

Stock Prediction using ARMA

Harsha Saxena¹, Ainapurapu Venkata Anurag², Nihal Chirayath³, Rahul Bendale⁴ and Sidharth Kaul⁵

¹Assistant Professor, Department of Computer Engineering, Ramrao Adik Institute of Technology, Navi Mumbai, INDIA

^{2,3,4,5}Student, Department of Computer Engineering, Ramrao Adik Institute of Technology, Navi Mumbai, INDIA

²Corresponding Author: avanurag96@gmail.com

ABSTRACT

A stock market is an institution where humans and computers buy and sell shares of companies. For many people, that is the first thing that comes to mind for investing. The goal is to buy the stock, hold it for a time, and then sell the stock for more than you paid for it. In the stock market, prices rise and fall every day. When you invest in the stock market, you are hoping that over the years, the stock will become much more valuable than the price you paid for it. But as data collected over the years shows that an individual's stock has 90% more chance to increase in value at some point over the period of investment. The project basically aims at encouraging the people about stock market investments. It works on an algorithm which predicts the most profitable stocks to invest in different companies thereby making it easier for the investors to invest wisely. It provides reliable information regarding the percent profit earned by any company and its expected gains according to studied and analysed trends. The project works on machine learning and data science.

Keywords— Stock Market, Prediction, Artificial Neural Networks, ARMA, Sentiment Analyser

I. INTRODUCTION

The stock markets are very interesting and can fetch you good returns when invested smartly.

The technical terms used in stocks are bull-profit, chicken- no profit, no loss and bear-loss. Recent data survey by Gartner Rich and Co. has projected that investment made over long periods with an appropriate amount fetch good results and returns.

Currently, stocks prices depend on factors like Company sentiment, economy and key events of the company. People depend on stock brokers and temporary trends while investing which may not produce expected returns.

The technology used in this project is Artificial Neural Networks (Multi-Layer Perceptron Model), Auto regressive Moving Average for predicting stock prices. Both models predict the prices separately and an average value is calculated. For other external factors, overall

company sentiment is analyzed by using NLP- Natural Language Processor. As an output the user is presented with Stock Price Prediction and company sentiment in terms of Positive, Neutral or Negative. Hence, this technology is much more efficient and accurate as past stock data set and trends are analyzed along with current data set and trends. The main objective of this project is to enlighten users and present them with some smart investment options to make more better and informed decisions, especially the current generation who represent the new face of any country. Many people who are not educated about Share markets have a negative review of it and tend to not invest money in it and rather prefer to put it in banks. That notion is not true once anybody is educated about stocks.

Infact benefits to its users are as follows:

1. Returns on Stocks are not taxed as long as they are long term investments (>12 months).
2. Investing in stocks means owning a part of a listed company, major stock holders have to ability to make critical decisions in an invested company.
3. Help to the country's economy and GDP growth
4. More returns than a conventional FD at banks and amounts can be liquidated any time investor desires.
5. Helps any investor with even small amounts of money like Rs 10000 to invest.
6. Process of investing in stocks is easy and doesn't involve much time and effort, further managing can be done using online demat accounts or a stock managing agent.

II. LITERATURE REVIEW

For any project planning activity, a good reading of existing projects, ideas and technology is needed. The following subsections provide much needed excerpts from important research papers and literary documents related to Stocks, Artificial Neural Network, Auto Regressive Moving Average and Natural Language Processing. Factors of change for Stock Markets:

1. Stock Prices: Stock Prices or share prices change every day depending on the market, economy and company's financial performance and outlook.

2. Sentiment: The news headlines, important statements made by key people of a company, announcements all contribute to stock values in a positive, negative or neutral way.

Market and Environmental Factors: Market and Environmental factors like Economic boom, depression, political scenario, natural disasters etc. contribute to the performance of stock market in general. Before deciding on the project, lot of research papers of the domain and other articles related to stocks, share markets and machine learning. Following inferences were made from these papers.

Formula for calculation of Stock Prices: The methodology used in this study considered the short-term historical stock prices as well as the day of week as inputs. The overall procedure is governed by the following equation:

$$y(k) = f(y(k-1), y(k-2), y(k-3), \dots, y(k-n), D(k))$$

where $y(k)$ is the stock price at time k , n is the number of historical days, and $D(k)$ is the day of week.

Selection of Algorithm: -Three algorithms were tested with input datasets of Google Inc. for the Nov to Dec 15. Tested models were Random Forest, Logical Regression and Multi-Layer Perceptron.

1. Random forest[7]: - Random forest is a faster calculation method but is the least accurate when compared with the other two algorithms. Random forest explores all the possible nodes of a tree and selects the maxima. Accuracy:50-60%. Good for short term predictions and gains.

2. Logical Regression[7]: -The logical regression method makes a graph of the dataset present. It converts the graph to an equation form and substitutes values. Accuracy: 60-64%. Good for midrange predictions.

3. Artificial Neural Network using Multi-Layer Perceptron Classifier[8]:-The Neural network (Multilayer Perceptron Model) works in layers and data is processed in each layer. The final layer produces the output. This is the best algorithm for long term predictions ranging from 300 days and above. Accuracy: 75% and above.

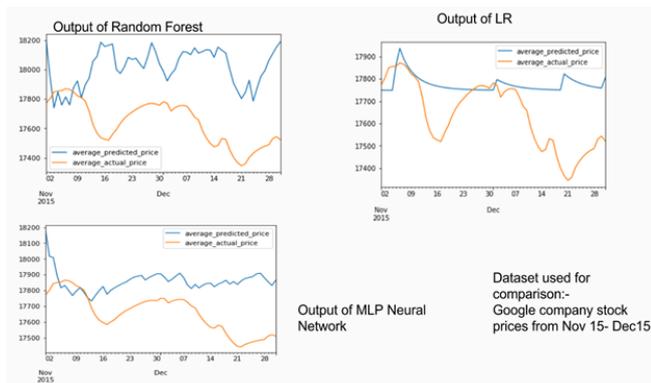


Fig-1 Different Prediction Algorithms and their accuracy

III. SIMILAR EXISTING PROJECT COMPARISON

We have drawn inspiration from the following applications:

1. ICICI Bank Direct mobile: The app allows the user to trade, view latest quotes, track personal portfolio and get live research calls. The user needs to be a registered demat a/c holder with ICICI to trade stocks.

2. Stock Trainer: A gamified app with real companies and virtual currency. The app provides a short 1 min tutorial at the beginning and gives the user 1lakh rupees of game currency. No suggestions are given and user is free to experiment his/her wishes

IV. EXISTING SYSTEM

We intend to use the following technologies in conjunction with each other so as to obtain proper

Use of API:- The yahoo stocks is an open source stock prices application program interface. It can be embedded in the backend of projects using its open source code to receive closing prices of stocks.

Use of MLP Classifier for Neural network:- A multilayer perceptron (MLP) is a class of feedforward artificial neural network. An MLP consists of at least three layers of nodes. Except for the input nodes, each node is a neuron that uses a nonlinear activation function. MLP utilizes a supervised learning technique called backpropagation for training. Its multiple layers and nonlinear activation distinguish MLP from a linear perceptron. It can distinguish data that is not linearly separable.

Auto Regressive Moving Average(ARMA): ARMA model is a tool for understanding and, perhaps, predicting future values in this series. The model consists of two parts, an autoregressive (AR) part and a moving average part. The Auto Regressive part involves regressing the variable on its own lagging(past) values. The Moving Average part involves modelling the error term as a linear combination of error terms occurring at the same time period and at various times in the past. ARMA is appropriate when a system is a function of a series of unobserved shocks (the MA or moving average part) as well as its own behaviour like the stock market as stock prices may be shocked by fundamental information as well as exhibiting technical trending and mean-reversion effects due to market participants. The notation $AR(p)$ refers to autoregressive model of order p . Hence, the model is an ideal benchmarking for predicting stocks using neural network.

It is written as the equation below.

$$X_t = c + \sum_{i=1}^p \phi_i X_{t-i} + \epsilon_t$$

Sentiment Intensity Analyser (Natural Language Processing module):- This is a tool present under the NLP Libraries of various development tools in Machine Learning like Ipython or Jupyter notebooks. What the analyser does is takes sentences or strings of meaningful English language words as input and produces the output

as negative, positive or neutral sentiment. This is effectively used in our project to determine the sentiment of a listed company.

V. PROPOSED SYSTEM

As studied in Literature review we can say that stock market prediction is very difficult and most existing systems only tend to use a single model to predict stock prices, the intention is to use Autoregressive Moving Average and Neural Network using Multi-Layer Perceptron together to obtain a more accurate prediction. Along with that for investor to obtain a better understanding of companies' impression on peoples mind the system also should have a sentiment analyser.

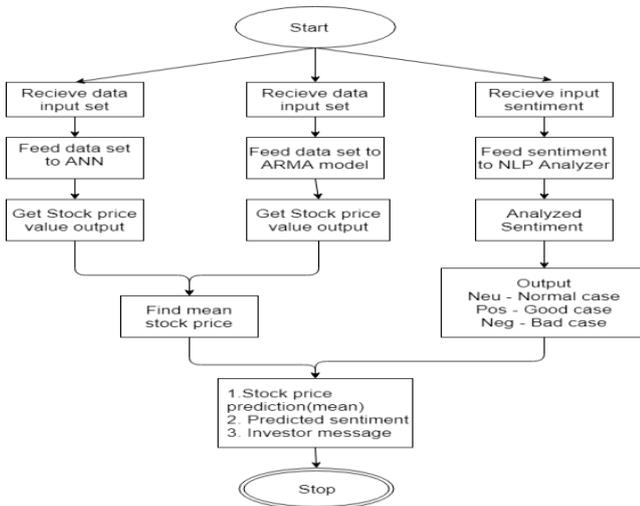


Fig-2: Data Flow diagram of system

As shown in Fig-2 there will be two Parallel process running in the system one will be for stock price prediction, while the other shows the sentiment of people for the company. The stock price prediction will be implemented using two modules that is Autoregressive Moving Average and Neural Network using Multi-Layer Perceptron, which will run independently and the final prediction shown will be a median of the both modules. Data set of previous stock prices of the company will serve as input for MLP and ARMA models. The prediction will be accompanied by a score dignifying the sentiment of company in public eye. The sentiment analysis will be performed by picking up tweets from twitter using the company hashtags and then the sentiment of sentence will add to the score of the company.

VI. EXPECTED RESULTS AND ANALYSIS

The application will collect input data from the user. Input will be amount to be invested, time period and preferences like specific rate of growth and if looking for any specific companies to invest in. The Natural language processor already has a precomputed sentiment of the

various companies included in the application. Upon receiving user input of investment and time period and rate of growth expected (complementary input condition) the MLP classifier computes the stock prices of listed companies over the future time period and user receives output from single to multiple companies to invest in(along with their predicted prices) that matches the user's criteria, ie only companies whose current stock prices match with user given investment input but have high predicted rate of returns over a long period of investment will be displayed to the user along with their predicted stock prices and sentiment.. User can still narrow down to specific companies by firing queries or sorting through her/his updated preferences. NLP specifies whether the predicted price is of good, neutral or worst-case scenario depending on that specific company's current sentiment.

VII. CONCLUSION

Being profitable on the stock market is a challenging task. The future price forecasting based on the historical data could be considered as a method mainly falling into the technical analysis domain, where the evaluation of the securities is derived from the statistics generated by market activities. MLP classifier neural network algorithm would be used. The proposed model aims at predicting the future stock prices and the company's outlook based on recent sentiment analyzed. Our initial study has shown the various technical indicators used by investors and study of financial crisis is done. Moreover to fulfil the concern of people investing wisely in stocks we have designed the system which keeps all the user (people) based constraints in mind.

ACKNOWLEDGMENT

We would like to recognize the sincere efforts of our professors especially Dr. Leena Ragha (H.O.D-Computer Engineering), Mrs. Aditi Chhabria(Project Coordinator), Mrs. Harsha Saxena for their step by step guidance , as well as Dr. Ramesh Vasappanavara (Principal) for their suggestions, time to time help which helped us achieving goals and overcoming difficulties.

REFERENCES

- [1] Mahezabin Shaikh & Mrs. G. J. Chhajer. (2015). Financial forecasting using neural network. *International Journal of Advanced Research in Computer Science and Software Engineering*, 5(1), 66-70.
- [2] C. Lee & C. Ho. (2011). Short-term load forecasting using lifting scheme and ARIMA model. *Expert System with Applications*, 38(5), 5902-5911.
- [3] M. Khashei, M. Bijari & G. A. R. Ardali. (2012). Hybridization of autoregressive integrated moving average (ARIMA) with probabilistic neural networks. *Computers and Industrial Engineering*, 63(1), 37-45.
- [4] N. Merh, V.P. Saxena, & K.R. Pardasani. (2010). A comparison between hybrid approaches of ANN and ARIMA for Indian stock trend forecasting. *Journal of*

Business Intelligence, 3(2), 23-43.

[5] S. Fallahi, et al. (2011). Applying GMDH-type neural network and genetic algorithm for stock price prediction of Iranian cement sector. *Applications and Applied Mathematics: An International Journal*, 6(2), 572 – 591.

[6] Contreras, J., Espinola, R., Nogales, F. J., & Conejo, A. J. (2003). ARIMA models to predict next-day electricity

prices. *Power Systems, IEEE Transactions on Power Systems*, 18(3), 1014-1020.

[7] J.J. Wang, J.Z. Wang, Z.G. Zhang & S.P. Guo. (2012). Stock index forecasting based on a hybrid model. *Omega* 40(6), 758-766.