East African Scholars Journal of Engineering and Computer Sciences
Abbreviated Key Title: East African Scholars J Eng Comput Sci
ISSN: 2617-4480 (Print) & ISSN: 2663-0346 (Online)
Published By East African Scholars Publisher, Kenya

OPEN ACCESS

Volume-8 | Issue-2 | Mar-Apr-2025 | DOI: https://doi.org/10.36349/easjecs.2025.v08i02.001

Original Research Article

Forecasting United States Dollar to Tanzania Shillings Exchange Rate Using Comparable LSTM and BiLSTM Deep Learning Models

Isakwisa Gaddy Tende[1*]
[1]Department of Computer Studies, Dar es Salaam Institute of Technology, P.O. Box 2958, Dar es Salaam, Tanzania

Article History
Received: 22.01.2025
Accepted: 27.02.2025
Published: 01.03.2025

Journal homepage:
https://www.easpublisher.com

Quick Response Code

Abstract: Tanzania heavily depends on United States Dollar (USD) foreign currency to import various goods and services into the country. Failure to correctly forecast exchange rates between USD and Tanzanian Shillings (TZS) may pose risks such as inability to import intended goods and services, possibility of losing money in stock exchange markets and other investment businesses in case of unexpected currency appreciation or depreciation as well as poor investment decisions in foreign exchange markets. To address this, this study has developed and comparatively evaluated performances of LSTM (Long Short-Term Memory) and BiLSTM (Bidirectional LSTM) deep learning models for forecasting daily USD to TZS exchange rates. The findings reveal that, BiLSTM model outperforms LSTM in forecasting daily USD to TZS exchange rates, achieving a MAPE (Mean Absolute Percentage Error) score of 0.363 on test set (unseen data) compared to a MAPE score of 1.471 achieved by LSTM model. This study recommends to the prospective Artificial Intelligence (AI) researchers and software developers to use BiLSTM instead of LSTM model to forecast (predict) USD to TZS exchange rates. Also, this study has developed USD to TZS exchange rates dataset which can be used by AI researchers, saving them time and costs involved with creating datasets from scratch. This study has also developed ready to use BiLSTM and LSTM models which can be used by Tanzanian business men and women involved in stock exchange markets, foreign exchange markets and other businesses, to predict daily USD to TZS exchange rates and make appropriate business and investment decisions.

Keywords: United States Dollar (USD), Tanzanian Shillings (TZS), Exchange Rate Forecasting, LSTM, BiLSTM, Deep Learning.

1. INTRODUCTION

Tanzania relies heavily on United States Dollar (USD) for importing goods and services such as fuel, motor vehicles, electronic appliances and other goods and services (Nyamrunda *et al.*, 2014). BoT (2025) reveals that, in 2023, Tanzania imported goods and services worth of 16,058.71 million USD, showing importance of having access to strong foreign currency such as USD for importing goods and services in Tanzania.

Failure to accurately forecast exchange rate between USD and Tanzania Shillings (TZS), might pose several risks, such as inability to import intended goods and services and risk of losing money in stock exchange markets and other financial investment businesses in case of unexpected currency appreciation or depreciation. Also, this forecasting failure might lead to poor investment decisions in foreign exchange markets. To address these risks, it is important to have in place effective models which can accurately forecast daily USD to TZS exchange rates.

Deep Learning (DL), has been showing high effectiveness in accurately forecasting time series variables, which are variables spanning over a period of time (Sejnowski, 2020). Long Short-Term Memory (Hochreiter *et al.*, 1997) abbreviated as LSTM and its variant Bidirectional Long Short-Term Memory abbreviated as BiLSTM are two of the DL models which have been widely used to forecast time series variables and there is a number of studies which have developed these two DL models for forecasting various time series variables.

***Corresponding Author:** Isakwisa Gaddy Tende
Department of Computer Studies, Dar es Salaam Institute of Technology, P.O. Box 2958, Dar es Salaam, Tanzania

Lei *et al.*, (2024) developed LSTM based method to predict wind power with results showing the method was effective with the Mean Absolute Percentage Error (MAPE) on test set (unseen data) being reduced by of 9.3508%. Abbas *et al.*, (2024) developed LSTM based model for predicting power system voltage stability with the findings revealing the LSTM model achieved the highest accuracy target of 99.5%. He *et al.*, (2023) developed LSTM model to predict tunnel surrounding rock deformation with the results showing the LSTM model was effective, achieving the Root Mean Square Error (RMSE) score of 2.64. Wang (2024) developed LSTM based system to early detect motor failure in chemical plants, with the results showing the model was effective achieving results of 98% agreement between the predicted and actual anomalies over three months period. Iordan (2024) developed LSTM model for predicting software development efforts, achieving good results of RMSE score of 6.560 between predicted and actual values on test set.

On the other hand, several studies have developed BiLSTM models to predict time series variables. Fan *et al.*, (2024) developed a model based on BiLSTM and Local Attention Mechanisms for predicting sensor data, achieving effective RMSE enhancements of 17.96%. Marjani *et al.*, (2024) developed BiLSTM based model for near-real-time daily wildfire spread prediction, achieving effective F1 Score of 0.73. Zhang *et al.*, (2023) developed a network intrusion detection model based on BiLSTM with multi-head attention mechanism, achieving high accuracy of 99.08%. Staffini (2023) developed BiLSTM based architecture for macroeconomic time series forecasting, achieving RMSE score of 1.5165 in forecasting inflation. Wang *et al.*, (2023) developed BiLSTM model for predicting post-construction subsoil settlement under embankment, achieving prediction accuracy of 92%. Xu *et al.*, (2024) developed a novel trajectory prediction method based on BiLSTM model, achieving effective Mean Absolute Error (MAE) value of 1.401 in angle prediction.

Despite good work from the reviewed studies, there is still a wide research gap, as to which model among the two (LSTM and BiLSTM) is more effective, especially in the context of forecasting USD to TZS exchange rate, because of the uniqueness of USD to TZS exchange rate trend and pattern. Because of this uniqueness, it is not possible to just generalize from other studies and assume either LSTM or BiLSTM is more effective.

Hence, this study aims to fill this research gap by answering one key research question; which model, among LSTM and BiLSTM is more effective in forecasting USD to TZS exchange rate? The objectives of the study are twofold; first to develop two LSTM and BiLSTM deep learning models for predicting USD to TZS exchange rate and second, to evaluate comparatively the performances of the two models in forecasting USD to TZS exchange rate. The model with best performance among the two models (LSTM and BiLSTM) will be recommended as the better model for forecasting USD to TZS exchange rate. To the best of my knowledge, there is no study which has attempted to forecast USD to TZS exchange rate by using comparable LSTM and BiLSTM models.

2. MATERIALS AND METHODS

2.1 Data Collection

The dataset used for forecasting USD to TZS exchange rate was downloaded from Investing website (Investing, 2025), an online and reliable resource storing various historical exchange rates. Although the dataset contains several parameters, I only selected and kept four daily parameters in the dataset; Price which indicates the exchange rate, Open which indicates the opening exchange rate in the day, High which indicates the highest exchange rate in the day and Low which indicates the lowest exchange rate in the day. The daily records of these four parameters spanning a total of 12 years (from January 3rd, 2011 to December 30th, 2022) were downloaded and saved in CSV format.

2.2 Input Features Correlation Analysis

To select the best input feature (s) to use for training LSTM and BiLSTM deep learning models, I performed correlation analysis of the four parameters using Pearson correlation coefficient r (Benesty *et al.*, 2008). The results are shown in the correlation heatmap in Figure 1.

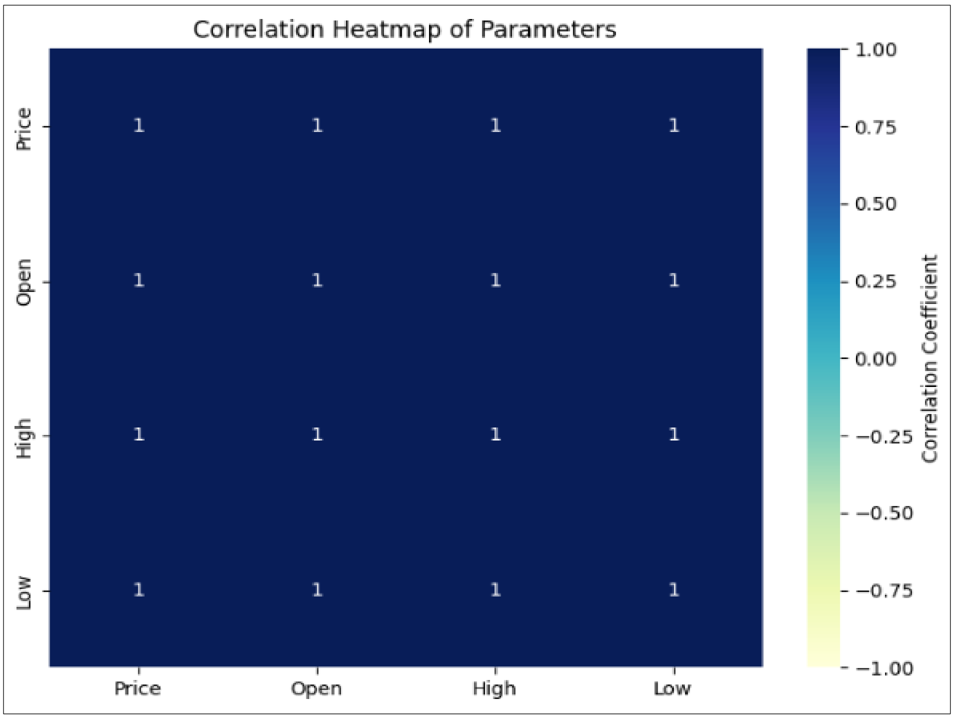

Figure 1: Correlation Heatmap of Input Parameters

From the results of Figure 1, it can be seen that, all of the parameters are strongly correlated (Pearson r=1), and this implies there is no need of using more than one parameter as input features, as the additional features will not add any values in training accuracy of the models. Because of this reason, I chose Price as the only input feature to the LSTM and BiLSTM deep learning models.

2.3 Data Preprocessing

I used several steps to preprocess input data before feeding the data into the LSMT and BiLSTSM deep learning models for training and testing purposes.

- **Sorting:** The raw dataset consisted of values ordered in descending order (2022 to 2011). However, to facilitate easy processing and understanding, the dataset was sorted in ascending order (from 2011 to 2022).
- **String into Numeric Data Conversion:** Since the raw dataset consisted of values in string format, including symbols such separators, the sorted dataset was converted into numeric (float) format so that they can be processed by the LSTM and BiLSTM deep learning models. Figure 2 shows the daily USD to TZS exchange rate (Price) spanning 12 years.
- **Data Scaling:** For effective training of deep learning models, the numeric dataset (daily exchange rate (Price)) was then scaled down to fit within a range of 0 to 1.
- **Train-Validation-Test Dataset Split:** The 12 years scaled dataset was then split into training set which (the first 60% of the scaled dataset), validation set (the next 20% of the scaled dataset) and the test set (the last 20% of the scaled dataset, which is also called unseen data, with a purpose of evaluating prediction accuracies of the LSTM and BiLSTM models).
- **Lagging:** Since only single variable (Price, alternatively called exchange rate) from the dataset was selected as input feature and this study is based on supervised deep learning, I had to create labels for the values in the input feature. The labels allow the models to learn how to correctly predict values which are close to the true values. To create labels, I used lagging where the exchange rates (input feature) of the five previous days are associated with the exchange rate of the next day (label). This allows the model to eventually learn pattern in the previous days exchange rates and be able to predict the next day exchange rate.

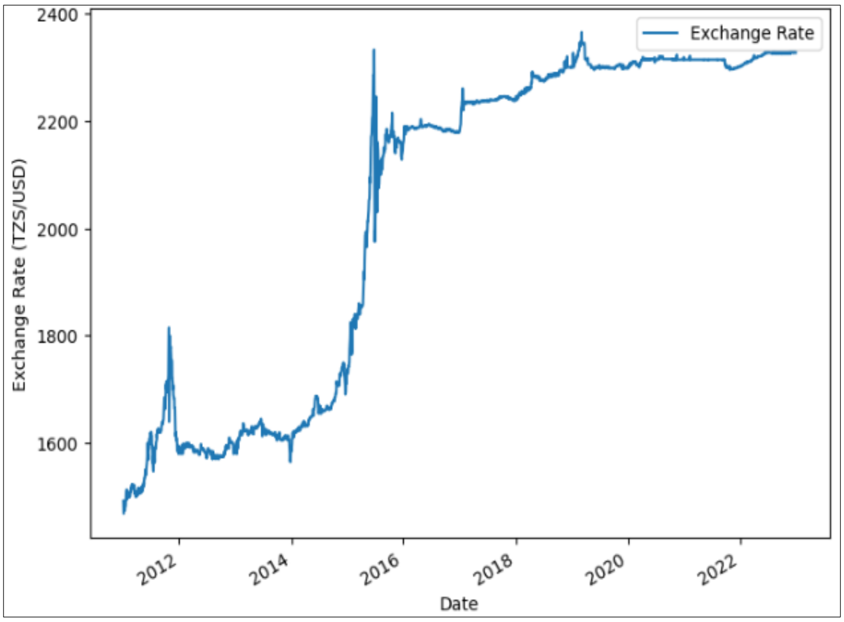

Figure 2: Daily USD to TZS Exchange Rates (January 3, 2011 to December 30th, 2022)

2.4 LSTM and BiLSTM Architectures

2.4.1 LSTM Architecture

LSTM architecture is shown in Figure 3 and consists of cell state c_t, forget gate f_t, input gate i_t and output gate o_t, with every component having its role in helping the LSTM network learning the pattern of previous days exchange rate and how to map it with next day exchange rate. LSTM learns how next day exchange rate depends on previous days (long-term information).

- **Cell State:** Cell c_t stores information over many timesteps and the contents can be added or removed from it through use of gates. Through this, LSTM network gains the ability to either retain or discard information based on how relevant it is. This is the reason why LSTM network has capability of remembering useful (relevant) information over long duration (many timesteps).
- **Forget Gate:** Forget gate f_t determines what information to not keep (discard) from the cell state through joining the previous hidden state h_{t-1} and the current input x_t and then give an output value ranging between 0 and 1 for every item in the cell state. The sigmoid function σ produces an output ranging between 0 and 1 when given any input, with an output value close or equal to 0 meaning 'don't keep' and output value close or equal to 1 meaning "keep". Equation (i) shows the forget gate, with weight matrices and bias vector parameters of the forget gate represented as W_f and b_f respectively, both of which are learned during training of the LSTM deep learning model.
  $f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$ (i)
- **Input Gate:** The input gate $\mathrm{i_t}$ determines what part of the new information should be added into the cell state. The input gate is made up of two components; the sigmoid function σ and the $tanh$ function (which is used to produce a candidate cell state $\tilde{c}_t$ whose values could be added to the cell state). The input gate components are shown in equations (ii) and (iii) with weight matrices and bias vector parameters of the input gate represented as W_i and b_i respectively and weight matrices and bias vector parameters of the candidate cell state represented as W_c and b_c respectively, all of which are learned during training of the LSTM deep learning model.
  $i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$ (ii)
  $\tilde{c}_t = tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$ (iii)
- **Output Gate:** The role of output gate o_t (refer to equation (iv)) is to control the LSTM unit output by joining the previous hidden state h_{t-1} and the current input x_t and make decision on which part of the cell state to output as the next hidden state. Weight matrices and bias vector parameters of the output gate are represented as W_o and b_o respectively, both of which are learned during training of the LSTM deep learning model.
  $o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$ (iv)
- **Updates:** The cell state and hidden state are finally updated. The forget gate and input gate are used to update the cell state as shown in equation (v). Afterwards, the cell state and the output gate are combined to update the hidden state (output of LSTM unit) as it can be seen in equation (vi).
  $c_t = f_t * c_{t-1} + i_t * \tilde{c}_t$ ....(v)
  $h_t = o_t * \tanh(c_t)$ ....(vi)

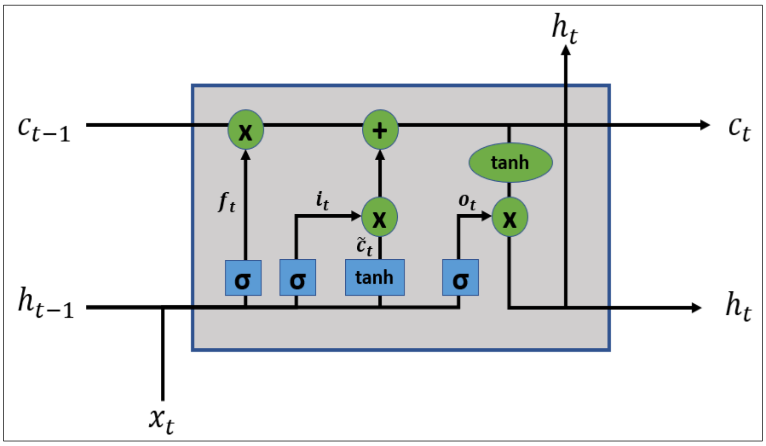

Figure 3: LSTM Architecture

2.4.2 BiLSTM Architecture

BiLSTM (refer to Figure 4) is a variant of the LSTM network. However, unlike LSTM unit which processes information in one direction (forward), the BiLSTM processes information in both forward and backward directions. The BiLSTM architecture consists of two LSTM units, forward LSTM and backward LSTM. The outputs from both forward and backward LSTM layers are then aggregated to produce a single output.

- **Forward LSTM:** This LSTM unit processes the input sequential data from the first timestep to the last timestep.
- **Backward LSTM:** The backward LSTM processes the information in reverse order to forward LSTM.

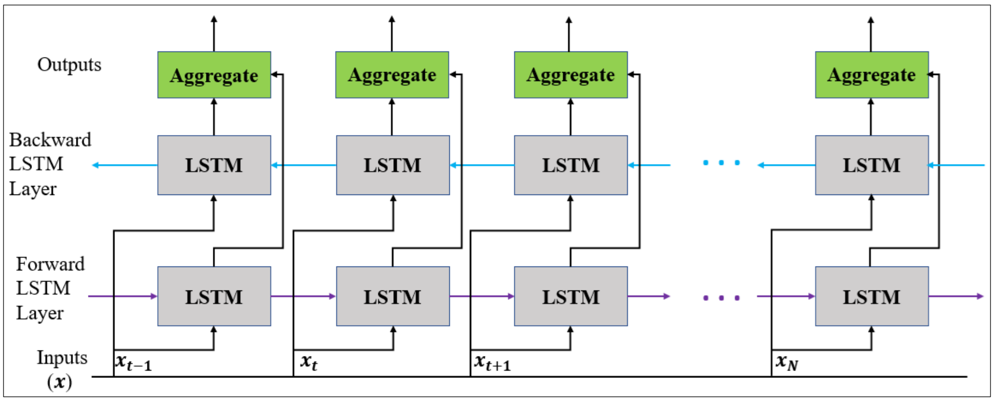

Figure 4: BiLSTM Architecture

2.5 Proposed Deep Learning Models

The proposed LSTM and BiLSTM deep learning models are shown in Figure 5 and Figure 6 respectively. The proposed LSTM model consists of two LSTM layers for learning the pattern of previous days exchange rates and a single Dense layer for predicting a single numerical value as the next day exchange rate. On the other hand, the proposed BiLSTM model consists of two BiLSTM layers for learning the pattern of previous days exchange rate and a single Dense layer for predicting a single numerical value as the next day exchange rate.

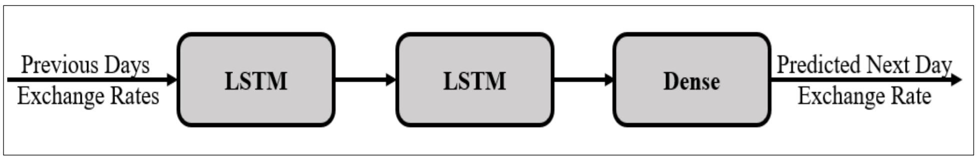

Figure 5: Proposed LSTM Deep Learning Model

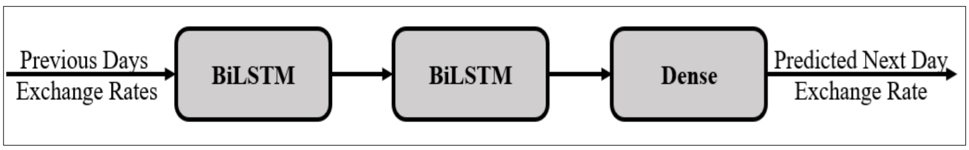

Figure 6: Proposed BiLSTM Deep Learning Model

2.6 Evaluation Metrics

2.7 Loss Function

Loss function is an important metric used to evaluate performance of the model during training. Loss function measures the prediction error between true value y_i and predicted value $\hat{y}_i$. Loss function is an important metric when updating weights of the model as it enables the model to understand how correctly it is in prediction task by comparing predicted and actual values. Through this way, the model is trained how to predict (forecast) values which are as close as possible to the actual (true). In this study, Mean Squared Error (MSE) described in equation (vii) was used as a loss function.

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

2.8 Performance Evaluation Metric

Performance evaluation is used when testing performance of the model (prediction capability) on unseen data (test set) and measure its capability to generalize when fed with completely new data. In this study, Mean Absolute Percentage Error (MAPE) described in equation (viii) was used as a performance evaluation metric, and it computes prediction error between predicted values ($\hat{y}_i$) and actual values (y_i).

$MAPE = 100\frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i-\hat{y}_i}{y_i}\right|$ ….(viii)

3. RESULTS AND DISCUSSION

3.1 Experiments Environment

All training and performance evaluation experiments with both LSTM and BiLSTM deep learning models were conducted in Google Colab, a cloud computation environment developed by Google with allocation of Random Access Memory (RAM) of 12.7 GB and Hard Disk space of 107.7 GB My study also utilized several deep learning libraries available in Colab such as TensorFlow which is used for mathematical computations and Keras which consists of instances LSTM and BiLSTM networks.

3.2 Hyperparameters Tuning

Hyperparameters play a crucial role in shaping performance of deep learning models. Tuning and obtaining best hyperparameters contribute to having the best performing model. In this study, hyperparameters tuning was conducted by having several rounds of training both LSTM and BiLSTM models, changing hyperparameters in each training round until the best hyperparameters were obtained by looking at decreasing loss (MSE) scores and how closely training loss correlates with validation loss. After several training rounds, the following best hyperparameters were chosen; two layers of both LSTM and BiLSTM units, output-dimensionality of 100 for the first LSTM and BiLSTM layers, output-dimensionality of 200 for the second LSTM and BiLSTM layers, ReLU as an activation function for both LSTM and BiLSTM units as well as batch-size of 16, learning rate of 0.01 and 200 training epochs for both LSTM and BiLSTM models.

3.3 Training of LSTM and BiLSTM Deep Learning Models

Finally, after hyperparameters tuning, both LSTM and BiLSTM models were effectively trained using the best hyperparameters, with results in Figure 6 showing lower MSE scores between predicted and actual exchange rates on training set and results in Figure 7 showing lower MAPE scores between predicted and actual exchange rates on training set. This demonstrates effective training of both LSTM and BiLSTM models.

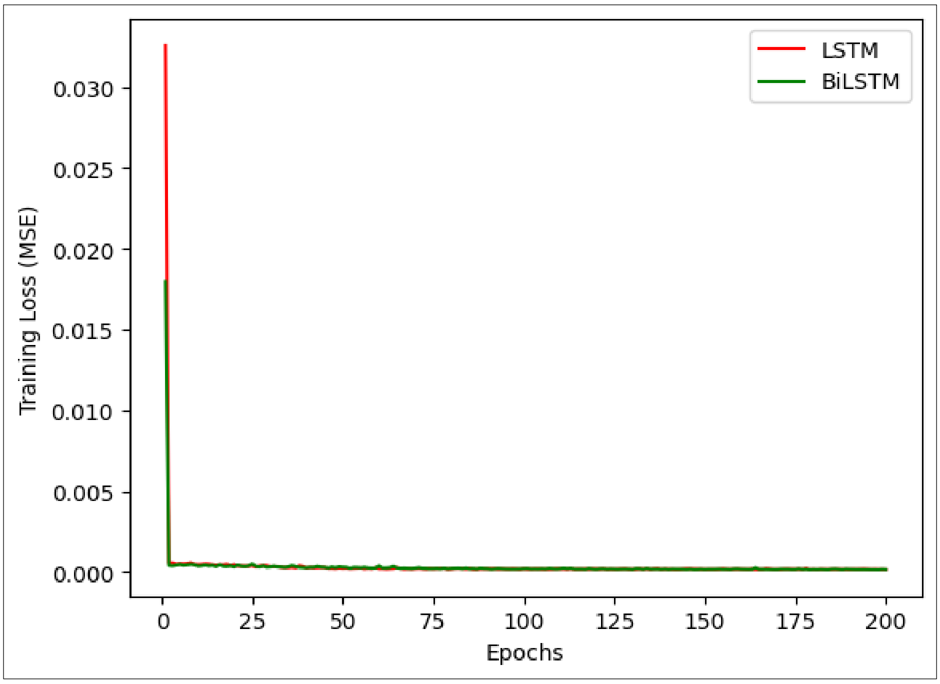

Figure 6: Training MSE

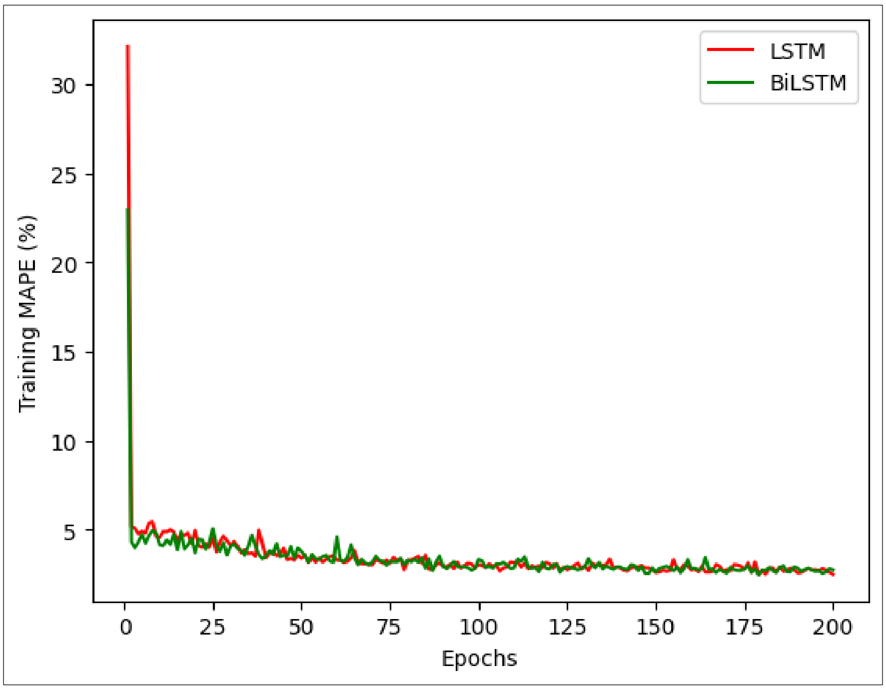

Figure 7: Training MAPE

3.4 Performance Evaluation

After training, performance evaluation of both LSTM and BiLSTM was done by evaluating their capabilities to predict next day exchange rate on test set (the data they have never seen before) and the results (refer to Table 1 and Figure 8) reveal that, BiLSTM model outperforms LSTM model in forecasting daily USD to TZS exchange rates, with BiLSTM model achieving a test MAPE score of 0.363% compared to a test MAPE score of 1.471% achieved by LSTM model. Afterwards, both LSTM and BiLSTM models were saved into .h5 format in order to facilitate future inference.

Table 1: Test MAPE Scores

Deep Learning Model	MAPE Score on Test Set (%)
BiLSTM	0.363
LSTM	1.471

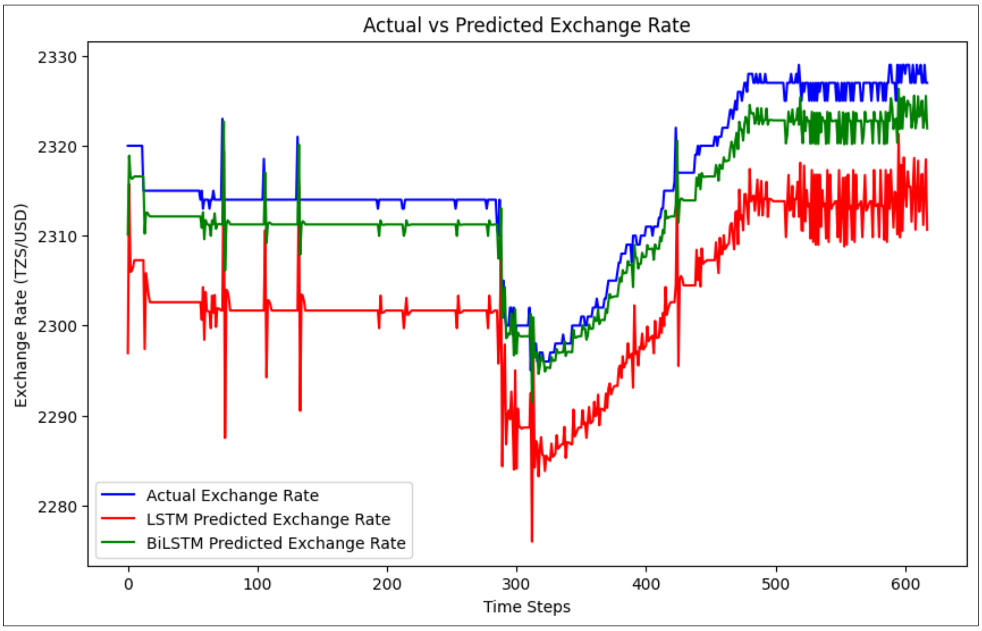

Figure 8: Predicted versus Actual Exchange rates on Test Set

3.5 DISCUSSION

These results indicate that, BiLSTM is a better deep learning model than LSTM when it comes to forecasting daily USD to TZS exchange rates. The findings align with results in literature, which also reveal superiority of BiLSTM model over LSTM in other forecasting tasks as evident in a study by Abduljabbar *et al.*, (2021) which predicted short-term traffic in Australia, with results revealing BiLSTM outperformed LSTM model and a study by Liu *et al.*, (2022) which predicted stock closing prices in China using six types of air pollutants as input features, with results revealing BiLSTM model outperformed LSTM model.

This study has the following major contributions:

- **Dataset Creation:** This study has developed a dataset which is ready to be used by future AI researchers and software developers. This dataset will save them time and cost involved with colleting and preprocessing of datasets from scratch, and instead they can focus on other research or software development activities. In future, the dataset will be hosted in open cloud environments like Zenodo (Sicilia *et al.*, 2027) or GitHub where they can freely and easily be accessed by anyone.
- **Ready to use Models:** This study has developed two novel deep learning models, LSTM and BiLSTM. These models can be used by business men and women in Tanzania who are involved in stock exchange markets, foreign exchange markets and other businesses to predict daily USD to TZS exchange rates and then take appropriate and favorable business and investment decisions and avoid losing money, for instance buying or selling stocks before they appreciate or depreciate and make profits.
- **Filling the Research Gap:** The findings of this study will help to fill the research gap on comparative performance of LSTM and BiLSTM deep learning models in forecasting daily exchange rates.

4. CONCLUSION

This study has developed LSTM and BiLSTM deep learning models and comparatively evaluate their performances in forecasting daily USD to TZS exchange rates, with the results revealing BiLSTM outperforming LSTM in terms of test MAPE scores. I therefore recommend to prospective AI researchers and software developers in Tanzania to use BiLSTM instead LSTM in predicting USD to TZS exchange rates while developing software and information systems involving prediction of USD to TZS exchange rates due to its superior performance.

5. Acknowledgements: The author would like to thank the Dar es Salaam Institute of Technology for supporting this research.

REFERENCES

- Abbass, M. J., Robert, L., & Waldemar, R. (2024). A Predictive Model Using Long Short-Time Memory (LSTM) Technique for Power System Voltage Stability. *Applied Sciences*, *14*(16), 7279. DOI: https://doi.org/10.3390/app14167279
- Abduljabbar, R., Dia, H., & Tsai, P. (2021). Unidirectional and bidirectional LSTM models for short-term traffic prediction. *Journal of Advanced Transportation*, *1*, 5589075. DOI: https://doi.org/10.1155/2021/5589075
- Bank of Tanzania [BoT]. (2025). External Sector Statistics, Total Imports of Goods and Services. Retrieved February 22, 2025, from https://www.bot.go.tz/Statistics/externalstatistics?code=EXS3&TypeOption=Imports&variableOption=Total%20imports%20of%20goods%20and%20services
- Benesty, J., Chen, J., & Huang, Y. (2008). On the importance of the Pearson correlation coefficient in noise reduction. *IEEE Transactions on Audio, Speech, and Language Processing*, *16*(4), 757-765. DOI: https://doi.org/10.1109/TASL.2008.919072
- Fan, Y., Qian, T., Yangming, G., & Yifei, W. (2024). BiLSTM-MLAM: A Multi-Scale Time Series Prediction Model for Sensor Data Based on Bi-LSTM and Local Attention Mechanisms. *Sensors*, *24*(12), 3962. DOI: https://doi.org/10.3390/s24123962
- He, Y., & Qiunan, C. (2023). Construction and Application of LSTM-Based Prediction Model for Tunnel Surrounding Rock Deformation. *Sustainability*, *15*(8), 6877. DOI: https://doi.org/10.3390/su15086877
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*, 1735–1780.
- Investing. (2025). USD/TZS - US Dollar Tanzanian Shilling, USD/TZS Historical Data. Retrieved February 22, 2025 from https://www.investing.com/currencies/usd-tzs-historical-data
- Iordan, A. (2024). An Optimized LSTM Neural Network for Accurate Estimation of Software Development Effort. *Mathematics*, *12*(2), 200. DOI: https://doi.org/10.3390/math12020200
- Lei, P., Fanglan, M., Changsheng, Z., & Tianyu, Li. (2024). LSTM Short-Term Wind Power Prediction Method Based on Data Preprocessing and Variational Modal Decomposition for Soft Sensors. *Sensors*, *24*(8), 2521. DOI: https://doi.org/10.3390/s24082521
- Liu, B., Yu, Z., Wang, Q., Du, P., & Zhang, X. (2022). Prediction of SSE Shanghai Enterprises index based on bidirectional LSTM model of air pollutants. *Expert Systems with Applications*, 204, 117600. DOI: https://doi.org/10.1016/j.eswa.2022.117600

- Marjani, M., Masoud, M., & Fariba, M. (2024). CNN-BiLSTM: A Novel Deep Learning Model for Near-Real-Time Daily Wildfire Spread Prediction, *Remote Sensing*, *16*(8), 1467. DOI: https://doi.org/10.3390/rs16081467
- Nyamrunda, G., Mbogela, C. (2014). Impacts of lower exchange rates on exports, imports, and National Output of Tanzania. *ACRN Journal of Finance and Risk Perspectives*, *3*(2), 1-9.
- Sejnowski, T. (2020). The unreasonable effectiveness of deep learning in artificial intelligence. *Proceedings of the National Academy of Sciences*, *117*(48), 30033-30038. DOI: https://doi.org/10.1073/pnas.1907373117
- Sicilia, M., García-Barriocanal, E., & Sánchez-Alonso, S. (2017). Community curation in open dataset repositories: insights from Zenodo. *Procedia Computer Science*, *106*, 54-60. DOI: https://doi.org/10.1016/j.procs.2017.03.009
- Staffini, A. (2023). A CNN–BiLSTM Architecture for Macroeconomic Time Series Forecasting. *Engineering Proceedings*, *39*(1), 33. DOI: https://doi.org/10.3390/engproc2023039033
- Wang, C. (2024). T2-LSTM-Based AI System for Early Detection of Motor Failure in Chemical Plants. *Mathematics*, *12*(17), 2652. DOI: https://doi.org/10.3390/math12172652
- Wang, L., Taifeng, L., Pengcheng, W., Zhenyu, L., & Qianli, Z. (2023). BiLSTM for Predicting Post-Construction Subsoil Settlement under Embankment: Advancing Sustainable Infrastructure. *Sustainability*, *15*(20), 14708. DOI: https://doi.org/10.3390/su152014708
- Xu, Y., Quan, P., Zengfu, W., & Baoquan, H. (2024). A Novel Trajectory Prediction Method Based on CNN, BiLSTM, and Multi-Head Attention Mechanism. *Aerospace*, *11*(10), 822. https://doi.org/10.3390/aerospace11100822
- Zhang, J., Xin, Z., Zhaojun, L., Fa, F., Yihan, J., & Fei, Xu. (2023). A Network Intrusion Detection Model Based on BiLSTM with Multi-Head Attention Mechanism. *Electronics*, *12*(19), 4170. DOI: https://doi.org/10.3390/electronics12194170

Cite This Article: Isakwisa Gaddy Tende (2025). Forecasting United States Dollar to Tanzania Shillings Exchange Rate Using Comparable LSTM and BiLSTM Deep Learning Models. *East African Scholars J Eng Comput Sci, 8*(2), 15-23.