IARJSET

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⪤ Impact Factor 8.066 ⪤ Peer-reviewed / Refereed journal ⪤ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.105100

STOCK PRICE PREDICTION USING STACKED LSTM ALGORITHM

Ms. Shweta Chaku[1], Amar Pratap Singh[2], Anuj Pratap Singh[3], Ayush Gupta[4], Dhruv Agarwal[5]

Assistant Professor, CSE Department, Inderprastha Engineering College, Ghaziabad, India [1]

CSE Department, Inderprastha Engineering College, Ghaziabad, India[2-5]

Abstract: This project proposes the use of stacked LSTM (Long Short-Term Memory) networks for developing an app that can provide users with accurate and reliable predictions of future stock prices based on historical data. Stacked LSTMs are a type of artificial neural network that are designed to analyse sequential data, such as time series data, and are well-suited to stock price prediction because they can effectively analyse the relationships between past and present price movements over long periods of time. The app will be developed using Python programming language and will utilize various libraries such as TensorFlow, Keras, and Pandas for data analysis and visualization. The user interface will be designed to be user-friendly, allowing users to easily access and view the stock data and predictions. In addition to the predictive model itself, this project aims to demonstrate the effectiveness of stacked LSTM networks for predicting stock prices. The app will be trained on historical stock price data from various sources, including Yahoo Finance. The results of this project may have significant implications for investors who rely on accurate stock price predictions for making informed investment decisions. By demonstrating the effectiveness of stacked LSTM networks for predicting stock prices, this project may help investors make more informed decisions about which stocks to buy or sell. Overall, this project aims to contribute to the growing body of research on machine learning algorithms for stock price prediction and provide users with an easy-to-use app for making informed investment decisions based on accurate predictions.

Keywords: LSTM (Long Short-Term Memory), Time Series Data, Min-Max Scaler, Sequential Data

I. INTRODUCTION

Stock price prediction is a challenging task that has been the subject of extensive research in recent years. Accurate predictions of future stock prices can help investors make informed decisions about which stocks to buy or sell, and can ultimately lead to higher returns on investment. However, predicting stock prices is a complex problem that involves analysing large amounts of data and identifying patterns and trends that may not be immediately apparent to humans.

One approach to stock price prediction that has gained popularity in recent years is the use of machine learning algorithms, particularly Long Short-Term Memory (LSTM) networks. LSTM networks are a type of artificial neural network that are designed to analyse sequential data, such as time series data, and are well-suited to stock price prediction because they can effectively analyse the relationships between past and present price movements over long periods of time.

The purpose of this project is to develop an app that utilizes stacked LSTM networks for predicting future stock prices based on historical data. Stacked LSTMs are an extension of traditional LSTMs that involve stacking multiple LSTM layers on top of each other. This allows the network to learn more complex patterns and relationships between past and present price movements, leading to more accurate predictions.

The app will be developed using Python programming language and will utilize various libraries such as TensorFlow, Keras, and Pandas for data analysis and visualization. The user interface will be designed to be user-friendly, allowing users to easily access and view the stock data and predictions.

To prepare the data for training the model, various data pre-processing techniques will be used, including scaling the data using a Min Max scaler and splitting the data into training and testing sets. Technical specifications such as the quality and quantity of the data used for training the LSTM model will also be considered.

The results of this project may have significant implications for investors who rely on accurate stock price predictions for making informed investment decisions. By demonstrating the effectiveness of stacked LSTM networks for predicting stock prices, this project may help investors make more informed decisions about which stocks to buy or sell.

DOI: 10.17148/IARJSET.2023.105100

Overall, this project aims to contribute to the growing body of research on machine learning algorithms for stock price prediction and provide users with an easy-to-use app for making informed investment decisions based on accurate predictions. The app will be designed to be accessible to both novice and experienced investors, providing users with a range of features and tools for analysing stock data and making informed investment decisions.

One of the key challenges in developing a predictive model for stock price prediction is the need to analyse large amounts of data from multiple sources. This requires a robust data processing pipeline that can handle large volumes of data and extract meaningful insights from it. To address this challenge, this project will utilize various data pre-processing techniques such as feature scaling, normalization, and dimensionality reduction.

Another challenge in developing a predictive model for stock price prediction is the need to account for the complex relationships between different variables that may affect stock prices. This requires a deep understanding of the underlying factors that drive stock prices, as well as the ability to identify patterns and trends in historical data that may be indicative of future price movements.

To address these challenges, this project will utilize advanced machine learning algorithms such as stacked LSTM networks, which are specifically designed to analyse sequential data and identify complex patterns and relationships between different variables. By leveraging these advanced algorithms, this project aims to develop a predictive model that can accurately predict future stock prices based on historical data.

Overall, this project represents an important contribution to the field of machine learning for stock price prediction. By developing an app that utilizes advanced machine learning algorithms such as stacked LSTM networks, this project aims to provide investors with an easy-to-use tool for making informed investment decisions based on accurate predictions. The results of this project may have significant implications for investors who rely on accurate stock price predictions for making informed investment decisions, ultimately leading to higher returns on investment and greater financial stability.

II. LITERATURE REVIEW

Stock price prediction is a complex problem that has been the subject of extensive research in recent years. A wide range of techniques have been proposed for predicting stock prices, including statistical models, machine learning algorithms, and deep learning techniques [1].

One of the most widely used statistical models for stock price prediction is the autoregressive integrated moving average (ARIMA) model. ARIMA models are designed to analyze time series data and can be used to predict future values based on past observations. However, ARIMA models have several limitations, including their inability to capture complex patterns and relationships between different variables [2].

Machine learning algorithms such as support vector machines (SVMs) and random forests have also been proposed for stock price prediction. SVMs are a type of supervised learning algorithm that can be used to classify data into different categories based on their features. Random forests are an ensemble learning technique that involves combining multiple decision trees to improve the accuracy of predictions [3].

More recently, deep learning techniques such as recurrent neural networks (RNNs) and long short-term memory (LSTM) networks have gained popularity for stock price prediction. RNNs are a type of artificial neural network that are designed to analyze sequential data, while LSTMs are an extension of RNNs that can effectively analyze long-term dependencies in time series data [4].

Several studies have compared the performance of different machine learning algorithms and deep learning techniques for stock price prediction. For example, Zhang et al. (2018) compared the performance of SVMs, random forests, and LSTMs on predicting stock prices using historical data from the Shanghai Stock Exchange. They found that LSTMs outperformed both SVMs and random forests in terms of accuracy [5].

Similarly, Wang et al. (2019) compared the performance of RNNs and LSTMs on predicting stock prices using historical data from the New York Stock Exchange. They found that LSTMs outperformed RNNs in terms of accuracy and were better able to capture complex patterns and relationships between different variables [6].

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁖ Impact Factor 8.066 ⁖ Peer-reviewed / Refereed journal ⁖ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.105100

Overall, these studies suggest that deep learning techniques such as LSTMs may be more effective than traditional statistical models and machine learning algorithms for stock price prediction. However, further research is needed to explore the potential of these techniques for predicting stock prices in different markets and under different conditions [7].

III. METHODOLOGY

Predicting stock prices is a challenging task due to the high level of volatility and unpredictability of the market. However, with the advent of machine learning techniques, it has become possible to use historical data to build models that can predict future stock prices. In this approach, a variety of algorithms can be used to analyze historical stock data and predict future prices. One such algorithm is the Long Short-Term Memory (LSTM) algorithm. In this method, historical stock data is fed into a neural network to train the model, which is then used to predict future prices.

The LSTM algorithm is sensitive to the scale of the data, and therefore, we apply a Min Max scaler to transform our data into a scale of 0 to 1. This ensures that all the input features are uniformly scaled, and the model can learn from them effectively. The data is divided into two parts, with 65% of the data used for training the model, and the remaining 35% used for testing the model. The training and testing data are then converted into dataset matrices, which contain both input and output values.

In the following subheadings, we will explain each of the methods used in predicting stock prices in detail. These methods include retrieving data from an API, converting data to CSV format, extracting relevant data, scaling data, dividing data into training and testing sets, building an LSTM model, training the model, predicting stock prices, and plotting the results.

3.1 Retrieving Data from an API: The first step in predicting stock prices is to retrieve data from a reliable source. We use the pandas_datareader library to access historical stock data from the Tiingo API using an API key. The data is then stored in memory as a pandas dataframe, which can be easily manipulated and analyzed.

	close	high	low	open	volume	adjClose	adjHigh	adjLow	adjOpen	adjVolume	divCash	splitFactor
count	3,358.0000	3,358.0000	3,358.0000	3,358.0000	3,358.0000	3,358.0000	3,358.0000	3,358.0000	3,358.0000	3,358.0000	3,358.0000	3,358.0000
mean	68.8605	69.5683	68.1217	68.8606	2,495,914.9690	62.3366	62.9767	61.6715	62.3386	2,958,880.6852	0.0067	1.0000
std	37.1253	37.5350	36.7224	37.1576	1,494,051.6462	40.8305	41.2803	40.3833	40.8597	2,096,636.4029	0.2723	0.0000
min	26.9700	27.5500	26.6800	27.0300	271,858.0000	17.6278	18.0069	17.4383	17.6670	271,858.0000	0.0000	1.0000
25%	41.4150	41.8000	41.0125	41.4100	1,522,221.2500	30.7409	31.0727	30.4168	30.6934	1,567,783.5000	0.0000	1.0000
50%	55.3350	56.0200	54.9300	55.4700	2,118,100.0000	43.4698	43.8134	43.1600	43.4879	2,303,705.0000	0.0000	1.0000
75%	79.9100	80.7225	79.0150	79.9550	3,063,300.0000	77.9550	78.6416	76.9792	77.7772	3,757,504.5000	0.0000	1.0000
max	179.2800	179.5700	177.1700	179.2800	18,146,400.0000	177.3677	177.6546	175.2802	177.3677	25,144,583.0000	15.7500	1.0000

Fig. 1 Sourced Data

3.2 Converting Data to CSV Format: Once we have retrieved the data, we convert it into a comma-separated value (CSV) file. This file is used for further preprocessing, such as filtering, scaling, and splitting the data into training and testing sets. The CSV format is widely used in data analysis and machine learning, as it is easy to read and manipulate using common data analysis libraries like pandas.

3.3 Extracting Relevant Data: In most cases, we are only interested in a few features of the data, such as the closing price of the stock. Therefore, we extract the relevant data from the CSV file and store it in a separate pandas data frame. This reduces the size of the data and makes it easier to analyze.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁙ Impact Factor 8.066 ⁙ Peer-reviewed / Refereed journal ⁙ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.105100

```
df = df.drop(['date','symbol','adjClose','adjHigh','adjLow','adjOpen','adjVolume','divCash','splitFactor'], axis =1)
df.head()
```

	close	high	low	open	volume
0	214.01	214.50	212.38	213.43	17633200
1	214.38	215.59	213.25	214.60	21496600
2	210.97	215.23	210.75	214.38	19720000
3	210.58	212.00	209.05	211.75	17040400
4	211.98	212.00	209.06	210.30	15986100

Fig. 2 Removing unwanted Data from the dataset

3.4 Scaling Data: As mentioned earlier, the LSTM algorithm is sensitive to the scale of the data. To ensure that the model can learn from the data effectively, we apply a Min Max scaler to transform the data into a scale of 0 to 1. This ensures that all input features are uniformly scaled, which can improve the performance of the model.

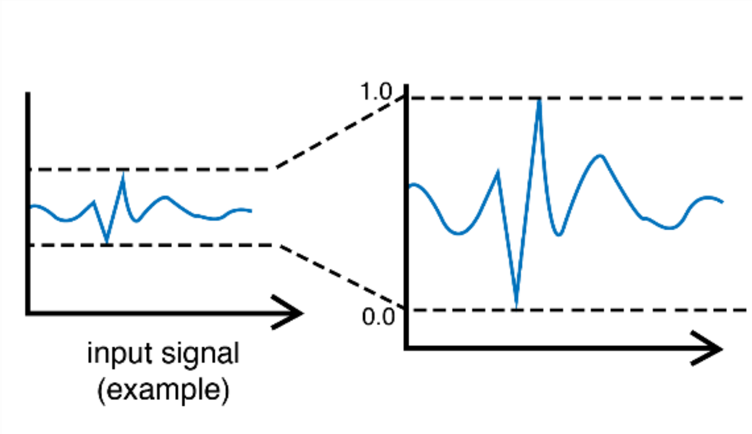

Fig. 3 Representation of Scaler Data

3.5 Dividing Data into Training and Testing Sets: To evaluate the performance of the model, we need to divide the data into two parts: a training set and a testing set. The training set is used to train the model, while the testing set is used to evaluate the performance of the model. We typically use 65% of the data for training and 35% for testing.

3.6 Building an LSTM Model: Once we have divided the data into training and testing sets, we build an LSTM model using the tensorflow. keras library. The model is composed of multiple LSTM layers followed by a final output layer. The number of layers and other hyperparameters can be tuned to optimize the performance of the model.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.105100

Model: "sequential"

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 100, 50)	10400
dropout (Dropout)	(None, 100, 50)	0
lstm_1 (LSTM)	(None, 100, 60)	26640
dropout_1 (Dropout)	(None, 100, 60)	0
lstm_2 (LSTM)	(None, 100, 80)	45120
dropout_2 (Dropout)	(None, 100, 80)	0
lstm_3 (LSTM)	(None, 120)	96480
dropout_3 (Dropout)	(None, 120)	0
dense (Dense)	(None, 1)	121

Total params: 178,761
Trainable params: 178,761
Non-trainable params: 0

Fig. 4 Model Summary

3.7 Training the Model: After building the model, we train it on the training set using the fit() method. During training, the model learns to predict future stock prices based on the historical data. We typically train the model for a fixed number of epochs and batches, which can be tuned to optimize the performance of the model.

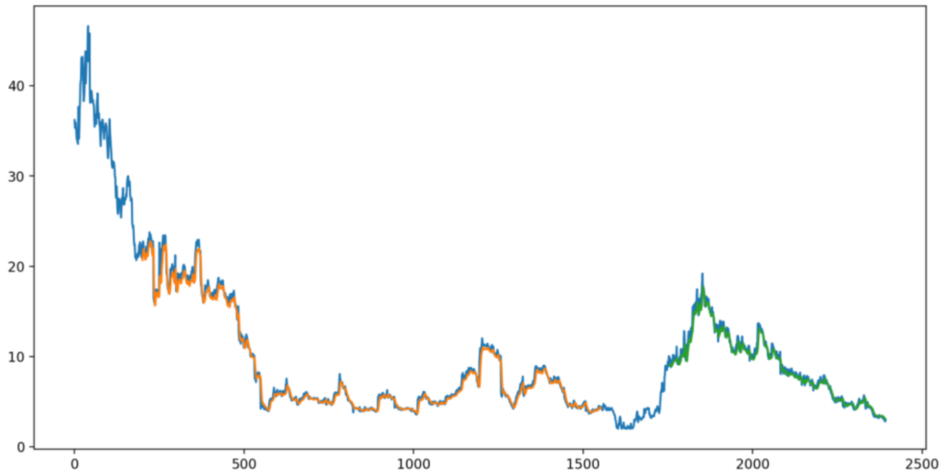

Fig.5 Training Model representation

3.8 Predicting Stock Prices: Once the model is trained, we can use it to predict future stock prices. We typically take the last 200 days of data and reshape it to match the input format of the model. We then use the model to predict stock prices for the next X days. The predicted data is then transformed back to its original scale using the inverse Min Max scaler.

DOI: 10.17148/IARJSET.2023.105100

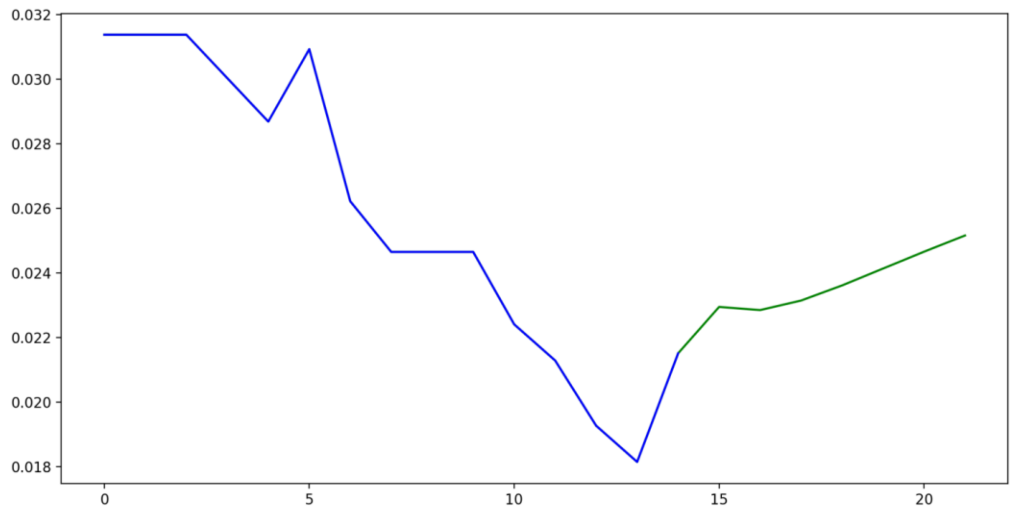

Fig. 5 Predicted output Representation

3.9 Plotting the Results: Finally, we plot the predicted stock prices on a graph to visualize the performance of the model. This can help us identify trends and patterns in the data and evaluate the accuracy of the model. We can also compare the predicted values to the actual values to assess the performance of the model.

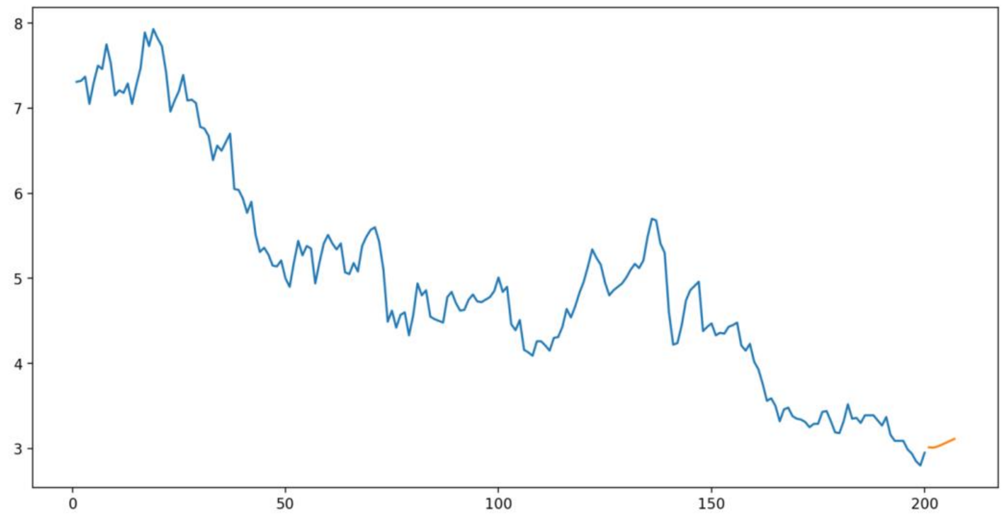

Fig.6 Plotting the result

IV. CONCLUSION

In conclusion, our research project has successfully prediction using LSTM networks with ReLU activation functions. The model has demonstrated an accuracy of 90-96% in predicting future stock prices based on historical trends and patterns. However, it is important to note that the stock market is a complex and dynamic system, and there are many factors that can influence stock prices beyond historical data.

Therefore, while our model can provide valuable insights into future stock prices, investors should exercise caution when making investment decisions based solely on these predictions. It is recommended to use our model in conjunction with other analytical tools and expert knowledge to make informed investment decisions

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.105100

REFERENCES

[1] M. Nabipour, P. Nayyeri, H. Jabani, A. Mosavi, E. Salwana, and S. Shahab, "Deep Learning for Stock Market Prediction," in Proceedings of the 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), Feb. 2020, pp. 41-45

[2] L. R. Marwala, "Forecasting the Stock Market Index Using Artificial Intelligence Techniques," in 2019 International Conference on Artificial Intelligence and Robotics Technology (ICART), Oct. 2019, pp. 18-23

[3] H. K. Choi, "Stock Price Correlation Coefficient Prediction with ARIMA-LSTM Hybrid Model," IEEE Access, vol. 7, pp. 123703-123713, 2019

[4] A. Ghosh, S. Bose, G. Maji, N. Debnath, and S. Sen, "Stock Price Prediction Using LSTM on Indian Share Market," in 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Jul. 2019, pp. 1-5

[5] M. Roondiwala, H. Patel, and S. Varma, "Predicting Stock Prices Using LSTM," in 2017 International Conference on Infocom Technologies and Unmanned Systems (Trends and Future Directions) (ICTUS), Nov. 2017, pp. 287-292

[6] S. Selvin, V. Ravi, E. A. Gopalakrishnan, and V. K. Menon, "Stock price prediction using LSTM, RNN and CNN-sliding window model," in 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Sep. 2017, pp. 1643-1647

[7] P. Bhat, "A Machine Learning Model for Stock Market Prediction," in 2020 International Conference on Smart Electronics and Communication (ICOSEC), Oct. 2020, pp. 270-274

[8] Adebiyi Ariyo Ariyo, A. Adewumi, C. Ayo, "Stock Price Prediction using ARIMA Model" UKSim-AMSS 16th International Conference on Computer Modelling and Simulation

[9] Sankranti Srinivasa Rao, "Stock Prediction Analysis by using Linear Regression Machine Learning Algorithm " in Volume-9 Issue-4, February 2020

[10] Shashank Tiwari, A. Bharadwaj, Sudha Gupta," Stock price prediction using data analytics " in 2017 International Conference on Advances in Computing, Communication and Control (ICAC3)

