A Project Report

on

Stock Market Prediction Using Deep Learning

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

BACHELOR OF ENGINEERING
IN
COMPUTER SCIENCE & ENGINEERING

Under The Supervision of
Ms.Vaishali Gupta
(Assistant Professor)

Submitted By:

Ayush Agrawal
Enroll no: 19021011530
Adm no: 19SCSE1010346

Prabhat Ranjan Rai
Enroll no: 19021011483
Adm no: 19SCSE1010295

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING GALGOTIAS UNIVERSITY, GREATER NOIDA, INDIA

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING
GALGOTIAS UNIVERSITY, GREATER NOIDA

CANDIDATE'S DECLARATION

I/We hereby certify that the work which is being presented in the thesis/project/dissertation, entitled **"Stock Market Prediction Using Deep Learning"** in partial fulfillment of the requirements for the award of the **BACHELOR OF TECHNOLOGY IN COMPUTER SCIENCE AND ENGINEERING** submitted in the **School of Computing Science and Engineering** of Galgotias University, Greater Noida, is an original work carried out during the period of month, **JULY-2021 to DECEMBER-2021**, under the supervision of **Ms.Vaishali Gupta, Assistant Professor, Department of Computer Science and Engineering** of School of Computing Science and Engineering , Galgotias University, Greater Noida

The matter presented in the thesis/project/dissertation has not been submitted by me/us for the award of any other degree of this or any other places.

Ayush Agrawal – 19SCSE1010346

Prabhat Ranjan Rai – 19SCSE1010295

This is to certify that the above statement made by the candidates is correct to the best of my knowledge.

Supervisor

Ms. Vaishali Gupta,

Assistant Professor

CERTIFICATE

The Final Thesis/Project/ Dissertation Viva-Voce examination of **Ayush Agrawal – 19SCSE1010346, Prabhat Ranjan Rai – 19SCSE1010295** has been held on _____ and his/her work is recommended for the award of **BACHELOR OF TECHNOLOGY IN COMPUTER SCIENCE AND ENGINEERING.**

Signature of Examiner(s) **Signature of Supervisor(s)**

Signature of Project Coordinator **Signature of Dean**

Date:

Place:

ABSTRACT

Strategies of the stock market are widely complex and rely on an enormous amount of data. Hence, predicting stock markets has always been a challenge for many researchers and investors. Much research has been done, and many machine learning techniques have been developed to solve complex computational problems and improve predictive capabilities without being explicitly programmed. This research attempts to explore the capabilities of Long Short Term Memory a type of Recurrent Neural Networks in the prediction of future stock markets. Long Short-Term Memory variations with single and multiple feature models are created to predict the value of S&P 500 based on the earnings per share and market to earnings ratio.

We aim to construe the Stacked Long–Short term memory (LSTM) and Multi-layered perceptron intended for the NSE-Stock Market prediction. Stock market prediction can be instrumental in determining the future value of a company stock.It is imperative to say that a successful prediction of a stock's future market could yield significant profit which would be beneficial for those who invested in the pipeline of things including stock market prediction. The model uses the information pertaining to the stocks and contemplates the previous model accuracy to innovate the approach used in our paper. The experimental evaluation is based on the historical data set of National Stock Exchange (NSE). The proposed approach aims to provide models like Stacked LSTM and MLP which perform better than its contemporaries which have been achieved to a certain extent. This can be verified by the results embedded in the paper . The future research can be focused on adding more variables to the model by fetching live data from the internet as well as improving model by selecting more critical factors by ensemble learning.

Contents

List of Figures

CHAPTER-1
Introduction

The stock market can be seen as the public marketplace, where shares and other financial instruments are being sold and bought everyday. Each share represents a portion of a company's ownership, and S&P 500 constitutes shares of the five hundred most important United States companies.

From the appearance of markets, investors explored ways to acquire more knowledge of the companies listed in the market, and further tried to keep up with the enormous number of news feed in the world. With the increase of market size and the speed at which trades are executed investors became less capable on relying on personal experience to identify market patterns. As technology progressed, investors and researches have developed many techniques and various models to solve problems that arise.

Furthermore, most studies are conducted using a single time scale feature of the stock market index, it is therefore reasonable for studying multiple time scale features to determine a more accurate model outcome. It is important to note that markets are affected by many elements such as political, industrial development, market news, social media and economic environments. One reason for the luck of predictability is that appropriate variables to model are unknown and hard to acquire.

1.1 Motivation For Work

Businesses primarily run over customer's satisfaction, customer reviews about their products. Shifts in sentiment on social media have been shown to correlate with shifts in stock markets. Identifying customer grievances thereby resolving them leads to customer satisfaction as well as trustworthiness of an organization. Hence there is a necessity of an un biased automated system to classify customer reviews regarding any problem. In today's environment where we're justifiably suffering from data overload (although this does not mean better or deeper insights), companies might have mountains of customer feedback collected; but for mere humans, it's still impossible to analyse it manually without any sort of error or bias. Oftentimes, companies with the best intentions find themselves in an insights vacuum. You know you need insights to inform your decision making and you know that you're lacking them, but don't know how best to get them. Sentiment analysis provides some answers into what the most important issues are, from the perspective of customers, at least. Because sentiment analysis can be automated, decisions can be made based on a significant amount of data rather than plain intuition.

1.2 Problem Statement

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).

CHAPTER-2

Literature Reviews

Stock market prediction can be predicted using AI and machine learning models in machine learning fields. Using the SVM model for stock market prediction. SVM is one of the machine learning algorithms which works on classification algorithms. It is used to get a new text as an output. Applying Multiple Linear Regression with Interactions to predict the trend in stock markets (Osman Hegazy et al. 2013; V Kranthi Sai Reddy, 2018; a Banerjee et al. 2020; Lufuno Ronald Marwala). Random Walk Hypothesis which is proposed by Horne, j. C et al 1997 which is used to predict stock markets, Horne j.c said that the stock values are changes random and the past market values are not dependent on current values. EMH is different from the Random walk hypothesis but the EMH works mainly on Short term patterns for predicting stock markets. Manh Ha Duong Boris's Siliverstovs, 2006 search the abstraction between equity markets and combined finances in Key Eu nations like UK and Germany. Acceleration in Eu nations investments is apt to results successful even Stronger correlation between the different Eu nations and equity markets. This operation may also lead to a merge in financial development between EU nations, if advancements in stock markets affect real financial instruments, such as investing and Consuming. Fahad Almudhaf et al, 2012, tests the weak-form market efficiency of CIVETS over the period 2002–2012. The random walk hypothesis process is used in CIVETS. In an efficient stock market, the equity values must follow a random walk hypothesis, when it comes to the future market, the values are changing randomly and unpredictable. Everyday returns for rising and improved markets have been tested for random walks.

LSTM algorithm consists of a Recurrent Neural network to encode data. The algorithm inputs are economic news headings infusion From Bloomberg and Reuters. Long Short- term Memory with embedded layer and the LSTM with the automatic encoder in the stock market for predicting stock values. The Xiongwen Pang et al. Used an automatic encoder and embedded layer to vectorizing the values by using LSTM layers. Correlation coefficients in stocks are selected randomly and predicted using ARIMA and the neural network approach. In this RNN and LSTM algorithms are implemented. M. Nabipour et al. Used different machine learning and deep learning algorithms for predicting stockvalues such as random forest, decision tree and neural networks. LSTM gives the most accurate results and it has the best ability to fit. LSTM gives the best results while predicting stock markets with the least error rate (Hyeong Kya Choi,2018; Huicheng Liu, 2018; M. Nabipour et al,2020; Xiongwen Pang et al, 2020).

The LMS filter is a type of adaptive filter which is used for solving linear problems. The idea of the filter is to find the filter coefficients and to minimize a system by reducing the least mean square of the error value (Asep Juarna, 2017; Eleftherios Giovanis, 2018). They used a hybrid model for predicting the stock values by using deep learning and ML methodologies and they built a model using deep regression based on CNN. Here they used CNN for parameters, thereby increase the no of loops will stabilize the validation loss. They also tested using DL and a hybrid ML algorithm for stock market prediction. Vivek Rajput and Sarika Bobde used sentiment analysis from online posts or multimedia and data mining is used. In sentiment analysis, they are trying to get emotion either positive or negative based on the textual information available on social networks. sentiment analysis for predicting the stock market to get more accurate and efficient results.

2.1 Existing System

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 Forecasting the Stock Market Index Using Artificial Intelligence Techniques

The research work done by Lufuno Ronald Marwala A dissertation submitted to the Faculty of Engineering and the Built Environment, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science in Engineering. The weak form of Efficient Market hypothesis (EMH) states that it is impossible to forecast the future price of an asset based on the information contained in the historical prices of an asset. This means that the market behaves as a random walk and as a result makes forecasting impossible. Furthermore, financial forecasting is a difficult task due to the intrinsic complexity of the financial system. The objective of this work was to use artificial intelligence (AI) techniques to model and predict the

future price of a stock market index. Three artificial intelligence techniques, namely, neural networks (NN), support vector machines and neuro-fuzzy systems are implemented in forecasting the future price of a stock market index based on its historical price information. Artificial intelligence techniques have the ability to take into consideration financial system complexities and they are used as financial time series forecasting tools.

Two techniques are used to benchmark the AI techniques, namely, Autoregressive Moving Average (ARMA) which is linear modelling technique and random walk (RW) technique. The experimentation was performed on data obtained from the Johannesburg Stock Exchange. The data used was a series of past closing prices of the All Share Index. The results showed that the three techniques have the ability to predict the future price of the Index with an acceptable accuracy. All three artificial intelligence techniques outperformed the linear model. However, the random walk method out performed all the other techniques. These techniques show an ability to predict the future price however, because of the transaction costs of trading in the market, it is not possible to show that the three techniques can disprove the weak form of market efficiency. The results show that the ranking of performances support vector machines, neuro-fuzzy systems, multilayer perceptron neural networks is dependent on the accuracy measure used.

2.1.3 Indian stock market prediction using artificial neural networks on tick data

The research work done by Dharmaraja Selvamuthu, Vineet Kumar and Abhishek Mishra Department of Mathematics, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India. A stock market is a platform for trading of a company's stocks and derivatives at an agreed price. Supply and demand of shares drive the stock market. In any country stock market is one of the most emerging sectors. Nowadays, many people are indirectly or directly related to this sector. Therefore, it becomes essential to know about market trends. Thus, with the development of the

stock market, people are interested in forecasting stock price. But, due to dynamic nature and liable to quick changes in stock price, prediction of the stock price becomes a challenging task. Stock m Prior work has proposed effective methods to learn event representations that can capture syntactic and semantic information over text corpus, demonstrating their effectiveness for downstream tasks such as script event prediction. On the other hand, events extracted from raw texts lacks of common-sense knowledge, such as the intents and emotions of the event participants, which are useful for distinguishing event pairs when there are only subtle differences in their surface realizations. To address this issue, this paper proposes to leverage external common-sense knowledge about the intent and sentiment of the event.

Experiments on three event-related tasks, i.e., event similarity, script event prediction and stock market prediction, show that our model obtains much better event embeddings for the tasks, achieving 78% improvements on hard similarity task, yielding more precise inferences on subsequent events under given contexts, and better accuracies in predicting the volatilities of the stock market1. Markets are mostly a nonparametric, non-linear, noisy and deterministic chaotic system (Ahangar et al. 2010). As the technology is increasing, stock traders are moving towards to use Intelligent Trading Systems rather than fundamental analysis for predicting prices of stocks, which helps them to take immediate investment decisions. One of the main aims of a trader is to predict the stock price such that he can sell it before its value decline, or buy the stock before the price rises. The efficient market hypothesis states that it is not possible to predict stock prices and that stock behaves in the random walk. It seems to be very difficult to replace the professionalism of an experienced trader for predicting the stock price. But because of the availability of a remarkable amount of data and technological advancements we can now formulate an appropriate algorithm for prediction whose results can increase the profits for traders or investment firms.

Thus, the accuracy of an algorithm is directly proportional to gains made by using the algorithm.

2.1.4 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.

CHAPTER-3
Proposed Model

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).

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. While information enters the LSTM's network, it can be selected by rules. Only the information conforms to the algorithm will be left, and the information that does not conform will be forgotten through the forgetting gate.

The experimental data in this paper are the actual historical data downloaded from the Internet. Three data sets were used in the experiments. It is needed to find an optimization algorithm that requires less resources and has faster convergence speed.

• 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.

• The accuracy of this LSTM model used in this project is 57%

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.

Hardware Requirements

• RAM: 4 GB

• Storage: 500 GB

• CPU: 2 GHz or faster

• Architecture: 32-bit or 64-bit

Software Requirements:

• Python 3.5 in Google Colab is used for data pre-processing, model training and prediction.

• Operating System: windows 7 and above or Linux based OS or MAC OS

FUNCTIONAL REQUIREMENTS:

The prediction shall abide by the following functional requirements:

1. Prior to application of stock recommendations, the database is updated by the latest values.

2. The charts and comparison of the companies would be done only on the latest data stock market data.

3. The user is provided with a login, logging into which enables the user to view his past stock purchases and future recommendations.

4. The user can look previous data Information which was collected.

5. Each user has a friend list and can also be recommended on their buying patterns.

6. The user can also be recommended on the basis of the trending stocks which would require the data regarding the stocks.

NON-FUNCTIONAL REQUIREMENTS:

Product properties

• Usability:

It defines the user interface of the software in terms of simplicity of understanding the user interface of stock prediction software, for any kind of stock trader and other stakeholders in stock market.

• Efficiency:

maintaining the possible highest accuracy in the closing stock prices in shortest time with available data.

• Performance:

It is a quality attribute of the stock prediction software that describes the responsiveness to various user interactions with it.

SYSTEM ARCHITECTURE

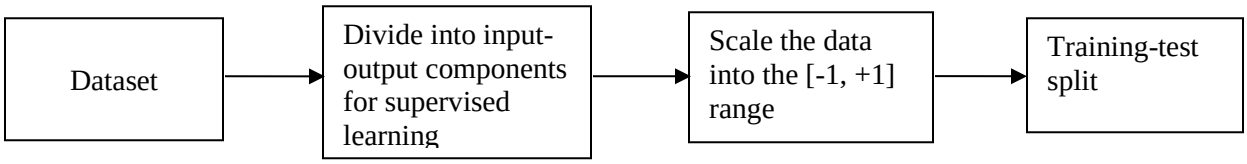

Fig. Pre-processing of data

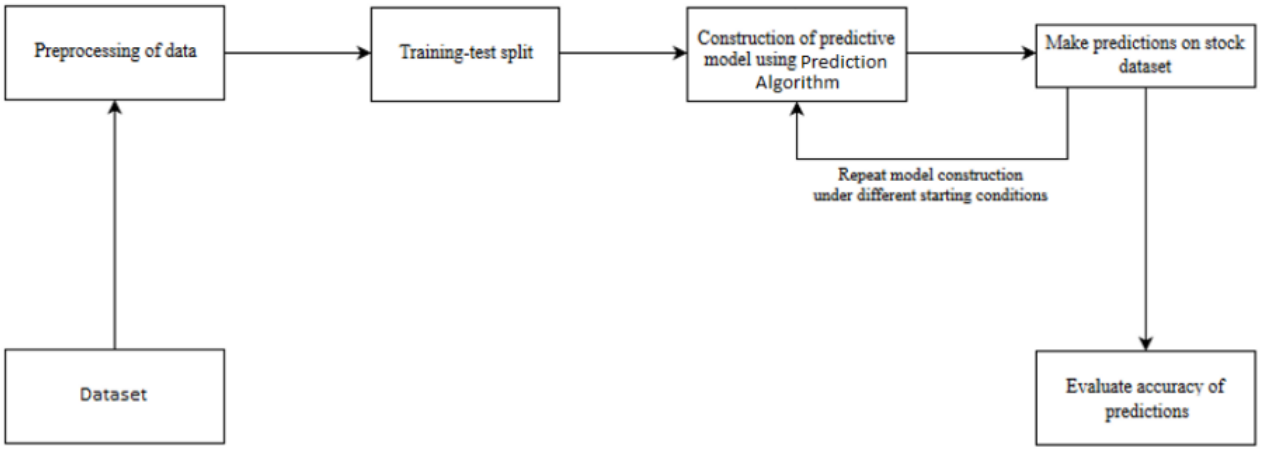

Fig. Overall Architecture

CHAPTER-4

Design

UML Diagram

A UML diagram is a partial graphical representation (view) of a model of a system under design, implementation, or already in existence. UML diagram contains graphical elements (symbols) - UML nodes connected with edges (also known as paths or flows) - that represent elements in the UML model of the designed system. The UML model of the system might also contain other documentation such as use cases written as templated texts.

Flow Chart Diagram

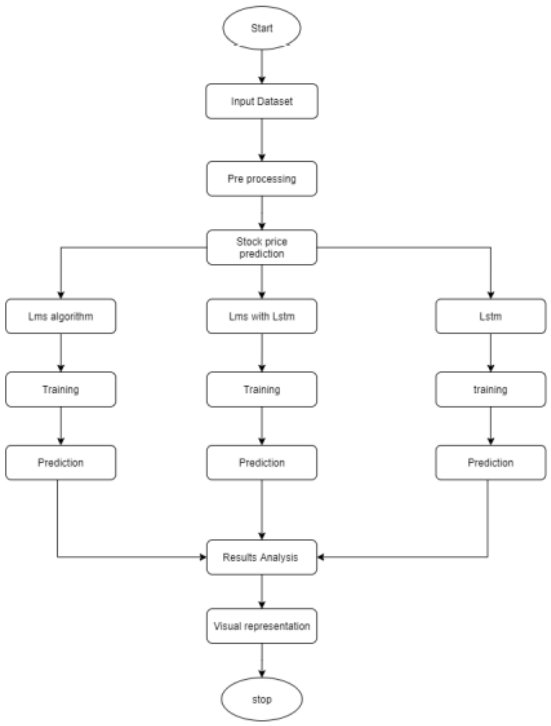

Fig. Flow chart diagram

Use Case Diagram

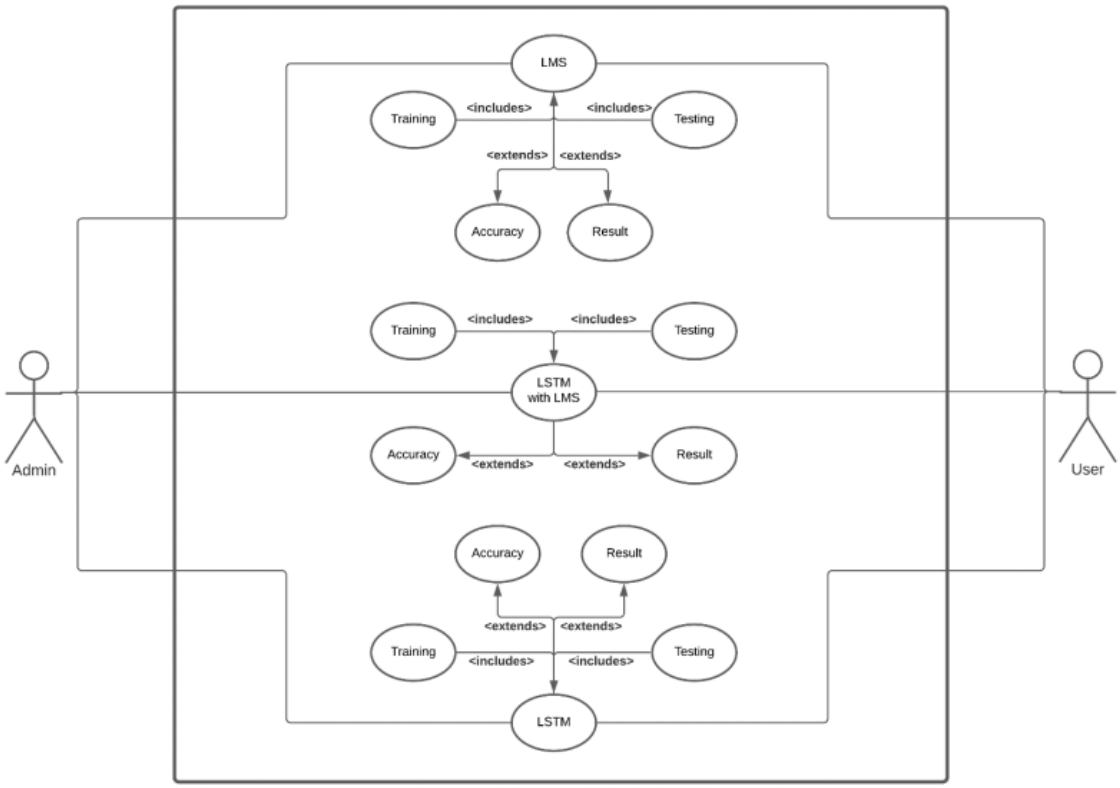

Fig. Use Case Diagram

Activity Diagram

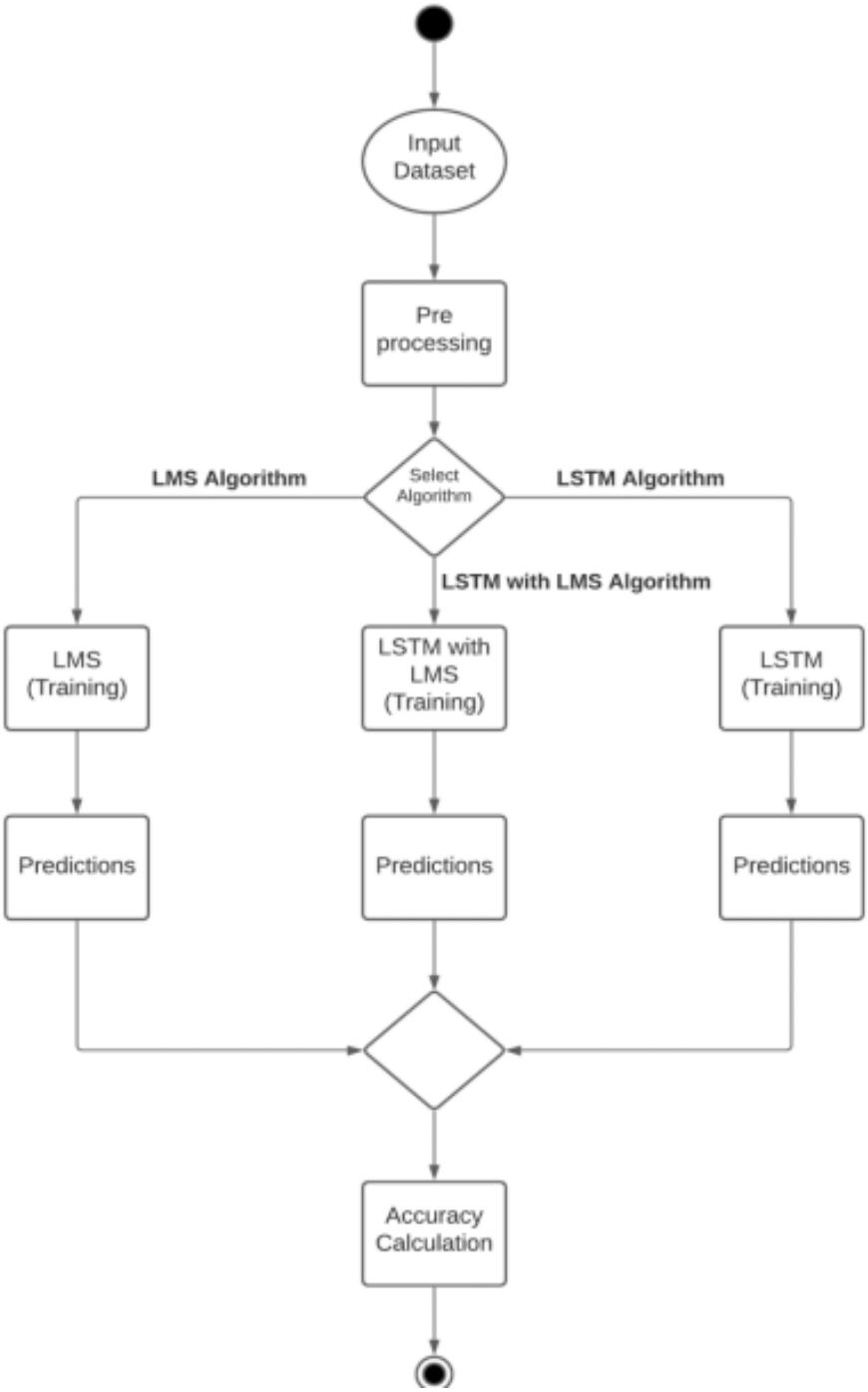

Fig. Activity Diagram

CHAPTER-5
Conclusion

In this project, we are predicting closing stock price of any given organization, we developed a web application for predicting close stock price using LMS and LSTM algorithms for prediction. The testing results confirm that the LSTM variants are capable of tracing the evolution of closing price for long term transactions leaving much room for improvement of daily transactions. We have applied datasets belonging to Google, Nifty50, TCS, Infosys and Reliance Stocks and achieved above 95% accuracy for these datasets.

Future Work

- We want to extend this application for predicting cryptocurrency trading.
- We want to add sentiment analysis for better analysis.

Reference

[1] Stock Price Prediction Using LSTM on Indian Share Market by Achyut Ghosh, Soumik Bose1, Giridhar Maji, Narayan C. Debnath, Soumya Sen

[2] S. Selvin, R. Vinayakumar, E. A. Gopalkrishnan, V. K. Menon and K. P. Soman, "Stock price prediction using LSTM, RNN and CNN-sliding window model," in International Conference on Advances in Computing, Communications and Informatics, 2017.

[3] Murtaza Roondiwala, Harshal Patel, Shraddha Varma, “Predicting Stock Prices Using LSTM” in Undergraduate Engineering Students, Department of Information Technology, Mumbai University, 2015.

[4] Xiongwen Pang, Yanqiang Zhou, Pan Wang, Weiwei Lin, “An innovative neural network approach for stock market prediction”, 2018

[5] Ishita Parmar, Navanshu Agarwal, Sheirsh Saxena, Ridam Arora, Shikhin Gupta, Himanshu Dhiman, Lokesh Chouhan Department of Computer Science and Engineering National Institute of Technology, Hamirpur – 177005, INDIA - Stock Market Prediction Using Machine Learning.

[6] Pranav Bhat Electronics and Telecommunication Department, Maharashtra Institute of Technology, Pune. Savitribai Phule Pune University - A Machine Learning Model for Stock Market Prediction.

[7] Anurag Sinha Department of computer science, Student, Amity University Jharkhand Ranchi, Jharkhand (India), 834001 - Stock Market Prediction Using Machine Learning.