# Predictive time series analysis of stock prices using neural network classifier

Abhinav Pathak,
National Institute of Technology, Karnataka,
Surathkal, India
abhi.pat93@gmail.com

**Abstract— The work pertains to developing financial forecasting systems which can be used for performing an in-depth analysis of the stocks prices, downloading/importing data from the various locations and analyzing that data and producing charts to determine statistical trends. There on it describes to perform a time series predictive analysis of the stocks data that we have and plot the various opening and closing prices of the stocks and then convert it to time series data so that we can proceed and perform a time series predictive analysis thereby predicting the h-days closing prices of a certain stock using the neural networks classification algorithm. The implementation is done using the open source software R & WEKA thereby aiming to reduce the analytics cost for any organization.**

**Keywords—**Stock market predictions, neural networks, data mining classification algorithms.

## I. Introduction

Stock markets have always been an area of massive interest for people especially research scholars and engineers. Stock markets and prices are undoubtedly a huge aggregation of buyers and sellers. It covers participation from the whole world and decides entirely on the success or failure of any venture in the most factual way. Therefore forecasting stock prices or financial markets has always been a great challenge to the experts of computer science, artificial intelligence, economics, mathematics, etc. Various technical and statistical methods/ models have been developed and used with varying outcome results but none of them have shown great success and which makes this area a subject of research.

Data mining is a great asset to research in this domain and provides a huge range of solutions to deal with such problems of predictive analysis. It is able to uncover hidden patterns and predict future trends and behaviors in financial markets. It creates opportunities for companies to make proactive and knowledge-driven decisions in order to gain a competitive advantage. Data mining has been applied to a number of financial applications, including development of trading models, investment selection, loan assessment, portfolio optimization, fraud detection, bankruptcy prediction, real-estate assessment, and so on. Recently, various open source statistical software have also been developed which simplify the process of such an analysis to a great extent.

The purpose of an analysis like this is developing a financial forecasting system which can incorporate the presently happening changes as well. Moreover it also gives a system to perform an in-depth of the stocks downloading/importing data from the various locations and analyzing that data and producing charts to determine statistical trends. Financial analytics is primarily used by firms to assess the elements like profitability, solvency, liquidity, stability, etc. It also helps in predictive technology, where we can make the predictions about the future statistics of any product on the basis of the past and present statistics. It is also used greatly in stocks and trading companies.

Various researches have shown that the neural networks have a great capability in machine learning, pattern recognition and predictive analysis problems such as classification and regression, etc. The primitive methods of predictive analysis included models based on linear and logistic regression whose efficiency depends on the type of model used for the regression. Later came the neural network classifiers which are less of model dependent and more of training dependent and hence show better and self-adjusting results in various analyses and are greatly useful for analyses in which the predictive model is not fundamentally clear in its essence. Neural networks are well applied to the problems in which extracting the relationships among data is really difficult. Neural networks have been claimed to be general function approximates. It is proved that an neural network can approximate any complex continuous function that enables us to learn any complicated relationship between the input and the output of the system.

## II. Related Work

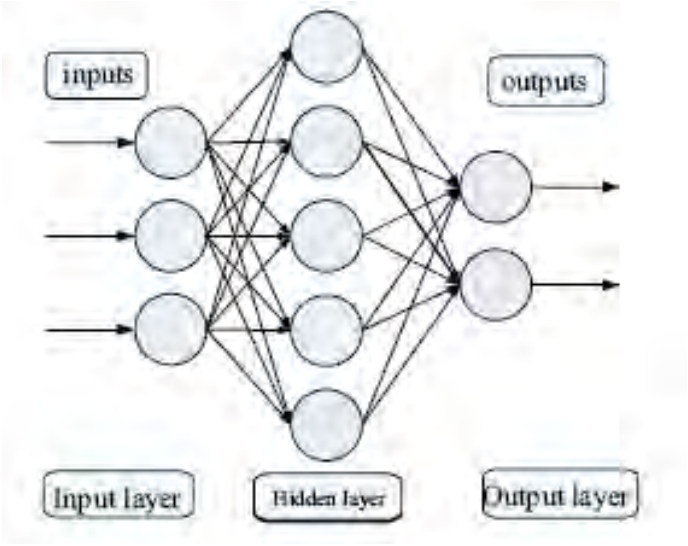
Since stock market prediction is an area of huge interest, there has been a constant research to develop more and more efficient ways of such predictive analyses. The most prominent technique apart from the neural networks involves the use of Genetic Algorithms, support vector machines, etc. All these methods and techniques have shown varying efficiency of the results produced greatly dependent on the type of input provided and which makes the comparison of these techniques to judge the more efficient one a hard and unfeasible task as of now. Therefore this becomes a subject of ongoing research to develop more and more efficient ways of such predictive analyses with great efficiencies.

## III. Neural Networks

A neural network is formed by a network of computing units (the neurons) linked to each other. Each of these connections has an associated weight. Constructing a neural network consists of using an algorithm to find the weights of the connections between the neurons. A neural network has its neurons organized in layers. The first layer contains the input neurons of the network. The cases of the problem we are addressing are presented to the network through these input neurons. The final layer contains the predictions of the neural network for the case presented at its input neurons. In between we have one or more "hidden" layers of neurons. The weight updating algorithms, like for instance the back-propagation method, try to obtain the connection weights that optimize a certain error criterion that is the weights which ensure that the network output is in accordance to the cases presented to the model.

The use of neural networks simulates how human brain functions, by feeding computers with massive data to mimic human thinking. The most common form of neural networks in use for stock market prediction is the feed forward network utilizing the backward propagation of errors algorithm to update the network weights. These networks are commonly referred to as Back propagation networks. Another form of neural networks that is more appropriate for stock prediction is the time recurrent neural network (TRN) or time delay neural network (TDNN). Examples of TRN and TDNN are the Elman, Jordan, and Elman-Jordan networks.

For stock prediction with neural networks, there are usually two approaches taken for forecasting different time horizons: independent and joint. The independent approach employs a single ANN for each time horizon, for example, 1-day, 2-day, or 5-day. The advantage of this approach is that network forecasting error for one horizon won't impact the error for another horizon—since each time horizon is typically a unique problem. The joint approach, however, incorporates multiple time horizons together so that they are determined simultaneously. In this approach, forecasting error for one time horizon may share its error with that of another horizon, which can decrease performance. There are also more parameters required for a joint model, which increases the risk of over fitting.




## IV. Implementation

The implementation of such a system can be easily done using the open source software keeping in mind cost effective analytics for any organization. For example, it can be implemented easily using the various packages available in R. R directly downloads the stocks data from the online resources and plots it accordingly as asked by the user.

There on we perform a time series predictive analysis of the stocks data that we have. We use the online data from yahoo finance to plot the various opening and closing prices of the stocks and then convert

it to time series data so that we can proceed and perform a time series predictive analysis thereby predicting the h-days closing prices of a certain stock.

The type data we have for this case study is usually known as a time series. The main distinguishing feature of this kind of data is the existence of a time tag attached to each observation, meaning that order between cases matters. In the case of our stocks data we have what is usually known as a multivariate time series, because we have several variables being recorded at the same time tags, namely the Open, High, Low, Close and Volume. The usual approach in financial time series analysis is to focus on predicting the closing prices of a stock.

We create an R function to obtain the h-days returns of a vector of values, here the closing prices of a stock. To create this function we have used the function diff (). This R function calculates lagged differences of a vector. We

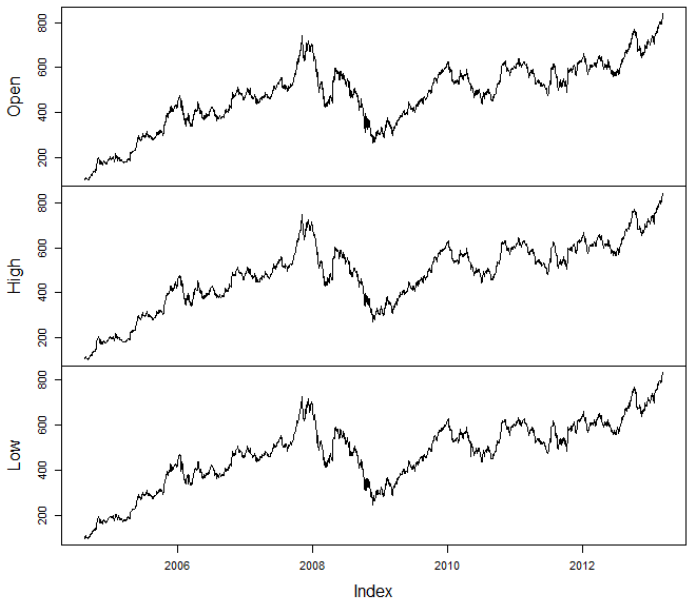
will generate a data set using this function, which will then be used to obtain a model to predict the future h-days returns of the closing price of IBM stocks.

The model used with the goal of predicting the 6-8-day ahead returns of closing prices will be a neural network. Neural networks are among the most frequently used models in financial predictions experiments because of their ability to deal with highly non-linear problems. The package “nnet” implements feed forward neural nets in R. These types of neural networks are among the most frequently used ones.

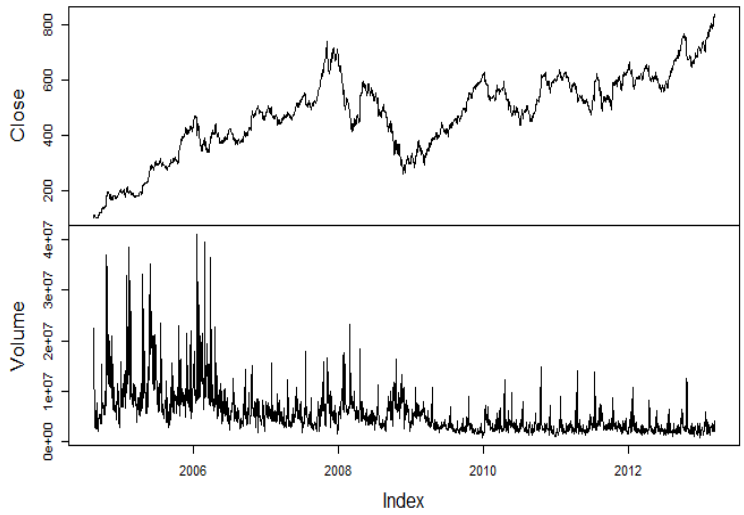
## V. Simulation Results

This section presents the whole simulation process of stocks prediction using the neural networks classifier. Initially the classifier is trained using the already existing stock prices available online which can be directly imported for use from online libraries like Yahoo Finance, Google Finance, etc.

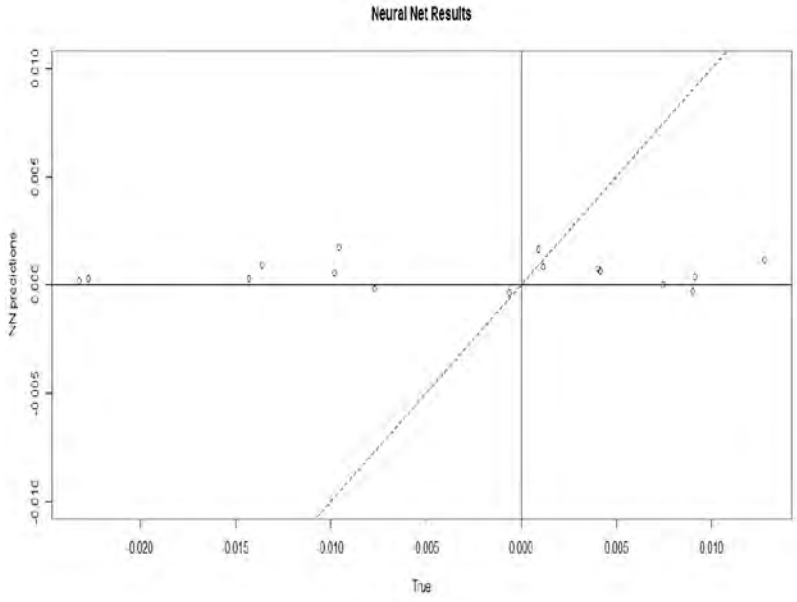
The stock index variables Open, High, Low, Close and Volume are obtained from those online libraries and they provide as the input data or the training data for the neural networks classifier.




These are the various plots of the varying prices of the stocks which is then converted to time series data so that we can apply classification algorithms to it and produce useful predictive analysis out of it.




These inputs after being converted to time series data are fed as input or training data for the classification algorithms and after sufficient training of the classifier, it is used to predict the closing prices future and which is plotted against the actual closing prices to evaluate the efficiency and accuracy of the classifier.




Due to the time dependence between observations the evaluation procedures for time series prediction models are different from standard methods. The latter are usually based on resampling strategies (for instance bootstrap or cross validation), which work by obtaining random samples from the original unordered data. The use of these methodologies with time series could lead to undesirable situations like using future observations of the variable for training purposes, and evaluating models with past data. In order to avoid these problems we usually split the available time series data into time windows, obtaining the models with past data and testing it on subsequent time slices.

## VI. Conclusion & Further Scope

This is a beginning of the use of neural networks classifier for predicting the time series data or the stock prices. It can be modified and optimized in a lot of ways to produce much better and much more efficient and accurate results.

The common approaches which can act as supplementary to these include using multiple regression techniques like the projection pursuit regression or the Multivariate adaptive regression splines (MARS) which give slightly better outputs and hence are slightly more efficient. In research perspectives, the MARS model has shown the best accuracies but there is still a great scope to further enhance these algorithms for even more efficient outcomes.

## REFERENCES


[1] Data Mining with R: learning by case studies, Luis Torgo, LIACC-FEP, University of Porto, Portugal
[2] Stock Price Prediction and Trend Prediction using Neural Networks, Mr. Pritam R. Charkha, Assistant Professor, Department of Computer Science & Engineering, Marathwada Institute of Technology, Aurangabad.
[3] Stock Market Value Prediction Using Neural Network, Mahdi Pakdaman Naeini, Hamidreza Taremian, Islamic Azad University, Parand Branch, Tehran, Iran and Homa Baradaran Hashemi, School of Electrical and Computer Engineering, University of Tehran, Tehran, Iran
[4] Stock Prediction – A Neural Network Approach, Karl Nygren, karlnyg@kth.se
[5] Textual Analysis of Stock Market Prediction Using Breaking Financial News: The AZFinText System, Robert P. Schumaker and Hsinchun Chen, Artificial Intelligence Lab, Department of Management Information Systems, The University of Arizona, Tucson, Arizona 85721, USA
[6] Advanced Topics in Time Series, Econometrics Using R, ZONGWU CAI
[7] Credit rating analysis with support vector machines and neural networks: a market comparative study, Zan Huanga,*, Hsinchun Chena, Chia-Jung Hsua, Wun-Hwa Chenb, Soushan Wuc, a Department of Management Information Systems, Eller College of Business and Public Administration, The University of Arizona, Rm. 430, McClelland Hall, 1130 E. Helen Street, Tucson, AZ 85721, USA, b Department of Business Administration, National Taiwan University, Taiwan, c College of Management, Chang-Gung University, Taiwan
[8] Financial time series forecasting using support vector machines
[9] Kyoung-jae Kim∗ Department of Information Systems, College of Business Administration, Dongguk University, 3-26, Pil-dong, Chung-gu, Seoul 100715, South Korea
[10] Stock Market Forecasting Using Machine Learning Algorithms, Shunrong Shen, Haomiao Jiang, Department of Electrical Engineering, Stanford University, Tongda Zhang, Department of Electrical Engineering, Stanford University.
[11] http://en.wikipedia.org/wiki/Artificial_neural_network
[12] http://en.wikipedia.org/wiki/Stock_market_prediction
[13] http://en.wikibooks.org/wiki/Data_Mining_Algorithms_In_R/