

Stock Market Prediction

Asiya Salim¹ and Shyma Kareem²

Student, Department of Computer Applications¹

Assistant Professor, Department of Computer Applications²

Musaliar College of Engineering and Technology, Pathanamthitta, Kerala, India

Abstract: *The aim is to create software that analyzes previous stock data of certain companies, with help of certain parameters that affect stock value. I am going to implement these values in data mining algorithms and we will be able to decide which algorithm gives the best result. This will also help me to determine the values that particular stock will have in near future. I will determine the patterns in data with help of Machine learning techniques like support vector and linear regression.*

Keywords: Machine Learning (ML), Support Vector Regression (SVR), Linear Regression (LR), Support Vector Machine (SVM).

I. INTRODUCTION

Stock market prediction is the act of attempting to anticipate the future value of a stock. In the world of finance, stock trading is one of the most significant activities. The majority of stock brokers employ technical, fundamental, or time series analysis when making stock predictions. The trend in a stock market prediction is not a new thing and yet this issue is kept being discussed by various organizations. Before purchasing a stock, investors can do one of two types of stock analyses. The first is a fundamental analysis in which they examine the stock's intrinsic worth and performance of the industry, economy, political climate etc. to decide that whether to invest or not. On the another side, the technical analysis it is an evolution of stocks by the means of studying the market-generated statistics like historical pricing and volume. Here, the Machine Learning (ML) method is used to foretell a specific company's future trend, that will be trained from the available stocks data and gain intelligence and then uses the acquired knowledge for an accurate prediction. Machine learning technique used here is called Support Vector Regression (SVR) and Linear Regression to predict stock prices for the large and small capitalizations, employing prices with both daily and monthly time series. The scope of the work assists the economist for analyzing the future of trading sector and giving the support in the decision making in stock market. The predict process reduce the risk in the stock investment. The predicted results assist to manage the company resources and walk on the way to achieve the successful business.

II. PROPOSED SYSTEM

In the proposed system, future value can be found out by using the four prices that are recorded during the day. The four prices which includes open, low, high and close. Open: this is the first price at which a trade gets executed when markets open in the morning. High: This is the highest price at which a trade could be executed during the trading day. Low: this is the lowest price at which a trade could be executed during the trading day. Close: This is the last price at which the markets got closed. If the open price is lower than the close price, the day is considered to be a positive day and if the close price is lower than the open price, the day is considered to be a negative one. In this system Linear Regression and SVR is used for predicting the future trend of the company. After performing the Linear Regression, R^2 value is calculated and based on this value and a particular date which the user applied is used to forecasting the prediction module.

III. LITERATURE SURVEY

T. Mankar proposes a Machine learning and artificial intelligence techniques that are being used in conjunction with data mining to solve the real world problems. These methods have shown to be quite efficient, producing maximum accuracy with little financial outlay and also saving a significant amount of time.

The stock exchange is a highly composite system in addition often concealed with secrecy, consequently, it is highly challenging to consider all the relevant elements before making a choice. K. S. Umadevi created a stock market prediction model that takes into account many stock-specific factors. After getting the stock scores, analysis is done. This analysis involves visualization of stock scores in the form of various plots and prediction of the scores using a time series model known as ARIMA (auto regressive moving average). The findings demonstrate that the time series model successfully predicted market scores with a high degree of accuracy.

H. L. Siew suggested a regression method for forecasting stock price trends utilizing transformed data in an ordinal data format. The original pre-transformed data source includes data of several heterogeneous data kinds that are used to handle financial ratios and exchange rates. The forms of the data for financial ratios and currency values offer a method for calculating stock prices. Only a standardized ordinal data type, which offers a method to rank stock price trends, is included in the modified data set.

R. A. Kamble proposed a method that is designed to predict price trends in the stock market. The first goal is to use oscillators and indicators, such as the Moving Average Convergence Divergence (MACD), Relative Strength Index (RSI), Stochastic Oscillator (KDJ), and Bollinger Band (BB), to improve short-term stock price trend prediction. It is feasible to increase the short-term trend forecast accuracy rate by applying the right pre-processing method and machine learning model.

N. Powell proposes a method which is based on classification techniques like k-means clustering. Since forecasts of future prices, indices, volumes, and a variety of other values are frequently incorporated into the process of making economic decisions, stock forecasting is a crucial component. Although there are many different approaches which includes unsupervised classification techniques such as k-means clustering

IV. STEPS FOR STOCK VALUE PREDICTION

Step 1: This step is important for the download data from the net. We are predicting the financial market value of any stock. So that the share value up to the closing date are download from the site.

Step 2: In the next step the data value of any stock that can be converted into the CSV file(Comma Separate Value) so that it will easily load into the algorithm.

Step 3: In the next step in which GUI is open and when we click on the selected market button it will show the window from which we select the stock dataset value file.

Step 4: After selecting the stock dataset file from the folder it will show graph Stock before mapping and stock after mapping.

Step 5: The next step algorithm calculated the RMSE value and R² value for minimizing error. So, it will predict the graph for the dataset value efficiently.

Step 6: In final step algorithm display the predicted value graph of select stock which shows the original value and predicted value of the stock.

V. TECHNOLOGIES FOR PREDICTING STOCK DATA

For predicting the stock market, Machine learning techniques like SVR and Linear Regression is used.

Linear Regression

The most basic machine learning algorithm that can be implemented on this data is linear regression. A link between the independent variables (x) and the dependent variable (y) is established by the linear regression model's output equation.

The equation for linear regression can be expressed as:

$$Y = \theta_1 X_1 + \theta_2 X_2 + \dots + \theta_n X_n$$

Here, x_1, x_2, \dots, x_n represent the independent variables while the coefficients $\theta_1, \theta_2, \dots, \theta_n$ represent the weights.

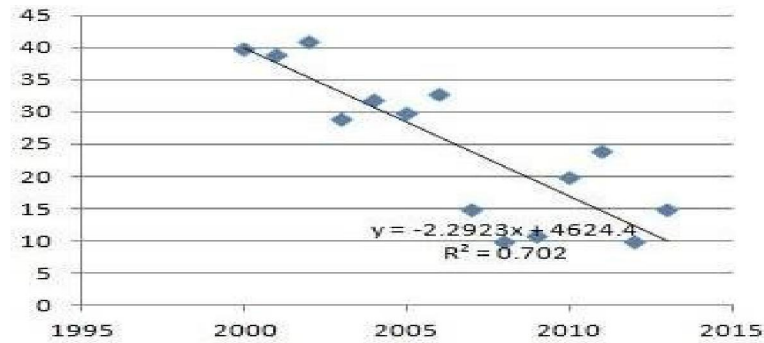


Figure 1: LR Graph

Compared to stock trend, the exact increment in stock index may provide more information for investment strategy. This means the classification problem now evolves to a regression problem. For finding the performance of this model, square root of mean square error (RMSE) is used as a criteria, which is defined as The variance of the residuals (prediction errors). The distance between the data points and the regression line is measured by residuals, and the spread of these residuals is measured by RMSE. In other words, it tells you how concentrated the data is around the line of best fit. The formula is:

$$RMSE = \sqrt{\frac{\sum (f - o)^2}{N}}$$

Where,

f = forecasts (expected values or unknown results),

o = observed values (known results).

The bar above the squared differences is the mean (similar to \bar{x}). The same formula can be written with the following,

By linear regression R^2 value is calculated for the next step for forecasting the prediction module.

We now want to conduct an experiment, which entails an iterative procedure that linearizes my data by minimizing a cost function.

$$RMSE_{fo} = \left[\sum_{i=1}^N (z_{fi} - z_{oi})^2 / N \right]^{1/2}$$

Support Vector Regression (SVR)

Support Vector Machine can be used as a regression method, maintaining all the main features that characterize the algorithm (maximal margin). The Support Vector Regression (SVR) for classification adheres to the same concepts as the SVM. The information at hand, which has limitless possibilities, becomes incredibly difficult to forecast when the output is a real number. In the case of regression, a tolerance margin (epsilon) is established roughly in line with what the SVM would have already asked for from the problem. But besides this fact, there is a problem that the algorithm requires more effort. As a result, it should customize the hyperplane to maximize the margin and then calculate the minimize error that is utilized to fit the point on the graph.

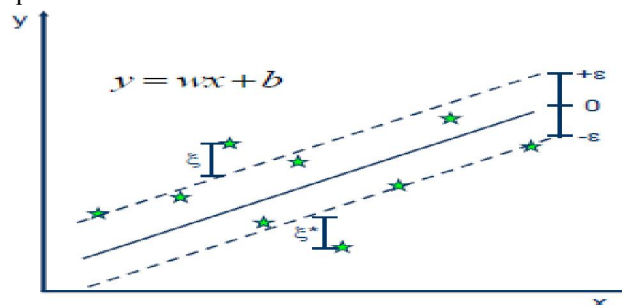


Figure 2: SVR Graph

Linear SVR

$$y = \sum_{i=1}^N (\alpha_i - \alpha_i^*) \cdot \langle x_i, x \rangle + b$$

Non-Linear SVR

To enable perfolinear separation, the kernel functions convert the data into a higher dimensional feature space.

$$y = \sum_{i=1}^N (\alpha_i - \alpha_i^*) \cdot \langle \varphi(x_i), \varphi(x) \rangle + b$$

VI. CONCLUSION

Stock Market Prediction, a system for predicting the future trend of a particular company. Here, the data is collected from different global financial markets with machine learning algorithms in order to predict the stock index movements which contains the different attributes. Selecting algorithm is mainly depend on the dataset therefore SVM algorithm is taken here, it works based on the large dataset value ie (historical data) which contains bundle of data's which is collected from internet. Also, SVM does not give a problem of over fitting. Various machine learning based models are proposed for predicting the daily trend of Market stocks. Numerical results suggest the high efficiency. Compared to other machine learning methods, the model yields better profit. Support vector regression and linear regression are very good at predicting Market Fluctuations if trained with sufficient amount of data. The results are very accurate and the error rate is minimal. The proposed system with SVR based prediction method the system's error rate is close to not more than 10% which implies a very high 90 % prediction accuracy.

REFERENCES

- [1]. M. Usmani, S. H. Adil, K. Raza, and S. S. A. Ali, "Stock market prediction using machine learning techniques," 3rd International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, 2016, pp. 322-327. doi: 10.1109/ICCOINS.2016.7783235
- [2]. T. Mankar, T. Hotchandani, M. Madhwani, A. Chidrawar and C. S. Lifna, "Stock Market Prediction based on Social Sentiments using Machine Learning", 2018 International Conference on Smart City and Emerging Technology (ICSCET), Mumbai, 2018, pp. 1-3. doi: 10.1109/ICSCET.2018.8537242
- [3]. P. Somani, S. Talele, and S. Sawant, "Stock market prediction using Hidden Markov Model," 2014 IEEE 7th Joint International Conference on Information Technology and Artificial Intelligence, Chongqing, 2014, pp. 89-92. doi: 10.1109/ITAIC.2014.7065011
- [4]. K. Raza, "Prediction of Stock Market Performance by Using Machine Learning Techniques," 2017 International Conference on Innovations. in Electrical Engineering and Computational Technologies (ICIEECT), Karachi, 2017, pp. 1-1. doi: 10.1109/ICIEECT.2017.7916583
- [5]. M. Usmani, M. Ebrahim, S. H. Adil and K. Raza, "Predicting Market Performance using a Hybrid Model", 3rd International Conference on Emerging Trends in Engineering, Sciences, and Technology, 2018, (ICEEST), Karachi, Pakistan, 2018, pp. 1-4. doi: 10.1109/ICEEST.2018.8643327