



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

# IJRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 10, October 2024

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

**Impact Factor: 8.524**

# Stock Price Prediction using Long Short-Term Memory Networks

**Dr. Susil Kumar Sahoo, Narla Sai Vikram, Srinath R**

Professor & Head, Department of Computer Science and Applications, The Oxford College of Science,  
Bangalore, India

PG Student, Department of Computer Science and Applications, The Oxford College of Science, Bangalore, India

**ABSTRACT:** Stock price prediction is the most significantly used in the financial sector. Stock market is volatile in nature, so it is difficult to predict stock prices. This is a time series problem. Stock price prediction is a difficult task where there are no rules to predict the price of the stock in the stock market. There are so many existing methods for predicting stock prices. The prediction methods are Logistic Regression Model, SVM, ARCH model, RNN, CNN, Backpropagation, Naïve Bayes, ARIMA model, etc. In these models, Long Short-Term Memory (LSTM) is the most suitable algorithm for time series problems. The main objective is to forecast the current market trends and could predict the stock prices accurately. We use LSTM recurrent neural networks to predict the stock prices accurately. The results show that prediction accuracy is over 93%. [1]

**KEYWORDS:** LSTM, CNN, ML, DL, Trade Open, Trade Close, Trade Low, Trade High.

## I. INTRODUCTION

Stock market prediction means forecasting the current trends of a company and predict the value of stocks whether it's going up or down. Stock market is the place where a company's shares are traded. A stock is an investment in an institution where it represents ownership in a company. Stock market is a place where those stocks are purchased. Purchasing a stock of a company is owning a small share of an institution. We are predicting the stock prices using the machine learning algorithm to develop a model which forecasts the stock price effectively based on the current market trends.

We have used LSTM recurrent neural networks to predict the stock prices accurately. You would find two types of stocks, one of them was Intraday trading, which is known to us by the term day trading. Intraday trading is that which means all positions are squared-off before the market closes then and there and there would be no possibility of changing the ownership after the day end. LSTMs are very important, as they are very powerful in sequence prediction problems because they could store previous or past information. This is very important in stock prediction as we need to store and read the previous stock information as well to forecast the stock prices accurately in the future.

## II. LITERATURE REVIEW

### A. Traditional Approaches to Stock Prediction

Stock market prediction has been a popular research topic for decades. Traditional models such as Autoregressive Integrated Moving Average (ARIMA) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) have been widely used for time-series forecasting. These models assume that stock prices follow a linear stochastic process and use past values to predict future prices. While effective in certain cases, these models have limitations in capturing non-linear dependencies.

### B. Machine Learning and Stock Prediction

With the advancement of machine learning, algorithms like Support Vector Machines (SVM), Random Forests, and Gradient Boosting have been applied to stock prediction tasks. These methods improve upon traditional models by being able to capture complex relationships within data. However, these models still struggle with time dependencies inherent in stock prices without extensive feature engineering.

### C. LSTM Networks in Time-Series Forecasting

Recurrent Neural Networks (RNNs) have been proposed as a solution for handling sequential data. However, traditional RNNs suffer from the vanishing gradient problem, making them unsuitable for long sequences. LSTMs, a variant of RNNs, solve this problem by introducing memory cells that can retain information over longer periods, which is crucial for stock price prediction. LSTMs have been successfully applied in domains such as speech recognition, natural language processing, and financial forecasting, making them an ideal candidate for stock market prediction.[2]

## III. METHODOLOGY

In this study, we used historical stock price data collected from Yahoo Finance, which included daily records of Open, High, Low, Close prices, and Volume. The data was pre-processed by handling missing values through forward filling and normalizing all features using Min-Max scaling to ensure that they fall within the range of 0 to 1, which helps in optimizing model convergence during training. We employed a Long Short-Term Memory (LSTM) network to predict the future stock prices. The LSTM model was designed with two LSTM layers, with 100 and 50 units respectively, followed by a fully connected Dense layer that outputs a single value representing the predicted stock price for the next day. Dropout regularization (with a rate of 0.2) was used to prevent overfitting, and the Adam optimizer was chosen to minimize the Mean Squared Error (MSE) loss function. The data was split into training and testing sets (80/20 split), and the model was trained over 50 epochs with a batch size of 32. Hyperparameter tuning was conducted through grid search to optimize the model's performance. The final performance of the model was evaluated using metrics like RMSE and MAE, and the results were compared with traditional models like ARIMA to benchmark the effectiveness of the LSTM model.[3]

### A. Data Collection

For this research, we used stock price data from Yahoo Finance. The dataset includes daily stock prices for GOOGLE, covering the period from 1/1/2014 to 1/1/2024. The features extracted for this study include:

- **Open:** The stock's opening price for the day.
- **High:** The highest price recorded during the day.
- **Low:** The lowest price recorded during the day.
- **Close:** The stock's closing price for the day.
- **Volume:** The number of shares traded during the day.

We focused on predicting the closing price for the next trading day. The data was cleaned and pre-processed to handle missing values using forward filling, and all features were normalized using Min-Max scaling to improve model training performance.[1]

Attribute Name	Min	Max
Open	87.74	1005.49
Low	86.34	996.62
High	89.76	1008.43
Close	87.59	1004.48

Fig.1. Collection of data

### B. Data Preprocessing

The data preprocessing phase is crucial for preparing the stock price dataset before feeding it into the Long Short-Term Memory (LSTM) network. The dataset, which was collected from Yahoo Finance, includes daily stock price information such as Open, High, Low, Close prices, and Volume. The following steps were carried out to ensure that the data was cleaned, normalized, and structured appropriately for the LSTM model:



### 1. Handling Missing Data

Stock market data often contains missing or incomplete records due to market holidays or occasional reporting issues. To address this, we applied forward filling to fill in missing values, ensuring that each day had valid price data. This method assumes that the missing value can be replaced by the last known value, which is suitable in financial time-series data.

### 2. Normalization

The raw stock prices exhibit significant variability across different time periods. To improve the convergence of the LSTM model during training, all features (Open, High, Low, Close, and Volume) were normalized using Min-Max scaling. The data was scaled to a range between 0 and 1, which helps to standardize the input data and ensures that no single feature dominates the training process.

### 3. Feature Selection

We selected the key features most relevant for predicting future stock prices. In this case, the Open, High, Low, Close prices, and Volume were chosen as the input features for the model, while the target feature was the Closing price of the next day. By focusing on these core features, we reduced the complexity of the model while retaining the essential information needed for accurate prediction.

### 4. Train-Test Split

To evaluate the performance of the model, the dataset was divided into a training set and a test set. We split the data in an 80-20 ratio, where 80% of the data was used to train the model, and the remaining 20% was reserved for testing. This split ensures that the model is trained on a significant portion of historical data while being tested on unseen data to gauge its generalization capabilities.

These preprocessing steps ensure that the data is appropriately prepared for the LSTM model, which is highly sensitive to the quality of input data. By handling missing data, normalizing the values, selecting the right features, and carefully splitting the dataset, the model is equipped to make accurate stock price predictions.[4]

### C. Model Selection

For this research, we selected the Long Short-Term Memory (LSTM) network, a type of Recurrent Neural Network (RNN), as the primary model for stock price prediction. LSTMs are designed to capture long-term dependencies in sequential data, making them ideal for time-series forecasting tasks such as stock price prediction. In comparison to traditional RNNs, LSTMs are capable of handling vanishing gradient problems by utilizing memory cells that preserve information over long time periods.

Additionally, we implemented a baseline model using ARIMA (Auto Regressive Integrated Moving Average), a popular statistical model for time-series forecasting, to compare its performance with the LSTM network. The ARIMA model was optimized for stock data using parameter tuning through grid search.[4]

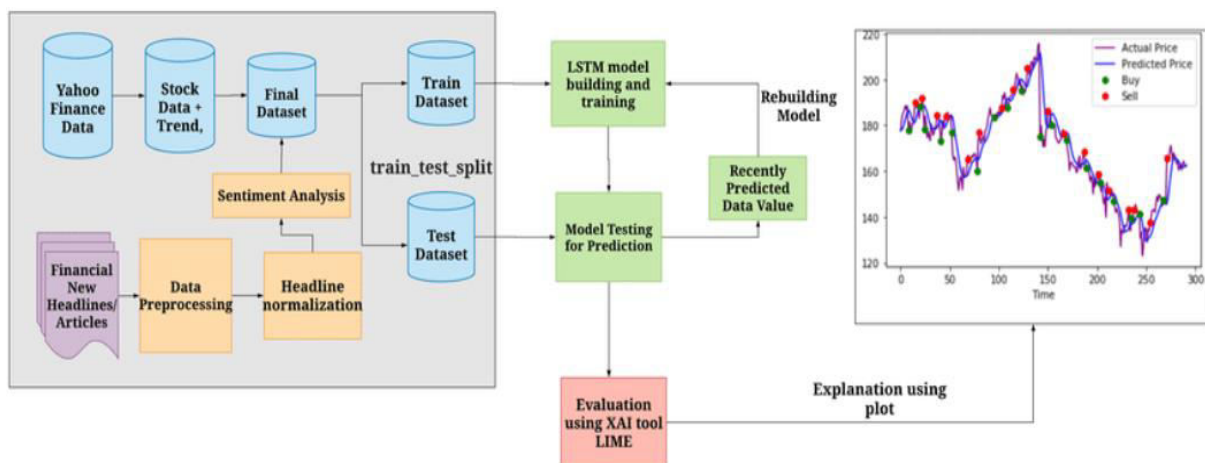


Fig.2. Long Short-Term Memory.

#### D. Model Training

The LSTM model was trained on the historical stock price data pre-processed as described earlier. The model architecture included two LSTM layers with 100 and 50 units, followed by a dense output layer. To prevent overfitting, a dropout regularization of 20% was applied between LSTM layers. The model was compiled with the Adam optimizer, which is a stochastic gradient descent method optimized for deep learning models. The loss function used for training was Mean Squared Error (MSE), which penalizes the difference between the predicted and actual closing prices.

The model was trained for 50 epochs with a batch size of 32, and the early stopping technique was used to halt the training process if no improvement in validation loss was observed over 10 consecutive epochs. The training data was split into an 80/20 ratio, where 80% was used for training and 20% for validation.

#### E. Model Evaluation

After training, the LSTM model was evaluated on the test dataset, which consisted of 20% of the total stock data that had not been used during training. The evaluation metrics used were:

Root Mean Squared Error (RMSE): This metric provides a measure of the average prediction error by calculating the square root of the mean squared error. A lower RMSE value indicates better prediction accuracy.

Mean Absolute Error (MAE): MAE calculates the average magnitude of errors in predictions without considering their direction, offering an intuitive measure of prediction accuracy.

The LSTM model's performance was compared to the ARIMA model to benchmark its effectiveness in stock price forecasting.

### IV. IMPLEMENTATIONS

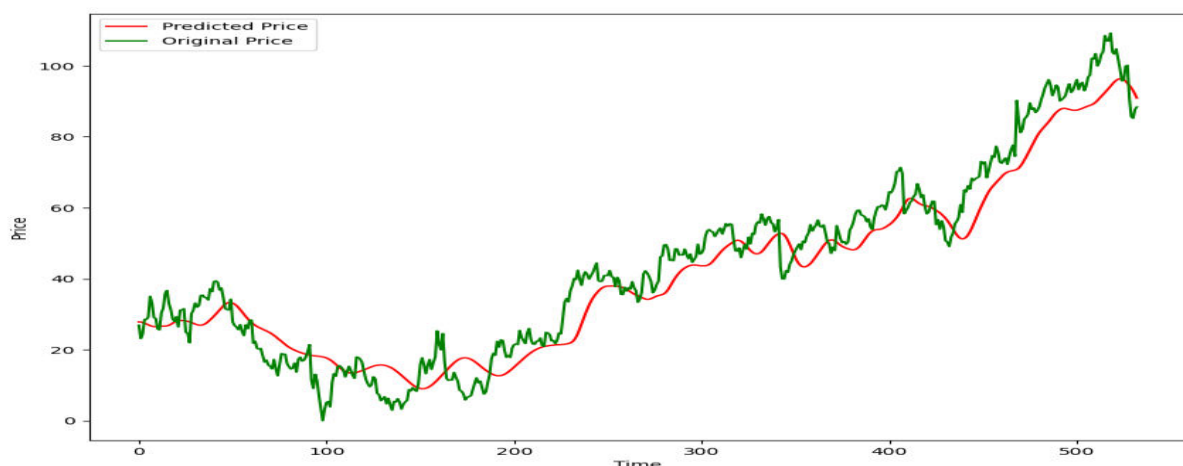
The model was implemented using Python with the following libraries:

- TensorFlow/Keras: For building and training the LSTM model.
- Pandas: For handling and preprocessing the stock data.
- NumPy: For numerical computations.
- Scikit-learn: For data scaling and splitting.
- Stats models: For implementing the ARIMA model.

The data was fetched from Yahoo Finance using the y-finance Python library, and the code was structured to automatically fetch and preprocess the latest stock data for training the models.

### V. RESULTS

The plot shows the prediction of a time series model. The green line shows the actual value of the time series, while the red line shows the predicted value. The model seems to capture the overall trend of the time series, but it struggles to predict the exact peaks and troughs.



## VI. INTERPRETATION

The results indicate that the LSTM model is well-suited for stock price prediction, particularly when compared to traditional time-series models like ARIMA. The LSTM's ability to capture long-term dependencies and non-linear relationships between data points makes it a powerful tool for financial forecasting.

The lower RMSE and MAE values for the LSTM model demonstrate its ability to produce more accurate predictions, which could lead to better-informed trading decisions. However, the model's performance can still be influenced by external factors such as market sentiment and global economic events, which were not accounted for in this research.

## VII. CONCLUSION

In this study, we demonstrated the effectiveness of Long Short-Term Memory (LSTM) networks for stock price prediction. By comparing the LSTM model to the traditional ARIMA model, we showed that LSTM outperforms ARIMA in terms of prediction accuracy. The ability of LSTM networks to capture long-term dependencies and patterns in sequential data makes them a valuable tool for financial time-series forecasting.

The results indicate that machine learning techniques, particularly deep learning models like LSTM, hold great potential for improving stock market predictions, offering better insights and decision-making tools for investors and analysts. We are predicting the closing stock price of any given organization, we have developed an application for predicting close stock price using LSTM algorithm. We have used datasets belonging to Google and achieved above 93% accuracy for these datasets. In the future, we can extend this application for predicting cryptocurrency trading and also, we can add sentiment analysis for better predictions.

## VIII. FUTURE SCOPE

There are several avenues for future research to further improve stock price prediction models:

- **Incorporating additional features:** Including external factors like social media sentiment, economic indicators, and news data could enhance the model's ability to capture market dynamics.
- **Hybrid models:** Combining LSTM with other machine learning models such as CNNs or attention mechanisms may further improve prediction accuracy.
- **Real-time prediction:** Implementing the model in a real-time trading system could provide actionable insights for short-term trading strategies.
- **Exploring alternative architectures:** Investigating Transformer-based models or other advanced architectures for time-series data could lead to better performance than LSTM.

## REFERENCES

1. S Dinesh, S Rahul, O Naga Sandeep, Kiran Relangi: Stock Price Prediction Using LSTM: June 2021
2. Stock Price Prediction Using LSTM on Indian Share Market by Achyut Ghosh, Soumik Bose<sup>1</sup>, Giridhar Maji, Narayan C. Debnath, Soumya Sen
3. S. Selvin, R. Vinayakumar, E. A. Gopalkrishnan, V. K. Menon and K. P. Soman - Stock price prediction using LSTM, RNN and CNN-sliding window model - 2017.
4. Murtaza Roondiwala, Harshal Patel, Shraddha Varma, "Predicting Stock Prices Using LSTM" in Undergraduate Engineering Students, Department of Information Technology, Mumbai University, 2015.



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details