PREDICTION OF STOCK MARKET

ABSTRACT:

Prediction of stock/index price accurately is essential for market investors to maximize their profits and avoid losses. Support vector machine (SVM) is the model which is used to predict stock market daily trends i.e. ups and downs. Support vector regression (SVR) has helped in many ways in prediction problems in the stock market. This paper proposes a framework of Support Vector Machine (SVM) and Natural language processing (NLP) with approach for stock market prediction. The main objective of this paper is to get the depth knowledge in the stock market using these models and tools such as predicting closing price, opening price of the stock market for the available data. This model uses SVM with different functions to predict profit or loss, and the output of SVM helps to get the accuracies from the training set to predict future of stock value in the future.

Keywords :Support Vector Machines, Natural language processing, Regression, Stock Market

INTRODUCTION:

Stock market prediction is one of the important exertions in business and finance. There are many uncertainties involved in the movement of the market. There are many factors that affect the stock market including political events, general economic conditions. This leads to predict the stocks in the market. Stock market prediction determines the company’s future by taking the historical data into consideration, make predictions of the crashes that’s going to occur and minimize the investor’s losses. It majorly includes the company performance in the recent years based on factors such as political consequences and climate changes. Toronto stock exchange is one of the major stock exchanges in the world and deals with many company equities, major oil and gas company equities. It has a valued market capital of 3.5 trillion.

PROBLEM STATEMENT:

Stock market prediction determines the company’s future by taking the historical data into consideration, make predictions of the crashes that’s going to occur and minimize the investor’s losses. By observing the past data, we predict the rise and fall of the stocks and help the investors to know where to invest. The main goal of our project is to study the data and apply different machine learning models on our data to get the outcome. The outcomes will be useful to the investors for the future investment . There are a number of further directions can be investigated starting from this project. The first one is to explore other creative and effective methods that might yield even better performance on stock market forecasting.by using this machine learning model we can maximize the profits and minimize the loses. This gives the idea to future investors to invest in which equities.

DATASET:

Dataset of our project,” Stock market prediction” is extracted from NASDAQ (National Association of Securities Dealers Automated Quotations) website. Our dataset consists of information about 10 different companies’ stocks. The data is of last ten years. We can extract the data from this website till the current date. Dataset consists of Volume column which explains the number of shares that changed hands during a given day. Open, close columns explain starting and ending period of price at which a stock trades upon the exchange on a given day. High, low columns in the context of stocks means the highest and lowest prices of a stock recorded on a given day. The column company indicates the name of the organization. Dataset consists of 24000 rows and 7 columns.

FLOWCHART:

ETHICAL PRINCIPLES:

For any project, before proceeding, the ethical principles or guidelines are to be followed to make the project successful in ethical manner.

Consent:

The dataset we choose is open data which is openly available and easily accessible by anyone. It doesn’t consist any personal information, so it doesn’t affect anyone’s privacy.

Clarity:

Data will be used to analyze ups and downs in given stocks and forecast the crashes.

Consistency:

Our data is openly accessible by public therefore, there is no chance for consistency issues.

Control:

The data we extracted is from NASDQ website, all the rights are acquired by the organization and they have the complete control to access the data. If any of the data in our dataset will be restricted in the future, we can always remove it from our dataset.

Consequences:

As our data doesn’t involve any individual’s personal information, it is unlikely less harmful to the public. The data will not be shared to any third-party vendors without user consent.

REFERENCES:

Stock Prediction Using Twitter Sentiment Analysis (2020). Retrieved 24 February 2020, from <http://cs229.stanford.edu/proj2011/GoelMittalStockMarketPredictionUsingTwitterSentimentAnalysis.pdf>

Hybrid Machine Learning System for Stock Market Forecasting <https://www.researchgate.net/profile/Kumkum_Garg2/publication/238747905_A_Hybrid_Machine_Learning_System_for_Stock_Market_Forecasting/links/00b7d53b4cfc215d81000000/A-Hybrid-Machine-Learning-System-for-Stock-Market-Forecasting.pdf>

Stock Market Prediction Using Artificial Neural Networks (2020). Retrieved 24 February 2020,from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.134.1114&rep=rep1&type=pdf>

Equity forecast: Predicting long term stock price movement using machine learning (2020). Retrieved 24 February 2020, from <https://arxiv.org/ftp/arxiv/papers/1603/1603.00751.pdf>

Improving Long Term Stock Market Prediction with Text Analysis (2020). Retrieved 24 February 2020, from <https://ir.lib.uwo.ca/cgi/viewcontent.cgi?article=6267&context=etd>