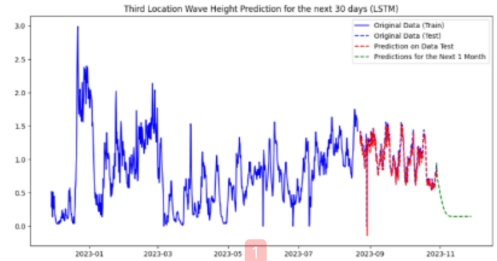


Figure 8. Next 30-Day Wave Height Prediction

Based on the evaluated model, a prediction of the wave height for the next 30 days was then made using the historical data in the test data. The prediction is done to predict the wave height at the first location every 6 hours, which is presented in **Fig. 8**. The green line shows the pattern of the predicted wave height for the next 30 days, which tends to decrease when compared to the previous period.

4th Location (-6.26°S;112.20°E)

The tuning results using grid search for the fourth location obtained a negative mean square error value of -0.00203 or RMSE value of 0.04506 with the best parameters using batch size 32; epoch 100; activation tanh; learning rate 0.001; number of neurons 30; optimizer Adam. The best parameters were then used to make wave height predictions, obtained MAPE evaluation results of 7.386%; MAE 0.025; and RMSE 0.042 for the fourth location. Wave height prediction at the fourth location (**Fig. 9**) shows the prediction plot (red line) tends to follow the pattern and height of the actual data, when compared to the visualization results of the other three locations, the prediction plot on the fourth location test data is more fit to the actual data.

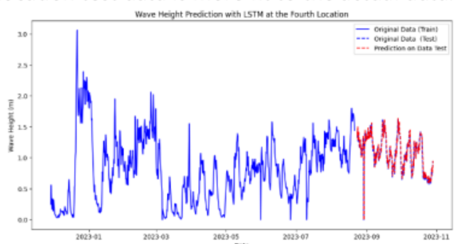


Figure 9. Comparison of Actual and Predicted Values Location 4

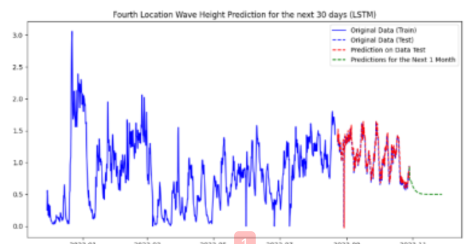


Figure 10. Next 30-Day Wave Height Prediction

Based on the evaluated model, a prediction of the wave height for the next 30 days was made using the historical data in the test data. The prediction is done to predict the wave height at the first location every 6 hours, which is presented in **Fig. 10**. The green line shows the pattern of the predicted wave height for the next 30 days, which tends to decrease when compared to the previous period.

Extreme Gradient Boosting (XGBoost)

1st Location (-6.21°S;111.80°E)

By using grid search, the best parameter combination is colsample bytree 1; learning rate 0.05; max depth 3; n estimator 150; reg lambda 0.1; subsample 0.6 with a negative mean square error value of -0.00406 or RMSE value of 0.06376. The best parameters are then used to make wave height predictions, obtained MAPE evaluation results of 9,022%; MAE 0.031; and RMSE 0.046 for the first location. The prediction results show that the test data (blue line) (**Fig. 12**) has a higher value when compared to the prediction plot, however the prediction plot pattern is quite following the pattern of the actual data.

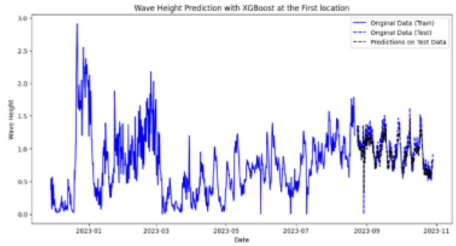


Figure 11. Comparison of Actual and Predicted Values Location 1

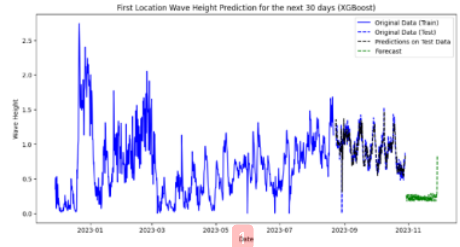


Figure 12. Next 30-Day Wave Height Prediction

Based on the evaluated model, a prediction of the wave height for the next 30 days was made using the historical data in the test data. The prediction is done to predict the wave height at the first location every 6 hours, which is presented in **Fig. 12**. The green line shows the pattern of the predicted wave height for the next 30 days, which tends to decrease when compared to the previous period.

2nd Location (-6.22°S;112.04°E)

The best parameter combination obtained with grid search is colsample bytree 1; learning rate 0.15; max depth 3; n estimator 100; reg lambda 0.5; subsample 0.6 with a negative mean square error value of -0.00374 or RMSE value of 0.06122. The best parameters are then used to make wave height predictions, obtained MAPE evaluation results of 11.05%; MAE 0.036; and RMSE 0.050 for the second location. A visualization of the prediction results is presented in **Fig. 13**. The prediction plot (black line) has higher values at the end of August and mid-September when compared to the actual values. Overall, it can be observed that the plot of the predicted data does not follow the pattern of the actual data.

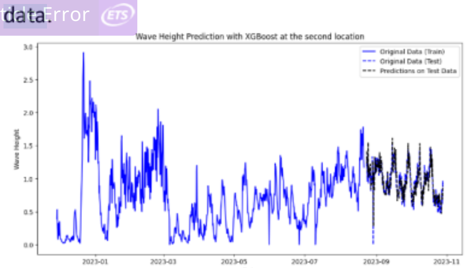


Figure 13. Comparison of Actual and Predicted Values Location 2

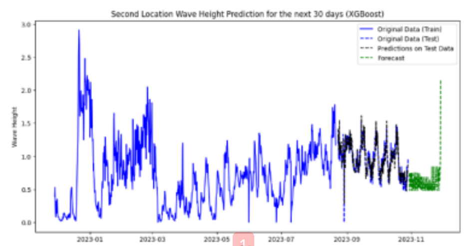


Figure 14. Next 30-Day Wave Height Prediction

Based on the evaluated model, the wave height prediction for the next 30 days was then carried out using the historical data in the test data. The prediction is done to predict the wave height at the first location every 6 hours, which is presented in **Fig. 14**. The green line shows the pattern of the wave height prediction for the next 30 days, which tends to decrease when compared to the previous period.

3rd Location (-6.26°S;112.20°E)

The best parameter combination for the third location is colsample bytree 1; learning rate 0.15; max depth 3; n estimator 50; reg lambda 0; subsample 0.6 with a negative mean square error value of -0.00267 or an RMSE value of 0.05171. The best parameters are then used to make wave height predictions, obtained MAPE evaluation results of 8,160%; MAE 0.028; and RMSE 0.043 for the third location. Based on **Fig. 15**, the prediction plot (black line) has a higher value in early September when compared to the actual value. Then in October the prediction plot also shows lower values than the actual data. However, overall the plots of the actual and predicted data have almost the same pattern and height.

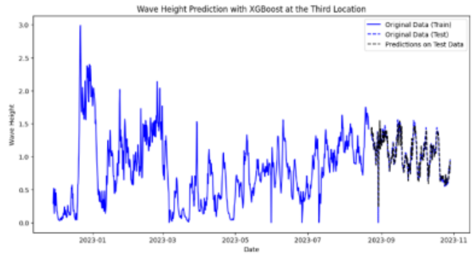


Figure 15. Comparison of Actual and Predicted Values Location 3

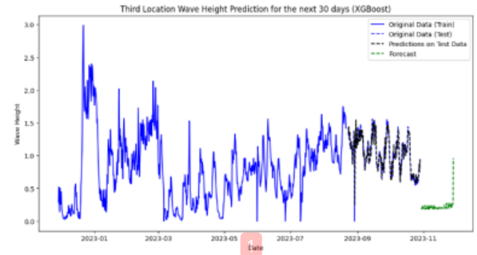


Figure 16. Next 30-Day Wave Height Prediction

Based on the evaluated model, a prediction of the wave height for the next 30 days was then made using the historical data in the test data. The prediction is done to predict the wave height at the first location every 6 hours, which is presented in **Fig. 16**. The green line shows the pattern of the predicted wave height for the next 30 days, which tends to decrease when compared to the previous period.

4th Location (-6.26°S;112.20°E)

The best parameter combination for the fourth location is colsample bytree 1; learning rate 0.05; max depth 3; n estimator 100; reg lambda 0; subsample 1 with a negative mean square error value of -0.00275 or an RMSE value of 0.05249. The best parameters were then used to make wave height predictions, obtained MAPE evaluation results of 7.286%; MAE 0.025; and RMSE 0.040 for the fourth location. Based on **Fig. 17** the prediction plot (black line) has a higher value in early September when compared to the actual value. In October the actual data plot tends to be higher when compared to the prediction data plot.

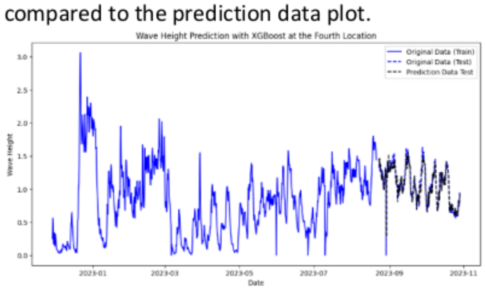


Figure 17. Comparison of Actual and Predicted Values Location 4

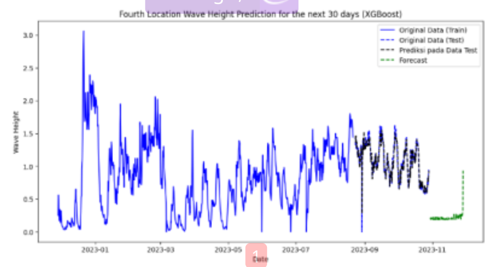


Figure 18. Next 30-Day Wave Height Prediction

Based on the evaluated model, a prediction of the wave height for the next 30 days was made using the historical data in the test data. The prediction is done to predict the wave height at the first location every 6 hours, which is presented in **Fig. 18**. The green line shows the pattern of the predicted wave height for the next 30 days, which tends to decrease when compared to the previous period.

Comparison of LSTM and XGBoost Prediction Results

The comparison results of the two models can be used to determine the effectiveness of the model in predicting wave heights in the Tuban Regency with four location coordinates. The comparison results of the two models are presented in **Table 3**

Table 3. Method Comparison Results

Location	Coordinate	LSTM			XGBoost		
		RMSE	MAE	MAPE	RMSE	MAE	MAPE
1st Location	-6.21;111.80	**0,045**	**0,029**	**8,671%**	0,046	0,031	9,022%
2nd Location	-6.22;112.40	0,051	**0,035**	**10,64%**	**0,050**	0,036	11,05%
3rd Location	-6.26;112.20	0,044	**0,027**	**7,773%**	**0,043**	0,028	8,160%
4th Location	-6.31;112.27	0,042	0,025	7,386%	**0,040**	**0,025**	**7,286%**

Based on Table 1, the results of the comparison of the two models in predicting wave heights at the first location show that the LSTM method is more optimal when compared to the XGBoost method, the RMSE value of the LSTM model is 0.045, MAE is 0.029 and MAPE is 8.671%, but the difference between the two methods is relatively small. **Fig. 19** shows the results of the comparison plot of the two methods, it can be seen that the prediction plot of the LSTM method has a value that tends to be lower when compared to the prediction results with the XGBoost method. However, the LSTM and XGBoost methods have fairly good accuracy results, this is known from the plots of the two methods which have almost the same pattern as the actual data plot.

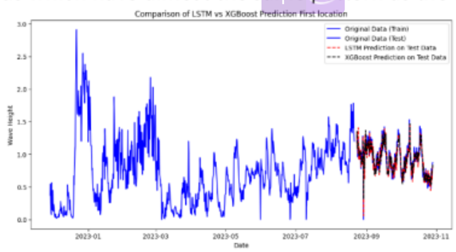


Figure 19. Comparison of LSTM and XGBoost at Location 1

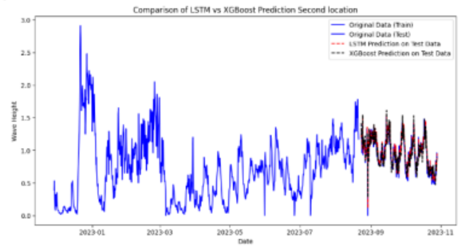


Figure 20. Comparison of LSTM and XGBoost at Location 2

The results of the comparison of the two models for the second location also have relatively small difference values for the RMSE, MAE, and MAPE values for the two methods of 0.001; 0.001; and 0.41, respectively. The predicted value of the LSTM method has MAE and MAPE values that tend to be better when compared to the XGBoost method, namely 0.035 and 10.64%. The XGBoost method has a better RMSE value when compared to LSTM which is 0.050. **Fig. 20** shows the results of the comparison plot of the two methods, It can be seen that the prediction plot of the XGBoost method has a value that tends to be higher when compared to the prediction results with the LSTM method. However, the LSTM and XGBoost methods have fairly good accuracy results, this is known from the plots of the two methods which have a pattern that is quite similar to the actual data plot of the second location.

The evaluation results of the LSTM method at the third location have a better value when compared to the XGBoost method. However, the difference between the evaluation results of the two methods is relatively small. The difference for RMSE, MAE, and MAPE values for both methods are 0.001; 0.001; and 0.387. From **Fig. 21**, the XGBoost prediction results show higher results in early September 2023 when compared to the actual data and LSTM prediction plots. However, from the results of the comparison plot of the two methods, it can be seen that both method plots have almost the same pattern and height.

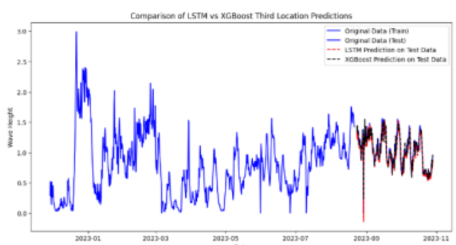


Figure 21. Comparison of LSTM and XGBoost at Location 3

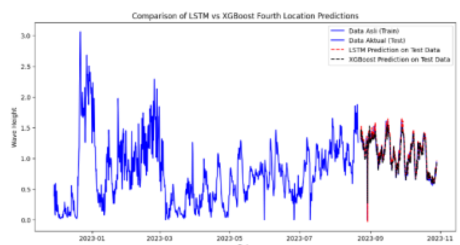


Figure 22. Comparison of LSTM and XGBoost at Location 4

The results of the comparison of the two models for the fourth location obtained the same MAE value between the two methods of 0.25, while the RMSE and MAPE values between the two methods have relatively small differences. The difference between the RMSE and MAPE values for the two methods is 0.002 and 0.100, respectively. The LSTM and XGBoost methods have quite good accuracy results, this is known from the plots of both methods (**Fig. 22**) which have almost the same pattern as the actual data plot. In addition, the plot results of both methods at the fourth location also show results that are more fitted to the actual data when compared to plots from the first, second, and third locations.

Conclusions

Based on the research results that have been explained, there are several conclusions in this research:

1. The LSTM method is superior in predicting wave height at the first location (RMSE 0.045; MAE 0.029; MAPE 8.671%), second location (RMSE 0.051; MAE 0.035; MAPE 10.64%) and third location (RMSE 0.044; MAE 0.027; MAPE 7.773%). Meanwhile, XGBoost gave better results on the fourth location data (RMSE 0.040; MAE 0.025;

MAPE 7.286). Even though they have relatively small differences, both methods have quite optimal results in predicting sea wave heights in Tuban Regency.

2. Prediction results for the next 30 days show that wave heights in all locations are below the safe limit set by BMKG, namely less than 1.25 meters. Therefore, fishermen in Tuban Regency can continue their shipping activities by paying attention to the safety guidelines provided.

Limitations

Based on the results that have been described, several suggestions can be recommended to readers or future researchers:

1. Future researchers can conduct comparisons by exploring prediction models other than LSTM and XGBoost or combining (hybrid) models that may be relevant and can provide better accuracy results.
2. Future researchers can add other factors such as rainfall, wind direction, and speed related to sea wave height, the use of other factors is expected to improve the results of accuracy or understanding related to the prediction of sea wave height.
3. Future researchers can involve spatial data such as seabed topography such as depth and sea slope. Integration of these data can provide a better understanding of the physical conditions of the region and increase the number of features that can be used by the model.

Ethics statements

As an expert scientist and along with co-authors of the concerned field, the paper has been submitted with full responsibility, following due ethical procedure, and there is no duplicate publication, fraud, plagiarism, or concerns about animal or human experimentation.

CRediT author statement

Riswanda Ayu Dhiya'ulhaq: Conceptualization, Formal analysis, Investigation, Writing— original draft, Writing—review and editing, Software, **Anisya Safira**: Conceptualization, Formal analysis, Investigation, Writing— original draft, Writing—review and editing, Software, **Indah Fahmiyah**: Supervision, Conceptualization, Methodology, Writing–original draft, Writing– review & editing, Validation, **Mohammad Ghani**: Supervision, Conceptualization, Methodology, Writing– original draft, Writing– review & editing, Validation.

Acknowledgments

The data source used in this research is secondary data obtained from Perak Meteorological Station Surabaya II

Declaration of interests

The author declares that the author does not have competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Adytia, D., Yonanta, A. R., & Subasita, N. (2019). Wind Wave Prediction by using Autoregressive Integrated Moving Average model : Case Study in Jakarta Bay. *International Journal on Information and Communication Technology (IJoICT)*, *4*(2), 33. https://doi.org/10.21108/ijoict.2018.42.300

[2] Anggoro, D. A., & Mukti, S. S. (2021). Performance Comparison of Grid Search and Random Search Methods for Hyperparameter Tuning in Extreme Gradient Boosting Algorithm to Predict Chronic Kidney Failure. *International Journal of Intelligent Engineering and Systems*, *14*(6), 198–207. https://doi.org/10.22266/ijies2021.1231.19

[3] Asmaul Husnah, Abdillah, A., Vera Mandailina, Syaharudin, S., & Mehmood, S. (2023). Wind Speed Regression Model in Forecasting Wave Height in the Shipping Channel Zone. *JST (Jurnal Sains Dan Teknologi)*, *12*(1), 30–38. https://doi.org/10.23887/jstundiksha.v12i1.50981

[4] Belete, D. M., & Huchaiah, M. D. (2022). Grid search in hyperparameter optimization of machine learning models for prediction of HIV/AIDS test results. *International Journal of Computers and Applications*, *44*(9), 875–886. https://doi.org/10.1080/1206212X.2021.1974663

[5] Campos, R. M., Costa, M. O., Almeida, F., & Soares, C. G. (2021). Operational wave forecast selection in the atlantic ocean using random forests. *Journal of Marine Science and Engineering*, *9*(3). https://doi.org/10.3390/jmse9030298

[6] Chang, Z., Zhang, Y., & Chen, W. (2018). Effective Adam-optimized LSTM Neural Network for Electricity Price

Forecasting. *IEEE International Conference on Software Engineering and Service Sciences (ICSESS)*, 245–248.

[7] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, *13-17-Augu*, 785–794. https://doi.org/10.1145/2939672.2939785

[8] de Amorim, L. B. V., Cavalcanti, G. D. C., & Cruz, R. M. O. (2023). The choice of scaling technique matters for classification performance. *Applied Soft Computing*, *133*, 1–37. https://doi.org/10.1016/j.asoc.2022.109924

[9] Ding, E. (2022). Regularization. In *Data Science Resources* (pp. 1–8). https://s3.amazonaws.com/kajabi-storefronts-production/file-uploads/sites/2147512189/themes/2150624317/downloads/5355dc-82-8b2-2b53-dd6a85bcb82_Regularization.pdf

[10] Eni. (2017). Proceeding The 1st International Seminar on Sustainability in the Marine Fisheries Sector 2017.Establishing Sustainable Marine and Fisheries Sector to Support Food Security Within ASEAN Economic Community Framework. *Angewandte Chemie International Edition, 6(11), 951–952.*, *Mi*, 5–24.

[11] Ghosh, A., Bose, S., Maji, G., Debnath, N. C., & Sen, S. (2019). Stock price prediction using lstm on indian share market. *EPiC Series in Computing*, *63*(September 2019), 101–110. https://doi.org/10.29007/qgcz

[12] Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

[13] Hu, H., van der Westhuysen, A. J., Chu, P., & Fujisaki-Manome, A. (2021). Predicting Lake Erie wave heights and periods using XGBoost and LSTM. *Ocean Modelling*, *164*(November 2020), 101832. https://doi.org/10.1016/j.ocemod.2021.101832

[14] ILO. (2016). Fishers first - Good practices to end labour exploitation at sea. *International Labour Office, Fundamental Principles and Rights at Work Branch (FUNDAMENTALS), Sectoral Policies Department (SECTOR), Geneva*, 82.

[15] Ismanto, E., & Effendi, N. (2023). An LSTM-based prediction model for gradient-descending optimization in virtual learning environments. *Computer Science and Information Technologies*, *4*(3), 199–207. https://doi.org/10.11591/csit.v4i3.p199-207

[16] Jabbar, H. K., & Khan, R. Z. (2015). *Methods to Avoid Over-Fitting and Under-Fitting in Supervised Machine Learning (Comparative Study). December 2014*, 163–172. https://doi.org/10.3850/978-981-09-5247-1_017

[17] Marco, P. (2022). Time Series Forecasting in Python. Simon and Schuster.

[18] Malik, S., Harode, R., & Singh Kunwar, A. (2020). *XGBoost: A Deep Dive into Boosting Feb 3 · 12 min read. February*. https://doi.org/10.13140/RG.2.2.15243.64803

[19] Meng, Z. F., Chen, Z., Khoo, B. C., & Zhang, A. M. (2022). Long-time prediction of sea wave trains by LSTM machine learning method. *Ocean Engineering*, *262*(August), 112213. https://doi.org/10.1016/j.oceaneng.2022.112213

[20] Murty, S. V., & Kiran Kumar, R. (2019). Accurate liver disease prediction with extreme gradient boosting. *International Journal of Engineering and Advanced Technology*, *8*(6), 2288–2295. https://doi.org/10.35940/ijeat.F8684.088619

[21] Prakash, K., Kannan, R., S.A, A., & G.R, K. (2021). Advanced Deep Learning for Engineers and Scientists. In *EAI/Springer Innovations in Communication and Computing.* . https://books.google.com.pk/books?id=MDsNzgEACAAJ%0Ahttps://link.springer.com/10.1007/978-3-030-66519-7

[22] Pramesti, D. D., Novitasari, D. C. R., Setiawan, F., & Khaulasari, H. (2022). Long-Short Term Memory (Lstm) for Predicting Velocity and Direction Sea Surface Current on Bali Strait. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, *16*(2), 451–462. https://doi.org/10.30598/barekengvol16iss2pp451-462

[23] Putri Udiani, N. M. R., Darma Putra, I. K. G., & Arya Sasmita, G. M. (2020). Forecasting of Arabica Coffee Production in Bali Province Using Support Vector Regression. *International Journal of Computer Application Technology & Research*, *9*(2), 41–46. https://doi.org/10.7753/ijcatr0902.1001

[24] Staudemeyer, R. C., & Morris, E. R. (2019). *Understanding LSTM -- a tutorial into Long Short-Term Memory Recurrent Neural Networks*. 1–42. http://arxiv.org/abs/1909.09586

[25] Tarwidi, D., Pudjaprasetya, S. R., Adytia, D., & Apri, M. (2023). An optimized XGBoost-based machine learning method for predicting wave run-up on a sloping beach. *MethodsX*, *10*(December 2022), 102119. https://doi.org/10.1016/j.mex.2023.102119

[26] Vijayaprabakaran, K., & Sathiyamurthy, K. (2022). Towards activation function search for long short-term model network: A differential evolution based approach. *Journal of King Saud University - Computer and Information Sciences*, *34*(6), 2637–2650. https://doi.org/10.1016/j.jksuci.2020.04.015

MethodsX Riswanda Safira Indah Ghani

ORIGINALITY REPORT

18%	11%	13%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Tsukasa Hokimoto. "Chapter 14 Prediction of Wave Height Based on the Monitoring of Surface Wind", IntechOpen, 2012 Publication	2%
2	www.mdpi.com Internet Source	1%
3	Muhammad Ikhlas, Lika Jafnihirda. "Comparative Analysis of Strategic Location Selection Decisions for MSMEs (UMKM) Using the MFEP and SAW Method", 2021 International Conference on Computer Science and Engineering (IC2SE), 2021 Publication	1%
4	www.researchsquare.com Internet Source	1%
5	ojs3.unpatti.ac.id Internet Source	1%
6	blog.csdn.net Internet Source	1%
7	export.arxiv.org	

Internet Source

1%

8	Submitted to University of Northampton	Student Paper	<1%
9	Submitted to University of Sydney	Student Paper	<1%
10	www.frontiersin.org	Internet Source	<1%
11	journal.ugm.ac.id	Internet Source	<1%
12	ejurnal.ung.ac.id	Internet Source	<1%
13	hdl.handle.net	Internet Source	<1%
14	iptek.its.ac.id	Internet Source	<1%
15	Kaoutar Dachri, Mohamed Bouabidi, Khalid Naji, Kalthoum Nouar et al. "Predictive insights for copper recovery: A synergistic approach integrating variability data and machine learning in the geometallurgical study of the Tizert deposit, Morocco", Journal of African Earth Sciences, 2024	Publication	<1%

16	Wahyutama Fitri Hidayat, Muhammad Fahmi Julianto, Yesni Malau, Ahmad Setiadi, Sriyadi. "Implementation of LSTM and Adam Optimization as a Cryptocurrency Polygon Price Predictor", 2023 International Conference on Information Technology Research and Innovation (ICITRI), 2023 Publication	<1 %
17	Submitted to Angeles University Foundation Student Paper	<1 %
18	Submitted to National University of Singapore Student Paper	<1 %
19	Qian Zhang, Junjie Wu, Yuan Ma, Guoli Li, Jinhui Ma, Can Wang. "Short-term load forecasting method with variational mode decomposition and stacking model fusion", Sustainable Energy, Grids and Networks, 2022 Publication	<1 %
20	Submitted to University of Nottingham Student Paper	<1 %
21	Submitted to Australian Institute of Higher Education Student Paper	<1 %
22	Submitted to Concordia University Student Paper	<1 %

23	Destinee M. Ditton, Colin R. Marchus, Aimee L. Bozeman, Alleyna C. Martes, Michele R. Brumley, Nathan R. Schiele. "Visualization of rat tendon in three dimensions using micro-Computed Tomography", MethodsX, 2024 Publication	<1 %
24	Lian Liu, Santhi Davedu, Ayumi Fujisaki-Manome, Haoguo Hu, Christiane Jablonowski, Philip Y. Chu. "Machine Learning Model-Based Ice Cover Forecasting for a Vital Waterway in Large Lakes", Journal of Marine Science and Engineering, 2022 Publication	<1 %
25	arxiv.org Internet Source	<1 %
26	repository.upy.ac.id Internet Source	<1 %
27	etheses.lib.ntust.edu.tw Internet Source	<1 %
28	ltu.diva-portal.org Internet Source	<1 %
29	Jie Ding, Fangyu Deng, Qi Liu, Jichao Wang. "Regional forecasting of significant wave height and mean wave period using EOF-EEMD-SCINet hybrid model", Applied Ocean Research, 2023 Publication	<1 %

30	Mohammad Azad, Md. Alhaz Uddin. "Prediction of Offshore Wave at East Coast of Malaysia—A Comparative Study", Electronics, 2022 Publication	<1 %
31	Xiaozhou Lei, Zixiang Fei, Wenju Zhou, Huiyu Zhou, Minrui Fei. "Low-Light Image Enhancement Using the Cell Vibration Model", IEEE Transactions on Multimedia, 2023 Publication	<1 %
32	Yehezkiel K. A. Purba, Deni Saepudin, Didit Adytia. "Prediction of Sea Level by Using Autoregressive Integrated Moving Average (ARIMA): Case Study in Tanjung Intan Harbour Cilacap, Indonesia", 2020 8th International Conference on Information and Communication Technology (ICoICT), 2020 Publication	<1 %
33	iaesprime.com Internet Source	<1 %
34	Dalei Song, Miao Yu, Zhifeng Wang, Xuegong Wang. "Wind and wave energy prediction using an AT-BiLSTM model", Ocean Engineering, 2023 Publication	<1 %
35	Senthil Murugan Nagarajan, V. Muthukumaran, R. Murugesan, Rose Bindu	<1 %

Joseph, Munirathanam Meram, A. Prathik. "Innovative feature selection and classification model for heart disease prediction", Journal of Reliable Intelligent Environments, 2021
Publication

36	ebin.pub Internet Source	<1 %
37	repository.riteh.uniri.hr Internet Source	<1 %
38	unsworks.unsw.edu.au Internet Source	<1 %
39	voycn.com Internet Source	<1 %
40	Abbas Yeganeh-Bakhtiary, Hossein EyvazOghli, Naser Shabakhty, Soroush Abolfathi. "Machine learning prediction of wave characteristics: Comparison between semi-empirical approaches and DT model", Ocean Engineering, 2023 Publication	<1 %
41	Eka Pandu Cynthia, Suzani Binti Mohamad Samuri, Wang Shir Li, Edi Ismanto, Liza Afriyanti, M. Imam Arifandy. "Improved Machine Learning Algorithm for Heart Disease Prediction Based on Hyperparameter Tuning", 2023 IEEE International Conference	<1 %

on Artificial Intelligence in Engineering and Technology (IICAIET), 2023
Publication

42	Hongbo Ding, Xue Feng, Qi Yang, Yichang Yang et al. "A Risk Prediction Model for Efficient Intubation in the Emergency Department: A Five-Year Single-Center Retrospective Analysis", Research Square Platform LLC, 2023 Publication	<1 %
43	Jaimin Shah, Darsh Vaidya, Manan Shah. "A comprehensive review on multiple hybrid deep learning approaches for stock prediction", Intelligent Systems with Applications, 2022 Publication	<1 %
44	Zhe Wang, Filippo Ubertini, Simon Laflamme. "Ensemble of long short-term memory recurrent neural network for semi-active control of tuned liquid wall damper", Engineering Structures, 2022 Publication	<1 %
45	Zhijian Qu, Lin Zhu, Shuaijun Ma, Boyu Zhang. "Remaining Useful Life Prediction of High-Speed Railroad Contact Network Based on Stacking Integrated Attention-LSTM-CNN Deep Learning", Arabian Journal for Science and Engineering, 2024 Publication	<1 %

46	dr.ntu.edu.sg Internet Source	<1 %
47	inass.org Internet Source	<1 %
48	repository.uam.ac.id Internet Source	<1 %
49	www.igpm.rwth-aachen.de Internet Source	<1 %
50	Nurul Asyikin Mohd Najib, Vasanthi Sethu, Senthil Kumar Arumugasamy, Anurita Selvarajoo. "Artificial Neural Network (ANN) Modelling of Palm Oil Mill Effluent (POME) Treatment with Natural Bio-coagulants", Environmental Processes, 2020 Publication	<1 %
51	"Advanced Deep Learning for Engineers and Scientists", Springer Science and Business Media LLC, 2021 Publication	<1 %
52	Qin-Rui Luo, Hang Xu, Long-Hu Bai. "Prediction of significant wave height in hurricane area of the Atlantic Ocean using the Bi-LSTM with attention model", Ocean Engineering, 2022 Publication	<1 %

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		

Missing "," Review the rules for using punctuation marks.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Missing "," Review the rules for using punctuation marks.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

S/V This subject and verb may not agree. Proofread the sentence to make sure the subject agrees with the verb.

P/V You have used the passive voice in this sentence. You may want to revise it using the active voice.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

PAGE 2

Article Error You may need to use an article before this word.

S/V This subject and verb may not agree. Proofread the sentence to make sure the subject agrees with the verb.

S/V This subject and verb may not agree. Proofread the sentence to make sure the subject agrees with the verb.

Article Error You may need to use an article before this word.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Missing "," Review the rules for using punctuation marks.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Word Error Did you type **the** instead of **they**, or have you left out a word?

Article Error You may need to remove this article.

Proper Nouns You may need to use a capital letter for this proper noun.

Proofread This part of the sentence contains an error or misspelling that makes your meaning unclear.

Run-on This sentence may be a run-on sentence.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Article Error You may need to use an article before this word.

Article Error You may need to use an article before this word.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Article Error You may need to use an article before this word.

Article Error You may need to use an article before this word.

PAGE 6

Article Error You may need to remove this article.

Proofread This part of the sentence contains an error or misspelling that makes your meaning unclear.

Article Error You may need to use an article before this word.

Run-on This sentence may be a run-on sentence.

Article Error You may need to use an article before this word.

Article Error You may need to use an article before this word.

PAGE 7

Article Error You may need to use an article before this word. Consider using the article **the**.

Article Error You may need to use an article before this word.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Article Error You may need to use an article before this word.

Missing "," Review the rules for using punctuation marks.

Missing "," Review the rules for using punctuation marks.

PAGE 8

Article Error You may need to use an article before this word.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Article Error You may need to use an article before this word.

Article Error You may need to remove this article.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Missing "," Review the rules for using punctuation marks.

Missing "," Review the rules for using punctuation marks.

PAGE 9

Article Error You may need to use an article before this word. Consider using the article **the**.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Missing "," Review the rules for using punctuation marks.

Missing "," Review the rules for using punctuation marks.

Article Error You may need to remove this article.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Article Error You may need to use an article before this word.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Run-on This sentence may be a run-on sentence.

Article Error You may need to remove this article.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Run-on This sentence may be a run-on sentence.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

P/V You have used the passive voice in this sentence. You may want to revise it using the active voice.

Article Error You may need to use an article before this word.

Missing "," Review the rules for using punctuation marks.

Missing "," Review the rules for using punctuation marks.

Article Error You may need to use an article before this word.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.

Sp. This word is misspelled. Use a dictionary or spellchecker when you proofread your work.