International Journal of Advanced Research, Volume 8, Issue 1, 2025
Article DOI: https://doi.org/10.37284/ijar.8.1.2977

International Journal of Advanced Research
ijar.eanso.org
Volume 8, Issue 1, 2025
Print ISSN: 2707-7802 | Online ISSN: 2707-7810
Title DOI: https://doi.org/10.37284/2707-7810

EAST AFRICAN NATURE & SCIENCE ORGANIZATION

Original Article

Development of GRU Deep Learning Model for Predicting Daily United States Dollar to Tanzanian Shilling Exchange Rate Using Comparable Time-Lags Inputs

Isakwisa Gaddy Tende[1]*

[1] Dar es Salaam Institute of Technology, P. O. Box 2958, Dar es Salaam, Tanzania.
* Author for Correspondence ORCID ID; https://orcid.org/0009-0000-2264-7815; Email: isakwisa.tende@dit.ac.tz

Article DOI: https://doi.org/10.37284/ijar.8.1.2977

Publication Date:

09 May 2025

Keywords:

United States Dollar (USD), Tanzanian Shilling (TZS), Exchange Rate Prediction, Gated Recurrent Unit (GRU), Time-Lag, Deep Learning.

ABSTRACT

To import goods and services into the country, Tanzania relies on foreign currencies, specifically the United States Dollar (USD). Failure to timely predict accurate USD to TZS exchange rates may result in several problems, including failure to import into the country critical services and goods timely manner, losses in foreign exchange markets and bad decisions in investments. To address these challenges, this study has developed a Gated Recurrent Unit (GRU) Deep Learning model to predict the next day's USD to TZS exchange rate (output) using three different inputs (time-lags) of previous days' exchange rates. This study has also developed a Web User Interface (UI) which is integrated with the developed GRU Deep Learning model. The Web UI receives the previous days' exchange rates entered by a user as inputs, predicts the next day's exchange rate (output) and displays it to the user. The findings reveal that, 5-days time-lag (input) is the optimal (best performing) time-lag with a Mean Absolute Percentage Error (MAPE) score of 0.11%, followed by 10 days time-lag with a MAPE score of 0.20% and 15 days time-lag with a MAPE score of 1.12%, suggesting that the shorter the time-lag (input), the better the performance of the GRU model in predicting the next day's USD to TZS exchange rate (output). Therefore, this study recommends that Artificial Intelligence (AI) researchers and software developers use an optimal 5-day time-lag input when predicting the USD to TZS exchange rate using previous days' exchange rates using the GRU Deep Learning model. This study's major contributions include an operational GRU model and Web User Interface (UI) for allowing users to predict daily USD to TZS exchange rates and a pre-processed 12-year-long daily USD to TZS exchange rates dataset ready and suitable for usage in AI research and software development activities.

APA CITATION

Tende, I. G. (2025). Development of GRU Deep Learning Model for Predicting Daily United States Dollar to Tanzanian Shilling Exchange Rate Using Comparable Time-Lags Inputs. *International Journal of Advanced Research*, *8*(1), 186-199. https://doi.org/10.37284/ijar.8.1.2977

 This work is licensed under a Creative Commons Attribution 4.0 International License

Article DOI: https://doi.org/10.37284/ijar.8.1.2977

CHICAGO CITATION

Tende, Isakwisa Gaddy. 2025. "Development of GRU Deep Learning Model for Predicting Daily United States Dollar to Tanzanian Shilling Exchange Rate Using Comparable Time-Lags Inputs". *International Journal of Advanced Research* 8 (1), 186-199. https://doi.org/10.37284/ijar.8.1.2977.

HARVARD CITATION

Tende, I. G. (2025) "Development of GRU Deep Learning Model for Predicting Daily United States Dollar to Tanzanian Shilling Exchange Rate Using Comparable Time-Lags Inputs". *International Journal of Advanced Research*, 8(1), pp. 186-199. doi: 10.37284/ijar.8.1.2977

IEEE CITATION

I. G., Tende "Development of GRU Deep Learning Model for Predicting Daily United States Dollar to Tanzanian Shilling Exchange Rate Using Comparable Time-Lags Inputs", *IJAR*, vol. 8, no. 1, pp. 186-199, May. 2025.

MLA CITATION

Tende, Isakwisa Gaddy. "Development of GRU Deep Learning Model for Predicting Daily United States Dollar to Tanzanian Shilling Exchange Rate Using Comparable Time-Lags Inputs". *International Journal of Advanced Research*, Vol. 8, no. 1, May. 2025, pp. 186-199, doi:10.37284/ijar.8.1.2977

INTRODUCTION

The United States Dollar (USD) is an important foreign currency in Tanzania as it is used to import various goods and services into the country (Moussaoui, 2022), with these imports ranging from cars and tractors to food imports like wheat. Report from the Bank of Tanzania (BoT, 2025) shows that, goods and services worth 16,674.120266 million USD were imported in Tanzania in 2022, indicating the critical role of USD foreign currency when importing goods and services in Tanzania.

Since most of Tanzanians have access to and use TZS in their daily life, when the exchange rate between TZS and USD is not accurately predicted, there can be several problems including difficulty of importing critical services and goods timely and possibility of financial loss in stock and foreign exchange markets when the USD or TZS currency abruptly appreciates or depreciates. Therefore, it is important to develop accurate models which can accurately predict the USD to TZS daily exchange rate and help to address these problems.

Over recent years, Deep Learning has been showing very high accuracy in predicting timeseries parameters such as rainfall, temperature, water deficit and energy consumption (Torres et al., 2021) with Gated Recurrent Unit denoted as GRU (Wang et al., 2018) and Long Short-Term Memory denoted as LSTM (Hochreiter et al., 1997) being two of those Deep Learning models often used in prediction of different timeseries parameters accurately. On the other hand, time-lagging is concept of using previous timesteps (time-lags) values of a parameter as input feature to machine learning model to learn pattern of previous timesteps values and be able predict next timestep value of a parameter, for instance using previous values of rainfall (the last 30 days daily rainfall values) as input to the model for predicting the next day (31^{st} day) rainfall value. The following studies have applied Deep Learning models to predict various time series parameters using a time-lagging approach.

Chen et al. (2023) used a GRU model combined with reconstructed datasets and a 15-hour time-lag of historical stock price data to predict the 16^{th} day stock price, with the results indicating improvement of stock price prediction accuracy across different industries. Dip et al. (2024) used an Encoder–Decoder GRU model to predict stock and cryptocurrency prices by using a time-lag of 120 previous timesteps of data to predict the next timestep price, with results showing the proposed GRU model is significantly effective in forecasting prices. Hussain et al. (2021) used an optimised GRU model to predict traffic flow in California, United States, using input of historical traffic flow data and a 6-hour lag-time, achieving an MAPE score of 5.93. Kristiani et al. (2022) used GRU and

Article DOI: https://doi.org/10.37284/ijar.8.1.2977

LSTM models to predict air pollution (PM2.5) in Taiwan using an 8-hour lag-time of air pollutant factors and meteorological factors, with results indicating LSTM and GRU models achieved effective Root Mean Squared Error (RMSE) scores of 1.8643 and 2.6398, respectively. Zhang et al. (2023) used spatiotemporal Attention Gated Recurrent Unit (STA-GRU), LSTM and GRU models and a 12-hour lag-time of water level, discharge and precipitation to predict flood (water level), with results indicating an effective R-squared (R^2) value of STA-GRU of 0.9215, with also reduced RMSE score, and at the same time outperforming LSTM and GRU models in flood prediction. Ren et al. (2022) predicted runoff in the Yangtze River basin in China using GRU, LSTM, Recurrent Neural Network (RNN) Deep Learning models and a 7-day lag-time of past runoff data, with results indicating LSTM and GRU models outperforming RNN models in terms of RMSE scores of runoff prediction. Caicedo-Vivas et al. (2023) used an LSTM model and a 52-week lag-time of historical load values in Colombia to predict short-term load for a grid operator in Colombia, with results indicating an achievement of the best MAPE score of 1.65%. Bouktif et al. (2019) used GRU and LSTM models to predict electric load in France using 1-day and 1-week time-lags of electricity consumption data merged with meteorological variables, with results indicating that GRU and LSTM achieved the most accurate and stable prediction results. Yu et al. (2012) proposed a spatiotemporal Convolutional GRU (Conv-GRU) model to predict Mean Wave Period Field in the South China Sea using bivariate Mean Wave Period (MWP) and Significant Wave Height (SWH) as input features, with results showing the Conv-GRU model outperforms the single GRU model as it achieves an RMSE score of 0.121 metres for a 1-hour lead time. Riaz et al. (2022) used baseline GRU and LSTM models to predict multistep traffic speed in the United States using a 30-minute time-lag, achieving MAPE scores of 3.69 and 3.90, respectively.

Although reviewed studies show high effectiveness in predicting time series variables using time-lagging approach, there is still a research gap on what is the optimal time-lag to use as input in GRU Deep Learning model especially in the context of predicting daily USD to TZS exchange rate whose pattern is unique and different from other exchange rates. This study aims to fill this gap by answering one key research question: what is the optimal time-lag to use as input to the GRU Deep Learning model when predicting the USD to TZS exchange rate? This is achieved by first developing a novel GRU model and then experimenting with three time-lags in the GRU model to find the optimal (most accurate) time-lag to use as input in the GRU model. This study aims to answer one key research question: Does a shorter time-lag improve the GRU model's USD to TZS exchange rate prediction accuracy? To answer this research question, this study has three objectives; first, to develop novel GRU deep learning model which utilizes 3 time-lags to predict daily USD to TZS exchange rate, second, to comparatively evaluate performance of the GRU model in predicting USD to TZS exchange rate with the 3 different time-lags, and third, to develop a Web based UI which helps users to easily predict daily USD to TZS exchange rates, for instance based on 5 days time-lag input, on Friday a user can enter the exchange rates of Monday, Tuesday, Wednesday, Thursday and Friday itself on the Web UI and be able to view the predicted exchange rate for Saturday.

MATERIALS AND METHODS

Dataset Acquisition

The dataset was accessed from the Investing Website (Investing, 2025), a reputable source of financial information. The daily USD to TZS exchange rates from January 3rd, 2011, to December 30th, 2022, covering 12 years of exchange rates data were downloaded and saved in CSV file. The downloaded data contains several parameters; *Price* (the closing exchange rate of the day), *Open* (the

Article DOI: https://doi.org/10.37284/ijar.8.1.2977

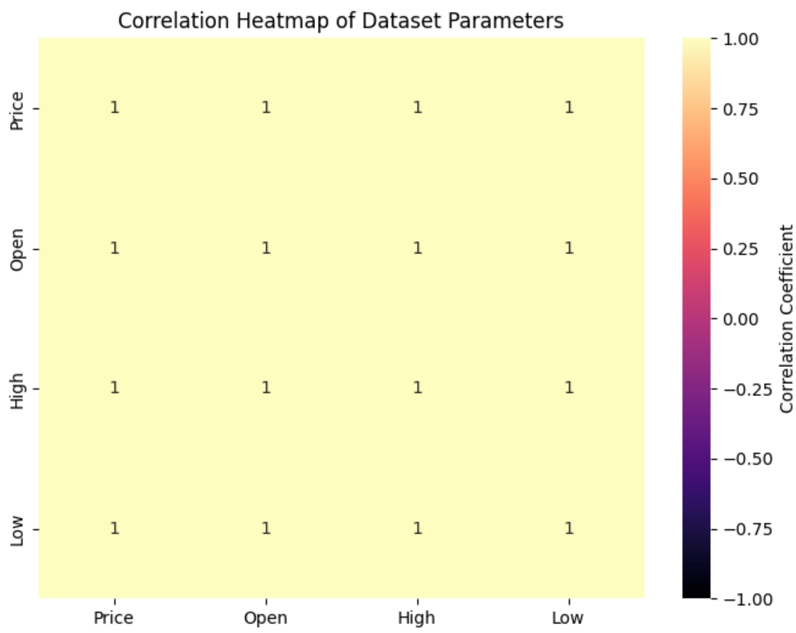
opening exchange rate of the day), *High* (the highest exchange rate of the day), *Low* (the lowest exchange rate of the day), *Vol* (volume of the daily exchange rates) and *Change %* (percentage of change from previous day exchange rates).

Input Features Selection

Not all features are relevant in prediction tasks of time series datasets by Deep Learning models (Bezerra et al., 2024). In order to identify the relevant features in a dataset to use as input to a Deep Learning model for prediction tasks, it is important to perform correlation analysis among the features. To identify the relevant input features for predicting the USD to TZS exchange rate by the GRU model, Pearson correlation analysis (Liu et al., 2023) was performed between the four key features (*Price*, *Open*, *High* and *Low*) of the dataset. The results of the Pearson correlation analysis are shown in the correlation heatmap in *Figure 1*. Correlation analysis results show that all four input features (parameters) are strongly correlated with each other (Pearson correlation coefficient (r) = 1), suggesting only one feature can be used in prediction task in GRU model and the other features will add no value in the prediction task. Since *Price*, which represents the closing (final) exchange rate of the day, presents the best indication of the exchange rate of a particular day, it was chosen as the only feature to be used in the GRU model for predicting the USD to TZS exchange rate.

Although incorporation of several input features in GRU model's prediction task can sometimes enhance noise resilience of the GRU model and help capture volatility of the market, the other three features (*Open*, *High* and *Low*) show perfect linear correlation (Pearson correlation coefficient (r) = 1) with the *Price* feature (closing exchange rate), meaning this is a statistical redundancy and these three features (*Open*, *High* and *Low*) do not contribute any additional unique information to the task of predicting daily USD to TZS exchange rates. Inclusion of perfectly correlated features would just increase GRU model's complexity without increasing its predictive accuracy and at the same time add possibility of overfitting and computation overhead. By using only *Price* feature, the GRU model remains simple and efficient while still leveraging all useful predictive information present in the USD to TZS daily exchange rates dataset.

Figure 1: Correlation Heatmap



 This work is licensed under a Creative Commons Attribution 4.0 International License.

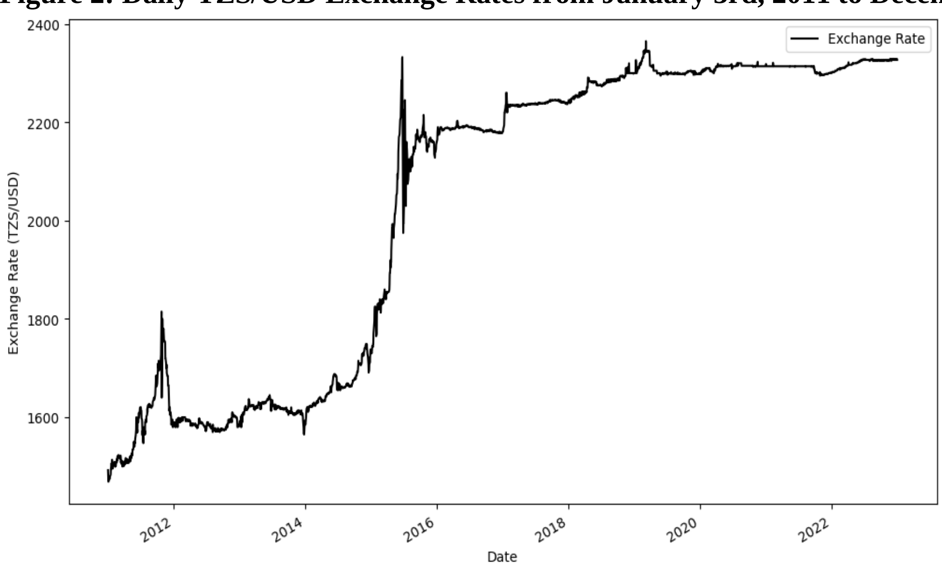
Article DOI: https://doi.org/10.37284/ijar.8.1.2977

Pre-processing of Data

Deep Learning models process numeric data, and this requires the input data to be pre-processed accordingly. This section describes the process involved in pre-processing the data.

- Analysis of Data**:** The 12 years dataset of USD to TZS exchange rates (*Price*) was visualized (refer to *Figure* 2) and analysed with the following observations: a count of 3111 records representing total number of daily USD to TZS exchange rates over a period of 12 years, a mean value of 2035.571919 TZS/USD representing an average daily exchange rate for 12 years, a standard deviation of 313.537845 TZS/USD for 12 years of daily exchange rates, a minimum of 1469 TZS/USD daily exchange rate in 12 years and a maximum of 2365 TZS/USD exchange rate in 12 years.
- Data Normalization: The daily USD to TZS exchange rate (*Price*) data for the 12 years was scaled down to fit in a range of between 0 and 1, suitable for training process in GRU model as scaled data helps faster convergence during model's training. Min-Max Scaler shown in equation (1) was used to normalise exchange rate data, where X, X_{Max}, X_{Min} and X_S represent actual, maximum, minimum and scaled values respectively.
- Training-Validation-Test Data Split: Splitting the data used in Deep Learning models is essential as it ensures prediction accuracy of a model and its ability to perform well on unseen (never seen before) data. In this study, the 12-year daily exchange rate data was split into a training set (the first 60% of the data), validation set (the next 20% of the data) and test set (the last 20%) of the data. Validation set is used to evaluate the accuracy of the GRU model during the training process and ensures proper selection of hyperparameters, while test set (unseen data) is the data used to evaluate prediction performance of the model and its ability to generalize when fed with input data that it has never seen before.
- Time-Lags and Labels: GRU model learns how to predict output (next day's exchange rate) by looking at the pattern of input time-lags (previous days exchange rates). Three time-lags (5 days, 10 days and 15 days) were selected as input. The 3 selected time-lags were alternately used as input data and the next day's exchange in a particular time-lag was used as the output (label) in each of the 3 sets of data.

Figure 2: Daily TZS/USD Exchange Rates from January 3rd, 2011 to December 30th, 2022



This work is licensed under a Creative Commons Attribution 4.0 International License.

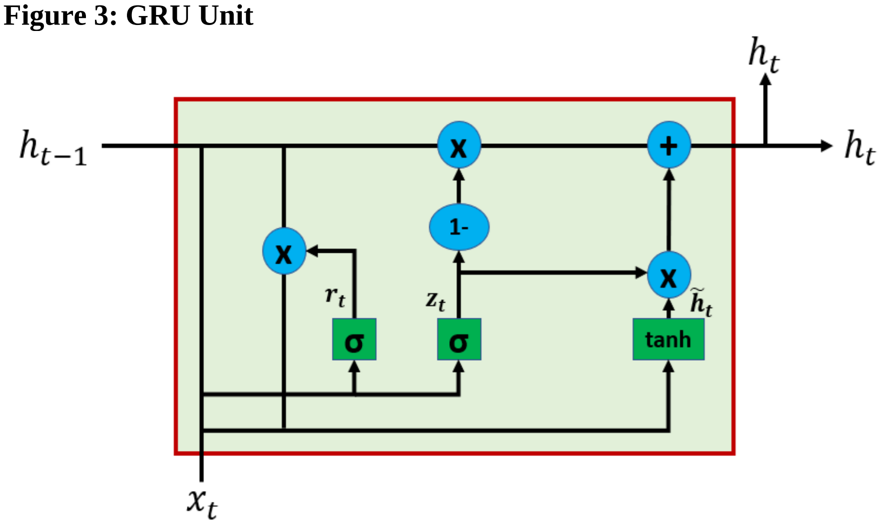
Article DOI: https://doi.org/10.37284/ijar.8.1.2977

$$X_S = \frac{X - X_{Min}}{X_{Max} - X_{Min}} \qquad (i)$$

GRU Deep Learning Model Architecture

GRU is a variant of Recurrent Neural Network (RNN), normally used to process time series variables. Just like LSTM, GRU is also used to address the problem vanishing gradients present in classical RNN models, where the model becomes unable to remember important information from previous timesteps while being trained, the information which could be useful in future timesteps. To address this problem, GRU uses several gates to retain important information from past timesteps which could be useful in future timesteps. However, unlike LSTM which consists of many parameters, GRU has relatively fewer parameters, making it more efficient computationally compared to LSTM. GRU's architecture is shown in *Figure 3*.

Figure 3: GRU Unit



At every timestep t, the GRU unit processes the current input vector x_t and the previous hidden state h_{t-1} and outputs a new hidden state h_t as described in the following steps.

- Reset Gate: The reset gate r_t is used to make decisions on how much information from the previous hidden state should be forgotten before computation of the new candidate hidden state as shown in equation (iv), where σ is the Sigmoid activation function (refer to equation (i)), W_r is the weight matrix of the input, U_r is the weight matrix of the previous hidden state and b_r is the bias term. Since the Sigmoid activation function produces values in a range of between 0 and 1, when r_t is close to 0, most of the information in the previous hidden state is forgotten, and when r_t is close to 1, most of the information in the previous hidden state is retained.
- Update Gate: The update gate z_t is used to control how much information in the previous hidden state is carried over to the next timestep as shown in equation (v), where W_z is the weight matrix of the input, U_z is the weight matrix of the previous hidden state and b_z is the bias term. When z_t is close to 0, most of the information in the previous hidden state is discarded and when z_t is close to 1, most of the information in the previous hidden state is retained.
- Candidate Hidden State: The candidate hidden state $\tilde{h}_t$ is used to control how much of the new information is injected into the current hidden state, and it does this by looking at current input and past information from the reset gate as shown in equation (vi) where $tanh$ (refer to equation

This work is licensed under a Creative Commons Attribution 4.0 International License.

Article DOI: https://doi.org/10.37284/ijar.8.1.2977

(ii)) is the activation function with e being the Euler's number, W_h is the weight matrix of the input, U_h is the weight matrix of the previous hidden state, b_h is the bias term, and $\odot$ is the element-wise multiplication. If r_t is close to 1, the past hidden state is retained, allowing the candidate hidden state to remember previous information, otherwise if r_t is close to 0, the previous hidden state is ignored and the candidate hidden state depends on the new input only.

- Final Hidden State: The final hidden state h_t is calculated using the update gate which is used to decide if the previous hidden is kept or not kept (replaced by the candidate hidden state) as shown in equation (vii). If z_t is close to 1, the hidden state remains the same, while if z_t is close to 0, the hidden state is replaced by the candidate hidden state.

$$\sigma(x) = \frac{1}{1+e^{-x}} \qquad (ii)$$

$$tanh(x)\ \frac{e^x - e^{-x}}{e^x + e^{-x}} \qquad (iii)$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \qquad (iv)$$

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \qquad (v)$$

$$\tilde{h}_t = tanh(W_h x_t + U_h(r_t \odot h_{t-1}) + b_h) \qquad (vi)$$

$$h_t = (1 - z_t) \odot \tilde{h}_t + z_t \odot h_{t-1} \qquad (vii)$$

Proposed GRU Model

Proposed GRU model's architecture (refer to *Figure 4*) contains two GRU layers and one last Dense layer with role of GRU layers being learning previous days' exchange rates (time-lags) and how to map them with the output (next day's exchange rate) while role of Dense layer being producing the single output value which acts as the predicted output of the GRU model.

Figure 4: Proposed GRU Model

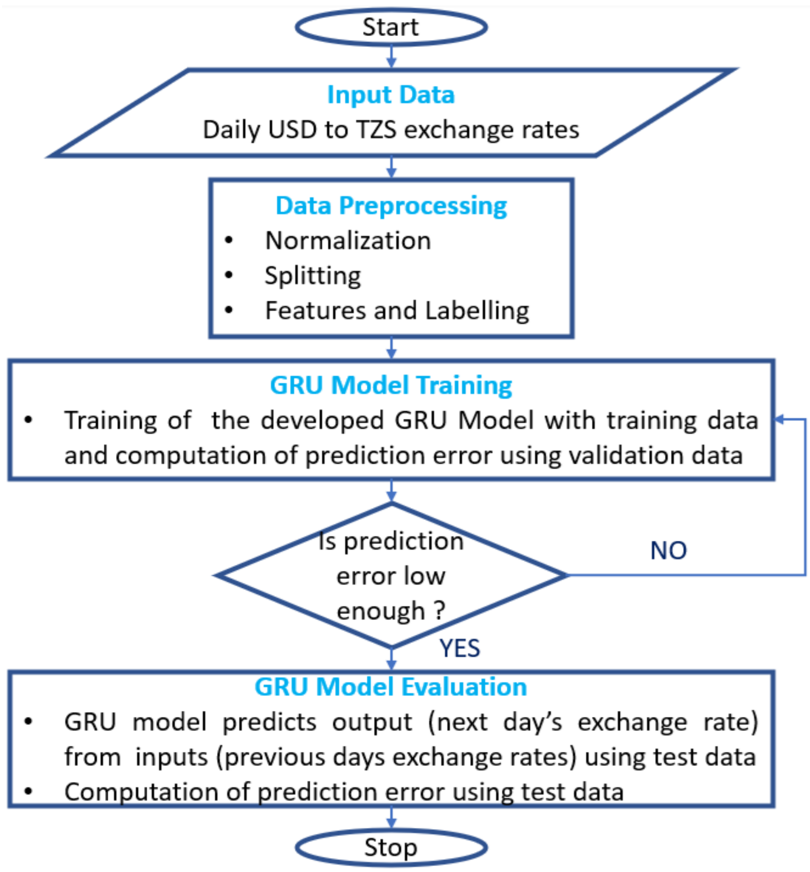
Methodology Flowchart

The flowchart in *Figure 5* summarizes the methodology used to predict the next day's USD to TZS exchange rate (output) by using inputs of previous days exchange rates by using GRU Deep Learning model as described in previous sections. The same methodology was used for each of the three input time-lags (5 days, 10 days and 15 days).

 | This work is licensed under a Creative Commons Attribution 4.0 International License.

Article DOI: https://doi.org/10.37284/ijar.8.1.2977

Figure 5: Methodology Summary



Loss Function and Performance Evaluation Metrics

The role of the Loss Function is to compute the error between the true (actual) value y and the predicted value $\hat{y}$ during training process of the GRU model, which helps the model to update its weights accordingly so as to reduce the error by predicting exchange rates which are close as possible to the actual values. Mean Squared Error (MSE) (refer to equation (viii)) is used as a Loss Function. In order to evaluate performance of the GRU model and measure its ability to generalize, it is important to test its prediction performance on test data (unseen data) by measuring the error between true (actual) and predicted values. In this study, Mean Average Percentage Error (MAPE) (refer to equation (ix)) is used as a performance evaluation metric.

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

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

Web User Interface

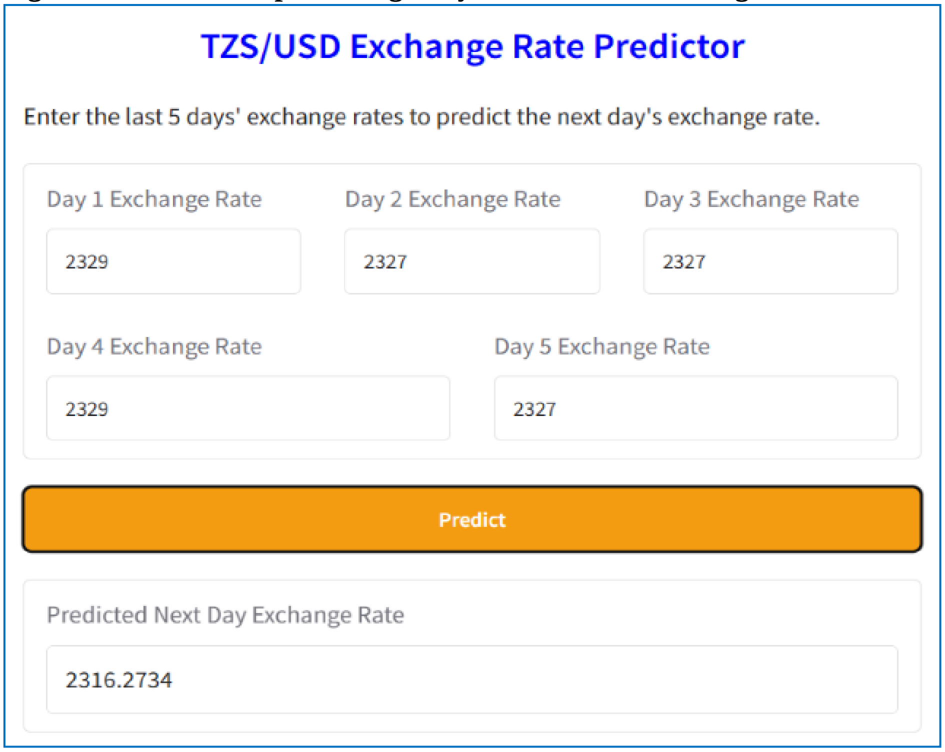
The Web UI (refer to *Figure 6*) was created by using Gradio (Ferreira et al., 2024) library, which provides CSS styled Web page for users to enter 5 previous days TZS/USD exchange rates and displays the predicted next day's exchange rate. Gradio imports (.h5) pretrained GRU Deep Learning model, takes the previous 5 days exchange rates entered by a user, passes the exchange rates to the GRU model as

This work is licensed under a Creative Commons Attribution 4.0 International License.

Article DOI: https://doi.org/10.37284/ijar.8.1.2977

input, waits for the GRU model to predict next day's exchange rate and displays the predicted next day's exchange rate back to the user.

Figure 6: Web UI for predicting daily USD to TZS Exchange Rates



RESULTS

Hyperparameters Tuning Experiments

Hyperparameters are important when finetuning the GRU model during training process as they help to get the best-performing model. The following hyperparameters were chosen after a number of hyperparameters tuning experiments: 2 GRU layers, 1 Dense layer, 1st GRU layer's output-dimensionality of 100, 2nd GRU layer's output-dimensionality of 200, batch-size of 16, learning rate of 0.001, Adam as an optimizer and 100 training epochs.

Training Experiment Results

Three training experiments were conducted for the GRU model by using a different time-lag and identical hyperparameters in each experiment. The chosen time-lags (5-days, 10-days and 15-days) were used alternately. In each training experiment the input was the time-lag (for instance 5 previous days exchange rates) and output was the next day's exchange rate (for instance 6th day's exchange rate). *Figure 7* shows the first few pairs of inputs and outputs when using input of 5-days time-lag. At the end of each training experiment, each instance of GRU model was saved in (.h5) format to facilitate future inference by users in the Web UI. The GRU model was developed in IPython notebook in Google Colab cloud environment (Bisong, 2019) with allocated run-time environment of 107.7 GB Hard Disk space and 12.7 GB RAM. *Figures 8* and *9* show the training loss (MSE) and MAPE of the GRU model. The lower training MSE and MAPE values indicate effective training of the GRU model.

 This work is licensed under a Creative Commons Attribution 4.0 International License.

Article DOI: https://doi.org/10.37284/ijar.8.1.2977

Figure 7: Sample Input and Output Pairs in 5-days Time-Lag

```
X[0]: ['1492.5', '1477.0', '1469.0', '1470.0', '1475.0'] -> y[0]: 1475.0
X[1]: ['1477.0', '1469.0', '1470.0', '1475.0', '1475.0'] -> y[1]: 1475.0
X[2]: ['1469.0', '1470.0', '1475.0', '1475.0', '1475.0'] -> y[2]: 1475.0
X[3]: ['1470.0', '1475.0', '1475.0', '1475.0', '1475.0'] -> y[3]: 1475.0
X[4]: ['1475.0', '1475.0', '1475.0', '1475.0', '1475.0'] -> y[4]: 1480.0
```

Figure 8: Training MSE for GRU Model

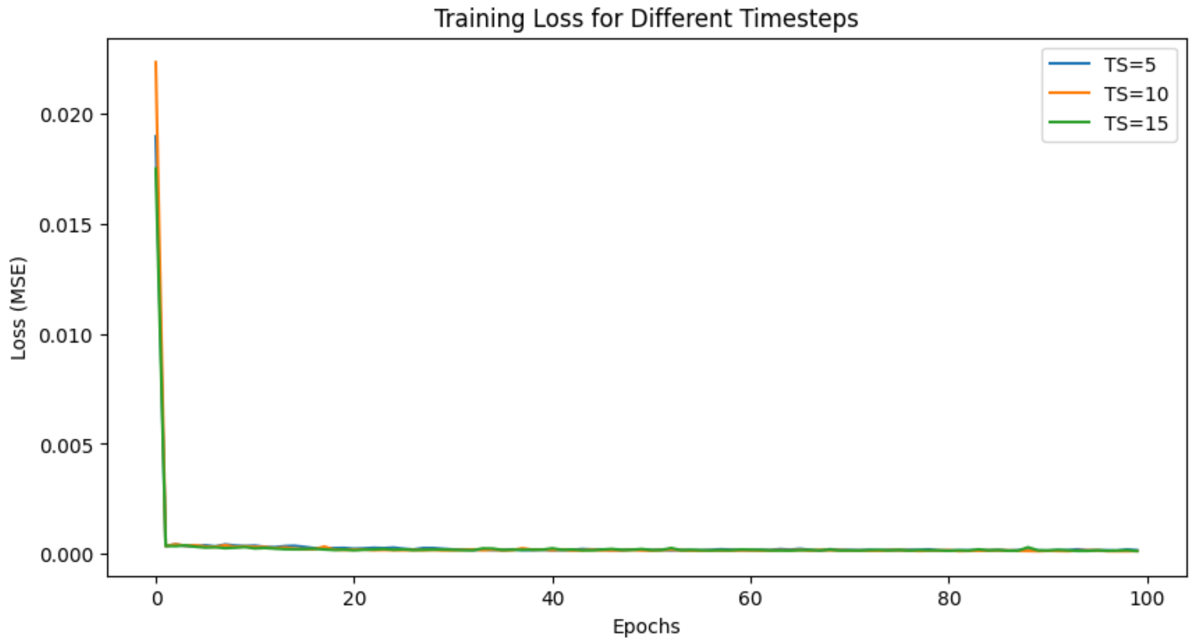
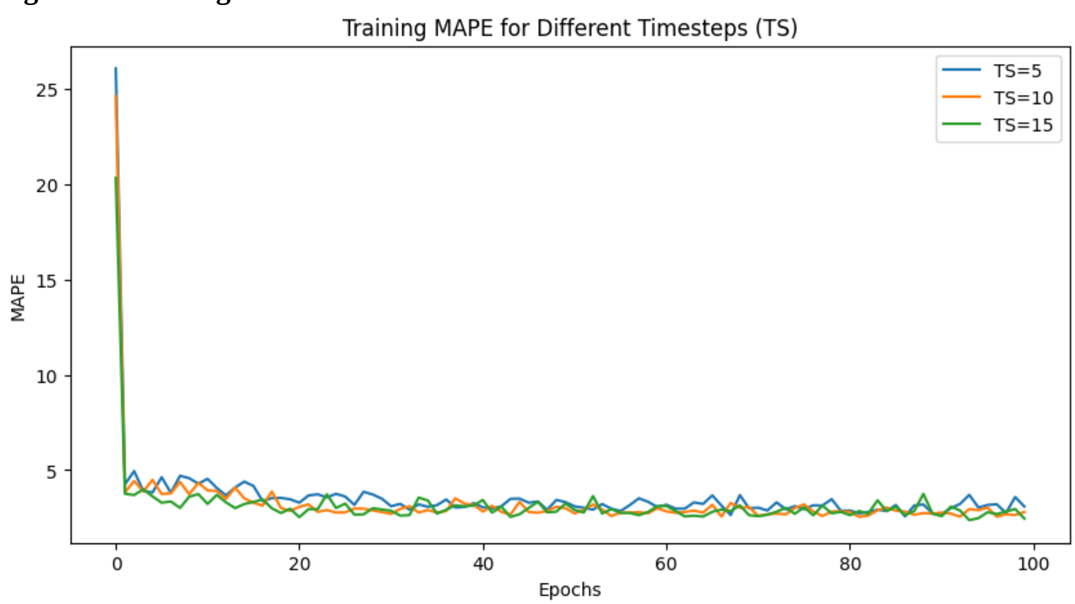


Figure 9: Training MAPE for the GRU Model



This work is licensed under a Creative Commons Attribution 4.0 International License.

Article DOI: https://doi.org/10.37284/ijar.8.1.2977

Performance Evaluation Results

After finishing training of the GRU model, each instance of the GRU model (the 3 instances correspond to 3 different inputs (time-lags)) was used to evaluate GRU's model performance on test set (unseen data). *Figure 10* shows true (actual) TZS/USD exchange rate against predicted TZS/USD exchange rate by different instances of the GRU model and *Table 1* contains test MAPE scores of the three instances of the GRU model.

Figure 10: True (Actual) vs Predicted Exchange Rates by GRU Model

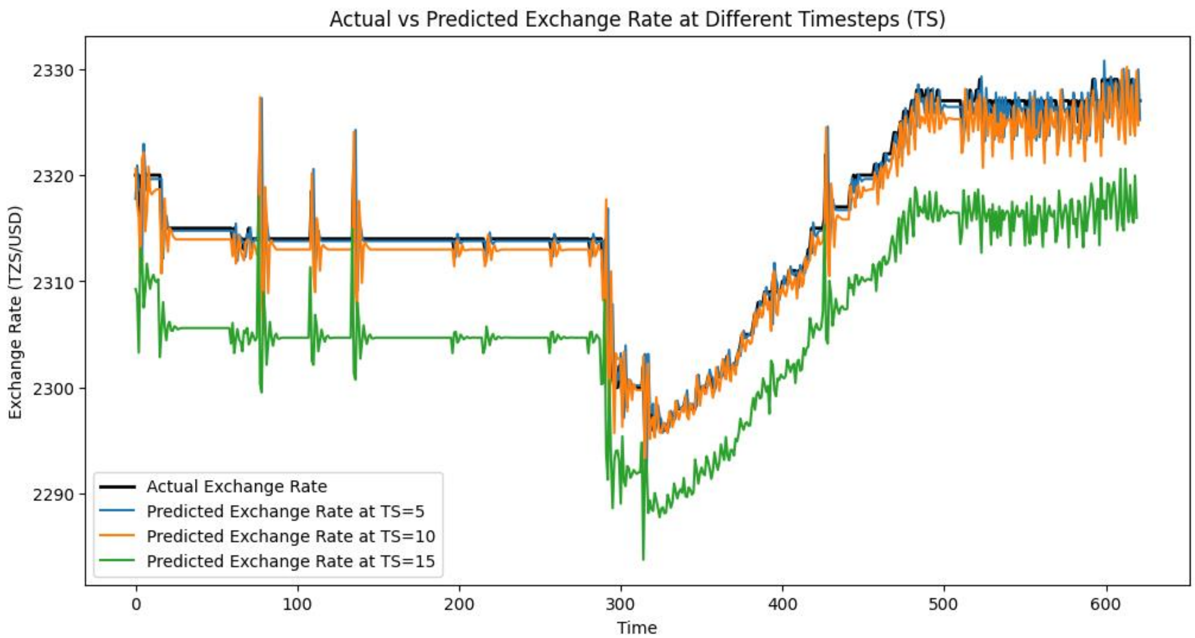


Table 1: Test MAPE Scores of GRU Model with Different Time-Lag Inputs

Time-Lag Input	Test MAPE (%)	Training Time (s)
5-Days	0.11	288
10-Days	0.20	540
15-Days	1.12	692

Results shown in *Figure 10* and *Table 1* show the performances GRU model with different time-lags when forecasting the output (next day's TZS/USD exchange rate) using input of previous days TZS/USD exchange rates (time-lags). *Figure 10* shows actual exchange rates versus exchange rates predicted by the 3 instances of the GRU model, with each instance representing a different input (time-lag). *Table 1* contains MAPE scores (Test MAPE) results of GRU model instances on the test set, revealing that 5 days is the optimal (best performing) time-lag with MAPE score of 0.11, followed by 10 days time-lag with MAPE score of 0.20 and 15 days time-lag with a MAPE score of 1.12, indicating the shorter the time-lag the better the performance of the GRU model in predicting daily USD to TZS exchange rates. *Table 1* also contains results of time taken to train each instance of the GRU model, revealing that, 5 days time-lag, 10 days time-lag and 15 days time-lag instances of the GRU model took 288, 540 and 692 seconds respectively to train. This suggests that, apart from improving prediction accuracy, a shorter time-lag is also computationally more efficient.

This work is licensed under a Creative Commons Attribution 4.0 International License.

Article DOI: https://doi.org/10.37284/ijar.8.1.2977

DISCUSSION

The findings reveal that, 5-days is the optimal time-lag to use as the input to the GRU model when predicting daily USD to TZS exchange rate, as it outperforms 10 days and 15 days time-lags, indicating that the shorter the time-lags, the better the performance of the GRU model in predicting daily USD to TZS exchange rates. These results align with the findings from literature which suggest the shorter the time-lag the better the prediction performance of Deep Learning models, as evident in a study by Zhang et al. (2023) which revealed increasing lag-time from 6 to 12 to 24 hours decreased performance of the GRU model in flood prediction resulting into increased Mean Absolute Error (MAE) scores of 0.0290, 0.0351 and 0.0607 respectively. Also, a study by Bouktif et al. (2019) revealed that increasing the time-lag from 2 weeks to 1 month reduced GRU model's performance in forecasting electric load, with Root Mean Squared Error (RMSE) increasing from 326.48 to 434.19 respectively.

The major contributions of this study include the following:

- Novel GRU Model: A novel GRU model has been developed, which uses the optimal input of 5 days time-lag to predict the daily USD to TZS exchange rate. The trained model is saved in the (.h5) format to facilitate future inference.
- Web UI: Web UI has been developed and integrated with the pre-trained (.h5) GRU model for allowing users to easily predict daily USD to TZ exchange rates.
- Preprocessed Dataset: The 12-years daily USD to TZ exchange rate dataset containing of '*Price*' parameter has been preprocessed by several methods including string to numerical data conversion, normalization and splitting into training, validation and test sets and saved in (.pkl) format and will be shared in cloud environment such as GitHub for free access by the general public interested in AI research and development.
- Filling the Research Gap: The findings fill the existing gap on which is the optimal time-lag to use as input to GRU model when predicting the daily USD to TZ exchange rate.

This study has the following limitations:

- Limited Dataset: This study uses USD to TZS exchange rate dataset which spans from 2011 to 2022, excluding recent geopolitical disruptions and market volatility. This may limit the ability of the developed GRU model to generalise to more recent market dynamics.
- Lack of Economic Context Integration: This study did not integrate macroeconomic indicators such as inflation rate and GDP growth which are known to influence the exchange rates of currencies. By excluding macroeconomic indicators, the ability of the developed GRU model to account for the basic drivers of currency movement is limited.
- Lack of Real-Time Data Integration: The developed GRU model was trained on historical (static) data with no integration of real-time exchange rate data. This may limit the ability of the developed GRU model to adapt to fast-changing market conditions.

CONCLUSION

GRU Deep Learning model for predicting daily USD to TZ exchange rate has been developed in this study and evaluated with three different inputs (time-lags) with the results revealing 5 days time-lag input is the optimal (best performing) time-lag input, followed by 10 days time-lag input, and 15 days time-lag input when predicting daily USD to TZS exchange rates, suggesting the shorter the time-lag the better the performance of the GRU model in predicting daily USD to TZ exchange rates.

 This work is licensed under a Creative Commons Attribution 4.0 International License.

Article DOI: https://doi.org/10.37284/ijar.8.1.2977

To further ensure forecasting of USD to TZS exchange rate is accurate and robust, future work will explore using ensemble models to capture diverse temporal patterns and reduce the variance of the models. Future work will also explore integrating real-time exchange rates data in prediction task, helping to develop models which are dynamic enough to respond to market changes. Future work will also explore integrating macroeconomic indicators such as inflation rates and GDP (Gross Domestic Product) growth to allow the developed models accommodate broad economic factors which influence currency movements.

Recommendations

The 5 days time-lag is recommended as the optimal input to use in GRU Deep Learning model to predict the output (next day's USD to TZ exchange rate).

Acknowledgements

The author thanks the Dar es Salaam Institute of Technology for supporting this research.

REFERENCES

Moussaoui, H. (2022). The impact of Exports and Imports on Economic Growth in Tanzania. *Asian Journal of Management, Entrepreneursh ip and Social Science*, 2(04), 150-160.

Bank of Tanzania [BoT]. (2025). External Sector Statistics, Total Imports of Goods and Services. Retrieved March 18, 2025, from https://www.bot.go.tz/Statistics/externalstatistics?code=EXS3&TypeOption=Imports&variableOption=Total%20imports%20of%20goods%20and%20services

Torres, J., Hadjout, D., Sebaa, A., Martínez-Álvarez, F., Troncoso, A. (2021). Deep learning for time series forecasting: a survey. *Big data*, *9*(1), 3-21.

Wang, Y., Liao, W., Chang, Y. (2018). Gated recurrent unit network-based short-term photovoltaic forecasting. *Energies*, *11*(8), 2163.

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

Chen, C., Xue, L., Xing, W. (2023). Research on Improved GRU-Based Stock Price Prediction Method. Applied Sciences, 13(15), 8813. DOI: https://doi.org/10.3390/app13158813

Dip Das, J., Thulasiram, R., Henry, C., Thavaneswaran, A. (2024). Encoder–Decoder Based LSTM and GRU Architectures for Stocks and Cryptocurrency Prediction. *Journal of Risk and Financial Management*, 17(5), 200. DOI: https://doi.org/10.3390/jrfm17050200

Hussain, B., Afzal, M., Ahmad, S., Mostafa, A. (2021). Intelligent traffic flow prediction using optimized GRU model. *IEEE Access*, 9, 100736-100746.

Kristiani, E., Lin, H., Lin, J., Chuang, Y., Huang, C., Yang, C. (2022). Short-Term Prediction of $PM_{2.5}$ Using LSTM Deep Learning Methods. *Sustainability*, 14(4), 2068. DOI: https://doi.org/10.3390/su14042068

Zhang, Y., Zhou, Z., Van Griensven Thé, J., Yang, S., Gharabaghi, B. (2023). Flood Forecasting Using Hybrid LSTM and GRU Models with Lag Time Preprocessing. *Water*, *15*(22), 3982. DOI: https://doi.org/10.3390/w15223982

Ren, Y., Zeng, S., Liu, J., Tang, Z., Hua, X., Li, Z., Song, J., Xia, J. (2022). Mid- to Long-Term Runoff Prediction Based on Deep Learning at Different Time Scales in the Upper Yangtze River Basin. *Water*, *14*(11), 1692. DOI: https://doi.org/10.3390/w14111692

This work is licensed under a Creative Commons Attribution 4.0 International License.

Article DOI: https://doi.org/10.37284/ijar.8.1.2977

Caicedo-Vivas, J., Alfonso-Morales, W. (2023). Short-Term Load Forecasting Using an LSTM Neural Network for a Grid Operator. *Energies*, 16(23), 7878. DOI: https://doi.org/10.3390/en16237878

Bouktif, S., Fiaz, A., Ouni, A., Serhani, M. (2019). Single and Multi-Sequence Deep Learning Models for Short and Medium Term Electric Load Forecasting. *Energies*, 12(1), 149. DOI: https://doi.org/10.3390/en12010149

Yu, T., Wang, J. (2021). A Spatiotemporal Convolutional Gated Recurrent Unit Network for Mean Wave Period Field Forecasting. *Journal of Marine Science and Engineering*, *9*(4), 383. DOI: https://doi.org/10.3390/jmse9040383

Riaz, A., Rahman, H., Arshad, M., Nabeel, M., Yasin, A., Al-Adhaileh, M., Eldin, E., Ghamry, N. (2022). Augmentation of Deep Learning Models for Multistep Traffic Speed Prediction. *Applied Sciences*, 12(19), 9723. DOI: https://doi.org/10.3390/app12199723

Investing. (2025). USD/TZS - US Dollar Tanzanian Shilling, USD/TZS Historical Data. Retrieved March 20, 2025 from https://www.investing.com/currencies/usd-tzs-historical-data

Bezerra, F., Oliveira Neto, G., Cervi, G., Francesconi Mazetto, R., Faria, A., Vido, M., Lima, G., Araújo, S., Sampaio, M., Amorim, M. (2024). Impacts of Feature Selection on Predicting Machine Failures by Machine Learning Algorithms. *Applied Sciences*, 14(8), 3337. DOI: https://doi.org/10.3390/app14083337

Liu, C., Zhang, A., Xue, J., Lei, C., Zeng, X. (2023). LSTM-Pearson Gas Concentration Prediction Model Feature Selection and Its Applications. *Energies*, 16(5), 2318. DOI: https://doi.org/10.3390/en16052318

Ferreira, R., Canesche, M., Jamieson, P., Neto, O., Nacif, J. A. (2024). Examples and tutorials on using Google Colab and Gradio to create online interactive student-learning modules. *Computer Applications in Engineering Education*, 32(4), e22729. DOI: https://doi.org/10.1002/cae.22729

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.

 This work is licensed under a Creative Commons Attribution 4.0 International License.