COEN 240: Machine Learning
Winter 2023
Santa Clara University

Prediction of Foreign Exchange Rates using a variety of Machine Learning Algorithms

Benita Rego

Meghshanth Sara

Under supervision of

Prof. Alex Sumarsono

March 17, 2023

ABSTRACT

One of the greatest financial marketplaces is currency exchange. Exchange rates are impacted by a variety of variables, including psychological, political, and economic ones. Every day, exchange rates vary, which has an impact on people's income, businesses, and even a nation's economy. Hence, forecasting currency exchange rates has various benefits for both an individual and a nation. Regression in machine learning is utilized for predicting currency exchange rates. We aim to build an efficient machine learning model (comparison based on our research through various technical publications) and check that to predict the currency exchange rates. Although Artificial Neural Networks (ANN) also suits best for this prediction, we chose Recurrent Neural Networks Algorithm variants as well as it showed better accuracy in our observed research. We will be using the live currency dataset which contains data of many time series with varying foreign exchange rates of different countries.

TABLE OF CONTENTS

TABLE OF FIGURES

1. Introduction

The greater accessibility of data, the increase in processing power, and the acceptance of machine learning techniques have revolutionized the financial sector. The Wall Street Journal (2017b) reported that quantitative hedge funds accounted for 27% of all trading activity in 2017, which is comparable to the 29% reflected by all private investors[1]. The majority of these organizations use machine learning to guide their investment decisions.

Despite the emergence of data-driven strategies, there is a paucity of research on machine learning techniques for financial forecasting. The majority of articles in this area are devoted to market return forecasting. The first thorough method for measuring the impact of applying machine learning (ML) to the forecast of monthly stock returns is presented by Gu, Kelly, and Xiu (2018) [2].

One of the largest financial marketplaces is the trade of currencies. At the moment, 73.02 Indian rupees are worth 1 US dollar. The economic, governmental, and even psychological variables are just a few that have an impact on exchange rates. The prediction of the currency rate is a challenging issue due to these reasons.

When you deal in foreign exchange, you sell one money and purchase another. You make money if the value of the currency you purchase rises relative to the value of the money you offer for sale, and you do this online via an intermediary as a consumer trader using the meta trader trading tool. It is one of the most difficult jobs because only 2% of individual dealers can accurately anticipate changes in currencies in the forex market. In the quickly evolving area of market forecasting, machine learning as well as its variants or mixed methods are gaining ground.

The FOREX market is a highly liquid and expansive trading platform that provides dealers and buyers with a vast array of commodities that can be exchanged for monetary value [3]. The average daily transaction volume in the forex market in April 2016 was 5.1 trillion dollars, based on the Federal Reserve Bank of International Settlements.

The foreign exchange market, also known as FOREX, offers a diverse range of valuable for trading, including foreign currencies, precious jewels, and oil.Among the most expensive things in the world is gold. Gold was first offered by investors as a selling commodity in opposition to foreign currencies. In the realm of FOREX dealing, machine learning (ML) is typically used to forecast future FOREX prices.

2. Objective

Our main motivation was to compare the Machine Learning algorithms: ANN, LSTM and GRU and compare them to check which algorithm provides the best accuracy in prediction for the provided database based on the time-series trend. This project helps to find the currency rate comparison in future for better understanding and also useful for international transfer of money and in the banking sector. To announce methods for forecasting currencies in the forex market that produce the greatest precision and the fewest errors. These are the USD/INR values that we are striving for. With the aid of its tools, such as keras, matplotlib, etc. The Python programming language was used to create various supervised machine learning models on Google Collaboratory Platform to compile and test our machine learning models line by line.

After conducting our research, we have determined that Artificial Neural Networks (ANNs) and variants of Recurrent Neural Networks (RNNs) are the most closely efficient techniques for foreign exchange prediction. However, these models work differently as per the data collected which we will observe in our project.

3. Problem Statement

This project aims to analyze historical trends of exchange rates between two currencies: USD against INR currency and further, predict the trend of the values. We have web scrapped live currency data from https://in.investing.com/currencies [2]. For this project, we have scrapped data from January 1, 2016 to January 31, 2023.

The project further focuses on interpreting a suitable model for this dataset and a larger set of data in order to solve a real world problem when it comes to prediction in the real foreign exchange market.

We will be testing these machine learning algorithms:

- Artificial Neural Networks (ANNs)
- Recurrent Neural Networks (RNNs)
    - Long Term Short Memory (LSTM)
    - Gated Recurrent Unit (GRU)

4. Literature Review

4.1 Paper 1: Forex Market Forecasting using Machine Learning: Systematic Literature Review and Meta-Analysis

The forex market presents a significant challenge in predicting currency movement, with a mere 2% of retail traders achieving success. Therefore, the domain of market forecasting has witnessed significant progress, with machine learning along with its byproduct or combination models gaining momentum. The academic community has extensively scrutinized the forecasting techniques employed by researchers in the forex market[4]. However, there is still a pressing need to investigate the effectiveness of machine learning and artificial intelligence methodologies in predicting forex market trends.

The analysis findings indicate that the evaluation metrics that are prevalently utilized encompass MAE, RMSE, MAPE, and MSE against EUR/USD which are the most probably utilized across the globe. The machine learning algorithms that are most commonly utilized for FX market prediction are LSTM and Artificial Neural Network. The study has also brought to light a number of unresolved matters and obstacles that the scientific community must tackle in the times ahead.

4.2 Paper 2: Foreign Exchange Prediction using Machine Learning Approach: A Pilot Study

The realm of Foreign Exchange (FOREX) trading extends beyond foreign currencies to encompass various commodities, including Gold, Silver, and Oil. Given its high value, Gold has emerged as a preferred trading material against foreign currencies. This pilot research seeks to identify a machine learning model that can forecast FOREX with a reasonable amount of precision. This experimental research makes use of historical data from the investing.com database; the FOREX data used is FOREX XAU/USD data, covering the years 2019 through 2021. The Moving Average of Convergence/Divergence (MACD) technical evaluation is the indicator used[5].

The average accuracy achieved after training the Tree model is 86.3%, whereas the SVM model and the Ensemble model attain 86.6% and 86.55%, respectively. Upon testing, all three models exhibit an identical level of accuracy, which is 88.3%.

4.3 Paper 3: Currency Exchange Rate Forecasting using Machine Learning Techniques

The effect of machine learning algorithms for predicting time series data has been widely recognized. A multitude of investigations have been carried out regarding this subject matter, utilizing diverse algorithms. The purpose of this investigation is to evaluate and differentiate four of the proposed models, specifically Long Short-Term Memory (LSTM), Radial Basis Function (RBF), Backpropagation, and Support Vector Regression (SVR). The development of four distinct models to predict the prices of three currency pairs in the Forex market was the focus of this study.

Based on a thorough literature review, additional models were suggested. Based on the findings, it can be inferred that the Support Vector Regression (SVR) model demonstrated superior performance in comparison to the remaining three models. Conversely, the Backpropagation model in neural networks exhibited the lowest level of performance among the models.[6].

5. Theory and Experiments

5.1 Visualization and Analysis of Dataset

The dataset is from a live currency update website called https://in.investing.com/ [7]. where we web scrapped the required foreign exchange data needed for our analysis and model testing. For this project, we are observing USD (currency of the United States of America) against INR (currency of India) over the time period of January 1, 2016 to January 31, 2023 i.e 7 years, up to date.

The dataset consists of the dates with the price for that particular date along with the percentage of change observed by the website which includes the increase and decrease in the rates.

The sample data is shown below:

	Date	Price	Open	High	Low	Volume	Chg%
0	2023-01-31	0.01223	0.01227	0.01227	0.01219	0	-0.30%
1	2023-01-30	0.01227	0.01227	0.01228	0.01224	0	0.00%
2	2023-01-29	0.01227	0.01227	0.01227	0.01227	0	0.00%
3	2023-01-27	0.01227	0.01228	0.01228	0.01225	0	-0.07%
4	2023-01-26	0.01228	0.01227	0.01229	0.01226	0	0.07%

Figure 1: Head of the scraped data

We must change our first column, date, to index in order to solve the time-series issue.

There are two ways to accomplish this; in one, you can say index_col='Date' and =True while reading a CSV file in Pandas. If you provide these two parameters, Pandas will preprocess the Timestamp column (i.e., Date) while importing and allocate it to the relevant index of the info frame.

The subsequent method involves post-reading, wherein you may alter the data type of the date column to DateTime subsequent to reading the data, and subsequently allocate it to the data frame's number. The plot below depicts it after data pre-processing which explains the trend over the 7 years.

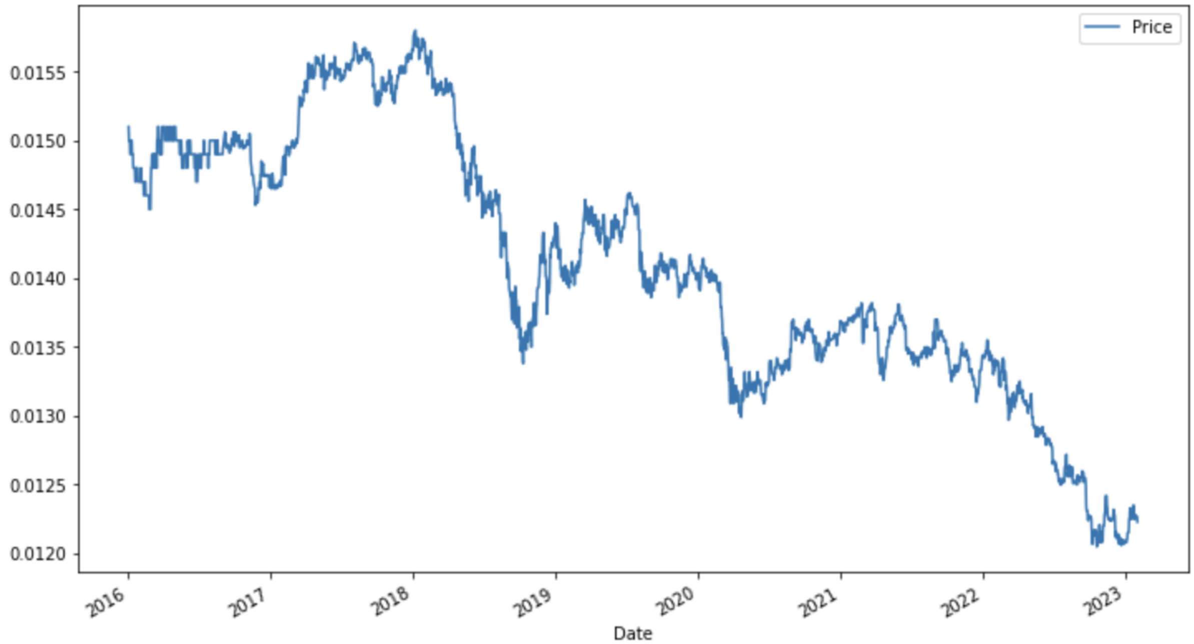

Figure 2: Data Visualization

5.2 Statistics and Training and Testing Split

After summarizing our statistics, we observe that mean for the entire dataset and choose it to distribute our data for training and testing. We chose to divide it into 80% training data i.e January 2016 to January 2021 which is approximately 5 years data and 20% testing data i.e January 2021 to January 2023. Figure 4 depicts the train test split visualization.

	Price
count	2217.000000
mean	0.014097
std	0.000940
min	0.012050
25%	0.013420
50%	0.014040
75%	0.014900
max	0.015800

Figure 3: Statistics

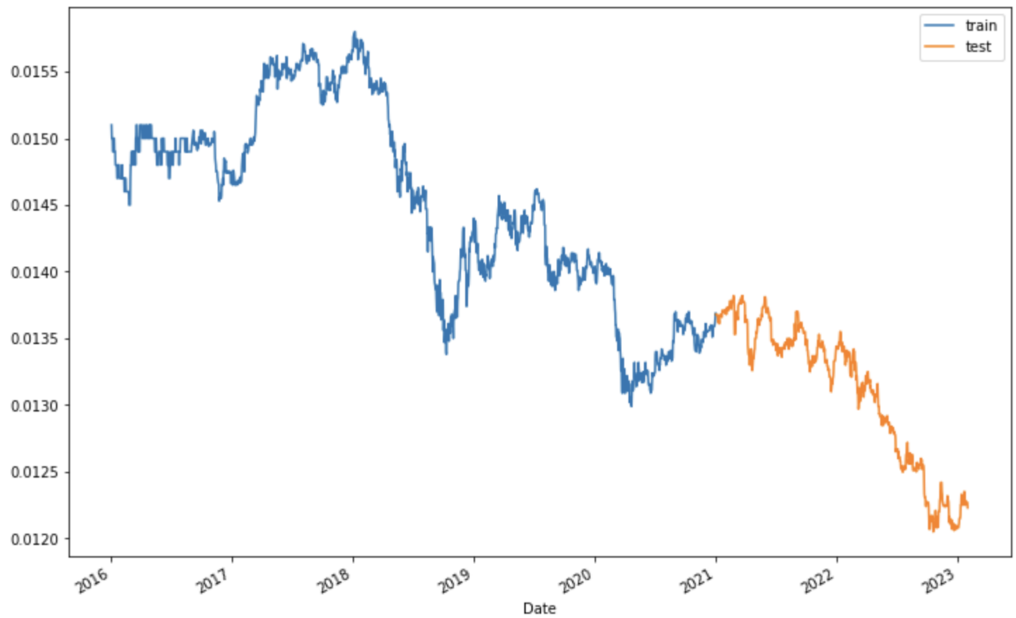

Figure 4: Train-Test Split Data Visualization

5.3 Machine Learning Algorithms

Based on our research through various technical publications, our aim is to develop a competent machine learning model that can be utilized to forecast currency exchange rates. It has been determined that Artificial Neural Networks (ANN) are the most effective method for this prediction.

5.3.1 Artificial Neural Networks (ANNs)

ANNs, is a type of machine learning model, designed to replicate the configuration and functionality of biological neural networks observed in the human brain. They are made up of numerous interconnected processing units, known as neurons, that collaborate to process incoming data and provide predictions as output.

Three primary sorts of layers make up an ANN's basic structure:

- Input layer: This layer receives input data and transmits it to the following layer. At the input layer, each neuron stands in for one feature or input variable.
- Hidden layer(s): The processing of the input data and the extraction of pertinent characteristics is done by one or more hidden layer(s). The hidden layer neurons acquire input from the layer above, evaluate a weighted sum of the inputs, employ an activation

function, and then dispatch the output to the layer above.

- Output layer: The ultimate forecast or output is produced by this layer. One possible output class or value is represented by each neuron in the output layer.

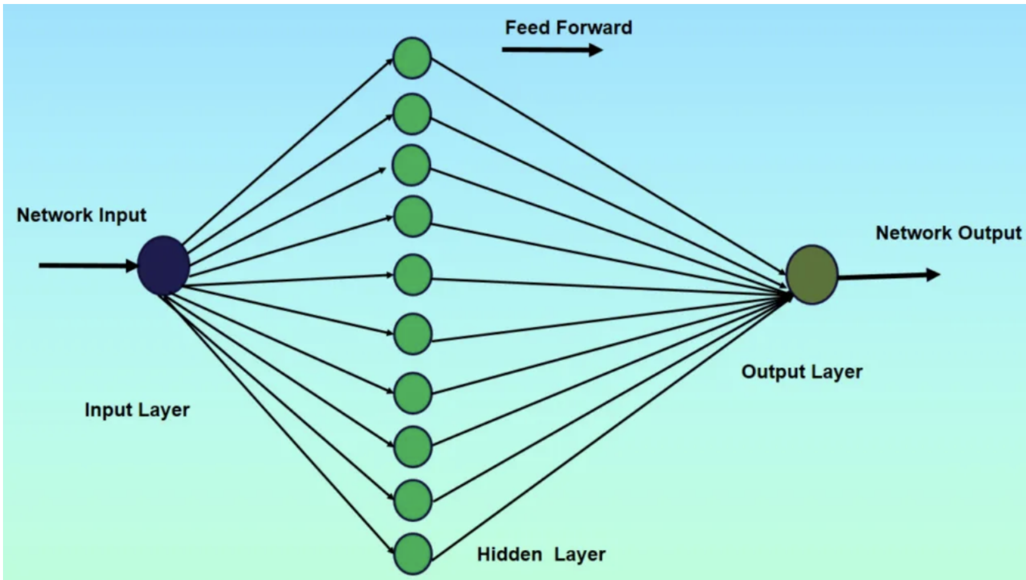

Figure 5: Artificial Neural Networks (ANNs)

Our ANN model comprises a total of 37 parameters, rendering it a lightweight model that mitigates the risk of overfitting. The loss function employed in this regression problem is mean_squared_error. The following function computes the mean squared error (MSE), which is a measure of the squared average of the error. At its core, it gauges the mean of the squared deviation between the estimated values produced by the model and the factual values. The mathematical expressions for the mean squared error are provided below.

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

With respect to the metric, the Root Mean Square Error (RMSE) has been incorporated, which is the square root of the average squared error. Essentially, this represents the standard deviation of the residuals, which indicate the deviation between the predicted and true values. The residuals act as a measure of the gap between the set of points and the true regression line.

The implementation of early stopping involves terminating the training phase of a model when it no longer demonstrates any progress in loss reduction. The monitor parameter is utilized to indicate the hyper-parameter that we intend to track throughout the training process of the model. As per the simulation, it is apparent that the model terminated its training process after 13 epochs since no further progress was observed.

```
The R2 score on the Train set is:        0.962
The  MAE on the Train set is:    0.155
The  RMSE on the Train set is:   0.194
The Adjusted R2 score on the Train set is:       0.962

The R2 score on the Test set is:         0.792
The Adjusted R2 score on the Test set is:        0.791
The  MAE on the Test set is:    0.245
The  RMSE on the Test set is:   0.332
```

Figure 6: Model Evaluation for ANN

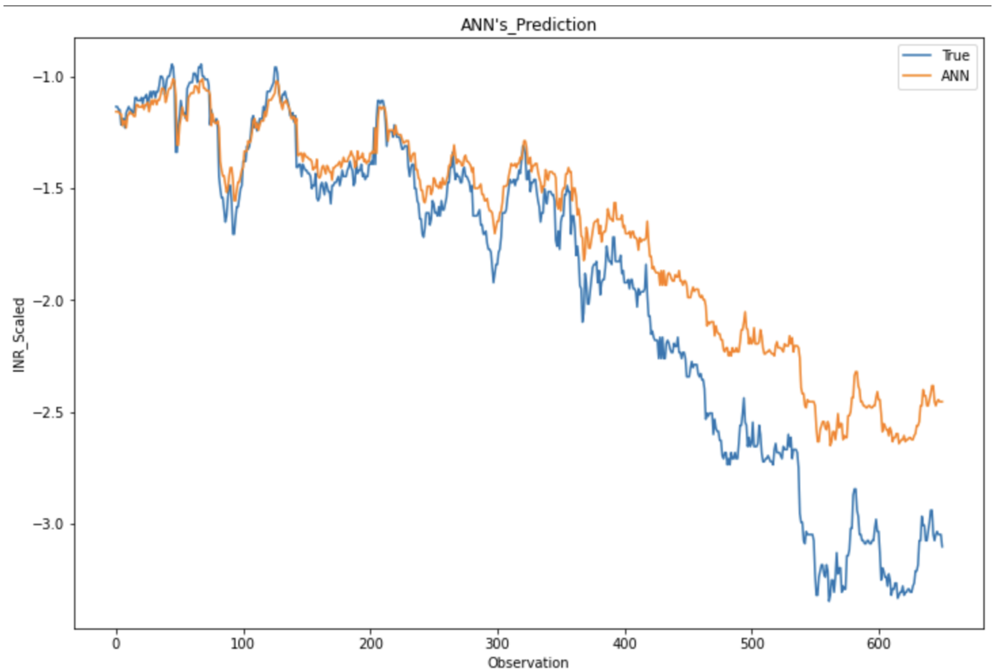

Figure: 7: ANN Prediction Graph

5.3.2 Recurrent Neural Networks (RNNs)

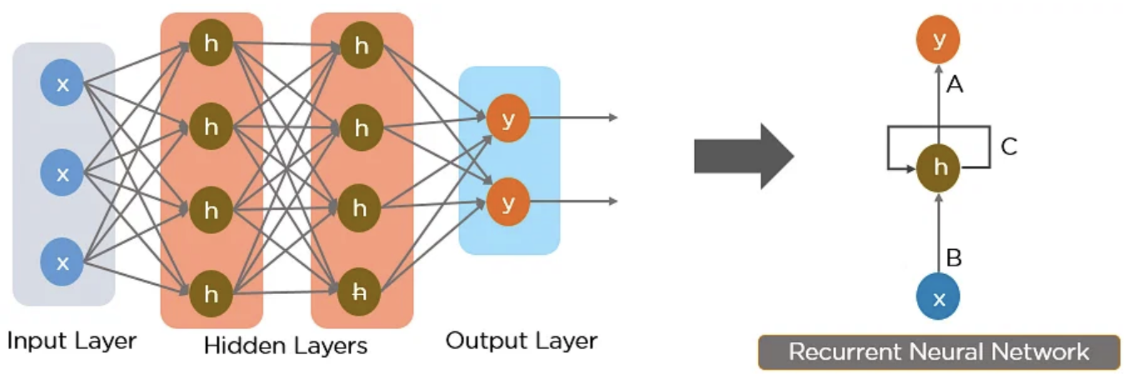

Figure 8: Recurrent Neural Networks (RNNs)

Sequential data is processed by RNNs, a neural network type, through the use of feedback loops. In contrast to feedforward neural networks, which analyze input data in a single pass, RNNs may process input sequences of varying lengths by leveraging their internal state to keep track of previous inputs.

Three essential parts make up an RNN's fundamental architecture:

- Input layer: At every time interval, the input layer acquires the information and transmits it to the subsequent layer.
- Hidden layer(s): Recurrent connections in the hidden layer enable the network to keep track of previous inputs. A weighted sum of the input and the previous hidden state is computed by each neuron in the hidden layer using data from the time step used previously as input. An activation function is utilized to compute the sum and generate the neuron's output and new hidden state.
- Output layer: The ultimate prediction or output is generated by the output layer based on the hidden state at the current time step.

Backpropagation through time (BPTT), a variant of the conventional backpropagation algorithm, can be used to train RNNs. In BPTT, the weights are updated at each time step while the error gradient is propagated back through time between the input and output layer.

5.3.2.1 Long Short-Term Memory (LSTM)

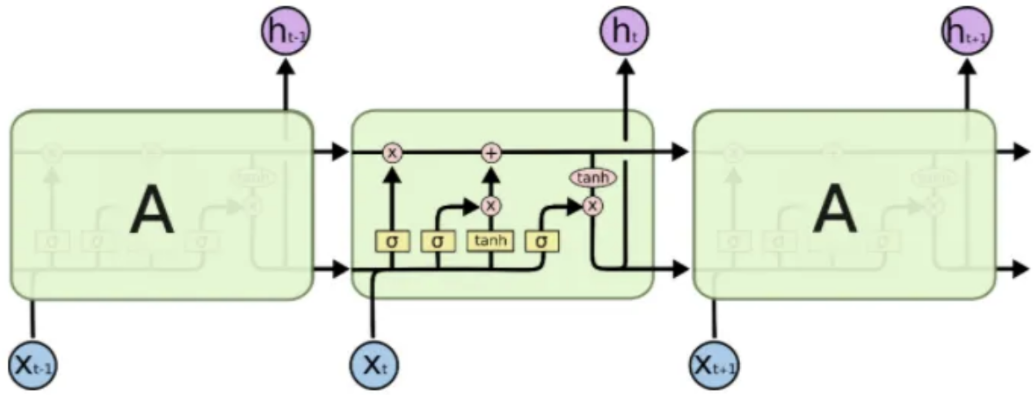

Figure 9: Long Short-Term Memory (LSTM)

The Long Short-Term Memory (LSTM) is a category of Recurrent Neural Network (RNN) that was created to overcome the obstacle of gradient vanishing in customary RNNs. LSTMs excel at processing sequential data with long-term dependencies, such as speech, language, and music.

Three gates plus a memory cell make up the fundamental structure of an LSTM:

- Forget gate: This gate produces a value between 0 and 1 for each component of the memory cell based on the inputs of the current input and the prior concealed state. The information contained in the memory cell at that element will either be kept or forgotten depending on this value.
- Input gate: The input gate produces a numerical value ranging from 0 to 1 for every element of the memory cell. The amount of fresh data to be stored in the memory cell at that element will depend on this value.
- Output gate: The current input and the previous hidden state are taken as input by this gate, which then generates a value between 0 and 1 for every element in the memory cell. The amount of data output to the following layer is determined by this value from the information stored in the memory cell.
- Memory cell: The LSTM's current state is kept in this vector. At each time step, the update of the input gate, forget gate, and previous hidden state are utilized to modify the current hidden state.

The LSTM cell was configured with 50 neurons, utilizing the relu activation function, lecun_uniform kernel initializer, and a return sequence of False. The LSTM model's training process was halted after 40 epochs owing to early stopping. From the outcomes derived from the LSTM model, it can be deduced that the shallow ANN network outperforms the LSTM model in this particular scenario.

Based on the plot presented below, it is evident that LSTM performs better in the initial observations, however, it deteriorates in comparison to ANN in the later stages. It is advisable to investigate alternative hyper-parameter combinations for LSTM in order to ascertain whether its performance can be improved.

```
The R2 score on the Train set is:        0.917
The  MAE on the Train set is:    0.213
The  RMSE on the Train set is:   0.288
The Adjusted R2 score on the Train set is:       0.917

The R2 score on the Test set is:         0.603
The Adjusted R2 score on the Test set is:        0.602
The  MAE on the Test set is:     0.326
The  RMSE on the Test set is:    0.459
```

Figure 10: Model Evaluation for LSTM

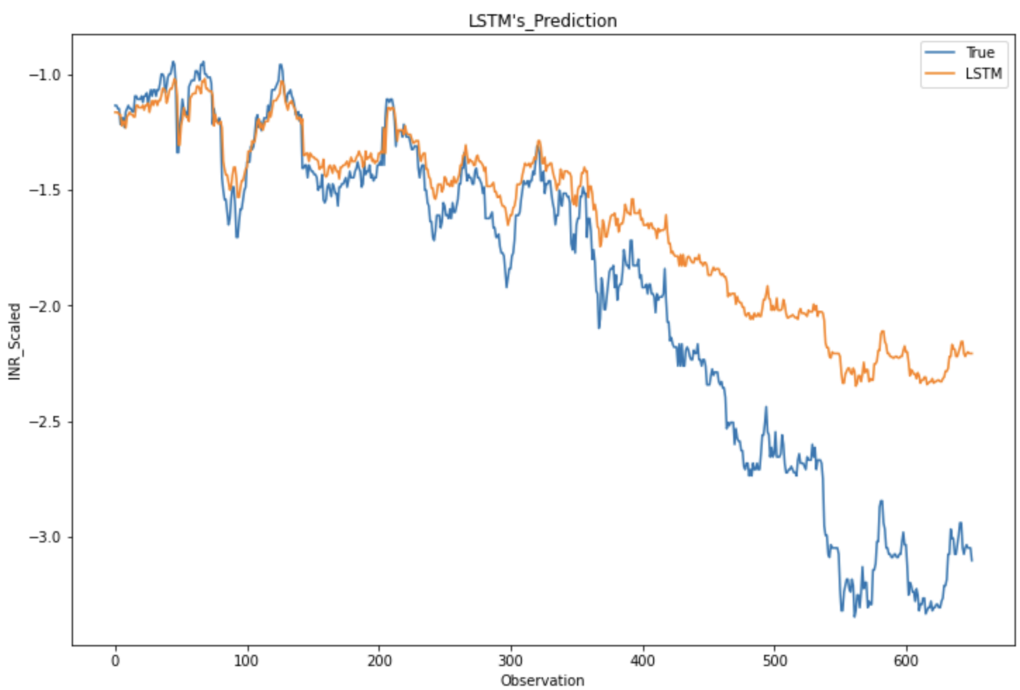

Figure: 11: LSTM Prediction Graph

5.3.2.2 Gated Recurrent Unit (GRU)

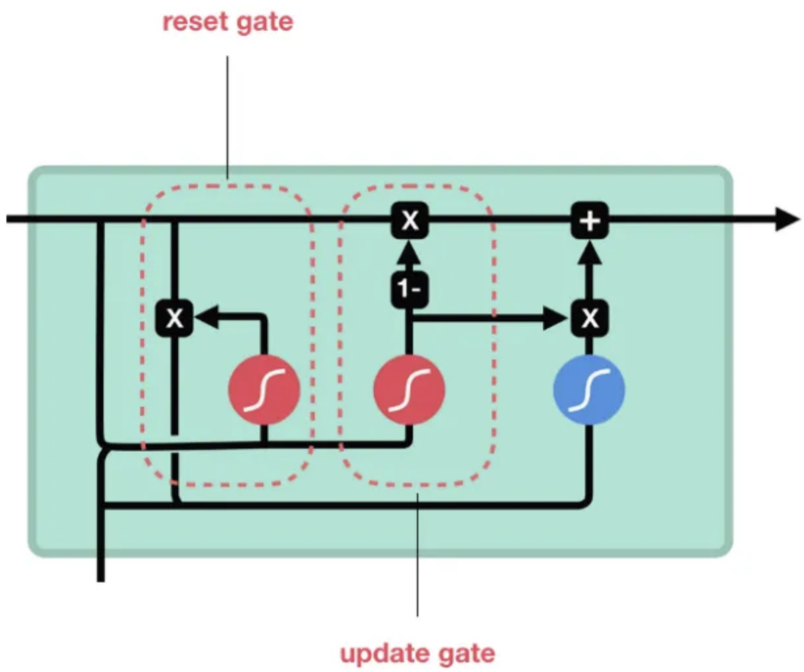

Figure 12: Gated Recurrent Unit (GRU)

A type of recurrent neural network (RNN) called gated recurrent units (GRUs) is intended to solve the vanishing gradient problem and enhance the training of long-term dependencies in sequential data. Similar in structure to LSTMs, GRUs have fewer parameters than LSTMs, which increases their computing efficiency.

The two main gates of GRU are update gate and reset gate that GRUs utilize to manage the information flow through the network:

- Update gate: The present gate generates an output ranging from 0 to 1 for every component of the concealed state, while taking in the current input and the previous hidden state as inputs. The value at hand determines the degree to which the preceding concealed state is preserved and fresh input is assimilated into the novel concealed state.
- Reset gate: The present gate is designed to produce an output ranging from 0 to 1 for every component of the hidden state. It takes in the current input and the previous hidden state as inputs. The value of the previous hidden state to be reset and the value of the fresh input to be included in the new hidden state are determined by this output..

A lightweight GRU model has been constructed for our project, featuring a hidden layer consisting of only 7 neurons. This has led to a total parameter count of 197. As per our training findings, it has been ascertained that the model discontinues its training procedure after accomplishing 44 epochs, which is most likely due to early cessation.

```
The R2 score on the Train set is:       0.953
The  MAE on the Train set is:   0.168
The  RMSE on the Train set is:  0.217
The Adjusted R2 score on the Train set is:      0.953

The R2 score on the Test set is:        0.773
The Adjusted R2 score on the Test set is:       0.773
The  MAE on the Test set is:    0.247
The  RMSE on the Test set is:   0.347
```

Figure 13: Model Evaluation for GRU

Based on the outcome obtained from the GRU model, it can be inferred that the performance of GRU is significantly inferior to that of the shallow ANN network. Furthermore, it is comparable to the performance of the LSTM network in terms of Forex rate prediction.

We can see from the figure that the GRU forecast is very far off from the actual values, but that after 400 observations, it has a much smaller deviation from the actual price than an ANN does. Therefore, it is evident from the findings above that our chosen model for predicting foreign exchange rates is ANN. As we had changed the real numbers during standardization using the standard scaler, we further performed an inverted transform of the expected and actual values.

Figure 14: GRU Prediction Graph

6. Result

After conducting a thorough analysis of the outcomes generated by the three models, it has been determined that ANN has exhibited superior performance in comparison to LSTM and GRU for the given task, thereby showcasing praiseworthy results.

	Price	ANN_prediction	RMSE
count	652.000000	652.000000	34.000000
mean	0.013096	0.013270	0.082457
std	0.000536	0.000363	0.011932
min	0.012050	0.012576	0.075346
25%	0.012580	0.012936	0.075832
50%	0.013320	0.013397	0.076135
75%	0.013490	0.013572	0.076624
max	0.013820	0.013772	0.103639

Figure 15: Result Statistics

7. Conclusion

The present study involves a comprehensive examination of the Forex dataset pertaining to the USD/ INR currency exchange rate spanning across a time frame of 7 years. Following essential pre-processing, the ANN model was adopted as the fundamental model, and subsequently, LSTM and GRU were employed.

Upon careful evaluation of the results produced by the three models, it has been ascertained that ANN has surpassed both LSTM and GRU in terms of performance for the assigned task. However, a significant observation was done while performing the analysis. We tried and tested our model for two datasets, one being the original dataset for the project which spans for 7 years and the other which is over a period of 10 years. As per our observations, it has been observed that ANN delivers better results with significantly lesser data, whereas RNN variants outperform with a larger dataset.

Therefore, we can conclude that the machine learning algorithms chosen for the project, work differently for varying sizes of datasets. This would not only aid while conducting research but also help with real world prediction products where it plays a vital role in the foreign exchange market across the globe.

8. References

[1] Christian Gonzalez Rojas , Molly Herman, *Foreign Exchange Forecasting via Machine Learning*, https://cs229.stanford.edu/proj2018/report/76.pdf

[2] Jonathan Felix Pfahler, *Exchange Rate Forecasting with Advanced Machine Learning Methods,* MDPI - Journal of Risk and Financial Management, December 21, 2021, https://www.mdpi.com/1911-8074/15/1/2

[3] Chikashi Tsuji, *Exchange Rate Forecasting via a Machine Learning Approach, Scientific Research*, Vol. 14 No. 3, Chuo University, Tokyo Japan, September 2022, https://www.scirp.org/journal/paperinformation.aspx?paperid=118439

[4] Michael Ayitey Junior, Peter Appiahene, Obed Appiah, Christopher Ninfaakang Bombie, *Forex market forecasting using machine learning: Systematic Literature Review and meta-analysis,* SpringerOpen - Journal of Big Data, Published January 28, 2023 , Article - 9, https://journalofbigdata.springeropen.com/articles/10.1186/s40537-022-00676-2

[5] Sudimanto, Yaya Heryadi, Lukas, Antoni Wibowo, *Foreign Exchange Prediction Using Machine Learning Approach: A Pilot Study,* IEEE, October 20, 2021, Yogyakarta, Indonesia, https://ieeexplore.ieee.org/document/9563998

[6] Wazir Mohammadi, *Currency Exchange Rate Forecasting using Machine Learning*, Graduate School of Applied Science, Near East University, Nicosia 2019, http://docs.neu.edu.tr/library/6721800683.pdf

[7] Project Dataset - https://in.investing.com/currencies/inr-usd-historical-data