

Improving Prediction of Dengue Outbreaks Using Attention-based LSTM Model with Honey Badger Optimization for Hyperparameter Tuning

Pudadera, Sombero, Dollaga, Sueno

Computer Studies Department Notre Dame of Marbel University Koronadal City, South Cotabato, Philippines

DOI: <https://doi.org/10.51583/IJLTEMAS.2026.150600192>

Received: 05 July 2026; Accepted: 10 July 2026; Published: 21 July 2026

ABSTRACT

Climate Change Poses a Significant Challenge to the Current Dynamics of Disease Outbreaks. This Study Improves Outbreak Prediction Using an Attention-Based LSTM Model Optimized by the Honey Badger Algorithm (HBA) for Hyperparameter Tuning.

Using Disease, Climate, and Geographic Data From 2015–2024 in Different Barangays in Koronadal, South Cotabato, the Model Predicts Incidence Over 1-, 3-, 6-, and 12-Month Horizons. Attention Mechanisms Enhanced Long-Term Pattern Detection, While HBA Reduces Overfitting and Boosts Accuracy. Results Show the HBA-LSTM Reduces Mean Squared Error by 43.7% Over Standard LSTM and 22.2% Over Attention Models. Similar Reductions are Seen in RMSE, MAE, and MAPE. Though Effective, Further Tuning and Alternative Architectures are Suggested for Improved Generalization.

Keywords: LSTM, overfitting, attention mechanism, deep learning, honey badger, optimization, fine-tuning

INTRODUCTION

Climate change, characterized by increasing global temperatures, alterations in precipitation patterns, and a higher frequency of extreme weather events, alters the distribution and transmission of pathogens, vectors, and hosts. Research highlights how climatic factors, such as the Indian Ocean Dipole (IOD) and the El Niño-Southern Oscillation (ENSO), impact the seasonal patterns of vector-borne diseases like dengue and malaria. [3].

Developing predictive models that effectively utilize climate data is essential, as they significantly enhance public health interventions. Deep learning is particularly well-suited for processing complex, high-dimensional data, making it ideal for modeling interactions between climate and disease. Auto-Regressive Integrated Moving Average (ARIMA) and Gated Recurrent Units (GRU) have been applied to predictive disease outbreaks using climate data ([23]; [33]).

LSTM models are subsets of Recurrent Neural Networks that overcome vanishing gradients through their memory cells. However, they face challenges in overfitting and long-term prediction accuracy. In stock market return forecasting, LSTM models demonstrate potential but still struggle with generalization to unseen data, indicating significant overfitting([32]).

Overfitting can be addressed through various techniques, including grid search and L2 regularization ([7]) and ant colony optimization for structure optimization ([18]). The attention mechanism, on the other hand, allows the model to focus on essential time steps, which could reduce noise and improve the model's generalizability [37]. However, it can still suffer from longer sequences due to issues with backpropagation through time and the selection of hyperparameters.

Honey Badger Optimization (HBO) offers a promising solution to overcome these limitations. Inspired by the adaptive foraging behavior of honey badgers, HBO effectively balances exploration and exploitation. This optimization technique fine-tunes model parameters to search for relevant values within the search space. When faced with local optima, it can escape from these traps, enhancing the overall optimization process ([13]:[14]). Since outbreak data is irregular and non-stationary, attention mechanisms may not effectively capture these trends. The researchers propose an Attention-based LSTM optimized for hyperparameters using the Honey Badger Algorithm. This approach aims to improve forecasting with lags of 1, 3, 6, and 12 months, comparing it to stacked LSTM and Attention stacked-LSTM models.

REVIEW OF RELATED LITERATURE

Relationship between Climatic and Geographic Features in disease incidence

Climatic and geographic factors significantly influence the incidence of vector-borne diseases, such as dengue fever, malaria, and chikungunya. Temperature, humidity, and

rainfall have a direct impact on mosquito breeding and the transmission of viruses. Higher temperatures accelerate mosquito development, increased rainfall creates more breeding sites, and higher humidity extends mosquito lifespan and infection rates ([21]; [6]).

Long Short-Term Memory

LSTM (Long Short-Term Memory) models are effective for predicting zoonotic diseases like schistosomiasis, echinococcosis, and leptospirosis due to their ability to retain relevant historical information while filtering out noise. Integrating LSTM with time-series data significantly enhances prediction accuracy. Metrics such as RMSE, MAE, and MSE demonstrate that these hybrid models reduce prediction errors and generate smoother incidence curves that align closely with actual data. (**Error! Reference source not found.**)

Attention Mechanism

Recent research on spatiotemporal attention mechanisms has enhanced prediction accuracy in multivariate time series forecasting by capturing complex dependencies among variables. Models like LSTNet utilize spatiotemporal self-attention for improved results.[29] Advances such as STAM ([11]) and symmetric kernel attention mechanisms ([25]) further refine long— and short-term dependencies. Integrating attention modules into LSTM models improves memory retention and predictive accuracy ([22]), while multi-head attention enhances spatiotemporal relationship analysis ([36]).

Hyperparameter Optimization in Prediction

While manual tuning relies on domain expertise, it is inefficient for large search spaces, and automated methods like grid search and random search suffer from slow performance. Studies show HBO outperforms traditional methods in optimizing CNNs and LSTMs, reducing training time while improving model performance ([15]), highlighting its potential in developing robust predictive models.

Honeybadger Optimization

The Honey Badger Algorithm (HBA) is a bio-inspired metaheuristic optimization technique that mimics honey badgers' foraging behavior to balance exploration and exploitation, making it effective for solving complex optimization problems ([15]; [14]). In deep learning, HBA fine-tunes weights, minimizes loss, and enhances generalization. The Global Optimization Honey Badger Algorithm (GOHBA) boosts convergence and search diversity by optimizing hyperparameters like learning and dropout rates to prevent overfitting [16].

Model Evaluation and Validation Metric

Mean Absolute Error

The Mean Absolute Error (MAE) measures the average absolute difference between actual and predicted cases. It provides an intuitive, easily interpretable metric that quantifies how much the model's predictions deviate from fundamental values [28].

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (1)$$

Mean Squared Error

MSE in a regression task squares the errors before averaging, giving more weight to larger deviations. This is particularly useful in outbreak prediction, as it penalizes extreme errors, ensuring the model is sensitive to sudden spikes in disease cases. [31].

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2)$$

Mean Absolute Percentage Error (MAPE)

Root Mean Square Error (RMSE) is obtained by taking the square root of the Mean Square Error (MSE). This transformation makes the error more interpretable because it expresses the error in the same units as the original data.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100 \quad (3)$$

Root Mean Squared Error (RMSE)

RMSE, derived from MSE by taking the square root, maintains interpretability while penalizing larger errors more heavily.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (4)$$

Concept of the Study

Data Acquisition

Disease data was sourced from the Department of Health Center for Health Development SOCCSKSARGEN, while climatic and geographic data were obtained from Google Earth via JavaScript Script Functions. NAMRIA provided shapefiles for Koronadal barangay boundaries.

Data Preprocessing

Irrelevant columns were removed from the dataset to focus solely on the relevant features. Since there were no missing values, imputation was not required. Categorical variables were converted to numerical values using label encoding with the sci-kit-learn library. These preprocessing steps effectively structured the dataset for training purposes.

Feature Engineering

Key climatic factors such as temperature, precipitation, and humidity, along with geographic features like Land Surface Temperature and NDVI, were preserved. Lagged outbreak cases were created at 1, 3, 6, and 12-month

intervals to capture temporal dependencies.

Feature Scaling

LSTM models are sensitive to input values, so feature scaling was used to ensure equal contribution from all variables. MinMaxScaler normalized input features to a range of [0,1].

Materials

I. METHODOLOGY

$$x_{\text{normalized}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (4)$$

Train-Test Split

The dataset was split into 70% for training, 15% for validation, and 15% for testing. The training and validation sets were used to fit the model and track hyperparameters, while the test set evaluated the model's generalization to unseen data by monitoring training and validation loss.

Model Development

The model consists of three main models for comparison: the Stacked-LSTM, Stacked-LSTM-Att, Stacked-LSTM-Att-HBO model. This was designed to process disease, climatic, and geographic data separately using distinct LSTM layers. The outputs of these layers were then combined and passed through fully connected dense layers, which generated the final predictions.

Hyperparameter Optimization

The Honey Badger Optimization (HBO) algorithm was used to fine-tune the model's hyperparameters, reducing prediction errors. HBO optimizes parameters through two main phases:

Exploration Phase: The algorithm applies a digging formula to search broadly within the solution space, identifying a wide range of potential hyperparameter combinations.

$$[.oneX_i^{t+1} = X^* + F \cdot \alpha \cdot rand \cdot X^*] \quad (5)$$

Where:

- X^* is the current best solution
- $F \in [-1, 1]$ is a random direction factor
- α is a control parameter that decreases over iterations
- $rand$ is a random vector.

Exploitation Phase: A honey formula is used to refine the search, focusing on the most promising solutions (Hyperparameters). Key climatic factors such as temperature, precipitation, and humidity, along with geographic features like elevation and NDVI, were preserved.

$$[d_i = rand \cdot \beta \cdot I \cdot X^* + F \cdot rand \cdot \alpha \cdot X_i^t][X_i^{t+1} = X^* + d_i] \quad (6)$$

Where:

- β is a scaling parameter,

I is the intensity, calculated as: $[I = rand()]$

This study utilized personal computers with Intel® Core™ i5 processors (2.4 GHz) and Apple M1 chips, offering sufficient power for data preprocessing, model training, and validation.

Software

Python was chosen for its versatility and libraries like Pandas, NumPy, Matplotlib, TensorFlow, and Keras, which streamline these tasks. Google Colab served as the primary development platform, providing free GPU resources for faster model training and easy integration with Google Drive and GitHub.

Data

Infectious disease data was sourced from the Department of Health Center for Health Development SOCCSKSARGEN, while climatic and geographic information came from Google Earth. Land Surface Temperature (LST) was extracted from MOD21C3, and climate variables, including temperature and precipitation, were gathered from ERA5-Land. The Normalized Difference Vegetation Index (NDVI) was derived from Landsat 8/9, and barangay boundary information was provided by NAMRIA shapefiles.

Procedures

Pre-processing

Data from the DOH, Google Earth Engine, and NAMRIA were combined using a Python script to create a single dataset. These data were aggregated into monthly values to ensure compatibility for training the model. A total of 2918 rows and 19 columns are prepared for the model. Selected barangays were only chosen due to risk in context of outbreak. Irrelevant columns were eliminated, leaving only the pertinent features. As there were no missing values in the dataset, no imputation was needed. Categorical variables were transformed into numerical values

1) Building the Model

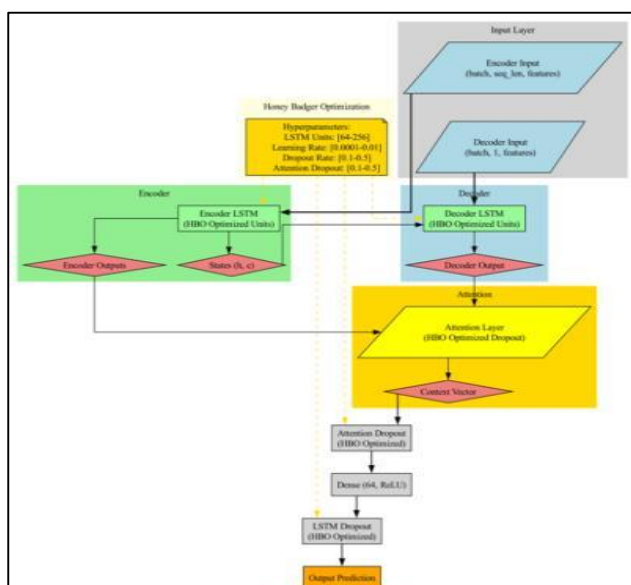


Fig. 1. LSTM Attention Framework with Honey Badger Optimization

Input Layers: Each input sequence represents a historical time window of monthly data for a barangay. The features include feature-engineered variables such as cases from previous months, and climatic and geographic data encoded in the monthly format as input.

Encoder Layer (LSTM): The Encoder LSTM layer processes the input sequence, learning temporal dependencies between past disease cases, climate, and geography. The final hidden and cell states (h , c) from the encoder are passed to the decoder. The encoder outputs store feature-rich representations for attention processing.

Decoder Layer (LSTM Layer): The Decoder LSTM layer generates future predictions based on the encoder's final states. It processes the decoder input step by step, predicting the next month's disease cases.

Attention Layer: The attention Layer assigns different importance levels to the encoder's outputs, allowing the model to focus on the most relevant past months. The context vector is computed based on the weighted outputs of the encoder. This mechanism helps the decoder make better predictions by focusing on past trends that are highly correlated with outbreaks.

Fully Connected Layers & Prediction: To enhance generalization and mitigate overfitting, the model incorporates dropout layers at multiple stages, including an attention dropout and an LSTM dropout. A fully connected dense layer and a rectified linear activation function further refine the extracted features before generating the final output.

LSTM-Attention Model : After the output from the LSTM-Attention Model, it will be trained using the default hyperparameter for the LSTM Model.

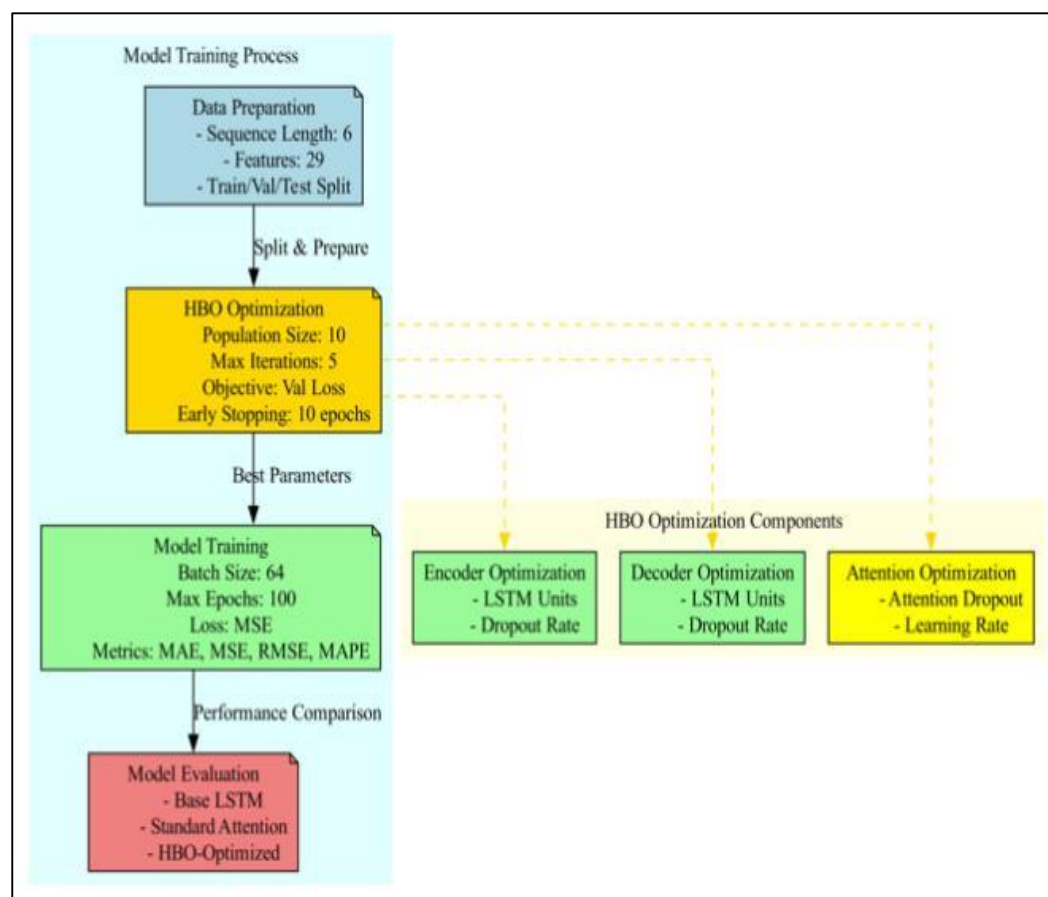


Fig. 2. Honey Badger Optimization Method

HBO for Hyperparameter Tuning

Exploration Phase: The HBO algorithm starts by generating a diverse population of candidate hyperparameter sets (LSTM units, learning rate, dropout, attention dropout, embedding dimension). Each candidate is evaluated by training an Attention LSTM model and measuring its validation loss (fitness). Randomness in candidate generation and updates ensures the search covers a broad region of the

hyperparameter space, helping to avoid getting stuck in local minima.

Exploitation Phase: In the exploitation phase, the focus is on fine-tuning hyperparameters through localized adjustments and reduced randomness. This involves refining the encoder and decoder LSTM units, as well as the dropout rates and attention parameters, to achieve optimal model performance. By utilizing smaller step sizes and targeted changes, the algorithm efficiently converges while balancing exploration and optimization, resulting in a solid set of hyperparameters for final model training and evaluation.

RESULTS & DISCUSSIONS

TABLE I. Performance Metrics of models on lookback

Lookback window	Model	MSE	RMSE	MAE	MAPE ($\times 10^{16}$)
1 months	LSTM	24.67	4.95	2.98	8.77
	LSTM-Attention	17.83	4.21	2.53	7.38
	HBO LSTM Attention	13.88	3.72	2.23	6.60
3 months	LSTM	18.71	4.24	2.66	18.17
	LSTM-Attention	13.51	3.60	2.26	15.37
	HBO LSTM Attention	10.51	3.18	1.99	13.68
6 months	LSTM	28.04	5.19	2.94	10.30
	LSTM-Attention	20.28	4.42	2.51	8.71
	HBO LSTM Attention	15.75	3.89	2.20	7.71
12 months	LSTM	22.49	4.73	3.78	42.51
	LSTM-Attention	16.25	4.02	3.22	36.05
	HBO LSTM Attention	12.65	3.57	2.84	31.89

The HBO-LSTM Attention model consistently outperforms both S-LSTM and Attention models across all lookback windows (1, 3, 6, and 12 months) in terms of error metrics, achieving the lowest values for MSE, RMSE, MAE, and MAPE. Specifically, in the 1-month lookback, HBO reduced MSE by 43.7% compared to S-LSTM and by 22.2% compared to the Attention model. Similar trends are observed in all error metrics, with the 3-month lookback showing a 43.8% reduction in MSE versus LSTM. This pattern of superior performance continues through the 6-month and 12-month windows.

Fig. 3. Time Series Visualization of S-LSTM, S-Att-LSTM-HBO LSTM

The sample plot for Barangay Zone II shows that during critical periods, actual case numbers frequently spike to between 25 and 35 cases, while the predictions (represented by colored lines) rarely exceed 10 to 15 cases. Although the HBO-optimized Attention LSTM model (shown in red) shows slight improvements, achieving the lowest RMSE values across all time windows (4.04, 3.34, 4.30, and 3.15), this improvement is not enough to accurately capture the extreme variability in dengue transmission.

TABLE II. Best HBO configurations per lookback window on hyperparameter optimization

Window Size	HBO Bounds	HBO Parameters	Best Configuration	Optimization Time (s)	Best Test RMSE	Best Test MAE	Best Test MAP
(1)	Dropout: [0.2,0.4], Att Dropout: [0.2,0.4], Embedding Dim: [64,256] Learning Rate: [0.0001,0.005] LSTM Units: [128,512]	c : 1.2 beta : 3 max_iterations: 15 population_size: 20	Dropout: 0.36, Att Dropout: 0.26, Emb Dim: 153, LR: 0.0023, LSTM Units: 141	395	21.58	17.38	62.76
(3)	Dropout: [0.2,0.4], Att Dropout: [0.2,0.4], Embedding Dim: [64,256] Learning Rate: [0.0001,0.005] LSTM Units: [128,512]	c : 1.2 beta : 3 max_iterations: 15 population_size: 20	Dropout: 0.25, Att Dropout: 0.38, Emb Dim: 78, LR: 0.0035, LSTM Units: 201	482	22.68	17.84	52.99
(6)	Dropout: [0.2,0.4], Att Dropout: [0.2,0.4], Embedding Dim: [64,256] Learning Rate: [0.0001,0.005]	c : 1.2 beta : 6 max_iterations: 5 population_size: 20	Dropout: 0.36, Att Dropout: 0.23, Emb Dim: 216, LR: 0.0041,	787	21.19	19.29	59.15

	LSTM Units: [128,512]	ze: 10	LSTM Units: 467				
(12)	Dropout: [0.2,0.4], Att Dropout: [0.2,0.4], Embedding Dim: [64,256] Learning Rate: [0.0001,0.005]	c : 1.2 beta : 6 max_iteratio ns: 5 population_si ze: 10	Dropout: 0.22, Att Dropout: 0.20, Emb Dim: 151, LR: 0.0007, LSTM Units: 128	707	26.61	22.2	59.47
	LSTM [128,512]						

Table 2 shows the optimal Honey Badger Optimization configurations across various lookback windows. The analysis demonstrates a strategic tradeoff between window size and optimization parameters, with smaller windows (1, 3) requiring more extensive exploration (population size: 20, iterations: 15), while larger windows (6, 12) perform better with focused search strategies (population size: 10, iterations: 5). Performance metrics vary significantly across configurations, with window size 6 achieving superior prediction accuracy (RMSE: 21.19), window size 1 minimizing absolute error (MAE: 17.38), and window size 3 offering the most proportionally accurate forecasts (MAPE: 52.99%). However, computational demands escalate with window size, evidenced by increasing optimization times (395s for window one vs. 787s for window 6), reflecting the complexity of modeling longer temporal dependencies. The optimal embedding dimensions exhibit substantial variation, ranging from 78 to 216, indicating that representation learning requirements vary significantly based on the temporal context length.

CONCLUSION AND RECOMMENDATION

The evaluation of various deep learning approaches for forecasting Dengue using LSTM forecasting models provides valuable insights into the potential and limitations of predictive epidemiological modeling, especially using hyperparameter optimizations. It highlights fundamental constraints within the current modeling methods, likely stemming from a lack of sufficient data on non-linear interactions among environmental factors, vector dynamics, and human movement patterns.

These results underscore the need for better modeling techniques that consider additional factors, such as climate variability, population immunity, and vector abundance. Using multi-modal ensemble frameworks could also improve understanding of the complex relationships driving dengue transmission.

The researchers recommend conducting a more detailed analysis of outbreak dynamics, particularly focusing on non-linear relationships such as mobility data, internet trends, and changes in land use.

When modeling, especially with deep learning models, it is crucial to choose the appropriate hyperparameters and architecture to ensure that the model's accuracy is clearly defined. Explore using different architectures, such as hybrid ensemble models or transformer models.

REFERENCES

1. Alangari, N., El Bachir Menai, M., Mathkour, H., & Almosallam, I. (2023). Exploring Evaluation Methods for Interpretable Machine Learning: A Survey. *Information* 2023, Vol. 14, Page 469, 14(8), 469.
2. <https://doi.org/10.3390/INFO14080469>
3. Barboza, M. F. X., Monteiro, K. H. de C., Rodrigues, I. R., Santos, G. L., Monteiro, W. M., Figueira, E. A. G., Sampaio,

4. V. de S., Lynn, T., & Endo, P. T. (2022). Prediction of malaria using deep learning models: A case study on city clusters in the state of Amazonas, Brazil, from 2003 to 2018. *Revista Da Sociedade Brasileira de Medicina Tropical*, 55, e0420-2021. <https://doi.org/10.1590/0037-8682-0420-2021>
5. Barrera, R., Acevedo, V., Amador, M., Marzan, M., Adams,
6. L. E., & Paz-Bailey, G. (2023). El Niño Southern Oscillation (ENSO) effects on local weather, arboviral diseases, and dynamics of managed and unmanaged populations of *Aedes aegypti* (Diptera: Culicidae) in Puerto Rico. *Journal of Medical Entomology*, 60(4), 796.
<https://doi.org/10.1093/JME/TJAD053>
8. Bellone, R., & Failloux, A. B. (2020). The Role of Temperature in Shaping Mosquito-Borne Viruses Transmission. *Frontiers in Microbiology*, 11, 584846. <https://doi.org/10.3389/FMICB.2020.584846/PDF>
9. Bergstra, J., & Bengio, Y. (2012). Random Search for Hyper-Parameter Optimization Yoshua Bengio. In *Journal of Machine Learning Research* (Vol. 13). <http://scikit-learn.sourceforge.net>.
10. Brown, J. J., Pascual, M., Wimberly, M. C., Johnson, L. R., & Murdock, C. C. (2023). Humidity – The overlooked variable in the thermal biology of mosquito-borne disease. *Ecology Letters*, 26(7), 1029. <https://doi.org/10.1111/ELE.14228>
11. Budiarto, Z., Listiyono, H., & Karim, A. (2024). Optimizing LSTM with Grid Search and Regularization Techniques to Enhance Accuracy in Human Activity Recognition. *Journal of Applied Data Sciences*, 5(4), 2002–2014. <https://doi.org/10.47738/jads.v5i4.433>
12. Chowdhury, A. A., Das, A., Hoque, K. K. S., & Karmaker, D. (2022). A Comparative Study of Hyperparameter Optimization Techniques for Deep Learning (pp. 509–521). https://doi.org/10.1007/978-981-19-0332-8_38
13. Chen, C., He, Z., Zhao, J., Zhu, X., Li, J., Wu, X., Chen, Z.,
14. Chen, H., & Jia, G. (2024). Zoonotic outbreak risk prediction with long short-term memory models: a case study with schistosomiasis, echinococcosis, and leptospirosis. *BMC Infectious Diseases*, 24(1), 1062.
<https://doi.org/10.1186/s12879-024-09892-y>
16. Dianne Ligue, K. B., & Joy Ligue, K. B. (n.d.-a). Deep Learning Approach to Forecasting Dengue Cases in Davao City Using Long Short-term Memory (LSTM). <https://doi.org/https://doi.org/10.56899/151.03.01>
17. Gangopadhyay, T., Tan, S. Y., Jiang, Z., Meng, R., & Sarkar,
18. S. (2020). Spatiotemporal Attention for Multivariate Time Series Prediction and Interpretation. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2021-June, 3560–3564.
<https://doi.org/10.1109/ICASSP39728.2021.9413914>
20. Hashim, F. A., Houssein, E. H., Hussain, K., Mabrouk, M. S., & Al-Atabany, W. (2022a). Honey Badger Algorithm: New metaheuristic algorithm for solving optimization problems. *Mathematics and Computers in Simulation*, 192, 84–110. <https://doi.org/10.1016/j.matcom.2021.08.013>
21. Hashim, F. A., Houssein, E. H., Hussain, K., Mabrouk, M. S., & Al-Atabany, W. (2022b). Honey Badger Algorithm: New metaheuristic algorithm for solving optimization problems. *Mathematics and Computers in Simulation*, 192, 84–110. <https://doi.org/10.1016/j.matcom.2021.08.013>
22. Hassan, I. H., Abdullahi, M., Isuwa, J., Yusuf, S. A., & Aliyu,
23. T. (2024). A comprehensive survey of the honey badger optimization algorithm and meta-analysis of its variants and applications. *Franklin Open*, 8(August), 100141.
<https://doi.org/10.1016/j.fraope.2024.100141>
25. Houssein, E. H., Emam, M. M., Singh, N., Samee, N. A., Alabdulhafith, M., & Çelik, E. (2024). An improved honey badger algorithm for global optimization and multilevel thresholding segmentation: real case with brain tumor images. *Cluster Computing*. <https://doi.org/10.1007/S10586-024-04525-0>
26. Huang, Y., Lu, S., Liu, Q., Han, T., & Li, T. (2025). GOHBA:
27. Improved Honey Badger Algorithm for Global Optimization. *Biomimetics* 2025, Vol. 10, Page 92, 10(2), 92.
<https://doi.org/10.3390/BIOMIMETICS10020092>
29. Kavi Priya, S., & Pon Karthika, K. (2023a). EliteVec: Feature Fusion for Depression Diagnosis Using Optimized Long Short-Term Memory Network. *Intelligent Automation and Soft Computing*, 36(2), 1745–1766. <https://doi.org/10.32604/iasc.2023.032160>
30. Kinoyama, R., Perez, E. A. M., & Iba, H. (2021). Preventing Overfitting of LSTMs using Ant Colony

- Optimization. IIAI International Conference on Advanced Applied Informatics, 343–350. <https://doi.org/10.1109/IIAI-AAI53430.2021.00061>
31. Kurnianingsih, N., Wirasatriya, A., Lazuardi, L., Kubota, N., & Ng, N. (2020). IOD and ENSO-Related Time Series Variability and Forecasting of dengue and malaria incidence in Indonesia.
32. Landsat Normalized Difference Vegetation Index | U.S. Geological Survey. (n.d.). Retrieved March 24, 2025, from <https://www.usgs.gov/landsat-missions/landsat-normalized-difference-vegetation-index>
33. Liu, Z., Zhang, Q., Li, L., He, J., Guo, J., Wang, Z., Huang,
34. Y., Xi, Z., Yuan, F., Li, Y., & Li, T. (2023). The effect of temperature on dengue virus transmission by Aedes mosquitoes. *Frontiers in Cellular and Infection Microbiology*, 13, 1242173. <https://doi.org/10.3389/FCIMB.2023.1242173>
35. Majeed, M. A., Shafri, H. Z. M., Wayayok, A., & Zulkafli, Z. (2023a). Prediction of dengue cases using the attention-based long short-term memory (LSTM) approach. *Geospatial Health*, 18(1). <https://doi.org/10.4081/gh.2023.1176>
36. Nguyen, V. H., Tuyet-Hanh, T. T., Mulhall, J., Van Minh, H.,
37. Duong, T. Q., Van Chien, N., Nhung, N. T. T., Lan, V. H.,
38. Minh, H. B., Cuong, D., Bich, N. N., Quyen, N. H., Linh, T.
39. N. Q., Tho, N. T., Nghia, N. D., Anh, L. V. Q., Phan, D. T. M.,
40. Hung, N. Q. V., & Son, M. T. (2022a). Deep learning models for forecasting dengue fever based on climate data in Vietnam. *PLoS Neglected Tropical Diseases*, 16(6). <https://doi.org/10.1371/journal.pntd.0010509>
41. Qasem, S. N. (2024). A novel honey badger algorithm with a multilayer perceptron for predicting COVID-19 time series data. *The Journal of Supercomputing*, 80(3), 3943–3969. <https://doi.org/10.1007/s11227-023-05560-1>
42. Roy, S. B., Yuan, M., Fang, Y., & Sett, M. K. (2022). Spatio-Temporal Attention with Symmetric Kernels for Multivariate Time Series Forecasting. 2022 IEEE 17th Conference on Industrial Electronics and Applications (ICIEA), 21–26. <https://doi.org/10.1109/ICIEA54703.2022.10006303>
43. Sankalpa, D., Dhou, S., Pasquier, M., & Sagahyroon, A. (2024). Predicting the Spread of a Pandemic Using Machine Learning: A Case Study of COVID-19 in the UAE. *Applied Sciences* 2024, Vol. 14, Page 4022, 14(10), 4022.
44. <https://doi.org/10.3390/APP14104022>
45. Shih, S.-Y., Sun, F.-K., Lee, H., Borgwardt, K., Loh, P.-L., Terzi, E., & Ukkonen Shun-Yao Shih, A. B. (2019). Temporal pattern attention for multivariate time series forecasting. *Machine Learning*, 108, 1421–1441. <https://doi.org/10.1007/s10994-019-05815-0>
46. Ud Deen Shafique, H. N., Ali, N., Hussain, M. Z., Hasan, M. Z., Mustafa, M., Khalid, A., Awan, R., Ashraf, F., Khan, Z. A., & Javaid, A. (2024). Dengue Cases Prediction and Analysis using LSTM and Region-specific Data. 2024 OPJU International Technology Conference on Smart Computing for Innovation and Advancement in Industry 4.0, OTCON 2024. <https://doi.org/10.1109/OTCON60325.2024.10688374>
47. Wang, D., & Chen, C. (2023). Spatiotemporal Self-Attention-Based LSTNet for Multivariate Time Series Prediction. *International Journal of Intelligent Systems*, 2023. <https://doi.org/10.1155/2023/9523230>
48. Wang, G. H., Gamez, S., Raban, R. R., Marshall, J. M.,
49. Alphey, L., Li, M., Rasgon, J. L., & Akbari, O. S. (2021). Combating mosquito-borne diseases using genetic control technologies. *Nature Communications* 2021 12:1, 12(1), 1–12. <https://doi.org/10.1038/s41467-021-24654-z>
50. Wang, P., Zheng, X., Ai, G., Liu, D., & Zhu, B. (2020). Time series prediction for the epidemic trends of COVID-19 using the improved LSTM deep learning method: Case studies in Russia, Peru and Iran. *Chaos, Solitons & Fractals*, 140, 110214. <https://doi.org/10.1016/J.CHAOS.2020.110214>
51. Yu, S. (2024). Advancing Stock Market Return Forecasting with LSTM Models and Financial Indicators. <https://doi.org/10.54254/2754-1169/122/2024.17734>
52. Yu, T., & Zhu, H. (2020a). Hyper-Parameter Optimization: A Review of Algorithms and Applications. <http://arxiv.org/abs/2003.05689>
53. Yu, T., & Zhu, H. (2020b). Hyper-Parameter Optimization: A Review of Algorithms and Applications. <http://arxiv.org/abs/2003.05689>
54. Zhao, Y., Liang, Z., Du, J., Zhang, L., Liu, C., & Zhao, L. (2021). Multi-Head Attention-Based Long

Short-Term Memory for Depression Detection From Speech. *Frontiers in Neurorobotics*, 15.
<https://doi.org/10.3389/fnbot.2021.684037>

55. Zhao, Y., & Lu, J. (2024). CCC Publications Spatiotemporal Sequence Prediction Based on Spatiotemporal Self-Attention Mechanism. 1–15.
56. Zheng, H., Lin, F., Feng, X., & Chen, Y. (2021). A Hybrid Deep Learning Model with Attention-Based Conv-LSTM Networks for Short-Term Traffic Flow Prediction. *IEEE Transactions on Intelligent Transportation Systems*, 22(11), 6910–6920. <https://doi.org/10.1109/TITS.2020.2997352>