(Department of Computer Science and Technology)

Mini Project Report

On

Stock Price Prediction using Machine Learning Techniques

*Presented by*

ISHAN CHAUHAN

Roll No: 2016781

*Under the guidance of*



Mr. Kireet Joshi

INTRODUCTION

1.1 Context

This project has been done as a part of my course for the B.Tech.

Computer Science and Technology at Graphic Era (Deemed to be

University). Supervised by Mr. Kireet Joshi Sir. I had 4-5 months to complete this project.

1.2 Motivations

Being keenly interested in learning everything that is related to

Machine Learning, the independent project was a great occasion to give me the time to learn and confirm my interest for this field. The fact that we can predict stock price according to our needs whether it is for our private use or for the public, brainstorming the ideas. Learn a lot about Machine Learning Algorithms.

That’s why I decided to conduct my project about the Stock Price Prediction using Machine Learning.

1.3 Idea

This project was motivated by my desire to investigate about the

prediction of stock price, how stock market works and how it can be predict in near future.

That’s why I chose to work on this project.

THE PROJECT

Basic idea of the project -

Long short term memory (LSTM) is a model that increases the memory of recurrent neural networks. Recurrent neural networks hold short term memory in that they allow earlier determining information to be employed in the current neural networks. For immediate tasks, the earlier data is used.  Unlike standard [feedforward neural networks](https://en.wikipedia.org/wiki/Feedforward_neural_network), LSTM has feedback connections.  We may not possess a list of all of the earlier information for the neural node. In RNNs, LSTMs are very widely used in Neural networks. Their effectiveness should be implemented to multiple sequence modelling problems in many application domains like video, NLP, geospatial, and time-series.

One of the main issues with RNN is the vanishing gradient problem, and it emerges due to the repeated use of the same parameters, in RNN blocks, at each step. We must try to use different parameters to overcome this problem at each time step.

We try to find a balance in such a situation. We bring novel parameters at each step while generalizing variable-length sequences and keeping the overall amount of learnable parameters constant. We introduce gated RNN cells like LSTM and GRU.

Gated cells hold internal variables, which are Gates. This value of each gate at each time step depends on the information at that time step, including early states. The value of the gate then becomes multiplied by the different variables of interest to influence them. Time-series data is a series of data values gathered over time interims, allowing us to trace differences over time. Time-series data can trace progress over milliseconds, days, and years.

Early, our perspective of time-series data meant more static; the everyday highs and lows under temperature, the opening and closing amount of the stock market. Now we will go to the coding part. We will implement LSTM on the stocks dataset.

* Methodology followed:
* Language used: Python
* Tool used: Jupyter Notebook
* Installation: We need to install Libraries of Python:
* Pandas
* Numpy
* Matplotlib
* Sklearn
* Keras

**Now, let us try to understand the code:**

## Reading data:

## 

## 2-Exploring Dataset:

## 

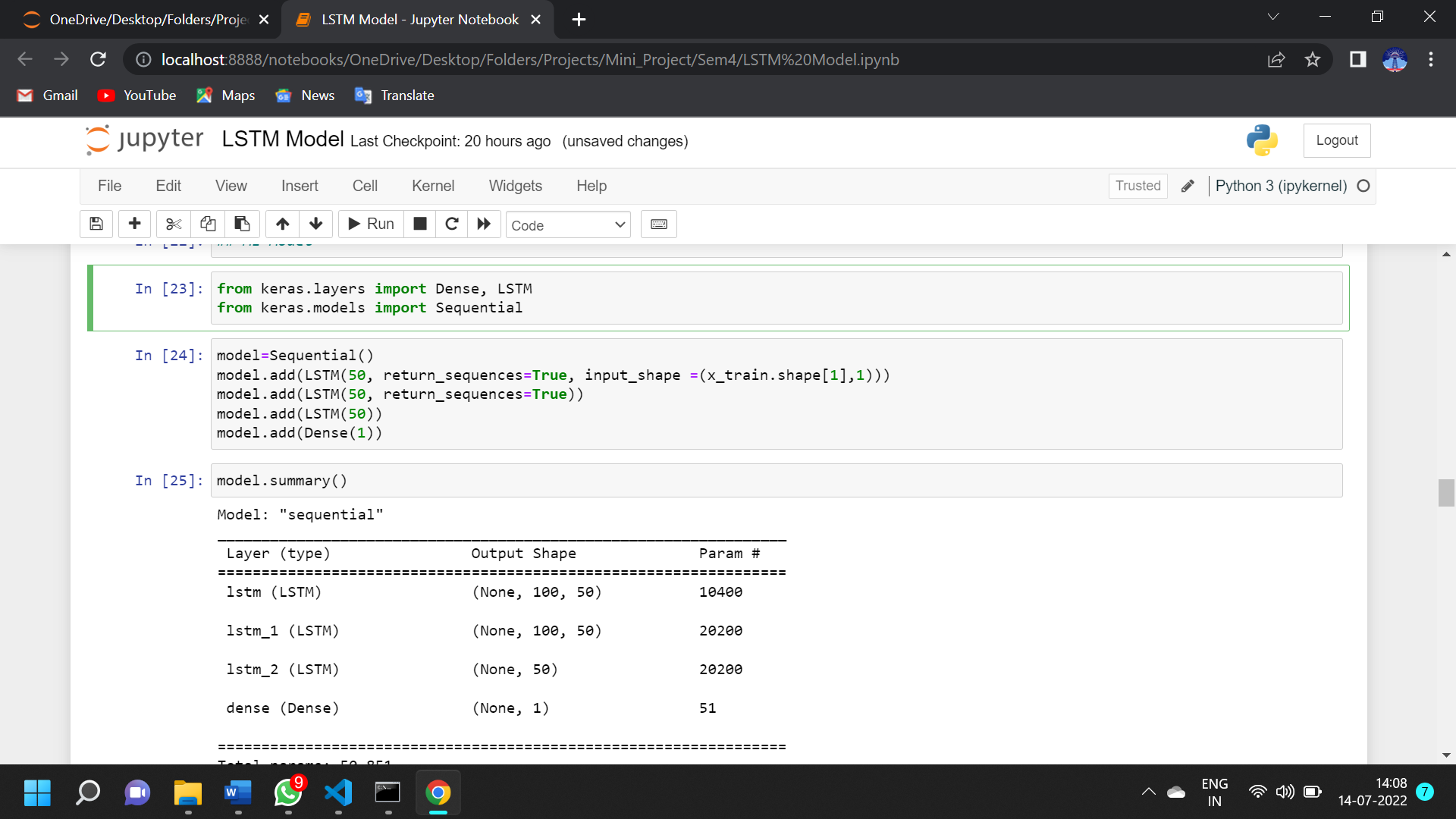
**3- Data Pre-processing:**

I have used Min-max scaler from sklearn for scaling the data so that we can bring all the price values to a common scale. I have than use 70 % data for training and the rest 30% for testing and assign them to separate variables.

## 

**4-Implementation of LSTM model:**

In the next step, from keras I imported Dense and LSTM model.  In this model designing I have used the Sequential model imported from Keras and required libraries are imported.

****

I have used 3 LSTM layers in my model. Then I have used Adam optimizer for the calculating the Mean squared error. It is the loss function for optimizing the problem. Mean absolute error is the metric used in our LSTM network as it is associated with time-series data.

**Prediction:**

## After fitting the data with our model we use it for prediction. We must use Every Detail from the processed values of given datasets and then put it into our model to get back the original value with the transformed function. Now we can use this data to visualize the prediction.

## 

## 

## Conclusion

In this project, we explored LSTM and stock price using LSTM. We then use the moving average to skip first 100 days values and then at last, visualized the closing price value after using LSTM.

**Reference:**

https://the-learning-machine.com/article/dl/long-short-term-memory

\_\_\_\_\_\_\_\_\_\_\_\_\_\_Thankyou\_\_\_\_\_\_\_\_\_\_\_\_\_