

Stock Market Value Prediction Using DBSCAN Algorithm

¹Rinkesh Shah, ²JairajShettigar, ³SmeetPurohit, ⁴Sweta Jain

^{1, 2, 3}Student, IT, Rajiv Gandhi Institute of Technology, Mumbai, India.

⁴Assistant Professor, IT, Rajiv Gandhi Institute of Technology, Mumbai, India

Abstract—A stock market or equity market is a private or public market for the trading of company stock; these are securities listed on a stock exchange as well as those only traded privately. This paper focus on developing an android based application in which user can analysis the stock market by viewing graph of previous year, can view the predicted result and can see the live stock quote. The paper propose DBSCAN algorithm to predict stock market. By taking the historical data and analyzing it, we will be forming cluster and will be applying to live stock market. As Stock market is an integral part of society where almost everyone invest in it and earn a profit, this project will be huge beneficial to all.

Index Terms— Android Studio, Clustering, Classification, Data mining, R, Time Series Analysis and Forecasting,

I. INTRODUCTION

Stock market is one of the major part in an individual life. Individual, investor, or business person always want to analysis the past data and predict the outcome of stock price and invest in it and earn a profit. But due to unavailability of time an individual is unable to do it. This paper focuses on Stock market prediction, an android application so that almost everyone can access it. Our motive is to analysis the past record of data, extract meaningful information from it and predict the stock market.

Below is the activity diagram of the user. A user needs to register first and after that he can login into application. Once he login then he can select the Stock which he want to Analysis or predict or want to see live Stock quote.

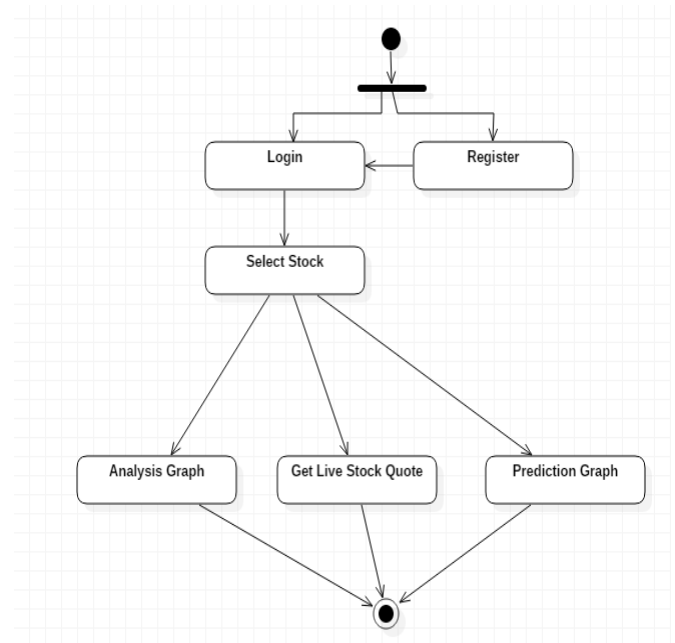


Fig.1 Activity Diagram

Analysis Graph will help user to see the high and low value of share of previous year so that he can analysis the meaningful information from it. Prediction Graph will predict the data by analysis the historical data and applying suitable algorithm. Get live Stock Quote will display the live stock price along with day's high, low and change percentage.

II. EXISTING SYSTEM

Investing in stock market without having any prior knowledge is risk. Stock market is the source where investor can make money faster what he dream of. But taking huge risk can result in huge loss.

People hire professional expert or invest in professional firm which invest into stock market. But by doing so there's still no guarantee of huge profit. Sometime it can also take years to earn a little profit. Investor also have to pay some amount of profit to the professional expert of

firm which again is lost to investor and also expert from professional isn't guarantee a profit, can also result a loss.

Also some website give a mock investment to investor on how stock market looks like based on live market. This website also provides prediction of stock market which helps investor to invest. However website is difficult to access through mobile and also they are payed ones which work on subscription.

III. PROPOSED WORK

Below figure shows the block diagram of the stock market prediction. Firstly, we will obtain data in raw from which by processing by using various algorithm will convert into organized form data and store in the database. Now by using Normalization process we will convert the huge amount of data in small size which would be easier to process and apply various algorithms. The purpose of this step is to reduce the complexity and increase the accuracy of the process.

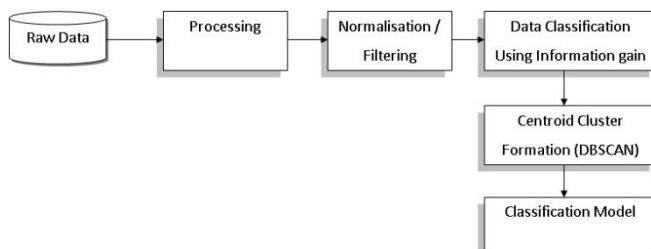


Figure.2 Basic model of a Stock market prediction

After we obtained organized and well-structured data from raw data, next step is to select the relevant data from the whole structure by using information gain of each attribute and thus we classify the given set of data. Next step is to form cluster based on the predict value of the data. We will be using Regression or time series analysis to predict the value of a stock and by using centroid clustering based algorithm we will be forming cluster based on the attribute of the data. Centroid algorithm using DBSCAN is a process in which we select a point regardless of the point being in the dataset or not and draw a circle of radius 'r' and the data present in the cluster would perform cluster having same attribute and predicted value.

IMPLEMENTATION DETAILS :

This explains the steps that should be followed for the real life implementation of stock market value prediction using DBSCAN algorithm.

Data collection: The first phase of our implementation is to collect a real time database of stock market. The database should be dynamic. There are many source of real time stock data like moneycontrol, yahoo finance or can be given by a stock broker which can be extracted by various languages like python.

Pre-Processing: The second phase of this implementation is to process the data. Now it is the most important step as to store data which can be use efficiently. The data which we got are in form of raw data which contain irrelevant and redundant information present or noisy and unreliable data. So by applying various pre-processing steps like cleaning, normalization, transformation, feature extraction and selection we can convert the database into well-structured which can be proved efficient.

Normalization: Normalization means rescaling of attribute. It is one of the stages of data preprocessing where we find new range from an existing one which can effective and useful in forecasting or prediction purpose. Normalization also improve the accuracy and efficient of data mining algorithm.

Time series analysis and forecasting: Time series is basically a sequence of all data point of Stock. Time series analysis involves method for analyzing past time series data in order to extract meaningful data and forecasting is use for predicting the future value based on past value.

Centroid based Clustering (DBSCAN): This is final phase of implementation where we form cluster based on the attribute and analysis or prediction we got from above step. Clustering means grouping of similar kind of data who have similar kind of attribute value.

IV. RESULT

A. Live Stock Quote

Below is the Screenshot of get live stock quote of Google.

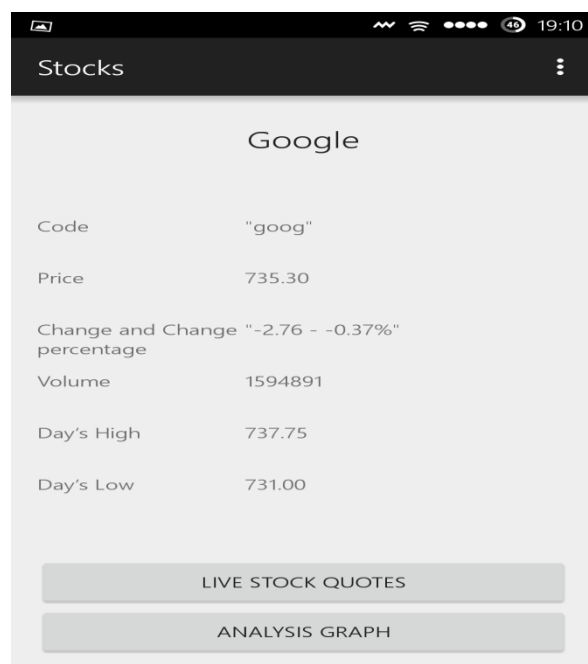


Figure.3 Live Stock Quote

Live Stock quote will return real time data of google like price, high, low, change, change percentage, volume etc. This will help user to keep intact with real world price. This work as getting data from yahoo finance using php language then storing in MySQL table and then sending to android phone.

B. Analysis Graph

Below is the screenshot of analysis graph of Google

Showing data of Google

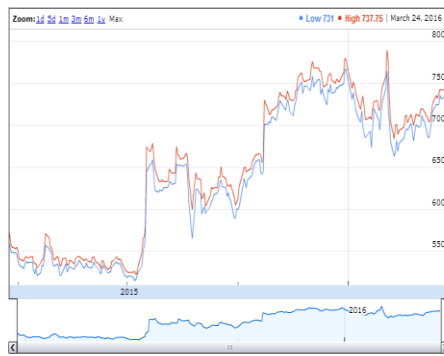


Figure.4 Analysis Graph

This graph will help user to analysis past year record of company so that users can understand various aspect of company. This is implemented using php and JavaScript. Php script downloads recent data from yahoo and store in table. Using JavaScript it plots the graph and return the graph and using Webview it display the result on android

C. Prediction Graph

Below is the Screenshot of Prediction graph of Google. This Graph shows the final predicted value so that it helps user for long time investment.

Prediction

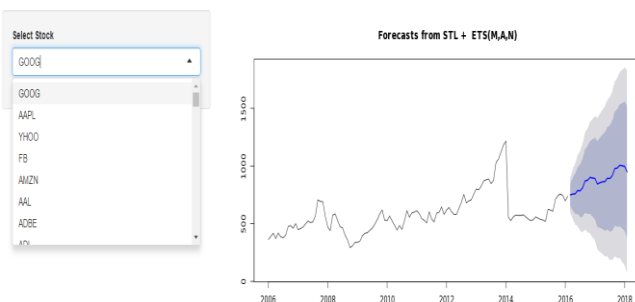


Figure.5 Prediction Graph

This is implemented using R language. It takes the past record of data and using time series forecasting it predict the result. This script is connected to cloud and using shiny UI it is being displayed in web which is connected through android via Webview.

V. CONCLUSION

The technical performance of this application will be predicting stock values on basis of past values, hence results can be expected to be precise, this application won't be charging any fees or subscription charges to the users, using this application will be easy and portable. Stock market prediction in this particular application will done using data mining, this prediction will be done by making clusters, and clusters formed will help to group the units.

REFERENCES

- [1] CicilFonseka "A Data mining algorithm to analyse stock market data using lagged correlation"
- [2] Fan Cai: "Toward a New Classification Model for Analysing Financial Datasets"
- [3] Jiayi Yao: "The Application of Stream Data Time-Series Pattern Reliance Mining in Stock Market Analysis"
- [5] Kim, Suk-Joong and McKenzie, Michael D., "Conditional Autocorrelation and Stock Market Integration" 2006, Available at SSRN: <http://ssrn.com/abstract=943970>
- [6] T. Wittman. (2002, December). Time-Series Clustering and Association Analysis of Financial Data
- [7] J. Han and M. Kamber, Data Mining: Concept and Techniques. Morgan Kaufmann publishers, 2nd Eds., Nov. 2005.
- [8] P-N. Tan, M. Steinbach, and V. Kumar, Introduction to Data Mining, Addison Wesley, 2006, pp.150-172