East African Journal of Engineering
eaje.eanso.org
Volume 8, Issue 1, 2025
Print ISSN: 2707-5389 | Online ISSN: 2707-5397
Title DOI: https://doi.org/10.37284/2707-5397

EAST AFRICAN NATURE & SCIENCE ORGANIZATION

Original Article

Forecasting Maximum Temperature Using Comparable Optimizers in LSTM Deep Learning Model: A Case of Koga Mango Farm, Mkuranga District, Tanzania

Isakwisa Gaddy Tende[1] & Mbazingwa Elirehema Mkiramweni[1]*

[1] Dar es Salaam Institute of Technology, P. O. Box 2958 Dar es Salaam, Tanzania.
[*] Correspondence Email: isakwisa.tende@dit.ac.tz

Article DOI : https://doi.org/10.37284/eaje.8.1.2959

Publication Date:

07 May 2025

Keywords:

Maximum Temperature Forecast, LSTM, Deep Learning, Optimizers, Tanzania.

ABSTRACT

Mango farming is an important economic activity in Tanzania, contributing to the economy through exports of mango fruits and products and acting as a primary source of income for many farmers. Maximum temperature is one of the critical weather variables affecting the growth of mango, having an impact both on flowering stages and fruits, so failure to correctly forecast extreme maximum temperature and take appropriate measures may pose challenges such as poor quality of mango fruits and hence low income to farmers. Long Short-Term Memory (LSTM) is one of the famous deep learning models used for forecasting time-series variables such as temperature. In the LSTM model, an optimizer is a very important component as it is used to minimize loss during model training. Despite there being a number of optimizers, which can be used in the LSTM model, there is still a research gap, on which one is the best-performing optimizer in forecasting tasks, especially in the context of forecasting maximum temperature in Koga farm, a mango farm located in Mkuranga district, Pwani region, Tanzania which has unique climatic conditions and has a small geographical area. This study aims to fill this gap by comparing the performances of common LSTM optimizers and developing an LSTM model for helping Koga farm officials forecast daily maximum temperature using the best-performing optimizer. The experimental findings reveal that Adam and Adamax are the two best-performing optimizers with both having Root Mean Squared (RMSE) values of 0.089 on the test set (unseen data). The performance of the remaining optimizers on the test set with their RMSE values in brackets are as follows; RMSprop (0.091), Adagrad (0.099), SGD (0.102) and Adadelta (0.107). This study recommends that software developers and researchers use either Adam or Adamax optimizer in LSTM models when forecasting temperature in environments which resemble that of the Koga farm in Tanzania.

APA CITATION

Tende, I. G. & Mkiramweni, M. E. (2025). Forecasting Maximum Temperature Using Comparable Optimizers in LSTM Deep Learning Model: A Case of Koga Mango Farm, Mkuranga District, Tanzania. *East African Journal of Engineering*, *8*(1), 190-203. https://doi.org/10.37284/eaje.8.1.2959

CHICAGO CITATION

Tende, Isakwisa Gaddy and Mbazingwa Elirehema Mkiramweni. 2025. "Forecasting Maximum Temperature Using Comparable Optimizers in LSTM Deep Learning Model: A Case of Koga Mango Farm, Mkuranga District, Tanzania". *East African Journal of Engineering* 8 (1), 190-203. https://doi.org/10.37284/eaje.8.1.2959.

Article DOI: https://doi.org/10.37284/eaje.8.1.2959

HARVARD CITATION

Tende, I. G. & Mkiramweni, M. E. (2025) "Forecasting Maximum Temperature Using Comparable Optimizers in LSTM Deep Learning Model: A Case of Koga Mango Farm, Mkuranga District, Tanzania", *East African Journal of Engineering*, 8(1), pp. 190-203. doi: 10.37284/eaje.8.1.2959.

IEEE CITATION

I. G., Tende & M. E., Mkiramweni "Forecasting Maximum Temperature Using Comparable Optimizers in LSTM Deep Learning Model: A Case of Koga Mango Farm, Mkuranga District, Tanzania" *EAJE*, vol. 8, no. 1, pp 190-203, May. 2025.

MLA CITATION

Tende, Isakwisa Gaddy & Mbazingwa Elirehema Mkiramweni. "Forecasting Maximum Temperature Using Comparable Optimizers in LSTM Deep Learning Model: A Case of Koga Mango Farm, Mkuranga District, Tanzania" *East African Journal of Engineering*, Vol. 8, no. 1, May. 2025, pp. 190-203, doi:10.37284/eaje.8.1.2959.

INTRODUCTION

Mango farming is a primary economic activity for many farmers in Tanzania, acting as their major source of income (William et al., 2015; Baltazari et al., 2020). Findings from the World Bank (World Bank, 2022) reveal that Tanzania's exports of Guavas, mangoes and mangosteens (fresh or dried) in 2022 were valued at 235,000 USD, with a total exported quantity of 38,818,000 Kg. On the other, findings from the Food and Agriculture Organization of the United Nations (FAO, 2023) reveal that Tanzania ranked 19th in the World in 2023 for producing mangoes, guavas and mangosteens, all showing mango farming is an important economic activity to many Tanzanian farmers.

Maximum Temperature is an important climatic parameter when it comes to growing mango (Rajan, 2012), having an impact on the mango flowering process and hence impacting the quality of the mango fruits (Khalifa and Abobatta, 2023). Failure to accurately forecast maximum temperature in Mango farms might pose several challenges such as poor mango yields, and hence low income for farmers. To address these challenges, it is important to have in place an effective technological model which can accurately forecast maximum temperature in mango farms, specifically Koga Farm, a mango farm located in Mkuranga district, Pwani region, Tanzania.

The use of Artificial Intelligence (AI), specifically Deep Learning for forecasting tasks has been gaining popularity over recent years due to the effectiveness of these Deep Learning models to accurately forecast/predict different variables. One example of such Deep Learning models is Long Short-Term Memory (LSTM) which has been used in several studies to forecast different parameters. Li et al. (2023) used LSTM and Adam optimizer to accurately forecast daily air temperature in Tabriz city, Iran with results showing an effective coefficient of determination (R^2) value of 0.93 on the test set. Sowmya et al. (2020) used LSTM and Stochastic Gradient Descent (SGD) optimizer to detect fake news with greater accuracy. John-Africa and Emmah (2022) used LSTM and RMSprop optimizer to detect spam messages in email, with results showing high detection accuracy of 94%. Anh et al. (2023) used LSTM and Adagrad optimizer for accurate rainfall-runoff modelling with results showing Adagrad optimizer outperformed other optimizers during testing. Karabıyık (2023) used LSTM and Adadelta optimizer to forecast Brent oil prices effectively with a performance of a Mean Average Error (MAE) score of 1.1239657. Afan et al. (2024) used LSTM and Adamax optimizer to forecast the drought index for Anbar Province, Iraq with results showing a high accuracy of 90.61%. Teixeira et al. (2024) used LSTM and Adam optimizer to forecast the average temperature in Portugal, with results showing that MSE (Mean Squared Error) decreased by 97% on the test set. Chai et al. (2024) used LSTM and Adam optimizer to accurately predict runoff in the Xijiang River Basin in China with results showing an effective daily runoff prediction determination coefficient (R^2) of 0.971. Waqas et al. (2024) used LSTM and Adam optimizer for accurate Seasonal Precipitation Forecasting (SPF) in Eastern Thailand, with results showing a coefficient of determination (R^2) value of 0.91. Yan et al. (2024) used LSTM and Adam optimizer to predict soil

Article DOI: https://doi.org/10.37284/eaje.8.1.2959

temperature in China, with results showing an accurate Nash-Sutcliffe Efficiency (NSE) value of 0.92.

Although the reviewed studies reveal good performances and powerfulness of LSTM-based models in forecasting tasks, there is a research gap on which optimizer is most effective to use in LSTM models forecasting tasks, especially in the context of forecasting maximum temperature in Koga farm which has unique characteristics such as small geographical area and unique climatic condition. This uniqueness calls for a study to comparatively evaluate the performances of these optimizers since an optimizer for the LSTM model cannot just be selected, assuming it will have the best performance. This is due to the fact that climatic conditions in Tanzania are unique and different from the conditions in other parts of the World also, as it has been revealed in the literature, different optimizers perform well under different conditions, so performances of these optimizers can't just be assumed.

Therefore, this study has two objectives; first to comparatively evaluate the performances of different optimizers for the LSTM model in forecasting daily maximum temperature in Koga farm and second to develop the LSTM model for forecasting daily maximum temperature in Koga farm by using the best-performing optimizer. The developed model will help farm officials in Koga farm to predict next-day maximum temperature and take precautionary measures to protect mango flowers and fruits in case of extreme maximum temperature. This study aims to answer one research question; what is the best LSTM optimizer to use for forecasting maximum temperature in Koga farm? To the best of our knowledge, there is no study which has comparatively evaluated the performance of LSTM optimizers in Mkuranga district, Pwani region, Tanzania. The findings of this study will fill the existing information gap of performance comparison of LSTM model optimizers in forecasting maximum temperature, especially in the context of Tanzania's unique climatic conditions.

MATERIALS AND METHODS

Study Area

The area under this study is Koga farm, a mango farm located in Mkuranga district, Pwani region, Tanzania. The study area's climate is of modified equatorial type (Majule, 2012) with bimodal (two) rainy seasons, short season (vuli) and long season (masika). The study area also experiences variations of both maximum and minimum temperatures as it is located in a coastal region. *Figure 1* shows the study area.

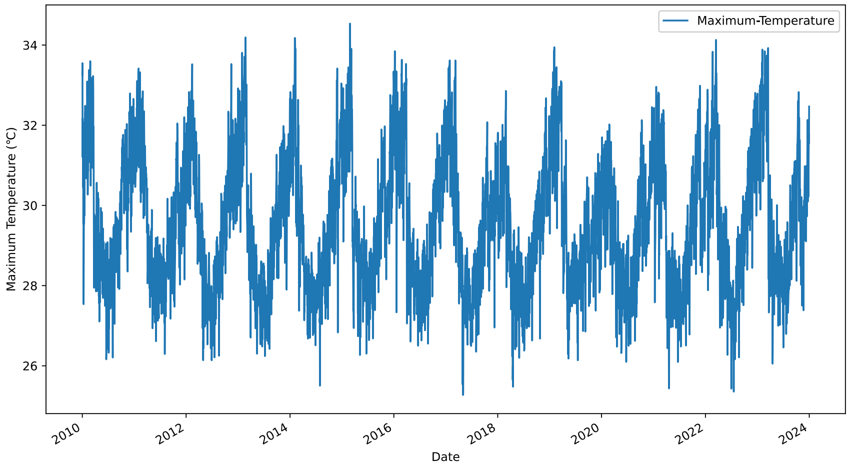
Figure 1: Koga Farm

Article DOI: https://doi.org/10.37284/eaje.8.1.2959

Dataset

This study used the European Reanalysis (ERA5 Ag - 9.6km Daily) global dataset to download time-series daily maximum temperature data for Koga farm (refer to *Figure 1*). The polygon drawn in *Figure 1* shows the map of Koga farm. The downloaded data were limited to the area enclosed in the polygon. The daily maximum temperature data in CSV format for a total of 14 years (from January 1st, 2010 to December 31st, 2023) for Koga farm were downloaded from Google Earth Engine (GEE) which hosts the ERA 5 Ag – 9.6km Daily dataset. GEE is a cloud platform developed by Google (Tamiminia et al., 2020). The CSV data were downloaded from GEE with the help of the Climate Engine application (Huntington et al., 2017). *Figure 2* shows the pattern of the daily maximum temperature data for Koga farm for 14 years, from January 1st, 2010 to December 31st, 2023.

Figure 2. Daily Maximum Temperature for Koga Farm from January 1st, 2010 to December 31st, 2023



Sampling Criteria for LSTM Optimizers

The LSTM optimizers evaluated in this study (Adam, SGD, RMSprop, Adadelta, Adamax and Adagrad) were purposively selected based on the following sampling criteria:

- Popularity: The mentioned LSTM optimizers are popular and have been widely used in LSTM models for forecasting and classifying different parameters. They are among the most commonly used optimizers as evidenced by their inclusion and implementation in popular Deep Learning frameworks such as TensorFlow, Keras and PyTorch.

- Diversity in Optimizer Theory: Each of the six optimizers uses a distinct gradient-based optimization approach for adjusting weights and biases of a Deep Learning model in order to minimize loss (difference between actual and predicted values). SGD adjusts model parameters (weights and biases) using a fixed learning rate and gradients from individual batches. Adam adjusts model parameters by combining momentum and adaptive learning rates using moving averages of gradients and their corresponding squares. RMSprop adjusts model parameters by adapting the learning rate based on recent gradient magnitudes using an exponential moving average. Adagrad adjusts model parameters by adjusting the learning rate based on the sum of squared gradients of a parameter. Adadelta adjusts model parameters by modifying Adagrad to use a moving window of squared gradients instead of accumulating all past values. Adamax adjusts model parameters by modifying Adam to use the infinity norm (focusing on the largest absolute

Article DOI: https://doi.org/10.37284/eaje.8.1.2959

gradient value) instead of the standard L2 (Euclidean) norm.

- Relevance in Empirical Literature: The six optimizers have been widely studied in the literature with results suggesting their performances vary from scenario to scenario and depend on the nature of the problem, model architecture and dataset used. Comparative evaluation helps to understand the strengths and weaknesses of each optimizer in different scenarios.

- Availability of Libraries: Each of the six optimizers is readily available in major Deep Learning frameworks such as TensorFlow, Keras and PyTorch. This helps to ensure consistent implementation and reproducibility of results.

Data Pre-processing

The LSTM model requires data to be pre-processed before being fed into the model. The data pre-processing process involved several steps as described in the following section.

- Data Analysis**:** All daily data of maximum temperature for Koga farm for a period of 14 years was analyzed (refer to *Figure 3*) with the data analysis revealing a count (total datapoints) of 5109, an average (mean) of 29.553057 ℃, a standard deviation (std) of 1.668063 ℃, a lowest (min) value of 25.270400 ℃, first quartile (25%) at 28.242200 ℃, second quartile (50%) at 29.395300 ℃, third quartile (75%) at 30.732600 ℃ and a highest (max) value of 34.538400 ℃.

- Scaling of Data: The daily maximum temperature data for Koga farm was scaled down to within a range of 0 and 1 to enhance the training of the LSTM model.

- Data Split: To have an effective LSTM model, it is important to split the data into train, validation and test set (refer to *Figure 4*). The train and validation sets are normally used during the training of the LSTM model while the test set (unseen data) is normally used for evaluating the LSTM model's performance and test its ability to generalize when fed with completely new data, for instance, the already trained LSTM model which has been saved in .h5 format can be fed with the first 30 days of January 2025 (from January 1st to January 30th) as input data and asked to forecast the maximum temperature of January 31st). The downloaded 14-year data was split into a training set (60%, from 2010 to 2017), a validation set (20%, from 2018 to 2020) and a test set (20%, from 2021 to 2023).

- Input Features and Labels: Since the downloaded data consists of only daily maximum temperature data, we needed to create input features and their corresponding labels in order for the LSTM model to be correctly trained and learn to map input features to their corresponding labels. This study decided to select a sequence of the previous 30-day maximum temperature as the input feature and the next-day maximum temperature as its label. We did this for all three sections of the data (training, validation and test sets).

East African Journal of Engineering, Volume 8, Issue 1, 2025
Article DOI: https://doi.org/10.37284/eaje.8.1.2959

Figure 3: Data Analysis Results

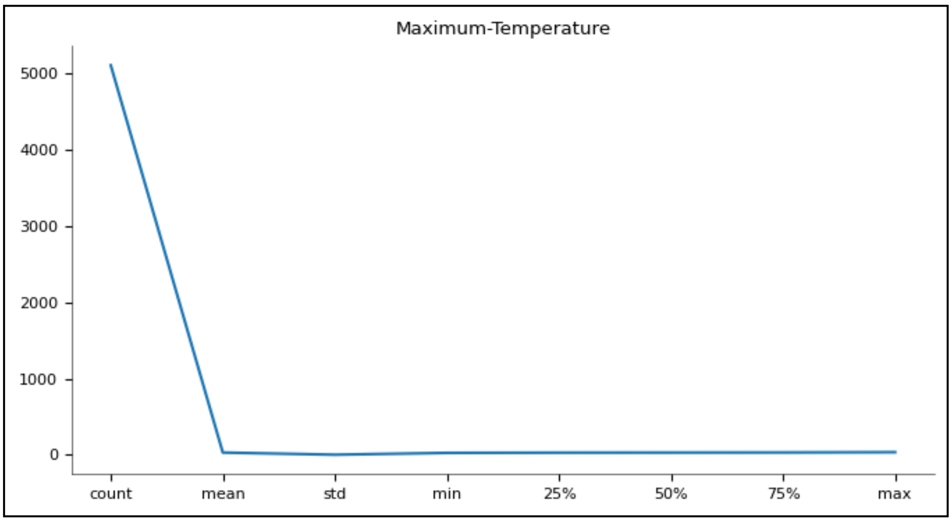
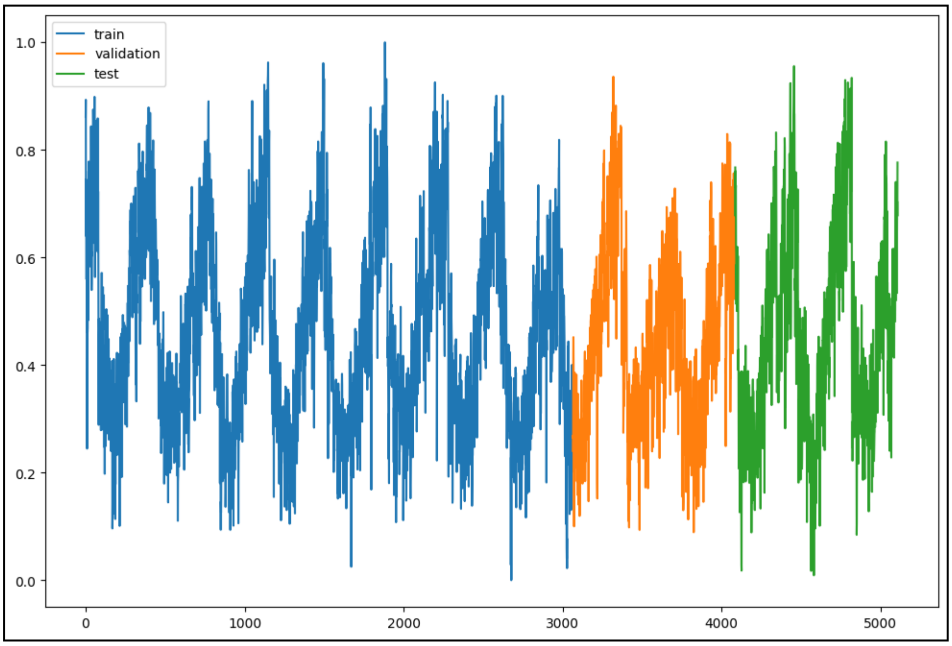


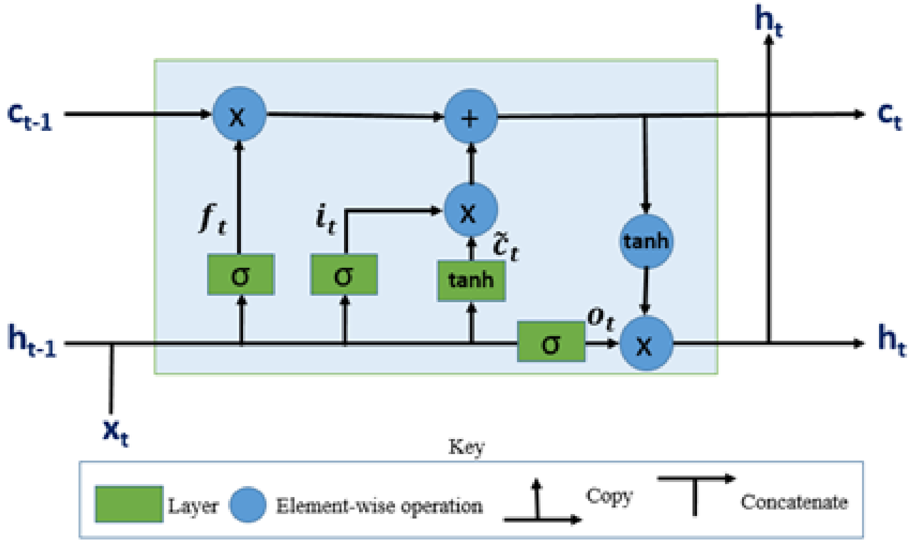
Figure 4: Training, Validation and Test Sets



LSTM Architecture

LSTM (Hochreiter and Schmidhuber, 1997) is a type of Recurrent Neural Network (RNN), normally suited for processing sequential data such as time series variables. However, unlike traditional RNNs which are faced with the challenge of vanishing gradients (failure to remember information from earlier time steps), a scenario happening during training of RNN by back-propagation, LSTM addresses the issue of vanishing gradients, by having a unique ability to keep holding necessary information over many time steps, making the information available when it needs to be utilized later. For this reason, LSTM is very useful in forecasting time series variables such as daily temperature. The architecture of the LSTM unit is shown in *Figure 5*.

Figure 5: LSTM Unit



The LSTM unit consists of a cell state c_t and three gates namely; forget gate f_t, input gate i_t and output gate o_t.

- Cell State: The purpose of the cell c_t is to store information across time steps. The cell is usually updated (adding or removing information) by forget and input gates. This helps the LSTM network to either retain or not retain information based on its relevance. This feature allows LSTM to remember relevant information over many time steps.

- Forget Gate: The purpose of the forget gate f_t is to make decisions on what information should be discarded (not kept) from the cell state. The forget get does this by combining two inputs; the previous hidden state h_{t-1} and the current input x_t and then producing an output value whose range is between 0 and 1 for every component in the cell state, with an output value of 0 indicating "completely discard" and an output value of 1 indicating "completely keep". The role sigmoid neural network layer σ is to produce an output value of between 0 and 1 when given any input. The forget gate is shown in equation (i) with W_f and b_f representing weight matrices and bias vector parameters of the forget gate respectively, all of which are learned during the training process.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad \text{(i)}$$

- Input Gate: The purpose of the input gate i_t is to make decisions on which new information should be added in the cell state. The input gate consists of two elements; the sigmoid neural network layer σ and a $tanh$ layer whose function is to create a candidate vector $\tilde{c}_t$whose values could be added to the cell state. The components of the input gate are shown in equations (ii) and (iii) with W_i and b_i being weight matrices and bias vector parameters of the input gate respectively, W_c and b_c being weight matrices and bias vector parameters of the candidate cell state respectively, all of which are learned during the training process.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad \text{(ii)}$$

$$\tilde{c}_t = tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad \text{(iii)}$$

- Output Gate: The output gate o_t (refer to equation (iv)) controls the LSTM unit output. The output gate combines the previous hidden state h_{t-1} and the current input x_t and decides which part of the cell state to output as the next hidden state, with

Article DOI: https://doi.org/10.37284/eaje.8.1.2959

W_o and b_o being weight matrices and bias vector parameters of the output gate respectively.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad \text{(iv)}$$

Finally, the cell state and hidden state need to be updated. The cell state is updated by the forget gate and input gates as shown in equation (v) and then the updated cell state and the output gate are used to update the hidden state (LSTM unit output) as shown in equation (vi).

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t \quad \text{(v)}$$

$$h_t = o_t * \tanh(c_t) \quad \text{(vi)}$$

Proposed LSTM Model

The architecture of the proposed LSTM model is shown in *Figure 5*. The model consists of two LSTM layers and one dense layer. The role of LSTM layers is to learn the pattern of the input data (previous day's maximum temperature) so that it can correctly predict the next day's maximum temperature while the role of a Dense layer is to output a single numerical value which acts as the forecasted (predicted) maximum temperature. A total of six optimizers; Adam, SGD, RMSprop, Adadelta, Adamax and Adagrad were chosen for this study since they are among the commonly used optimizers in LSTM models. The chosen optimizers were alternately used during the training of the LSTM model shown in *Figure 5*.

Figure 5: Proposed LSTM model

Loss Function and Evaluation Metric

Loss function is normally used to measure the error between the actual value y and predicted value $\hat{y}$. The loss function is an important function during the training of LSTM models since it is used to examine how accurately the model can forecast the values which are close to true (actual) values. For this study, the Mean Squared Error (MSE) shown in equation (vii) is used as a loss function. On the other hand, evaluation metrics are used to evaluate (test) the performance of the LSTM model on a test set (unseen data) and measure its ability to generalize on new data that it has never seen before. For this study, the Root Mean Squared Error (RMSE) shown in equation (viii) is used as an evaluation metric.

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

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

RESULTS

Tuning of Hyperparameters

Hyperparameters play a crucial role in finetuning the LSTM model during training to maximize the performance of the model. After several rounds of training the following hyperparameters were chosen for the LSTM model for each of the six optimizers; two layers of LSTM unit, output-dimensionality of 100 for the first LSTM layer, output-dimensionality of 200 for the second LSTM layer, batch-size of 16, learning rate of 0.01 and 100 epochs of training.

Training of LSTM Model

This study conducted a total of six training experiments for the same LSTM model, using a different optimizer and the same hyperparameters in each training experiment. After each training experiment of the LSTM model with an optimizer, the trained LSTM model was saved in .h5 format. This resulted in six different saved versions of the LSTM model, one version for each optimizer. The LSTM model was developed in IPython

Article DOI: https://doi.org/10.37284/eaje.8.1.2959

Notebook and all training and testing experiments of the LSTM model were conducted in the Google Colab Cloud Platform (Bisong, 2019) with the following runtime environment; system RAM of 12.7 GB and Hard Disk space of 107.7 GB. *Figure 6* and *Figure 7* show the training loss (MSE) and training RMSE respectively of the LSTM model using six different optimizers.

Figure 6: Training Loss (MSE) for LSTM Optimizers

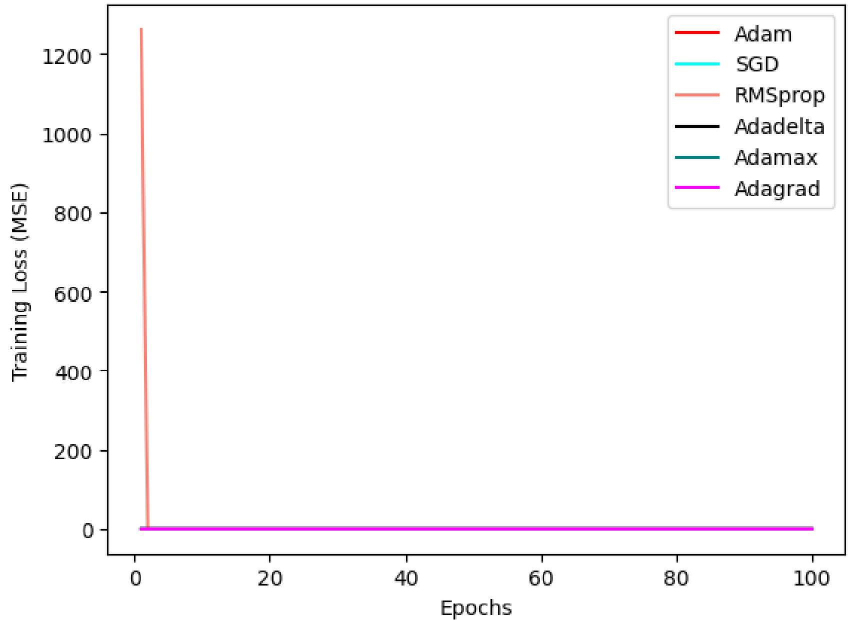
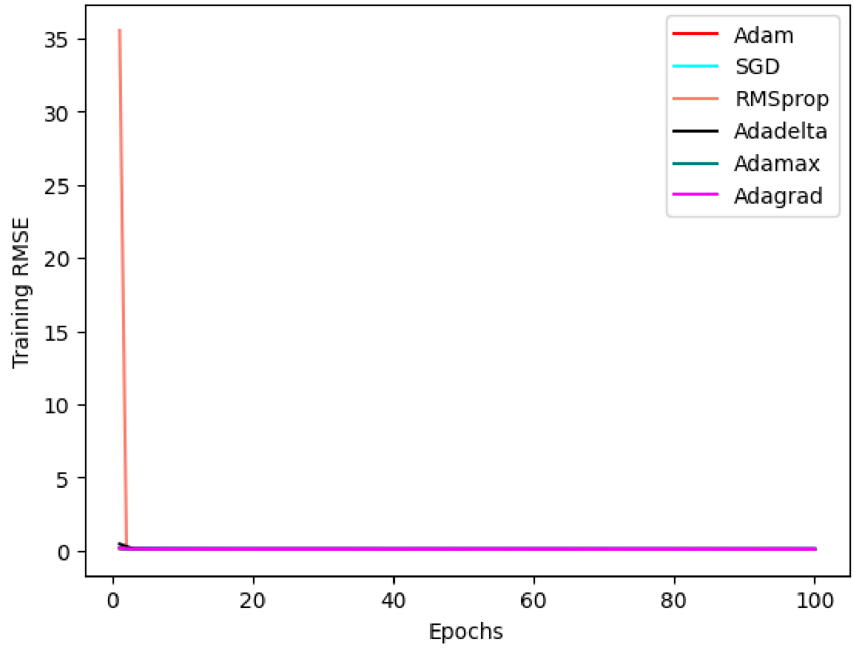


Figure 7: Training RMSE for LSTM Optimizers



Performance Evaluation of LSTM Optimizers

Each version of the LSTM model was used to evaluate the performance of the LSTM optimizer used by testing it against a test set (unseen data which have not been seen by the LSTM model). Each version of the LSTM model was given test data as input and asked to forecast the maximum temperature. *Figure 8* shows the actual maximum temperature against the predicted maximum temperature by different versions of the LSTM model (each version representing the LSTM model with a different optimizer). *Table 1* shows

Article DOI: https://doi.org/10.37284/eaje.8.1.2959

the results of test RMSE scores for all six versions of the LSTM model with Adam and Adamax optimizers achieving the best performances on test data with both attaining RMSE and MSE scores of 0.089 and 0.00792 respectively, followed by RMSprop with RMSE and MSE scores of 0.091 and 0.00828 respectively, followed by Adagrad optimizer with RMSE and MSE scores of 0.099 and 0.0098 respectively, followed by SGD optimizer with RMSE and MSE cores of 0.102 and 0.0104 respectively and lastly followed by Adadelta optimizer with RMSE and MSE scores of 0.107 and 0.0114 respectively from these results, it can be revealed that, for the case of Koga farm, Adam and Adamax are the two best-performing optimizers, followed by RMSprop, Adagrad, SGD and Adadelta optimizers.

Figure 8: Actual vs Predicted Maximum Temperature by LSTM Optimizers

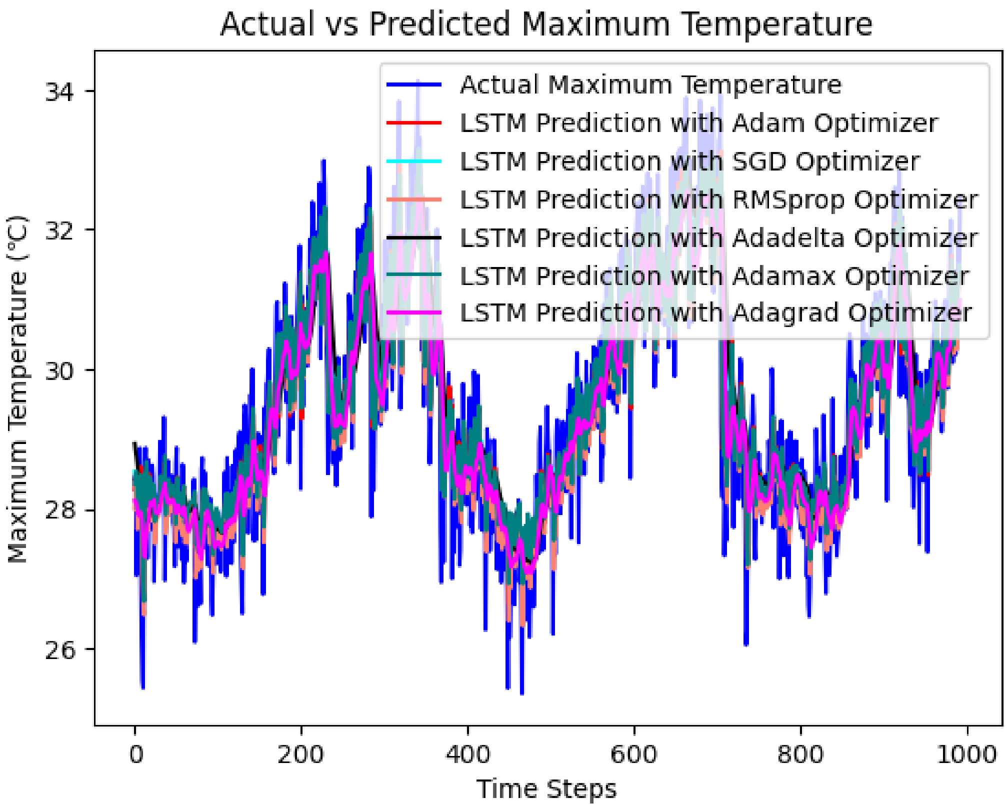


Table 1: Test RMSE and MSE Scores of LSTM Optimizers

LSTM Optimizer	Test RMSE	Test MSE
Adam	0.089	0.00792
SGD	0.102	0.0104
RMSprop	0.091	0.00828
Adadelta	0.107	0.0114
Adamax	0.089	0.00792
Adagrad	0.099	0.0098

DISCUSSION

Choice of LSTM Optimizer

Adam and Adamax achieved the best performance with the identical lowest Test RMSE and Test MSE scores. The strength of these two optimizers lies in their adaptive learning rate approaches combined with momentum. This enables these two optimizers to efficiently handle noisy and sparse gradients which is often the case in real-

Article DOI: https://doi.org/10.37284/eaje.8.1.2959

world data. The ability of these two optimizers to 'adapt' is the most likely reason which contributed to their superior generalization capability in predicting maximum temperature patterns in Koga farm and these results make these two optimizers the best choice when it comes to practical implementation and deployment. RMSprop was the next best-performing optimizer, likely because of its approach of using root mean square scaling of gradients, which also adapts the learning rate dynamically. Although the RMSprop optimizer did not outperform Adam and Adamax optimizers, its performance makes it an alternative choice for practical implementation and deployment in case Adam and Adamax optimizers cannot be implemented. Adagrad and SGD optimizers showed moderate performance and this is likely because Adagrad optimizer, though adaptive, normally reduces the learning rate too aggressively over time, which possibly impeded its ability to capture longer-term temperature trends and dependencies while SGD optimizer on the other hand which uses non-adaptive approach, was likely sensitive to the chosen learning rate and this probably affected its convergence speed and stability. Due to these reasons, Adagrad and SGD are unpreferred choices when it comes to practical implementation and deployment. Adadelta optimizer achieved the worst results, probably because of its reliance on accumulated gradient updates without good sensitivity to new patterns in data, resulting in failure to adjust appropriately to the dynamic nature of fluctuating maximum temperature. Due to this reason, Adadelta also becomes an unpreferred choice when it comes to practical implementation and deployment.

Comparison with Existing Studies

The findings of this study reveal that Adam and Adamax are the two best optimizers to use in the LSTM model when forecasting maximum temperature in Koga farm as well as in the environment with a small geographical area and climatic condition which resembles that of Koga farm, Mkuranga district, Pwani region, Tanzania. These findings align with results from the literature which reveal that the performance of the LSTM optimizer depends on the context such as the nature and pattern of the variable being predicted, the climate of particular case study area, etc. and thus cannot be generalized. This is evident in a study by Anh et al. (2023) which reveals that Adagrad is the best LSTM optimizer in forecasting rainfall-runoff outperforming RMSprop, Adadelta, and Adam optimizers and a study by Bakhashwain and Sagheer (2021) which reveals superior performance of Adamax optimizer over Adam optimizer in predicting stock prices, all of this suggesting different LSTM optimizers perform well under different conditions.

Theoretical Implications

These findings support established theories in Deep Learning regarding the effectiveness of adaptive gradient methods, specifically in real-world problems and scenarios. Adam and Adamax optimizers demonstrate how a combination of adaptive learning rates and momentum can lead to faster convergence and more accurate predictions in changing conditions and dynamic environments. These findings also imply that optimizers which are non-adaptive such as SGD may underperform when using fixed learning rates, specifically in experiments which involve some noisy or fluctuating data.

Practical Applications

This study's findings have implications for climate-smart agriculture and farm-level decision support systems. Prediction of accurate short-term maximum temperatures helps farm officials and farmers in the following areas:

- Implementation of Extreme Temperature Preventive Measures: Farm officials and farmers can take appropriate precautionary measures against predicted extreme maximum temperatures, measures such as installation of shade nets to cover young mango trees, adjusting the frequency of irrigation in order to reduce heat stress and use of mulching to retain moisture.

Article DOI: https://doi.org/10.37284/eaje.8.1.2959

- Agricultural Planning: Farm officials and farmers can schedule agricultural activities like planting and harvesting based on forecasted maximum temperatures.
- Efficiency in Resource Allocation: Farm officials and farmers can efficiently allocate resources and mitigate risks, especially in regions where variation in climatic conditions such as maximum temperature has the potential to affect crop yields and productivity.

Major Contributions

This study has the following three major contributions.

- *Ready-to-Use Dataset*: This study has created a pre-processed dataset which is ready to be used by prospective AI researchers in Tanzania. The dataset will later be hosted in a free cloud platform such as GitHub for free public access.
- *Ready to Use Model*: This study has developed a novel LSTM model integrated with Adam optimizer which is the best-performing optimizer to help farm officials in Koga farm and other similar farms to predict daily maximum temperature and take appropriate measures to protect mango flowers and fruits in case of forecasted extreme maximum temperature.
- *Filling the Gap*: This study has helped to fill the existing gap on which optimizer is the most effective to use in the LSTM model for forecasting temperature in an environment similar to Koga farm, Mkuranga district, Pwani region, Tanzania. This information is critical to AI software developers and researchers who aim to develop LSTM-based forecasting models in Tanzania.

Study Limitations

Although this study offers insightful practical findings, there are a few limitations which need to be acknowledged:

- Limited Metrics for Evaluation: This study used only RMSE as a performance evaluation metric and MSE as a loss function. Evaluating the performance of the LSTM model with additional metrics such as Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE) and R^2 would offer a more holistic view of the LSTM model performance in predicting maximum temperature in Koga farm.
- Single Dataset: This study used data from a single farm (Koga farm). This might limit its generalizability in other regions and countries with different climatic conditions.
- Focus on Predicting Short-Term Temperatures: This study has developed the LSTM model which focuses on predicting short-term temperatures and excluding prediction of long-term temperatures. Prediction of long-term temperature may require different modelling and optimization approaches.

CONCLUSION

This study has developed and comparatively evaluated the LSTM model with six commonly used optimizers for forecasting maximum temperature in Koga farm, Mkuranga district, Pwani region, Tanzania. The findings reveal that Adam and Adamax are the two best-performing optimizers in forecasting maximum temperature by the LSTM model in an environment resembling the geographical area and climatic conditions of Koga Farm.

Recommendations

This study recommends that AI software developers and researchers use either Adam or Adamax optimizer in LSTM-based temperature forecasting tasks in an environment resembling that of Koga farm.

Acknowledgements

The authors would like to give thanks to the Dar es Salaam Institute of Technology for supporting this research.

Article DOI: https://doi.org/10.37284/eaje.8.1.2959

REFERENCES

William, J., Hella, J., Lars, E., Offenberg, J., Mwatawala, M. & Rwegasira, G. (2015). Partial budgeting analysis of different strategies for management of insect pests in cashew and mango orchards in Tanzania. *International Journal of Sustainable Agricultural Research*, 2(4), 98-110

Baltazari A., Hosea M., Lucy C., Theodosy M., Abdul K., George T., Jaspa S., Gopinadhan P., Alan S., Jayasankar S. & Maulid M. (2020). Evaluation of post-harvest losses and shelf life of fresh mango (Mangifera indica L.) in Eastern zone of Tanzania. *International Journal of Fruit Science*, 20(4), 855-870

World Bank (2022). Tanzania Guavas, mangoes and mangosteens, fresh or dried exports by country in 2022. Retrieved January 31, 2025, from https://wits.worldbank.org/trade/comtrade/en/country/TZA/year/2022/tradeflow/Exports/partner/ALL/product/080450

Food and Agriculture Organization of the United Nations. (2023). Countries by Commodity Rankings. Retrieved January 31, 2025, from https://www.fao.org/faostat/en/#rankings/countries_by_commodity

Rajan, S. (2012). Phenological responses to temperature and rainfall: A case study of mango. *Tropical fruit tree species and climate change*, 71

Khalifa, S., & Abobatta, W. (2023). Climate Changes and Mango Production (Temperature). *IgMin Research*, 1(1), 043-046

Li, Y., Tongfei L., Wei L., Zhiyao L., & Junxian W. (2023). Prediction of Daily Temperature Based on the Robust Machine Learning Algorithms. *Sustainability* 15(12), 9289

Sowmya B., Nikhil C., Seema S., & Srinivasa, K. (2020). Fake News Detection using LSTM Neural Network Augmented with SGD Classifier. *Solid State Technology*, 63, 6

John-Africa, E. & Emmah, V. (2022). Performance Evaluation of LSTM and RNN Models in the Detection of Email Spam Messages. *European Journal of Information Technologies and Computer Science*, 2(6), 24-30

Anh, D., Thanh, D., Le, H., Sy, B., Tanim, A., Pham, Q., Dang, T., Mai, S., & Dang, N. (2023). Effect of gradient descent optimizers and dropout technique on deep learning LSTM performance in rainfall-runoff modeling. *Water Resources Management*, 37(2), 639-657

Karabıyık, M. (2023). A Framework for Parametric Model Selection in Time Series Problems. Gazi Mühendislik Bilimleri Dergisi, 9(4), 82-91

Afan, H., Atheer A., Basheer A., Faidhalrahman K., Alaa A., Md Munir K., Muhammad, M., & Ammar K. (2024). LSTM Model Integrated Remote Sensing Data for Drought Prediction: A Study on Climate Change Impacts on Water Availability in the Arid Region. Water, 16, 19, 2799

Teixeira, R., Adelaide C., Eduardo, P., & José Baptista. (2024). Enhancing Weather Forecasting Integrating LSTM and GA. *Applied Sciences* 14, 13, 5769.

Chai, Q., Shuting, Z., Qingqing, T., Chaoqiang, Y., & Lei, G. (2024). Daily Runoff Prediction Based on FA-LSTM Model, *Water* 16,16, 2216

Waqas, M., Usa H., Phyo, H., & Shakeel, A. (2024). Seasonal WaveNet-LSTM: A Deep Learning Framework for Precipitation Forecasting with Integrated Large Scale Climate Drivers. *Water* 16, 22, 3194

Yan, Y., Gan, L., Qingliang, L., & Jinlong, Z. (2024). Enhancing Hydrological Variable Prediction through Multitask LSTM Models. *Water* 16, 15, 2156

Majule, A. (2012). Implications of ecological and social characteristics to community

Article DOI: https://doi.org/10.37284/eaje.8.1.2959

livelihoods in the coastal areas of Tanzania. African Journal of Environmental Science and Technology, 6(1), 72-79

Tamiminia, H., Salehi, B., Mahdianpari, M., Quackenbush, L., Adeli, S. & Brisco, B. (2020). Google Earth Engine for geo-big data applications: A meta-analysis and systematic review. *ISPRS journal of photogrammetry and remote sensing*, 164, 152-170

Huntington, J., Hegewisch, K., Daudert, B., Morton, C., Abatzoglou, J., McEvoy, D. & Erickson, T. (2017). Climate engine: Cloud computing and visualization of climate and remote sensing data for advanced natural resource monitoring and process understanding. *Bulletin of the American Meteorological Society*, 98(11), 2397-2410

Hochreiter, S. & Schmidhuber, J. (1997). Long short-term memory. *Neural Comput*.,9, 1735–1780

Bisong, E. (2019). Google colaboratory. In *Building machine learning and deep learning models on Google Cloud Platform: a comprehensive guide for beginners*. Berkeley, CA: Apress.

Bakhashwain, N. & Sagheer, A. (2021). Online Tuning of Hyperparameters in Deep LSTM for Time Series Applications. *International Journal of Intelligent Engineering & Systems*, 14(1)

 |