Share Market Forecasting with LSTM Neural Network and Sentimental Trend prediction

by

Ismail Hossain Rony
15201048
Ahsan Ahmed Anik
13201014
Abdullah Al Asif
13301106
Sayeed Muhammad
19141022

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
Brac University
September 2019

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Ismail Hossain Rony 15201048	Ahsan Ahmed Anik 13201014
Abdullah Al Asif 13301106	Sayeed Muhammad 19141022

Approval

The thesis titled "Share Market Forecasting with LSTM Neural Network and Sentimental Trend prediction" submitted by

1. Ismail Hossain Rony(15201048)
2. Ahsan Ahmed Anik (13201014)
3. Abdullah Al Asif (13301106)
4. Sayeed Muhammad (19141022)

of Academic Year 2019 has been found as satisfactory and accepted as partial fulfillment of the requirement for the Degree of Bachelor of Science in Computer Science and Engineering

Examining Committee:

Supervisor:

____

Dr.Md.Iftekharul Mobin
Assistant Professor
Department of Computer Science And Engineering
BRAC University

Chairperson:

____

Dr.Mahbubul Alam Majumdar
Professor
Department of Computer Science and Engineering
Brac University

Abstract

Forecasting and predicting future trend is getting significant importance in stock market exponently as it is volatile in nature. Stock market is an extremely complicated , unstable and volatile place due to the fact that prediction is very difficult. Because of uncertainty and having scope of gaining financial profits, share market estimating and prediction has been a renowned matter in financial and academic studies. Advanced algorithms of machine learning is required as there is no persistently appropriate prediction tool. Many research works from various sector have been done to overcome this difficulties of predicting stock market. In machine learning sector a lot of research work already accomplished to predict share market. Many algorithms of Machine Learning have been utilized for this kind of prediction and the result was also satisfactory. In this thesis, we will extract all the real data from Dhaka Stock Exchange (DSE) using web scrapping and try to predict stock market price on a giving day, by approaching Long Short Term Memory(LSTM) Networks based on historical data mining method. The results of this paper show that Long Short Term Memory Networks can be applied for evaluation of historical stock pricing data and acquire valuable information by forecasting future trend with suitable financial indicators. Beside this, we will extract all the news opinions from the respective web pages (DSE, Lonkabangla financial port) and went through noise reduction, implementing algorithm and classifier to determine the sentiment polarity to come to a choice whether the stock price of a company are getting upward or downward trend. We are using naïve bayes classifier to examine the ratio of a sentence or phrase which can contain sentiment in from of positive, negative and neutral words. Using this model we can represent a status of some stock news.

Keywords:stock market, Long Short Term Memory Networks,Sentiment Analysis, prediction, forecasting, future trend

Acknowledgement

It is a matter of great pleasure to present this thesis report on "Share Market Forecasting with LSTM Neural Network and Sentimental Trend prediction ". We are grateful to our thesis supervisor Md.Iftekharul Mobin, who has been of invaluable assistance to our project with his advice and suggestions. We would like to thank our respected sir Mahbubul Alam Majumdar, Head of Department of Computer Science and Engineering, for willing to provide us with all the resource required to complete our thesis.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

Share market estimation is one of the important topic for researchers for last few decades but being highly volatile and complex in nature which make forecasting future trend a very difficult task[1]. Stock market also plays a crucial job in financial area of a nation. In spite of being a developing country share market has a great impact on the economy of Bangladesh. According to Dhaka Stock Exchange the market capital of listed companies of DSE is stood over than 40 billion in 2015. Stock market is directly involved with economic activities of a country. For this reason, appropriate prediction can both maximize the profit of investors and advise the government system to take necessary steps for possible economic downturns. Many technical and macro economical methods are utilized to estimate the nature of share market. All these procedures are related with different information's and patterns which are observed by human. As share market is dynamic, it is quite tough job to estimate the price. In our paper, we are approaching Long Short Term Memory Neural Networks which is a sort of recurrent network. LSTM has capability to solve the problem by distinguishing between present and past data using different weights for each memory to predict the forecast. Expect LSTM other recurrent neural networks are able to memorize short sequences But LSTM has capability to handle long sequences[2]. In our model we are using historic data based on Dhaka stock exchange (DSE) as source of information. After feeding the dataset the model will perform training, evaluating and will be attempting the future forecasting of a significant share price will climb up or fall down. Beside this stock price movement is dependent on other factors also and news sentiment is one of the important factor. As the growing rate of internet use and the popularity of news portals, sentiment analysis became one of the most popular concern in the field of research. By the help of the study of sentiment analysis, it can easily demonstrate the nature and trend of stock. It is possible to predict the stock market trend by using the Sentiment analysis .Sentiment analysis helps to decide the positive, negative or neutral views of news about stock market. Basically, sentiment analysis is not only the negative or positive polarity of sentences or words but also the syntactical analysis of those sentences. The main contribution of our work are giving a new forecasting model for share market using LSTM neural networks and validate the model using real data from Dhaka Stock Exchange and predict the train with the help of sentiment analysis. Our proposed model will help investors to invest in stock market based on the various factors.

Chapter 2

Related Work

Forecasting and prediction of stock price has become an noticeable field of research in the last quarter of the century. This is really challenging and exciting field of research. Being chaotic,complex and dynamic in nature,there has been a debate on the predictability of share market. E.F.Fama established a hypothesis that contains the present price of an stock reflects all historic information which are connected with it[3]. Another work by B.G.Malkiel mention his hypothesis that share prices are independent in nature and it hardly depends on its history[4]. These two hypothesis describe that accurate prediction of a stock price is quite difficult job. There are a lot of proposed models that have been used for share market forecasting. For time series learning Numerous customary techniques dependent on statistics implemented like linear regression[5], Moving Average (MA) and Auto-regression (AR). Box and Jenkins[6] proposed Auto-relapse Moving Average (ARMA) model that can manage the succession, some portion of which is Auto-relapse while another part is Moving Average. Auto-regression Integrated Moving Average which is known as ARIMA model is fit for managing non-stationary time seriesEngle[7] built up Auto-regressive Conditional Heteroskedasticity (ARCH) model, which can reproduce the variety of time series variable. From that point, Bolleralev[8] proposed Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, which is especially appropriate for the investigation and forecast of instability. These techniques are broadly utilized in the field of time arrangement forecasts and can accomplish a decent prescient impact when the time arrangement is Gauss dispersion. Some other customary strategies are additionally utilized for time series estimation. Elaal et al.[9] presented multivariate-factors fluffy time arrangement prediction model dependent on fluffy bunching to deal with certifiable multivariate prediction issues. In any case, there are as yet a couple of disadvantages with respect to the recently referenced models where they are not ready to take the flimsy and non-stationary components of the stock prediction into account[10]. Different application of Artificial Neural Network(ANN) is utilized for estimating the share value ANN[11] is one of the most exact strategies to estimate stock patterns. Up until this point, ANN has been generally utilized in share prediction. According to many researches. ANN beats factual models and discriminant methods[12]. ANN model are used by Yi, Jin, John and Shouyang[13] for predicting the financial volatility[14]. BP neural network to estimate the share value is used by Huarang and Yu[15]. Shen, Guo, Wu, and Wu[16] foresee stock files of Shanghai Stock Exchange with the model of outspread premise work neural system. Huarng and Yu[17] utilized back-proliferation neural system to foresee stock value.Some specialists consider stock value as time arrangement and utilize short-term memory model Recurrent Neural Network (RNN) to estimate time arrangement. RNN model is capable to handle short term memory. Hsieh, Hsiao, and Yeh applied RNN model and merged wavelet transformation for forecasting share price[18]. Depending on the existing above, these models exist some drawbacksMost importantly the conventional time arrangement models utilize former recorded stock information as the info factors. In any case, the cost of stock is influenced by countless components, for example, economic situations and natural impacts. The immediate utilization of complex recorded information in the conventional time arrangement models tends to decrease the estimating capacity and in this way the last outcomes. Along these lines, numerous specialists will in general diminish the multifaceted nature of the information, for example, head part investigation is proposed with the goal that the first information separate into more straightforward components or higher related factors to improve the capacity of the anticipating models. In addition these models just accept the previous collected data information as information factors, however not thinking about the effects of the condition of the stock on the prediction, for example, the enthusiastic inclination of financial specialists. Next to this, a few scientists do not consider stock value factors as a time sequences and think about them as discrete varieties. Different analysts considerably regard stock value varieties as a time arrangement, however they accept the memory of time arrangement is either short or fixed. However, as observed over, the long term dependence issues are not very much dealt with. As recurrent networks are not capable of handling long sequences data, S. Hochreiter and J. Schmidhuber, introduced LSTM[19].In our thesis work we proposed a system on LSTM to estimate the movement of stock price using input as different historic data. We use the dataset as wide range of technical indicators and apply different methods to get the accurate prediction rate. Our main target of this paper is giving a prediction system for stock market using LSTM combining with news Sentiment on stock market which gives a decision whether our predicting forecast on share market is accurate or over-fitted.

Chapter 3

Methodology and workflow

3.1 Methodology

Here we designed preferred model on LSTM networks for performing prediction of stock price movements for input number of stocks. Every day it feed data automatically by scrapping data from Dhaka Stock Exchange and then it attempted demo price for each stock and determined the next day price of the stock which was higher or lower than the previous day price of the stock and compared the demo stock price with the actual stock price. On the other hand in Sentiment analysis data collection was as same as LSTM (the scrapping method) but the data type was different. Here we extract all the news opinion by using the python script and made a excel sheet for the news sentiment. Here in this excel sheet the reaction of every news of the companies was included. After that we predicted the next day stock price was higher or lower than the previous stock price as like as LSTM trend prediction by analyzing the sentiment data. Finally, we justified the next day stock price prediction by these two methods.

3.1.1 Long Short Term Memory Approach

Data Source: Here our main target is to estimate the stock value and it will be done by stock's preceding cost and trends. It also emphasizes on data mining techniques with respect to the historical information of stock market. For getting historical information and quintessential data-sets, a trusted source is needed. For our primary supply of data we will be using Dhaka Stock Exchange website (DSE) and these data collected from https://www.dsebd.org/. DSE website categorizes into different features such as variety of share, opening value, closing value, best possible value, lowest value, extend or reduce rate of shares for each financial companies. It additionally gives the usual overall performance of Bangladesh Stock Exchange and performance of businesses of exceptional categories. The web site is updated on day by day groundwork and this previous recorded information is utilize for the estimation of future share values.
Data Prepossessing and feature exact: At first we normalized the data and integrate the data in data files.Then transformation process started and it divides the datasets into training and testing sets for evaluating process. Under transformation process datasets convert into clean and smooth datasets and then these historical information from these datasets categorize according to time series features such as

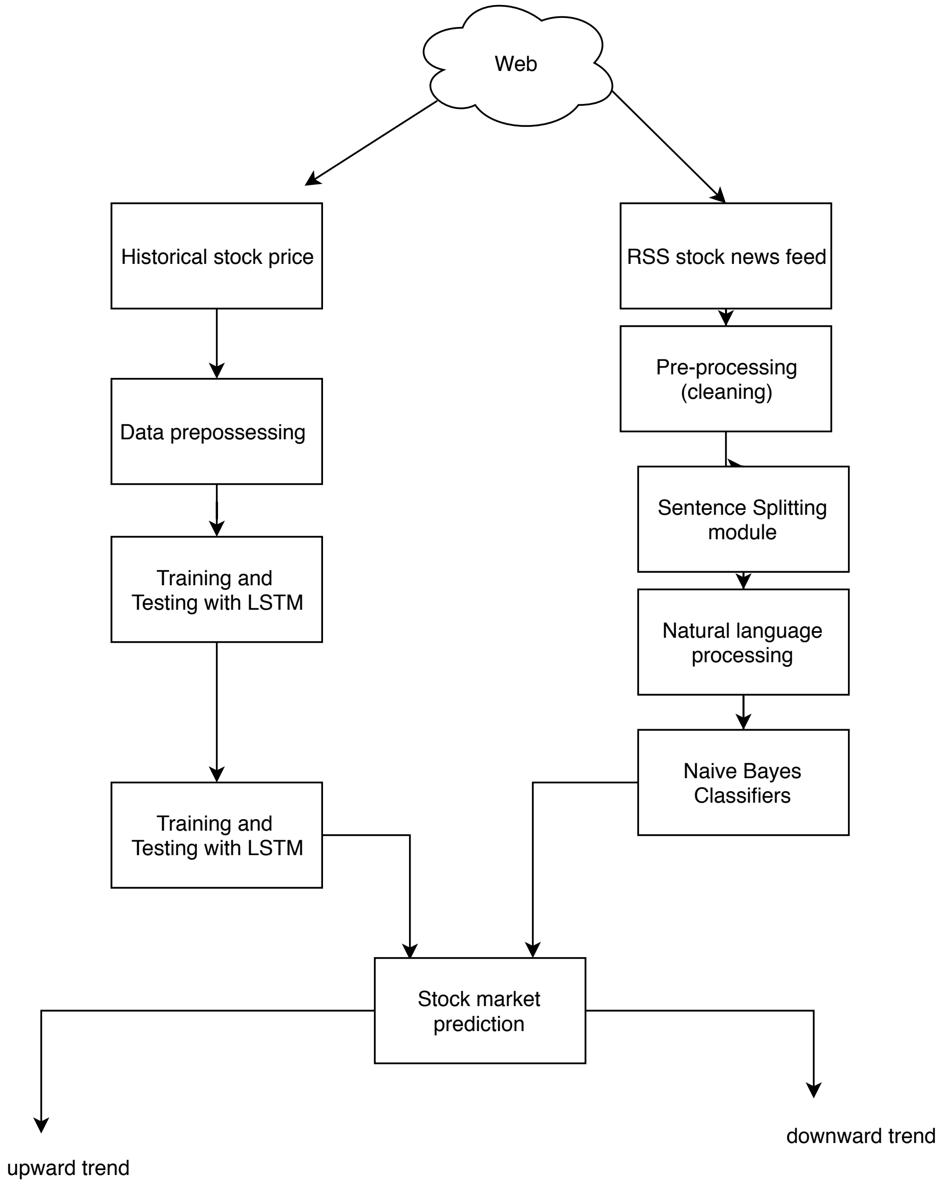

Figure 3.1: Activity diagram of methodology.

open, close, high, low and volume. For our thesis work, we gathered data from 2005 to 2018 for shares from Dhaka stock exchange (DSE).
Training Neural Network: All the information gathering form historical datasets then feed in to our proposed model and it evaluates all information according to random biases and weights and start the training process. Here this model is the collection of a constant input layer accompanied through two LSTM layers and dense layer with ReLU activation and then eventually a dense output layer with linear activation function. These training are carried out using ADAM optimizer. It operates with default parameters and learning rate for completing the training[20].

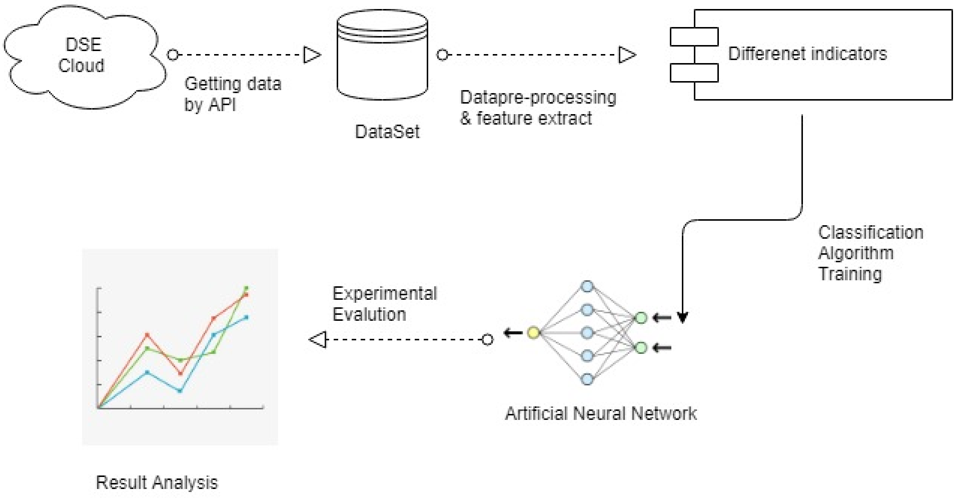

Figure 3.2: Methodolgy for LSTM.

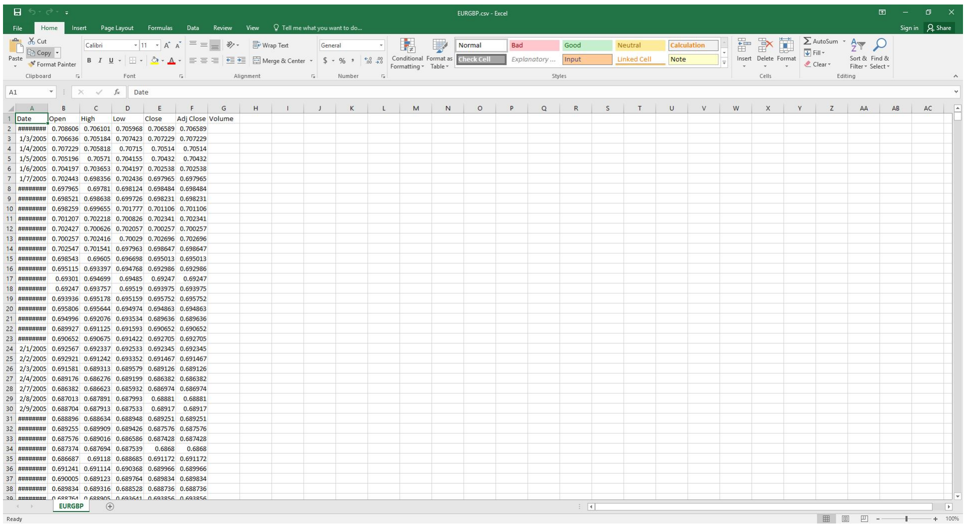

Figure 3.3: Data source for LSTM.

3.2 Long Short Term Memory Neural Network

Artificial Neural Network (ANN) is a scientific model which is utilized to replicate the data preparing capacity of the human neural system framework, and it has been broadly prominent in the example arrangement issue. As displaying in Fig3.2 it is dependent on a hierarchical structure that contains an input layer, one or more hidden layers and one output layer. These units are associated with one another, every neuron weighted entirety up the sources of info, applies an enactment capacity to the total and the yield is transmitted to the following layer. Multilayer Perceptron[21] are organized in layers, with the associations feeding onward from the layer to the following layer. Data is oriented into different layers such as input layer, hidden layer and output layer. The initial layer ended into yield layer passing through the hidden layer. System with these similarity known as forward pass which is actually passing from one step to another. The yield of a MLP is just identified with the present input and doesn't rely upon past information or future input, so MPL is reasonable for example characterization. A solitary MLP can execute

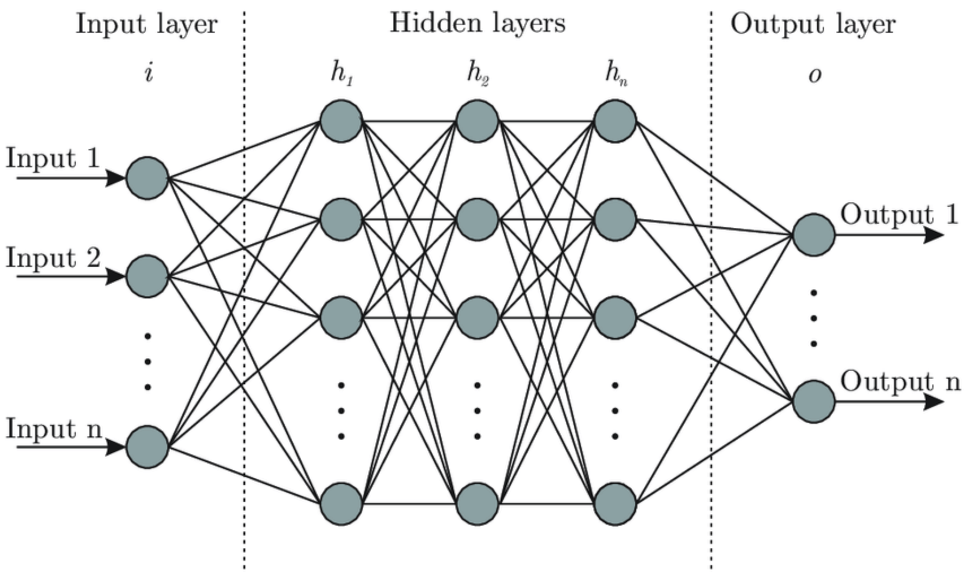

Figure 3.4: Artificial Neural Network.

well in many experimentation fields on basis of different weights. As a matter of fact Hornik[22] demonstrated that a MLP with an enormous number of units can surmised any persistent capacity to a discretionary exactness. As a result MLP is also known as universal function approximation. ANN has some varieties and

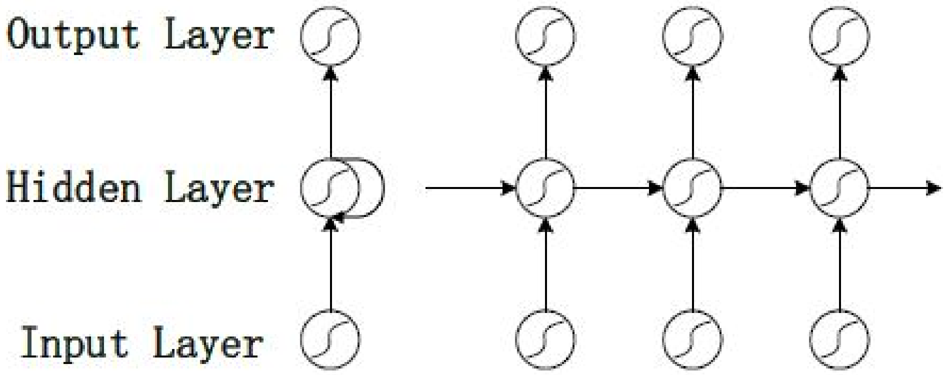

Figure 3.5: Recurrent Neural Network.

Recurrent Neural Networks(RNN) is one of the types of it. Distinct from the feed-forward neural networks instigated by MLP, RNN enables the association of the system to shape a cycle. The impact of RNN[23] on time arrangement learning is significant though the difference between RNN and MLP is very less . To be sure, comparable to the all-inclusive capacity guess, RNN with layer and adequate number of concealed units can inexact any quantifiable succession to-arrangement mapping to a subjective exactness[24]. For stock forecasting , RNN is a good learning model and like as Elman network[25] and Jordan network[26] different form of RNN have been implemented. The formation of the RNN is appeared in the left piece of Fig3.3 The repetitive associations of concealed layer help the system to "remember" the information condition of the past system. The right portion of Fig3.3 demonstrates an unfurled recurrent system, each concealed node consider to a state at a time step. The capacity to delineate the whole history of contributions to each yield is an important attribute of RNN . As the impacts decomposes after some time as the sources of info overwrite the initiations of the concealed layer, RNN is known as short-term memory model. Long Short Term Memory networks are a special

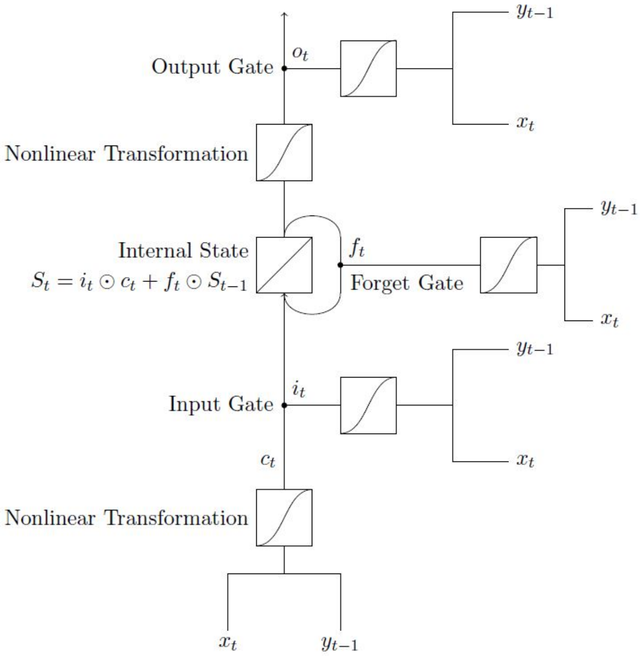

Figure 3.6: Long short term memory

form of Recurrent Neural Networks which is also recognized as LSTM Networks. Avoiding long term dependency is the main factor of LSTM networks. This network system designed in a specific way that can solve the issues of RNN by vanishing and exploding gradients and learn long term dependencies[27]. LSTM contained with hidden layers which is known as memory cells. These memory cells are adjusted and maintained by the combination of input , output and forget gates. These gates are defined with different works such as input gate indicates which data is appended to the cell, output gate determines which data from the cell state is right now utilized and forget gate characterizes which data is expelled from the cell entryway[28]. The capacity of the information entryway is to control the information signal that can change the condition of the memory cell or remove it. The capacity of the yield door is to enable the condition of memory cell to affect different units or intercept it. The forget gate is the self-association of the memory cell in order to allow the to cell recall or overlook past state. Because of the multiplicative entryways, LSTM can get to more setting data than RNN, which mitigate the disappearing angle issue.

3.2.1 Batch Size

Batch size represents the number training data we used in a network. In research we have to use large number of datasets for training the model. It is difficult to handle such huge number of datasets. For solution we divide the whole data into several parts, each part is known as a batch. Beside this, iteration number is depend on total number of batches. For example, assume that we have 100 training datas. We divide the hundred training datas into 20 batches. It means that for one round of training five iterations would be needed. In our model we use 50 as our initial batch size.

3.2.2 Epoch

Epoch is the key to know the behavior of different networks. It deals with the model and is the indication of the dataset that is fed to model only once. These datasets passed through a single network for multiple times. Each time it operates with different updated weights and its goal is to get more accurate and better result. Gradient descent are used in different deep learning algorithms in different models. For a given number of epoch these Gradient descent computes the gradient of the loss function according to features for the training values. The number of the rounds of epoch is not fixed. For getting optimal weights different rounds of epochs required to prepare a model with the equivalent dataset. Distinctive datasets show diverse conduct and as a result different epoch might be expected to ideally prepare their systems uniquely.

3.2.3 Root Mean Square Error

The standard deviation of the residuals is defined by Root Mean Square Error (RMSE) .RMSE is used for inspecting the productivity of the model. By proper utilizing RMSE value the contrast between the target and the acquired output value can be minimized. RMSE is the square root of the mean average of the square of all of the error. RMSE is the standard deviation of the residuals. How far from the regression line data points are defined by the measurement of residuals. The utilization of RMSE is exceptionally normal and it makes a magnificent broadly useful blunder metric for numerical forecasts. For verifying experimental result RMSE used as common tool such as in climatology, future estimation, and regression analysis[29].RMSE intensifies and seriously rebuffs enormous mistakes like Mean Absolute Error.

$$\mathbf{RMSE}_{fo} = [\sum_{i=1}^{N} (z_{f_i} - z_{o_i})^2 / N]^{1/2}$$

Figure 3.7: RMSE.

3.3 Sentiment analysis approach

In this model, we collect data from news of different stock exchanges with the assistance of RSS news feed. From the predetermined website pages, RSS feed reader reads the required news content in the arrangement of XML report. XML document prepare on basis of news headlines from different websites. The last data of depiction is snatched subsequent to setting up the site parameter. At last it parses the XML archive from RSS channel list. This RSS channel gathers the share market exchange news as an informational collection of data set.
Pre-processing:

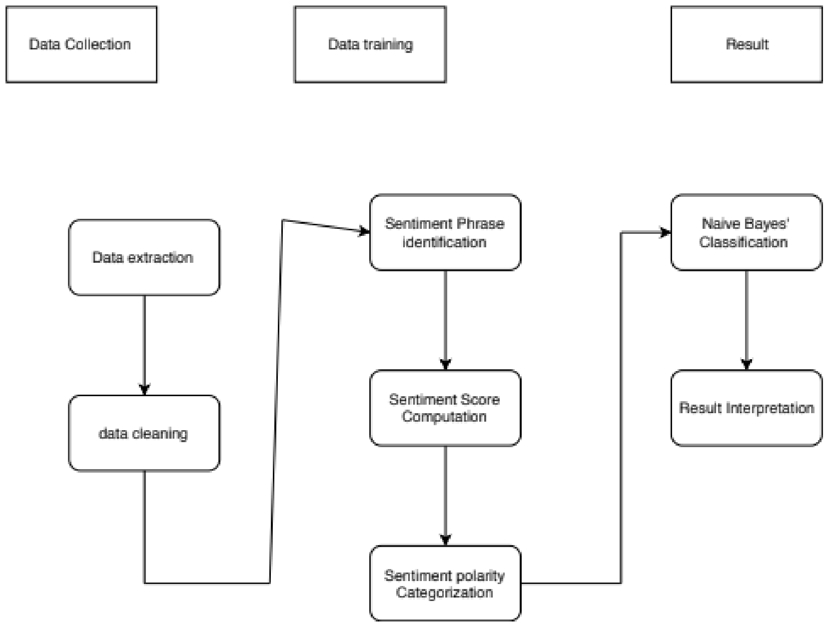

Figure 3.8: Sentiment flow chart.

In this stage it cleans data with removing the incorrect, incomplete, improperly formatted, or duplicated value. By filling missing datas ,leveling the noisy information, recognizing and expelling the exceptions it performs noise reduction. After completing this step, the pre-processed data handed to next step.

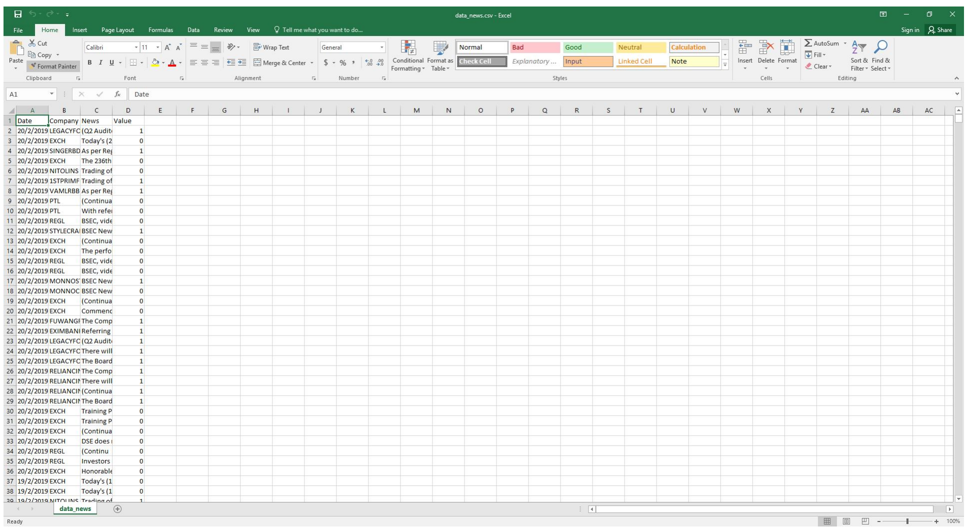

Figure 3.9: Data source for Sentiment Analysis.

Sentence splitting module :
Sentence splitting modules with splitting the prepossessed data in different section. These data are collected for testing purpose in a text document.

3.3.1 Natural language processing module

Sentiment analysis is the procedure utilize to regulate the nature of a particular study. Both natural language processing and text analytic are used to recognize and

extricate emotional data in source materials. This procedure not just discover the words that are indicative of sentiment, yet additionally to discover the connections between words so this assists to accurately recognize. For determining sentiment for the words having a positive/negative/neutral sentiment scaling system is used. Diverse procedure like part-of-speech tagger, dictionary based approach and Sentence Sentiment Score Algorithm are utilized to discover the extremity of the sentence, to close whether the sentence is positive, negative and unbiased. Positive and negative sentences have diverse estimation as positive feeling and negative notion.
Part-of-speech—tagger:
Syntactic labeling or word-classification disambiguation are known as, a Part-Of-Speech tagger. It is utilized in numerous Natural Language Processing assignments that do not require a full body electorate parse tree yet some data about the classes of words. Part-Of-Speech is a bit of programming that peruses message in some language and allots Part-Of-Speech to each word. The mix of noun-adverb has been decided for the best outcome for this module.
Dictionary based approach:
For regulate the placement of sentences dictionary based approach is utilized, which an effective method. In this process dictionary is used for the determination of words and their contradiction. It also use equivalent words, antonyms and chains of importance in to determine word sentiments.

3.3.2 Naive Bayes' classifiers

For text classification need high dimensional data to train and Naïve Bayes is the most popular method to do this. Here the theorem of probability likes Bayes being use. It is fast, effective and makes more accurate prediction. The naïve Bayes classifiers maintain the Bayes theorem but the naïve assure that the occurrence of any feature has no correlation among the class. P(A—B) = P(B—A) P(A) / P(B)................................(1) In conditional probability Bayes theorem gives the probability on the basis of previous circumstances. There are two circumstances A and B. P(A) represents the circumstance A and P(B) represent the circumstance B. For example Naïve Bayes classifier, there are corpus data based on news portal where the data set is full of positive and negative words. If we stored these words in two classes one is for positive words and one for negative words by natural language processing. Then we can find the probability of specific class by measuring the positive and negative sentiments. To train the data we need to preclassify the data set on the basis on sentiment. So the variable in the dataset can be represented as sentiment and the variable of the probability that a variable will occur given the evidence in the sentence can be represented as, P (sentiment — sentence) = P (sentence — sentiment) X P(sentiment)P(sentence) ....(2) Now if we think the word in a sentence as token and P(sentence— sentiment) as a result of P(token —sentiment) Now, for counting the number of times of a word we use token. P(token —sentiment)=count(this token in class)+1/count(all tokens in class)+count(all tokens) Here adding of 1 is called add one smoothing which will reduce the possibility of zero multiplication. However, the classifier calculates the prior probability. This probability represents the probability of positive and negative prior on the basis of positiveness and negativity of words which were taken as examples. Then the token are multiplied and lastly the final result of positive and negative ratio of news comes out.

3.3.3 Training Data

Here , for determining either a news article has positive or negative feedback for a stock we emphasize on Naïve Bayes classifier . Our main target is a news article and extract news segment. After that we organized these news segment on the basis of data set segment such as training and testing. We pulled around 1000 news segment from the news portal and prepared up to half of the news segment. We organized the information two polarity such as positive and negative. Two exhibits are set up for prescient positive words and prescient negative words. At that point, we check the recurrence of all these positive and negative words in the sentences. As indicated by the event of a prescient positive word for the genuine positive sentences and in the genuine negative sentences we utilize the mean factual procedure to discover a load for each word. For this, the preparation informational index will work all the more proficiently. Sentence level sentiment score algorithm:
For exploring the overall result we used Sentence Level sentiment score algorithm. Here for each individual sentence the investigation approach is connected lastly their outcomes are abridged to give the general consequence of the archive. As common, the score ranges from 0.0 to 1.0 and their whole is 1.0 for each syn set. Here, at first the Part-of-speech tagger is connected for each word and it indicates the tag as thing and intensifier in such a way which are equal to words. For finding sum of that sentence we appoint the score an incentive to each expression of a sentence.If the score value of segments is depended on positiveness and negativeness of that sentence.If it is 0.0, then regards as a negative segment and result is 1 then regards as positive segment.

Table 3.1: Example of Sentiment analysis data set

Date	**Company**	**News**	**value**
6-8-19	company A	positive	1
6-8-19	company B	negative	0
6-8-19	company C	positive	1
6-8-19	company D	negative	0
7-8-19	company A	negative	0
7-8-19	company B	positive	1
7-8-19	company C	positive	1
7-8-19	company D	negative	0

3.4 Software and Hardware

The dataset are prepared and organized in Python 3.5, This model use lots of packages like numpy[30] and pandas. Beside this we make use of sci-kit learn. The whole model is developed with keras and tensorflow which is a power full library and used for large scale heterogeneous systems.

Chapter 4

Result

4.1 Result Analysis

The RMSE is the standard deviation of the residuals and utilize to calculate accuracy which differentiates between estimated price and the actual price of a procedure. RMSE is very important as it is necessary to minimize the error of any model. For a standard model the differentiate rate of RMSE between testing data and training data should be minimum. Otherwise the model performed worse in testing and training and there is a high chance that the model over fit.

4.1.1 Comparative Results Using Different Parameters, batch and Epochs

Table 4.1: Result Analysis Using different Epoch Batch where Neuron is 50

Epoch	Batch Size	Train RMSE	TEST RMSE
50	50	0.00822	0.00888
100	50	0.00596	0.00673
200	50	0.00572	0.00657
400	50	0.00638	0.00804
50	100	0.00758	0.00907
100	100	0.00656	0.00721
200	100	0.00580	0.00645
400	100	0.00706	0.00796
50	200	0.00896	0.01010
100	200	0.00740	0.00849
200	200	0.00633	0.00702
400	200	0.00569	0.00626
50	400	0.01095	0.01275
100	400	0.00877	0.01012
200	400	0.00816	0.00911
400	400	0.00639	0.00694

In table number 1, we analyze with 50 Neurons and different number of parameters such as epoch and batch size.Performing various simulation we have observed that by taking 400 epoch and 400 batches we acquire result with instructing RMSE of 0.00639 and trailing RMSE of 0.00694.

4.1.2 Comparative Results Using Different Parameters, batch and Epochs

Table 4.2: Result Analysis Using different Epoch Batch where Neuron is 100

Epoch	Batch Size	Train RMSE	TEST RMSE
50	50	0.00652	0.00745
100	50	0.00618	0.00808
200	50	0.00632	0.00844
400	50	0.00512	0.00876
50	100	0.00758	0.00866
100	100	0.00635	0.00720
200	100	0.00591	0.00741
400	100	0.00537	0.00782
50	200	0.00885	0.01058
100	200	0.00774	0.00888
200	200	0.00609	0.00693
400	200	0.00668	0.00789
50	400	0.01118	0.01315
100	400	0.01002	0.01186
200	400	0.00697	0.00809
400	400	0.00629	0.00738

In table number 2, we analyze with 100 Neurons and different number of parameters such as epoch and batch size.Performing various simulation we have observed that by taking 200 epoch and 200 batches we obtain the result with instructing RMSE of 0.00609 and trailing RMSE of 0.00693.

4.1.3 Comparative Results Using Different Parameters, batch and Epochs

In table number 3, we analyze with 150 Neurons and different number of parameters such as epoch and batch size.Performing various simulation we have observed that by taking 100 epoch and 100 batches we acquire result with instructing RMSE of 0.00612 and trailing RMSE of 0.00680.

4.2 Historical data result analysis

For breaking down the proficiency of the framework we are utilized the Root Mean Square Error (RMSE). The contrast between the objective and the target value is limited by utilizing RMSE esteem. RMSE is the square root of the mean/normal of the square of the majority of the mistake. The utilization of RMSE is exceptionally normal and it makes an incredible universally useful blunder metric for numerical

Table 4.3: Result Analysis Using different Epoch Batch where Neuron is 150

Epoch	Batch Size	Train RMSE	TEST RMSE
50	50	0.00638	0.00732
100	50	0.00647	0.00714
200	50	0.00555	0.00766
400	50	0.00551	0.01034
50	100	0.00705	0.00801
100	100	0.00612	0.00680
200	100	0.00556	0.00662
400	100	0.00530	0.00772
50	200	0.01232	0.01243
100	200	0.00892	0.00992
200	200	0.00600	0.00672
400	200	0.00606	0.00696
50	400	0.01089	0.01234
100	400	0.00787	0.00944
200	400	0.00689	0.00764
400	400	0.00785	0.00846

expectations. On Comparison mind Mean Absolute Error, RMSE enhances and rebuffs enormous blunders.
In the wake of performing different reproductions with an alternate number of parameters,epochs,number of neurons and batch size,we have observed that evaluating with 50 neurons and by taking 400 epochs and 400 batch size we obtain the more accurate results with instructing RMSE of 0.00639 and trailing RMSE of 0.00694.

4.3 Sentimental result analysis

Three methods are used for processing sentiment analysis which are Accuracy, Precision and Recall. These methods are very effective for sentiment analysis. The result mainly depends on accuracy although we will use other two methods for more precise result. We used precision to see how much precise our classifier is. If the precision is higher it means there is less false positive on the other hand if the precision is lower it means we have more false positive. By decreasing Recall we can improve precision. In our model, we used Recall to measure completeness of a classifier. If the recall is high it means less false negatives, on the contrary if the recall is low there are more false negatives. In model, we find-out the accuracy, precision and recall percentage for both positive and negative set of words. We prepared the dataset with different parameters like positive, true positive, negative and true negative for calculating all percentages with different formula and we will follow comparison. Overall results are given in the below tables using different numbers of training dataset.

Table 4.4: Accuracy result on test dataset.

Number of trained dataset	**Accuracy**
50	59.6
100	66.8
150	69.2
200	72.9

Table 4.5: Precision positive result on test dataset.

Number of trained dataset	**Precision for positive**
50	56.6
100	59.2
150	64.3
200	66.8

Table 4.6: Precision negative result on test dataset.

Number of trained dataset	**Precision for negative**
50	44.6
100	58.2
150	66.3
200	74.8

Table 4.7: Recall positive result on test dataset.

Number of trained dataset	**Recall for positive**
50	69.1
100	66.1
150	68.6
200	73.2

Table 4.8: Recall negative result on test dataset.

Number of trained dataset	**Recall for negative**
50	59.6
100	66.8
150	69.2
200	72.9

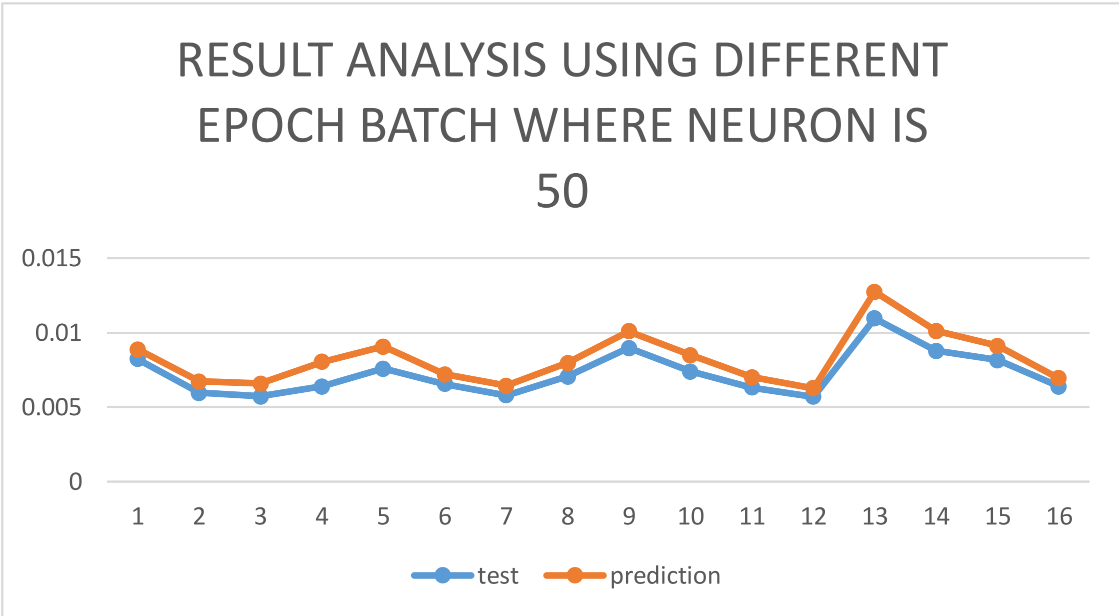

Figure 4.1: Evaluating with 50 neurons and by taking 400 epochs and 400 batch size

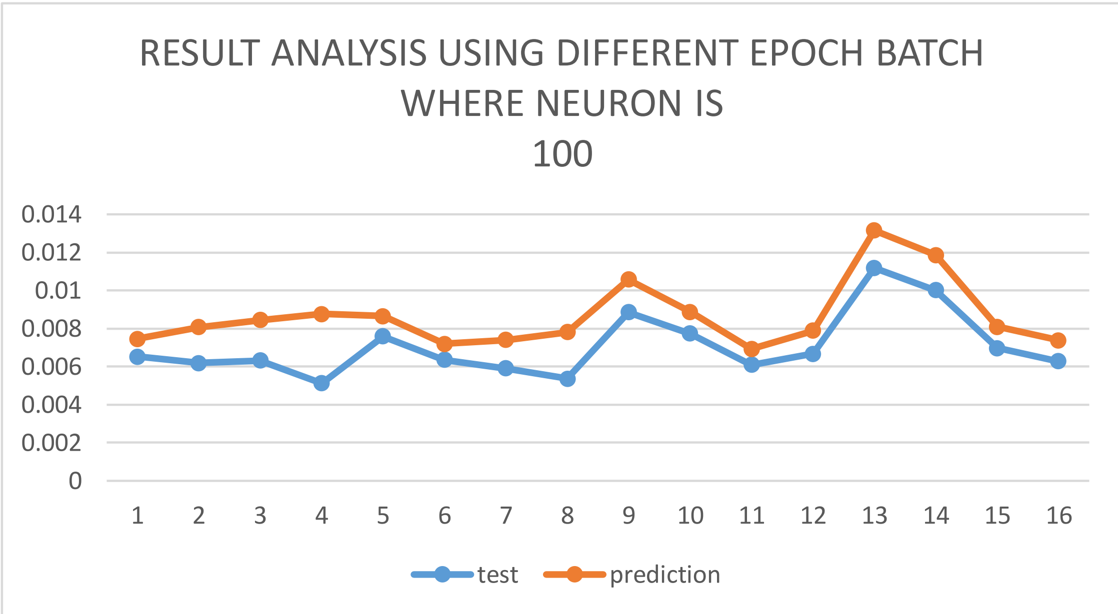

Figure 4.2: Evaluating with 100 neurons and by taking 200 epochs and 200 batch size

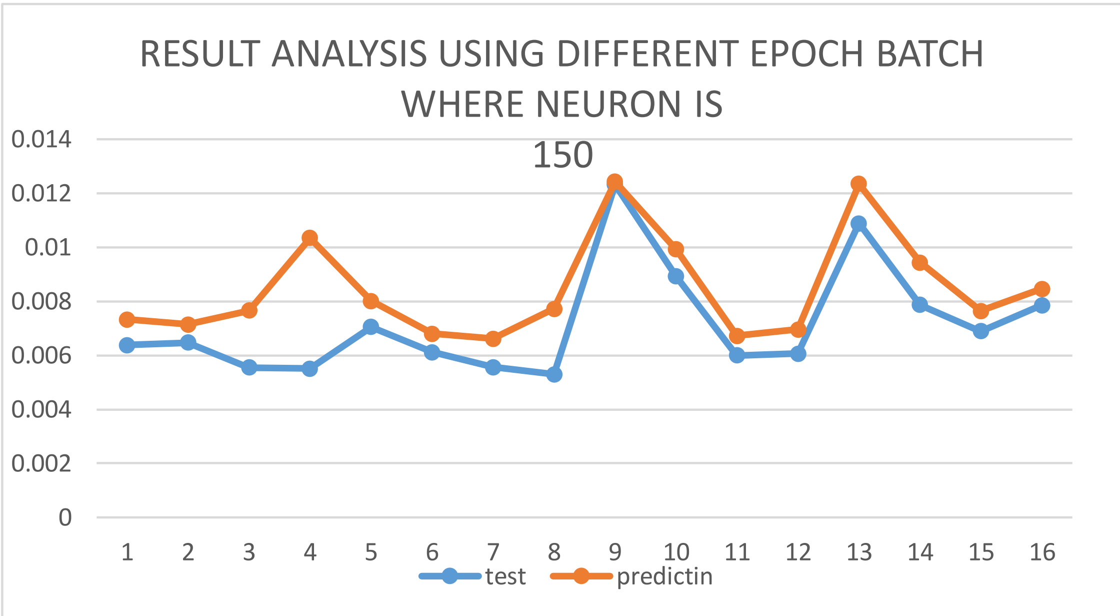

Figure 4.3: Evaluating with 150 neurons and by taking 100 epochs and 100 batch size

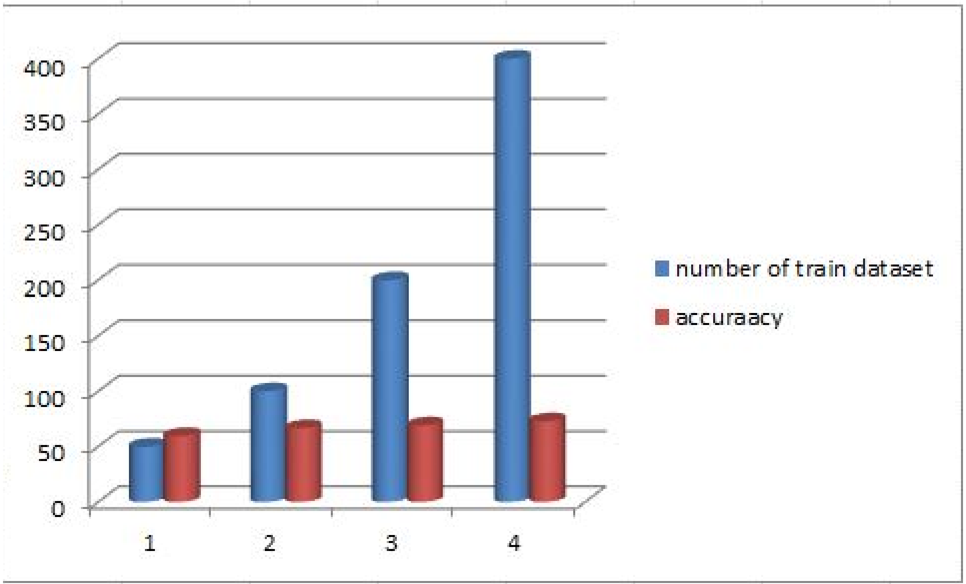

Figure 4.4: Accuracy result.

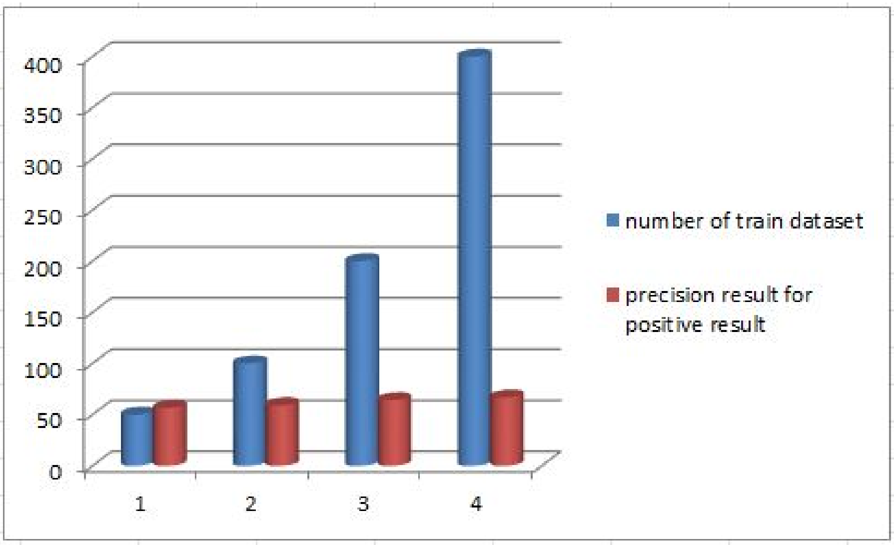

Figure 4.5: Precision positive result.

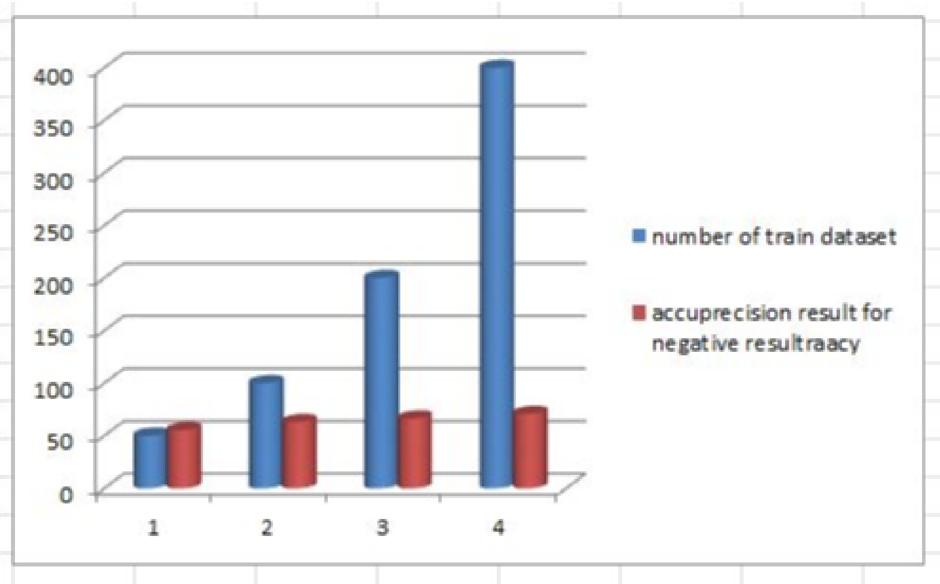

Figure 4.6: precision negative result.

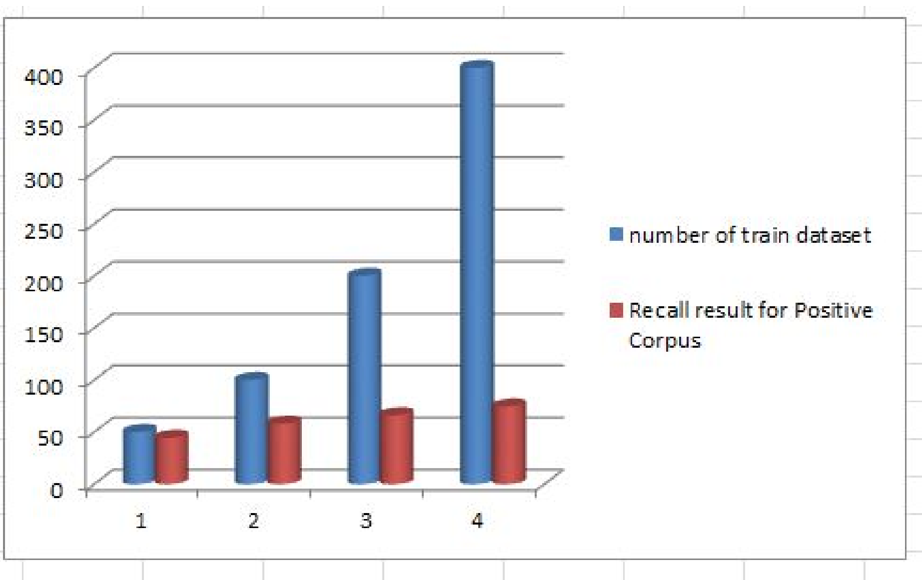

Figure 4.7: Recall positive result.

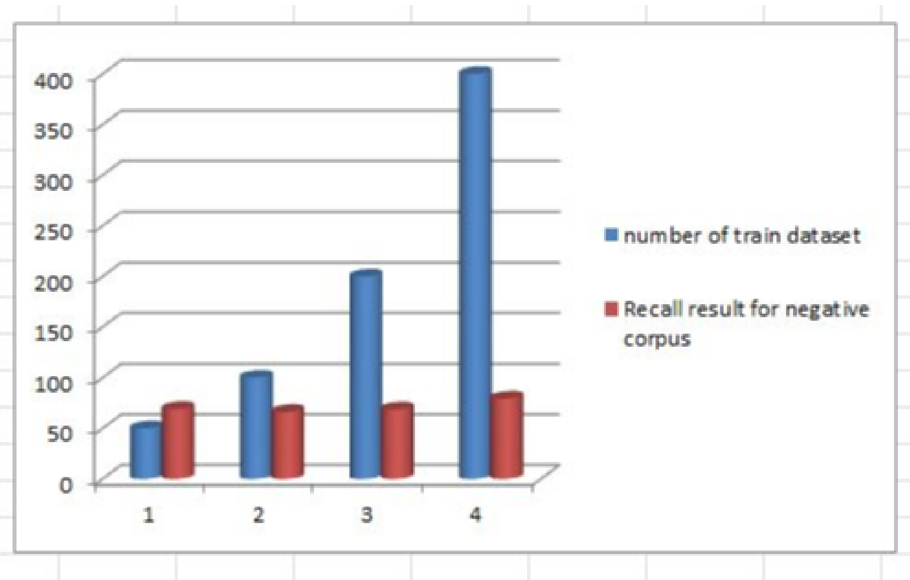

Figure 4.8: Recall negative result.

Chapter 5

Conclusion and Future Scope

5.1 Discussion

The outcome we got from or model is moderate. We are expecting that the efficiency and the accuracy can be enhanced with some more historical data of stock market. Our model would perform better knowledge if we can represent it by web application. We are arranging this for future work. In this paper, we connected an effective and precise yet simple model which gives an over-view of a company's stock price and provide knowledge to investors or any person who is interested in investing. Moreover, our model can provide the wanted news sentiment of a company's stock price that can help a investor or any person who wants to invest.

5.2 Motivation

There are many exiting model of share market which provide a overall prediction on stock price. But it is also important to know about news trend of share market. Our model provides both future prediction and news sentiment of a stock price. Usually for the complex and volatile nature of share market, people don't want to invest in share market because of the fear of loss. Our model can motivate many investors as well as business owners about share market. Moreover, it can be more efficient and accurate, if more better algorithm is used and find ways to collaborate with sentimental analysis which could be encouraging for researchers.

5.3 Future Scope

In this work, we have considered real data and news segment for analyzing share market prediction. As our main target is investor and we want to provide them a overview of share forecasting and trend prediction through a web application in future.

5.4 Limitation

It is very hard process to train and LSTM as it required a lot of time. LSTM requires a large amount memory bandwidth for computation. 4 different linear layer(MLP layer) per cell is required in LSTM for each time-step. This Linear layers needs large memory for computation. Often this layers could not use many computation unit because the system may not has enough memory bandwidth to feed the computational unit. It is also makes very difficult for designers to built that required large memory bandwidth. It often require more data to train than other models, and have lots of input parameters to tune.

5.5 Conclusion

Stock market trend is up rising in recent market news, which attracts many researchers to find out new methods for the prediction using new techniques. The forecasting method is not only helping the researchers but it also gives an overview of stock trend to investors and any person which may help them to understand the market. In order to help predict the stock indices, a forecasting model with good accuracy is required along with sentiment analysis on stock news. In our model, we worked with the most accurate forecasting technology LSTM (Long Short-Term Memory) neural network which is capable to solve the problem by distinguishing between present and past value. Investors, analysts or any person interested in investing in the stock market, our model could provide them a desired knowledge about stock price and market trend. In our work we try to establish a more accurate model to predict market movement.

References

[1] E. F. Fama, *Efficient capital markets: A review of theory and empirical work," journal of finance, 25*, 1970.

[2] B. G. Malkiel, *A random walk down wall street. w. w*, 1973.

[3] R. F. Engle, "Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation", *Econometrica: Journal of the Econometric Society*, pp. 987–1007, 1982.

[4] D. E. Rumelhart, "Hinton, ge and williams, r", *Learning representations by backpropagation error. nature*, vol. 323, pp. 533–536, 1986.

[5] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators", *Neural networks*, vol. 2, no. 5, pp. 359–366, 1989.

[6] J. L. Elman, "Finding structure in time", *Cognitive science*, vol. 14, no. 2, pp. 179–211, 1990.

[7] M. I. Jordan, *Attractor dynamics and parallelism in a connectionist sequential machine, artificial neural networks: Concept learning*, 1990.

[8] Y. Yoon, G. Swales Jr, and T. M. Margavio, "A comparison of discriminant analysis versus artificial neural networks", *Journal of the Operational Research Society*, vol. 44, no. 1, pp. 51–60, 1993.

[9] S. Haykin, *Neural networks: a comprehensive foundation.* Prentice Hall PTR, 1994.

[10] A. N. Refenes, A. Zapranis, and G. Francis, "Stock performance modeling using neural networks: A comparative study with regression models", *Neural networks*, vol. 7, no. 2, pp. 375–388, 1994.

[11] S. Hochreiter and J. Schmidhuber, "Long short-term memory", *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[12] B. Hammer, "On the approximation capability of recurrent neural networks", *Neurocomputing*, vol. 31, no. 1-4, pp. 107–123, 2000.

[13] H. Zhang, *The optimality of naive bayes, flairs conference*, 2004.

[14] K. Huarng and T. H.-K. Yu, "The application of neural networks to forecast fuzzy time series", *Physica A: Statistical Mechanics and its Applications*, vol. 363, no. 2, pp. 481–491, 2006.

[15] T. E. Oliphant, "Python for scientific computing", *Computing in Science & Engineering*, vol. 9, no. 3, pp. 10–20, 2007.

[16] W. McKinney *et al.*, "Data structures for statistical computing in python", in *Proceedings of the 9th Python in Science Conference*, Austin, TX, vol. 445, 2010, pp. 51–56.

[17] T.-J. Hsieh, H.-F. Hsiao, and W.-C. Yeh, "Forecasting stock markets using wavelet transforms and recurrent neural networks: An integrated system based on artificial bee colony algorithm", *Applied soft computing*, vol. 11, no. 2, pp. 2510–2525, 2011.

[18] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, *et al.*, "Scikit-learn: Machine learning in python", *Journal of machine learning research*, vol. 12, no. Oct, pp. 2825–2830, 2011.

[19] W. Shen, X. Guo, C. Wu, and D. Wu, "Forecasting stock indices using radial basis function neural networks optimized by artificial fish swarm algorithm", *Knowledge-Based Systems*, vol. 24, no. 3, pp. 378–385, 2011.

[20] S. Van Der Walt, S. C. Colbert, and G. Varoquaux, "The numpy array: A structure for efficient numerical computation", *Computing in Science & Engineering*, vol. 13, no. 2, p. 22, 2011.

[21] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization", *arXiv preprint arXiv:1412.6980*, 2014.

[22] H. Sak, A. Senior, and F. Beaufays, "Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition", *arXiv preprint arXiv:1402.1128*, 2014.

[23] Y. Xiao, J. Xiao, J. Liu, and S. Wang, "A multiscale modeling approach incorporating arima and anns for financial market volatility forecasting", *Journal of Systems Science and Complexity*, vol. 27, no. 1, pp. 225–236, 2014.

[24] A. M. Rather, A. Agarwal, and V. Sastry, "Recurrent neural network and a hybrid model for prediction of stock returns", *Expert Systems with Applications*, vol. 42, no. 6, pp. 3234–3241, 2015.

[25] S. S. Girija, "Tensorflow: Large-scale machine learning on heterogeneous distributed systems", *Software available from tensorflow. org*, 2016.

[26] H. Jia, "Investigation into the effectiveness of long short term memory networks for stock price prediction", *arXiv preprint arXiv:1603.07893*, 2016.

[27] F. Chollet *et al.*, *Keras https://github. com/fchollet/keras*, 2017.

[28] M. Roondiwala, H. Patel, and S. Varma, "Predicting stock prices using lstm", *International Journal of Science and Research (IJSR)*, vol. 6, no. 4, pp. 1754–1756, 2017.

[29] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions", *European Journal of Operational Research*, vol. 270, no. 2, pp. 654–669, 2018.

[30] Y. Li, "Deep reinforcement learning", *arXiv preprint arXiv:1810.06339*, 2018.