Stock Price Forecasting Using LSTM Algorithm

Nagarjun H M[1], Mohan Kumar H P[2]
[1, 2] Dept of MCA
[1, 2] PESCE Mandya, Karnataka, India

Abstract- *Investing in any financial assets including stock market is an extremely hard and complex task. The stock market is also heavily influenced by the external factors such as elections , interest rate hikes etc[8]. Because of all these factors, predicting the stock price is difficult. To forecast a specific stock's next thirty days, we'll use the LSTM algorithm. Basically, we predict the following day's data using the previous day's data, and we'll forecast the price for the next thirty days using the sliding window process.*

Keywords- Stock Market, LSTM algorithm , Price Prediction , RNN algorithm, Nifty50,cupid,NiftyBank,Deep learning

I. INTRODUCTION

Stock Market of a country can be used as a indicator to judge the economy of that country. Through stock market involvement, people might significantly increase their financial profits or significantly decrease them. Stock Market in majority consists of buyers and sellers ,for our simple understanding we can think it like an supply and demand of a product. As the supply of a product goes up and the demand for that product reduces , the price for that product will automatically reduces. And if the demand is high and the supply is low the the price goes high because of the limited availability. In the same way the stock market works , if the supply goes higher then the price comes down if the demand is high then the price goes up. And it is quite difficult to be on the right side of the trade. Based on the time frame you chose you will classified into different types of traders[2].In this study, we use the LSTM algorithm to predict the stock price for the next thirty days. Predicting a set of time series is mostly done to build simulation models of future values based on their historical values.

II. LITERATURE SURVEY

[1]   discusses the use of the LSTM model to forecast stock prices in addition to using the algorithm to forecast the price of the Tata Motors stock, which is traded on both the NSE and BSE stock exchanges.

In [2], here the paper discusses about the significant impact of both the length of the data used for the predicting

the future price of the financial asset ,the number of epochs used to train the model.

In [3] article, the LSTM method was used to the stock of Alcoa Corp. The model was trained using different bulk sizes, including 512 and 50 events, 64 and 100 events, and 256 and 50 events, and the forecasts were highly accurate.

In[4] paper they have implemented the RNNLSTM on stocks in order to predict the future prices, here they have used different data points to train and test the data. These different data points are open/close , open/close , open/close/high/low and RSME scores are compared with each other.

An attention-based long short-term memory model to forecast stock price trend is proposed in paper [5]. The input layer, hidden layer, attention layer, and output layer are the four components of the model. In order to comply with the model's input specifications, the input layer cleans the input data. Through the LSTM unit, the hidden layer is linked to the line model network. The feature vector was weighted by the attention layer. The calculated results are sent to the output layer. The gradient descent algorithm is used to solve the model training.

The time series behaviour of the India VIX and its relationship with various equities portfolio returns are examined in [6] research, along with the time-series characteristics of the VIX and its links with Nifty index performances.

They looked into how stability differences affected LSTM stock price prediction in paper [7]. The predictability of experimental outcomes has only a small impact, but it has a small bearing on how quickly the algorithm converges. The speed of convergence will increase as the P value increases. Finally, they ran an experiment to compare the performance of the LSTM model and the ARIMA model, and came to the conclusion that the LSTM algorithm performs better in prediction and has less mistakes.

The historical data of the Chinese stock market were converted into 30 day sequences with 10 features and 3 day earning rate labelling for the purpose of price prediction in

[10] article. Through training on 900000 sequences, the model was fitted.

In [11] paper, they used.Net to implement their solution, and Microsoft Visual Studio 2012 to design the user interface. They used a back propagation technique to train the data set, and when the error reached a predetermined level, the training was terminated. With the help of the trained data set, they were able to estimate the value of the stock.

In [12], the author proposes a multilayer LSTM model to forecast stock price and stock returns and uses NLP technical tools to analyse and quantify public sentiment. They then compare the predictive accuracy of the model using both news sentiment score and historical stock technical indicators as input and the model based solely on historical stock technical indicators.

In [13] the stock prediction is done using the regression models,the LSTM and bi-directional LSTM and comparison has been made between these two models also it has been found there is a requirement of fine tuning the parameters in order to get a low RSME value.

In [14] in this paper they have mainly focused on LSTM model to reach the accuracy of 60- 65% range , They considered the single dimensional and multi dimensional LSTM for the comparison purpose and have tried to get the optimal price prediction model.

In [15] in this paper they have built the two model using the LSTM, one model is univariant and the other model is the multivariant LSTM, it is seen in this paper that the multivariant has lower MSE value , lower the MSE precise the model is.

In [16] paper they have used 3 measuring techniques RMSE ,MAPE, MSE to measure the accuracy of the prediction also they have experimented with different number of hidden layers , different number of epochs to get more accurate model for price prediction

III. METHODOLOGY

Our data consists of day wise stock price of 1500 data points, each data points one trading day.It includes date , open , high , low , close , volume .Here we have taken the cupid stock which comes under the NIFY_50 index . To acquire the stock price, pre-processing was applied to the data of this stock after it was retrieved from the available data.

This work uses the sliding window method for estimating mid-term prices. The window size is set at 100 periods, which translates to 100 days with a 70-day overlap in information and projection for the next 30 days. Price of the cupid stock ranges from 2000 to 3000 , since the LSTM model is quite sensitive of the input data that is feed into the we need to normalize the data from 0 - 1 range .

Steps for implementing the LSTM algorithm

Input : Historic stock data

Output : Predicting the price for next 30 days

Step 1 : Start.

Step 2 : Using the nsepy Python library ,data is obtained and then preprocessing is done   on that data and then it normalized   between 0 to 1.

Step 3 : Splitting the data set into 65% training and 35% for the testing purposes.

Step 4 : Next step is to build the LSTM model and then pass train data and test data to the model, fig. 1.1 provides a summary of the model.

Step 5 : Adding the Adam optimization and mean squared error during LSTM compilation.

Step 6 : using chart tools to visualize the results and make predictions.

The figure below represents the summary of our LSTM model which has been built using 100 epochs and 3 LSTM layers
Fig 3.1

```
Model: "sequential"
_____
 Layer (type)                Output Shape              Param #
=================================================================
 lstm (LSTM)                 (None, 100, 50)           10400

 lstm_1 (LSTM)               (None, 100, 50)           20200

 lstm_2 (LSTM)               (None, 50)                20200

 dense (Dense)               (None, 1)                 51

=================================================================
Total params: 50,851
Trainable params: 50,851
Non-trainable params: 0
```
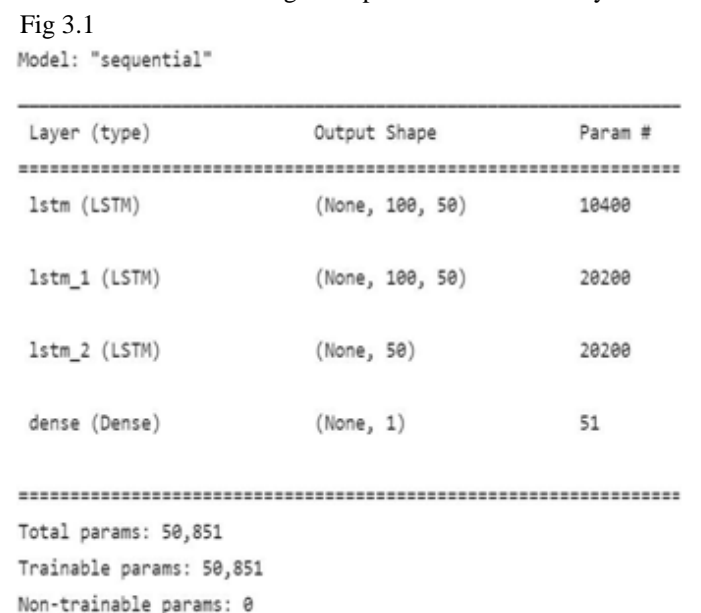
Fig 3.1

Also in this application we have included the two other methods of analysis that traders use, they are the pre-open market and post-market data.

Pre-open market data

The pre-open market is the market window which is before the normal market timings.The duration of the pre-open market is from 9:00 AM to 9:15 AM that is 15 min before the trading session starts and is conducted on both major Indian stock exchanges that is NSE and BSE.

The information that we can extract from the pre open market is the stocks that are about to gap up and gap down .so in this application users can find the stocks that going to open either gap up more than 1.5 percent or gap down more than 1.5 percentage and also it displays the number of advances ,declines , unchanged and this information will be updated at 9:10 AM in NSE website which is then reflected in our application.The fig 3.2 below represents how the GUI looks like

Fig 3.2

Post market data

Post market data is the data we receive after the end of the trading session.This data basically includes the summary of the trading session like top gainer , top loser of the trading day.The fig below represents how the GUI looks like all of those are described in the below table 3.3.

Table 3.3

	RSME of training data	RSME of testing data
LSTM with epochs 50	203.46	234.43
LSTM with epochs 100	202.8	231.68
LSTM with epochs 150	195.89	223

It is also seen that as we increase the number of epochs the RSME value decrease along with the increase in time required to train the model.

After implementing the model, the below graph fig 3.4 shows how the training and testing data is plotted against the original close price.
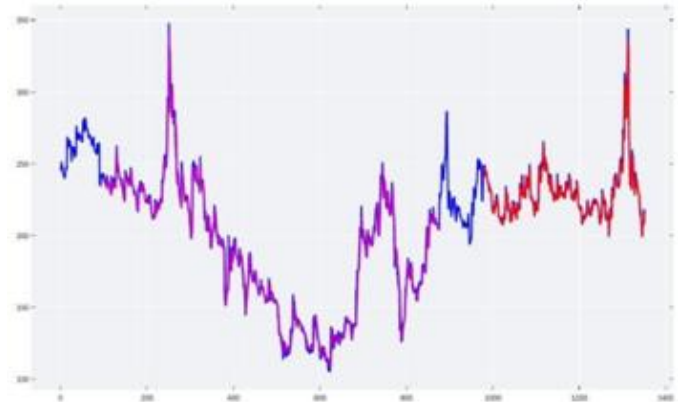

Fig 3.4

And then we have used the sliding window in order to predict the price for the next 30 days of the cupid stock.Fig 3.5

Prediction of future stock price and converting it into a web app by using streamlit library is a very simply and efficient way which is also mentioned in [9], we have built this application using the streamlit library.

IV. RESULT

Implementation of the suggested LSTM model that uses past data to forecast the share price of the particular stock. Below is a visual representation of the stock price of cupid. In our paper the implementation of the algorithm is done in such a way that it uses the previous 100 days of data to predict the value of the next 30 days. Our algorithm's predictions for the cupid share are displayed in the graph below. The results of our algorithm in structured form utilizing LSTM are displayed in the graph below. Here, the

accuracy of our model is assessed using the RSME, or Root Mean Square Error.

Basically we consider 3 different LSTM model with 50 epochs, 100 epochs and 150 epochs. We can see
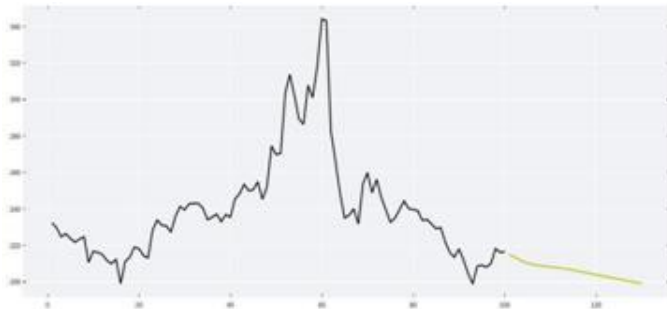

Fig 3.5

In the fig 3.5 the black line represents the past 100 days of data and the yellow line represents the future 30 days data of the cupid stock.

What we have also found is that the father into the future you predict, the price it becomes less and less accurate.

For example:consider we are predicting price based on previous 5 data points .
[a , b , c , d , e] These are 5 data point using which LSTM algorithm predicts the 6th data point.
[a , b , c , d , e] = f.
So now the LSTM model uses the latest 5 values to predict the next future value
[b , c , d , e , f] = g

Here we can see that f is already a predicted value and we are using that value to predict the future value , because of which the father we are going to predict the price the less accurate it becomes.

V. CONCLUSION

In this study we have considered a cupid stock and the LSTM model has be implemented to predict the future 30 days data and also the model can be implemented to any and all the financial assets that are traded. Instead of trying to predict the exact price of the stock rather we should be interested in finding the general trend of the assets. And for the future enhancement we can find the future volatility of the asset and also to identify the sentiment of the market through social media.

REFERENCES

[1] Pramod, & Pm, Mallikarjuna. (2021). Stock Price Prediction Using LSTM. Test Engineering and Management. 83. 5246-5251.

[2] Moghar, A., & Hamiche, M. (2020). Stock Market Prediction Using LSTM Recurrent Neural Network. In Procedia Computer Science (Vol. 170, pp. 1168–1173). Elsevier B.V. https://doi.org/10.1016/j.procs.2020.03.049

[3] Roondiwala, Murtaza, Harshal Patel, and Shraddha Varma. "Predicting stock prices using LSTM." International Journal of Science and Research (IJSR) 6.4 (2017): 1754-1756.

[4] Roondiwala, Murtaza & Patel, Harshal & Varma, Shraddha. (2017). Predicting Stock Prices Using LSTM. International Journal of Science and Research (IJSR). 6. 10.212

[5] D. Wei, "Prediction of Stock Price Based on LSTM Neural Network," 2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM), 2019, pp. 544-547, doi:10.1109/AIAM48774.2019.00113.75/ART20172755.

[6] Mall, Manmohan & Mishra, S & Mishra, P.K. & Pradhan, Bibhuti B. (2014). A STUDY ON RELATION BETWEEN INDIA VIX AND NIFTY RETURNS. Intercontinental Journal of Banking, Insurance and Finance. 1. 01-07.

[7] F. Qian and X. Chen, "Stock Prediction Based on LSTM under Different Stability," 2019 IEEE 4th International Conference on Cloud Computing and Big Data Analysis (ICCCBDA),2019, pp. 483-486, doi: 10.1109/ICCCBDA.2019.8725709.

[8] Chu, Thuy & Thuy, Thu. (2020). ANALYSIS OF INTERNAL AND EXTERNAL FACTORS AFFECTING TO THE STOCK MARKET PRICE OF NON- FINANCIAL JOINT STOCK COMPANIES LISTED ON THE VIETNAM'S STOCK MARKET.

[9] S. Shukla, A. Maheshwari and P. Johri, "Comparative Analysis of Ml Algorithms & Stream Lit Web Application," 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), 2021, pp. 175-180, doi: 10.1109/ICAC3N53548.2021.9725496.

[10] K. Chen, Y. Zhou and F. Dai, "A LSTM-based method for stock returns prediction: A case study of China stock market," 2015 IEEE International Conference on Big Data (Big Data),2015,pp. 2823-2824, doi:10.1109/BigData.2015.7364089.3.

[11] Dhiraj Mundada, Gaurav Chhaparwal, Sachin Chaudhari, and TruptiBhamare― Stock Value Prediction System,‖ International Journal on Recent and Innovation Trends in Computing and Communication, April 2015

[12]   Y. Guo, "Stock Price Prediction Based on LSTM Neural Network: the Effectiveness of News Sentiment Analysis," 2020 2nd International Conference on Economic Management and Model Engineering (ICEMME), 2020, pp. 1018-1024, doi: 10.1109/ICEMME51517.2020.00206.

[13]   M. A. Istiake Sunny, M. M. S. Maswood and A. G. Alharbi, "Deep Learning-Based Stock Price Prediction Using LSTM and Bi-Directional LSTM Model," 2020 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES), 2020, pp. 87-92, doi: 10.1109/NILES50944.2020.9257950.

[14]   Y. Wang, Y. Liu, M. Wang and R. Liu, "LSTM Model Optimization on Stock Price Forecasting," 2018 17th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES), 2018, pp. 173-177, doi: 10.1109/DCABES.2018.00052.

[15]   J. Du, Q. Liu, K. Chen and J. Wang, "Forecasting stock prices in two ways based on LSTM neural network," 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), 2019, pp. 1083-1086, doi: 10.1109/ITNEC.2019.8729026.

[16]   T. Damrongsakmethee and V. -E. Neagoe, "Stock Market Prediction Using a Deep Learning Approach," 2020 12th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), 2020, pp. 1-6, doi: 10.1109/ECAI50035.2020.9223142.