

Foundations of Intelligent Systems

Final Report

Currency Prediction System

Gevaria, Harnisha
hgg5350@rit.edu

Hingu, Dharmendra
dph7305@rit.edu

Gevaria, Kushal
kgg5247@rit.edu

April 17, 2017

1 Abstract:

This report focuses 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. 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.

2 Introduction:

The market of foreign exchange has been on significant growth since 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. For a common person, the foreign exchange comes into picture when the person is traveling to another country for business or vacation. The exchange rate fluctuations don't matter much as the amount would be less. But when international trades takes place the amount is very huge, and the minor difference, even decimal change, in the exchange rate can have a significant impact. Consider business deal finalized between two parties amounting to one billion dollars buyer residing in USA and seller residing in Europe. A slight variation in the exchange rate can have significant impact and seller may have to pay more than settled amount.

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. There are four ways to make the prediction about currencies: Purchasing power parity (PPP), Relative economic strength approach, Econometric models, and Time series analysis [1]. Our focus is to perform time series analysis for all major currencies against USD. 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.

As a foreign student, we almost every day check the currency value of our home country against the dollar value for fees and other types of usages. For a student, it is beneficial if the exchange rate is minimum. If we can, somehow, come up with a system, that can predict what would be the exchange rate in future, which is reliable, we don't have to make any assumption about it. Plus, there are many factors which determine the worth of currency on a global scale, which is not known to everyone. Using time series model like 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. 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).

3 Proposed method:

The proposed method involves using/building and evaluating different models to predict the currency values for the available time series data. The time series data has a special property that the data is captured at regular time intervals. We consider the daily data containing the exchange rates for the target currency against the currency of United States, i.e., USD. The time series data that we are using can be found at, <http://investexcel.net/automatically-download-historical-forex-data-into-excel/>. The historical data can be retrieved using this Microsoft Excel sheet by entering the standard currency abbreviation, start date, and end date. Using the above API, we have collected data for the past 16 years starting from 2001 for all the major 31 currencies, but we are using only last 366 days of data for our analysis to avoid possible overfitting. Another reason to avoid too much reliance on the historical data is because exchange rate of today has nothing to do with exchange rate a decade ago.

3.1 ARIMA model:

For this project, the first model that we build was ARIMA (Autoregressive Integrated Moving Average) model. ARIMA model is a statistical model which is used to forecast the dataset that has time series nature. 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 q 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 r be the size of the moving average inputs needed to calculate the regression error. The values p , q and r 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 USA (USD) and India (INR) which is as follows,

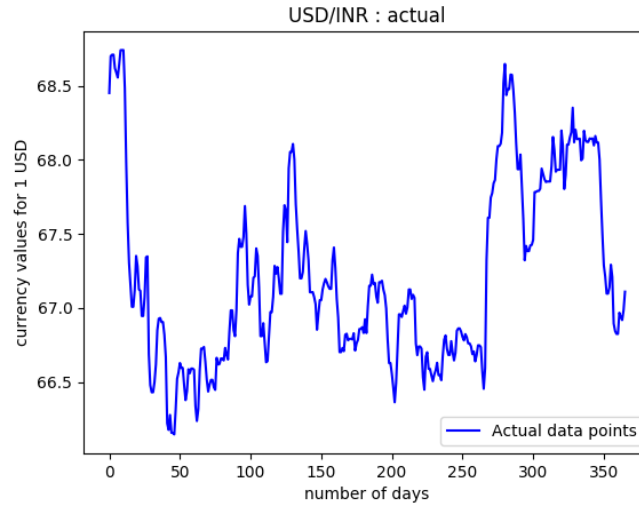


Figure 1: Plot of currency value for India against 1 USD

Observing the graph carefully, we see that the data of currency exchanges rates between USA and India is not stationary and this is true for any currency exchange rate. To make it stationary we need to set the integration factor which is parameter q . For this project, we are substituting 3 for parameter p and integration factor q as equal to 1. Also, we leave out the moving average factor by putting the value of r as 0. We used 70% of the data as training set and remaining 30% of the data as a testing set. We trained above model on the given training set using python *statsmodels* package. After training the model, we used our testing data set to find out how accurate the model is and calculated the mean square error which came out to be very low. The graph of actual currency exchange rate values and predicted values for testing set is as follows,

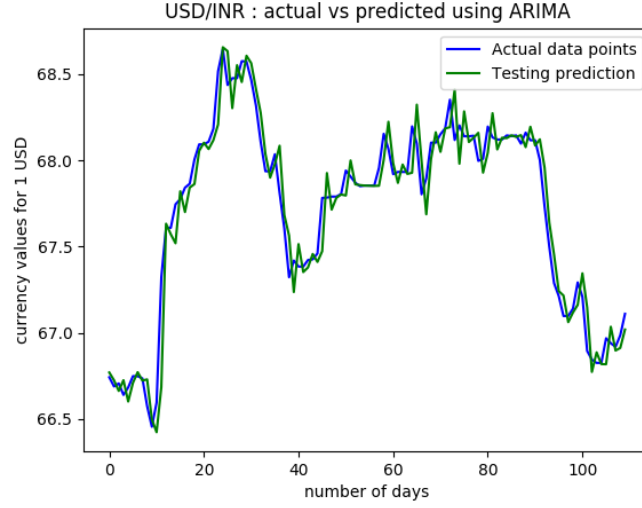


Figure 2: Plot of currency value: actual vs prediction using ARIMA

3.2 Neural network model:

The most intuitive way to use a neural network with time series data is to build a multi-layer perceptron and use the back propagation algorithm to learn the training set. The second model that we have built for trying to predict the currency, is a special type of neural network known as, recurrent neural network (RNN). RNN is a type of artificial neural network (ANN), where the inputs to the neurons at the first layer are not features unlike normal ANN, but time lag values i.e., previous values. We use python keras deep learning library to build the desired model.

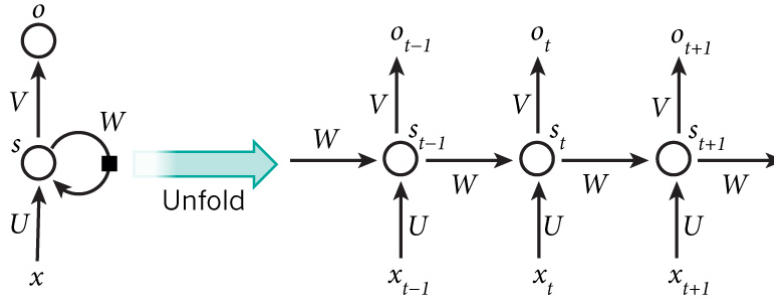


Figure 3: Recurrent neural network [2]

In our first attempt, we build an RNN with an input layer, output layer, and single hidden layer. Here we split the dataset into two sets training set, containing 70% of the data and testing set, containing 30% of the data. Training set helped model learn about the data and testing set was used to evaluate the performance of the model. Once we fit the model, we realized that relying on only the t data point to predict $t + 1$ is not an efficient method.

Later, to enhance the performance of the model we use multiple recent past steps instead of most recent step, i.e., to predict $t + 1$ we use t , $t - 1$ and $t - 2$ data points. Here again, we split the data into two sets, a training set containing 70% of data and testing set containing remaining 30% of data. We add another hidden layer, making two total hidden layers for this neural network. First hidden layer containing 12 neurons and the second hidden layer containing 8 neurons.

After training the model, we used testing data set to evaluate the performance of the model and calculated the mean square error which is very low. The graph of actual currency exchange rate values and predicted values for testing set is as follows,

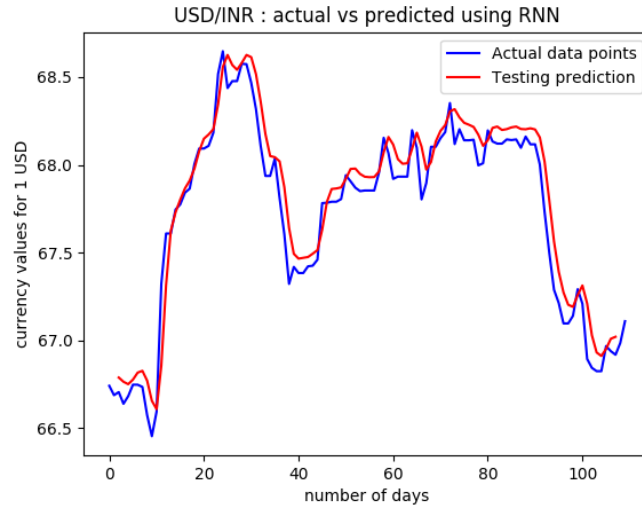
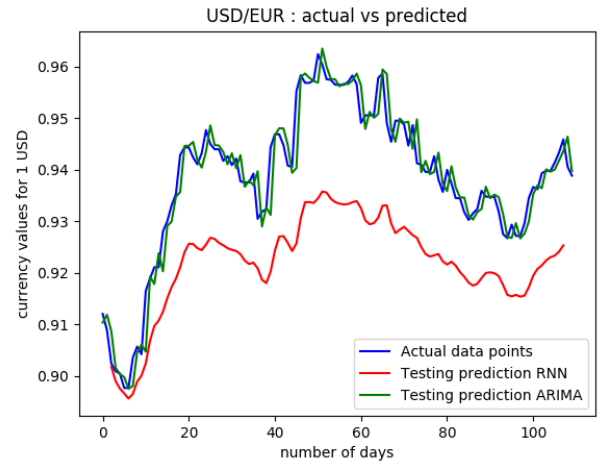
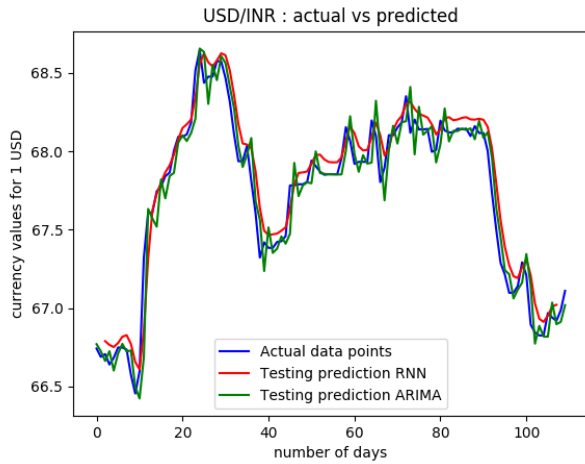
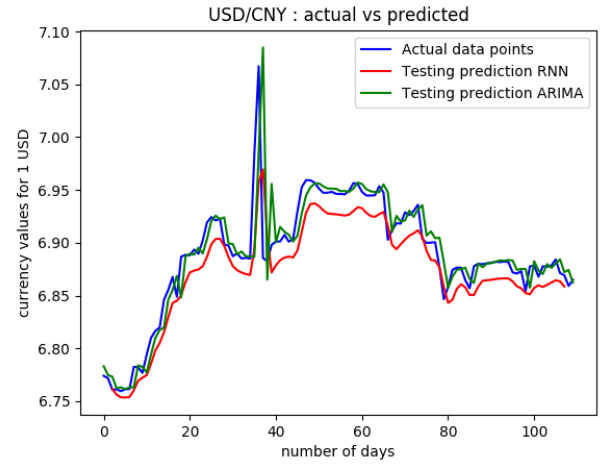
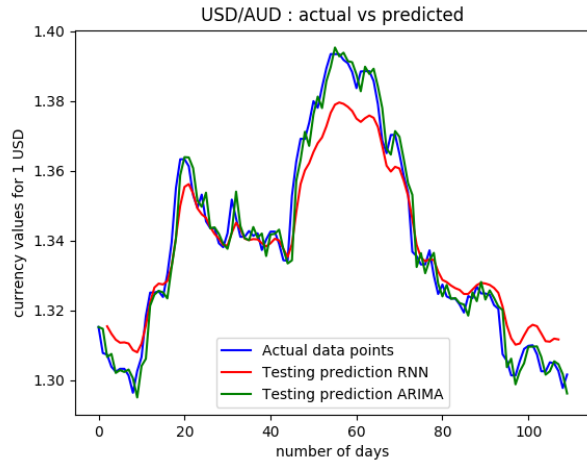


Figure 4: Plot of currency value: actual vs prediction using Neural Network

Experimental results and analysis

The two models, namely ARIMA model and Neural network model can be used to make prediction about future exchange rate for any currency. There are 35 different currencies present in our data set. For our analysis we have selected currency exchange rates of 4 countries against USD. We have selected USD (United States), being one of the standard currency and destination currencies can be any of INR (India), EUR (Europe), AUD (Australia) and CNY (China). The prediction for 2/18/2017 was carried out for exchange rates of USD against INR, CNY, AUD and EUR using the above two models for the data available from 2/17/2016 to 2/17/2017. Four graphs are plotted for the comparison between those two models on the above given currency exchange rates.





The table given below shows the mean square error obtained for predictions based on the given four currency exchange rates versus ARIMA model and neural network model,

	USD/INR	USD/CNY	USD/AUD	USD/EUR
ARIMA	0.01856	0.00067	0.00002	0.00001
RNN	0.04655	0.00102	0.00009	0.00032

Our goal is to be able to predict the exchange rate for the next day. For this we take into consideration past three days, for both the models, giving us very low MSE, as low as 0.00032 for EUR. We get this high accuracy because of two reasons. Firstly, in real world the rate fluctuations between the rates in consecutive days are not too considerable. Often the rate differences are significant over a period of time. So while predicting the future rate the model has to consider the past trend, analyses it whether the rate is of increasing trend or decreasing trend. Compute the possible change that can occur in the next day and accordingly give the prediction. This is the reason considering the past one day data doesn't work, because it becomes difficult for the model to analyze the trend between the rates and forecast appropriate value. Using the past few days data helps the model to analyze the data trend, for example consider that the rate was increasing since last 2 days and there was minor drop on 3rd day we wouldn't want to predict the drop on the 4th day just on the basis of the past one day. So here if we consider past few day's data, the model will appropriately predict the next days data by analyzing the trend. Secondly, data doesn't heavily depend on historical data as the exchange rate a decade ago has no significant impact on the exchange rate for today. Hence, we rely majorly on the past few days data to identify the trend and predict the values.

Past studies report that the neural network model being a complex model performed much better than the ARIMA model [7]. But after analyzing these two models on currency exchange rate predictions, it is clear that ARIMA model is performing better than the neural network model.

References

- [1] Joseph Nguyen. (n.d.). *4 Ways To Forecast Currency Changes*. Retrieved from <http://www.investopedia.com/articles/forex/11/4-ways-to-forecast-exchange-rates.asp>
- [2] Denny Britz. (2015, September 17). *RECURRENT NEURAL NETWORKS TUTORIAL, PART 1 INTRODUCTION TO RNNs*. Retrieved from <http://www.wildml.com/2015/09/recurrent-neural-networks-tutorial-part-1-introduction-to-rnns/>
- [3] Jason Brownlee. (2016, July 19). *Time Series Prediction With Deep Learning in Keras*. Retrieved from <http://machinelearningmastery.com/time-series-prediction-with-deep-learning-in-python-with-keras/>
- [4] Jason Brownlee. (2017, January 09). *How to Create an ARIMA Model for Time Series Forecasting with Python*. Retrieved from <http://machinelearningmastery.com/arima-for-time-series-forecasting-with-python/>
- [5] John Hunter, Darren Dale, Eric Firing, Michael Droettboom and the Matplotlib development team; 2012 - 2016 The Matplotlib development team. (2017, February 2017). *Pyplot tutorial*. Retrieved from http://matplotlib.org/users/pyplot_tutorial.html

[6] Chollet, François. (n.d.). *Keras*. Retrieved from <https://keras.io/models/sequential/>

[7] Babu AS, Reddy SK (2015) *Exchange Rate Forecasting using ARIMA, Neural Network and Fuzzy Neuron*. J Stock Forex Trad 4:155. doi:10.4172/2168-9458.1000155