

EED 301- APPLIED MACHINE LEARNING PROJECT (Phase 1) - SP2022

HARDIK GABA
1910110154

MEETENDRA SINGH
1910110232

NAMAN RASTOGI
1910110250

AUTOMATED INVESTMENT ADVISOR

Task

To identify market changes earlier than what is possible with the traditional methods to help people invest in the stock market.

What is the Stock Market?

The concept of stocks has existed since the 1600s. The first stock exchange was established in Amsterdam in 1602. Since then the strategies for putting your money in a company have evolved. Let's just say that with evolving technology people have shifted towards technical analysis rather than qualitative analysis but people still practice both. Stock market is characterized as dynamic, unpredictable and non-linear in nature. Predicting stock prices is a challenging task as it depends on various factors. In simple terms stocks are pieces of a company for which to own you need to give the company some 'x' amount of money. But this 'x' is where predicting stock value gets tricky, the value of a stock is a forever changing variable and it is difficult or rather impossible to predict an accurate stock price but we can still hit a close shot near the accurate value of a stock with some prediction methods.

Importance of Stock Market

- Helps Individuals and Companies to raise capital.
- The stock market plays an important role in the economy of a country in terms of spending and investment.
- Market serves as an indicator of the state of the economy
- Stock markets promote investment. The raising of capital allows companies to grow their businesses, expand operations and create jobs in the economy.

Ways of Predicting a stock price:

Technical Analysis: based on the premise that the future behavior of a financial time series is conditioned to its own past, thus technical indicators could help traders to interpret this behavior in their favor

Fundamental Analysis: based on company's financial reports and external information as political and economic factors. This information is taken from unstructured data as news articles, financial reports or even publishing in microblogs by analysts.

Financial Time series and Textual Data: uses different statistical techniques and artificial intelligence models to make a prediction based only on technical information

How does ML help in predicting Stock prices?

The internet says that ML was introduced in 1959 by Arthur Samuel. Machine learning is basically educating a machine to do a specific task and then putting it to use to help humans. Nowadays, advanced intelligent techniques based on either technical or fundamental analysis are used for predicting stock prices. Particularly, for stock market analysis, the data size is huge and also non-linear. To deal with this variety of data an efficient model is needed that can identify the hidden patterns and complex relations in this large data set. In the concept of prediction with Machine Learning what we do is teach a machine with some test cases of which we know the results and then use it to develop a system inside the machine to predict the future possibilities with a new input data. Hence we need to train the machine as best as possible for it to be as accurate as it could be and for that there are few methods that exist.

Types of Methods

While doing our survey for the project, we stumbled upon various methods to perform the task. They are as follows:

- Regression
- Random Walk Theory
- Moving Average Convergence / Divergence
- Autoregressive Moving Average
- Autoregressive Integrated Moving Average
- Long Short Term Memory (Neural Networks)

Model to be used

Keeping in mind the scope of this course, we will be restricting ourselves from using Deep Learning and focus mainly on Machine Learning to make our model. We will perform **regression analysis** using **Support Vector Machines** (SVMs) for our project.

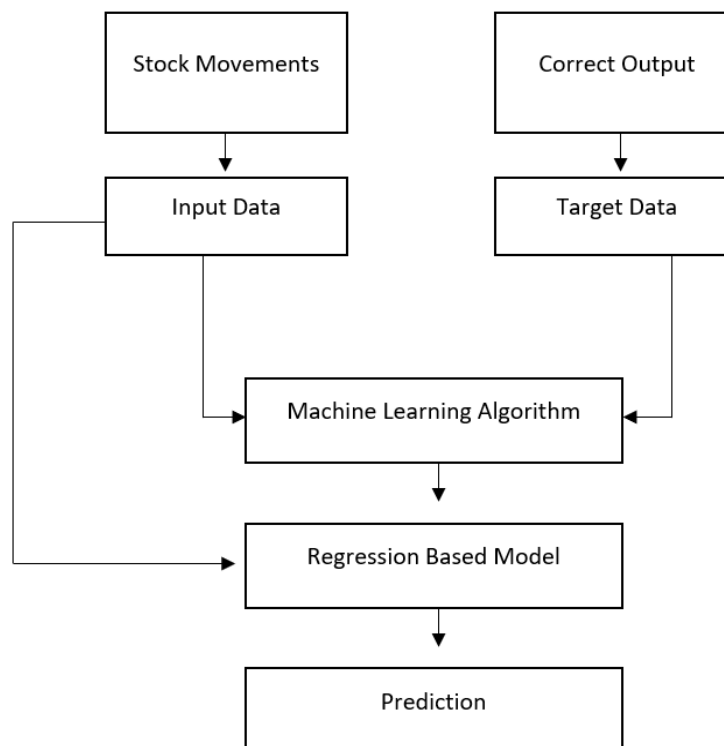
Our model will cater mainly to Indian investors. It will use various machine learning tools to forecast stock prices using the data from the past. We will select a few key features that determine the value of a stock like closing price, Volume, S&P close, NASDAQ index, INR to USD conversion rate, oil price, etc.

We will take past data about stocks of various companies whose closing price we already know, and feed it to our algorithm for training it. After this, we will test our model by giving only the input variables to the algorithm and compare the predictions with the known outputs. Once satisfactory results are obtained, we will use the final model constructed in real life and tabulate the results.

The data would be divided into 3 categories:

- Training
- Validation
- Testing.

We will be using k-fold cross validation technique to finalize our model.



Flow Chart for Regression Based Model

Data description

Data for the project will be collected from various sources, some of which are stated below:

Stocks	S&P, NASDAQ, Yahoo Finance, BSE, NSE
Currency	INR, USD
Commodity	Oil, Gold, Silver

* The features stated above are not definitive. They may be changed.

We will be taking data for the past ten years.

Since the markets are closed on holidays, we'll perform data alignment and fill in the missing data by linear interpolation.

The SVM algorithm is very sensitive to the size of training data. When the size of the training set is not large enough, the hyper-plane found by the SVM algorithm might not be able to split the data properly. Thus, feature selection is essential.

The performance of a stock market predictor heavily depends on the correlation between the data used for training and the current input for prediction. Intuitively, if the trend of stock price is always an extension to yesterday, the accuracy of prediction should be fairly high. To select input features with high temporal correlation, autocorrelation and cross-correlation of different market trends will be performed.

Literature survey and Key Points

1. Applied Machine Learning - M.Gopal
 - Basic concepts about the field of Machine Learning
 - How to train and test a ML model (techniques like k-fold cross validation)
 - Initial vision about how to go about tackling the problem at hand
2. Varsity by Zerodha
 - <https://zerodha.com/varsity/module/introduction-to-stock-markets/>
 - Basics of stock market
 - Importance of markets
 - What do different indices mean
 - Jargon of the field

3. Survey of Stock Market Prediction Using Machine Learning Approach | Ashish Sharma, Dinesh Bhuriya and Upendra Singh
<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8212715>
 - Importance of predicting stock prices
 - Various types of regression models - Polynomial, RBF, Sigmoid, Linear
4. Stock Market Forecasting Using Machine Learning Algorithms | Shunrong Shen, Haomiao Jiang and Tongda Zhang
<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.278.6139&rep=rep1&type=pdf>
 - External perturbations to the financial markets are no longer domestic. Globalization deepens the interaction between the financial markets around the world
 - How and where to collect the data from
 - Differences between SWM and MART that finally helped us conclude to choose the former.
 - How to process the data before feeding it to the algorithm (data alignment)
 - How to correctly test our model and tabulate the results
5. Stock Closing Price Prediction using Machine Learning Technique | Mehar Vijha , Deeksha Chandolab, Vinay Anand Tikkiwalb, Arun Kumar
<https://www.sciencedirect.com/science/article/pii/S1877050920307924>
 - What is the significance of stock
 - Machine learning methods
 - Stock price prediction and its accuracy
6. Stock Market Prediction Using Machine Learning | Ishita Parmar, Navanshu Agarwal, Sheirsh Saxena, et al
<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8703332>
 - How to form a proper model (flow of training, validation and testing)
 - Differences between Regression based model and Long Short Term Memory (LSTM) Network Based Model