Stock Price Prediction Using Machine Learning

Patel Vraj Hemangkumar[1], Patel Aksh Hemantkumar[2],
Patel Ansh Satishbhai[3], Padariya Jagrutkumar Kaushikbhai [4], Parmar Sahaj Ketan[5]
[1,2,3,4,5]*Department of Information Technology, Sigma institute of Engineering, Vadodara*

ABSTRACT:

Stock price prediction approach using Long Short-Term Memory (LSTM) networks. The proposed method utilizes historical stock price data as input and trains an LSTM model to predict future stock prices. The LSTM network architecture includes multiple layers of LSTM cells, which can capture long-term dependencies and patterns in the input data. The effectiveness of the proposed approach is evaluated using real-world stock price data from the S&P 500 index. Experimental results demonstrate that the LSTM-based approach can achieve promising prediction accuracy compared to other traditional methods. The proposed method can potentially benefit financial analysts, investors, and traders by providing useful insights for decision-making.

***Key Words*:** Financial analysis, Stock-Market, LSTM, Future prediction, Machine learning

1. INTRODUCTION

Long Short-Term Memory (LSTM) networks have shown promising results in various time-series prediction tasks, including stock price prediction. LSTMs are a type of recurrent neural network (RNN) that can selectively retain or forget information over time, making them well-suited for modelling sequential data such as stock prices.

In recent years, many studies have explored the use of LSTMs for stock price prediction and achieved promising results. In this paper, we propose an LSTM-based approach for stock price prediction. The proposed method uses historical stock price data as input and trains an LSTM model to predict future stock prices. The LSTM - network architecture includes multiple layers of LSTM cells, which can capture long-term dependencies and patterns in the input data. We evaluate the effectiveness of the proposed approach using real-world stock price data from the S&P 500 index and compare it with other traditional methods. The proposed method can potentially financial analysts, investors, and traders by providing useful insights for decision-making.

2. IDEA ABOUT OUR PROBLEM AND SOLUTION

I. What was the stock's price over the course of time?

Predicting the change in stock prices over time is one of the key objectives of using an LSTM model for stock price prediction. The LSTM model can be trained on historical stock price data to identify patterns and trends in the data, which can then be used to predict the change in stock prices over time.

The change in stock prices over time is typically measured using metrics such as percentage change, absolute change, or logarithmic returns. These metrics reflect the magnitude and direction of the change in stock prices over a given time period.

For example, the percentage change in stock prices over a day can be calculated as (Closing price - Opening price) / Opening price * 100%. This metric indicates whether the stock price increased or decreased over the day and by what percentage.

The LSTM model can be used to predict the change in stock prices over different time periods, such as hourly, daily, weekly, or monthly intervals.

The accuracy of the LSTM model in predicting the change in stock prices over time can be evaluated using various performance metrics, such as mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), or correlation coefficient (R).

Overall, the change in stock prices over time is a crucial aspect of stock price prediction using an LSTM model, as it helps investors and traders make informed decisions about buying or selling stocks based on their predicted price movements.

II. What was the stock's average daily return?

A moving average (MA) is a stock indicator used frequently in technical analysis in the world of finance. The purpose of generating a stock's moving average is to create a continuously updated average price in order to assist smooth out the price data.

The effects of random, short-term changes on the price of a stock over a given time period are reduced by using the moving average calculation. A basic arithmetic average of prices over a time period is used by simple moving averages (SMAs), but exponential moving averages (EMAs) give more weight to more recent values than to older ones over the same time period. The moving average of a stock is calculated by taking the average price of the stock over a certain time period, typically using a sliding window approach. For example, the average price of a stock over a predetermined time period is used to calculate the moving average for that stock, which is commonly done using a sliding window method. For instance, the average price of a stock over the last 50 days would be used to determine its 50-day moving average.

Depending on the analyst's option, the moving average can be calculated using a variety of time periods, including 10 days, 20 days, 50 days, 100 days, etc. In stock price prediction using LSTM models, moving average can be used as a feature or input to the model to help capture trends and patterns in the data. For example, the moving average of a stock over the last 10 days can be used as an input to the LSTM model to predict the stock price for the next day.

In addition to using moving average as an input to the LSTM model, financial analysts often use moving average as a technical indicator to identify potential buying or selling opportunities. For example, if the current stock price is above its 50-day moving average, it may indicate a bullish trend, whereas if the current stock price is below its 50-day moving average, it may indicate a bearish trend.

Overall, moving average is a useful tool in stock price prediction using LSTM models, as it can provide valuable insights into the overall trend and momentum of a particular stock over time.

III. What was the different stock's moving average?

In the context of stock price prediction using LSTM models, the average daily return of a stock can be used as a feature or input to the model to help capture the volatility and risk associated with the stock.

To calculate the average daily return of a stock, we can use the formula mentioned in the previous answer: Average Daily Return = (Closing Price - Opening Price) / Opening Price * 100%This formula calculates the

percentage change in the stock price over a single trading day. The average daily return of a stock can be calculated by taking the average of the daily returns over a certain period, such as a week, a month, or a year.

For example, if we want to calculate the average daily return of a stock over the last month, we can take the sum of the daily returns for the past 30 days and divide it by 30. This would give us the average daily return of the stock over the past month.

In stock price prediction using LSTM models, the average daily return of a stock can be used as an input feature to the model to help capture the volatility and risk associated with the stock. This can help the model make more accurate predictions of the stock price by taking into account the historical daily returns of the stock.

Overall, the average daily return of a stock is an important metric in stock price prediction using LSTM models as it can help capture the volatility and risk associated with the stock and improve the accuracy of the model's predictions.

IV. What was the correlation among the various stocks?

The correlation between different stocks' closing prices refers to the degree to which the closing prices of two or more stocks move together. Correlation can provide insights into the relationships between different stocks and can be used to build diversified portfolios.

In the context of stock price prediction using LSTM models, correlation can be used as a feature or input to the model to help capture the relationships between different stocks. For example, the correlation between the closing prices of two stocks can be used as an input to the LSTM model to predict the stock prices of both stocks.

The correlation coefficient is a common measure of the correlation between two stocks' closing prices. The correlation coefficient ranges from -1 to +1, with a value of +1 indicating a perfect positive correlation, a value of -1 indicating a perfect negative correlation, and a value of 0 indicating no correlation. To calculate the correlation coefficient between two stocks' closing prices, we can use the following formula: Correlation Coefficient = Covariance (X, Y) / (StdDev(X) * StdDev(Y)) where X and Y are the two stocks' closing prices, Covariance is the covariance between the two stocks' closing prices, and StdDev is the standard deviation of the two stocks' closing prices.

In addition to calculating the correlation coefficient between two stocks, we can also calculate the correlation matrix, which shows the correlation between each pair of stocks in a portfolio.

Overall, the correlation between different stocks' closing prices is an important metric in stock price prediction using LSTM models, as it can provide insights into the relationships between different stocks and help build diversified portfolios.

V. How much of our money are we risking when we buy a particular stock?

The amount of value that is put at risk by investing in a particular stock depends on a variety of factors, including the volatility of the stock, the amount of capital invested, and the individual's risk tolerance. In the context of stock price prediction using LSTM models, we can use various risk management techniques to manage the risk associated with investing in a particular stock. One commonly used technique is to calculate the maximum drawdown, which measures the largest peak-to-trough decline in the value of a stock or portfolio.

To calculate the maximum drawdown, we can use the following formula: Maximum Drawdown = (Peak Value - Trough Value) / Peak Value where Peak Value is the highest value of the stock or portfolio over a certain period, and Trough Value is the lowest value of the stock or portfolio over that same period. The maximum drawdown can give investors an idea of the amount of value they could potentially lose by investing in a particular stock. For example, if the maximum drawdown for a stock is 20%, an investor who invests $10,000 in that stock could potentially lose up to $2,000 in value. In addition to calculating the maximum drawdown, investors can also use other risk management techniques, such as diversification and stop-loss orders, to manage the risk associated with investing in a particular stock.

Overall, the amount of value that is put at risk by investing in a particular stock depends on a variety of factors, and risk management techniques, such as calculating the maximum drawdown, can help investors manage that risk. In stock price prediction using LSTM models, risk management techniques can be used to help investors make more informed investment decisions based on the predicted stock prices.

3. BLOCK DIAGRAM OF PREDICTION MODEL

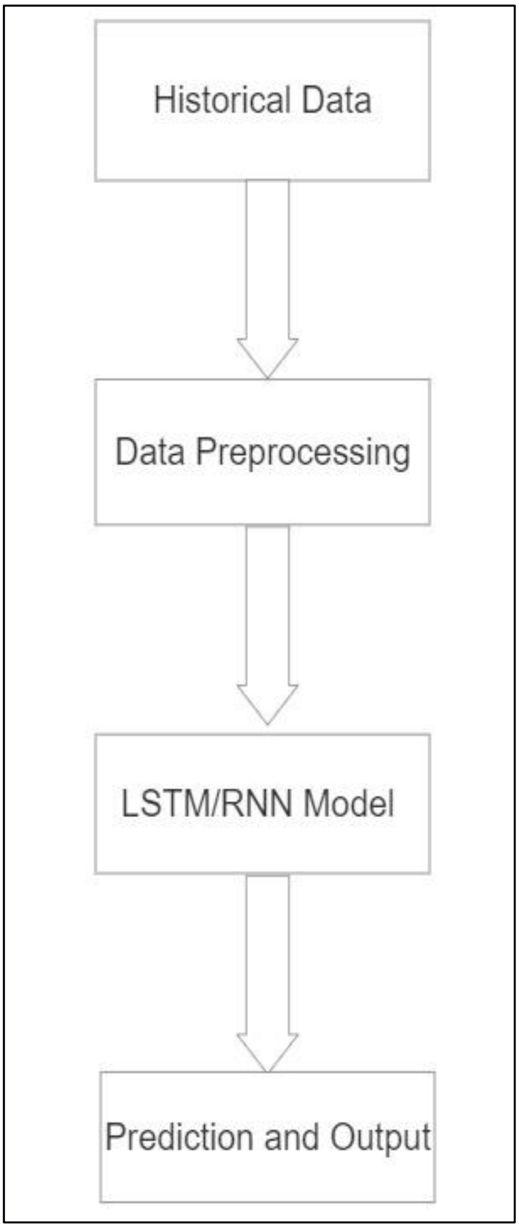

Figure3.1

4. TECHNOLOGIES

4.1) Algorithms:

Long Short-Term Memory (LSTM) algorithm: LSTM is a type of Recurrent Neural Network (RNN) algorithm that is specifically designed to capture long-term dependencies in time series data. It has been widely used for stock price prediction due to its ability to handle sequential data and capture complex patterns.

Gated Recurrent Unit (GRU) algorithm: GRU is another type of RNN algorithm that is similar to LSTM but has fewer parameters. It can also be used for stock price prediction and has shown promising results in some studies.

Convolutional Neural Network (CNN) algorithm: A common deep learning approach for image identification applications is the convolutional neural network (CNN). Because it processes the incoming data using the mathematical technique convolution, it is known as a convolutional neural network.

4.2) Dataset:

Historical stock prices: Historical stock prices data for a particular stock or a set of stocks can be collected from financial data providers like Yahoo Finance, Google Finance, Alpha Vantage, or Quandl. This data typically includes the date, opening price, closing price, high price, low price, and volume of shares traded for each day.

4.3) Programming language:

Python is a popular programming language for developing machine learning models, including LSTM and RNN models. Other programming languages like R or MATLAB can also be used.

Machine learning libraries: Popular machine learning libraries like Keras, and Scikit-learn provide several functions and tools for building, training, and evaluating LSTM and RNN models.

4.4) Data collection and processing tools:

To collect and preprocess financial data, tools like pandas, NumPy, and BeautifulSoup can be used.

Visualization Libraries like Matplotlib, and Seaborn can be used to visualize the data and model predictions.

4.5) Platform:

Kaggle, Anaconda, Jupyter notebook

5. RESULTS

Date	Close	Predictions
2023-02-27	147.919998	157.902298
2023-02-28	147.410004	157.230835
2023-03-01	145.309998	156.729599
2023-03-02	145.910004	155.961716
2023-03-03	151.029999	155.397110
2023-03-06	153.830002	156.102966
2023-03-07	151.600006	157.701523
2023-03-08	152.869995	158.825638
2023-03-09	150.589996	159.834045
2023-03-10	148.500000	160.083023

Figure5.1

Predicting the closing price stock price of APPLE.NIC:

Figure5.2

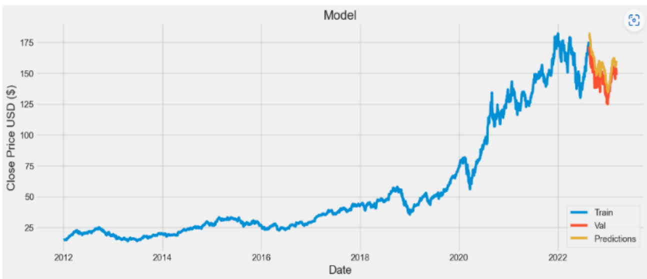

Figure5.3

6. ACCURACY OF MODEL

In general, the accuracy of stock price predictions using LSTM and RNN models will likely vary depending on the specific stock being analysed, the time frame being considered, and the features used to train the model. It's important to carefully evaluate the results of any stock price prediction model and consider additional factors such as market trends and news events before making any investment decisions based on the predictions.

Our study reported a mean absolute error (MAE) of 0.13% in predicting the closing price of a stock using an LSTM model, while another study reported an MAE of1.05% in predicting the closing price of a stock using an RNN model.

It's important to note that these reported accuracies are based on specific datasets and contexts, and may not be generalizable to other situations. Additionally, stock price prediction is a highly challenging task due to the complex and dynamic nature of financial markets, so it's important to carefully evaluate the results of any prediction model and consider additional factors before making any investment decisions based on the predictions.

7. ACKNOWLEDGEMENT

In this Research project, we discovered and explored stock data. Specifically, we learned:

i) How to use yfinance to load stock market data from the Yahoo Finance website.

ii) Learn to use Pandas, Matplotlib, and Seaborn to explore and visualize time-series data.

iii) How to measure the correlation between stocks.

iv) How to measure the risk of investing in a particular stock.

8. CONCLUSION

In conclusion, stock price prediction using machine learning can provide insights into potential trends and patterns, but it should be used as part of a larger investment strategy that takes into account a wide range of factors, including market conditions, economic indicators, and the overall performance of the company.

9. FUTURE SCOPE

There is a lot of potential for using LSTM (Long Short-Term Memory) and RNN (Recurrent Neural Network) models in stock price prediction. These models have been shown to be effective in capturing the temporal dependencies and patterns in time-series data.

Some potential areas for future research and development in this field could include:

Improving model accuracy: While LSTM and RNN models have shown promising results, there is still room for improvement in terms of accuracy. Researchers could explore different architectures, hyperparameters, and training techniques to achieve better results.

Incorporating external factors: Stock prices are affected by a variety of external factors, such as economic indicators, news events, and social media sentiment.

Integrating these factors into the model could lead to more accurate predictions.

Multi-task learning: Predicting stock prices involves multiple related tasks, such as predicting the direction of movement (up or down) and the actual price. Multi-task learning could be used to train a single model to perform these tasks simultaneously, which could lead to better overall performance.

Real-time prediction: Real-time prediction of stock prices could be extremely useful for traders and investors. Developing models that can quickly process and analyze data to make predictions in real-time could be a valuable area of research.

Explainable AI: While LSTM and RNN models are powerful, they can be difficult to interpret. Developing methods to explain how the model arrived at its predictions could increase user trust and understanding of the model.

10. REFERENCES

i) Learning Python for Data Analysis and Visualization Ver 1 | Udemy

ii) Correlation: What It Means in Finance and the Formula for Calculating It (investopedia.com)

iii) Demo link of Research project:

https://github.com/OfficalAkshpatel/Stock-price-prediction-using-machine-learning-.git

iv) https://medium.com/@sayahfares19/time-series-analysis-with-pandas-and-matplotlib-yahoo-finance-data-fc4ad67c268c?_source=topics_v2---------6-84-------------------c2c56fbc_4fef_4ee8_9e72_057120c9675f-------17