iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** IV **Month of publication:** April 2022

DOI: https://doi.org/10.22214/ijraset.2022.41846

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue IV Apr 2022- Available at www.ijraset.com

Visualizing and Forecasting Stocks Using Machine Learning

Ajitkumar Jagtap[1], Yash Patil[2], Darshan Oswal[3]
[1, 2, 3]*Department of Computer Engineering, TSSM's Bhivarabai Sawant College of Engineering and Research affiliated with SPPU*

Abstract: ***India's stock market is exceedingly changing and reductionism, which has a countless number of features that control the directions and trends of the stock price; therefore, prediction of uptrend and downtrend is a complex process. This paper point of view to demonstrate the use of recurrent neural network in finance to prediction of the closing price of a selected stock and analyse opinions around it in real-time. By combining both techniques, the submitted model can give buy or sell recommendation.***
In Stock Market Prediction, the aim is to predict the upcoming future value of the financial stocks of the company. The latest trend in stock market prediction technologies is the use of machine learning which makes predictions depending on the values of current stock market indices by training on their previous stock values. Machine learning itself use different models to make prediction easier and authentic. The paper focuses on the use of Regression and LSTM based Machine learning to prediction of stock values. Factors for stocks considered are open, close, low, high and volume.
Keywords: Machine learning, Stock Market, Long Short-Term Memory, Recurrent Neural Network.

I. INTRODUCTION

Stock market is characterized as active, unforeseeable and nonlinear in universe. Prediction of stock prices is a very difficult task as it depends on various things including so many things, global economy, financial reports of company and performance etc. Traditionally, two important methods have been made for prediction of the stock price of an company. Technical analysis
method uses recorded price of stocks like closing and opening price, volume trade, adjacent closing values etc. of the stock for predicting the coming price of the stock.
The second type of analysis is qualitative analysis, which is performed on the basis of exterior things like company's bio, market circumstance, profitable and political factors, descriptive information in the form of financial new articles, social media and even blogs by economical analyst.
What is Stock Price prediction?
Stock market prediction is the act of trying to find the upcoming future value of a company stock or other financial instrument traded on an exchange. The successful prediction of a stock's future price could yield notable profit. The systematic market hypothesis suggests that stock prices reflect all currently available information and any price changes that are not depends on newly revealed information thus are inherently unpredictable. Others disagree and those with this viewpoint possess a lot of methods and technologies which purportedly allow them to gain future price information.
What is Machine Learning?
Machine learning is a branch of artificial intelligence and computer science which focuses on the use of data and algorithms to emulate the way that human beings learn, gradually enhancing its accuracy. IBM has a rich history with machine learning. One of its own, Arthur Samuel, is credited for forging the term, “machine learning” with his research (PDF, 481 KB) (link resides outside IBM) all over the game of checkers. Robert Nealey, the proclaimed checkers master, played the game on an IBM 7094 computer in year 1962, and he lost to the computer.
Compared to what can be done today, this feat almost seems insignificant, but it’s considered a major achievement within the field of artificial intelligence. Over the next couple of decades, the technological developments around storage and processing power will allow some innovative products that we know and love today, such as Netflix’s recommendation engine or self-driving cars.
This paper's Pattern is categorized as follows:
The first section including of the Abstract and Introduction followed by the second Section presents the other pieces of literature review that we referred to before working on this project. In the next section i.e. In Section three, we define the problem statement we were resolving and in Section four the system methodology and scheduled that we followed in given.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue IV Apr 2022- Available at www.ijraset.com

II. LITERATURE REVIEW

A. Title: Forecasting directional movements of stock prices for intraday trading using LSTM and random forests:
Authors: Pushpendu Ghosh, Ariel Neufeld, Jajati Keshari Sahoo

They sign up both random forests and LSTM networks as training methodologies to examine their effectiveness in forecasting out-of-sample directional movements of constituent stocks of the S&P 500 from January 1993 till December 2018 for intraday trading. We initiate a multi-feature setting including not only of the returns with respect to the starting prices, but also with respect to the opening prices and intraday returns. As trading plan of action, we use Krauss et al. and Fischer & Krauss as benchmark. On each trading day, we buy the 10 stocks with the maximum probability and sell short the 10 stocks with the minimum probability to outperform the market in terms of intraday returns – all with equal financial weight. Our observed results show that the multiple feature setting obtains a regular return, prior to transaction costs, of 0.64% using LSTM networks, and 0.54% using random forests. Hence, we outperform the single-feature setting in Fischer & Krauss and Krauss et al. consisting only of the regular returns with respect to the closing prices, having corresponding regular returns of 0.41% and of 0.39% with respect to LSTM and random forests, respectively.

B. Title: Stock Closing Price Prediction using Machine Learning Techniques
Authors: Mehar Vijay, Deeksha Chandolab, Vinay Anand Tikkiwalb.

This paper they have recorded data for the five companies has been gathered from Yahoo Finance. The data contains
information about the stock such as Adjacent close, High, Low, Open, Close and Volume. Only the day-wise closing price of the stock has been pulled out. They have Six new variables have been created to predict closing stock price. These variables have been used to train the model properly. They have used ANN data mining technique to recognize elemental trend from data. They executed Random Forest for performing both regression and classification task.
Artificial Neural Network is used for predicting the next day closing price of the stock and for a relative analysis, RF is also implemented. The relative analysis depending on RMSE, MAPE and MBE values clearly designate that ANN gives
better prediction of stock prices by comparing RF. For future work, deep learning models should be proposed which consider financial news articles along with financial factors for possibly better results.

C. Title: A Survey of Forex and Stock Price Prediction Using Deep Learning:
Author: Zexin Hu , Yiqi Zhao and Matloob Khushi

Predictions of stock and Forex have always been a trending and profitable area of study. Deep learning applications have been approved to submit better accuracy and return in the field of financial prediction and forecasting. In this survey, we selected research papers from the Digital Bibliography & Library Project (DBLP) database for comparison and analysis. We separated papers according to different type of deep learning methods, which mentioned Convolutional neural network (CNN); Long Short-Term Memory (LSTM); Deep neural network (DNN); Recurrent Neural Network (RNN); Reinforcement Learning; and other deep learning methods such as Hybrid Attention Networks (HAN), self-paced learning mechanism (NLP), and Wave net. Furthermore, this paper examines the dataset, variable, model, and results of each one article. The survey used represents the results through the most used performance models: Root
Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Mean Square Error (MSE), exactness, keen ratio, and return rate. We recognized that recent models combining LSTM with other methods, for example, DNN, are widely researched. Reinforcement learning and other deep learning methods submitted great
returns and performances. We conclude that, in previous recent years, the trend of using deep-learning-based methods for financial modelling is increasing exponentially.

D. Title: Stock Price Prediction Using Machine Learning and LSTM-Based Deep Learning Models
Authors: Sidra Mehtab, Jaydip Sen and Abhishek Dutta

Prediction of stock prices has been an important area of research since a long while ago. While supporters of the well-organized market hypothesis accept that it is impossible to predict stock prices exactly, there are formal propositions exhibiting that exact modelling and designing of appropriate variables may lead to models using which stock prices and stock price movement patterns can be very exactly can predicted. Researchers have also worked on technical analysis of stocks with an aim of recognizing patterns in the stock price variations using advanced data mining techniques. In this work, we propose a way of hybrid modelling for stock price prediction building different machine learning and deep learning-based models.

For the motivation of our study, we have used NIFTY 50 index values of the National Stock Exchange (NSE) of India, during the period December 29, 2014 till July 31, 2020. We have built eight regression models by using the training values that consisted of NIFTY 50 index records during December 29, 2014 till December 28, 2018.

Using these regression models, we predicted the opening values of NIFTY 50 for the period December 31, 2018 till July 31, 2020. We, then, increase the predictive power of our forecasting framework by built four deep learning-based regression models using long-and short-term memory (LSTM) networks with a novel approach of walk-forward validation. Using the grid-searching technique, the hyperparameters of the LSTM models are optimized so that it is ensured that confirmation losses stabilize with the increasing number of epochs, and the connection of the validation.

accuracy is achieved. We utilize the power of LSTM regression models in forecasting the future NIFTY 50 open values using four different models that differ in their architecture and in the structure of their input data. Huge results are presented on various metrics for all the regression models. The results clearly designate that the LSTM-based univariate model that uses one-week earlier data as input for predicting the next week's open value of the NIFTY 50 time series is the most exact model.

III. PROBLEM DEFINITION

Developing an interface between human and the system using the machine learning algorithms and python language to predict the prices of the stock using Long Short-Term Memory (LSTM) networks.

This is totally based on Machine Learning Algorithm to proceed and provide an effective result. Getting the data and processing it and generating a forecast is the problem statement that we worked on.

IV. METHODOLOGY

While working on the use the machine learning in the prediction of stock market the different types of the neural network can be proposed for prediction. Here we are using Long short -term memory (LSTM) which is an artificial recurrent neural network (RNN).

A. Recurrent Neural Network

A recurrent neural network (RNN) is a part of artificial neural network. RNN uses the sequential data or time series data. RNN works on principle of saving the output of the specific layer and feeding this back to the input in order to forecast the output of the layer. The most important characteristic of the Recurrent neural network is hidden state, which consider data about sequence.

B. Long Short-Term Memory (LSTM):

The Long Short-Term Memory is a special type of RNN, it is efficient of catching the data from previous Stages and use it for upcoming future predication. LSTMs are widely used for sequence prediction problems and have proven to be extremely successful. The reason behind is they work so well is because LSTM is able to store historic information that is important and forget the information that is not. LSTM has three gates:

1) The input gate: The input gate adds details to the cell state.
2) The forget gate: It removes the details that is no longer required by the model.
3) The output gate: Output Gate at LSTM selects the details to be shown as output

C. Process Flow

1) *Collecting the Data:* Gathering the data is the initial step. This deals with group of the right dataset. we have collected the historical data from yahoo finance website. And using this historical data to predict the stock prices.
2) *Data Pre-processing:* Data pre-processing is the type of data mining. This requires transforming the raw data into the required format. The data pre-processing involves getting the data for specific time frame, looking for the particular features which are required to teach the model, feature scaling, Dividing the dataset into training and testing.
3) *Feature Extraction:* In the feature extraction the feature that are required to get expected output are picked. This selected feature is provided to the algorithm to get desired output. In this we are selecting time duration and Closing price of the Stock on that respective Date.
4) *Training the Machine:* In this the Data is fed to the machine learning algorithm (LSTM) to train the machine for making the future predictions. The training of machine is included of the cross-validation where we get balanced evaluated performance of the machine learning model using training data.

5) *Output Generation:* The stock symbol enters by the user performed as the input to the system. Based on that user has shown previous ten-year data. After that based on the hundred days and two-hundred days moving average machine learning model made the predictions. The analysis of input data takes place and after that the for that particular time period the predictions are created as output.

V. IMPLEMENTATION AND RESULT

After getting the relative stock data hundred days moving average and two-hundred days moving average is calculated to predict the closing price of the particular share over the given period.

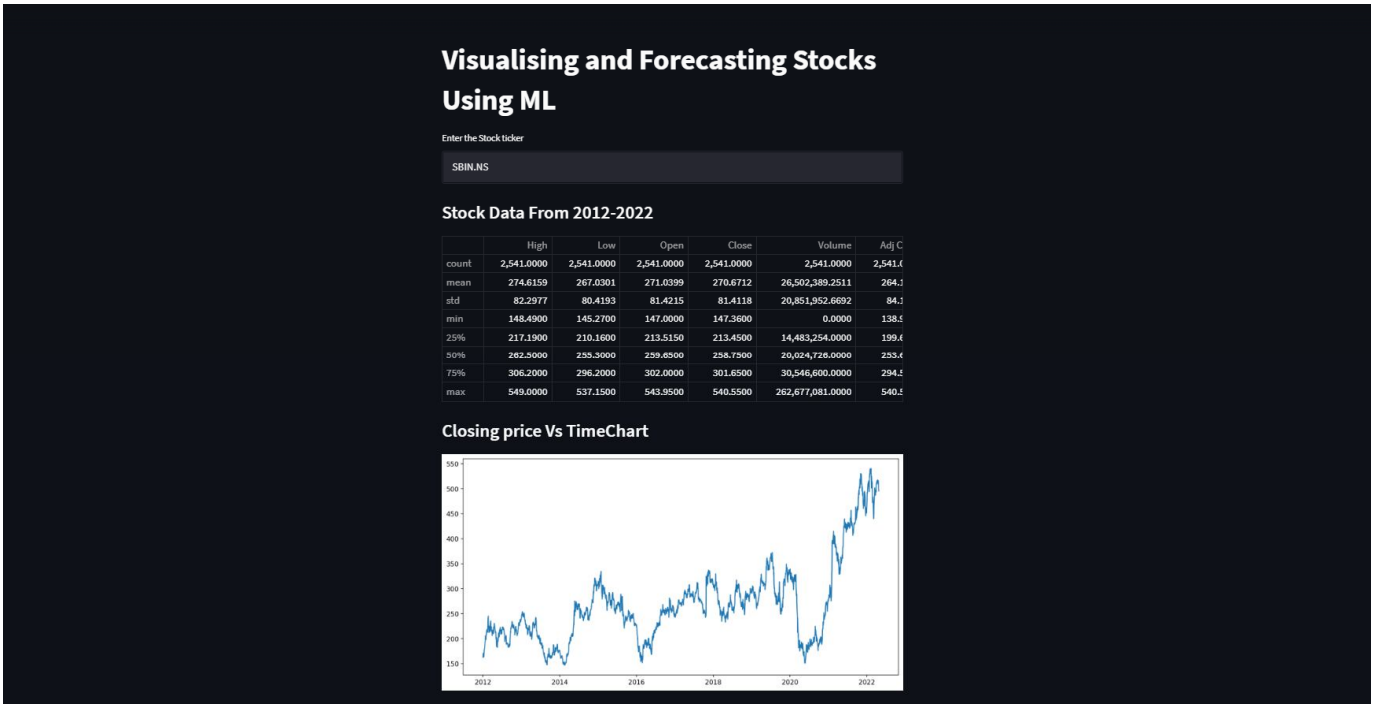


Fig. Stock data from 2012-2022

After getting the relative stock data hundred days moving average and two-hundred days moving average is calculated to predict the closing price of the particular share over the given period.

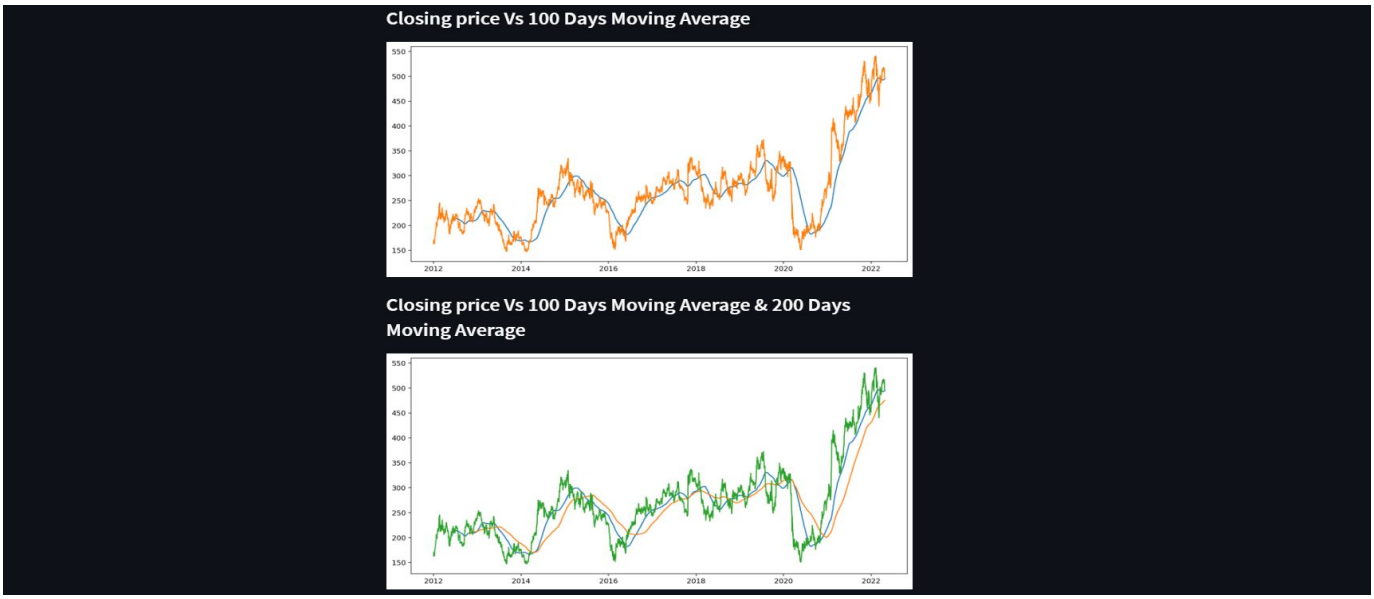


Fig. 100 Days and 200 Days Moving Average

The machine Learning model takes the hundred days moving average and two hundred moving average and predict the stock price for selected time period. The model takes first hundred days price and predict the next day price. By using the moving averages LSTM predict stock price. The prediction accuracy of the LSTM model improves with increase in the size of the dataset and makes is more efficient.

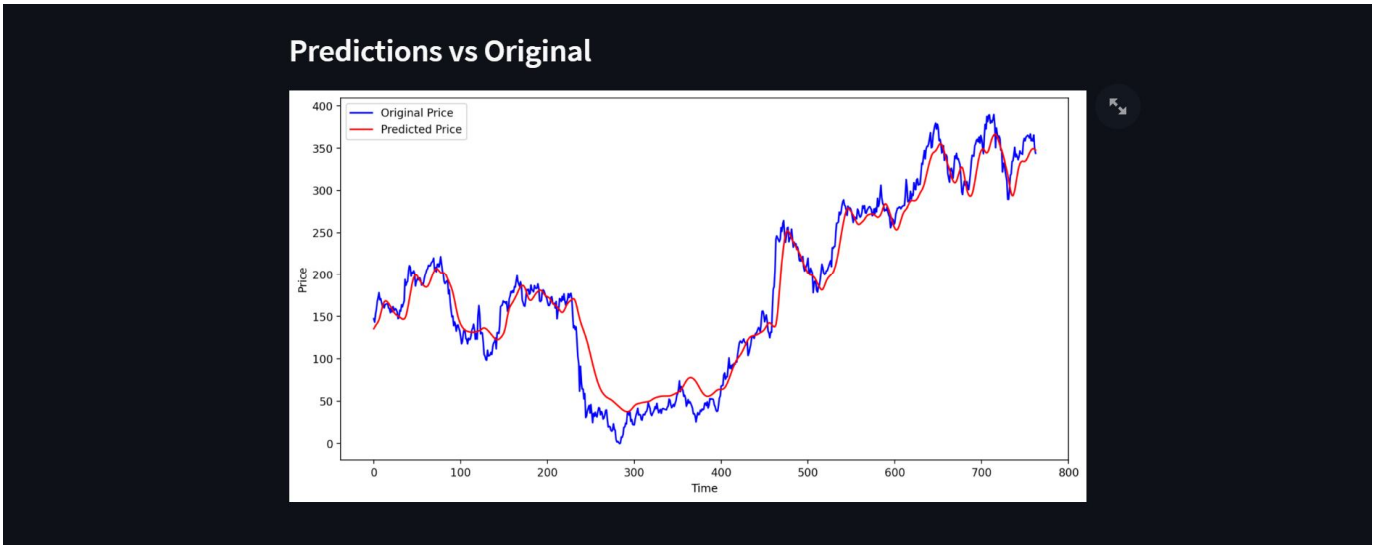


Fig. Predictions Vs Original

VI. CONCLUSIONS

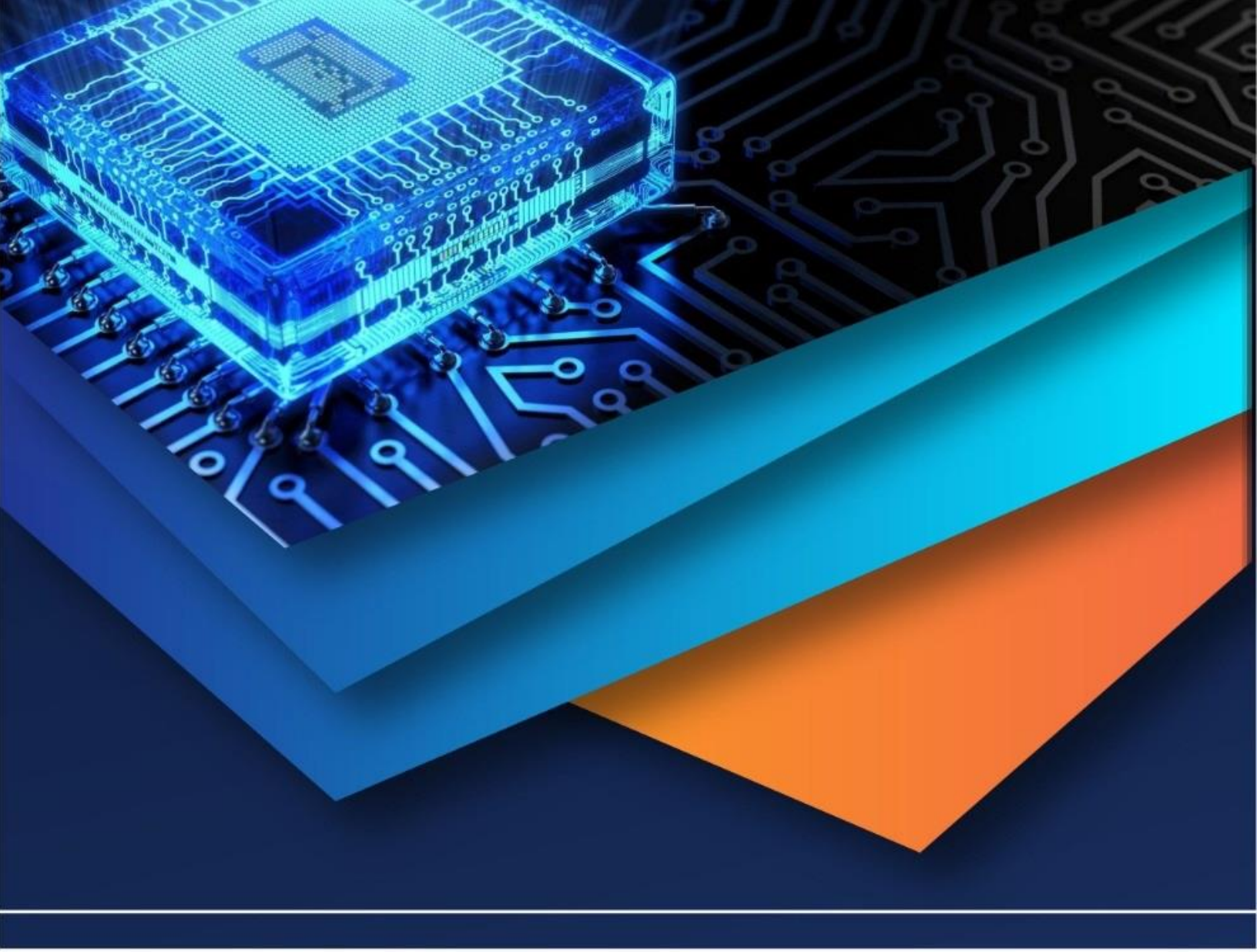
Predicting the stock market is a time-consuming and arduous procedure. However, with the introduction to Machine Learning and its various algorithms, the Stock Market Prediction advancements have begun to consist such approaches in considering stock market data. By calculating the accuracy of the different type of algorithms, we found that the most appropriate algorithm for predicting the market price of a stock based on various data points from the previous past data is the Long-Short Term Memory (LSTM) algorithm. The algorithm will be a great asset for brokers and investors for investing money in the stock market since it is trained on a vast collection of historical data and has been chosen after being tested on a trial data. The project demonstrates the machine learning model to predict the stock price with more accuracy as compared to other machine learning models. The sentiment analysis of social media can be used with the LSTM to better train weights and further improve machine learning model. It has led to the conclusion that it is possible to predict the stock market with more accuracy and efficiency using machine learning techniques.

VII. ACKNOWLEDGMENT

We would like to show gratitude to our project guide Prof. Shubhangi Bhagat who motivated and guided us to make this project successful. Our project guide enabled us to continue with her professional guidance, kind advice, suggestions and timely inspiration which has helped us decide on our project. We also extend our deepest gratitude to our HOD Dr. S. N. Gujar. We would like to thank him for simulating suggestions and encouragement, along with areas for improvement which helped us for implementation and writing of this dissertation.

REFERENCES

[1] Pritam Ahire, Hanikumar Lad, Smit Parekh, Saurabh Kabrawala, “LSTM Based Stock Prediction,” International Journal of Creative Research Thoughts(IJCRT), vol. 9, pp. 5118-5122, Feb. 2021.
[2] Ya Gao, Rong Wang, and Enmin Zou, “Stock Prediction Based on Optimized LSTM and GRU Models,” Hindawi, vol. 2021, pp. 1-8, Sept. 2021.
[3] Adil Moghar and Mhamed Hamiche, “Stock Market Prediction Using LSTM Recurrent Neural Network,” Sciencedirect, vol. 170, pp.1168-1173, Apr. 2020.
[4] Mallikarjuna Shastry P. M. and Pramod B S, “Stock Price Prediction Using LSTM,” ResearchGate, vol. 83, pp. 5246-5251, May. 2020.
[5] Ishita Parmar, Navanshu Agarwal, Sheirsh Saxena, Ridam Arora, Shikhin Gupta, Himanshu Dhiman, Lokesh Chouhan, “Stock Market Prediction Using Machine Learning,” in First Inter- national Conference on Secure Cyber Computing and Communications, 2018, doi: 10.1109/ICSCCC.2018.8703332
[6] Uttamraj K R, Raghav Nandakumar, Vishal R, Y.V Lokeswari, “Stock Price Prediction Using Long Short Term Memory,” International Research Journal of Engineering and Technology (IRJET), vol. 05, pp. 3342-3348, Mar. 2018.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY