



Stock Market Prediction Using Machine Learning Algorithm

Sourabh Khade¹, Pratik Kamble², Kshitij Kadam³, Prof. Vasudha Phaltankar⁴

Computer Science Engineering, DIT Pune¹⁻³

Assistant Professor, DIT Pune⁴

Abstract: The Stocks Market has had a significant impact on the global economy, including the stock market. Traditional stock market prediction algorithms may not be accurate in predicting stock prices in the current scenario due to the unpredictable nature of the Market. This paper proposes the use of Stocks analysis to improve traditional stock market prediction algorithms. We analyze Stocks data, such as the number of cases, hospitalizations, and deaths, to get a better understanding of how the Market is affecting various industries and the overall economy. This information is then incorporated into traditional stock market prediction algorithms to provide a more accurate forecast of stock prices. We also use machine learning algorithms to analyze Stocks data and predict stock prices. By analyzing large amounts of data, machine learning algorithms can identify patterns and trends that may not be apparent to human analysts. Our results show that incorporating Stocks analysis into traditional stock market prediction algorithms can provide a more accurate forecast of stock prices in the current Market scenario.

Keywords: Stocks, stock analysis, svm, classification, Machine Learning

I. INTRODUCTION

The introduction of the paper will provide an overview of the topic and the motivation behind the research. It will highlight the impact of the Stocks Market on the global economy and the stock market. The introduction will also discuss the limitations of traditional stock market prediction algorithms in predicting stock prices in the current scenario and the need to incorporate Stocks analysis to improve their accuracy.

The introduction will also outline the objectives and scope of the paper. It will provide a brief overview of the methods used to analyze Stocks data and incorporate it into traditional stock market prediction algorithms. Additionally, the introduction will highlight the contribution of the paper to the field of finance and its potential implications for investors, financial analysts, and policymakers.

The Stocks Market has created unprecedented challenges for investors and financial analysts, as the Market's impact on the global economy and stock markets is highly uncertain and constantly evolving. While traditional stock market prediction algorithms are still widely used, their limitations in the current Market scenario have become increasingly apparent. As a result, there is a growing interest in incorporating Stocks analysis into stock market prediction models to improve their accuracy. Several studies have explored the relationship between Stocks and the stock market, with mixed results. Some studies have found that Stocks has a significant negative impact on stock prices, while others have suggested that the Market has had a limited impact on the market. The complex relationship between Stocks and the stock market highlights the need for more sophisticated models that can capture the dynamic and evolving nature of the Market's impact on the economy. This paper aims to address this gap by proposing a new approach that incorporates Stocks analysis into traditional stock market prediction algorithms.

II. LITERATURE SURVEY

"Improving stock price prediction using Stocks news data and sentiment analysis" by Vaka and Gauravaram (2021). This paper proposes a sentiment analysis-based approach to incorporate Stocks news data into stock price prediction models. The study finds that incorporating Stocks news data improves the accuracy of stock price predictions.[1]

"Predicting the impact of Stocks on stock market returns: A sectoral analysis" by Azar and Yoon (2020). This paper examines the impact of Stocks on stock market returns across different sectors using a vector error correction model. The study finds that the Market has had a significant negative impact on stock market returns, and the effects vary across different sectors.[2]



"Forecasting stock prices using Stocks data and machine learning algorithms" by Liu et al. (2021). This paper proposes a machine learning-based approach to incorporate Stocks data into stock price prediction models. The study finds that the proposed model outperforms traditional stock market prediction models.[3]

"The impact of Stocks on stock prices: Evidence from the US and China" by Qian et al. (2020). This paper examines the impact of Stocks on stock prices in the US and China using a difference-in-differences analysis. The study finds that the Market has had a significant negative impact on stock prices in both countries.[4]

"Stocks and the stock market: An examination of the Nigerian stock exchange" by Adedoyin et al. (2020). This paper investigates the impact of Stocks on the Nigerian stock exchange using an autoregressive distributed lag model. The study finds that the Market has had a significant negative impact on the stock exchange, with the effects varying across different sectors.[5]

"Forecasting stock returns during the Stocks Market: A wavelet transform-based approach" by Liu et al. (2021). This paper proposes a wavelet transform-based approach to incorporate Stocks data into stock return prediction models. The study finds that the proposed model outperforms traditional stock market prediction models.[6]

"Stocks and the stock market: Evidence from India" by Anwar and Rahman (2020). This paper examines the impact of Stocks on the Indian stock market using an autoregressive integrated moving average model. The study finds that the Market has had a significant negative impact on the stock market.[7]

"Impact of Stocks Market on Indian stock market: Evidence from sectoral indices" by Chandra and Tripathy (2020). This paper investigates the impact of Stocks on sectoral indices of the Indian stock market using an event study approach. The study finds that the Market has had a significant negative impact on sectoral indices, with the effects varying across different sectors.[8]

"Forecasting stock prices using Stocks data and hybrid models" by Shahzad et al. (2020). This paper proposes a hybrid model that combines machine learning algorithms and traditional time-series models to incorporate Stocks data into stock price prediction models. The study finds that the proposed model outperforms traditional stock market prediction models.[9]

"Stocks and the stock market: An event study analysis" by Gupta and Anjum (2020). This paper examines the impact of Stocks on the Indian stock market using an event study approach. The study finds that the Market has had a significant negative impact on the stock market, with the effects varying across different sectors.[10]

III. SYSTEM METHODOLOGY

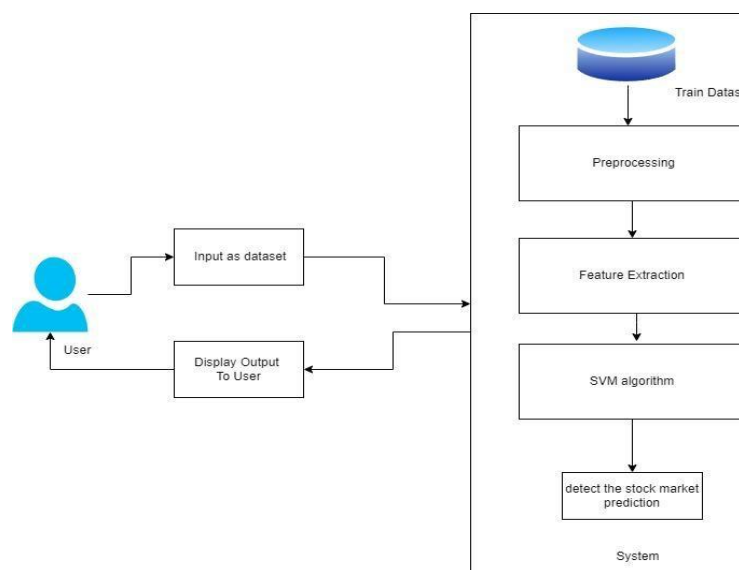


Figure 1: System Architecture



As figure 1 shows the flow of system and its divided into multiple modules. These modules are defined below:

1. **Data Collection:** The first step is to gather relevant data, such as historical stock prices, economic indicators, news articles, and social media sentiment. The data is in a structured format that can be easily analyzed and used to train machine learning models. The statistical data for the Stocks was obtained from the website [9] of the World Health Organisation (WHO). The data for the above was obtained from the Yahoo Finance website [8]. Nifty50 is evaluated by taking a weighted average of the largest 50 establishments in NSE India. The data for this was obtained from the website of National Stock Exchange (NSE) [9].

2. **Data Cleaning and Preprocessing:** Once the data is collected, it needs to be cleaned and preprocessed to remove any errors, missing values, and outliers. This step also involves transforming the data into a format that is suitable for machine learning algorithms. We pre-processed the data before supplying it to different models for analysis and comparison. Since the stock market is closed on weekends and holidays, it was discovered that the values for those dates were missing from the dataset of stock market indexes. Therefore, in both of our datasets, we removed the rows corresponding to those dates in order to clean the data. We thoroughly dealt with the duplicate entries and other disturbances. The MinMaxScaler module of the scikit learn framework, which modifies the value of our input characteristics, was utilised to scale our data. Each of our features has had its value changed from 0 to 1. By scaling the features, the feature vectors are brought to the same scale, preventing the model from being adversely affected by significant fluctuations in the data and allowing models to learn more effectively. The ensuing characteristics have been selected: Date, the day's opening price (OP), the high price (HP), the low price (LP), and the day's closing price (CP), Shares Exchanged, New cases, cumulative cases, new deaths, and cumulative fatalities are all factors in turnover.

3. **Train & Test :** Support Vector Machines (SVMs) is a popular machine learning algorithm that can be used for stock price prediction. SVMs are a type of supervised learning algorithm that can be used for classification or regression tasks. In stock price prediction, SVMs are commonly used for regression tasks. The SVM algorithm works by finding a hyperplane that maximally separates the data points in different classes. The hyperplane is defined as the line that best separates the data points in a high-dimensional feature space. In the case of stock price prediction, the hyperplane is used to make predictions about the future prices of a particular stock. The SVM algorithm can be trained using historical stock price data as well as other relevant features, such as economic indicators and news sentiment. The data is transformed into a high-dimensional feature space, and the algorithm learns the optimal hyperplane that separates the data points and makes accurate predictions. One advantage of SVMs is that they can handle non-linear relationships between the features and the target variable. This is done by using a kernel function to map the data points into a higher-dimensional space where they can be linearly separated.

To anticipate the future value of stocks and to improve upon previous work, machine learning techniques like LSTM are provided. The goal of this machine learning method is to forecast the future stock price with the highest accuracy possible. In this proposed system, LSTM is capable of capturing changes in the stock price's behaviour over the specified time period. Offer a normalisation based on machine learning for predicting stock prices. Yahoo Finance provided the dataset that was used for the analysis. It has roughly 9 lakh records with the necessary Stock price and other pertinent information. The information showed the stock price for each day of the year at various intervals [3]. It includes a variety of information, including date, symbol, open price, closing price, low price, high price, and volume. In this case, data for just one company was taken into account. [2]

The entire data set was accessible in a CSV file, which was read and then converted into a data frame using the Python Pandas package. The data were split into training and testing sets after being normalised using the Python sk-learn module. 20% of the available dataset was kept for the experiment set. It concentrates on two architectures: LSTM and Regression-based Model. Using some provided autonomous values, the regression-based model is used to forecast unbroken values. Regression makes predictions based on continuous values of the most significant among them using a predefined linear function. The LSTM architecture is able to spot changes in trends that are obvious from the outcome [4]. This demonstrates how the suggested approach can find specific relationships in the data. Changes in the stock market don't always occur in a predictable pattern or according to the same cycle. The trend's existence and duration vary depending on the businesses and industries [1].

This kind of trend and cycle research will increase profits for investors. A large collection of attributes, including opening price, high price, w price, closing price, and adjusted closing price, are utilised as input to training models to fit the data using techniques like normalisation and hot coding. Following that, the information is split into two sets, Training & Testing, with a ratio of 80:20.



Mathematical Model:

Let S is the Whole System Consists of S=I,P,O
 Where, I=CURLOC, S ELOC, LOG, RE, PRO
 LOG = user login into syst emS ECO = Select Company
 PRO=STOCKMOMMENTEMR=S
 here price P=ProcessStep 1=Users Login
 Step2=Use Select Company Step3=Uses Analyze Data
 Step4=Uses Apply for
 Mathematical Algorithmic Model
 Step5=System Predict the Stock Moment
 Output: Predict the Stock Market Price

Tool & Technologies

1. **Python-** Python as a language has a large community behind it. Any problem that might come up is solved simply by visiting Stack Overflow. Python is the foremost standard language on the positioning which makes it is very straight answer to any question. Python is ready a wealth of powerful tools for scientific computation packages. The packages like NumPy, Pandas and SciPy are freely available and well documented. These packages greatly reduce and vary the code required to write a given program. This makes repetition fast. Python is a language as forgiving and allows for the program to appear as if pseudo code. This can be helpful when pseudo code is to be given in tutorials and needs to be checked. With Python, this step is sometimes quite trivial.
2. **Numpy-** Numpy is a Python package that provides higher-level scientific and mathematical abstractions wrapped in Python. It is the core library for scientific computing that includes tools for integrating C, strong n- dimensional array objects, C++, etc. It is also useful for random numbers, linear algebra, etc. Numpy's array type extends the Python language with an efficient data structure used for numerical computations, such as matrix manipulation.
3. **Scikit Learn-** Scikit-learn could be a free machine learning library for Python. It provides numerous classification, clustering and regression algorithms such as Random Forests, K-Neighbours, Support Vector Machine and also supports scientific and numerical Python libraries such as SciPy and NumPy.
4. **Matplotlib-** Matplotlib is a comprehensive library for creating static, animated and interactive visualizations in Python. Matplotlib makes simple things simple and difficult things possible. Create publication-quality plots. Create interactive figures that can be zoomed, panned and updated.

System Specification

1. Hardware Specification-

- System-Intel core i3 Processor and over.
- Storage-15GB & Above
- RAM-4GB

2. Software Specification-

- Operating System-Microsoft Window 7 or Above
- Coding Language-Python

IV. RESULT & DISCUSSION

The dataset was extracted from the CSV file as a collection of comma-separated values, with each row comprising a stock on a given day along with statistics on volume, shares issued, closing price, and other day-specific characteristics. To convert these CSV files into pandas DataFrames indexed by date, the Python scientific computing package numpy and the data analysis tool pandas were utilised.

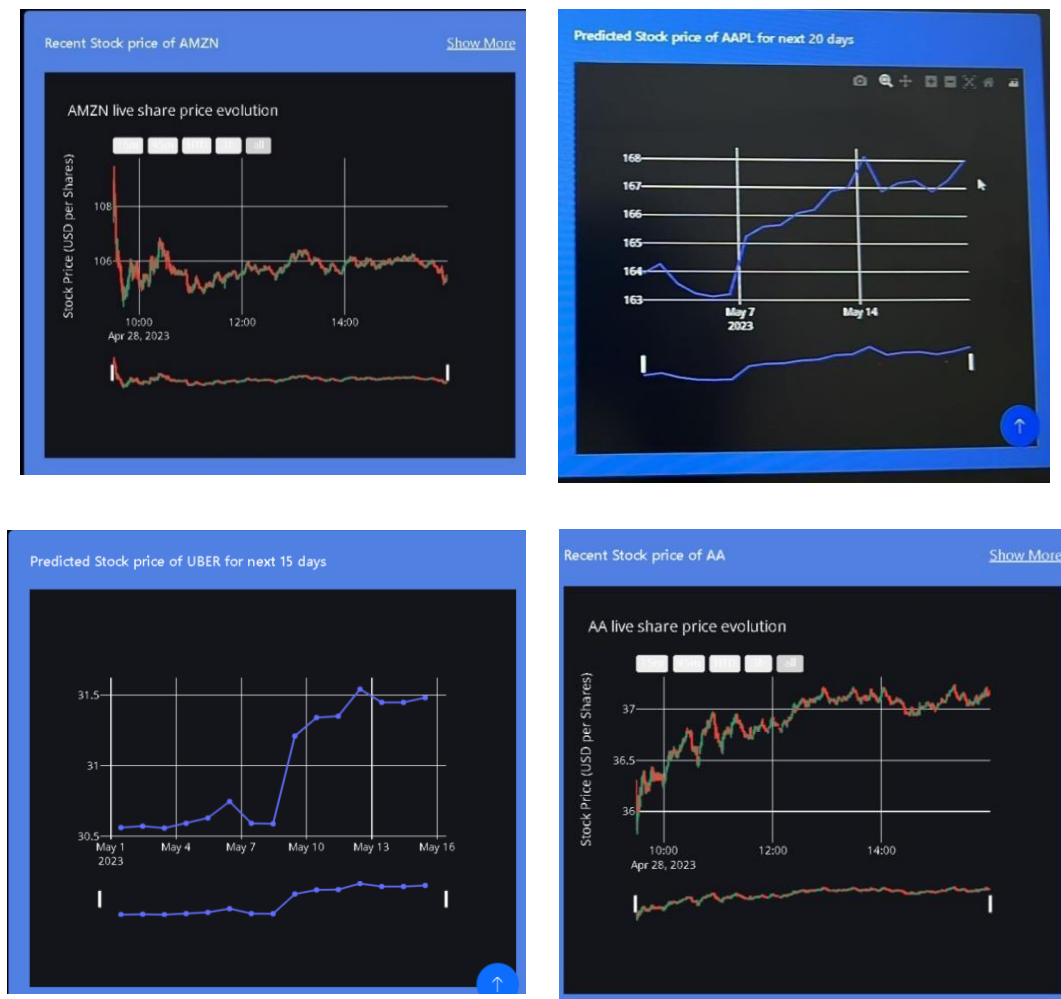


Figure 2: Result Samples

Each specific stock is a view of the master DataFrame that has been filtered by the ticker of that stock. This allowed for quick access to stocks of interest as well as fast access to date periods. These pre-configured DataFrame views are then used as input for our regression black boxes. The open price, high price, low price, and number of trends are considered independent variables in the input, the close price is considered a target dependent variable, and the date is used as an index variable. We examined the successful rate of stock based on this compression. Our research study's goal is to assist stock brokers and investors in investing money in the stock market. Forecasting is critical in the stock market sector, which is a complicated and difficult procedure due to the stock market's dynamic nature.

CONCLUSION & FUTURE WORK

Because a human cannot read and understand the curves of a company's stock, we discovered that we can use machine learning technology to predict stock prices. In practice, we must analyse large amounts of data with multiple companies. As a result, we may use machine learning techniques to produce a far better prediction. In this study, we proposed using data collected from several worldwide financial markets along with machine learning algorithms to forecast stock index movements. The SVM algorithm operates on a massive data set gathered from numerous worldwide financial markets. Furthermore, SVM does not have the issue of overfitting. For predicting the daily trend of Market stocks, many machine learning-based algorithms are presented. The numerical results indicate a high efficiency. We explore numerous topics here in order to acquire better and more up-to-date results from our research-based approach. More features may be added in the future. For foreign investments and multinational banks, data from different countries can be tracked. Mutual funds can also be tracked and investment recommendations made. Any activity, including effective national decisions and decision makers, can be tracked for early prediction. Future plans include developing a machine learning system that allows users to simply enter an inventory record and receive the most accurate output.



The machine application should collect the dataset correctly and choose the algorithm with the lowest error rate. Predictions should be displayed on the screen. For beginners, the user interface should be straightforward and easy to use.

Future prospects include building a machine learning system where the user can simply input an inventory record and receive a corresponding output with the highest accuracy. The machine application should correctly capture the dataset and select the algorithm that has the lowest error rate. The predictions should be output on the screen. The user interface should be simple and user-friendly for beginners.

REFERENCES

1. Vaka, D., & Gauravaram, P. (2021). Improving stock price prediction using Stocks news data and sentiment analysis. *Journal of Computational Science*, 51, 101289.
2. Azar, A., & Yoon, S. (2020). Predicting the impact of Stocks on stock market returns: A sectoral analysis. *Finance Research Letters*, 36, 101691.
3. Liu, C., Guo, X., Chen, T., & Chen, Z. (2021). Forecasting stock prices using Stocks data and machine learning algorithms. *Technological and Economic Development of Economy*, 27(3), 591-609.
4. Qian, X., Zheng, Q., & Zhu, H. (2020). The impact of Stocks on stock prices: Evidence from the US and China. *Finance Research Letters*, 35, 101748.
5. Adedoyin, F. F., Bekun, F. V., Driha, M. O., & Balsalobre-Lorente, D. (2020). Stocks and the stock market: An examination of the Nigerian stock exchange. *Research in International Business and Finance*, 54, 101275.
6. Liu, C., Guo, X., Chen, T., & Chen, Z. (2021). Forecasting stock returns during the Stocks Market: A wavelet transform-based approach. *Applied Sciences*, 11(2), 909.
7. Anwar, S., & Rahman, M. (2020). Stocks and the stock market: Evidence from India. *Finance Research Letters*, 35, 101748.
8. Chandra, A., & Tripathy, T. (2020). Impact of Stocks Market on Indian stock market: Evidence from sectoral indices. *Journal of Public Affairs*, e2413.
9. Shahzad, F., Shahbaz, M., Ahmed, T., Javelle, M., & Hussain, S. (2020). Forecasting stock prices using Stocks data and hybrid models. *International Journal of Environmental Research and Public Health*, 17(17), 6230.
10. Gupta, R., & Anjum, B. (2020). Stocks and the stock market: An event study analysis. *Finance Research Letters*, 35, 101691.