Daily Exchange Rates per Euro(2012-23)

Project Report

Avik Das(MDS202112) Meghna Mondal(MDS202123)

13/04/23

Abstract: We focus on building the model/s to predict the foreign exchange rates and then discuss and compare the empirical results. Two different approaches have been taken into consideration and then the results are evaluated in terms of mean squared error and R-square. The first model focuses on finding the relation between the data points using statistical methods, ARIMA (Autoregressive Integrated Moving Average) is a type of regression analysis that also takes into account the randomness in the data point for predicting the future values. The second model focuses on building a special type of neural network called recurrent neural network (RNN). Here, the inputs are not features but the past data, based on which future values are predicted. The results of the two models are then compared and based on the results, which model is accurate for currency prediction concludes the report.

Introduction: The market of foreign exchange has been on significant growth since the last few decades. Due to internationalization in businesses across the globe, the exchange rate is considered to be a crucial factor in majors business across the world.

The main objective of the project is to build a system that can predict the exchange rate for the next day or any of future day/s with the highest accuracy. This will help people to have a look ahead at currency exchange rate before making any kind of transaction.

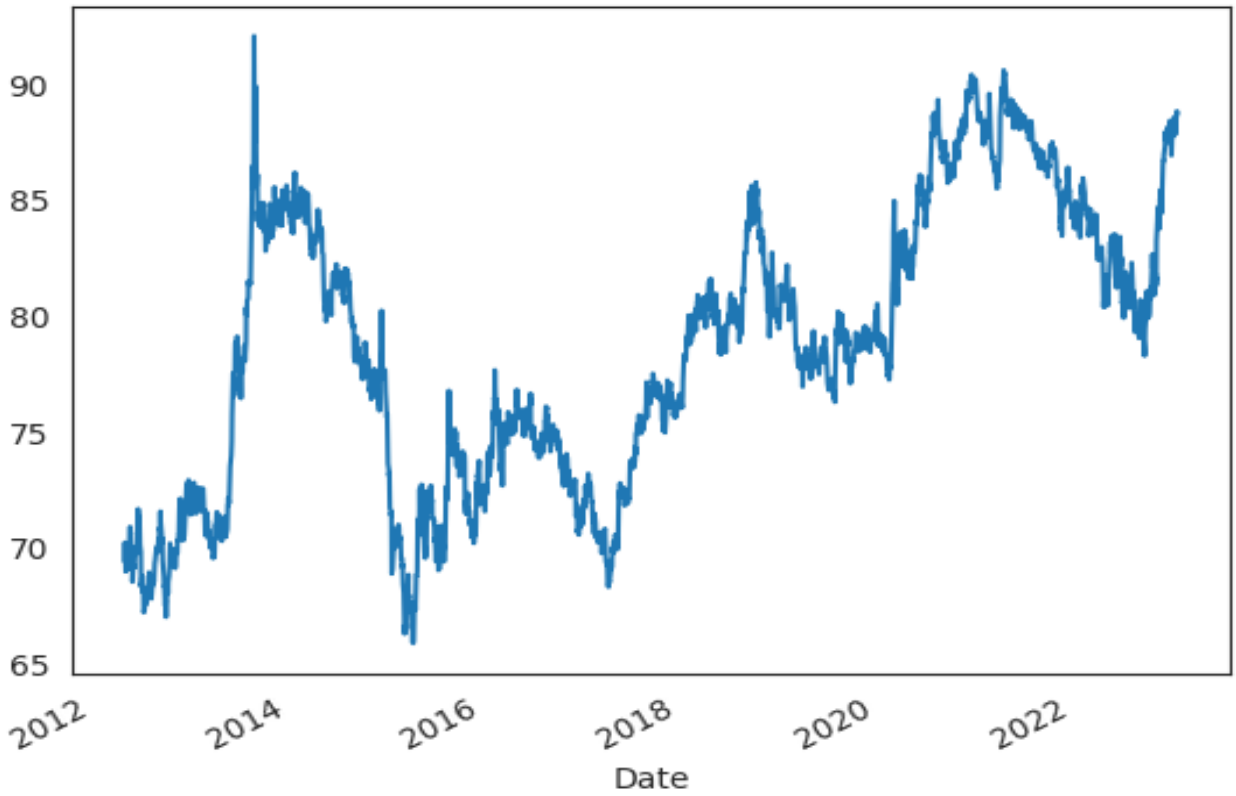
Our focus is to perform time series analysis for all major currencies against the Euro. By using the model/s we can predict the future values based on observed values, given in terms of time series data of currency exchange rate values of each day.

Using time series models like the ARIMA model can help achieve this goal of creating a currency prediction system. We have also implemented a neural network model and made a comparison between those two models to check which one performs and produces a better outcome. Our evaluation technique involves observing the actual currency value (for the day for which we estimated) and calculating the mean squared error and R-square for

ARIMA. We plan to achieve this by using Python as a programming language and using couple of libraries like Pandas (to handle time series data), Matplotlib (for plotting various graphs), Statsmodels (to use and evaluate the ARIMA model), Keras (deep neural network library).

About Dataset: Our dataset contains data points from 1st MAY,2012 - 27th JAN,2023 a total of 2752 observations five working days a week except saturday and sunday. Reference rates are euro foreign exchange rates observed on major foreign exchange trading venues at a certain point in time = they are the price of one currency in terms of another currency. The rates are usually updated around 16:00 CET on every working day.

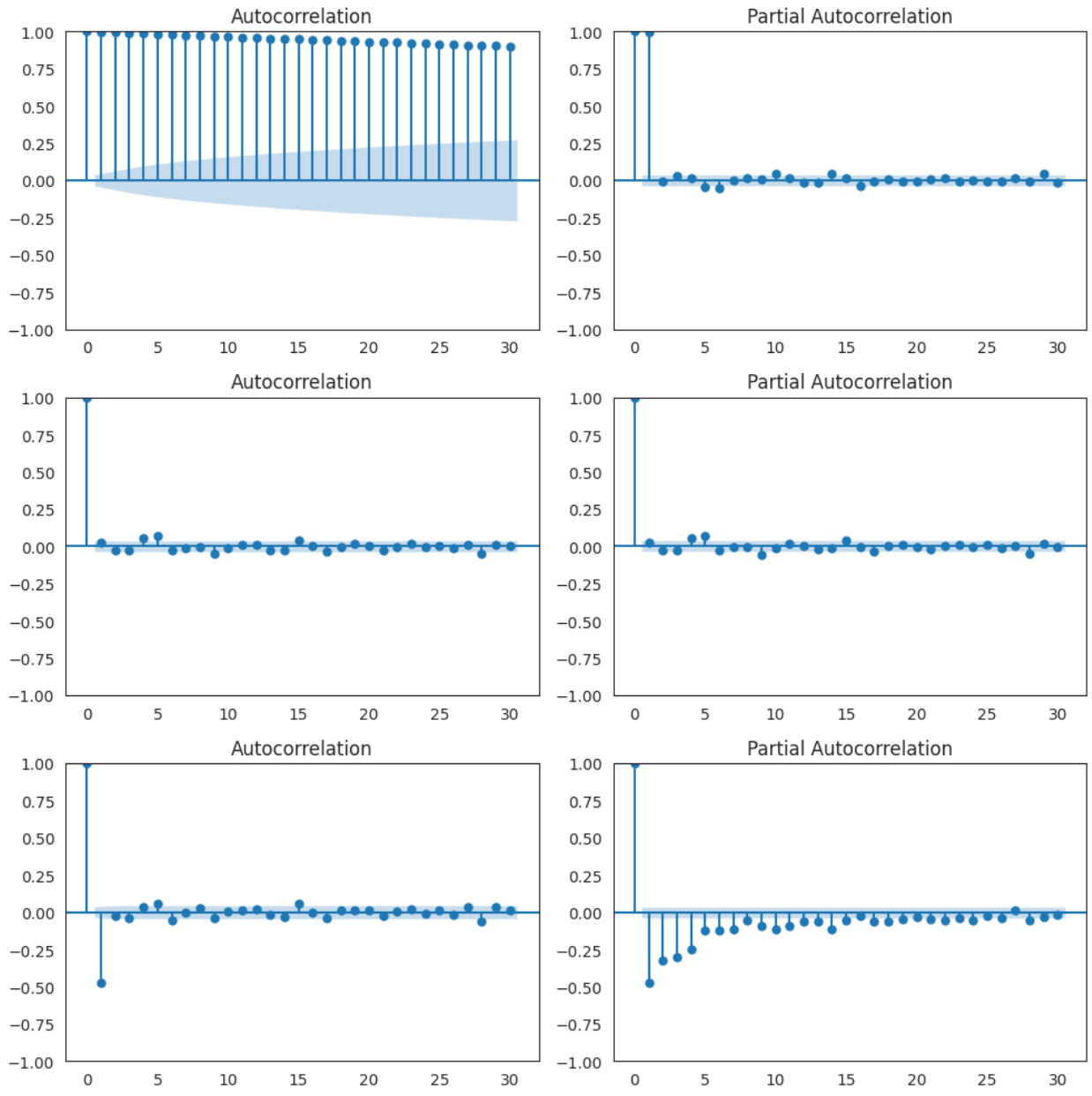
ARIMA model: For this project, the first model that we built was the ARIMA (Autoregressive Integrated Moving Average) model. ARIMA model is a statistical model which is used to forecast the dataset that has time series nature. The ARIMA model is nothing but a composition of three models namely Auto-regression(AR), Integration(I), and Moving Average(MA). The auto-regression part of the ARIMA model is used to do a regression on the current data point based on the previous data points mentioned. These previous data points are also called lagged points. Let p denote the number of lag points needed to do the regression. Secondly, the integration factor of the ARIMA model is used to make the total number of differences between the currently observed value and previous lag observations. This is called the degree of differencing. Let d denote the degree of differencing needed in the ARIMA model. Lastly, the moving average factor of the ARIMA model is used to calculate the regression error caused by moving average model when applied to the lagged observations. Let q be the size of the moving average inputs needed to calculate the regression error. The values p, d and q are called the parameters of the ARIMA model. These parameters can help to fit the ARIMA model and make future predictions appropriately. First of all, we need to identify what parameter values of the ARIMA model are required to fit the currency exchange dataset. We plot the graph of currency exchange rate between two countries namely Euro (Europe) and India (INR) which is as follows,



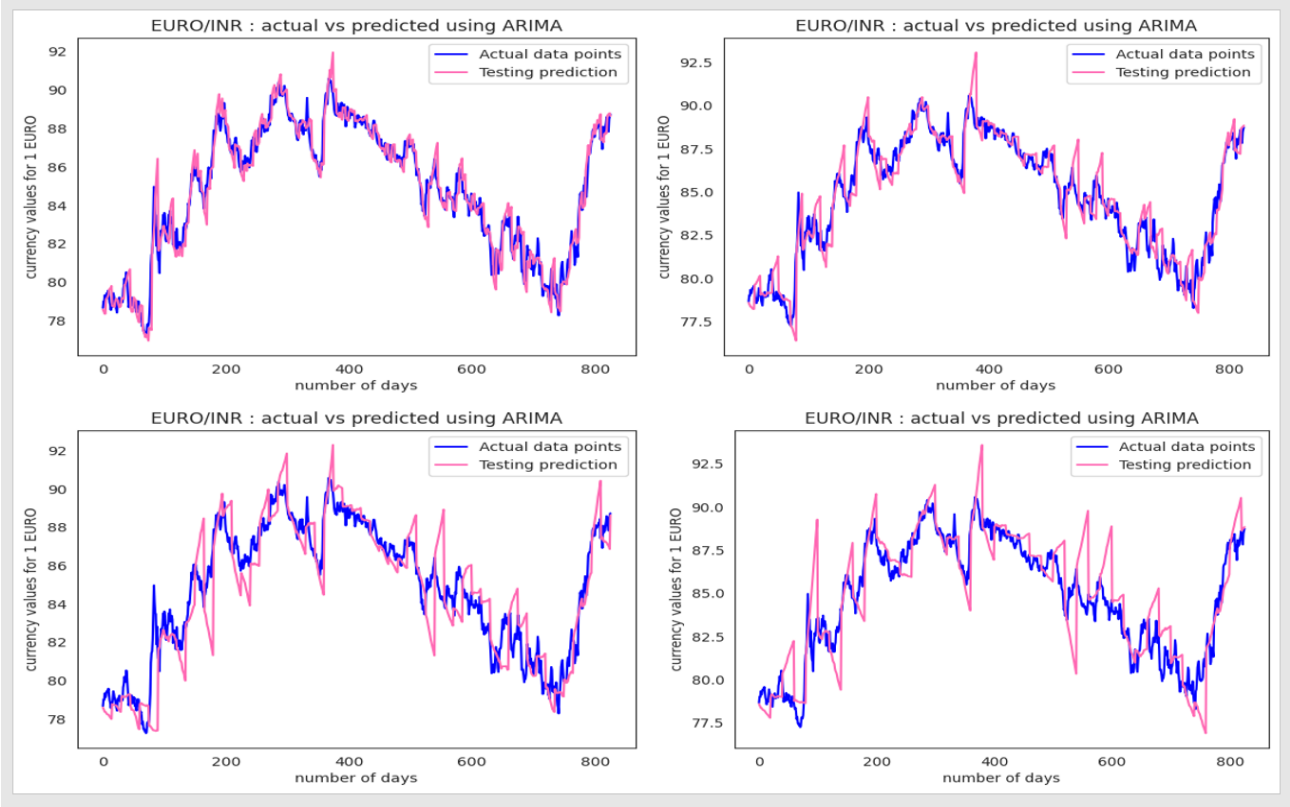
Observing the graph carefully, we see that the data of currency exchange rates between the Euro and India is not stationary. There is a slight upward trend, no seasonality and cyclicity of period 2.To make it stationary we need to set the integration factor d. For this project, we have done ADF tests and plotted ACF and PACF plots for determining the parameters of the ARIMA model.

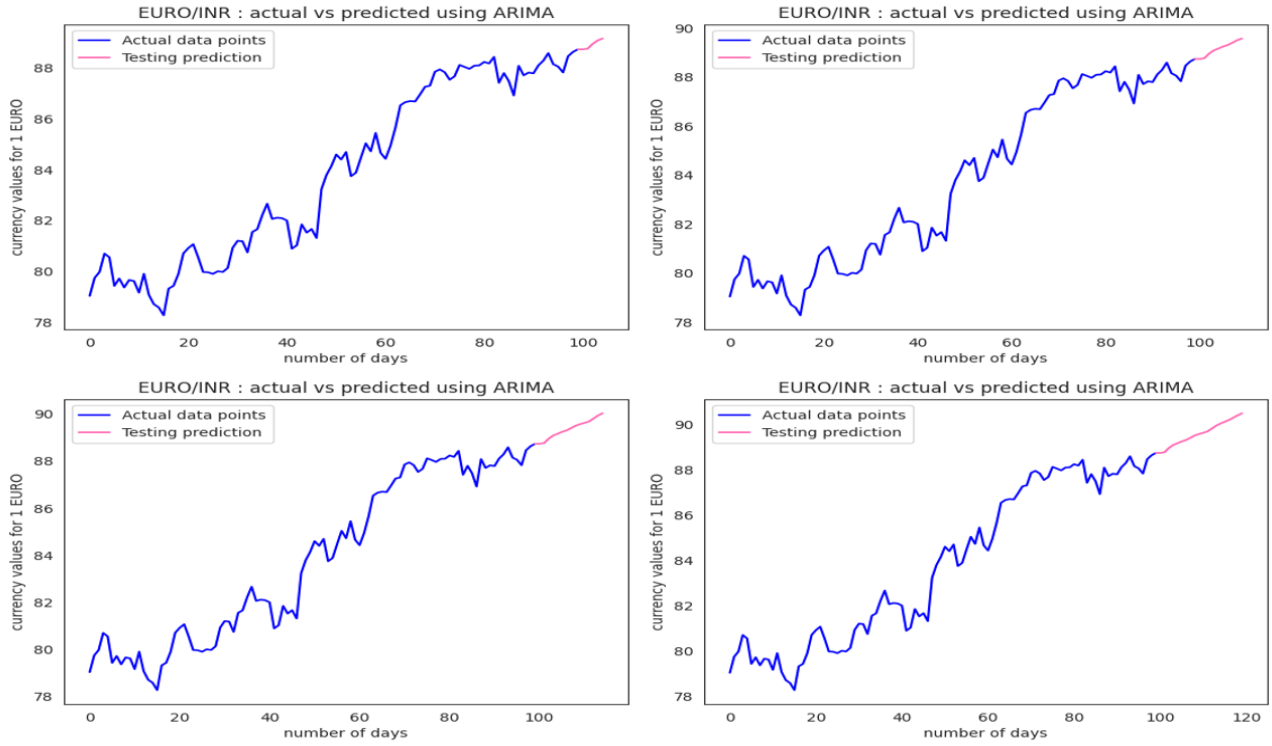
```
------------------------------------------------------------------
We are working with the series Original Series

statistical value of ADF of the precomputed tables: -2.867
statistical value of ADF: -1.8663913943871575

Level of significance to take the series as stationary: 0.05
p-value: 0.34803300847644647

------------------------------------------------------------------
We are working with the series Differenced Series

statistical value of ADF of the precomputed tables: -2.867
statistical value of ADF: -18.19639074730369

Level of significance to take the series as stationary: 0.05
p-value: 2.4120338625318466e-30

------------------------------------------------------------------
We are working with the series Twice differentiated Series

statistical value of ADF of the precomputed tables: -2.867
statistical value of ADF: -16.57263577818456

Level of significance to take the series as stationary: 0.05
p-value: 1.864356798871206e-29
```

We have seen that the p-value for the first order and second order difference is less than level of significance(0.05). Here is the ACF and PACF plots:

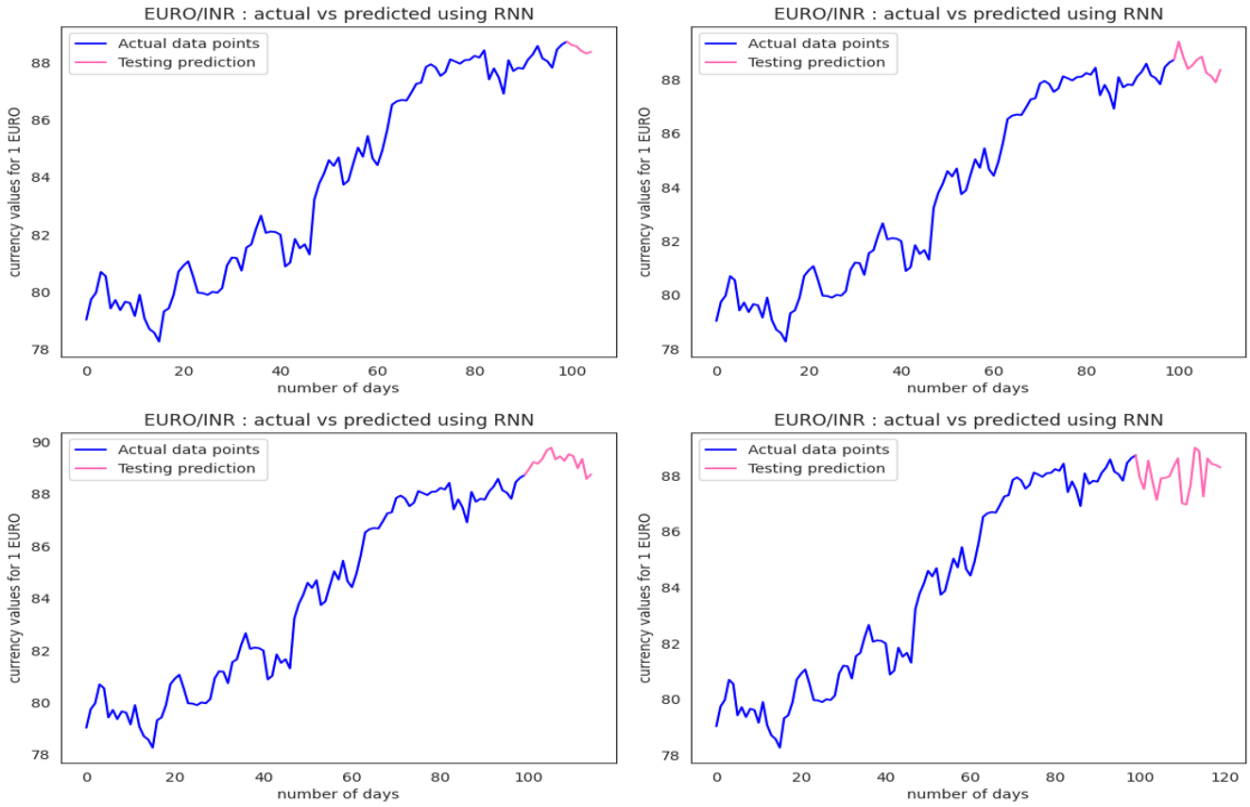


So, from the above plots, we can notice that the data gets stationary after 1st and 2nd order differencing, also we achieve significant PACF values after 2nd order differencing. Thus, we take our parameters p = 15, d = 2 and q = 0.
Here are our predictions and forecasts using ARIMA:


EURO/INR : actual vs predicted using ARIMA
Actual data points
Testing prediction
currency values for 1 EURO
number of days
EURO/INR : actual vs predicted using ARIMA
Actual data points
Testing prediction
currency values for 1 EURO
number of days
EURO/INR : actual vs predicted using ARIMA
Actual data points
Testing prediction
currency values for 1 EURO
number of days
EURO/INR : actual vs predicted using ARIMA
Actual data points
Testing prediction
currency values for 1 EURO
number of days


EURO/INR : actual vs predicted using ARIMA
Actual data points
Testing prediction
currency values for 1 EURO
number of days
EURO/INR : actual vs predicted using ARIMA
Actual data points
Testing prediction
currency values for 1 EURO
number of days
EURO/INR : actual vs predicted using ARIMA
Actual data points
Testing prediction
currency values for 1 EURO
number of days
EURO/INR : actual vs predicted using ARIMA
Actual data points
Testing prediction
currency values for 1 EURO
number of days

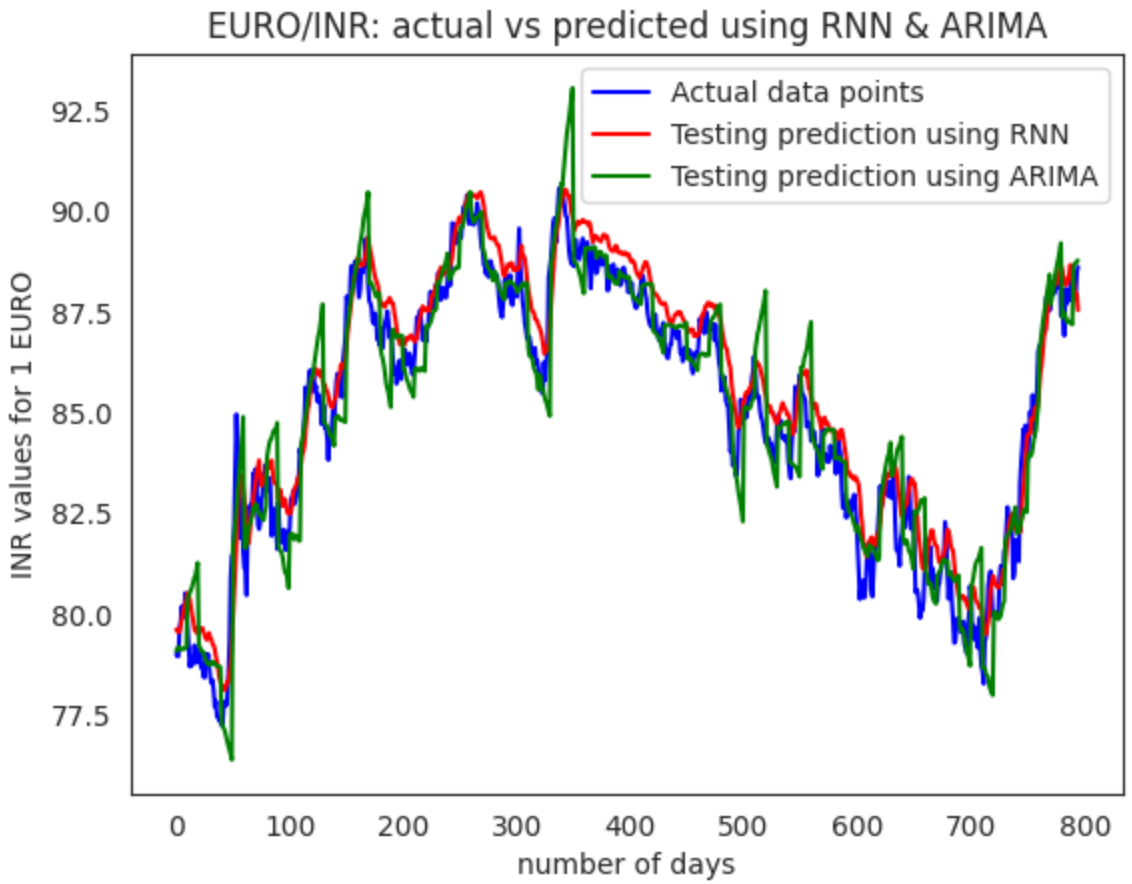
RNN model: In this project, we build an RNN with an input layer, output layer, and single hidden layer. Here we split the dataset into two sets, a training set, containing 70% of the data and a testing set, containing 30% of the data. Training set helped the model learn about the data and the testing set was used to evaluate the performance of the model. We have passed the sequence of the past 30 days and made predictions for the next 5/10/15/20 days. Also we have forecasted for the same. The graphs corresponding to the forecast are shown below.



Experimental results and analysis:
The table given below shows the mean square error obtained for predictions based on the above currency exchange rates versus ARIMA model and RNN model-

Prediction for-	5 days	10 days	15 days	20 days
ARIMA	0.5719	1.1205	2.3099	2.8723
RNN	0.7019	1.3348	2.1261	2.3959

Also we have plotted the next 10 days prediction given by ARIMA and RNN at a time in the same graph for comparing their performance.



We have also calculated R-square values for the ARIMA model to measure the goodness of the model.

Prediction for-	5 days	10 days	15 days	20 days
ARIMA	0.9499	0.9018	0.7977	0.7485

Thus, from the results we got, we can conclude that, for fewer days prediction at a time, ARIMA & RNN predict more or less good and MSE is close to each other. Though, in the long run, we can say that RNN has a better consistency than ARIMA.
Also, for the forecasting, RNN has performed better than ARIMA, though we can add a confidence interval for the forecasting given by the ARIMA model.