A Project/Dissertation Review Report

on

RULE THE MARKET

Submitted in partial fulfillment of
therequirementfortheawardofthedegreeof

Bachelor of Technology (CSE)

(Established under Galgotias University Uttar Pradesh Act No. 14 of 2011)

**Under The Supervision
of
Ms. Garima Pandey
Assistant Professor**

Submitted By

MOHAMMAD FAIZ

(19SCSE1010897)

MD IMRAN RAIN

(19SCSE1010887)

**SCHOOL OF COMPUTING SCIENCE AND ENGINEERING
DEPARTMENT OF COMPUTER SCIENCE AND
ENGINEERING GALGOTIAS UNIVERSITY,GREATER
NOIDA INDIA
DECEMBER , 2021**

Abstract

Top Class Stock Price Prediction Project through Machine Learning Algorithms for Google. Easy Understanding and Implementation.

Stock Market Analysis and prediction is a project for technical analysis, visualization, and estimation using Google Financial data. Seeing data from the market, especially some general and other software columns. Pandas used to take stock of the information, looked at different aspects of it, and finally looked at it in some way to assess the risk of a stock based on its recent performance history. Competing with the Monte Carlo method in anticipation of future prices.

Stock exchange analysis is only intended for the analysis of stock company data for various organizations. Using this method of data analysis, any organization can easily extract relevant information.

The main goal of my project is to analyze the data of all the institutions in which form we need. We evaluated two basic measurements of the analysis and found no conclusive evidence about their estimated value.

These predictions are also very long-lasting and will see a year in the future. Suggestions on this scale are not the main project time. Instead, we will focus on predicting daily market trends.

Because of these problems, we avoided basic analysis.

List of Figures

Table of Content

CHAPTER-1
Introduction

Purpose

The growing trend of people investing in Stock Market brought us to the unanimous decision of working on this project. The project will try to cover the need of the investors investing on various stock by helping them visualize the bricks and pieces of current market along with providing a future vision of what market will be. In short, it aims to help investors analyze the risk and invest accordingly.

Background

Forecasting the stock market due to its importance and popularity

among the masses and also small and large companies due to financial benefits and low risk is growing topic in research. It plays a crucial role in portfolio management, household finance decisions, business invest- ment planning, and policy formulation. Knowing what will be the stock price tomorrow will help investor decide the right time to buy or sell the stock.

With the advancement in technology we now have the ability to generate and collect enor- mous amount of data. There are many databases that are used for storing business data. We can readily use them for data mining application but the data information such as news are not stored in databases. The news articles are scattered across the various website of World Wide Web. That news is to be collected and classified for us to be able to use it in data min- ing application. Nowadays the stock market has been called for research in many fields due to its effects on financial challenging and capacity of predicting its various aspects through different scientific methods such as genetic algorithm, Artificial Neural Network(ANN) and other Meta heuristic algorithms. Many institutions and academic researchers are trying to propose a method for predicting next day behaviour of stock indexes in order to be better than other methods.

2

Despite the evolvement of high computational machines capable of performing incredible task, many economists are still reluctant to accept the validity of 'Computers forecasting the stock price Many have to say that the stock price for tomorrow are merely the results of people's senti- ment and hence has nothing to do with past records.

Having said all that, researchers haven't yet given up on forecasting the stock price.

Formulation of Problem

Before an investor invests in any stock, he needs to be aware how the stock market behaves. Investing in a good stock but at a bad time can have disastrous results, while investment in a mediocre stock at the right time can bear profits. Financial investors of today are facing this problem of trading as they do not properly understand as to which stocks to buy or which stocks to sell in order to get optimum profits. They are well aware that news moves financial markets. Hence analyzing news and other information about a particular stock before investing is the key here. Intelligent investors use Machine learning and Text mining techniques in predicting the stock market behavior which we believe will give more accurate results than analysis of numerical time series alone. This will allow financialÂaˇanalysts to foresee the behavior of the stock that they are interested in and thus act accordingly.

The world has come across some huge stock market crashes like:

- The Dutch Tulip Mania (aka ''Tulipomania'') of 1634-1637

- The South Sea Bubble (1716-1720)

- The Mississippi Bubble (1716-1720)

- The British "Railway Mania" Bubble

- The Florida Real Estate Bubble of the 1920s

- The Stock Market Crash of 1929 (US)

Such a sudden stock market crashes could potentially bring thousands of lives down to road and cause a chaos. An alternative to take precautions from such crisis could possibly be the use of available forecasting systems.

Component Diagram

Component diagram is a special kind of diagram in UML. The purpose is also different from all other diagrams discussed so far. It does not describe the

functionality of the system but it describes the components used to make those

functionalities.

Component diagrams are used in modeling the physical aspects of object-oriented systems that are used for visualizing, specifying, and documenting

component-based systems and also for constructing executable systems through forward

and reverse engineering. Component diagrams are essentially class diagrams that focus on

a system's components that often used to model the static implementation view of a system.

Tool and Technology Used

The system must also be able to handle heavy traffic smoothly. Security

Tools

- Django Web Framework

Django is a high-level Python Web framework that encourages rapid development and clean, pragmatic design. Built by experienced developers, it takes care of much of the hassle of Web development, so one can focus on writing one's app without

needing to reinvent the wheel. It is free and open source.

- AM Charts (JavaScript Charts)

AM Charts made it easy to display complex data visualizations. Combine various graph types on a single chart. Create clusters, or stacks, or clusters of stacks. Control the widths, open and close values, apply coloring based on value thresholds or changes, recalculate the values automatically. Use various value scales, including date and time. Those are just a few examples of what we can do.

.

Techniques

- **Moving Average:** The moving average (MA) is a simple technical analysis tool that smooths out price data by creating a constantly updated average price. The average is taken over a specific period of time, like 10 days, 20 minutes, 30 weeks, or any time period the trader chooses. Moving average strategies are popular throughout the world for stock anal- ysis and can be tailored to any time frame, suiting both long term investors and short-term traders.

- **Weighted Moving Average:** The Weighted Moving Average places more importance on recent price moves, therefore, the Weighted Moving Average reacts more quickly to price changes than the regular Simple Moving Average [1]. A basic example (3-period) of how the Weighted Moving Average is calculated is presented below:

CHAPTER-2
Literature Survey

Introduction

This chapter reviews on existing system, Trends analyzing methods, technologies and development methodologies.

Review on Existing System

Studies Using Artificial Neural Networks to Predict Stock Market Values The first set of articles includes studies that primarily focus on stock market prediction using artificial neural networks (ANNs). ANNs are computational models based on biological neural networks. In the network, sets of nodes are grouped into layers starting with an input layer and ending with an output layer. Signals are transmitted (propagated) through the connected nodes as they learn based on examples and attempt to reduce the level of prediction error. As the system is working to improve its performance, weights are adjusted for the signals between connected nodes. The following provides a brief description of each ANN-related study's unique research focus and findings. Jasic and Wood (2004) developed an artificial neural network to predict daily stock market index returns using data from several global stock markets. The focus is on trying to support profitable trading. A method is introduced based on univariate neural networks

using untransformed data inputs to provide.

1.User Interface: The user is required to select which company is he interested in amongst the various companies that have been provided.

2.Asp.net with SQL: As communication with user is performed in C# and data required for processing is in a Database, a connectivity has to be implemented between the Database and C# application

Analysis of stocks using data mining will be useful for new investors to invest in stock market based on the various factors considered by the software.

Stock market includes daily activities like sensex calculation, exchange of shares. The exchange provides an efficient and transparent market for trading in equity, debt instruments and derivatives.

Our software will be analyzing sensex based on company's stock value. The stock values of company depend on many factors, some of them are:

1> Demand and Supply:

Demand and Supply of shares of a company is a major reason price change in stocks. When Demand Increase and Supply is less, price rises.

and vice versa.

2> Corporate results: This will be regarding to the profits or progress of the company over a span of time say 3 months.

3> Popularity: Main Strength in hands of share buyer. Popularity of a company can effect on buyers. Like if any good news of a company, may result in rise of stock price. And a bad news may break dreams.

The stock value depends on other factors as well, but we are taking into consideration only these main factors.

1.User Interface: The user is required to select which company is he interested in amongst the various companies that have been provided.

2.Asp.net with SQL: As communication with user is performed in C# and data required for processing is in a Database, a connectivity has to be implemented between the Database and C# application

Analysis of stocks using data mining will be useful for new investors to invest in stock market based on the various factors considered by the software.

Stock market includes daily activities like sensex calculation, exchange of

shares. The exchange provides an efficient and transparent market for trading in equity, debt instruments and derivatives.

Our software will be analyzing sensex based on company's stock value. The stock values of company depend on many factors, some of them are:

1> Demand and Supply:

Demand and Supply of shares of a company is a major reason price change in stocks. When Demand Increase and Supply is less, price rises. and vice versa.

2> Corporate results: This will be regarding to the profits or progress of the company over a span of time say 3 months.

3> Popularity: Main Strength in hands of share buyer. Popularity of a company can effect on buyers. Like if any good news of a company, may result in rise of stock price. And a bad news may break dreams.

The stock value depends on other factors as well, but we are taking into consideration only these main factors.

Sentiment dataset

Sentiment dataset has been created by considering news dataset and

tweets. Both tweets and news has been collected for one year and sentiment analysis algorithm is applied on the same. Sentiment of the tweets and sentiment of the news are integrated on a daily basis.

The classification models are built for stock market data analysis. Performance of the model is evaluated using accuracy metric. Accuracy can be defined as proportion of true results in the test dataset.

Results

Daily prediction model

Label has been assigned to each transaction date by considering close price on a particular day with close price of the previous day. Total observation per company is 261 transaction dates. In these, 182 transaction dates data are used for training the model. The rest 79 transactions are used for testing the model. The graph in Fig. 3 depicts original trend versus predicted trend for all three sectors. The Table 1 shows the accuracy achieved in the daily prediction

short-term stock market index return predictions. The study uses the daily closing values of the Standard and Poor's 500 Index (S&P 500), the German DAX Index, the Japanese TOPIX index, and London's Financial Times Stock Exchange Index (FTSE All Share). The samples for the S&P 500, DAX and FTSE Index are from January 1, 1965 to November 11, 1999. The sample for TOPIX covers the period from January 1, 1969 to November 11, 1999 since data from earlier years was not available. The prediction performance for the neural network is evaluated against a benchmark linear autoregressive model and prediction improvement is confirmed when applied to the S&P 500 and DAX indices. Enke and Thawornwong (2005) use a machine learning information gain technique to evaluate the predictive relationships for numerous financial and economic variables. By computing the information gain for each model variable, a ranking of the variables is obtained. A threshold is determined to select only the strongest relevant variables to be retained in the forecasting models. Neural network

models for level estimation and classification are examined for their ability to provide an effective forecast of future values. A cross-validation technique is also employed to improve the generalizability of several models. The models are compared using S&P data from a 24-year period from March 1976 to December 1999. The results show that the trading strategies guided by the classification models generate higher risk-adjusted profits than the buy-and-hold strategy, the other neural network models, and the linear regression models.

System Flow-

Project Design

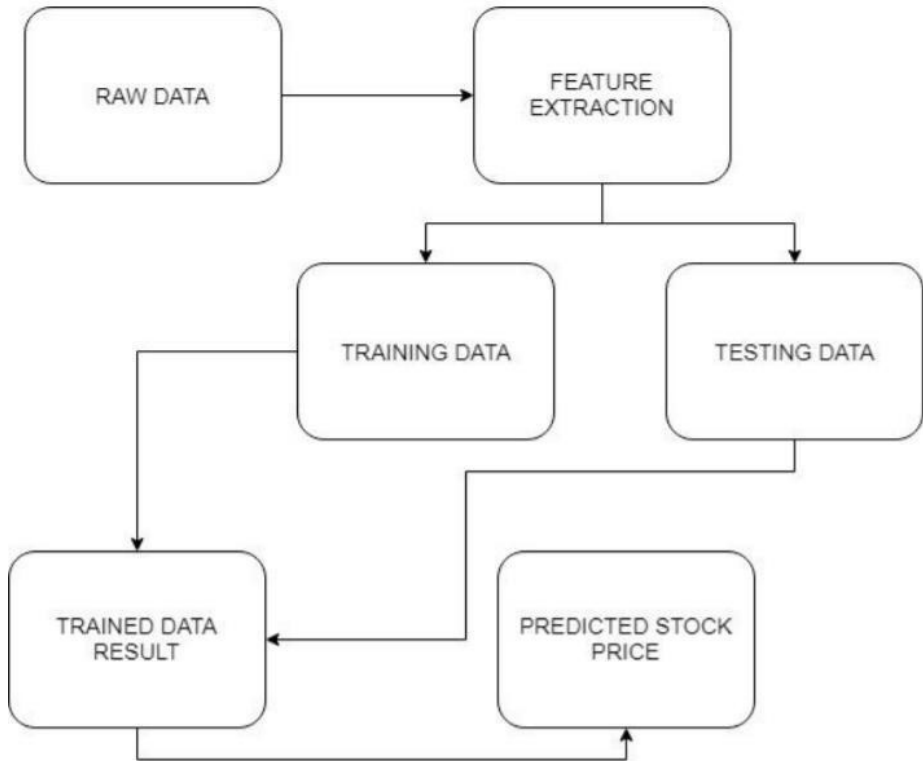FIGURE 2.1

The level 0 dataflow diagram is shown in Figure 2.2.

FIGURE 2.2

Use Case Diagram

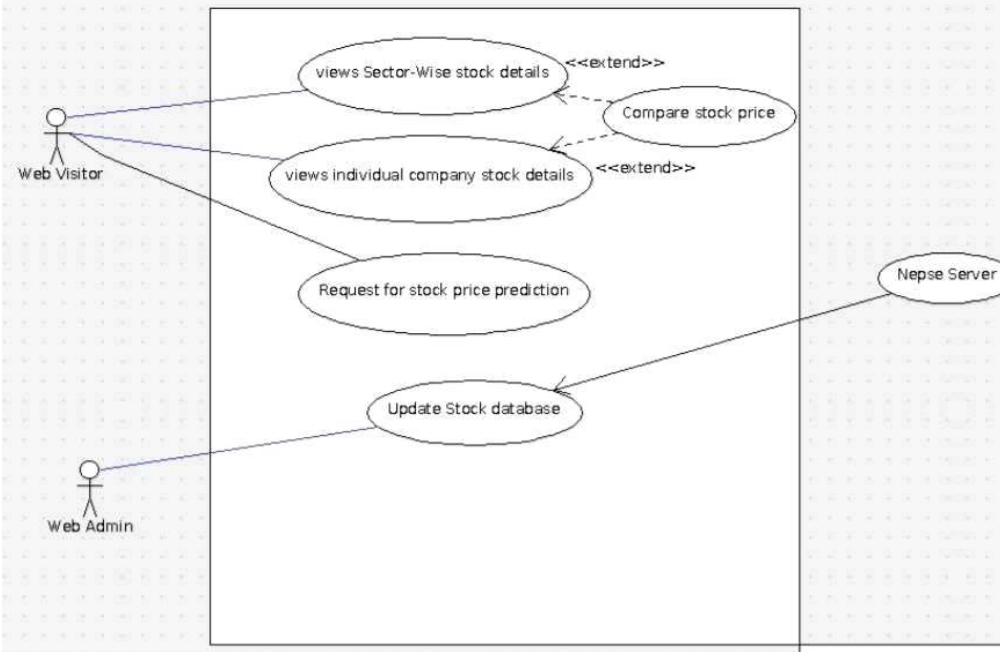A use case diagram shows the various actors that can act upon the system and what actions they can perform

FIGURE 2.3

CHAPTER-3
Functionality of Project

In the proposed method two datasets are used. The first one is historical price dataset, and second one is sentiment dataset of that company. Sentiment has been calculated from the obtained news and tweets information of the company. Companies are selected in such a way that it is from different sector. One of the company is from oil sector, another is from bank sector and last one is from mining sector. For daily prediction model, last one year's tweets and historical data is considered. It has almost 260 rows of data. For monthly prediction model, historical data from 2003 has been collected.

Historical prices are obtained from Yahoo Finance. Each transaction date consists of open price, close price, low price, high price, adjusted close price and volume traded on that day. Adjusted close price and close price depicts the close price of stock on a particular day. Adjusted close price will be adjusted for dividends and splits. Adjusted close price is considered as stock price as in other researches17, 18.

Our aim is to create software that analyzes previous stock data of certain companies, with help of certain parameters that affect stock value. We are going to implement these values in data mining algorithms. This will also help us to determine the values that particular stock will have in near future. We will determine the Month's High and Low with help of data mining algorithms.

Functional Requirements

1.User Interface: The user is required to select which company is he interested in amongst the various companies that have been provided.

2.Asp.net with SQL: As communication with user is performed in C# and data required for processing is in a Database, a connectivity has to be implemented between the Database and C# application

Analysis of stocks using data mining will be useful for new investors to invest in stock market based on the various factors considered by the software.

Stock market includes daily activities like sensex calculation, exchange of shares. The exchange provides an efficient and transparent market for trading in equity, debt instruments and derivatives.

Our software will be analyzing sensex based on company's stock value. The stock values of company depend on many factors, some of them are:

1> Demand and Supply:

Demand and Supply of shares of a company is a major reason price change in stocks. When Demand Increase and Supply is less, price rises. and vice versa.

2> Corporate results: This will be regarding to the profits or progress of the company over a span of time say 3 months.

3> Popularity: Main Strength in hands of share buyer. Popularity of a company can effect on buyers. Like if any good news of a company, may result in rise of stock price. And a bad news may break dreams.

The stock value depends on other factors as well, but we are taking into consideration only these main factors.

Sentiment dataset

Sentiment dataset has been created by considering news dataset and tweets. Both tweets and news has been collected for one year and sentiment analysis algorithm is applied on the same. Sentiment of the tweets and sentiment of the news are integrated on a daily basis.

The classification models are built for stock market data analysis. Performance of the model is evaluated using accuracy metric. Accuracy can be defined as proportion of true results in the test dataset.

Results

Label has been assigned to each transaction date by considering close price on a particular day with close price of the previous day. Total observation per company is 261 transaction dates. In these, 182 transaction dates data are used for training the model. The rest 79 transactions are used for testing the model. The graph in Fig. 3 depicts original trend versus predicted trend for all three sectors. The Table 1 shows the accuracy achieved in the daily prediction

CHAPTER-4
Conclusion

In the past few years, it has been observed that most of the people are investing in the stock market to make money easily. At the same time investor has high chance of losing all money invested. So an efficient predictive model is required for the user to

understand future markettrend.

There are many predictive models which tell about the market trend whether it is up or down, but they fail to give accurate results. An attempt has been made to build efficient predictive model of stock market where the trend for the next day is predicted. By considering various patterns like continuous up/down, volume traded per day and also including sentiment of the company a model has been built and tested with different stock market data available open source.

On the considered dataset, Decision Boosted Tree is performing better than Support Vector Machine and Logistic Regression.

The dataset which was been considered for sentiment analysis may be sparse which means we may not have news/tweet for a particular company for many days. In such cases Principle component analysis with multiple factors can be applied. The impact of intra day price movement for the next day stock price can be considered to improve the accuracy. sentiments.

REFERENCES

Datasets: Nifty 50 data , tw_spydata_raw(trading data)

[1] Stock Price Prediction Using LSTM on Indian Share Market by Achyut Ghosh, Soumik Bose1, Giridhar Maji, Narayan C. Debnath, Soumya Sen

[2] S. Selvin, R. Vinayakumar, E. A. Gopalkrishnan, V. K. Menon and K. P. Soman, "Stock price prediction using LSTM, RNN and CNN-sliding window model," in International Conference on Advances in Computing, Communications and Informatics, 2017.

[3] Murtaza Roondiwala, Harshal Patel, Shraddha Varma, "Predicting Stock Prices

Using LSTM" in Undergraduate Engineering Students, Department of Information Technology, Mumbai University, 2015.

[4] Xiongwen Pang, Yanqiang Zhou, Pan Wang, Weiwei Lin, "An innovative neural network approach for stock market prediction", 2018

[5] Ishita Parmar, Navanshu Agarwal, Sheirsh Saxena, Ridam Arora, Shikhin Gupta, Himanshu Dhiman, Lokesh Chouhan Department of Computer Science and Engineering National Institute of Technology, Hamirpur – 177005, INDIA - Stock Market Prediction Using Machine Learning.

[6] Pranav Bhat Electronics and Telecommunication Department, Maharashtra Institute

of Technology, Pune. Savitribai Phule Pune University - A Machine Learning Model for Stock Market Prediction.

[7] Anurag Sinha Department of computer science, Student, Amity University Jharkhand Ranchi, Jharkhand (India), 834001 - Stock Market Prediction Using Machine Learning.

[8] V Kranthi Sai Reddy Student, ECM, Sreenidhi Institute of Science and Technology,

Hyderabad, India - Stock Market Prediction Using Machine Learning.

[9] Asset Durmagambetov currently works at the mathematics, CNTFI. Asset does research in Theory of Computation and Computing in Mathematics, Natural Science, Engineering and Medicine. Their current project is 'The Riemann Hypothesis-Millennium Prize Problems' - stock market predictions.

59

[10] Mariam Moukalled Wassim El-Hajj Mohamad Jaber Computer Science

Department

American University of Beirut - Automated Stock Price Prediction Using Machine

Learning.

[11] Manh Ha Duong Boriss Siliverstovs June 2006 - The Stock Market and

Investment.

[12] Dharmaraja Selvamuthu , Vineet Kumar and Abhishek Mishra Department of

Mathematics, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016,

India - Indian stock market prediction using artificial neural networks on tick data.

[13] Lufuno Ronald Marwala A dissertation submitted to the Faculty of Engineering

and

the Built Environment, University of the Witwatersrand, Johannesburg, in

fulfilment of the requirements for the degree of Master of Science in Engineering -

Forecasting the Stock Market Index Using Artificial Intelligence Techniques.

[14] Xiao-Yang Liu1 Hongyang Yang,Qian Chen4,Runjia ZhangLiuqing Yang Bowen

Xiao Christina Dan Wang Electrical Engineering, 2Department of Statistics,

3Computer Science, Columbia University, 3AI4Finance LLC., USA, Ion Media

Networks, USA, Department of Computing, Imperial College, 6New York

University (Shanghai) - A Deep Reinforcement Learning Library for Automated

Stock Trading in Quantitative Finance.

[15] Pushpendu Ghosh, Ariel Neufeld, Jajati Keshari SahooDepartment of Computer

Science & Information Systems, BITS Pilani K.K.Birla Goa campus, India

bDivision of Mathematical Sciences, Nanyang Technological University, Singapore

cDepartment of Mathematics, BITS Pilani K.K.Birla Goa campus, India -
Forecasting directional movements of stock prices for intraday trading using LSTM
and random forests.

[16] Xiao Ding, Kuo Liao, Ting Liu, Zhongyang Li, Junwen Duan Research Center for
Social Computing and Information Retrieval Harbin Institute of Technology, China
- Event Representation Learning Enhanced with External Commonsense
Knowledge.

[17] Huicheng Liu Department of Electrical and Computer Engineering Queen's
University, Canada - Leveraging Financial News for Stock Trend Prediction with
Attention-Based Recurrent Neural Network.

[18] Hyeong Kyu Choi, B.A Student Dept. of Business Administration Korea
University
Seoul, Korea = Stock Price Correlation Coefficient Prediction with ARIMA-LSTM
Hybrid Model.