
Stock Price Prediction of Nifty 50

Submitted in partial fulfillment of the requirements

For the degree of

Bachelor of Engineering

by

Bhopatrao Nihal N.

Roll No. 5

Kumbhare Sanket R.

Roll No. 36

Wankhede Shashank D.

Roll No. 59

Under the supervision of

Prof. A. S. Kunte



DEPARTMENT OF INFORMATION TECHNOLOGY
KONKAN GYANPEETH COLLEGE OF ENGINEERING
Konkan Gyanpeeth Shaikshanik Sankul, Vengaon Road,
Dahivali, Karjat Dist. Raigad-410201
Academic Year- 2020-21

Certificate

This is to certify that the project entitled **Stock Price Prediction of Nifty 50** is a bonafide work of **Bhopatrao Nihal N.(Roll No. 5), Kumbhare Sanket R.(Roll No. 36), Wankhede Shashank D.(Roll No. 59)** submitted to the University of Mumbai in partial fulfillment of the requirement for the award of the degree of **Undergraduate in DEPARTMENT OF INFORMATION TECHNOLOGY.**

Supervisor/Guide

Prof. A. S. Kunte

Department of Information Technology

Head of Department

Prof. A. S. Kunte

Department of Information Technology

Principal

Dr. M. J. Lengre

Konkan Gyanpeeth College of Engineering

Project Report Approval for B.E.

This thesis / dissertation/project report entitled **Stock Price Prediction of Nifty 50** by **Bhopatrao Nihal N. (Roll No. 5), Kumbhare Sanket R.(Roll No. 36), Wankhede Shashank D. (Roll No. 59)** is approved for the degree of **DEPARTMENT OF INFORMATION TECHNOLOGY.**

Examiners

1. _____

2. _____

Date:

Place:

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original Sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in My submission. I understand that any violation of the above will be cause for disciplinary Action by the Institute and can also evoke penal action from the sources which have thus Not been properly cited or from whom proper permission has not been taken when needed.

Signature

Bhopatrao Nihal N. (Roll No. 5)

Signature

Kumbhare Sanket R. (Roll No. 36)

Signature

Wankhede Shashank D. (Roll No. 59)

Date:

Abstract

Stock market prediction is the act of trying to determine the future value of a Company stock or other financial instrument traded on an exchange. The successful Prediction of a stock's future price is always profitable. We will implement Machine learning algorithms to predict the future stock price of this company, Simple algorithms linear regression

Acknowledgement

Success is nourished under the combination of perfect guidance, care blessing. Acknowledgement is the best way to convey. We express deep sense of gratitude brightness to the outstanding permutations associated with success. Last few years spend in this esteemed institution has molded us into confident and aspiring Engineers. We express our sense of gratitude towards our project guide Prof. A. S. Kunte. It is because of his valuable guidance, analytical approach and encouragement that we could learn and work on the project. We will always cherish the great experience to work under their enthusiastic guidance. We are also grateful to our principle Dr. M.J. Lengare who not only supporting us in our project but has also encouraging for every creative activity. We extend our special thanks to all teaching and non-teaching sta, friends and well wishers who directly or indirectly contributing for the success of our maiden mission. Finally, how can we forget our parents whose loving support and faith in us remains our prime source of inspiration. Lastly we would like to thank all those who directly and indirectly helping to complete this project. We would also like to acknowledge with much appreciation the crucial role of the staff of Information Technology Department, who gave the permission to use the all required software/hardware and the necessary material to completing to the project.

Contents

Certificate	i
Project Report Approval	ii
Declaration	iii
Abstract	iv
Acknowledgements	v
Contents	vi
List of Figures	viii
List of Tables	ix
Abbreviations	x
1 INTRODUCTION.....	1
1.1 Introduction.....	1
1.2 Objectives.....	1
1.3 Purpose, Scope and Applicability.....	2
1.3.1 Purpose.....	2
1.3.2 Scope.....	2
1.3.3 Applicability.....	2

2	LITERATURE SURVEY AND PAPER REVIEW.....	3
2.1	Literature Survey.....	3
2.2	Paper Comparison.....	14
3	SURVEY OF MMETHODOLOG.....	15
3.1	Deep Learning.....	15
3.2	Artificial Neural Network.....	15
3.3	Time Series Analysis.....	16
3.4	Recurrent Neural Network.....	16
3.5	Long Short Term Memory.....	16
3.6	Python.....	17
3.7	Django.....	18
4	REQUIREMENTS AND ANALYSIS.....	20
4.1	Problem Definition.....	20
4.2	Requirements Specification.....	20
5.	SYSTEM DESIGN.....	21
5.1	Block Diagram.....	21
5.2	Algorithm.....	21
5.3	Class Diagram.....	22
6.	IMPLEMENTATION.....	23
6.1.	Experimental Results.....	23
6.1.1	Output with scaled Data.....	23
6.1.2.	Changing units in LSTM.....	24
6.1.3	Evaluating Model.....	26
6.2.	Implementation.....	26

6.2.1	Home screen.	27
6.2.2.	Signup/Registration Page.	27
6.2.3.	Stock selection Page.	28
6.2.4.	Next 30 days prediction.	28
6.2.5.	News section.	29
5	CONCLUSIONS.	30
5.1	Conclusion.	30
5.2	Limitation of Project.	30
5.3.	Future Scope.	30
	Bibliography	31

List of Figures

5.1	Block Diagram.	21
5.2	Algorithm.	21
5.3	Class Diagram.	22
6.1.1	Output with scaled Data.	23
6.1.2.	Changing units in LSTM.	24
6.1.3	Evaluating Model.	26
6.2.1	Home screen.	27
6.2.2	Signup/Registration Page.	27
6.2.3	Stock selection Page.	28
6.2.4	Next 30 days prediction.	28
6.2.5	News section.	29

List of Tables

Table no. 2.1 Comparison Table.	14
---	----

Abbreviation

ARIMA : Autoregressive integrated moving average

LSTM : Long Short Term Memory Networks

Chapter 1

INTRODUCTION

1.1 Introduction

The stock market is basically an aggregation of various buyers and sellers of stock. A stock (also known as shares more commonly) in general represents ownership claims on business by a particular individual or a group of people. The attempt to determine the future value of the stock market is known as a stock market prediction. The prediction is expected to be robust, accurate and efficient.

The system must work according to the real-life scenarios and should be well suited to real-world settings. The system is also expected to take into account all the variables that might affect the stock's value and performance. There are various methods and ways of implementing the prediction system like Fundamental Analysis, Technical Analysis, Machine Learning, Market Mimicry, and Time series aspect structuring.

Predicting how the stock market will perform is one of the most difficult things to do. There are so many factors involved in the prediction – physical factors vs. psychological, rational and irrational behaviour, etc. All these aspects combine to make share prices volatile and very difficult to predict with a high degree of accuracy.

In this project, we will work with historical data about the stock prices of a publicly listed company. We will implement a mix of machine learning algorithms to predict the future stock price of this company, starting with simple algorithms like averaging and linear regression, and then move on to advanced techniques like Auto ARIMA and LSTM.

1.2 Objective

1. To implement the LSTM model which will be trained on available historical data.
2. To forecast the next 30 days' price of stock of a particular NIFTY 50 listed company in which the user is interested.
3. To integrate this model in order to develop a simple system through which the user can interact and perform predictions as per his requirements i.e company name and number of predictions.
4. To provide insights of data in graphical format such as bar charts, line charts, comparison between predicted vs actual, accuracy of prediction etc.

1.3 Purpose, Scope & Applicability

1.3.1 Purpose

Machine learning algorithms allow computers to use available data to predict future outcomes and trends. For stocks it uses past performance data to forecast future prices of the stocks at different periods of time.

1.3.2 Scope

1. The project is focused on applications of model(LSTM) for stock price prediction.
2. In this project we are considering only NIFTY 50 listed companies as their historical data and real time data are easy to fetch.
3. To predict stock price users will have the option of uploading a dataset of that company .
4. Users will be able to make predictions for the next 30 days.

1.3.3 Applicability

Different organizations may consider it as an investment portal. Anyone who wants to buy stocks can take a look at this system. This helps us properly invest money in stocks with lesser chances of losing it.

Chapter 2

LITERATURE SURVEY

2.1 Literature Survey

1. Stock Market Prediction Using Artificial Neural Networks:

(Authors: Birgul Egeli, Meltem Ozturan, Bertan Badur. 2003)

In this paper it is described that the Prediction of stock market returns is an important issue in finance. Nowadays artificial neural networks (ANNs) have been popularly applied to finance problems such as stock exchange index prediction, bankruptcy prediction and corporate bond classification. An ANN model is a computer model whose architecture essentially mimics the learning capability of the human brain. The processing elements of an ANN resemble the biological structure of neurons and the internal operation of a human brain. Many simple interconnected linear or nonlinear computational elements are operating in parallel processing at multiple layers. In some applications it has been specified that ANNs have limitations for learning the data patterns. They may perform inconsistently and unpredictable because of the complex financial data used. Sometimes data is so voluminous that learning patterns may not work. Continuous and large volume of data needs to be checked for redundancy and the data size should be decreased for the algorithm to work in a shorter time and give more generalized solutions.

Artificial neural networks have been used in stock market prediction during the last decade. One of the first projects was by Kimoto and friends who had used ANN for the prediction of Tokyo stock exchange index. Mizuno and friends applied ANN

again to the Tokyo stock exchange to predict buying and selling signals with an overall prediction rate of 63%. Sexton and friends concluded in 1998 that use of momentum and start of learning at random points may solve the problems that may occur in the training process. Phua and friends applied a neural network with genetic algorithms to the stock exchange market of Singapore and predicted the market direction with an accuracy of 81%. In Turkey ANNs are mostly used in predicting financial failures . There has been no specific research for prediction of Turkish stock market values. The aim of this paper is to use ANNs to forecast Istanbul Stock Exchange (ISE) market index values.

2. Stock Market Forecasting Using Machine Learning Algorithms:

(Authors: Shunrong Shen, Haomiao Jiang, Tongda Zhang. 2007)

In this paper, we propose the use of global stock data in association with data of other financial products as the input features to machine learning algorithms such as SVM. In particular, we are interested in the correlation between the closing prices of the markets that stop trading right before or at the beginning of US markets. As the connections between worldwide economies are tightened by globalization, external perturbations to the financial markets are no longer domestic. It is our belief that data of oversea stock and other financial markets, especially those having strong temporal correlation with the upcoming US trading day, should be useful to machine learning based predictors, and our speculation is verified by numerical results.

The rest of the report is organized as follows. Section II presents our algorithm in detail, including the fundamental principle of our algorithm, data collection and feature selection. Numerical results are shown in Section III followed by analysis

and discussions. In Section IV, we established a simple trading model to demonstrate the capability of the proposed algorithm in increasing profit in NASDAQ. Section V summarizes the whole report.

3. Machine Learning Techniques for Stock Prediction:

(Author: Vatsal H. Shah. 2007)

In this paper, we discuss the Machine Learning techniques which have been applied for stock trading to predict the rise and fall of stock prices before the actual event of an increase or decrease in the stock price occurs. In particular the paper discusses the application of Support Vector Machines, Linear Regression, Prediction using Decision Stumps, Expert Weighting and Online Learning in detail along with the benefits and pitfalls of each method. The paper introduces the parameters and variables that can be used in order to recognize the patterns in stock prices which can be helpful in the future prediction of stocks and how Boosting can be combined with other learning algorithms to improve the accuracy of such prediction systems.

The main goal of the project was to study and apply as many Machine Learning Algorithms as possible on a dataset involving a particular domain, namely the Stock Market, as opposed to coming up with a newer (and/or better) algorithm that is more efficient in predicting the price of a stock.

4. Stock Market Value Prediction Using Neural Networks:

(Authors: Mahdi Pakdaman Naeini, Hamidreza Taremian, Homa Baradaran Hashemi. 2010)

This paper talks about the application of neural networks in prediction problems is very promising due to some of their special characteristics.

First, traditional methods such as linear regression and logistic regression are model based while Neural Networks are self-adjusting methods based on training data, so they have the ability to solve the problem with a little knowledge about its model and without constraining the prediction model by adding any extra assumptions. Besides, neural networks can find the relationship between the input and output of the system even if this relationship might be very complicated because they are general function approximators. Consequently, neural networks are well applied to the problems in which extracting the relationships among data is really difficult but on the other hand there exists a large enough training data sets. It should be mentioned that, although sometimes the rules or patterns that we are looking for might not be easily found or the data could be corrupted due to the process or measurement noise of the system, it is still believed that the inductive learning or data driven methods are the best way to deal with real world prediction problems.

Second, Neural Networks have generalization ability meaning that after training they can recognize the new patterns even if they haven't been in training set. Since in most of the pattern recognition problems predicting future events (unseen data) is based on previous data (training set), the application of neural networks would be very beneficial.

Third, neural networks have been claimed to be general function approximators. It is proved that an MLP neural network can approximate any complex continuous

function that enables us to learn any complicated relationship between the input and the output of the system.

5. Stock Market Prediction Using Hidden Markov Models:

(Authors: A Gupta, B Dhingra. 2012)

This paper briefly describes the Hidden Markov Model (HMM) approach was applied to this problem. The reason for using this approach is fairly intuitive. HMM's have been successful in analyzing and predicting time depending phenomena, or time series. They have been used extensively in the past in speech recognition, ECG analysis etc. The stock market prediction problem is similar in its inherent relation with time. Hidden Markov Models are based on a set of unobserved underlying states amongst which transitions can occur and each state is associated with a set of possible observations. The stock market can also be seen in a similar manner. The underlying states, which determine the behavior of the stock value, are usually invisible to the investor. The transitions between these underlying states are based on company policy, decisions and economic conditions etc. The visible effect which reflects these is the value of the stock. Clearly, the HMM conforms well to this real life scenario.

The choice of attributes, or feature selection is significant in this approach. In the past various attempts have been made using the volume of trade, the momentum of the stock, correlation with the market, the volatility of the stock etc. In our model we use the daily fractional change in the stock value, and the fractional deviation of intra-day high and low. The fractional change is necessary in order to make the

required prediction. Measuring the fractional deviation of both the intra-day high and low value is a good measure as it gives the direction of the volatility as well.

We use four different stocks for evaluating the approach - TATA Steel, Apple Inc., IBM Corporation and Dell Inc. A separate HMM is trained for each stock. The one constraint that the training set needs to have is suitable variability in the data. This is taken care of by taking appropriately large periods of time in which the stock value changes steadily yet significantly.

6. Deep Learning for Event-Driven Stock Prediction:

(Authors: Xiao Ding, Yue Zhang , Ting Liu , Junwen Duan. 2015)

In this paper it has been shown that the financial market is “informationally efficient” [Fama, 1965] — stock prices reflect all known information, and the price movement is in response to news or events. As web information grows, recent work has applied Natural Language Processing (NLP) techniques to explore financial news for predicting market volatility.

Pioneering work mainly uses simple features from news documents, such as bags-of-words, noun phrases, and named entities [Kogan et al., 2009; Schumaker and Chen, 2009]. Although useful, these features do not capture structured relations, which limits their potentials. For example, representing the event “Microsoft sues Barnes & Noble.” using term-level features {“Microsoft”, “sues”, “Barnes”, “Noble”} alone, it can be difficult to accurately predict the price movements of Microsoft Inc. and Barnes & Noble Inc., respectively, as the unstructured terms cannot differentiate the accuser (“Microsoft”) and defendant (“Barnes & Noble”).

Recent advances in computing power and NLP technology enables more accurate models of events with structures. Using open information extraction (Open IE) to obtain structured events representations, we find that the actor and object of events can be better captured [Ding et al., 2014]. For example, a structured representation of the event above can be (Actor = Microsoft, Action = sues, Object = Barnes & Noble). They report improvements on stock market prediction using their structured representation instead of words as features.

One disadvantage of structured representations of events is that they lead to increased sparsity, which potentially limits the predictive power. We propose to address this issue by representing structured events using event embeddings, which are dense vectors. Embeddings are trained such that similar events, such as (Actor = Nvidia fourth quarter results, Action = miss, Object = views) and (Actor = Delta profit, Action = didn't reach, Object = estimates), have similar vectors, even if they do not share common words. In theory, embeddings are appropriate for achieving good results with a density estimator (e.g. convolutional neural network), which can misbehave in high dimensions [Bengio et al., 2005]. We train event embeddings using a novel neural tensor network (NTN), which can learn the semantic compositionality over event arguments by combining them multiplicatively instead of only implicitly, as with standard neural networks.

For the predictive model, we propose to use deep learning [Bengio, 2009] to capture the influence of news events over a history that is longer than a day. Research shows diminishing effects of reported events on stock market volatility. For example, Xie et al. [2013], Tetlock et al. [2008] and Ding et al. [2014] show that the performance of daily prediction is better than weekly and monthly prediction. As shown in Figure 1, the influences of three actual events for Google Inc. in the year 2012 was the highest on the second day, but gradually weakened over time. Despite the relatively

weaker effects of long-term events, the volatility of stock markets is still affected by them. However, little previous work quantitatively models combined short-term and long-term effects of events. To fill in this gap, we treat history news as daily event sequences, using a convolutional neural network (CNN) to perform semantic composition over the input event sequence, and a pooling layer to extract the most representative global features. Then a feedforward neural network is used to associate the global features with stock trends through a shared hidden layer and a output layer.

Experiments on large-scale financial news datasets from Reuters and Bloomberg show that event embeddings can effectively address the problem of event sparsity. In addition, the CNN model gives significant improvement by using longer-term event history. The accuracies of both S&P 500 index prediction and individual stock prediction by our approach outperform state-of-the-art baseline methods by nearly 6%. Market simulation shows that our model is more capable of making profits compared to previous methods. To our knowledge, we are the first to use a deep learning model for event-driven stock market prediction, which gives the best reported results in the literature.

7. Stock Price Prediction using LSTM, RNN, CNN-Sliding Window Model:

(Authors: Sreelekshmy Selvin, Vinayakumar R, Gopalakrishnan E.A, Vijay Krishna Menon, Soman K.P. 2017)

This paper talks about Forecasting can be defined as the prediction of some future event or events by analyzing the historical data. It spans many areas including business and industry, economics, environmental science and finance. Forecasting problems can be classified as

- Short term forecasting (prediction for few seconds ,minutes, days, weeks or months)
- Medium term forecasting (prediction for 1 to 2 years)
- Long term forecasting (prediction beyond 2years)

Many of the forecasting problems involve the analysis of time . A time series data can be defined as a chronological sequence of observations for a selected variable. In our case the variable is stock price.It can either be univariate or multivariate. Univariate data includes information about only one particular stock whereas multivariate data includes stock prices of more than one company for various instances of time. Analysis of time series data helps in identifying patterns, trends and periods or cycles existing in the data. In the case of stock market, an early knowledge of the bullish or bearish mode helps in investing money wisely.Also the analysis of patterns helps in identifying the best performing companies for a specified period. This makes time series analysis and forecasting an important area of research.

The existing methods for stock price forecasting can be classified as follows

- Fundamental Analysis
- Technical Analysis
- Time Series Forecasting

Fundamental analysis is a type of investment analysis where the share value of a company is estimated by analyzing its sales, earnings, profits and other economic factors. This method is most suited for long term forecasting. Technical analysis uses the historical price of stocks for identifying the future price. Moving average is a commonly used algorithm for technical analysis. It can be considered as the unweighted mean of past n data points. This method is suitable for short term

predictions. The third method is the analysis of time series data. It involves basically two classes of algorithms, they are

- Linear Models
- Non Linear Models

The different linear models are AR, ARMA, ARIMA and its variations. These models use some predefined equations to fit a mathematical model to a univariate time series. The main disadvantage of these models is that, they do not account for the latent dynamics existing in the data. Since they consider only univariate time series, the inter dependencies among the various stocks are not identified by these models. Also the model identified for one series will not fit for the other. Due to these reasons, it is not possible to identify the patterns or dynamics present in the data as a whole

Non-linear models involve methods like ARCH, GARCH, TAR, Deep learning algorithms. In an analysis on the interdependence between stock price and stock volume for 29 selected companies listed in NIFTY 50 has been done. The proposed work focuses on the application of deep learning algorithms for stock price prediction. Deep neural networks can be considered as nonlinear function approximators which are capable of mapping nonlinear functions. Based on the type of application, various types of deep neural network architectures are used. These include multilayer perceptrons (MLP), Recursive Neural Networks (RNN), Long Short Term Memory (LSTM), CNN (Convolutional Neural Network) etc [9]. They have been applied in various areas like image processing, natural language processing, time series analysis etc.

2.2 Paper Comparison

Paper Title	Author	Year	Description
Stock Market Prediction Using Artificial Neural Networks	Birgul Egeli, Meltem Ozturan, Bertan Badur	2003	In this paper it is describe about ANN Methodology and its accuracy
Stock Market Forecasting Using Machine Learning Algorithms	Shunrong Shen, Haomiao Jiang, Tongda Zhang	2007	In this paper, we propose the use of global stock data in associate with data of other financial products as the input features to machine learning algorithms such as SVM.
Machine Learning Techniques for Stock Prediction	Vatsal H. Shah	2007	In this paper, we discuss the Machine Learning techniques which have been applied for stock trading to predict the rise and fall of stock prices before the actual event of an increase or decrease in the stock price occurs.
Stock Market Value Prediction Using Neural Networks	Mahdi Pakdaman Naeini, Hamidreza Taremian, Homa Baradaran Hashemi	2010	This paper talks about the application of neural networks in prediction problems is very promising due to some of their special characteristics.

Stock Market Prediction Using Hidden Markov Models	A Gupta, B Dhingra	2012	This paper briefly describes the Hidden Markov Model (HMM) approach was applied to this problem. The reason for using this approach is fairly intuitive.
Deep Learning for Event-Driven Stock Prediction	Xiao Ding, Yue Zhang , Ting Liu , Junwen Duan	2015	This paper describes Deep learning Methodology for Stock Prediction with example of stock of Google.Inc
Stock Price Prediction using LSTM, RNN, CNN-Sliding Window Model:	Sreelekshmy Selvin, Vinayakumar R, Gopalakrishnan E.A, Vijay Krishna Menon, Soman K.P	2017	This paper talks about Forecasting can be defined as the prediction of some future event or events by analyzing the historical data. It spans many areas including business and industry, economics, environmental science and finance.

Table No. 2.1 Comparison Table

Chapter 3

SURVEY OF METHODOLOGY

3.1 Deep learning:

Deep learning (also known as deep structured learning) is part of a broader family of machine learning methods based on artificial neural networks with representation learning. Learning can be supervised,

Deep learning architectures such as deep neural networks, deep belief networks, recurrent neural networks and convolutional neural networks have been applied to fields including computer vision, speech recognition, natural language processing, audio recognition, social network filtering, machine translation, bioinformatics, drug design, medical image analysis, material inspection and board game programs, where they have produced results comparable to and in some cases surpassing human expert performance.

3.2 Artificial Neural Network:

An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. ANNs, like people, learn by example. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons. This is true of ANNs as well. Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques.

3.3 Time series analysis:

A time series is a series of data points indexed (or listed or graphed) in time order. Time series analysis comprises methods for analyzing time series data in order to extract meaningful statistics and other characteristics of the data. Time series forecasting is the use of a model to predict future values based on previously observed values. While regression analysis is often employed in such a way as to test theories that the current values of one or more independent time series affect the current value of another time series, this type of analysis of time series is not called "time series analysis", which focuses on comparing values of a single time series or multiple dependent time series at different points in time. Interrupted time series analysis is the analysis of interventions on a single time series.

3.4 Recurrent Neural Network:

RNN is a neural network designed for analyzing streams of data by means of hidden units. In some of the applications like text processing, speech recognition and DNA sequences, the output depends on the previous computations. Since RNNs deal with sequential data, they are well suited for the health informatics domain where enormous amounts of sequential data are available to process.

In general, RNNs are provided with the input samples which contain more interdependencies. Also they have a significant representation for keeping the information about the past time steps. The output produced at time t_1 affects the parameter available at time $t+1$. In this manner, RNNs keep two kinds of input such as the present one and the past recent one to produce the output for the new data.

3.5 Long Short Term Memory:

Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. LSTM networks are well-suited to

classifying, processing and making predictions based on time series data, since there can be lags of unknown duration between important events in a time series.

LSTMs have an edge over conventional feed-forward neural networks and RNN in many ways. This is because of their property of selectively remembering patterns for long durations of time. The purpose of this article is to explain LSTM and enable you to use it in real life problems.

3.6 Python

Python is an interpreter, high-level, general-purpose programming Language. Created by Guido van Rossum and first released in 1991, Python's design philosophy emphasizes code readability with its notable Use of significant whitespace. Its language constructs and object-Oriented approach aim to help programmers write clear, logical code for Small and large-scale projects. Python is dynamically typed and Garbage-collected. It supports multiple programming paradigms, Including procedural, object-oriented, and functional programming. Python is often described as a "batteries included" language due to its Comprehensive standard library. Python was conceived in the late 1980s As a successor to the ABC language. Python 2.0, released in 2000, Introduced features like list comprehensions and a garbage collection System capable of collecting reference cycles. Python 3.0, released in 2008, was a major revision of the language that is not completely Backward- compatible, and much Python 2 code does not run unmodified on Python 3.

The Python 2 language was officially discontinued in 2020 (first planned for 2015), and "Python 2.7.18 is the last Python 2.7 release and therefore the last Python 2 release." No more security patches or other improvements will be released for it. With Python 2's end-of-life, only Python 3.5.x and later are supported. Python interpreters are available for many operating systems. A global community of programmers develops and maintains CPython, an open source reference implementation. A non-profit organization, the Python Software Foundation, manages and directs resources for Python and CPython development. Due to concern about the amount of code written for Python 2, support for Python 2.7 (the last release in the 2.x series) was extended to 2020.

Language developer Guido van Rossum shouldered sole responsibility for the project until July 2018 but now shares his leadership as a member of a five-person steering council. Python was conceived in the late 1980s by Guido van Rossum at Centrum Wiskunde & Informatica (CWI) in the Netherlands as a successor to the ABC language (itself inspired by SETL), capable of exception handling and interfacing with the Amoeba operating system. Its implementation began in December 1989. Van Rossum shouldered sole responsibility for the project, as the lead developer, until 12 July 2018, when he announced his "permanent vacation" from his responsibilities as Python's Benevolent Dictator for Life, a title the Python community bestowed upon him to reflect his long-term commitment as the project's chief decision-maker. He now shares his leadership as a member of a five-person steering council. In January 2019, active Python core developers elected Brett Cannon, Nick Coghlan, Barry Warsaw, Carol Willing and Van Rossum to a five-member "Steering Council" to lead the project.

Python 2.0 was released on 16 October 2000 with many major new features, including a cycle-detecting garbage collector and support for Unicode. Python 3.0 was released on 3 December 2008. It was a major revision of the language that is not completely backward-compatible. Many of its major features were back ported to Python 2.6.x and 2.7.x version series. Releases of Python 3 include the 2to3 utility, which automates (at least partially) the translation of Python 2 code to Python 3. Python 2.7's end-of-life date was initially set at 2015 then postponed to 2020 out of concern that a large body of existing code could not easily be forward-ported to Python 3.

3.7 Django

Django is a high-level Python Web framework that encourages rapid development and clean, pragmatic design. Built by experienced developers, it takes care of much of the hassle of Web development, so you can focus on writing your app without needing to reinvent the wheel. It's free and open source. Django was designed to help developers take applications from concept to completion as quickly as possible.

Django takes security seriously and helps developers avoid many common security mistakes. Some of the busiest sites on the Web leverage Django's ability to quickly and flexibly scale.

Chapter 4

REQUIREMENTS AND ANALYSIS

4.1 Problem Definition

A stock market, equity market or share market is the aggregation of buyers and sellers of stocks, which represent ownership claims on businesses. Our goal is to predict the price of a certain company's stock or value using time series analysis i.e deep learning. Time series data often arise when monitoring industrial processes or tracking corporate business metrics. In this project we are going to implement Long Short Term Memory network(LSTM) which is type of Recurrent Neural Network (RNN)

4.2 Requirements Specification

For implementation in software we will require the following software and hardware specifications:

- **Software Specification**

For implementation we will require applications such as Jupiter Notebook ,Python 3.6 or better , Spyder IDE,MS-Excel etc. That can be used to build an environment based on a neural network and to train machines based on it.

- **Hardware Specification**

To implement the project we will require a computer with specification such as multicore CPU, graphics card, hard disk upto 500GB, upto 8GB RAM. Input devices such as keyboard, optical mouse. Output device: LCD monitor.

Chapter 5

SYSTEM DESIGN

5.1 Block Diagram

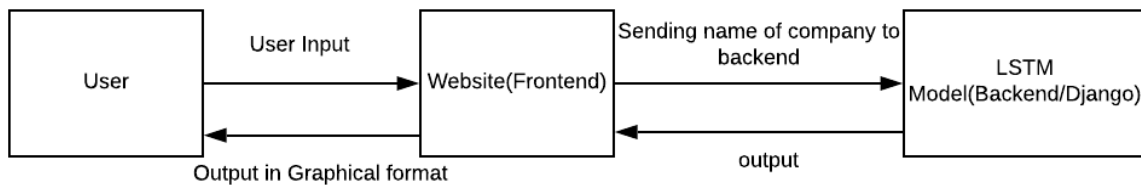


Fig. 5.1 Block Diagram

5.2 Algorithm

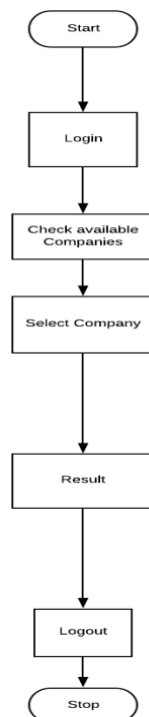


Fig 5.2 Algorithm

5.3 Class Diagram

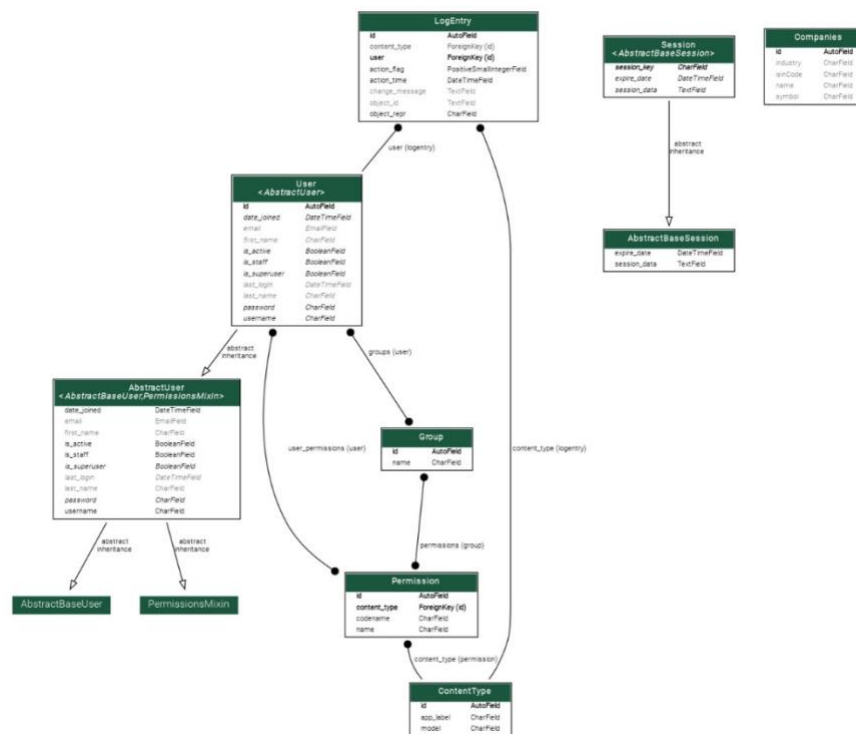


Fig 5.3 Class Diagram

Chapter 6

IMPLEMENTATION

6.1 Experimental Results

6.1.1 Output with Scaled Data

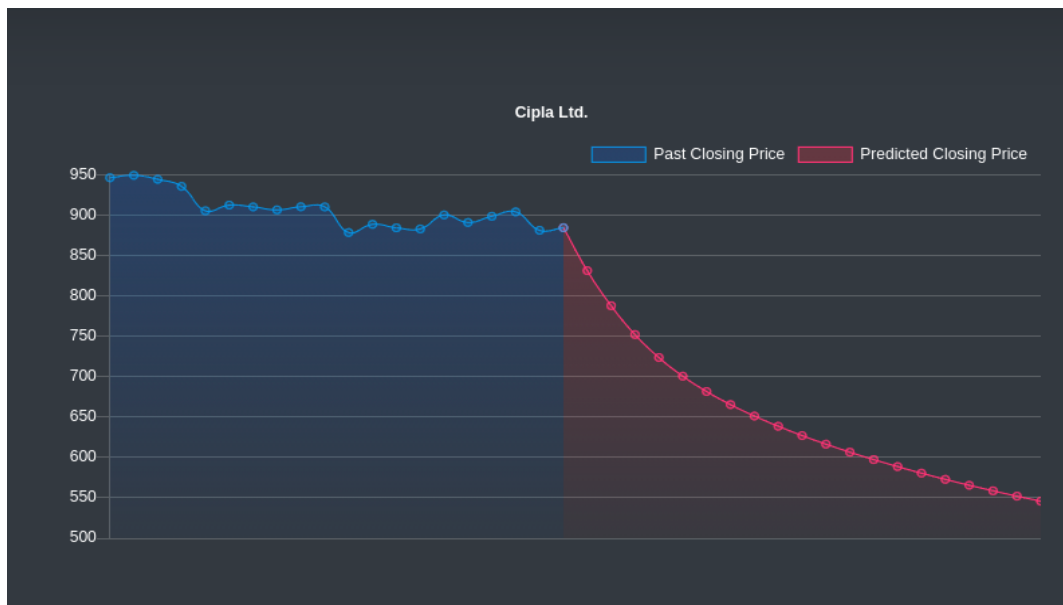


Fig. 6.1.1.1 Output with scaled Data 1

The above image shows the output of prediction which is decreasing continuously which was not expected as this model is not able to capture trends. After reading several blogs we got to know scaling in LSTM is not much useful. So we decided not to scaled data. After undoing scaling we got output which shows trend in prediction of future prices



Fig 6.1.1.2 Output with scaled Data 2

we get the above output without scaling input data .as we can see model is now able to capture trend

6.1.2. Changing units in LSTM



Fig 6.1.2.1 Changing units in LSTM 1

When we plot graph of actual close price against predicted close price of training set we got the above graph when we use 50 units in LSTM model for training the green

line is actual stock price of wipro and red line is predicted value of training set of wipro. but after changing units to 64 from 50 we got following graph which is more accurate on training set

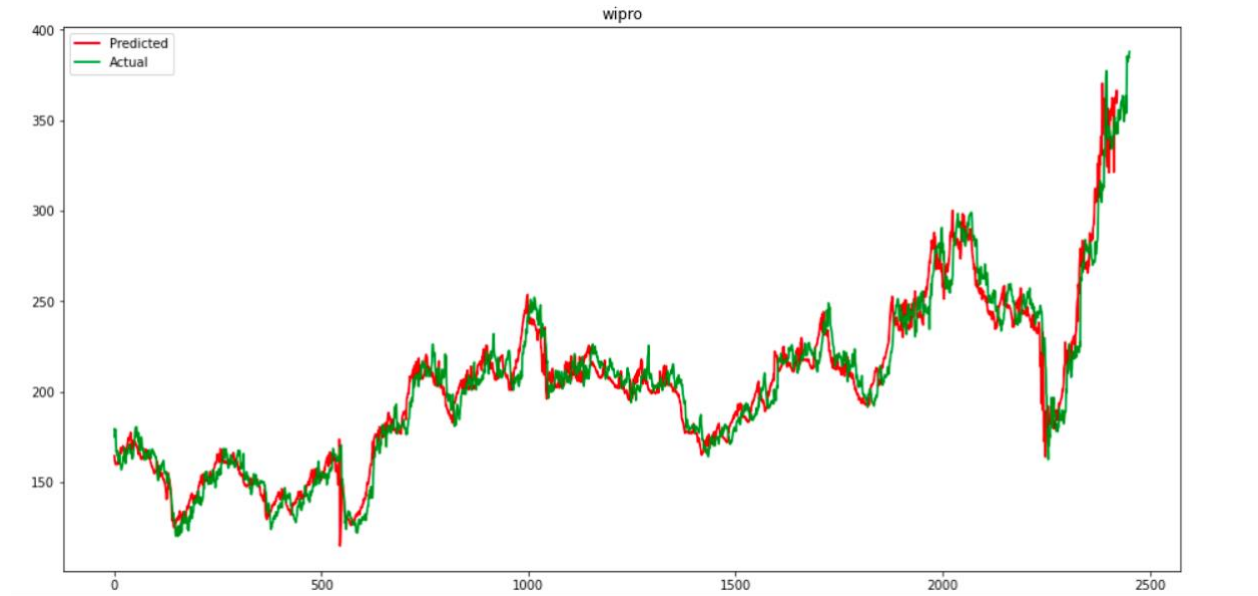


Fig. No. 6.1.2.2 Changing units in LSTM 2

6.1.3 Evaluating Model

```
[30]
compare_df = pd.DataFrame({'Actual':df2['Close'].values,
                           'Predicted':lst_output,
                           'Diff':df2['Close'].values-lst_output
                           })
compare_df
```

	Actual	Predicted	Diff
0	388.100006	371.843872	16.256134
1	396.399994	373.580780	22.819214
2	406.299988	379.351135	26.948853
3	406.399994	378.447083	27.952911
4	406.750000	379.498901	27.251099
5	430.200012	383.427185	46.772827
6	446.799988	384.531677	62.268311
7	457.700012	387.585114	70.114899
8	459.000000	388.239197	70.760803
9	454.350006	388.512024	65.837982
10	438.549988	382.440308	56.109680

Fig. 6.1.3.1 Evaluating Model

we predicted price of stock of Wipro for month of January 2021 and compare with actual data and evaluated using RMSE(Root mean squared error) which is 49.07.i.e mean difference between predicted and real price

6.2 Implementation

6.2.1 Home Page

- It is the main screen when the user enters the site.
- First the user views about live Nifty 50 chart.
- Then user should login for the prediction.

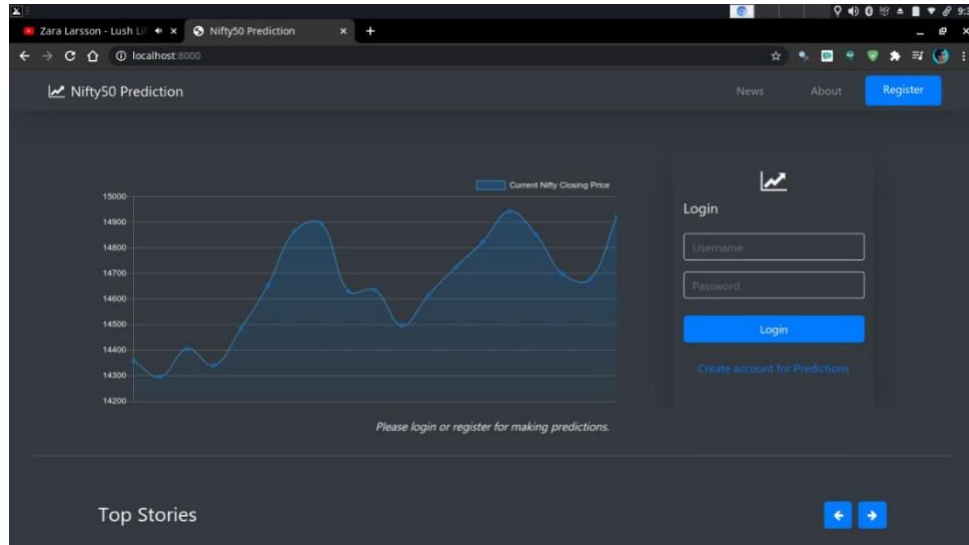


Fig. 6.2.1 Home Page

6.2.2 Signup/Registration Page

Fig. 6.2.2 Signup/Register Page

- In this registration module first the user has to registered to the site/app.
- The registration includes user id, email id, password, and re confirm the password.
- This is also called as a first-time registration form that is the user has to register for the first time and there after only need to login.

6.2.3 Stock selection page

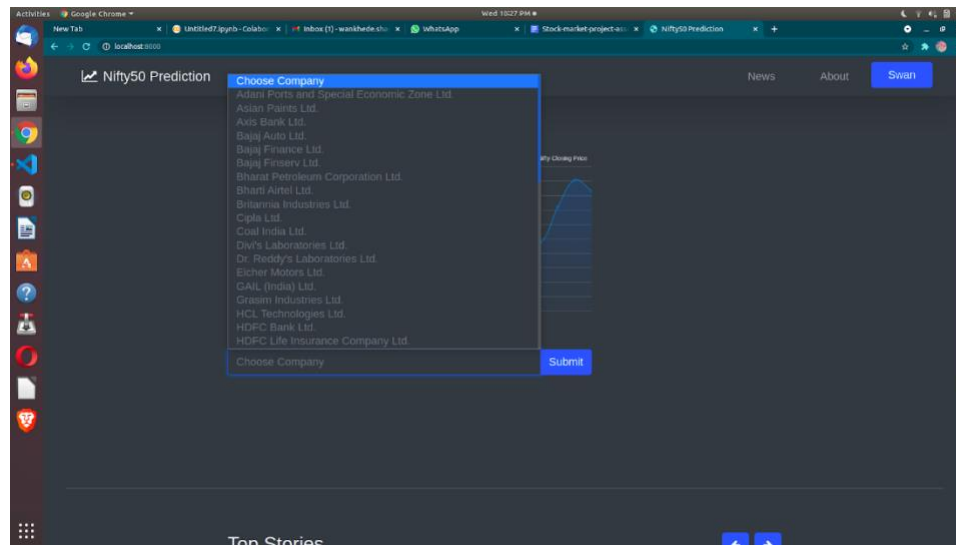


Fig 6.2.3 Stock selection page

- The user clicks on the scrollbar to select the stock and clicks on submit button.
- The display shows a predicted graph and table of next 30 days.

6.2.4 Next 30 day's prediction of Stock

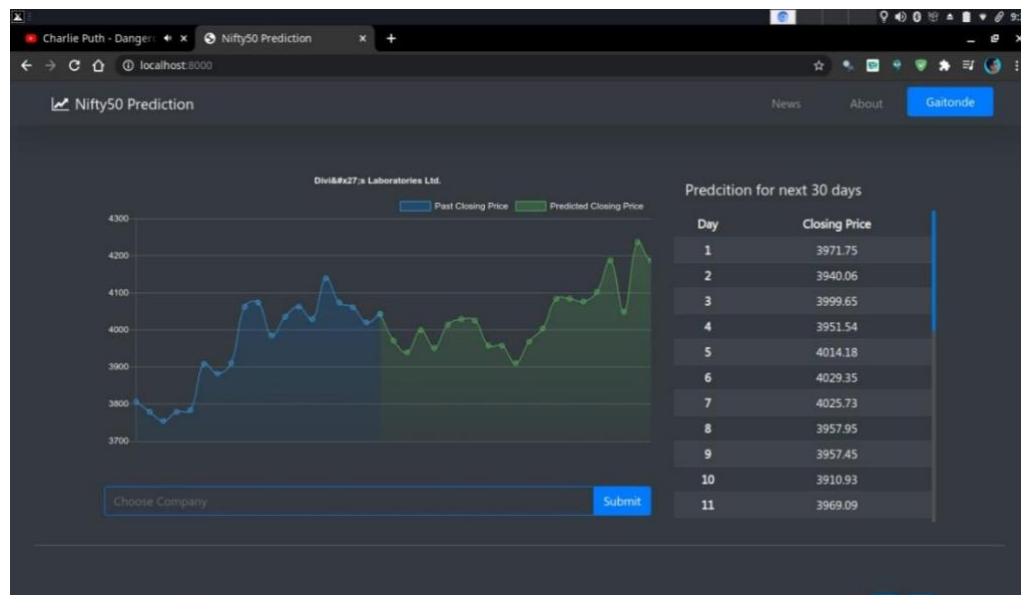


Fig. 6.2.4 Prediction of stock

- When the user clicks on the predict button the Prediction of next 30 days will be generated.
- On the screen user can watch the previous data and the predicted data in form of graph.
- On the right side we can watch the closing price of next 30 days.

6.2.5 News Section

- In news section user can read the latest news about the market when he or she clicks on the news.

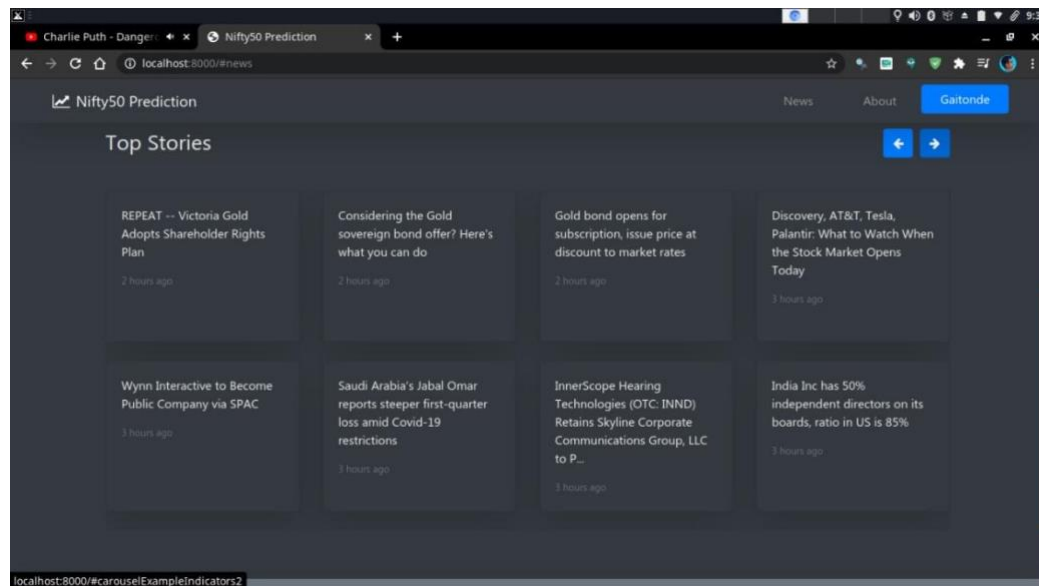


Fig. 6.2.5 News Section

Chapter 7

CONCLUSION

7.1 Conclusion

Conclusion After researching through various papers related to Time series analysis , stock market prediction. We have concluded that predicting the stock market is really a challenging task and also involves a lot of factors including natural factors, company's production or work. so it's impossible to predict accurately price of stock but A deep learning model can be developed that can predict the value of stocks based on previous values according to time or data which is just a mathematical model which can help us to see how market is going or we can take look of market direction i.e trends. Various experiments have been conducted using different methodologies, the best results are seen in the methods that are based on neural networks. and used a method with less error

7.2 Limitation of Project

- Since our project runs on different APIs these APIs are not much stable
- Project has only 50 company to choose
- Predicted and Actual price has significant error

7.3 Future scope

- Building more robust model for better accuracy
- More number of companies
- Better interactivity
- Custom data input

Bibliography

1. <http://cs229.stanford.edu/proj2012/ShenJiangZhang-StockMarketForecastingusingMachineLearningAlgorithms.pdf>
2. <https://bigquant.com/community/uploads/default/original/1X/5c6d3b9959a8556a533a58e0ac4568dfc63d6ff4.pdf>
3. http://people.cs.pitt.edu/~hashemi/papers/CISIM2010_HBHashemi.pdf
4. https://d1wqtxts1xzle7.cloudfront.net/52961747/Stock_market_prediction_using_hidden_markov_models.pdf?1493840215=&response-content-disposition=inline%3B+filename%3DStock_market_prediction_using_hidden_mar.pdf&Expires=1592668752&Signature=gKghtLJKcp7NqaJIygDZIBoZpa5QuPzlXJ92~AUVqFUVuft1KvDVoBbO5Qcu4Oyj0LfYMS7N-8Hqiswe~8UfOgda6sAozTPgRGw8DfMxjpvab~n71UGLvACTyucG~NDOAU4r9W8mNdNcy6k4C3jpm~BuuTcApP4hsGWWVAJxdtZDWcCl-8yZTD0~wJ6fSF9i-UoPWuFMMHlzw9-XH0Stsq46ZXSH3TJoMLXbGXUCga1VqhnoWrZW1K-jce06WQOIMNpst58qTyEpqAj5r5zQxCmkJ3Y4yKz8D4ziHmlMqf-u0BhvCm4yM0fbfY5FJqTnKuxLqLt0j6g3AZmYgGJkA_&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA
5. <https://www.aaai.org/ocs/index.php/IJCAI/IJCAI15/paper/viewFile/11031/10986>
6. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.134.1114&rep=rep1&type=pdf>
7. https://www.researchgate.net/profile/Vijay_Menon5/publication/325666831_NSE_Stock_Market_Prediction_Using_Deep-Learning_Models/links/5b287f67aca2727335b70828/NSE-Stock-Market-Prediction-Using-Deep-Learning-Models.pdf
8. <https://www.analyticsvidhya.com/>
9. https://en.wikipedia.org/wiki/Long_short-term_memory
10. <https://machinelearningmastery.com/what-is-deep-learning/>