

STOCK PRICE PREDICTION USING LSTM

A Project Work Synopsis

Submitted in the partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING

IN

ARTIFICIAL INTELLEGEENCE & MACHINE LEARNING

Submitted by:

VINITA RAI

20BCS6878

SHWETANSHU SOOD

20BCS6899

VAIBHAV SHARMA

20BCS6904

PRIYANSHU SINGH

20BCS6896

Under the Supervision of:

Mr. Sant Kumar Maurya



**CHANDIGARH
UNIVERSITY**

Discover. Learn. Empower.

CHANDIGARH UNIVERSITY, GHARUAN, MOHALI - 140413,

PUNJAB

OCTOBER, 2022

Table of Contents

Title Page

Abstract

List of Figures

- 1. INTRODUCTION***
 - 1.1 Problem Definition
 - 1.2 Project Overview/Specifications* (page-1 and 3)
 - 1.3 Hardware Specification
 - 1.4 Software Specification

- 2. LITERATURE SURVEY**
 - 2.1 Existing System
 - 2.2 Proposed System

- 3. PROBLEM FORMULATION**

- 4. RESEARCH OBJECTIVES**
- 5. METHODOLOGY**
- 6. CONCLUSION AND FUTURE SCOPE**
- 7. TENTATIVE CHAPTER PLAN FOR THE PROPOSED WORK**
- 8. REFERENCES**

ABSTRACT

In this project we attempt to implement machine learning approach to predict stock prices. Machine learning is effectively implemented in forecasting stock prices. The objective is to predict the stock prices in order to make more informed and accurate investment decisions. We propose a stock price prediction system that integrates mathematical functions, machine learning, and other external factors for the purpose of achieving better stock prediction accuracy and issuing profitable trades.

There are two types of stocks. You may know of inter-day trading by the commonly used term "day trading." Inter-day traders hold securities positions from at least one day to the next and often for several days to weeks or months. This project seeks to solve the problem of Stock Prices Prediction by utilizes Deep Learning models, Long-Short Term Memory (LSTM) Neural Network algorithm, to predict future stock values. LSTM(s) are very powerful in sequence prediction problems because they're able to store past information. This is important in our case because the previous price of a stock is crucial in predicting its future price. While predicting the actual price of a stock is an uphill climb, we can build a model that will predict whether the price will go up or down.

List of figures

- A) Graph of stock price
- B) Lstm Architecture
- C) Flow chart

INTRODUCTION

1.1 PROBLEM DEFINITION

In this project we attempt to implement machine learning approach to predict stock prices. Machine learning is effectively implemented in forecasting stock prices. The objective is to predict the stock prices in order to make more informed and accurate investment decisions. The system must be able to access a list of historical prices. It must calculate the estimated price of stock based on the historical data for the next 30 days. It must also provide an instantaneous visualization of the market index in a neatly formatted Python-Based Web App.

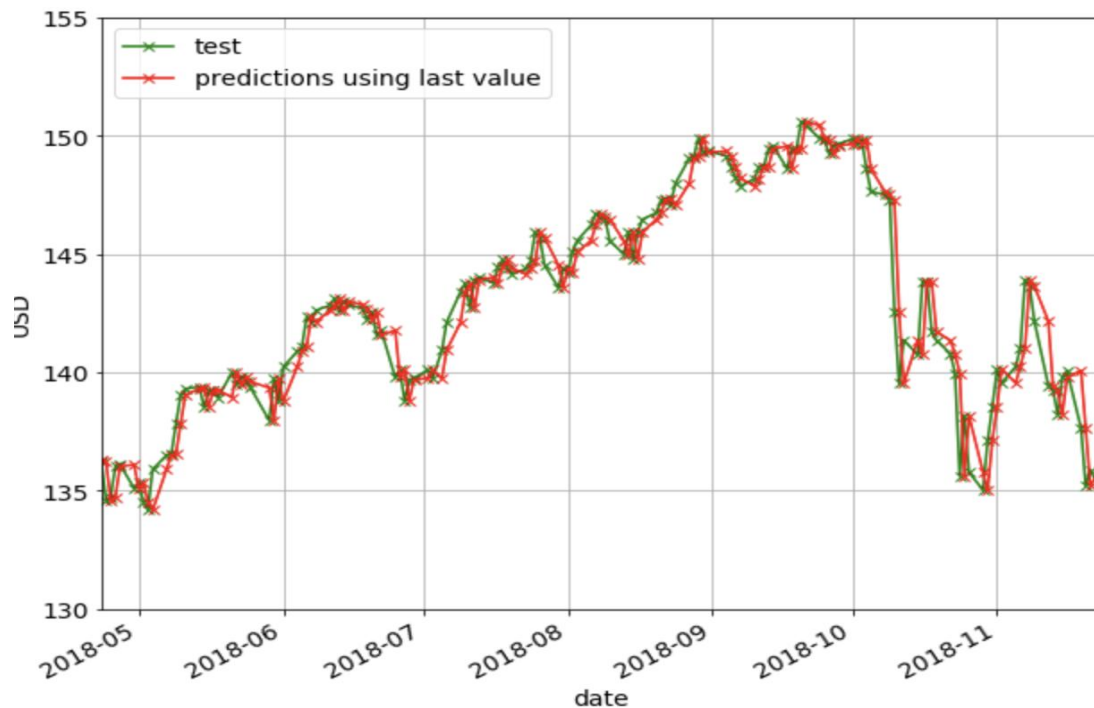
1.2 PROJECT OVERVIEW

1. Broadly, stock market analysis is divided into two parts – Fundamental Analysis and Technical Analysis.

- Fundamental Analysis involves analyzing the company's future profitability on the basis of its current business environment and financial performance.
- Technical Analysis, on the other hand, includes reading the charts and using statistical figures to identify the trends in the stock market

2. The vital part of machine learning is the dataset used. The dataset should be as concrete as possible because a little change in the data can perpetuate massive changes in the outcome. In this project, supervised machine learning is employed on a dataset obtained from Yahoo Finance. This dataset comprises of following five variables: open, close, low, high and volume. Open, close, low and high are different bid prices for the stock at separate times with nearly direct names. The volume is the number of shares that passed from one owner to another during the

3. Time Series forecasting & modelling plays an important role in data analysis. Time series analysis is a specialized branch of statistics used extensively in fields such as Econometrics & Operation Research. Time Series is being widely used in analytics & data science. Stock prices are volatile in nature and price depends on various factors. The main aim of this project is to predict stock prices using Long short term memory (LSTM).



a) Graph of stock price

Stock price example

As seen in the picture the prediction for the next day is done by using the data gathered on the previous days. The red lines which are the predictions by the model should closely resemble the actual data, so that model can become reliable to be used in actual trading and transactions.

1.3 HARDWARE SPECIFICATION

- 8 GB RAM
- Processor - 1.5–4.5x
- Monitor – 15.6”
- Keyboard - 2.4GHz USB wireless receiver 1.4

1.4 SOFTWARE SPECIFICATION

- Python Compiler with required Libraries and Modules
- Language: Python
- Operating System: Windows 7/8/10/11
- Modules like:
 - A) Yfinance
 - B) Plotly
 - C) Keras
 - D) Tensorflow
 - E) Pandas
 - F) Numpy
 - G) Dash

LITERATURE REVIEW

2.1 EXISTING METHODS

2.1.1 Stock Market Prediction Using Machine Learning The research work done by V Kranthi Sai Reddy Student, ECM, Sreenidhi Institute of Science and Technology, Hyderabad, India. In the finance world stock trading is one of the most important activities. Stock market prediction is an act of trying to determine the future value of a stock other financial instrument traded on a financial exchange. This paper explains the prediction of a stock using Machine Learning. The technical and fundamental or the time series analysis is used by the most of the stockbrokers while making the stock predictions. The programming language is used to predict the stock market using machine learning is Python. In this paper we propose a Machine Learning (ML) approach that will be trained from the available stocks data and gain intelligence and then uses the acquired knowledge for an accurate prediction. In this context this study uses a machine learning technique called Support Vector Machine (SVM) to predict stock prices for the large and small capitalizations and in the three different markets, employing prices with both daily and up-to-the-minute frequencies.

2.1.2 Automated Stock Price Prediction Using Machine Learning The research work done by Mariam Moukalled Wassim El-Hajj Mohamad Jaber Computer Science Department American University of Beirut. Traditionally and in order to predict market movement, investors used to analyse the stock prices and stock indicators in addition to the news related to these stocks. Hence, the importance of news on the stock price movement. Most of the previous work in this industry focused on either classifying the released market news as (positive, negative, neutral) and demonstrating their effect on the stock price or focused on the historical price movement and predicted their future movement. In this work, we propose an automated trading system that integrates mathematical functions, machine learning, and other external factors such as news' sentiments for the purpose of achieving better stock prediction accuracy and issuing profitable trades. Particularly, we aim to determine the price or the trend of a certain stock

for the coming end-of-day considering the first several trading hours of the day. To achieve this goal, we trained traditional machine learning algorithms and created/trained multiple deep learning models taking into consideration the importance of the relevant news. Various experiments were conducted, the highest accuracy (82.91%) of which was achieved using SVM for Apple Inc. (AAPL) stock.

2.2 PROPOSED SYSTEM

The prediction methods can be roughly divided into two categories, statistical methods and artificial intelligence methods. Statistical methods include logistic regression model, ARCH model, etc. Artificial intelligence methods include multi-layer perceptron, convolutional neural network, naive Bayes network, back propagation network, single-layer LSTM, support vector machine, recurrent neural network, etc. They used Long short-term memory network (LSTM).

Long short-term memory network: Long short-term memory network (LSTM) is a particular form of recurrent neural network (RNN). Long Short-Term Memory (or) LSTMs are widely used for sequence prediction problems and have proven to be extremely effective. The reason they work so well is because LSTM is able to store past information that is important and forget the information that is not. LSTM has three types of gates:

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

OBJECTIVES

1. To describe the stocks of different companies by downloading data live from yahoo finance and training using time series analysis and LSTM.
2. To develop an optimal way for giving some accurate future predictions of 1-5 days according to the data analysis.
3. To research the outcomes in form of graphs and statistics.

PROBLEM FORMULATION

The technology is growing day-by-day and still there exist many factors that have the scope of improvement. The existing systems have many cons few of which are solved and improved by our system. Mentioned below are few cons of existing system and the solution that our system provides to rectify these problems of existing system.

1) We use the StandardScaler, rather than the MinMaxScaler as you might have seen before. The reason is that stock prices are ever-changing, and there are no true min or max values. It doesn't make sense to use the MinMaxScaler, although this choice probably won't lead to disastrous results at the end of the day.

2) Stock price data in its raw format can't be used in an LSTM model directly; we need to transform it using our pre-defined `'extract_seqX_outcomeY'` function. For instance, to predict the 51st price, this function creates input vectors of 50 data points prior and uses the 51st price as the outcome value.

METHODOLOGY

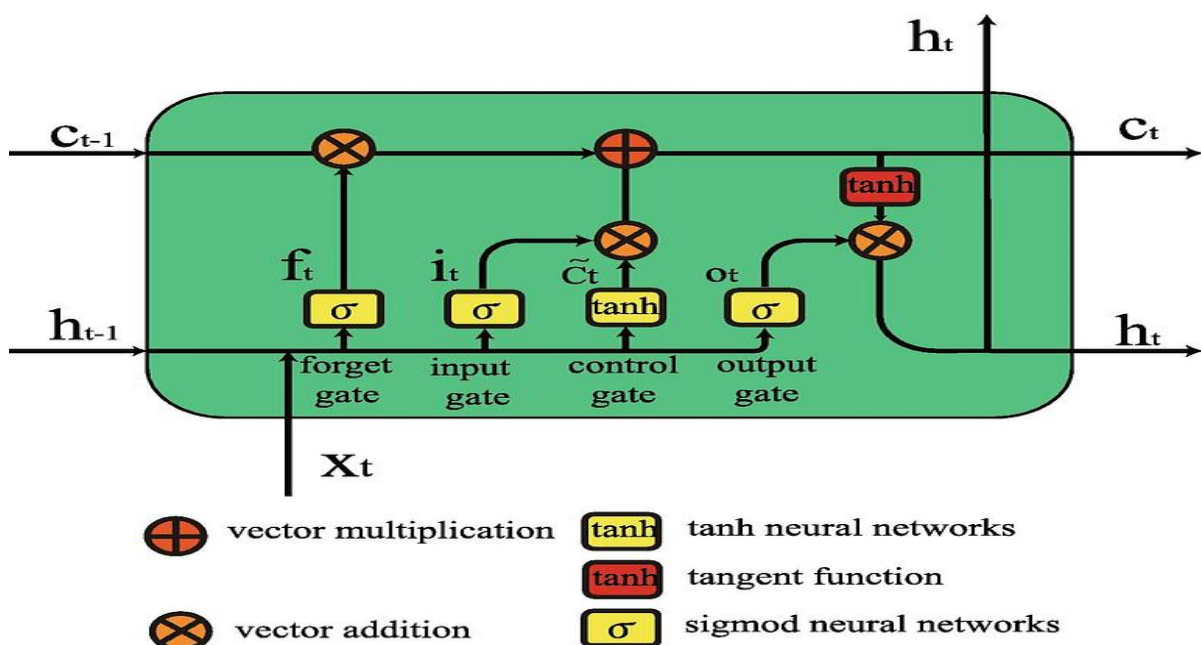
Long short-term memory network: Long short-term memory network (LSTM) is a particular form of recurrent neural network (RNN).

Working of LSTM:

LSTM is a special network structure with three “gate” structures. Three gates are placed in an LSTM unit, called input gate, forgetting gate and output gate. Used Long Short-term Memory (LSTM) with embedded layer and the LSTM neural network with automatic encoder.

- LSTM is used instead of RNN to avoid exploding and vanishing gradients.
- In this project python is used to train the model, MATLAB is used to reduce dimensions of the input. MySQL is used as a dataset to store and retrieve data.
- The historical stock data table contains the information of opening price, the highest price, lowest price, closing price, transaction date, volume and so on.

LMS filter: The LMS filter is a kind of adaptive filter that is used for solving linear problems. The idea of the filter is to minimize a system (finding the filter coefficients) by minimizing the least mean square of the error signal.

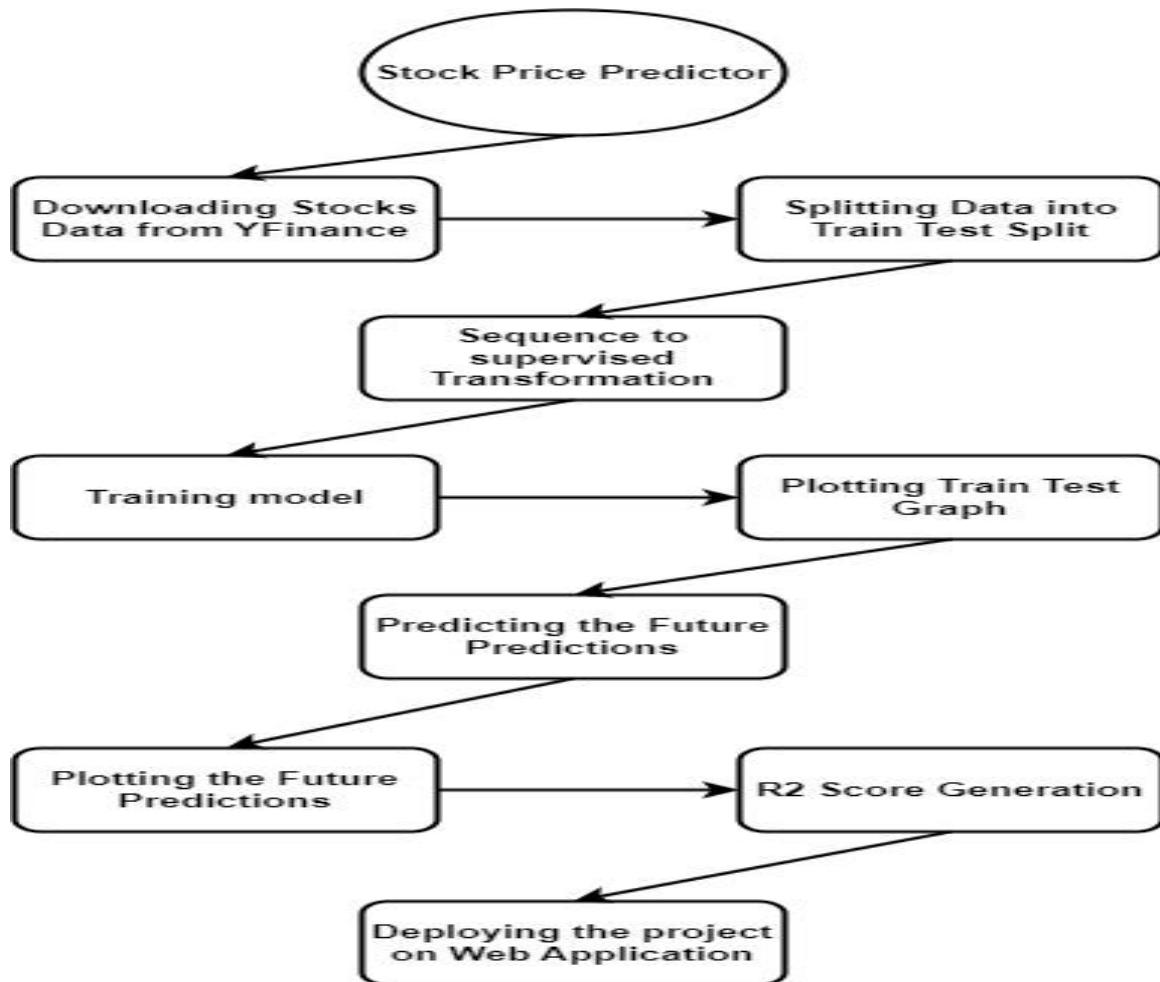


b)Lstm architecture

Procedure to perform this project

- The foremost step for training our stock price predictor is to gather the data for training. For this task rather than collecting the data manually through web scraping we are going to import our data from a python library called Yfinance.
- Yfinance is a python package that enables us to fetch historical market data from Yahoo Finance API. We can easily download historical stock data from yfinance.
- Next step is to split the data for training and testing respectively and in the same process transform the data using Normalization and scaling techniques to improve data quality.
- The final step of our first methodology process is to plot our training data using plotly library for easy visualization of the data we are working with.
- The plotly Python library is an interactive, open-source plotting library that supports over 40 unique chart types covering a wide range of statistical, financial, geographic, scientific, and 3-dimensional use-cases.
- After the plotting of our training data we will try to apply LSTM and Keras sequential mode to train our model on the training set. Long Short Term Memory networks generally called LSTM's is a variety of recurrent neural networks (RNNs) that are capable of learning long-term dependencies, especially in sequence prediction problems. LSTM has feedback connections, i.e., it is capable of processing the entire sequence of data.
- Next we plot our predictions to have a visual insight of how our model has predicted the stock to rise, fall or remain same for the upcoming days as seen in many stock data provider applications.
- Simultaneously we try to find out the R^2 score of our model to judge on the parameters of performance evaluation so that we can improve our model.
- R^2 is a measure of the goodness of fit of a model. In regression, the R^2 coefficient of determination is a statistical measure of how well the regression predictions approximate

the real data points. The final steps include developing a web application to host our model and present our visualizations to all the stock trading application users for easy and faster access



C)Flow Chart

CONCLUSION AND FUTURE SCOPE

6.1 CONCLUSION

- Successfully Constructed a working model that can predict the future stock price for most stocks with reasonable accuracy that can be used in the finance market for earning profits and minimize loss.
- Plotted multiple visualizations that can help investors visualize the rise and downfall of the price of stocks over time as well as the future stock prices in the form of line graph.
- Added R2 scores and actual values of stocks versus the predicted values to help user to see the accuracy and performance of the model to help deciding in purchasing or selling their stocks based on the model.
- Also, build a web application that can host the model and results and plots of the model for faster and easier access anywhere across the globe.

6.2 FUTURE SCOPE

- Ability to fetch and train with real time data updated every few seconds with the help application programming interfaces(API) to get the predictions one after other very quickly to help accelerate decision making on the transactions of stocks.
- More robust plots that can generate and update graphs quickly to see real time moving graphs as provided by some crypto trackers but with the added ability of updating predictions every few seconds.

Release a dedicated application based software that can be installed and can run independently without the help of any web browsers

TENTATIVE CHAPTER PLAN FOR THE PROPOSED WORK

CHAPTER 1: INTRODUCTION

This chapter will cover the overview of the project - “Windows Assistant for users”. It gives a brief knowledge about what the project is, what all it can do. It includes:

- Problem statement
- Importance of the system
- Aims and Objectives
- Introduction of study area

CHAPTER 2: LITERATURE REVIEW

This chapter includes the literature available for Windows Assistant. The findings of the researchers will be highlighted which will become basis of current implementation. It includes:

- The details like introduction, pros and about the existing system till date.
- The details about the proposed system.

CHAPTER 3: BACKGROUND OF PROPOSED METHOD

This chapter will provide introduction to the concepts which are necessary to understand the proposed system. This includes prerequisites, used modules and libraries.

CHAPTER 4: METHODOLOGY

This chapter will cover the technical details of the proposed approach. This gives the knowledge about the modules used, declaration and definition of functions, the logic behind the functions, algorithm and flowcharts (wherever necessary).

CHAPTER 5: EXPERIMENTAL SETUP

This chapter will provide information about the subject system and tools used for evaluation of proposed method. This will include the requirements of the project and the most suitable solution for it.

CHAPTER 6: RESULTS AND DISCUSSION

The result of proposed technique will be discussed in this chapter. This will display the whole output of the proposed work. It also contains alternative explanations about the output and significance of the project.

CHAPTER 7: CONCLUSION AND FUTURE SCOPE

The major finding of the work will be presented in this chapter. Also, directions for extending the current study will be discussed. The conclusion will include everything from top in brief. It contains introduction, needs, advantages, problems, scope and suggestions.

REFERENCES

- <https://www.analyticsvidhya.com/blog/2021/12/stock-price-prediction-using-lstm/>
- <https://towardsdatascience.com/lstm-for-google-stock-price-prediction-e35f5cc84165>
- <https://www.bajajfinserv.in/insights/share-trading>
- <https://www.intechopen.com/chapters/72381>
- <https://tradebrains.in/types-risk-involved-in-stocks-know/>
- <https://pypi.org/project/yfinance/>
- <https://plotly.com/python/getting-started/>
- https://en.wikipedia.org/wiki/Coefficient_of_determination