

STOCK MARKET ANALYZING AND PREDICTION USING MACHINE LEARNING TECHNIQUES

Vandana Navale¹, Priya Bhalchim², Akanksha Salunke³, Prathamesh Kadale⁴, Shridhan Badve⁵

¹Computer, AISSMS COE, No1, Kennedy Road, Pune, 411001, Maharashtra, India.

-----***-----

Abstract — In this research, we examine current approaches and then suggest fresh approaches for predicting the stock market. We tackle the issue from three separate angles: fundamental analysis, technical analysis, and machine learning application. The weak form of the Efficient Market Hypothesis is supported by research showing that historical prices do not contain relevant information but that out-of-sample data may be predictive. News that is pertinent to any company listed on the stock market has an impact on future stock movement. We demonstrate how machine learning and fundamental analysis could be used to help investors make decisions. Techniques for machine learning come in handy here. Intelligent investors can use machine learning techniques to predict the behaviour of the stock market since they are aware that the examination of numerical time series produces accurate findings. Utilize the SVM technique to train the dataset and forecast the stock market.

Keywords — Machine Learning, SVM, Feature Extraction, Classification, Data Analytics, Stock Market, Decision Tree

INTRODUCTION

One of the highest rated industries for producing money for middle class investors is investment in stocks. Then, high class investors and traders engage in actual trading. The most crucial factor for investors is a company's share price, which is constantly subject to ups and downs. To avoid financial loss and ultimately to make money, it is essential to keep an eye on the live share market price and make quick decisions.

You must research the company's past financial performance and planned course of action for this. You can choose to invest based on a general market and company investigation. However, research has its limits because it is impossible to determine with absolute certainty whether a study or analysis is accurate.

The company's financial trajectory and future plans must be researched for this. You can choose to invest based on a comprehensive review of the market and the company. However, you are limited in what you can learn because it is impossible to know for sure whether a study or analysis is accurate.

Discovering the future worth of business stock and other financial assets traded on an exchange is made possible with the aid of stock price prediction utilising machine learning. Gaining significant profits is the whole point of making stock price predictions. It's challenging to forecast how the stock market.

I. RELATED WORK

PUBLICATION	Year	Title	Overview	Positive aspects	Limitations
IEEE	2021	Improving Traditional Stock Market Prediction Algorithms Using Covid-19 Analysis	By taking into account the parameters linked to COVID-19, the proposed work seeks to improve the stock market forecast capacity of	The process of predicting stock market trends is difficult due to its dynamic and turbulent nature. The COVID-19 pandemic in recent years has made this work even more difficult. The market has never been more volatile because to the increase in COVID-19 cases	Due to the fact that these standard trend prediction algorithms do not take into consideration how the pandemic has affected stock market

			several popular prediction models.	worldwide.	movements, this has led to their poor performance.
IEEE	2019	Analysis of Investor Sentiment and Stock Market Volatility Trend Based on Big Data Strategy [1].	In order to create a comparable analysis index, this article gathers data on the web news emotion index, web search volume, social network emotion index, and social network heat index.	Prediction of stock market volatility and risk aversion in the financial sector.	Brings a fresh viewpoint and a new framework for measuring the risk of market volatility.
IEEE	2017	Combining of Random Forest Estimates using LSboost for Stock Market Index Prediction [2].	Each of the prediction models is selected with technical indicators as inputs.	The goal of this study is to use regression to forecast the future values of stock market indexes based on the stock prices in the past.	The proposed prediction model may also be used in other fields, such as forecasting GDP, energy use, or weather.
IEEE	2018	Literature review on Artificial Neural Networks Techniques Application for Stock Market Prediction and as Decision Support Tools [3].	This study shows that the LSTM's capability exhibits a consistent rate of stock market prediction accuracy.	The results of this study show that an LSTM can reliably display a stock's accuracy rate.	Future research may evaluate the results utilising a large database that is now available as well as examine them using statistical methods.
IEEE	2018	Predicting the Effects of News Sentiments on the Stock Market [4].	The effects of news feelings on the stock market were	A potent indicator of stock movements, sentiment scores gleaned from the analysis of news articles can be effectively leveraged to forecast	The system can be enhanced to operate on a broad scale.

			recovered, extracted, and examined in this paper.	short-term trends.	
--	--	--	---	--------------------	--

II. PROPOSED METHODOLOGY

Support Vector Machine (SVM):

The supervised learning method Support Vector Machine is used to address problems in both classification and regression. To swiftly classify further data points in the future, this SVM algorithm's main goal is to identify the best line or decision boundary that may divide n-dimensional space into classes. The name of this best decision boundary is a hyperplane.

Decision Tree Algorithm:

A supervised learning method called a decision tree can be used to solve classification and regression problems, but it is typically favoured for doing so. It is a tree-structured classifier, where each leaf node represents the classification outcome and inside nodes represent the features of a dataset. The two nodes in a decision tree are the Decision Node and Leaf Node. While Leaf nodes are the results of decisions and do not have any more branches, Decision nodes are used to create decisions and have numerous branches.

Data Preprocessing:

Data preparation is the process of transforming raw data into something a machine learning model can use. It is the first and most crucial step in the process of creating a machine learning model.

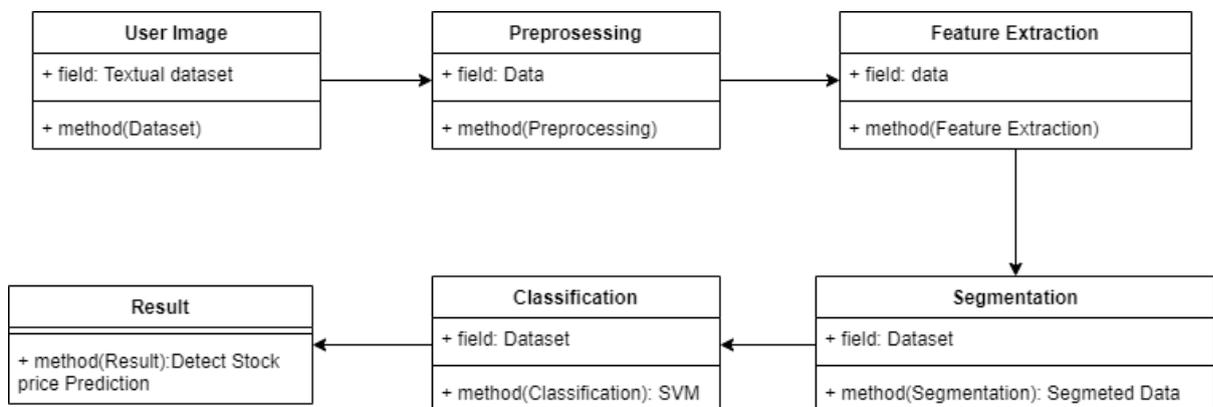
When working on a machine learning project, it is not always the case that we are presented with the clean and prepared data. Additionally, you must prepare and clean up your data every time you work with it.

It is not always the case that we come across the clean and prepared data when developing a machine learning project. Additionally, any time you work with data, you must clean it up and format it. Therefore, we use a data pretreatment activity for this.

Feature extraction:

Using the feature extraction method, a big quantity of raw data is divided into more manageable, more compact groups.

These huge data sets share the trait of having many variables that demand a lot of computational power to process. The term "feature extraction" refers to techniques for choosing and/or combining variables into features, which significantly reduces the amount of data that needs to be processed while properly and fully characterising the initial data set.



Segmentation:

This is made feasible via segmentation, a method of categorising clients into different groups based on characteristics or behaviour. Customer segmentation can be useful for other marketing tasks like upselling techniques, product recommendations, and pricing.

Classification:

It is possible to categorise the algorithm for supervised machine learning using regression and classification algorithms. We have forecasted the results for continuous values using regression techniques, but we require classification algorithms to predict the results for categorical values.

The Classification algorithm is a Supervised Learning method that is used to classify fresh observations on the basis of training data. In classification, a programme uses the dataset or provided observations to figure out how to divide new observations into different classes or groupings. Cat or dog, yes or no, 0 or 1, spam or not spam, etc. are a few examples. Classes may be denoted by targets, labels, or categories.

Unlike regression, classification produces a category as opposed to a value as the output variable, such as "Green or Blue," "fruit or animal," etc. Because the Classification method is a supervised learning technique and has input and output data, it needs labelled input data.

Mathematical Model:

Let S is the Whole System Consist of

$$S = I, P, O$$

Where,

I=CUR_LOC, SE_LOC, LOG, RE, PRO

LOG = userloginintosystem

SE_cO = SelectCompany

PRO = StockMovements

R = Shareprice

P = Process

Step1 : user will login

Step2 : User will select Company

Step3 : User will Analyze Data

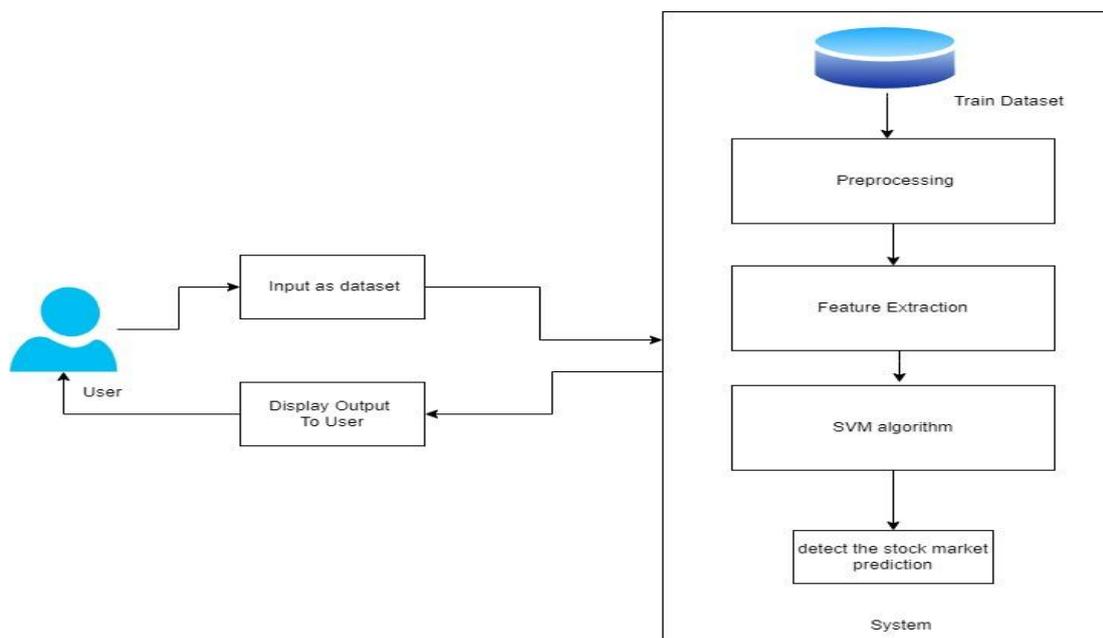
Step3 :User will Apply Mathematical Algorithmic Method

Step4 : System will give Prediction about further

Stock movement

OUTPUT : predictstockmarketprice

● **System Architectural Diagram**



Conclusion:

Here, we discovered that machine learning technology can be used to predict stock market movements. A graph of any company's stock price cannot be read in detail by a single individual. In the real world, we must analyse data at a massive scale with many different businesses. As a result, we may use machine learning techniques to make considerably better predictions. We may utilise the SVM algorithm to greatly increase the accuracy of our predictions.

References:

- 1) University of Bradford, 2017. H. Isah, "Social Data Mining for Crime Intelligence: Contributions to Social Data Quality Assessment and Prediction Methods."
- 2) University of Bradford, 2017. H. Isah, "Social Data Mining for Crime Intelligence: Contributions to Social Data Quality Assessment and Prediction Methods."
- 3) Expert Systems with Applications, vol. 83, pp. 187–205, 2017. E. Chong, C. Han, and F. C. Park, "Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies."
- 4) A unique data-driven stock price trend prediction system was developed by J. Zhang, S. Cui, Y. Xu, Q. Li, and T. Li in 2018. Expert Systems with Applications, vol. 97, pp. 60–69.
- 5) Expert Systems with Applications, vol. 105, pp. 11-22, 2018. L. S. Malagrino, N. T. Roman, and A. M. Monteiro, "Forecasting stock market index daily direction: A Bayesia Network approach."
- 6) "Stock Price Prediction Using Artificial Neural Network," by M. B. Patel and S. R. Yalamalle, International Journal of Innovative Research in Science, Engineering, and Technology, vol. 3, June 2014, pp. 13755–13762.
- 7) Wen Fang, Jie Wang, and Jun Wang. Elman Recurrent Random Neural Networks for Financial Time Series Prediction[J]. Neuroscience and Computational Intelligence, 2016(12):1–14
- 8) Twitter is a wonderful site to spot health concerns, according to V. M. Prieto, S. Matos, M. Alvarez, F. Cacheda, and J. L. Oliveira. 2014, PLoS ONE, vol. 9, no. 1, article e86191.
- 9) "Overcoming data scarcity of Twitter: Using tweets as bootstrap with application to autism-related topic content analysis," A. Beykikhoshk, O. Arandjelovi, D. Phung, and S. Venkatesh, in Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining, Paris, France, August 2015, pp. 1354–1361.
- 10) Deep learning- and word embedding-based heterogeneous classifier ensembles for text categorization, Z. H. Kilimci and S. Akyokus, Oct. 2018, Complexity, vol. 2018, art. no. 7130146.