ABHISHEK SUMAN - 20UCS007
HARSH AGARWAL - 20UCS073

ML PROJECT

RESEARCH PAPER

STOCK MARKET PREDICTION USING LINEAR REGRESSION AND SVM

OVERVIEW

INTRODUCTION

- The main reason behind using machine learning algorithms is to reduce the risks of tasks that have failed in the past to not occur in the future using the past information.
- In case of stocks data is very dynamic and depends on a lot of factors. There are a lot of machine learning algorithms to choose from.
- The machine learning algorithm used here are linear regression and SVM.

DATASET

For Stock Market Prediction, Web Scrapping technique was used in the research paper to get the data.

Python Module "A beautiful Soup" was used as tool for the same.

Fig. 1: Amazon stock plot

RSI Indicator

Relative Strength Index is momentum indicator which which keeps oscillating with with the market.

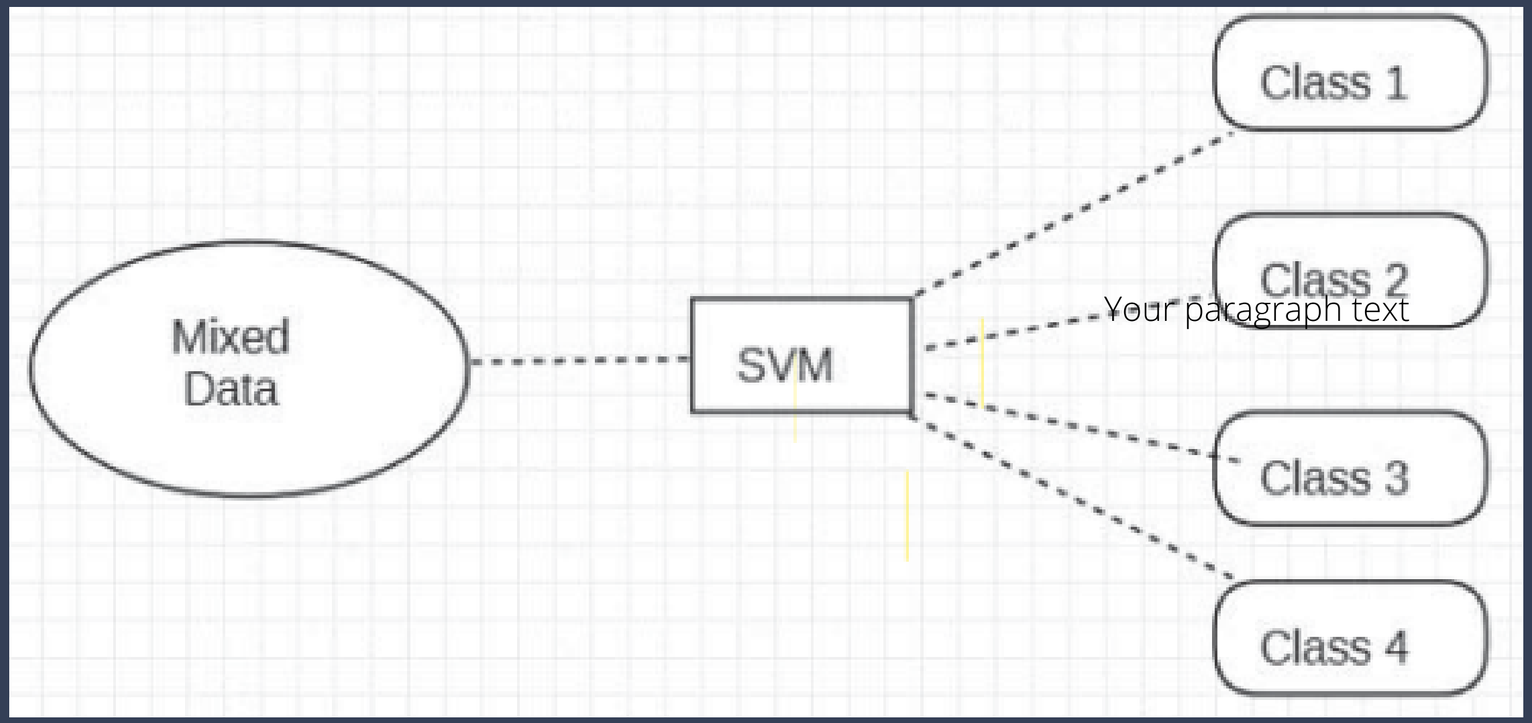
Overbought to Oversold ratio is (70+) : (30-)

RSI Indicator

Relative Strength Index is momentum indicator which which keeps oscillating with with the market.

Overbought to Oversold ratio is (70+) : (30-)

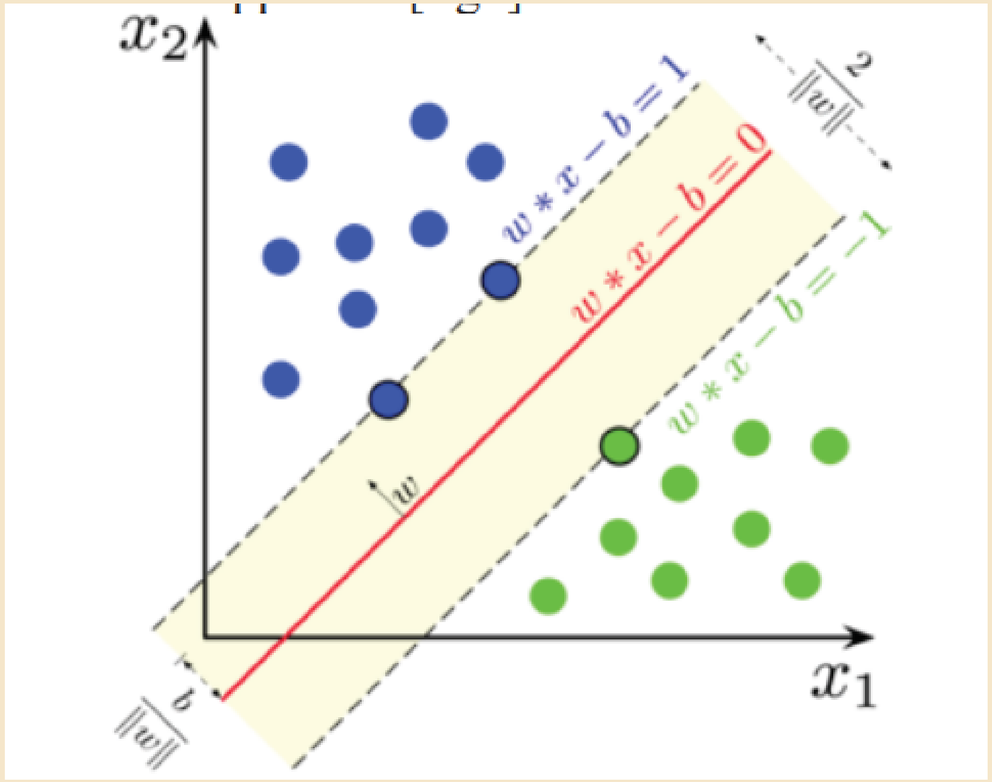
SVM



Your paragraph text

SVM Categorization

- SVM is primarily based on statistical learning approach.
- In SVM, we predict results, to find out a plane of seperation for any given input.



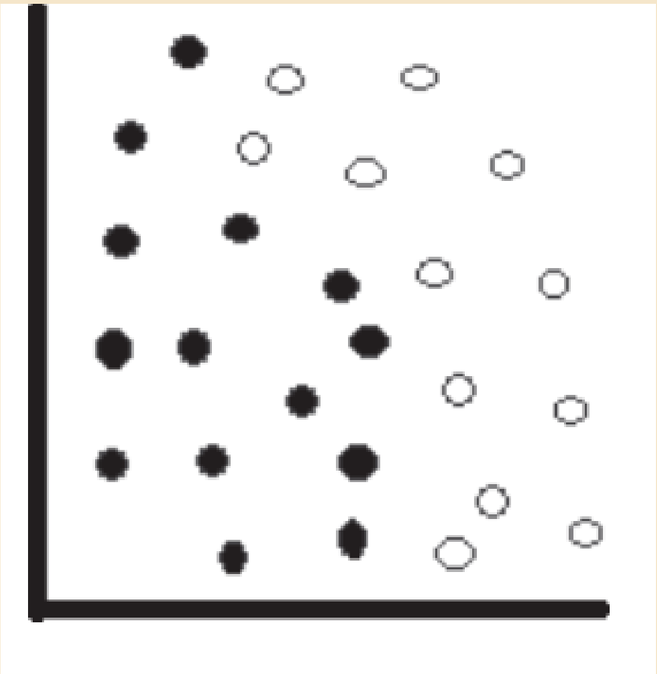
SVM ARCHITECTURE

- In SVM we analyze the two classes for dividing them with a hyperplane, it is a line that splits the data of the two classes such that value between the last points of the classes is maximum it is called maximum margin.

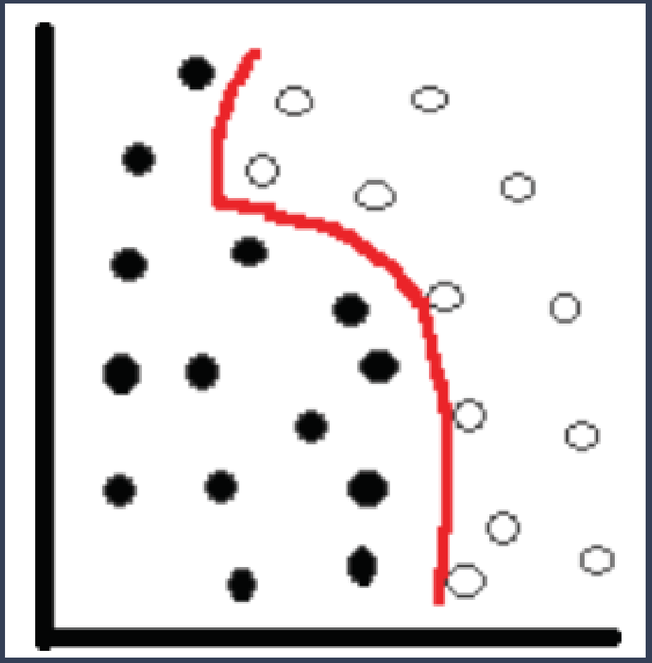
- w = vectors x= variables
- b- biased term(wo)
- w*x-b=1
- w*x-b=0
- w*x+b=1

HOW DOES SVM WORK ?

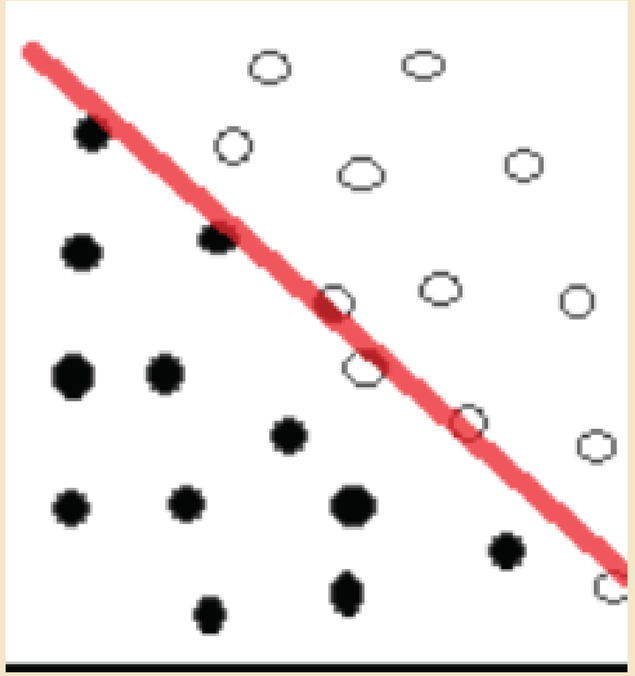
TWO CLASSES OF DATA



IMAGINARY SEPERATOR



SVM HYPERPLANE



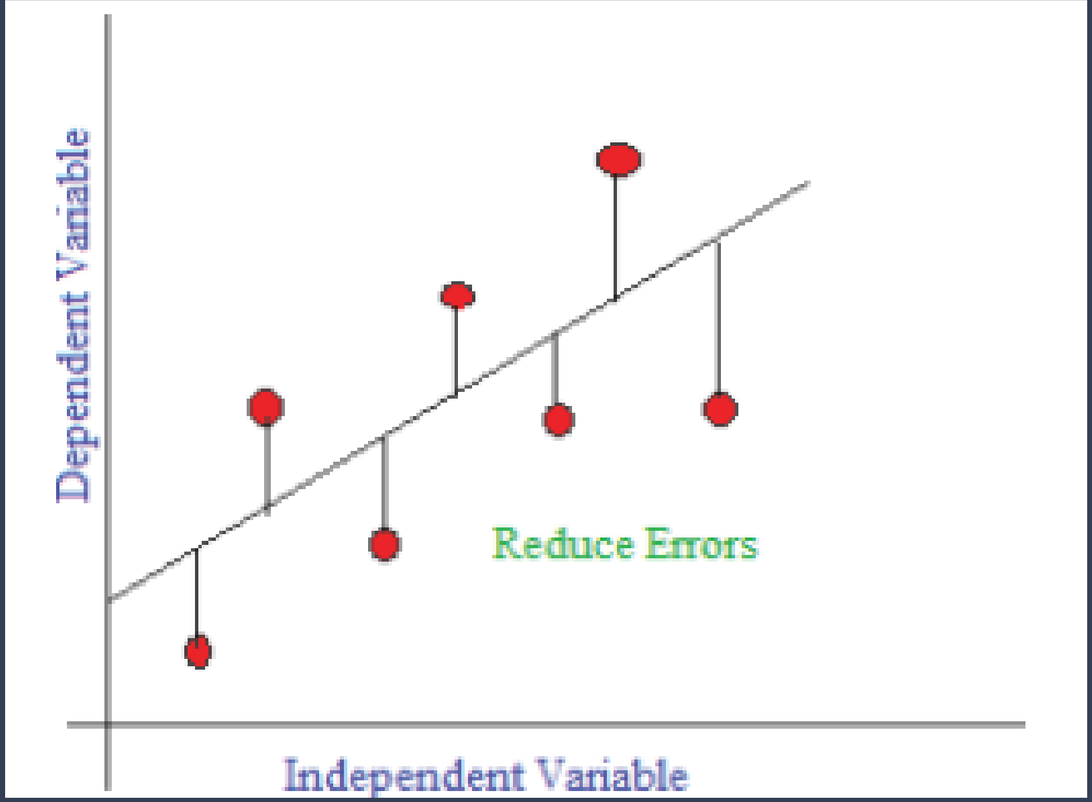
SOME APPLICATIONS OF SVM

- CREDIT CARD FRAUD DETECTION
- RECOGNIZE HANDWRITTEN ALBHABETS OR DIGITS
- PREDICTION OF STOCK PRICES
- EMAIL CLASSIFICATION

PROBLEMS IN SVM

- IT DOES NOT WORK WELL IF DATA IS NOISY AND UNLABELLED.
- SVM DOES NOT WORK WELL WITH VERY LARGE DATASETS.

LINEAR REGRESSION



- Linear Regression is a model that displays the relationship between the dependent and independent variables. Also, the variations involved in the dependent variable on changes being observed in the independent variable.

- Linear Regression is also used for predictive analysis, predict the needs of the customer based on its shopping list.

LINEAR REGRESSION ARCHITECTURE

y=B0+B1 *X1

- In the above equation, y is the dependent variable x is the independent variable bo is constant bl is coefficient.

- HOW DOES LINEAR REGRESSION WORK ?

  See graph on the left, a simple example.

  Salary = B0 + B1*EXPECTATIONS

SOME APPLICATIONS OF LINEAR REGRESSION

- PREDICT WEATHER
- ACADEMIC REPORT
- FORECASTING RESULT
- SPORTS ANALYSIS

RESULT

On analyzing the prediction results of SVR and Linear Regression the Amazon stock data Linear Regression aligned to an accuracy of 98.76% and SVR accuracy is 94.32%. So, it can be inferred that in stock market prediction Linear Regression is better than the State Vector Regression

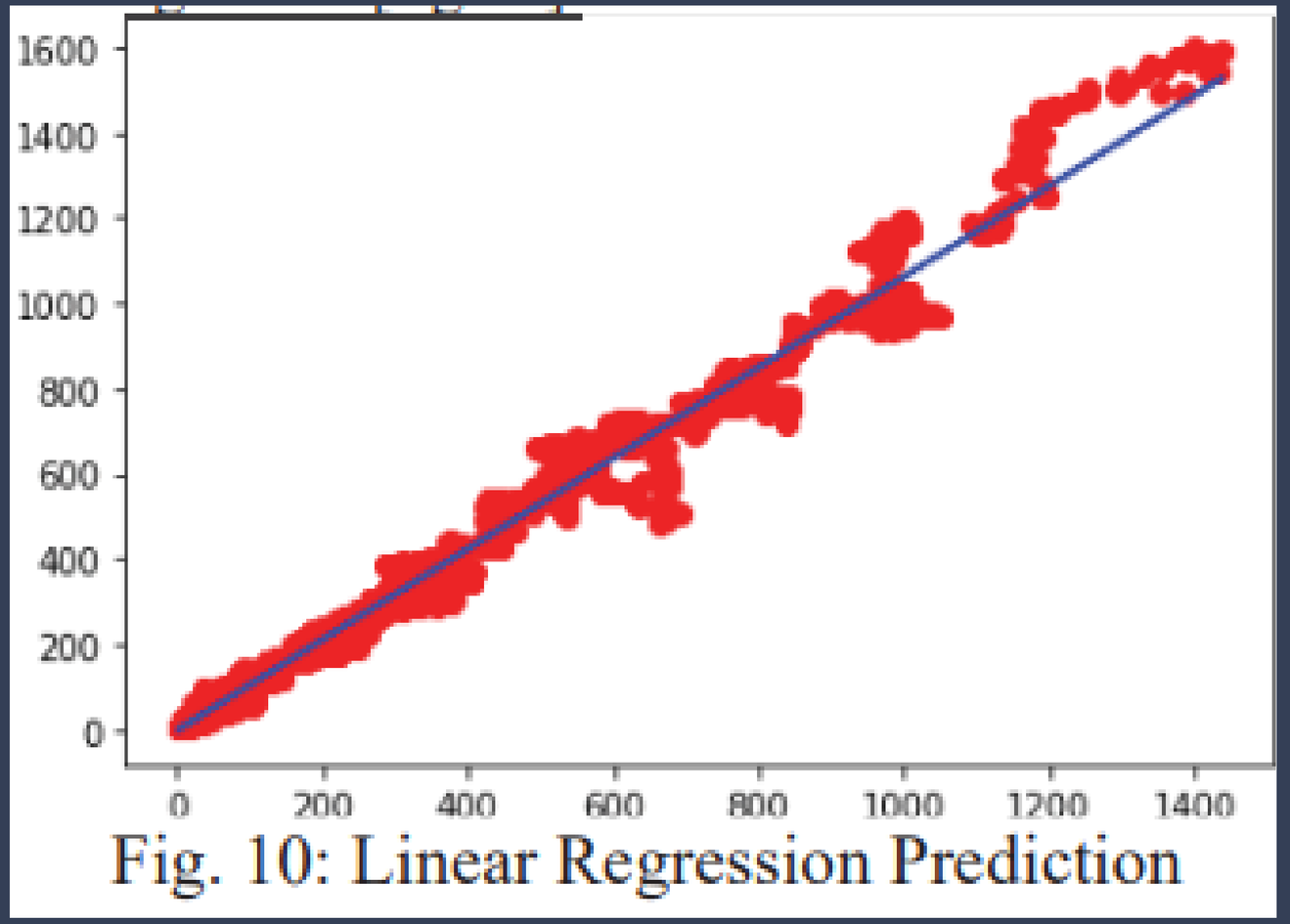


Fig. 10: Linear Regression Prediction

CONCLUSION

On observing the accuracy of both the algorithms on same the stock data it was observed linear regression predicted more accuracy as compared to the SVM model. Thus, it is proved that Linear Regression for stock data or stock market analysis is better than the SVM for the same.

LIMITATIONS

In the Research Paper, only SVM and LR classification techniques were provided.

Classification techniques like KNN were not included, although they are not optimal.

For all the different classification techniques, we cannot plot a 3-D model or with more features.

We needed stick with any two parameters, in our case, those were "High" & "Open - Close"

LIMITATIONS

Data quality and availability: The accuracy and reliability of stock market predictions heavily depend on the quality and availability of historical data. If the paper uses limited or insufficient data, it may affect the reliability of the results.

In the proposed dataset, we tried to overcome this limitation. Also created few intermediate parameters.

FUTURE WORK

Incorporating **textual and sentiment analysis** into prediction models is indeed a common practice in various fields, including finance and stock market prediction. By analyzing textual data from sources such as financial news articles, social media sentiment, or other textual sources related to the market, one can **capture market sentiment** and potentially **enhance the accuracy** of predictions. This approach leverages natural language processing (NLP) techniques to extract valuable information and sentiment from textual data

FUTURE WORK

Alternative machine learning algorithms: Besides linear regression and SVM, exploring other machine learning algorithms, such as -

1. Random forests
2. Gradient boosting machines (GBMs)
3. Deep neural networks

may provide insights into their comparative performance and suitability for stock market prediction tasks.

THANK YOU