STOCK PRICE PREDICTION USING LSTM

Sahil Chauhan
School of Computing Scienceand Engineering
Galgotias UniversityGreater Noida,
*Uttar Pradesh, India*sahilchauhan.sc38@gmail.com

Hitesh Shukla
School of Computing Scienceand Engineering
Galgotias University
Greater Noida,
Uttar Pradesh, India
shuklahitesh99999@gmail.com

***Abstract*— As we all know that the stock market is a very difficult market for investment and needs strong brainstormingbefore one shall start investing money in it. The field of stock price analysis requires immense research and deep knowledge of finance. Stock Market Analysis is one of the leading use of fundamentals of machine learning. [1]**

It is very difficult to find the future price of stock market because of the volatile nature of the market. Our work focuses on the LSTM model to predict the prices using strategic parameters of stock market like RMSE and MAPE. The lowerthe values of the parameter shows the higher accuracy of our model.

Keywords—Stock market, stock price analysis, machine learning technique, LSTM

I. INTRODUCTION

Stock Price Prediction using LSTM helps you to find out the future price of a particular company or an organization's stock. It is mandatory for the company to be listed on the verified stock exchange for any kind of analysis. The traditional integrated concept of predicting stock prices is toboost convincing benefits. Predicting future trends of a particular stock is a difficult task to do. [2]

The goal and sole purpose of this research paper are to investigate the drawback of the current system and build a robust model that would benefit everyone by implementing more efficient algorithms. These algorithms are purposely supposed to work on the time-series model of implementing algorithms.

A stock market is a place where the stocks or other financialassets like bonds, funds, etc. of public companies are traded. Each stock is dedicated to one company, although these companies can be a part of one group or work under a parent company. A prime example in the Indian Stock Exchange is;

1) **Adani Group**, which has varieties of companies like Adani Power, Adani Enterprises, etc. working under one parent group. Each company has its own stock and its stock price. This group has a bunch of stocks tradingin the market. Tata Group can be another example of this type of working methodology.

2) **Reliance Industries** has various companies like Reliance Retail, Reliance Life Sciences, Reliance Logistics, etc. working under one parent company. Although it has varieties of companies working but noone has individual stock trading in the market.

The stock price prediction model which we are building stores the previous data of stock rates of a particular company and applies the RNN technique of the machine learning named LSTM. The approach is to process the required past data of a particular stock and predict that stockon appropriate attributes.

The major attributes of any stock are the Close price, Open price, one-day Low, one-day High, the price of previous day, the trading date, the trading quantity and the total turnover. The stock price prediction model uses time series analysis to

get the predicted price of a particular stock for a required duration of time. Our model will consider the companies listed in yahoo finance.

II. LIBRARIES & MACHINE LEARNING MODEL

LIBRARIES

A. NumPy: NumPy is one of the free to use Python librarythat is used for analyzing the arrays. It can also work with linear algebra, Fourier transform, matrices, etc. It was developed and introduced by Travis Oliphant in 2005. [3]

B. Pandas: Pandas is one of the popular Python packages which can work with fast and flexible data structures. It works efficiently with data structures which are specially designed to make work with "relational" or "tagged" data. It is a basic, high-level building block to perform real-world data analysis using Python.

C. Matplotlib: Matplotlib is one of the data visualization libraries in Python to work with 2D boxplots. Matplotlib is specifically built on NumPy arrays. It was developed and introduced by John Hunter in 2002. [4]

D. Keras: Keras is a popular free-to-use open-source python package. It is capable to build deep models on smartphone software like iOS & Android. It supports Tensor-flow, PlaidML, Theano and MicrosoftCognitive Toolkit.

MACHINE LEARNING MODEL

A. LSTM: LSTM networks are suitable for making predictions which are based on time series data. It ispossible that there can be short delays between important events in the time series. It was developedand introduced by Hochreiter & Schmidhuber in theyear 1997. LSTMs are specifically known for computing long-term dependency solutions. It iscapable of remembering past information. [3]

B. Dense: A dense layer has important work to do in thismodel. It transfers all outputs generated by the preceding layer to the next neurons, then this neuron provides only single output to the upcoming layer.

C. Dropout: Dropout is a technique in which somehow the machine ignores randomly selected neurons duringtraining. It is a famous regularization technique whichwas proposed by Srivastava et al. in 2014. [4]

D. Sequential: It is a model which is not suitable for a multiple input for single stack layer, where each stacklayer comprises exactly one input and one output. It does not work on non-linear topology.

III. LITERATURE SURVEY

Sayavong Lounnapha et al., IEEE 2019 in [5], Research on "Stock Price Prediction Method Based on Convolutional Neural Network". This paper discussed about convolution neural network-based stock price prediction model, which is capable to learn by itself. It is shown in the result that the model based on Convolutional Neural Networks can easily forecast the trend in stock market price, providing significant allusion for stock price forecasting.

IEEE 2019-Soheila Abrishami et al. in [6], "Enhancing Profitby Predicting Stock Prices Using Deep Neural Networks". Predicting future trends of a particular stock is a difficult taskto do. The paper discussed on developing a deep neural learning system that uses various facts for particular stock listed on the New York based exchange named NASDAQ to predict its stock value.

"Share Price Prediction Using Machine Learning Technique", IEEE 2019, Jeevan B et al. in [7]. The paper discussed predicting stock prices on the National Stock Exchange with the help of Recurrent Neural Network and Long Short-Term Memory based on a variety of elements such as current market price and anonymous events.

Machine Learning Techniques for Stock Market Prediction IEEE 2020-Naadun Sirimevan et al. in [8]. In today's economy, stock market prices are critical. In various research, it is shown that popular social media platforms likeTwitter have the potential to influence anyone's decision- making process. Forecasts for the next day, week, and two weeks were made with great accuracy.

By combining methods from four well-known machine learning models, Weng et al. in [11] concentrated on short-term stock price prediction. Five sets of data make up the dataset for this study. These datasets were gathered from three open-source APIs and the TTR R program. With previous knowledge, the authors carefully evaluated five datasets before choosing eight technical markers for their investigation. The main contribution of this work is the development of an R-based platform for investors that does not require users to input their own data but instead uses an API to directly get the data from an internet source.

Additionally, Kara et al. in [12] made use of ANN and SVM to forecast changes in the stock price index. The Istanbul Stock Exchange's time span from January 2, 1997, to December 31, 2007, is covered by the data set they used. The key merit of this study is the comprehensive documentation of parameter tuning methods. The authors' failure to explain why their model beat other models in earlier research, the lack of novelty in either the technical indicator or the model structure, and these are the work's flaws—not its strengths.

IV. RELATED WORK

Kim and Han made use of both artificial neural networks (ANN) and genetic algorithms to develop a model, especially for predicting stock price indexes. They used technical indicators like one-day high, one-day low, total trade quantity, etc. in the daily Korean stock price index. They selected features and formulas based on data fromJanuary 1989 to December 1998, i.e. total of 2928 trading days.

First, the hidden layer has 12 fixed input and processing elements. The major limitation is that the authors only focused on two optimization factors during the ANN learning process. The features that have been chosen are referred to asour initialized feature pool [9].

The Hidden Markov Model (HMM) used by Hassan and Nath to forecast stock market prices for 4 different airlines. They categorization of the states of their model into the following 4 groups: opening prices, closing prices, highest prices, and lowest prices. The method used in this paper is unique in that it does not necessitate the use of higher knowledge in order to create a basic-level stock prediction model.

V. METHODOLOGY

In this area, we provide the suggested techniques and the plan for the proposed answer. Additionally, we highlight the framework plan with algorithmic and deployment particulars.

A. PROBLEM STATEMENT

We are analyzing the best way to forecast short-term price movements of certain stocks from two perspectives: feature engineering and forecasting algorithm. We then inquiredthree questions for each of these perspectives: How canfeature engineering enhance the accuracy of the model? Andwhich algorithm is the best for the prediction of short-term price movements?

The initial inquiry is regarding feature engineering. In this enquiry, we would try to understand how feature selection plays the important role in increasing the efficiency of the model.

The second inquiry we are addressing is which algorithms weshould use to model our data. We have divided the issue into two parts: firstly, the prediction of pattern and then the exact figure. In this paper, we are particularly focusing on predicting the pattern.

Our working strategy is to break down the complicated problem into subsets of tasks and solve them one after Another, then join the solutions into an efficient model to serve as a guide for investment decisions.

B. PROPOSED PLAN

We forecasted the future trends of particular stock in a stock market with the help of Long Short Term Memory. In Fig. 1. This architecture was chosen because LSTM fits in perfectly in this type of problem.

The LSTM model enables us to accurately predict the stock price of "x+1"Th days considering the previous data of "x" consecutive days. Keras, which is based on Tensor-flow, is used to create LSTM. [10

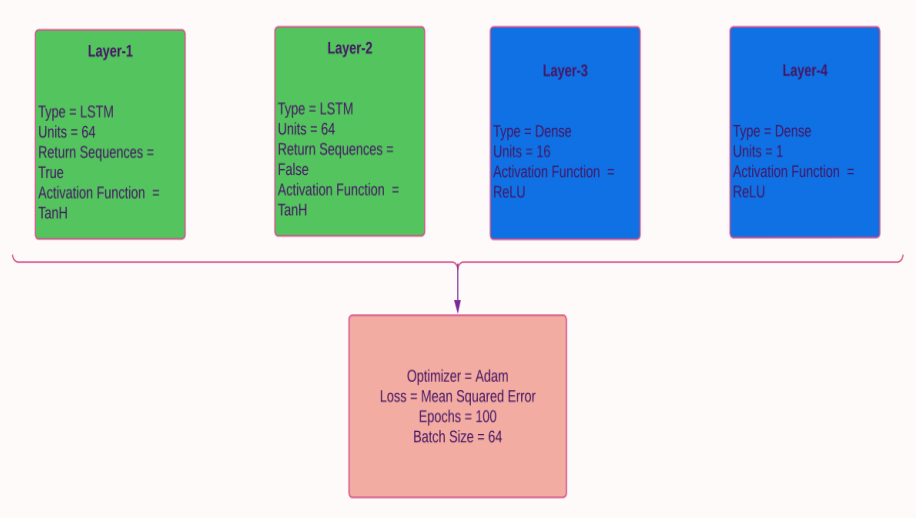

Fig. 1. The Architecture of Processed Recurrent Neural Network

We have taken use of "Min-Max Scaling" to scale the ten years of historical data and then concluded that "x" consecutive days of data predict the "x+1"th days during the preprocessing stage. Following that, our model is trained with an LSTM model of 4 layers, which consists of 64 units on the 1st layer and 64 units on the 2nd layer. The model is then converted into a 16 node Dense with one output node in between. With the help of this method, we were able to construct a solid connection between the inputs and outputs of our LSTM model and capture temporal dependencies.

With this architectural layout, we attempted to balance complexity and speed while preserving the capacity to precisely express temporal patterns and reducing the possibility of overfitting the model.

In order to optimize the model, the Adam optimizer is applied, and the loss is calculated using the "mean squared error" method. The model was trained over 100 epochs of a batch size of 64. The approximate loss following each training session is "1.243e04". The model is then used to forecast the closing price on the next trading day, resulting in the predicted closing
price.

Following model training, a forecast was made using the anticipated closing price for the following trading day. We evaluated the model's capacity to identify and forecast the underlying patterns and trends in the stock market data by utilizing this projected value as an estimate of the anticipated closing price.

The Flow diagram of our model is (Fig. 2.

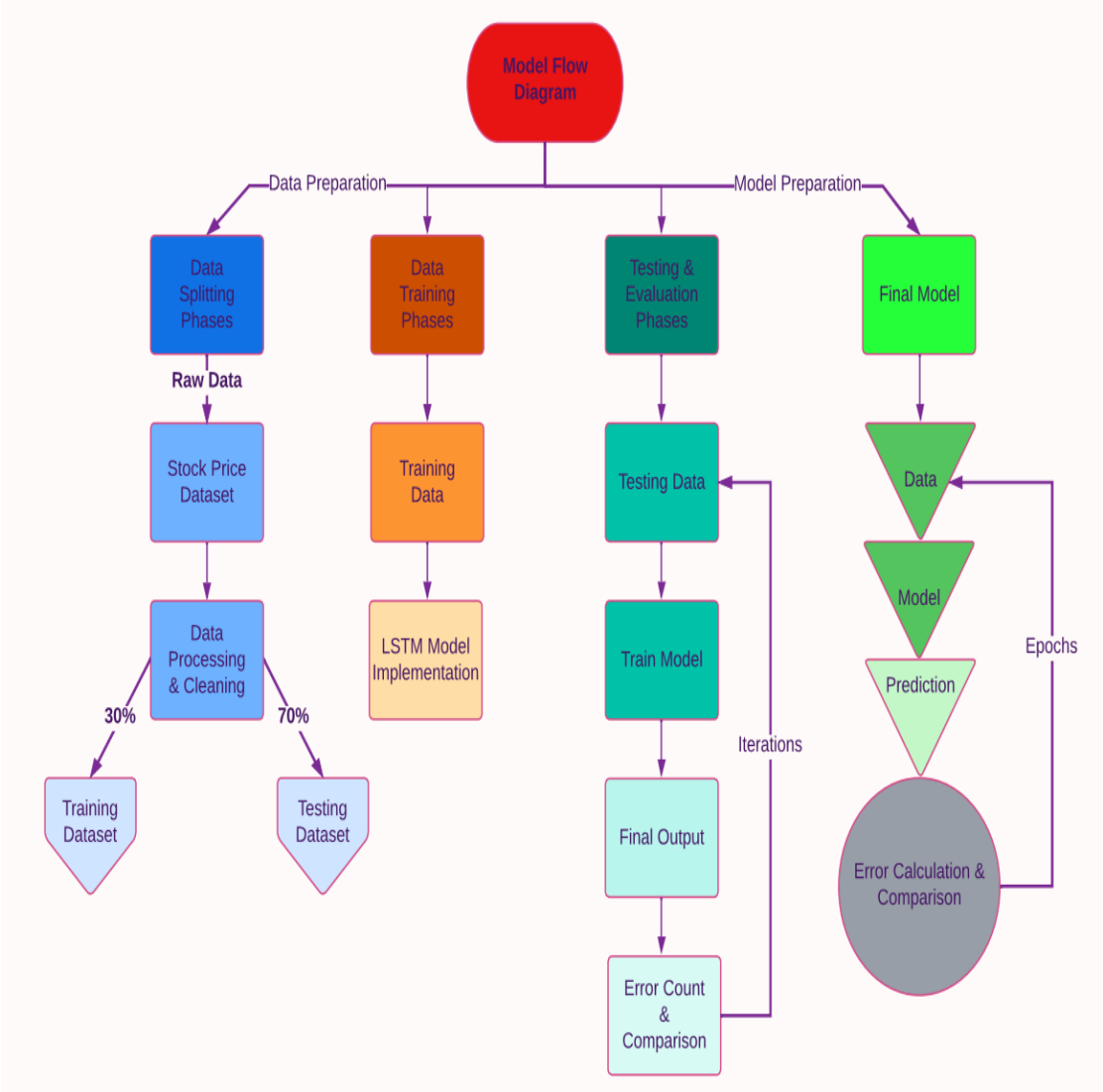

Fig. 2. The Model Flow Diagram of LSTM

VI. EXPERIMENTAL RESULT

The current section will be focused on the results of the proposed model. As a result, our model predicts the closing price for the day. To determine how accurately this model measures the performance of a specific stock, it must be compared to the actual closing price. The data set for training and testing model is directly extracted from 'finance.yahoo.com'.

In the following *Fig. 3.1*, *Fig. 3.2*, *Fig. 3.3,* and *Fig. 3.4* shows the Actual data, Training data, Predicted data and the difference between the predicted and actual data change fromNovember 2013 to December 2021 of 'Tata Consultancy Stock – TCS'.

Fig. 3.1. Actual data

Fig. 3.2. Training data

Fig. 3.3. Predicted data

Fig. 3.4. Price Prediction

The accuracy of our model is quite good for the given dataset. The Root Mean Squared Error (i.e., RMSE) is '13.35%' and Mean Absolute Percentage Error (i.e., MAPE) is '8.71%'. Hence,

We can find the accuracy of the model as:

$$100 - \text{Value of MAPE} = \text{Model Accuracy} \quad (1)$$

Thus, in this case:

$$\text{Accuracy} = (100\text{-}8.71)\ \%$$

$$\text{Accuracy} = 91.29\ \%$$

The Accuracy of our model for other stocks are (Fig. 4.):

TABLE I. ACCURACY OF MODEL OVER OTHER STOCKS

S. No.	Stock Name	Ticker ID[a]	Start Date	End Date	RMSE[b]	MAPE[c]	Accuracy[d]
1.	Tata Consultancy Services	TCS	2013/11/01	2021/12/31	13.35	8.71	91.29
2.	Tata Motors	TTM	2012/01/03	2021/12/31	32.33	3.7	96.3
3.	Infosys	INFY	2012/01/03	2021/12/31	7.43	14.18	85.82
4.	Apple	AAPL	2012/01/03	2021/12/31	28.75	2.8	97.2
5.	Tesla	TSLA	2012/01/03	2021/12/31	14.71	16.67	83.33
6.	Nike	NKE	2012/01/03	2021/12/31	49.63	7.38	92.62

Ticker ID: Unique ID of Stock given by Regulatory Board
b. RMSE: Root-Mean Squared Error
c. MAPE: Mean Absolute Percentage Error
d. Accuracy: 100 - MAPE

Fig. 4. Accuracy of Model over other Stocks listed in Yahoo Finance

VII. CONCLUSION

On comparing with previous works of same type of problems, our proposed solution isa kind of customization of an existing model because, instead of proposing new over-the-top LSTM model, we proposed a fully functional, efficient and custom deep learning prediction system and combined it with LSTM machine learning algorithm to predict future trends of stock market. By recommending a feature extension technique prior to recursive feature reduction, we narrow the difference between market investors and analysts and, by analyzing the results of earlier studies, we significantly increase model performance.

In the context of our research on the application of LSTM in stock price prediction, a number of constraints should be taken into consideration. First, the caliber and accessibility of our data affect how accurate our forecasts will be. Additionally, because of the volatility of the stock market, LSTM models might not be able to account for abrupt changes in investor behavior. Additionally, there's a chance of overfitting, particularly with short or noisy datasets, which might result in worse performance when applied to unobserved data.

The expected challenge of predicting stock price is that the model needs a variety of inputs in order to be able to anticipate things. It is possible to compare the stock price to important political and economic developments. Additionally, it is vital to include in the model challenging-to-represent elements like market psychology and mood. It's critical to know that human behaviors plays a major role in the market's volatility. Because it's a market, people tend to be impulsive and unpredictable.

VIII. ACKNOWLEDGMENT

We are grateful for our guide's assistance. We learned a lot about teamwork, leadership, and some important technical aspects of our project. Our guide Dr. Vipin Rai helpful suggestions are greatly appreciated.

IX. REFERENCES

[1] Banerjee, Sharanya & Dabeeru, Neha & Ra, Lavanya. (2020). StockMarket Prediction. International Journal of Innovative Technology and Exploring Engineering. 9. 2278-3075. 10.35940/ijitee.I7642.079920. https://www.researchgate.net/publication/343906044_Stock_Market_Prediction

[2] Stock Price Prediction Using Machine Learning: An Easy Guide! https://www.simplilearn.com/tutorials/machine-learning-tutorial/stock-price-prediction-using-machine-learning

[3] Kalkhoran, L.S., Tabibian, S. & Homayounvala, E. Detecting Persian speaker-independent voice commands based on LSTM and ontology in communicating with the smart home appliances. Artif Intell Rev (2022). https://doi.org/10.1007/s10462-022-10326-x

[4] Dropout Regularization with Keras - Dropout Regularization in Deep Learning Models with Keras - MachineLearningMastery.com

[5] N. Sirimevan, I. G. U. H. Mamalgaha, C. Jayasekara, Y. S. Mayuran and C. Jayawardena, "Stock Market Prediction Using Machine Learning Techniques," 2019 International Conference on Advancements in Computing (ICAC), 2019, pp. 192-197, doi: 10.1109/ICAC49085.2019.9103381. https://ieeexplore.ieee.org/document/9103381

[6] S. Singh, T. K. Madan, J. Kumar and A. K. Singh, "Stock Market Forecasting using Machine Learning:

Today and Tomorrow," 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), 2019, pp. 738-745, doi: 10.1109/ICICICT46008.2019.8993160. https://ieeexplore.ieee.org/document/8993160

[7] B. Jeevan, E. Naresh, B. P. V. kumar and P. Kambli, "Share Price Prediction using Machine Learning Technique," 2018 3rd International Conference on Circuits, Control, Communication and Computing (I4C), 2018, pp. 1-4, doi:10.1109/CIMCA.2018.8739647. https://ieeexplore.ieee.org/docment/8739647

[8] Mojtaba Nabipour, Pooyan Nayyeri, Hamed Jabani, Shahab S., (Senior Member, Ieee), and Amir Mosavi (2020). Predicting Stock Market Trends Using Machine Learning and Deep Learning Algorithms Via Continuous and Binary Data; a Comparative Analysis. DOI : 10.1109/ACCESS.2020.3015966

[9] Shen, J., Shafiq, M.O. Short-term stock market price trend prediction using a comprehensive deep learning system. J Big Data 7, 66 (2020). https://doi.org/10.1186/s40537-020-00333-6

[10] Ajinkya Rajkar , Aayush Kumaria , Aniket Raut , Nilima Kulkarni, 2021, Stock Market Price Prediction and Analysis, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH& TECHNOLOGY (IJERT) Volume 10, Issue 06 (June 2021) https://www.ijert.org/stock-market-price-prediction-and-analysis