**Mini Project Report on**



**STOCK PRICE PREDICTION PROJECT**



**Submitted in partial fulfilment of the requirement for the award of the degree of**

**BACHELOR OF TECHNOLOGY**

**IN**

**COMPUTER SCIENCE & ENGINEERING**

**Submitted by:**

**Student Name: Yash Singhal**  **University Roll No.: 2017587**

***Under the Mentorship of***

**Dr. Jyoti Agarwal**

**Assistant Professor**



**Department of Computer Science and Engineering**

**Graphic Era (Deemed to be University)**

**Dehradun, Uttarakhand**

**July-2023**



**CANDIDATE’S DECLARATION**

I hereby certify that the work which is being presented in the project report entitled **“Stock Price Prediction Project”** in partial fulfillment of the requirements for the award of the Degree of Bachelor of Technology in Computer Science and Engineeringof the Graphic Era (Deemed to be University), Dehradun shall be carried out by the under the mentorship of **Dr. Jyoti Agarwal, Assistant Professor**, Department of Computer Science and Engineering, Graphic Era (Deemed to be University), Dehradun.

Name: Yash Singhal University Roll no.: 2017587

**Table of Contents**

|  |  |  |
| --- | --- | --- |
| **Chapter No.** | **Description** | **Page No.** |
| Chapter 1 | Introduction | 1-3 |
| Chapter 2 | Literature Survey | 4-6 |
| Chapter 3 | Methodology | 7-9 |
| Chapter 4 | Result and Discussion | 10-11 |
| Chapter 5 | Conclusion and Future Work | 12-13 |
|  | References | 14 |

**Chapter 1**

**Introduction**

Stock price forecasting has been a key area of research for a very long time. Although proponents of the efficient market hypothesis claim that it is impossible to predict stock prices with any degree of accuracy, there are formal arguments that suggest accurate modelling and the design of relevant variables may lead to models that can be used to predict stock prices and stock price movement patterns with a high degree of accuracy.

**What exactly is a stock market?**

A stock market is a public market where you may buy and sell shares of companies that are traded openly. The stocks, also known as equities, represent ownership in the company. The stock exchange serves as the middleman that enables the buying and selling of shares.

* Stock markets encourage individual wealth and help firms raise capital.
* It promotes individual progress.
* An economic indicator of a country is the state of the stock markets.
* People frequently utilize it as a source when investing in businesses with strong growth prospects.

The stock market is known for being unexpected, non-linear, and dynamic. It can be difficult to predict stock values since they depend on a variety of variables, such as the political climate, the state of the global economy, the financial success of the firm, and more. Therefore, methods to anticipate stock values by looking at the pattern over the last several years could prove to be quite helpful for making stock market moves, maximizing profit and minimizing losses. The stock price of an organization has traditionally been predicted using two major methods. In order to forecast the future price of a stock, technical analysis examines previous stock prices, such as closing and opening prices, volume traded, adjacent close values, etc.

The second sort of analysis is qualitative, which is carried out using outside variables such as the firm profile, the market environment, political and economic issues, textual data in the form of financial news stories, social media, and even blogs written by economic analysts. Modern stock price forecasting methods involve sophisticated intelligent methodologies based on either technical or fundamental analysis. In particular, the data size is enormous and non-linear for stock market analysis. An effective model that can find the hidden patterns and intricate relationships in this massive data collection is required to handle this diversity of data. Compared to previous methodologies, machine learning techniques in this field have shown to increase efficiency by 60–85%.

**Understanding Stock Price Prediction**

Stock Price Prediction Using Machine Learning is the practice of predicting the future value of a stock traded on a stock market in order to generate money. Because there are so many factors to consider, it is challenging to anticipate stock values with high accuracy, which is why machine learning is so crucial.

It helps you determine the potential worth of stock in a firm and other financial asset traded on a market. Gaining significant profits is the whole point of making stock price predictions. It's challenging to forecast how the stock market will fare. Other variables, including biological and psychological ones, as well as rational and irrational behavior, are included in the prediction. These forces work together to create a volatile and dynamic market for shares. Because of this, it is quite challenging to create precise stock price predictions.

Discovering the future worth of business stock and other financial assets traded on an exchange is made possible with the aid of stock price prediction utilizing machine learning. Gaining significant profits is the whole point of making stock price predictions. It's challenging to forecast how the stock market will fare. Other variables, including biological and psychological ones, as well as rational and irrational behavior, are included in the prediction. These forces work together to create a volatile and dynamic market for shares. Because of this, it is quite challenging to create precise stock price predictions.

**Chapter 2**

**Literature Survey**

Based on the usage of variables and the method used to represent the problem, the current research on time series forecasting and stock price prediction may be roughly divided into three clusters. The models that use bivariate or multivariate regression on cross-sectional data make up the majority of the first type of work. These models frequently fail to yield extremely accurate findings due to their intrinsic simplicity and the falsity of the linearity assumptions they make. The hypotheses in the second category forecast stock prices using time series theory as well as other econometric tools like the Granger Causality Test, autoregressive distributed lag (ARDL), vector autoregression (VAR), and quantile regression.

Machine learning, deep learning, and natural language processing are all used in the third category of work, which contains assertions based on learning.

One of the primary flaws of the present theories proposed in the literature for stock price prediction, except for the category of work that makes use of learning-based approaches, is their inability to correctly forecast extremely dynamic and swiftly changing patterns in stock price movement. By utilizing the strength of machine learning and deep learning-based models to create a very solid, dependable, and accurate framework for stock index prediction, we aim to solve the issue in this study. We have specifically deployed an LSTM network-based deep learning model and evaluated how well it performs in forecasting future stock index values.

The use of machine learning techniques in finance is the subject of numerous studies at the moment. Some of these studies used tree-based models to forecast portfolio returns, while others employed deep learning to predict the future prices of financial assets. Additionally, various scholars reviewed the ADaBoost algorithm's predicting of returns.

Others continue to predict stock returns using a special decision-making model for day trading investments on the stock market that was established by the authors. In this model, portfolio selection is done using the mean variance (MV) approach and the support vector machine (SVM) method, respectively. In another work, deep learning algorithms for smart indexing were discussed.

Applications of LSTM includes return forecasting, portfolio construction, ethics, fraud detection, decision making, language processing, and sentiment analysis. Additionally, some studies have covered a wide range of trends and Applications of Machine Learning in Quantitative Finance. Due to the fact that these models don't rely on long-term memory (past sequences of data), a class of machine learning algorithms based on recurrent neural networks has proven to be particularly effective in forecasting and price prediction in the financial market. In a paper, the accuracy of the illustrative time series data forecasting methods, the LSTM and the autoregressive integrated moving average (ARIMA), is compared. These methods were applied to a collection of financial data, and the findings revealed that LSTM outperformed ARIMA by a wide margin.

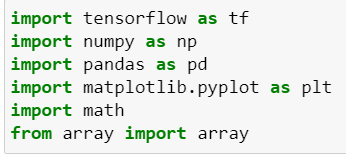
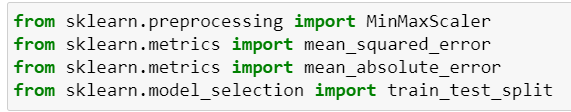
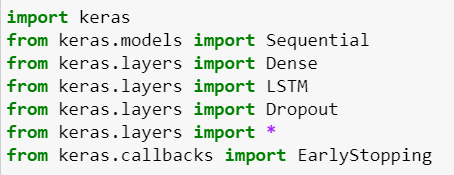
The majority of earlier research in this field has used traditional algorithms to predict stock prices, including linear regression, Random Walk Theory (RWT), Moving Average Convergence/Divergence (MACD), and autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA). Recent research demonstrates that machine learning can improve stock market prediction. methods like Random Forest (RF) and Support Vector Machine (SVM). Some neural network-based techniques, including Recurrent Neural Networks (RNN), and Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM), have also demonstrated promising outcomes.

**Chapter 3**

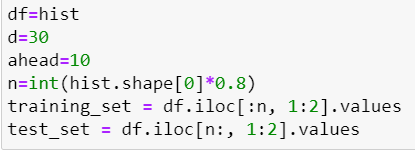
**Methodology**

The following steps are included in Stock Price Prediction using LSTM –

1. Importing the required libraries

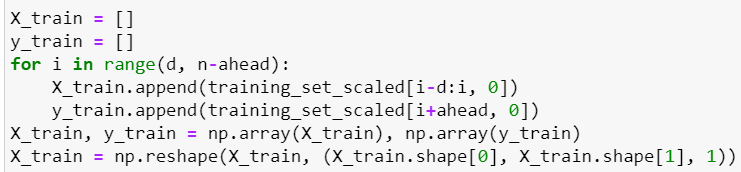
1. Preprocessing the Data

With real data, it is always a good to normalize or rescale the data within a given range. This helps the model to rapid convergence by preventing characteristics with higher numeric values from unfairly interfering with and biassing the model.

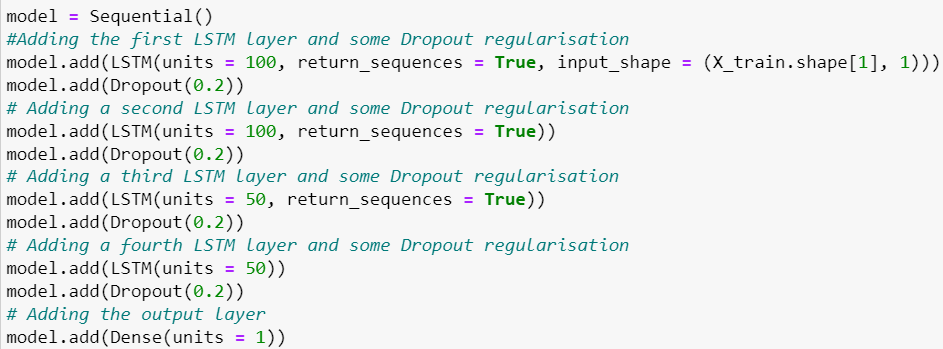
MinMaxScaler preprocessing package is also available through Scikit-learn. However, taking the surrounding circumstances into account, stock prices may peak or trough on various days, making it risky to use those numbers to sway others. Since neither of these solutions would significantly alter the values, we continue to use StandardScaler.

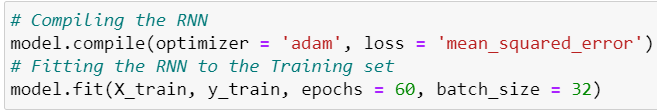
1. Create Training and Testing Sets (80:20 ratio)



1. Building the LSTM Model

We will create a basic, single-unit LSTM model using the Sequential and LSTM modules offered by Tensorflow Keras. Compile the model using and fit the model to the training data.





1. Finally predict the stock prices



**Understanding Long Short Term Memory Network (LSTM)**

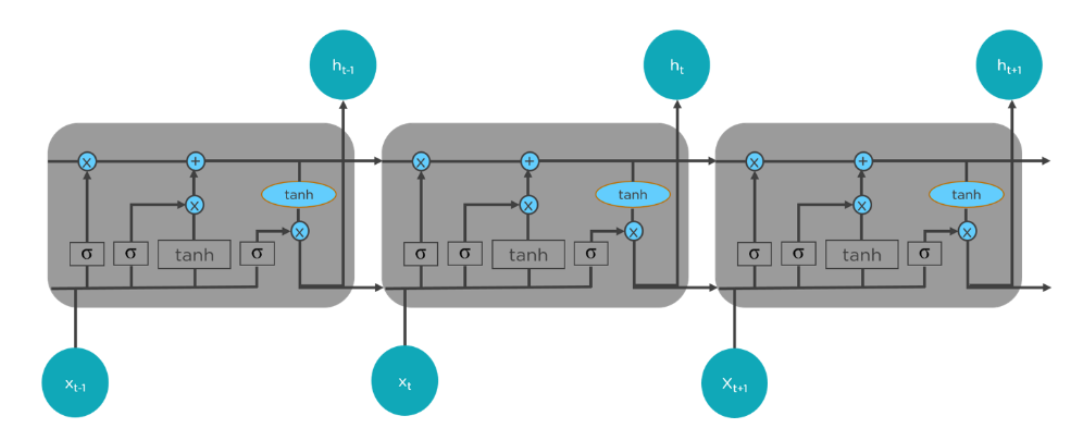


Fig.1 Detailed architecture of Long Short Term Memory Network (LSTM).

As in Fig. 2, LSTMs have a chain-like structure. There is only one neural network layer in general RNNs. On the other hand, LSTMs have four interacting layers that communicate incredibly well.

LSTMs work in 3 steps-

* 1st Step - The initial stage of LSTM is to choose which data should be left out of the cell at that specific time step. A sigmoid function is employed to make the decision. It computes the function while considering the current input xt, the previous state (ht-1), and both.
* 2nd Step - The second layer has two functions. The sigmoid function and the tanh function are the first and second, respectively. Which numbers to pass through are determined by the sigmoid function (0 or 1). The tanh function assigns the values passed weight, determining their relevance on a scale of -1 to 1.
* 3rd Step - The third phase entails choosing the final product. Run a sigmoid layer first, which chooses which components of the cell state are sent to the output. The cell state must then be multiplied by the output of the sigmoid gate after being passed through the tanh function to push values between -1 and 1.

Over time LSTM has largely contributed in transforming Language comprehension, predicting, speech and handwriting recognition, as well as a number of other applications that are now considered standard practice.

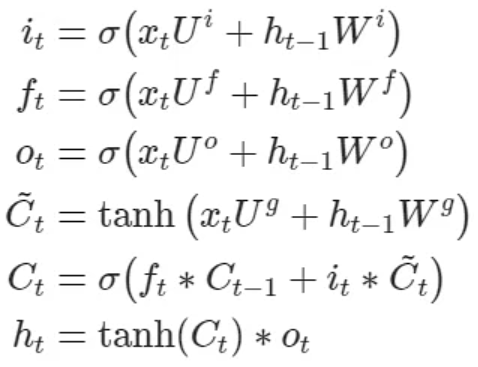
****

Fig.2 Equations computed inside each cell for each timestamp t,

considering the learned data from timestamp (t-1).LSTM.

**Chapter 4**

**Result and Discussion**

From our stock price dataset, we select the ‘Open’ and ‘Close’ columns as our training data, i.e., the first and second column. The "Close" column shows the highest price shares hit that day, while the "Open" column shows the opening price for shares that day. Our model is trained as per data in these columns. We plot the Stock Opening Price History Graph and Stock Closing Price History Graph to have a better understanding of the data in Fig 3.1 and Fig 3.2 shown below.

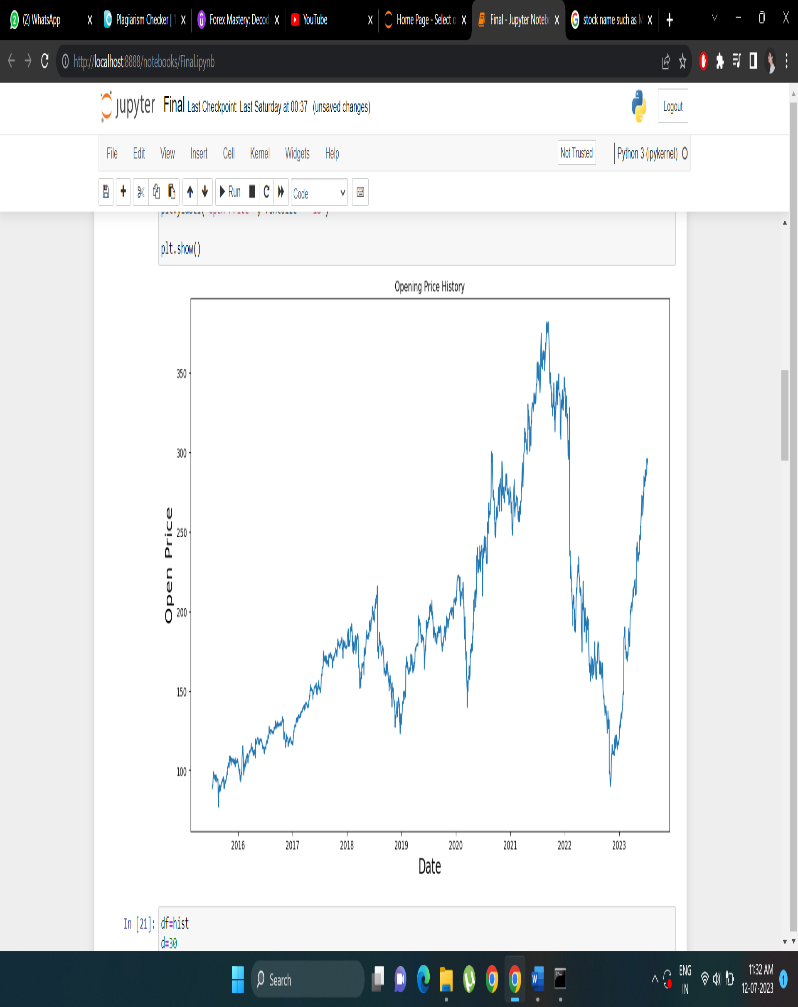
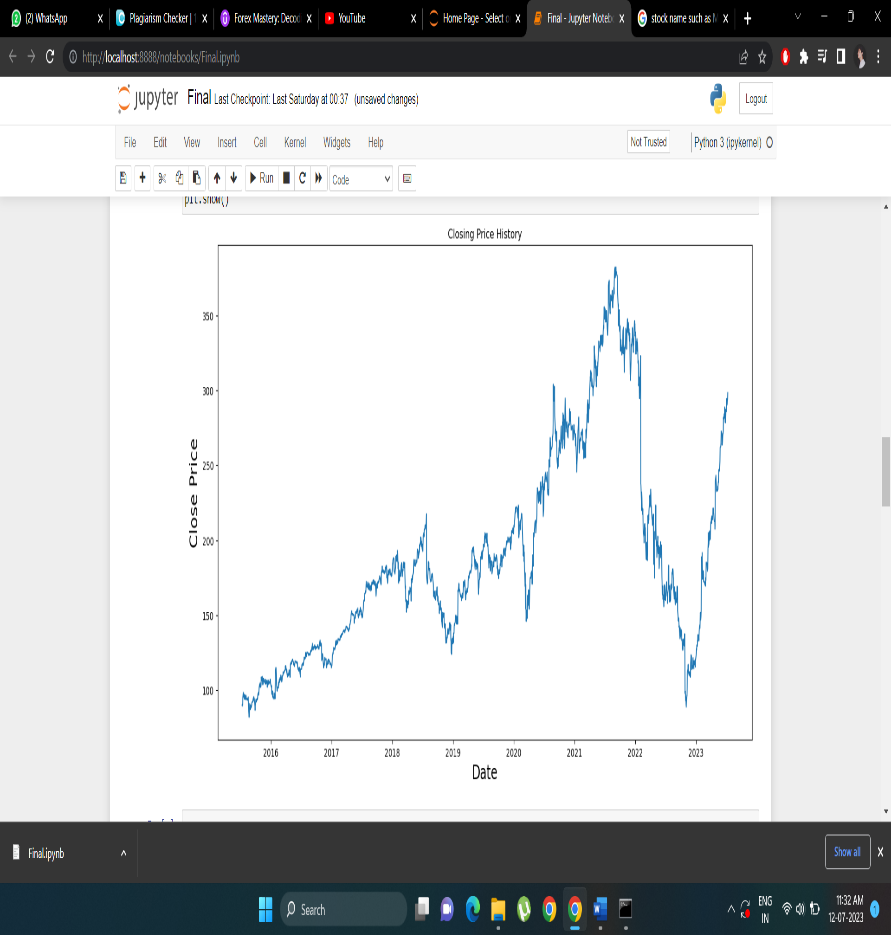
 

Fig.3.1 Graph plotted between Fig.3.2 Graph plotted between

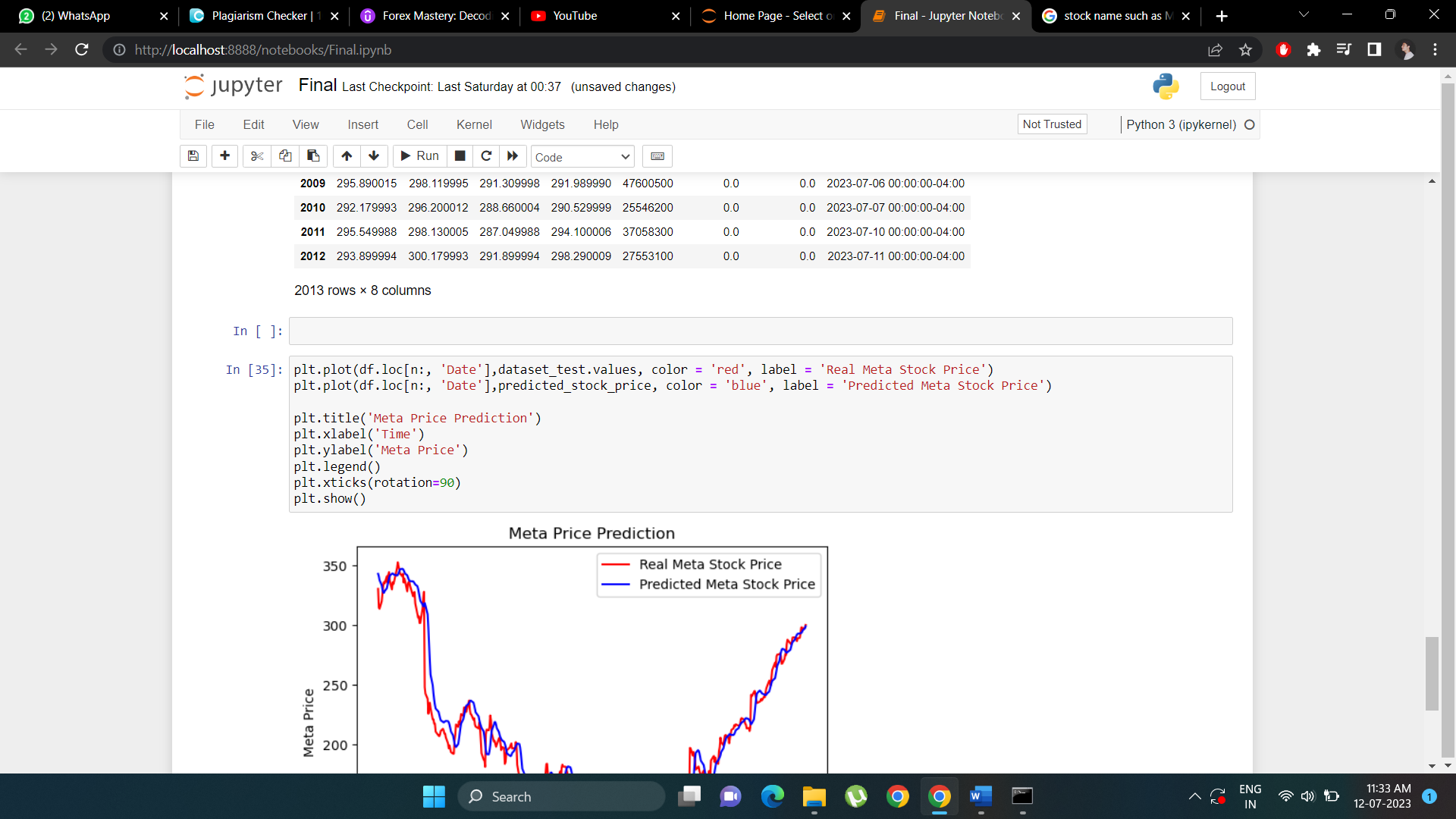
Open Stock Prices vs Date Close Stock Prices vs Date

To check the result and accuracy of the prediction of stock prices made by the LSTM Model, we plot a graph between the actual stock prices vs the predicted stock prices for META within a specified range of July 8, 2015 to July 7, 2023 (Fig 3.3).

We use matplotlib library to visualize the result wherein,

Red -> indicates Actual Stock Prices

Blue -> indicates Predicted Stock Prices

****

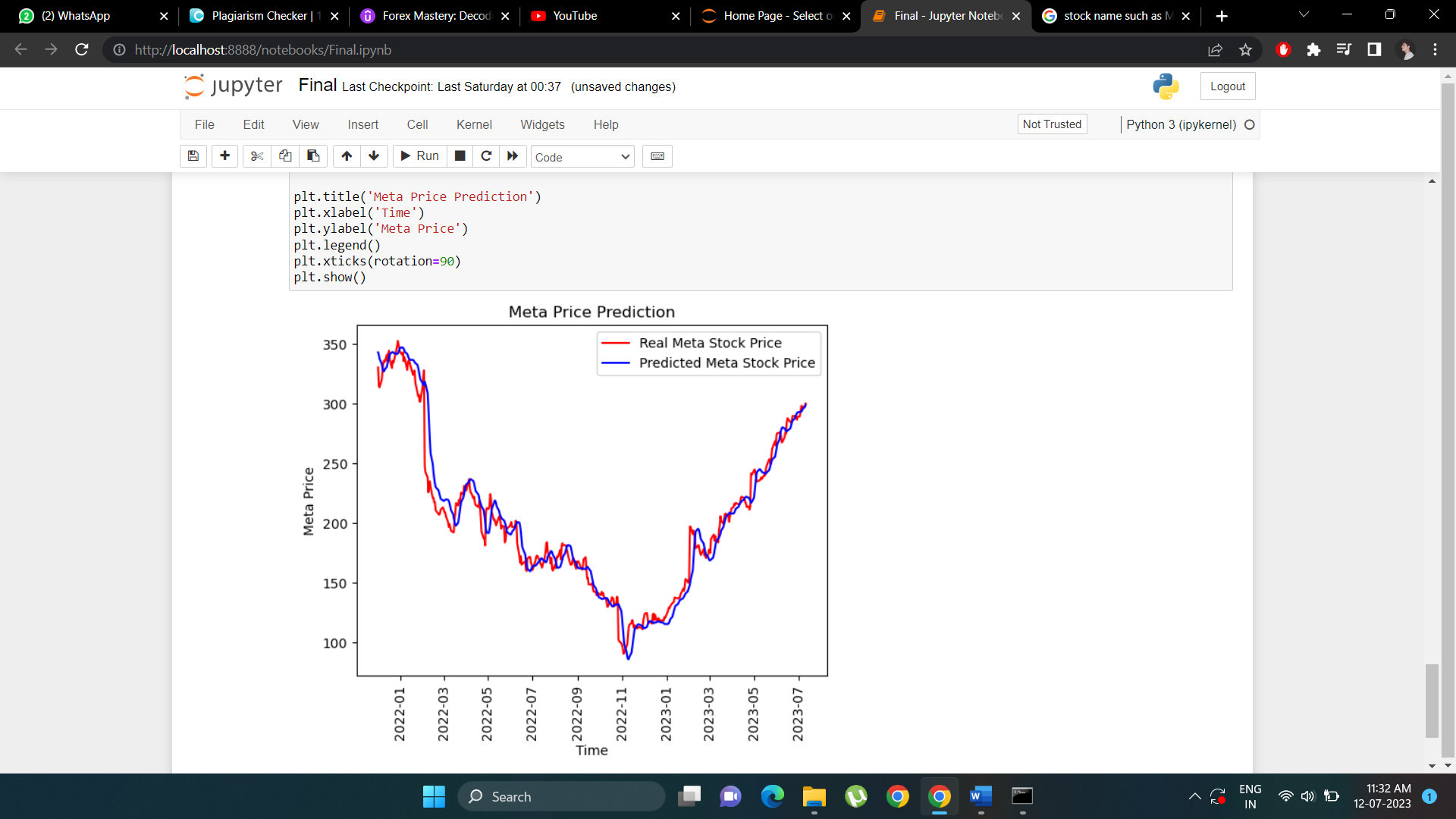
****

Fig.3.3 Graph plotted between Actual Stock Prices vs Predicted Stock Prices

While the particular price points from our anticipated price were not always close to the actual price in the beginning, our model shows very accurate forecasted pricing from the middle of the length to the finish as well as overall trends like moving up or down. The LSTMs can be fairly useful in times series forecasting, as this project has shown.

Also, by training the model with new data and adding more LSTM layers, the accuracy of the model can be improved.

**Chapter 5**

**Conclusion and Future Work**

The stock market has a remarkable impact on our daily lives and has a much bigger impact on the GDP growth of any country. Because stock values frequently fluctuate and depend on numerous factors that create complicated patterns, predicting stock market returns is a difficult undertaking. Only a few parameters, such as high, low, open, close, and adjacent close values of stock prices, as well as the volume of shares traded, are included in the historical dataset that is accessible on the company's website.

We created, optimized, and finally evaluated the predicted models using the data from Yahoo Finance, stock values of tech giants from July 8, 2015 to July 7, 2023. The raw data was subjected to pre-processing and data wrangling activities, and a set of derived variables was produced for use in the model-building process. The performances of the LSTM-based deep learning regression models were found to be excessively better than those of the machine learning-based predictive models among all regression models based on machine learning and deep learning. The study has convincingly supported our hypothesis that models built on deep learning are far more capable than models built on machine learning of extracting and learning the characteristics of time series data.

In order to anticipate future values for stock prices, RNN based on LSTM is utilized. The results of our model have been encouraging. The testing outcomes support the claim that our model can track the development of opening prices for both assets. In our future work, we'll look for the sets that best suit our resources and increase the accuracy of our forecasts in terms of data length and training epochs.

LSTMs appear to be the best first strategy for tackling the stock price prediction problem, according to the experiments done in this project. Other techniques can blend features input into a traditional ANN regressor with features taken from LSTM or Bi-LSTM models. This method might aid in recovering data that a straightforward LSTM regression model had previously overlooked. Recent studies employ reinforcement learning, graph neural networks, and multi-headed attention processes. Despite such complicated methods being offered frequently, LSTM-based approaches are also being actively employed and investigated lately—Visit Stock Market Prediction | Papers with Code to view the most recent articles in this field.

Further many more improvements and additions can be made to the existing models and techniques for stock price prediction. Deep learning models that take into account financial news stories and financial criteria like a closing price could be created for future work for conceivably superior outcomes, trading volume, profit and loss accounts, etc. Many decades' worth of historical stock price data can be fed into machine learning models, which can then be used to identify important trends and characteristics that characterize a company's stock performance. The model can forecast the performance of stocks in the future if these patterns, metrics, and features are effectively extracted.

**References**

[1] Mehar Vijh , Deeksha Chandola, Vinay Anand Tikkiwal, Arun Kumar (2019) “Stock Closing Price Prediction using Machine Learning Techniques”

[2] Selvin, Sreelekshmy, R. Vinayakumar, E. A. Gopalakrishnan, Vijay Krishna Menon, and K. P. Soman. (2017) “Stock price prediction using LSTM, RNN and CNN-s

[1] Derrick Mwiti, [Data and Notebook for the Stock Price Prediction Tutorial](https://github.com/mwitiderrick/stockprice)(2018), Github

[3] Yahoo Finance - Business Finance Stock Market News, [ Accessed on June 25, 2023]

[4] Simplilearn [Online]. https://www.simplilearn.com/tutorials/machine-learning-tutorial/stock-price-prediction-using-machine-learning [ Accessed on June 27, 2023]

[5] ScienceDirect [Online] <https://www.sciencedirect.com/science> [ Accessed on June 30, 2023]

[6] Arxiv [Online] <https://arxiv.org/ftp/arxiv/papers/2009f> [ Accessed on June 30, 2023]

[7] Huang, Wei, Yoshiteru Nakamori, and Shou-Yang Wang. (2005) “Forecasting stock market movement direction with support vector machine.” Computers & Operations Research 32 (10) : 2513-2522.