ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering
(An ISO 3297: 2007 Certified Organization)
Website: www.ijircce.com
Vol. 5, Issue 2, February 2017

Stock Trend Prediction Using Regression Analysis – A Data Mining Approach

Bhagyashree Nigade[1], Aishwarya Pawar[1], Varsha Bavane[1], Siddhant Navaratna[1], Prof. Sampada Kulkarni[2]

Student, Department of Computer Engineering, D.Y. Patil College of Engineering, Ambi, Savitribai Phule Pune University, Pune, India[1]

Professor, Department of Computer Engineering, D.Y. Patil College of Engineering, Ambi, Savitribai Phule Pune University, Pune, India[2]

ABSTRACT: Stock market prediction is the method of determining future values of a company's stocks and other financial values. Stock market prediction with the help of regression analysis is the most efficient combination to predict the stocks and the conditions of the market. Market lacks a efficient software where the proper suggestions of available stocks and the proper investment analysis are presented in an efficient way. The investors should be guided and encouraged to invest in the stocks in a coherent way. The development of a vibrant application for analyzing and predicting stock market prices is a basic tool aimed at increasing the rate of investors interest in stock markets. This paper explains the development and implementation of a stock price prediction application using machine learning algorithm and object oriented approach of software system development. The algorithm was used in training a set of market data collected for the period of one thousand, two hundred and three days.

KEYWORDS: Technical Key Words, Prediction methods, Stock markets, Mean Square Error, Regression Analysis.

I. INTRODUCTION

Trading in shares is big business in many economies. Based on the information on their websites, Stockbrokers do not seem to have any intelligent tool that can help them advise clients on which stocks are suitable for any buy or sale trade. These websites provide information that points to use of fundamental, technical and time series analysis methods. These prevalent methods show a trend on future movement and not the likely trade price for any stock in future. It is therefore desirable to have a tool that does not just point a direction of price movement, but also indicates the most likely price value of the stock itself. A regression analysis algorithm that is well tuned with the appropriate parameters can be used to develop such a predictive tool.Based on the historic data this algorithm predicts the future stocks and helps to study them. Hence the trouble off going through many rigidious courses and suggestions, user is able to study and understand the stocks and also obtain predictions for the same.

Motivation:
The realm of stock market is constantly thriving under the process of modifications and alterations. Considering the fluctuations it brings every day, making profit from it requires intensive planning. Since the inception of stock market, analysts have always struggled to predict the future prices of stocks because of its highly intricate complexity and profitability. The most reliable way to forecast the future is to try to understand the present but the amount of data available nowadays is huge and generally beyond human comprehension. Data analysis comes handy to solve this problem. Data analysis can be used to better understand the present scenario of the Stock market so as to understand and try to create a better future scope for investment.With Data analysis, we can add a degree of certainty to the unpredictable and volatile nature of stock prices. This certainty can go a long way to ensure that losses are minimized

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering
(An ISO 3297: 2007 Certified Organization)
Website: www.ijircce.com
Vol. 5, Issue 2, February 2017

and profits are maximized. Though the predictions can never be fully accurate even a minute increase in accuracy of prediction can help a lot in terms of profitability.

Regression Analysis

In statistical modeling, regression analysis is a statistical process for estimating the relationships among variables. It includes many techniques for modeling and analyzing several variables, when the focus is on the relationship between dependent variable and one or more independent variables (or 'predictors'). More Specifically, regression analysis helps one understand how the typical value of the dependent variable (or 'criterion variable') changes when any one of the independent variables is varied, while the other independent variables are held fixed. Most commonly, regression analysis estimates the conditional expectation of the dependent variable given the independent variables that is, average value of the dependent variable when the independent variables are fixed. Less commonly, the focus is on other location parameter of the conditional distribution of the dependent variable given the independent variables. In all cases, the estimation target is a function of the independent variables called the regression function. In regression analysis, it isalso of interest to characterize the variation of the dependent variable around the regression function which can be described by a probability distribution. A related but distinct approach is necessary condition analysis (NCA), which estimates the maximum (rather than average) value of the dependent variable for a given value of the independent variable (ceiling line rather than central line) in order to identify what value of the independent variable is necessary but not sufficient for a given value of the dependent variable.

II. LITERATURE SURVEY

1)STOCK PRICE PREDICTION USING REGRESSION ANALYSIS. AUTHORS:Dr. P. K. Sahoo, Mr. Krishna Charlapally

This paper we investigate to predict the stock prices using auto regressive model. The auto regression model is used because of its simplicity and wide acceptability. We have also conducted a study on the efiectiveness of auto regressive model. The Moore and Penrose technique is used to estimate the coefficients of the regression equation. We have also studied accuracy of the prediction by comparing the predicted values with the actual values over a period of time.

2) A LINEAR REGRESSION APPROACH TO PREDICTION OF STOCK MARKET TRADING VOLUME:A CASE STUDY. AUTHORS: Farhad Soleimanian Gharehchopogh, Tahmineh Haddadi Bonab and Seyyed Reza Khaze

In this paper, by applying linear regression for predicting behaviour of SP 500 index, we prove that our proposed method has a similar and good performance in comparison to real volumes and the stockholders can invest confidentially based on that.

3. MULTIPLEREGRESSION: A DATA MINING APPROACH FOR PREDICTING THE STOCK MARKET TRENDS BASED ON OPEN, CLOSE AND HIGH PRICE OF THE MONTH. AUTHORS: SACHINKAMLEY , SHAILESH JALOREE R. S. THAKUR

In this paper we applied well known efficient multiple regression approach to predict the stock market price from stock market data based on three variables. In future the results of multiple regression approach could be improved using more number of variables.

4. STOCK INDEX PREDICTION USING REGRESSION AND NEURAL NETWORK MODELS UNDER NON NORMAL CONDITIONS. AUTHORS: K. V. Sujatha , S. Meenakshi Sundaram

In this paper a survey is made to compare the predictive performances of the nonparametric models of closing prices of Stock Index data, where the data is non normal. Comparative studies with the existing statistical prediction models indicate that the proposed neural network model is very promising and can be implemented into real time trading system for stock price prediction.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering
(An ISO 3297: 2007 Certified Organization)
Website: www.ijircce.com
Vol. 5, Issue 2, February 2017

5. REGRESSION TECHNIQUES FOR THE PREDICTION OF STOCK PRICE TREND AUTHORS: Han Lock Siew ,Md Jan Nordin .
This study showed that the outcomes of regression techniques can be improved for the prediction of stock price trend by using a dataset in standardized ordinal data format.

6. TEXT MINING APPROACHES FOR STOCK MARKET PREDICTION AUTHORS :AzadehNikfarjam , EhsanEmadzadeh , Saravanan Muthaiyah.
In this paper, the main components of such forecasting systems have been introduced. In addition, different developed prototypes have been introduced and the way whereby the main components are implemented compared. Based on studied attempts, the potential future research has been activities suggested.

III. EXISTING SYSTEM APPROACH

Stock prediction using linear regression analysis:
Linear regression is used to predict data. Linear regression is used for modelling the relationship between a scalar dependent variable y and one or more explanatory variables denoted X. There are advances in this field but the limitations remain the same. Simple Linear Regression is the one where only one explanatory variable are used.

Disadvantages:

1. Considers only two columns of the dataset for analysis.
2. The open value and close value is considered.
3. But the accuracy given is not satisfactory

IV. PROPOSED SYSTEM APPROACH

Stock prediction using multiple regression technique:

Multiple regression is an extension of simple linear regression. It is mostly used when we want to predict the value of a variable based on the value of two or more other variables as we need to in the prediction process. The variable we want to predict is called the dependent variable (or sometimes, the outcome, target or criterion variable).

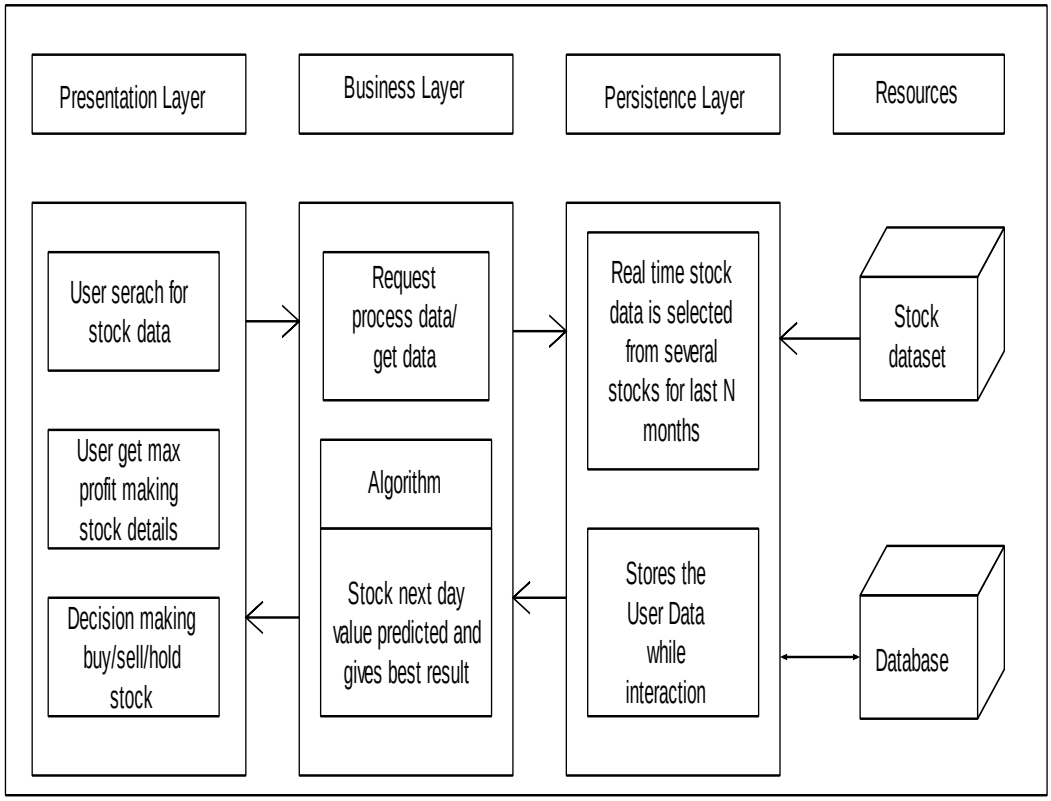
Advantages:

1. It is a flexible method
2. It can process more than one column in the dataset to give more accurate result
3. Multiple regression uses multiple independent variables, with each controlling for the others.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering
(An ISO 3297: 2007 Certified Organization)
Website: www.ijircce.com
Vol. 5, Issue 2, February 2017

Architecture



V. PROPOSED SYSTEM ALGORITHM

Regression Analysis Algorithm:

1. Regression predicts a numerical value.
2. Regression performs operations on a dataset where the target values have been defined already. And the result can be extended by adding new information.
3. The relations which regression establishes between predictor and target values can make a pattern. This pattern can be used on other datasets which their target values are not known.
4. We prepare the input data by using that historical data. In this model the input data is grouped into two sets as training data set and testing data set.
5. The training data set is used to train a model and to estimate the unknown coefficients of the auto regression equation. These coefficients are estimated.
6. The estimated coefficients are used to predict the future price of a stock.
7. Thus the coefficients are used to test the testing data set and the comparison is done between actual price and predicted price.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering
(An ISO 3297: 2007 Certified Organization)
Website: www.ijircce.com
Vol. 5, Issue 2, February 2017

VI. RESULT AND GRAPH COMPARISON

Linear Regression	Multiple Regression
Analyses data for only two columns in a dataset (X and Y axis)	Analyses data for more than two columns in a dataset
Less Efficient than Multiple Regression	More efficient than linear regression
Result accuracy is less in comparison	Result accuracy is more in comparison

The proposed model uses regression analysis as a data mining technique and developsa system for exploiting time series data in financial institution. A prediction system has been built that uses data mining technique to produce periodically forecasts about stock market prices. It can be very useful for the investors to use this to gain maximum profit.

VII.CONCLUSION AND FUTURE SCOPE

The proposed model uses regression analysis as a data mining technique and developsa system for exploiting time series data in financial institution. A prediction system has been built that uses data mining technique to produce periodically forecasts about stock market prices. The use of back propagation in neural network enables us to reduce errors and improve accuracy of the system. It can be very useful for the investors to use this to gain maximum profit. There are many research directions which might be considered in the future work. Improving the accuracy of the predictive models is one of them. Accuracy can be improved by considering an entirely different aspect i.e., human sentiments. As future scope of stock market is limitless, the demand for its data analysis will be ever increasing. By changing only the training data, the proposed system can be used for any stock markets of other countries. With few altercations, the system can be used for various purposes such as predicting prices of commodities like gold, predicting the fuel consumption of a vehicle as well as monitoring health of a patient.

REFERENCES

1)Dr. P. K. Sahoo, Mr. Krishna Charlapally, “Stock Price Prediction Using Regression Analysis”.

2)Sachin Kamley, Shailesh Jaloree R. S. Thakur, “A Linear Regression Approach To Prediction Of Stock Market Trading Volume:A Case Study”

3. Multipleregressions: A Data Mining Approach For Predicting The Stock Market Trends Based On Open, Close And High Price Of The Month.

4. K. V. Sujatha, S. Meenakshi Sundaram, “Stock Index Prediction Using Regression And Neural Network Models Under Non Normal Conditions”

5.Han Lock Siew, Md Jan Nordin, “Regression Techniques For The Prediction Of Stock Price Trend”

6.Azadehnikfarjam, Ehsanemadzadeh, Saravananmuthaiyah, “Text Mining Approaches For Stock Market Prediction”