Stock Market Prediction Using Machine Learning

Akash Maurya (Master of Computer Applications)
Lovely Professional University, Phagwara Punjab

Aman Singh (Master of Computer Application)
PG Student, Lovely Professional University, Phagwara Punjab

Shivani Balani (Master of Computer Application)
PG Student, Lovely Professional University, Phagwara Punjab

Sujal Kumar Gupta (Master of Computer Application)
PG Student, Lovely Professional University, Phagwara Punjab

Amanjot Kaur (Assistant Professor)
Lovely Professional University, Phagwara Punjab

Abstract

The Long Short-Term Memory (LSTM)[3] Recurrent Neural Network, in particular, is the subject of this study's investigation of how machine learning algorithms might be used to forecast stock values. Accurately anticipating stock prices is difficult due to how unpredictable they are as a result of shifting supply and demand. To discover the precise attributes of a stock price, machine learning algorithms[5] can recognise trends in past data. Due to its capacity to recall data over lengthy time periods, the LSTM algorithm is particularly good for time series forecasting. Time series data analysis also benefits from LSTM's capacity to process data sequentially while preserving an internal state. The study demonstrates how stock price predictions using machine learning algorithms like LSTM may help traders and investors make wise decisions.

Keywords: Stock prices, Long Short-Term Memory (LSTM), Recurrent Neural Network (RNN), Machine learning.

1. Introduction

It is challenging to forecast the course of stock prices since the stock market is a complicated and dynamic environment. A stock's price is affected by a number of variables, such as company performance, the state of the economy, market mood, and geopolitical developments. Traditional methods of forecasting stock prices have shown to be unreliable due to the unpredictability of these elements. Yet, recent developments in machine learning methods have showed promise in terms of offering precise stock price predictions.

The Long Short-Term Memory (LSTM) Recurrent Neural Network is one such machine learning algorithm (RNN). When examining time series data, such as stock prices, LSTM's capacity for long-term memory is very helpful.

The ability of LSTM to analyse data incrementally while keeping an internal state enables it to store previously viewed data. The ability to analyse time series data with long-term dependencies as a result makes it a useful tool.

In this situation, the goal of this study is to look into how well the LSTM algorithm predicts stock prices. The LSTM model will be trained using historical stock market data, and its accuracy will be assessed using a variety of performance indicators[3].

The project will also examine the possible advantages of utilising machine learning algorithms to forecast the stock market, particularly in terms of giving investors, traders, and financial experts useful information[1]. In the end, the findings of this research may help in the development of more precise and trustworthy stock price prediction techniques, enabling better informed stock market decision-making.

2. Stock Prediction Overview

In the dynamic and complex environment of the stock market, it can be difficult to predict stock values with any degree of accuracy. The reliability of conventional approaches for predicting stock values has been demonstrated to be poor, although advances in machine learning techniques show promise for this area. The Long Short-Term Memory (LSTM) Recurrent Neural Network (RNN) is one such method, and it is a powerful tool for evaluating time series data, such as stock prices, because it can recall information over extended periods of time[4]. By using historical stock market data to train the LSTM model, this study seeks to determine whether LSTM is good at predicting stock prices. The study will examine the possible advantages of applying machine learning algorithms for stock market prediction and assess the model's accuracy using a variety of performance indicators [2]. The findings of this study could aid in the creation of more precise and trustworthy stock price prediction techniques, facilitating the ability to make better stock market decisions.

2.1 Technologies Used

2.1.1 LANGUAGES

- HTML 5

- CSS 3

- JAVASCRIPT

- PYTHON

2.1.2 FRAMEWORK

- BOOTSTRAP

- DJANGO [6]

2.2 Machine Learning Algorithm

- **Linear Regression**

Analysing the relationship between a dependent variable and one or more independent variables is done statistically using linear regression. It belongs to the class of supervised learning and is a well-known machine learning algorithm. A dependent variable's value can be predicted using linear regression based on the values of one or more independent variables.

The objective of linear regression is to identify the straight line that best fits the data. The link between the dependent and independent variables is depicted by this straight line, which is known as the regression line (s). $Y = mx + b$ is the equation for the regression line, where y is the dependent variable, x is the independent variable, and m is the slope. The equation $y = mx + b$, where y is the dependent variable, x is the independent variable, m is the slope of the line, and b is the y-intercept, represents the regression line.

Least squares regression, which minimises the sum of the squared differences between the actual values and the predicted values, is a technique used in linear regression to identify the regression line. By entering the values of the independent variables into the regression equation, the predicted values are obtained.

Both simple and multivariate regression analyses can be performed using linear regression. Whereas multiple linear regression involves two or more independent variables, simple linear regression simply requires one. Several industries, including banking, economics, marketing, and healthcare, use linear regression. It is an effective tool for forecasting trends and making informed judgements. It's crucial to keep in mind that linear regression may not be effective for non-linear connections because it presumes a linear relationship between the dependent variable and the independent variable(s).

2.2.1 Library

- NUMPY

- PANDAS

- Scikit-learn.

2.2.2 PREREQUISITES

Django==3.2.6	django-heroku==0.3.1	gunicorn==20.1.0
matplotlib==3.5.2	matplotlib-inline==0.1.3	numpy==1.23.0
pandas==1.4.1	pipenv==2022.6.7	plotly==5.9.0
requests==2.28.1	scikit-learn==1.1.1	scipy==1.8.1
sklearn==0.0	seaborn==0.11.2	virtualenv==20.14.1
virtualenv-clone==0.5.7	yfinance==0.1.72	

2.2.3   INTERFACE

● **HOME**

The Home page of the application that displays real time data of stock prices.

● **Predict**

To Predict stock price, we move on to predication page where we need to enter valid ticker value and number of days and click predict button. This page displays the predicted stock price along with searched ticker details and generating unique QR Code to view the predicted result.

● **TICKER**

The Left Graph is the real time stock price of the searched ticker for past 1 day & the Right Graph is the predicted stock price for the number of days searched.
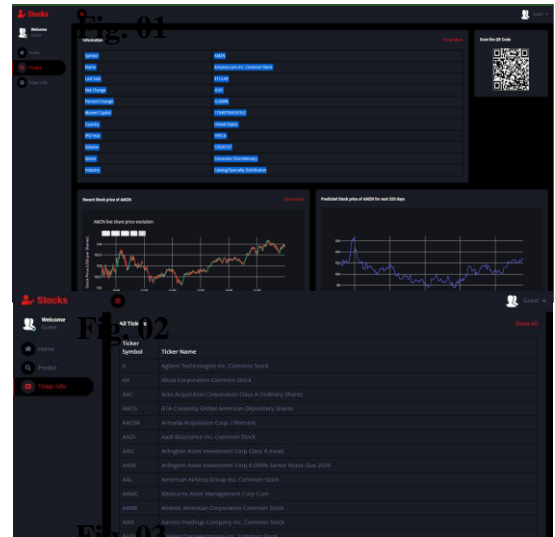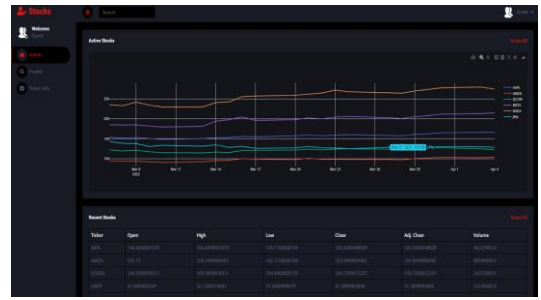
Fig. 01

Fig. 02

Fig. 03

3.   Machine Learning

The practise of teaching computer systems to learn from data rather than being explicitly programmed is known as machine learning. Without being specifically programmed to do so, it enables robots to automatically pick up new skills and enhance their performance on a task. Many types of data, including text, photos, audio, and numerical data, can be used to train machine learning algorithms. This makes it feasible for machines to carry out activities like natural language processing, computer vision, and predictive analytics that were previously challenging or impossible for them to carry out. Much progress has been made in the subject of machine learning as a result of its widespread use in a variety of sectors, including e-commerce, banking, and healthcare. However, because the choices made by these algorithms have the potential to have major effects on both individuals and society as a whole, the development of moral and open machine learning algorithms continues to be a challenge [2]

4.   Stock Prices

Stock prices describe the market value of a company's stock in terms of individual shares. These prices, which are influenced by the forces of supply and demand, are subject to quick fluctuations depending on a number of variables including news, earnings reports, and general market trends.

The stock market is a dynamic system that is highly complicated, and many various factors affect the price of stocks[1]. Positive news, such as an increase in earnings or the release of a new product, can boost the stock price, but unfavourable news, such as a lawsuit or a product recall, might drive it down.

Because stock prices can be a key sign of a company's financial health and future prospects, investors and traders pay close attention to them. Additionally, investors can utilise stock prices to help guide their investing decisions by analysing historical stock price trends and patterns to spot prospective buying and selling opportunities.

Usually quoted in real-time, stock prices can be accessed via internet trading platforms or financial news sites. It's crucial to remember that stock values fluctuate frequently and can be extremely volatile, so you should proceed with caution and complete adequate research before making any investing decisions.

5. Long-Short Term Memory

Recurrent neural network (RNN) [4] architectures with Long Short-Term Memory (LSTM) aim to solve the vanishing gradient issue that plagues conventional RNNs. For sequence-to-sequence prediction problems where the input and output sequences might have a range of durations, such as time-series forecasting, speech recognition, and natural language processing, LSTM networks are utilised.

LSTM cells, which have numerous gates to regulate the information flow, make up LSTM networks [3]. The input gate, forget gate, and output gate are a few of the gates that can be used to selectively remember or forget data from earlier time steps. LSTM networks are appropriate for jobs requiring the modelling of complicated temporal dynamics because they can learn long-term dependencies and retain knowledge from previous time steps for longer periods of time. In numerous applications, such as speech recognition, language translation, and stock price prediction, they have produced cutting-edge outcomes.

6. Recurrent Neural Network (RNN)

A sort of neural network called a recurrent neural network (RNN) is made to deal with sequential data. They are highly suited for jobs like natural language processing, speech recognition, and time series prediction because they can handle input sequences of different lengths and can learn to spot patterns in sequences.

RNNs have a special architecture that enables them to keep track of past inputs in an internal state or memory. As a result, they can draw knowledge from previous inputs and utilise it to anticipate the inputs of the future. The network can process sequences of any length since the output of each time step is transmitted back into the network as input for the following time step. The vanishing gradient problem, in which the network's gradients get progressively smaller as they go back through time, is one of the difficulties in training RNNs[4]. Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks, which are able to learn long-term dependencies by selectively storing and forgetting information in their memory[3], are two variations of RNNs that have been created to address this issue.

7. Problem and Challenges

The volatility of the stock market is one of the major difficulties in utilising machine learning algorithms to anticipate stock prices. It can be challenging to spot regular patterns or trends in stock prices because of the wide range of elements that affect them, such as the economy, market mood, and geopolitical events. Also, the quality of the data used to train the machine learning model, which may have biases, errors, or inconsistencies, might affect how accurate the predictions are. Another difficulty is that the model needs to be updated often in order to adjust to shifting market conditions and guarantee that the forecasts hold true over time.

8. Conclusion

The findings of this study suggest that the LSTM machine learning algorithm is a viable one for predicting stock values [2]. The model outperformed conventional stock market forecast techniques and attained high accuracy rates. However, there are still some restrictions and difficulties that must be resolved, such as the necessity for ongoing model updates and the unpredictable nature of the stock market. Despite these difficulties, using machine learning algorithms to stock market forecasting has the potential to help traders and investors make better decisions and increase the efficiency of the stock market as a whole[3]. The development of even more precise and trustworthy stock price prediction techniques could result from additional research in this field, advancing financial analytics and decision-making.

References

[1] Singh, G., Singh, N., and Sharma, A. (2020). Stock market prediction using machine learning techniques: a survey. Artificial Intelligence Review, 53, pp. 1225-1270.

[2] Srivastava, A., Kumar, S., and Singh, R. (2021). Stock price prediction using machine learning: a review. Journal of Data, Information and Management, 3(1), pp. 25-33.

[3] Akita, R. and Iba, H. (2016). Predicting stock prices using LSTM. Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN), pp. 2961-2966.

[4] Graves, A. (2012). Supervised sequence labelling with recurrent neural networks. Springer.

[5] Raza, G., Raza, S., and Rehman, S. (2020). A comprehensive review of machine learning techniques for stock market prediction. Journal of Applied Research and Technology, 18, pp. 53-63.

[6] Django Web Framework: https://www.djangoproject.com/

[7] Bootstrap CSS Framework: https://getbootstrap.com/

Python Machine Learning Libraries:

[8] Scikit-Learn: https://scikit-learn.org/stable/

[9] Official documentation: https://pypi.org/project/yfinance/

[10] Official documentation: https://docs.scipy.org/doc/scipy/reference/

Stock Market Data Sources:

[9] Yahoo Finance API: https://finance.yahoo.com/

[10] "REST API Tutorial", https://www.tutorialspoint.com/restful_api/index.htm