VISUALISING AND FORECASTING THE STOCK PRICE

Sowkya Irupaka[1], Salma Sahaba[2], Hadiya Fatima[3], Shivani Yadao[4]

Department of Computer Science and Engineering, Stanley College of Engineering and Technology for Women, Telangana, India

Abstract

Stock investments provide one of the highest returns in the market. Even though they are volatile in nature, one can visualize share prices and other statistical factors which helps the keen investors carefully decide on which company they want to spend their earnings on.

Developing this simple project idea using the Dash (a Python framework) to create web application, we can make dynamic plots of the financial data of a specific company by using the tabular data provided by yfinance python library and some machine learning models which will show company information and stock plots based on the user input stock code. On top of it, we can use a machine learning algorithm-Support Vector Regression (SVR) module from the sklearn library and then train the SVR model with the training dataset to predict the upcoming stock prices for the date inputted by the user.

Keywords:

Stock Price, Dash, Web Application, Dynamic Plots, yfinance python library, Machine Learning, Support Vector Regression, Sklearn library, Indicators, Forecasting.

1. Introduction

1.1 About Project

The stock market is defined to be the collection of markets and exchanges where regular activities of buying and selling of shares of publicly-held companies take place. It is a place where shares of pubic listed companies are traded. The primary market is the place where companies float shares to the general public in an initial public offering (IPO). This is done to raise capital. People mainly buy the stocks in the expectation that their price may rise in the future. But there is always an uncertainty in the stock market due to which people is not willing to invest their money in the stock markets. Thus we need a technique which can predict the stock market prices, so that people can invest their money in the best stocks.

The financial market is a dynamic and composite system where people can buy and sell currencies, stocks, equities and derivatives over virtual platforms supported by brokers. The stock market allows investors to own shares of public companies through trading either by exchange or over the counter markets. This market has given investors the chance of gaining money and having a prosperous life through investing small initial amounts of money, low risk compared to the risk of opening new business or the need of high salary career. Stock markets are affected by many factors causing the uncertainty and high volatility in the market. Although humans can take orders and submit them to the market, automated trading systems (ATS) that are operated by the implementation of computer programs can perform better and with higher momentum in submitting orders than any human. However, to evaluate and control the performance of ATSs, the implementation of risk strategies and safety measures applied based on human judgements are required. Many factors are incorporated and considered when developing an ATS, for instance, trading strategy to be adopted, complex mathematical functions that reflect the state of a specific stock, machine learning algorithms that enable the prediction of the future stock value, and specific news related to the stock being analysed.

Time-series prediction is a common technique widely used in many real-world applications such as weather forecasting and financial market prediction. It uses the continuous data in a period of time to

International Journal for Innovative Engineering and Management Research
A Peer Reviewed Open Access International Journal
www.ijiemr.org

predict the result in the next time unit. Many timeseries prediction algorithms have shown their effectiveness in practice.

1.2 Objectives of the Project

We will be creating a single-page web application using Dash (a python framework) and some machine learning models which will show company information (logo, registered name description) and stock plots based on the stock code given by the user to get predicted stock prices for the date inputted by the user. Also the ML model will enable the user to get predicted stock prices for the date inputted by the user.
Developing this simple project idea using the Dash library (of Python), we can make dynamic plots of the financial data of specific company by using the tabular data provided by yfinance python library. On top of it we can use a machine learning algorithm to predict the upcoming stock prices.
This project is good start for beginners in python/data science and a good refresher for professionals who have dabbled in python / ML before. This web application can be applied to any company (whose stock code is available) of one's choosing.

1.3 Scope of the Project

We can analyse various stock and study it. Also easily use to study and invest accordingly into various companies Scope of stock price determining the future scope of market. A system is essential to be built which will work with maximum accuracy and it should consider all important factors that could influence the result. Predicting the performance of a stock market is tough as it takes into account various factors. An input field for number of days is provided to forecast for the respective input.
Stock Market Analysis of stocks using data mining will be useful for new investors to invest in market based on the various factors considered by the software.

2. Literature Survey
2.1 Existing System

As many have invested their time and effort in this world trade for getting it closer and more reliable to the people for carrying out the resources and make their lifestyle more deliberate than the previous. In the past few years various strategies and the plans had been derived and deployed ever since it's continuation and the topic is still a point of research where people are coming up with ideas to solve.

Intelligence fascinates mankind and having one in machine and integrating on the same, is the hot key of research. There are various people contributing on the same research-ASheta tried its invention on two nonlinear process and had came up with TS which is used as a model for fuzzy sets.

All the learning system from the past are limited and are simplest in nature where learning of the simple algorithm for a computational mean is not enough which can even be done by human brain itself. The main motto of learning was limitized and learning model was not efficient.

The existing models can't cope up with the vulnerabilities and remove the rarest information that they can't process causing it a major data loss which creates a problem in forecasting. Observation is the integral part in the resource and prediction management. If the outcome can't be observed it's point of time estimation is compromised causing it less liable in market. Monitoring of the same is not possible in the existing system.

Loss of sights is a major problem in the existing system as the stock varies each days and the loss margin can be higher with respect to time. An initial instance is taken for prediction.

The existing system in stock market predictions are apparently biased because it consider a only source point for data source. Before the prediction of the data set a simple data retrieval should be generated and tested on the training data set which are more flexible and versatile in nature.

2.2 Proposed System

In proposed work we will be creating a single-page web application using Dash (a python framework) and some machine learning models which will show company information (logo, registered name and description) and stock plots based on the stock code given by the user. Also the ML model will enable the user to get predicted stock prices for the date inputted by the user.

Generating company's information including name, logo and description and visualizing graphs generated with respect to input provided.
Developing this simple project idea using the Dash library (of Python), we can make dynamic plots of the financial data of a specific company by using the tabular data provided by yfinance python library. On top of it, we can use a machine learning algorithm to predict the upcoming stock prices. Here the graphs are plotted dynamically.

Learning model forecast the upcoming stock prices for the date inputted by the user.
Provides the interface where the investor itself can interact with the application.

3. Proposed Architecture

Fig 3.1: Data flow diagram
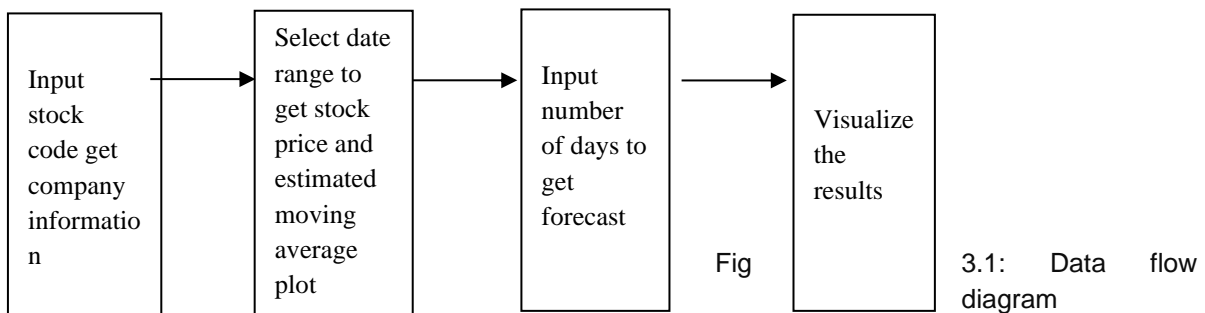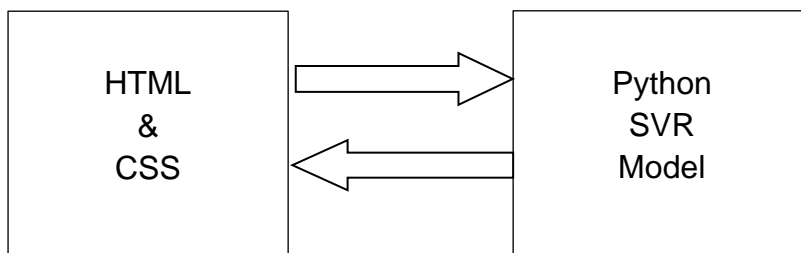
Front End   Back End

Fig 3.2: Block diagram

High Level Approach

Task 1

We are going to set up the project's environment by setting up the application's starter files/folders. Also the dependencies to be installed will be covered here.

Install necessary libraries using pip package installer.

Task 2

The basic layout of the application will be built using Dash in this task.

We mainly need two divisions (.Div) for the entire layout.

The first one is for our inputs like stock code, date range selector, number of days of forecast and buttons. These components are the ones which the user will be interacting with. They should be given appropriate IDs and class names

The second division will be for the data plots and company's basic information (name, logo, brief intro) only. Leave the divisions inside them as blank and give unique IDs because these will get updated.

Task 3

Using CSS we will style our webpage to make it look more neat and user friendly.

Enhance your webpage by exploring and improving the styling.

Task 4

We are going to use the yfinance python library to get company information (name, logo and description) and stock price history.

Dash's callback functions will be used to trigger updates based on change in inputs.

Task 5

We will build a machine learning model - Support Vector Regression (SVR) for predicting the stock prices.

4. Implementation

4.1 Algorithm

Support Vector Regression as the name suggests is a regression algorithm that supports both linear and non-linear regressions. This method works on the principle of the Support Vector Machine. SVR differs from SVM in the way that SVM is a classifier that is used for predicting discrete categorical labels while SVR is a regressor that is used for predicting continuous ordered variables. In simple regression, the idea is to minimize the error rate while in SVR the idea is to fit the error inside a certain threshold which means, work of SVR is to approximate the best value within a given margin called ε- tube.
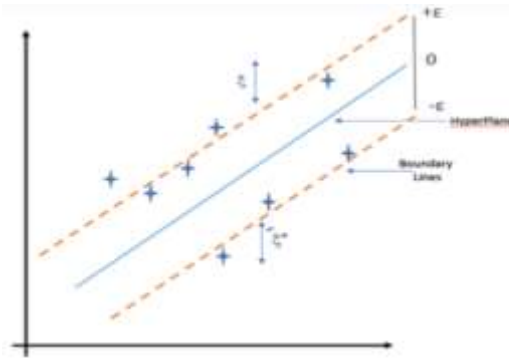 Look at the figure below to understand SVR

Fig 4.1: Support vector regression

To understand the figure you need to know the following term

1. Hyperplane: It is a separation line between two data classes in a higher dimension than the actual dimension. In SVR it is defined as the line that helps in predicting the target value.

2. Kernel: In SVR the regression is performed at a higher dimension. To do that we need a function that should map the data points into its higher dimension. This function is termed as the kernel. Type of kernel used in SVR is Sigmoidal Kernel, Polynomial Kernel, Gaussian Kernel, etc,

3. Boundary Lines: These are the two lines that are drawn around the hyperplane at a distance of ε (epsilon). It is used to create a margin between the data points.

4. Support Vector: It is the vector that is used to define the hyperplane or we can say that these are the extreme data points in the dataset which helps in defining the hyperplane. These data points lie close to the boundary.

.

Step 4:in this case, we pass 0 to the method to read from a webcam. We can just add the exact same paint interface for ease of usage

Step 5:find the contour-of-interest once we start reading the webcam ,our finger gets detected and we can start writing the content using 1 finger and eraser & choose another color using 2 fingers.

4.2 Code Implementation

Make the main website's structure using mainly Dash HTML Components and Dash Core Components.

Enhance the site's UI by styling using CSS

Generate plots of data using the plotly library of Python. The data is fetched using yfinance python library

Implement a machine learning model to predict the stock price for the dates requested by the user.

Deploy the project on Heroku to host the application live.

5. Result

The basic layout of the web page application will be as shown below in the below image

Welcome to the Stock Dash App!

Click on below link to view Stock Codes of various company

Link to view stock codes

Web Page Layout

View Stock Code

An illustrated example for Google (stock code is GOOGL) on submit you will be able to view Google info and its logo. Similarly you can find for other companies as well and here the requirement is Input which is basically a Stock Code of a company.

Input Stock Code

As soon as input is entered, it will display company info and the dynamic plotly graphs.
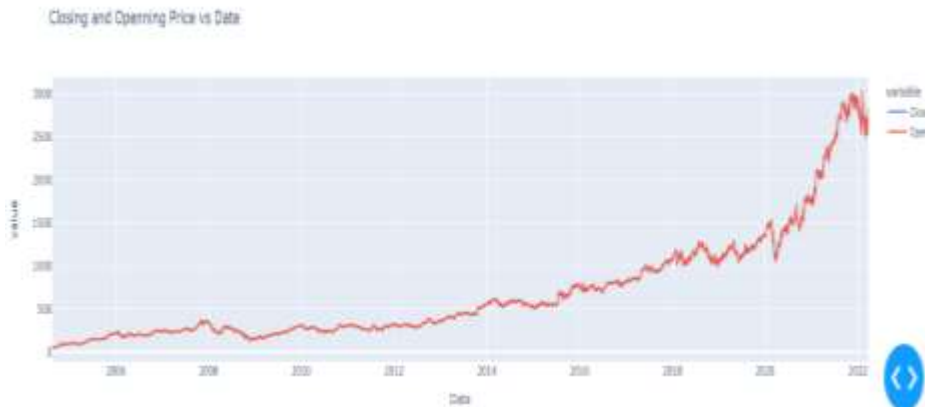Here the graph is for visualizing, based on inputs provided.
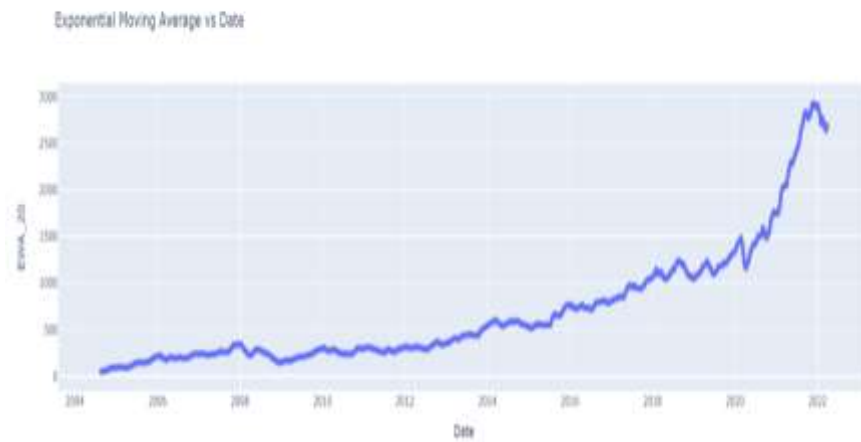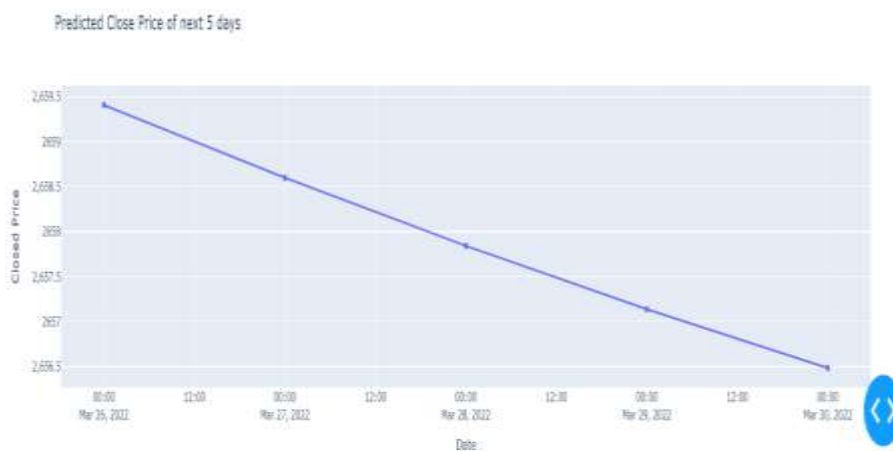
Fig 5.4 Stock Price

Fig 5.5 Indicators

Fig 5.6 Forecast for Number of Days

Fig 5.7 Start Date

6. Conclusion

We have created a single-page web application using Dash (a python framework) and some machine learning models which displays the company information (logo, registered name and description) and stock plots based on the stock code given by the user. Also the ML model will enable the user to get predicted stock prices for the date inputted by the user. The developed model and program for stock prices prediction mainly by using data from Yahoo finance. Efficient and accurate prediction systems for stock prices help traders, investors, and analyst by providing supportive information like the future direction of the stock market. We created the main website's structure using mainly Dash HTML Components and Dash Core Components and enhance the site's UI by styling using CSS. The model generate plots of data using the plotly library of Python. The data is fetched using yfinance python library. Finally, our web application can be accessed by anyone in the world.

We study that the use of Support Vector Regression to predict financial movement direction gave us better results. Support Vector Regression is a promising type of tool for financial forecasting. Support Vector Regression is superior to the other individual classification methods in forecasting daily movement direction. This is a clear message for financial forecasters and traders, which can lead to a capital gain. However, each method has its own strengths and weaknesses. In this model, the principal components identified by the Support Vector Regression are used along with internal and external financial factors in Support Vector Regression for forecasting. We also observed that the choice of the indicator function can dramatically improve /reduce the accuracy of the prediction system. Also a particular Machine Learning Algorithm might be better suited to a particular type of stock, say Technology Stocks, whereas the same algorithm might give lower accuracies while predicting some other types of Stocks, say Energy Stocks.

7. Future Scope
We have created a single-page web application using Dash (a python framework) and some machine learning models which displays the company information (logo, registered name and description) and stock plots based on the stock code given by the user. Also the ML model will enable the user to get predicted stock prices for the date inputted by the user. The developed model and program for stock prices prediction mainly by using data from Yahoo finance. Efficient and accurate prediction systems for stock prices help traders, investors, and analyst by providing supportive information like the future direction of the stock market. We created the main website's structure using mainly Dash HTML

Components and Dash Core Components and enhance the site's UI by styling using CSS. The model generate plots of data using the plotly library of Python. The data is fetched using yfinance python library. Finally, our web application can be accessed by anyone in the world.

We study that the use of Support Vector Regression to predict financial movement direction gave us better results. Support Vector Regression is a promising type of tool for financial forecasting. Support Vector Regression is superior to the other individual classification methods in forecasting daily movement direction. This is a clear message for financial forecasters and traders, which can lead to a capital gain. However, each method has its own strengths and weaknesses. In this model, the principal components identified by the Support Vector Regression are used along with internal and external financial factors in Support Vector Regression for forecasting. We also observed that the choice of the indicator function can dramatically improve /reduce the accuracy of the prediction system. Also a particular Machine Learning Algorithm might be better suited to a particular type of stock, say Technology Stocks, whereas the same algorithm might give lower accuracies while predicting some other types of Stocks, say Energy Stocks.

8. References

1. Achyut Ghosh, Soumik Bose1, GiridharMaji, Narayan C. Debnath, Soumya Sen-Stock Price Prediction Using LSTM on Indian Share Market

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

3. Asset Durmagambetov currently works at the mathematics, CNTFI- 'The Riemann HypothesisMillennium Prize Problems' - stock market predictions.

4. 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.

5. 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.

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

7. 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.

8. Lavanya Ra SRM Institute of Science and Technology | SRM · Department of Computer Science - Stock Market Prediction.

9. Kishor Kumar Reddy C and Vijaya Babu B, "ISLIQ-OC: Improved Supervised Learning in Quest using Optimal k-means Clustering Mechanism to Nowcast Snow/No-Snow", International Journal of Control Theory and Applications , 2017.

10. (http://serialsjournals.com/abstract/61260_09-c._kishor_kumar_reddy.pdf)

11. R Gangadhar Reddy, M. Srinivasa Reddy, P R Anisha, Kishor Kumar Reddy C, "Identification OF EARTHQUAKES USING WAVELET TRANSFORM AND CLUSTERING METHODOLOGIES", International Journal of Civil Engineering and Technology, 2017. (http://www.iaeme.com/MasterAdmin/UploadFolder/IJCIET_08_08_067/IJCIET_08_08_067.pdf)

12. B Subbarayudu, Srija Harshika D, E Amareswar, R Gangadhar Reddy, Kishor Kumar Reddy C, "Review and Comparison on Software Process Models", International Journal of Mechanical Engineering and Technology, 2017.

13. (http://www.iaeme.com/MasterAdmin/UploadFolder/IJMET_08_08_105/IJMET_08_08_105.pdf)

14. Kishor Kumar Reddy C and Vijaya Babu B, "ISLGAS: Improved Supervised Learning in Quest using Gain Ratio as Attribute Selection Measure", International Journal of Control Theory and Applications, 2016.

15. Kishor Kumar Reddy C and Vijaya Babu B "A Survey on Issues of Decision Tree and Non-Decision Tree Algorithms", International Journal of Artificial Intelligence and Applications for Smart Devices, 2016.

16. Kishor Kumar Reddy C, Rupa C H and Vijaya Babu B, "ISLIQ: Improved Supervised Learning in Quest to Nowcast the presence of Snow/No-Snow", WSEAS Transactions on Computers, 2016.

17. (http://www.wseas.org/multimedia/journals/computers/2016/a055705-872.pdf)

18. Kishor Kumar Reddy C and Vijaya Babu B, "ISPM: Improved Snow Prediction Model to Nowcast the Presence of Snow/No-Snow", International Review on Computers and Software, 2015.

19. (http://www.praiseworthyprize.org/jsm/index.php?journal=irecos&page=article&op=view&path%5B%5D=17055)