PROJECT REPORT

STOCK MARKET PREDICTION USING RNN & LSTM

**Padmajaa G ( 15Z235 )**

**Priyadarshini T ( 15Z240 )**

**Puja Saha ( 15Z242 )**

**Rajalakshmi P K ( 15Z243 )**

15Z611 - INTERNSHIP & INNOVATION PRACTICES LABORATORY

Branch: Computer Science & Engineering



**April 2018**

**COMPUTER SCIENCE & ENGINEERING**

##### PSG COLLEGE OF TECHNOLOGY

(Autonomous Institution)

COIMBATORE – 641 004

##### PSG COLLEGE OF TECHNOLOGY

(Autonomous Institution)

COIMBATORE – 641 004

**PROJECT REPORT**

STOCK MARKET PREDICTION USING RNN & LSTM

Bona fide record of work done by

**Padmajaa G ( 15Z235 )**

**Priyadarshini T ( 15Z240 )**

**Puja Saha ( 15Z242 )**

**Rajalakshmi P K ( 15Z243 )**

**15Z611 - INTERNSHIP & INNOVATION PRACTICES LABORATORY**

Branch: Computer Science & Engineering

**April 2018**

...……………………… …………..……………….  **Mr. P. Sampath Kumar Dr. R.Venkatesan**

Faculty In charge Head of the Department

Certified that the project was examined in the viva-voce examination held on ………………….

…………………….. ……………………..

(Internal Examiner) (External Examiner)

# SYNOPSIS

The purpose of this project is to predict the stock market using RNN(Recurrent Neural Network) and LSTM(Long Short-Term Memory). In the present scenario of financial market world, especially in the stock market, forecasting the trend or the price of the stocks using machine learning techniques such as artificial neural networks has become important. Unlike traditional machine learning models, the network learns from the examples by constructing an input-output mapping for the problem at hand. Neural networks are more noise tolerant and more flexible compared with traditional statistical models. Long short-term memory is a recurrent neural network designed to forecast, predict and classify time series data even long time lags between vital events happened before. This will be tested using OpenPOWER systems. This study seeks to use RNN to make predictions of the direction of the daily returns on NASDAQ index using the ‘technical analysis’ approach. The choice of RNN is motivated by the fact that ﬁnancial time series are noisy, chaotic and nonlinear in nature and hence, predicting its future behaviour requires a nonlinear approach of which RNN is. Researchers have taken different dimensions on predicting the returns on stocks with different approaches. Masoud (2014) used neural network to predict the direction of stock prices index movement and assessed the accuracy with the various error measurement techniques. Ou and Wang (2009) in an attempt to predict the changes in the daily closing prices of the Hand Seng index, using open price, low price, high prices, the S&P 500 index, US dollar- Hong Kong dollar exchange rate; reported various level of successes using different techniques, with all techniques achieving more than 80 percent hit-rate. Amin et al. (2014) achieved as high as above 89 percent accuracy in predicting the direction of stock price index in Tehran stocks. With this background, the results from this essay will be compared with the claims of other researches to observe how well our models are performing. Results from this essay would be useful for risk-averse investors, since it will help them to know when to invest and when to hold their resources.

CONTENTS

**CHAPTER Page No.**

# List of Figures (optional)………………………………………………...…………….. ()

# List of Tables (optional)……………………………………………………………… ()

**List of Symbols, Abbreviations or Nomenclature** (optional)**……………..…….. ()**

1. **…………………………………………………………...**
2. **…………………..…………………………………….**

:

l:

:

**N. CONCLUSION………………..…………………………………….……….…**

**CLIENT REPORT………………..…………………………………….……….………..**

**BIBLIOGRAPHY ..…………………………………………………………….…………**

**APPENDICES …………..………………………………………………………….…….**

1. **Introduction**

**1.1 Motivation**

[Machine learning](https://www.dezyre.com/article/what-is-machine-learning-and-how-it-is-remaking-our-world/191?utm_source=TextLink&utm_medium=DeZyre&utm_campaign=Blog_LearnMachineLearning_362" \o "What is Machine Learning?" \t "https://www.dezyre.com/article/why-you-should-learn-machine-learning/_blank) is the shining star of the moment. With every industry looking to apply AI in their domain, studying machine learning opens world of opportunities to develop cutting edge machine learning applications in various verticals – such as cyber security, image recognition, medicine, or face recognition. With several machine learning companies on the verge of hiring skilled ML engineers, it is becoming the brain behind business intelligence.

We as the students of computer science, want to explore this trending technology. As a beginner our knowledge on this field was casual. We started with an online course. By digging more and more into machine learning we got to know about deep learning. Our interests in artificial neural networks grew gradually and lead us to take up a project on that.

The financial sector has become one of the most important subjects of human life in recent years.Investors desire to invest their assets in the most viable manner. However, they want to achieve this at the lowest risk of loss. Information on the trend and behaviour of the market is therefore relevant.The use of algorithms to make trading decisions has become a prevalent practice in major stock exchanges of the world. Algorithmic trading, sometimes called high-frequency trading, is the use of automated systems to identify true signals among massive amounts of data that capture the underlying stock market dynamics. Machine Learning has therefore been central to the process of algorithmic trading because it provides powerful tools to extract patterns from the seemingly chaotic market trends.This project has given us the opportunity to explore the world of stock market.

**1.2 Problem Statement**

Stock market is very vast and difficult to understand. It is considered too uncertain to be predictable due to huge fluctuation of the market. The stock of a corporation constitutes the equity stack of owner. It represents the residual assets of the company that would be due to stockholders after discharge of all senior claims such as secured and unsecured debt.

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 could yield significant profit. Stock market prediction task is interesting as well as divides researchers and academics into two groups, those who believe that we can devise mechanisms to predict the market and those who believe that the market is efficient and whenever new information comes up the market absorbs it by correcting itself, thus there is no space for prediction.

Investing in a good stock but at a bad time can have disastrous result, while investing in a stock at the right time can bear profits. Financial investors of today are facing this problem of trading as they do not properly understand as to which stocks to buy or which stocks to sell in order to get optimum result. So, the purposed project will reduce the problem with suitable accuracy faced in such real time scenario.

**1.3 Objectives**

According to the Firm Foundation theory the market is defined from the reaction of the investors, which is triggered by information that is related to the “real value” of firms. The “real value” or else the intrinsic value is determined by careful analysis of present conditions and future prospects of a firm. On the other hand, according to the Castles in the Air theory the investors are triggered by information that is related to other investor’s behaviour. So for this theory the only concern the investor should have is to buy today with the price of 20 and sell tomorrow with the price of 30, no matter what the intrinsic value of the firm he (or she) invests in is. Therefore, the Firm Foundation theory favour the view that the market is defined mostly by logic, while the Castles in the Air theory supports that the market is defined mostly by psychology. Despite its prevalence, Stock Market prediction remains a secretive and empirical art. Few people, if any, are willing to share what successful strategies they have. A chief goal of this project is to add to the academic understanding of stock market prediction. The hope is that with a greater understanding of how the market moves, investors will be better equipped to prevent another financial crisis. The project will evaluate some existing strategies from a rigorous scientific perspective and provide a quantitative evaluation of new strategies.

The aims of this project are as follows:

* To identify factors affecting stock market
* To generate the pattern from large set of data of stock market for prediction of multiple companies
* To predict an approximate value of stock price
* To provide analysis for users through web application

**1.4 Scope of the Project**

The project will be useful for investors to invest in stock market based on the various factors. The project target is to create web application that analyses previous stock data of companies and implement these values in data mining algorithm to determine the value that particular stock will have in near future with suitable accuracy. These predicted and analyzed data can be observed by individual to know the financial status of companies and their comparisons. Company and industry can use it to breakdown their limitation and enhance their stock value. It can be very useful to even researchers, stock brokers, market makers, government and general people. The main feature of this project is to generate an approximate forecasting output and create a general idea of future values based on the previous data by generating a pattern. The scope of this project does not exceed more than a generalized suggestion tool.

Although vital to any investor operating in the real world, no attempt is made in this project at portfolio management. Portfolio management is largely an extra step done after an investor has made a prediction on which direction any particular stock will move. The investor may choose to allocate funds across a range of stocks in such a way to minimize his or her risk. For instance, the investor may choose not to invest all of their funds into a single company lest that company takes unexpected turn. A more common approach would be for an investor to 3 investors across a broad range of stocks based on some criteria he has decided on before. This project will focus exclusively on predicting the daily trend (price movement) of individual stocks. The project will make no attempt to deciding how much money to allocate to each prediction. More so, the project will analyse the accuracy of these predictions.

**2. System Requirements**

The requirement that the system needs is categorized into the functional, non-functional, software and hardware requirements. These requirements are listed below

**2.1 Functional Requirements**

Functional requirement are the functions or features that must be included in any system to satisfy the business needs and be acceptable to the users.

Based on this, the functional requirements that the system must require are as follows:

* The system should be able to generate an approximate stock price.
* The system should collect accurate data from the “**morning star**” website in consistent manner.

**2.2 Non-Functional Requirements**

Non-functional requirement is a description of features, characteristics and attribute of the system as well as any constraints that may limit the boundaries of the proposed system.The non- functional requirements are essentially based on the performance, information, economy, control and security efficiency and services.

Based on these the non-functional requirements are as follows:

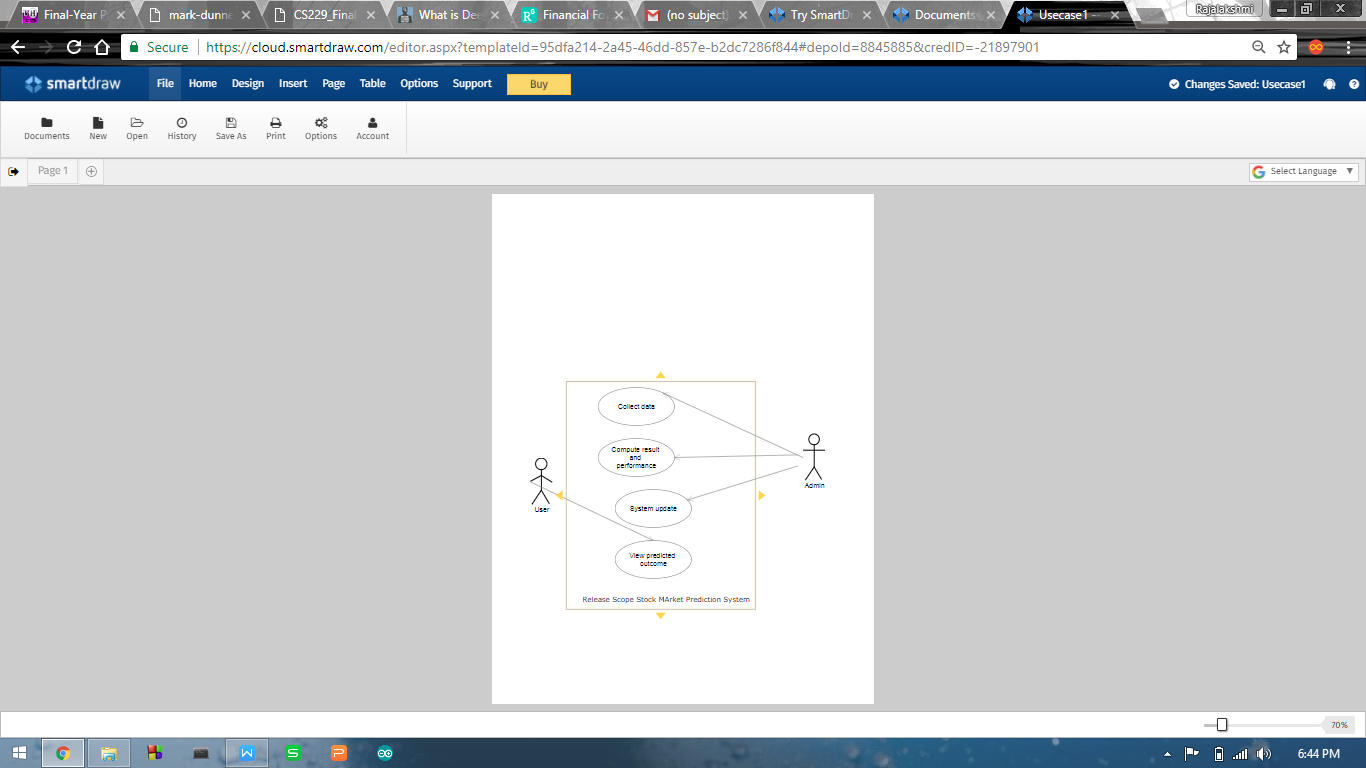
* The system should provide better accuracy
* The system should have simple interface for users to use
* To perform efficiently in short amount of time

**2.3 Software and hardware requirements**

|  |  |
| --- | --- |
| Software Requirements | Language : Python  IDE : Anaconda  Front end : Keras  Back end : Tensorflow |
| Hardware Requirements | OpenPOWER systems |

**3. System Design**

**3.1 Use case diagram:**



**3.1.1 Use case description:**

Use case ID:1

Use case name: Collect data

Description: Every required data will be available in “morning star”. Admin will be able to collect the data for system.

Use case ID:2

Use case name: Compute result and performance

Description: Prediction result will be handled and generated by admin. The system will be built, through which the result of prediction and system performance will be analyzed.

Use case ID: 3

Use case name: System update

Description: With the change of market and technology regular update of system is required. Beside that , predicted result of stock exchange and their actual price will be updated by admin in regular daily basis.

Use Case ID: 4

Use Case Name: View predicted outcome

Description: This use case is must important in whole project. The key feature of this project is to predict the stock value of companies. Thus, this will be available in user interface and viewer can observe them.

**4.System implementation and integration**

The purposed method for developing the system consists of mainly two main steps. Firstly, data is collected from web. Secondly, relevant features are selected. At last, an RNN with LSTM is designed and a suitable algorithm yielding best accuracy is chosen to predict the stock value.

**4.1 Data Source**

This project attempts to predict the stock value with respect to the stock’s previous value and trends. It requires historic data of stock market as the project also emphasizes on data mining techniques. So, it is necessary to have a trusted source having relevant and necessary data required for the prediction. We will be using Stock Exchange website (<http://www.morningstar.com> ) as the primary source of data. This website contains all the details such as: Opening value, Closing value, Highest value, Lowest value, number of shares, adjusted close value, adjusted opening value, increase or decrease in stock values for each financial companies. The site is updated on daily basis and it is also a repository for years of stock market data for all companies.

**4.2 Selection of Company**

The stock market is a very fluctuating market. There are many companies of different sectors and the values as well as parameters can vary differently in time. In this case, same rules or logic for constructing a prediction model may not apply to the all the companies. So, this project performs analysis and prediction on only the companies that fall in the IT sector.

**4.3 RNN Design and Training**

The main problem in predicting share market is that the share market is a chaos system. There are many variables that could affect the share market directly or indirectly. There are no significant relations between the variables and the price. We cannot draw any mathematical relation among the variables. There are no laws of predicting the share price using these variables. For this kind of chaotic system the neural network approach is suitable because we do not have to understand the solution. This is a major advantage of neural network approaches. On the other hand in the traditional techniques we must understand the inputs, the algorithms and the outputs in great detail. With the neural network we just need to simply show the correct output for the given inputs. With sufficient amount of training, the network will mimic the function. Another advantage of neural network is that during the tanning process, the network will learn to ignore any inputs that don’t contribute to the output. In our purposed system, there is a training phase where some parameters named weights are found from this section and Back-propagation Algorithm is used for this training phase. These weights are used in prediction phase using same equations which are used in training phase.

**4.3.1 Data-set Creation**

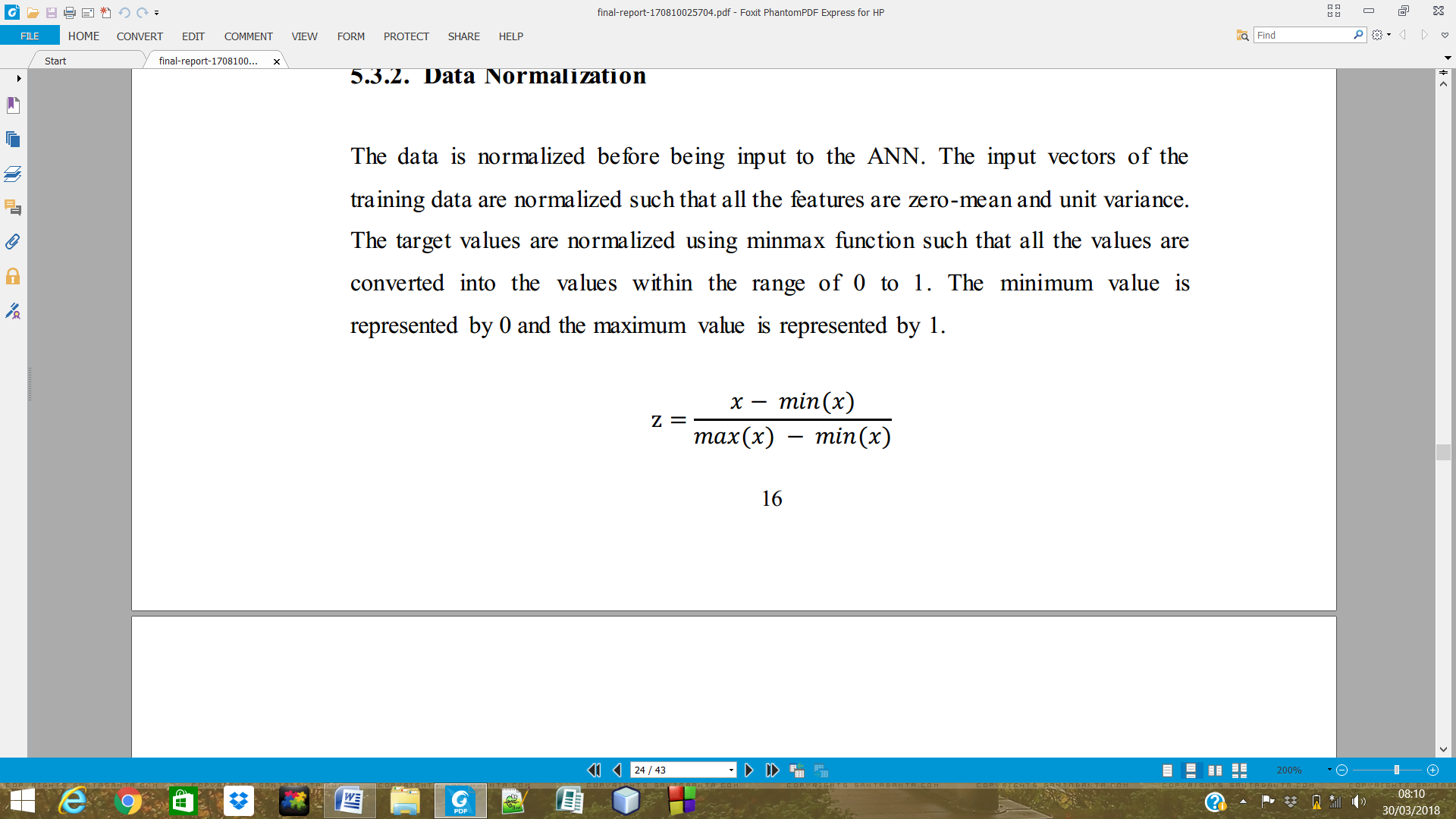
First of all, a data-set is created for training the Re-current neural network. 80% of the dataset is taken for training and remaining 20% for testing.

**4.3.2 Feature Selection**

There are so many features such that opening price, highest price, closing price, lowest price, volume, adjusted close price, split ratio, adjusted open price etc. We need to select the relevant features. We have chosen opening price, highest price, lowest price, volume as input features and closing price as output.

**4.3.3. Data Normalization**

The data is normalized before being input to the RNN. The input vectors of the training data are normalized such that all the features are zero-mean and unit variance. The target values are normalized using min-max function such that all the values are converted into the values within the range of 0 to 1 . The minimum value is represented by 0 and the maximum value is represented by 1.



**4.3.4. Backpropagation with Feed forward Neural Network**

The main steps using the Backpropagation algorithm as follows:  
**Step 1:** Feed the normalized input data sample, compute the corresponding output.  
**Step 2:** Compute the error between the output(s) and the actual target(s).  
**Step 3:** The connection weights and membership functions are adjusted.  
**Step 4:** IF Error > Tolerance THEN go to Step 1 ELSE stop

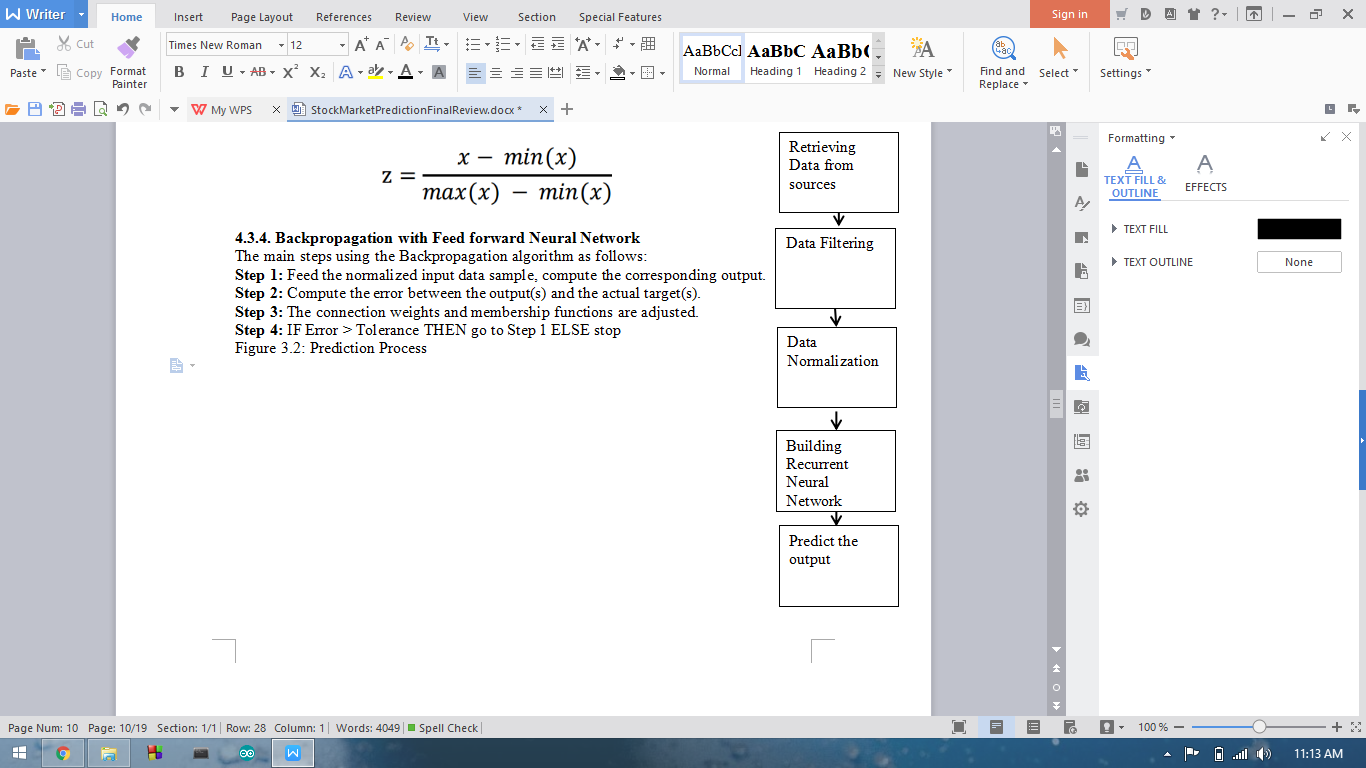
****

Figure 3.2: Prediction Process

**4.4. Training Parameters**

In order to select optimal parameters for the neural network, simulation is carried out. A model of a neural network is constructed and simulated using simulation tool. Test runs are carried out and the model yielding the best accuracy is selected for implementation. The best model so far has the following parameters:

Initial weights : Weights are initialized randomly

Learning rate : 0.01

Total Layers : 2

Input Neurons : No of days \* No of features (4)

[since , start date is Jan 1, 2010 and end date is Mar 27, 2018]

Hidden Layer Neurons : 50

Output Neurons : 1

Limit of epochs : 50000

**4.5. Execution and calculation of result**

After training the neural network, weights and bias are set accordingly to calculate the closing price. The network used as a feed forward network, which gives a certain output when given a set of inputs.

The feed forward network then calculates an output according to the weights and bias of the individual neurons.The output normalized data is then converted into normal form by using reverse min-max formula.

**4.6. Model Design**

We use recurrent neural network which has a input layer with (No of days\*4) neurons, a single hidden layer which has 50 neurons and a output layer with single neuron. The Back-propagagtion algorithm has been used for training the network.

1

2

n

3

1

2

50

3

Input layer

Hidden layer(LSTM)

Output layer

**4.7 Input & Output Features**

**4.7.1. Opening Price**

It is the first price of a company stock.

**4.7.2. Highest Value**

It is the highest value the share price of a company that has reached in the previous day.

**4.7.3. Lowest Value**

Similar to the highest value, it is the lowest value the share price of a company that has reached in the previous day.

**4.7.4. Share Volume**

Share volume can be calculated in two different types the daily share volume and the monthly share volume. the total number of share is sold in a particular day is called daily share volume. In monthly share volume is the sum of the trading volumes during that month.  
**4.7.5 Closing price**Closing price generally refers to the last price at which a stock trades during a regular  
trading session.

**4.8.Programming tools and External Dependencies**

The following tools are used

Xampp - server

Mysql - Database to store the predicted values

Canvas js - to draw chart

**5.Testing**

**5.1. Unit Testing**

Unit testing is carried out for testing modules constructed from the system design. Each part is compiled using inputs for specific modules. Every modules are assembled into a larger unit during the unit testing process. Testing has been performed on each phase of project design and coding. The testing of module interface is carried out to ensure the proper flow of information into and out of the program unit while testing. The temporarily generated output data is ensured that maintains its integrity throughout the algorithm's execution by examining the local data structure. Finally, all error-handling paths are also tested.

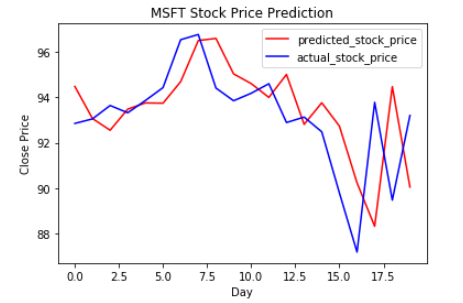
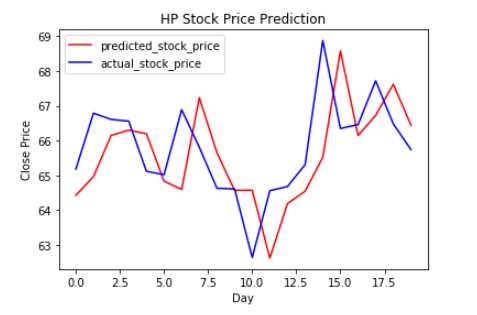
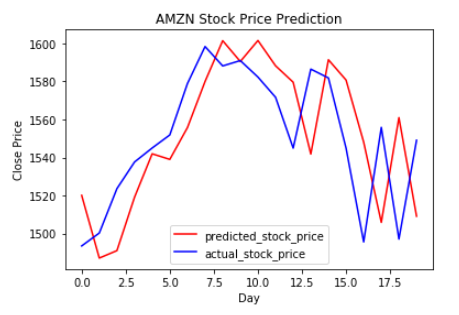
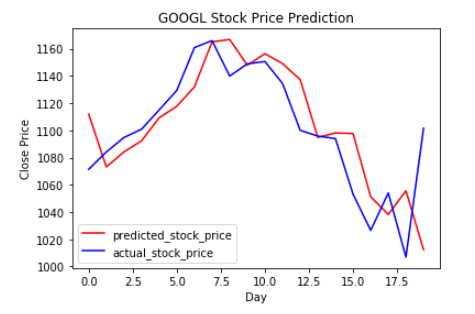
**5.2. Integration Testing**

We usually perform system testing to find errors resulting from unanticipated interaction between the sub-system and system components. Software must be tested to detect and rectify all possible errors once the source code is generated before delivering it to the customers. For finding errors, series of test cases must be developed which ultimately uncover all the possibly existing errors. Different software techniques can be used for this process. These techniques provide systematic guidance for designing test that exercise the internal logic of the software components and exercise the input and output domains of a program to uncover errors in program function, behaviour and performance. We test the software using two methods: White Box testing: Internal program logic is exercised using this test case design techniques. Black Box testing: Software requirements are exercised using this test case design techniques. Both techniques help in finding maximum number of errors with minimal effort and time.  
  
**5.3. Verification and Validation**

The testing process is a part of broader subject referring to verification and validation. We have to acknowledge the system specifications and try to meet the customer’s requirements and for this sole purpose, we have to verify and validate the product to make sure everything is in place. Verification and validation are two different things. One is performed to ensure that the software correctly implements a specific functionality and other is done to ensure if the customer requirements are properly met or not by the end product. Verification of the project was carried out to ensure that the project met all the requirement and specification of our project. We made sure that our project is up to the standard as we planned at the beginning of our project development.  
  
**6. Analysis and Results**

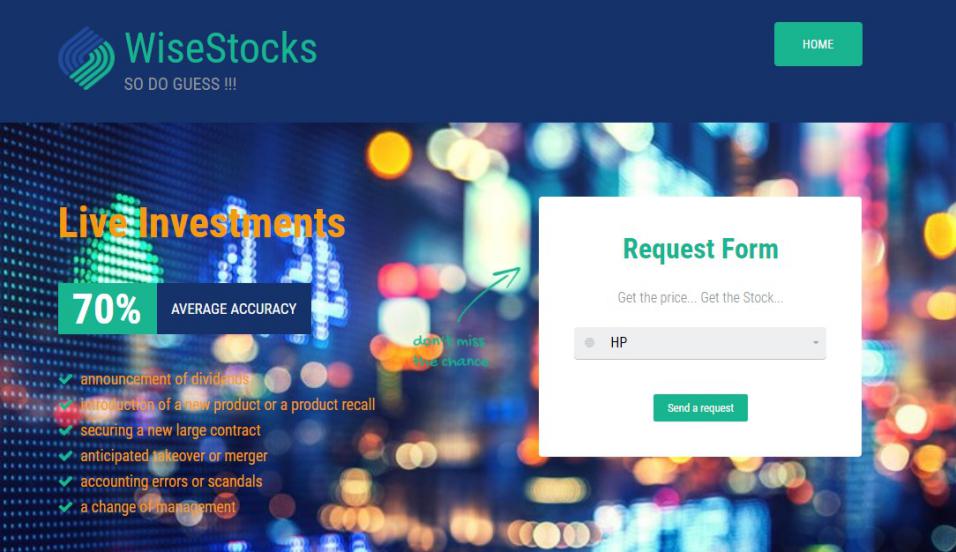
**6.1. Analysis**

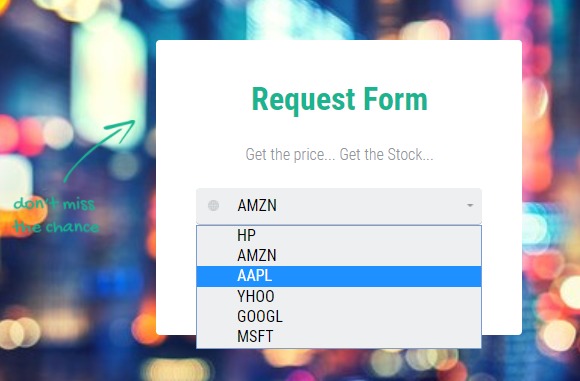
In this project, the factors that are taken into account for change in the closing price of a particular company over the past 20 days are: Opening price, highest value, lowest value, share volume and closing price. We performed analysis on obtained data to establish relation between our output parameters and the selected factors.

**  **

**6.2 Result**

After collection of data the next day’s stock price is predicted using neural network. The value is then compared the next day with the actual value. An website “WiseStocks” is created to display each company’s prediction trend in past 20 day’s including today.







**7. Conclusion**

This study presents different ways of predicting the direction of the returns on the NASDAQ stock index using Recurrent Neural Networks and LSTM. It has been observed that Artificial Neural Network performed relatively better than the other techniques. Although none of the models achieved a 100 percent accuracy, the result from the trained. ANN model suggests that investors can develop a trading strategy with the model with some level of confidence. Future works should consider predicting the returns on the index for longer periods and also the actual returns; taking the high and low prices during the day, the volume traded and the other economic indicators such as exchange rates and consumer index into account

**ACKNOWLEDGEMENT**

Our sincere gratitude goes to our professors Mrs.Sudha Sadhasivam, Mr.Sampath kumar and Ms.Bharti.

# BIBLIOGRAPHY

1. Machine learning on [www.Coursera.org](http://www.Coursera.org) by Andrew NG
2. Deep learning.ai on [www.Coursera.org](http://www.Coursera.org) by Andrew NG
3. Deep learning course on udemy

<https://www.udemy.com/deeplearning>

1. Multivariant time-series prediction <https://machinelearningmastery.com/multivariate-time-series-forecasting-lstms-keras/>
2. Storing the neural network model as .json file <https://machinelearningmastery.com/save-load-keras-deep-learning-models/>