TRIBHUVAN UNIVERSITY

INSTITUTE OF ENGINEERING

Himalaya College of Engineering

[Code No: CT755]

A

FINAL YEAR PROJECT

ON

STOCK MARKET ANALYSIS AND PREDICTION USING ARTIFICIAL NEURAL NETWORK

BY

Apar Adhikari (070/BCT/03)

Bibek Subedi (070/BCT/04)

Bikash Ghimirey (070/BCT/06)

Mahesh Karki (070/BCT/22)

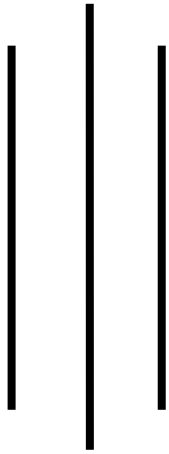
A REPORT SUBMITTED TO DEPARTMENT OF ELECTRONICS AND COMPUTER ENGINEERING IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR BACHELOR'S DEGREE IN COMPUTER ENGINEERING

DEPARTMENT OF ELECTRONICS AND COMPUTER ENGINNERING

LALITPUR, NEPAL

AUGUST, 2017

A
PROJECT REPORT
ON
STOCK MARKET ANALYSIS AND PREDICTION USING ARTIFICIAL NEURAL NETWORK



Prepared For

Department of Electronics and Computer Engineering
Himalaya College of Engineering
Chyasal, Lalitpur

Approved By

Er. Bal Krishna Nyaupane
(Project Supervisor)

Prepared By

Apar Adhikari (070/BCT/03)
Bibek Subedi (070/BCT/04)
Bikash Ghimirey (070/BCT/06)
Mahesh Karki (070/BCT/22)

August, 2017

ACKNOWLEDGEMENT

It is a matter of great pleasure to present this progress report on "Stock Market Analysis and Prediction using Artificial Neural Network". We are grateful to Institute of Engineering, Pulchowk and Himalaya College of Engineering management for providing us with this great opportunity to develop a web application for the major project.

We are also grateful to our project supervisor Er. Bal Krishna Nyaupane, who has been of invaluable assistance to our project with his advice and suggestions.

We would like to thank our respected sir Er. Ashok GM, Head of Department and Er. Alok Smarat Kafle, Major Project Coordinator, for willing to provide us with all the resource required to complete our project.

Group Members

Apar Adhikari

Bibek Subedi

Bikash Ghimirey

Mahesh Karki

ABSTRACT

Stock price forecasting is a popular and important topic in financial and academic studies. Share Market is an untidy place for predicting since there are no significant rules to estimate or predict the price of share in the share market. Many methods like technical analysis, fundamental analysis, time series analysis and statistical analysis, etc. are all used to attempt to predict the price in the share market but none of these methods are proved as a consistently acceptable prediction tool.

In this project we attempt to implement an Artificial Neural Network approach to predict stock market prices. Artificial Neural networks are very effectively implemented in forecasting stock prices, returns, and stock modeling, and the most frequent methodology is the Backpropagation algorithm. This project is for Nepalese users as the prediction is done on the listed companies of Nepal Stock Exchange Ltd. We outline the design of the Neural Network model with its salient features and customizable parameters. We select a certain group of parameters with relatively significant impact on the share price of a company. With the help of statistical analysis, the relation between the selected factors and share price is formulated which can help in forecasting accurate results. Although, share market can never be predicted, due to its vague domain, this project aims at applying Artificial Neural Network in forecasting the stock prices.

Keywords: ***Artificial Neural Network, Backpropagation, Hydropower, NEPSE***

ABBREVIATIONS

ANN : Artificial Neural Network

API : Application Programming Interface

CSS : Cascading Style Sheet

EHM : Efficient Market Hypothesis

HTML : Hyper Text Markup Language

NEPSE : Nepal Stock Exchange

SP500 : Standard and Poor's 500

SQL : Structured Query Language

LIST OF FIGURES

TABLE OF CONTENTS

1. INTRODUCTION

Stock Market prediction and analysis is the act of trying to determine the future value of a company stock or other financial instrument traded on an exchange. Stock market is the important part of economy of the country and plays a vital role in the growth of the industry and commerce of the country that eventually affects the economy of the country. Both investors and industry are involved in stock market and wants to know whether some stock will rise or fall over certain period of time. The stock market is the primary source for any company to raise funds for business expansions. It is based on the concept of demand and supply. If the demand for a company's stock is higher, then the company share price increases and if the demand for company's stock is low then the company share price decrease.

Another motivation for research in this field is that it possesses many theoretical and experimental challenges. The most important of these is the Efficient Market Hypothesis(EMH), the hypothesis says that in an efficient market , stock market prices fully reflect available information about the market and its constituents and thus any opportunity of earning excess profit ceases to exist. One of the example of big exchange is New York Stock Exchange.

The Nepal Stock Exchange Limited(NEPSE) is the only stock exchange of Nepal. It is established under the company act, operating under securities exchange act, 1983. It is located in Singha Durbar Plaza, Kathmandu Nepal. The basic objective of NEPSE is to import free marketability and liquidity to the government and corporate securities by facilitating transactions in its trading floor through member , market intermediaries, such as broker, market makers etc. NEPSE opened its trading floor on 13th January 1994.

Government of Nepal, Nepal Rastra Bank, Nepal Industrial Development Corporation and members are the shareholders of NEPSE. Members of NEPSE are permitted to act as intermediaries in buying and selling of government bounds and listed corporate securities. At present, there are 50 member broker and 2 market makers, who operate on the trading floor as per the Securities Exchange Act, 1983 rule and bye-law.

Due to involvement of many number of industries and companies, it contain very large sets of data from which it is difficult to extract information and analyze their trend of work manually. Stock market analysis and prediction will reveal the market patterns and predict the time to purchase stock. The successful prediction of a stock's future price could yield significant profit. This is done using large historic market data to represent varying conditions and confirming that the time series patterns have statistically significant predictive power for high probability of profitable trades and high profitable returns for the competitive business investment.

1.1. Statement of the problem

Stock market is very vast and difficult to understand. It is considered too uncertain to be predictable due to huge fluctuation of the market. 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.

1.2. Objectives

The aims of this project are as follows:

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

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.

1.3. System Overview

This system named "Stock Market Analysis and Prediction using Artificial Neural Networks" is a web application that aims to predict stock market value using Artificial Neural Network. This project is intended to solve the economic dilemma created in individuals that wants to invest in Stock Market.

1.4. System Features

1.4.1. Stock market prediction

Stock price movements are in somewhat repetitive in nature in the time series of stock values. The prediction feature of this system tries to predict the stock return in the time series value by training Neural Network which involves producing an output and correcting the error.

1.4.2. Market Analysis

A detailed analysis of Stock market is presented to the user. The analysis contains the performance of most of the listed companies for certain interval of days. The numbers and figures are represented in graphs and plots in the form of line charts.

2. LITERATURE REVIEW

In the last few decades forecasting of stock returns has become an important field of research. In most of the cases the researchers had attempted to establish a linear relationship between the input macroeconomic variables and the stock returns. After the discovery of nonlinearity in the stock market index returns, many literatures have come up in nonlinear statistical modeling of the stock returns, most of them required that the nonlinear model be specified before the estimation is done. But since stock market return is noisy, uncertain, chaotic and nonlinear in nature, ANN has evolved out to be better technique in capturing the structural relationship between a stock's performance and its determinant factors more accurately than many other statistical techniques.

In literature, different sets of input variables are used to predict stock returns. In fact, different input variables are used to predict the same set of stock return data. Some researchers used input data from a single time series where others considered the inclusion of heterogeneous market information and macro-economic variables. Some researchers even preprocessed these input data sets before feeding it to the ANN for forecasting.

2.1. Relevant Works

Wilson and Sharda [1] studied prediction firm bankruptcy using neural networks and classical multiple discriminant analysis, where neural networks performed significantly better than multiple discriminant analysis.

Min and Lee were doing prediction of bankruptcy using machine learning. They evaluated methods based on Support Vector Machine, multiple discriminant analysis, logistic regression analysis, and three-layer fully connected back-propagation neural networks. Their results indicated that support vector machines outperformed other approaches.

Lee was trying to predict credit rating of a company using support vector machines. They used various financial indicator and ratios such as interest coverage ratio, ordinary income to total assets, Net income to stakeholders' equity, current liabilities ratio, etc. and achieved accuracy of around 60%. Predicting credit rating of the companies were also studied using neural networks achieving accuracy between 75% and 80% for the United States and Taiwan markets.

Tsai and Wang [2] did a research where they tried to predict stock prices by using ensemble learning, composed of decision trees and artificial neural networks. They created dataset from Taiwanese stock market data, taking into account fundamental indexes, technical indexes, and macroeconomic indexes. The performance of Decision Tree + Artificial Neural Network trained on Taiwan stock exchange data showed F-score performance of 77%. Single algorithms showed F-score performance up to 67%.

Kim and Han [3] used a genetic algorithm to transform continuous input values into discrete ones. The genetic algorithm was used to reduce the complexity of the feature space. This paper proposes a novel evolutionary computing method called a genetic quantum algorithm. Genetic Quantum Algorithm is based on the concept and principles of quantum computing such as qubits and superposition of states. Instead of binary, numeric, or symbolic representation, by adopting bit chromosome as a representation Genetic Quantum Algorithm can represent a linear superposition of solutions due to its probabilistic representation. As genetic operators, quantum gates are employed for the search of the best solution.

There are many tools and software available out there that provide forecasting of stock market entities, share quantity and share value for a given financial organization. Most of them claim to predict the stock market with near to 100% accuracy but the opinions from the users vary. Some of the popular tools and software with their methodologies are mentioned as follows.

2.2. inteliCharts Predictive Stock Market Analytics

It is a quantitative modeling tool used for financial time series forecasting. The system is adaptive in its core as it learns the patterns and geometrical relationships defined by historical time series data points, which are unique for each individual stock, index, or another financial instrument.

2.3. Markettrak

Its stock market forecast system consists of two major parts: an extensive database and a forecast model. The forecast model reads the database and then makes a prediction of where the market is headed. From this prediction, it determines a trading position for the Dow Diamonds or the SP500 Spiders [5]. The database and forecast are updated daily at the close of trading.

It uses a neural network model in combination with a genetic algorithm to calculate the SP500 forecast. The calculations are somewhat complex but can be summarized by the following three procedural steps.

Step one: The genetic algorithm is used to find the optimum neural network structures and inputs. This calculation basically determines how the networks will be wired.

Step two: Using the information from the first step, a set of networks is initialized and then trained on about 75 percent of the market data (in-sample) in their database, which currently consists of about 7200 days of data. They use an evolutionary program to train the networks (i.e. to determine the neural network weights).

Step three: After training, the networks are rigorously tested on the remaining 25 percent of market data (out-of-sample). Networks that fail the test are discarded. Networks that pass the test are included in the library that they use to calculate the forecast. The number of neural networks currently in their library varies from day to day, but normally contains more than 400.

Input to the networks are technical and fundamental market data. The table below shows the types of data that are currently used by the model:

- Dow Jones Industrial Average closing value
- Dow Jones Industrial Average theoretical high value
- Dow Jones Industrial Average theoretical low value
- Dow Jones Transportation Average closing value
- Dow Jones Utility Average closing value
- New York Stock Exchange total volume
- New York Stock Exchange number of advancing stocks
- New York Stock Exchange number of declining stocks
- New York Stock Exchange number of new highs
- New York Stock Exchange number of new lows
- New York Stock Exchange advancing volume
- New York Stock Exchange declining volume
- SP500 closing value
- SP500 trailing earnings
- Yen-Dollar exchange rate
- Treasury bill discount rate
- Commodity Research Bureau index

The above data are filtered and normalized and certain functions of these data are computed. It currently computes 63 separate input variables at the close of each trading day.

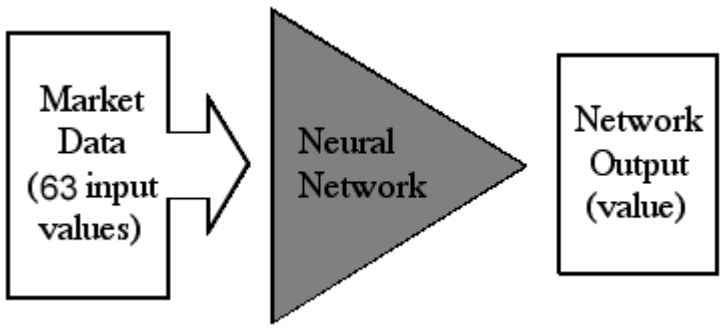


Figure 2.1: Forecasting method for Markettrak

The 63 inputs are applied to a neural network and after some number crunching the network outputs a value between -1.0 and +1.0, with -1.0 being a very strong down market signal and +1.0 being a very strong up market signal. A value near zero would indicate a neutral market signal. They apply the inputs to each network in our library and an average of their outputs is computed. This average network output is used with position set points to determine a trading position for the Dow Diamonds or the SP500 Spiders for the next trading day. When the computed value of the average network output is above the long position set point, a long position is indicated. When the value of the average network output is below the short position set point, a short position is indicated. When the average network output falls between these two set points, a cash position is indicated. Because of the timing of the update, trades may be made in the extended sessions or at open of the next trading day.

When computing our performance, all trades are assumed to take place at the session close. The current trading position along with recent average network output values and the average network output set points are shown on our forecast page.

2.4. Stock-Forecasting.com

www.stock-forecasting.com (Center of Mathematics & Science, Inc., Chicago, United States of America) provides innovative price-prediction technology for active Day Traders, Short- and Long-term Investors. They develop web-based software for stock market forecasting and analysis.

The artificial intelligence www.stock-forecasting.com software is based on neural network technology, advanced statistical methods and non-periodic stock price wave analysis. The Stock-Forecasting software predicts stock prices, generates trading "Buy-Hold-Sell" signals, computes the most profitable company to invest in and analyzes the accuracy of predictions.

3. REQUIREMENT ANALYSIS AND FEASIBILITY STUDY

3.1. Feasibility Study

Simply put, stock market cannot be accurately predicted. The future, like any complex problem, has far too many variables to be predicted. The stock market is a place where buyers and sellers converge. When there are more buyers than sellers, the price increases. When there are more sellers than buyers, the price decreases. So, there is a factor which causes people to buy and sell. It has more to do with emotion than logic. Because emotion is unpredictable, stock market movements will be unpredictable. It's futile to try to predict where markets are going. They are designed to be unpredictable.

There are some fundamental financial indicators by which a company's stock value can be estimated. Some of the indicators and factors are: Price-to-Earning (P/E) Ratio, Price-to-Earning Growth (PEG) Ratio, Price-to-Sales (P/S) Ratio, Price/Cash Flow (P/CF) Ratio, Price-to-Book Value (P/BV) Ratio and Debt-to-Equity Ratio. Some of the parameters are available and accessible on the web but all of them aren't. So we are confined to use the variables that are available to us.

The proposed system will not always produce accurate results since it does not account for the human behaviours. Factors like change in company's leadership, internal matters, strikes, protests, natural disasters, change in the authority cannot be taken into account for relating it to the change in Stock market by the machine.

The objective of the system is to give a approximate idea of where the stock market might be headed. It does not give a long term forecasting of a stock value. There are way too many reasons to acknowledge for the long term output of a current stock. Many things and parameters may affect it on the way due to which long term forecasting is just not feasible.

3.2. Requirement Analysis

After the extensive analysis of the problems in the system, we are familiarized with the requirement that the current system needs. The requirement that the system needs is categorized into the functional and non-functional requirements. These requirements are listed below:

3.2.1 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 share price.
- The system should collect accurate data from the NEPSE website in consistent manner.

3.2.2 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.

4. SYSTEM DESIGN AND ARCHITECTURE

4.1. Use Case Diagram

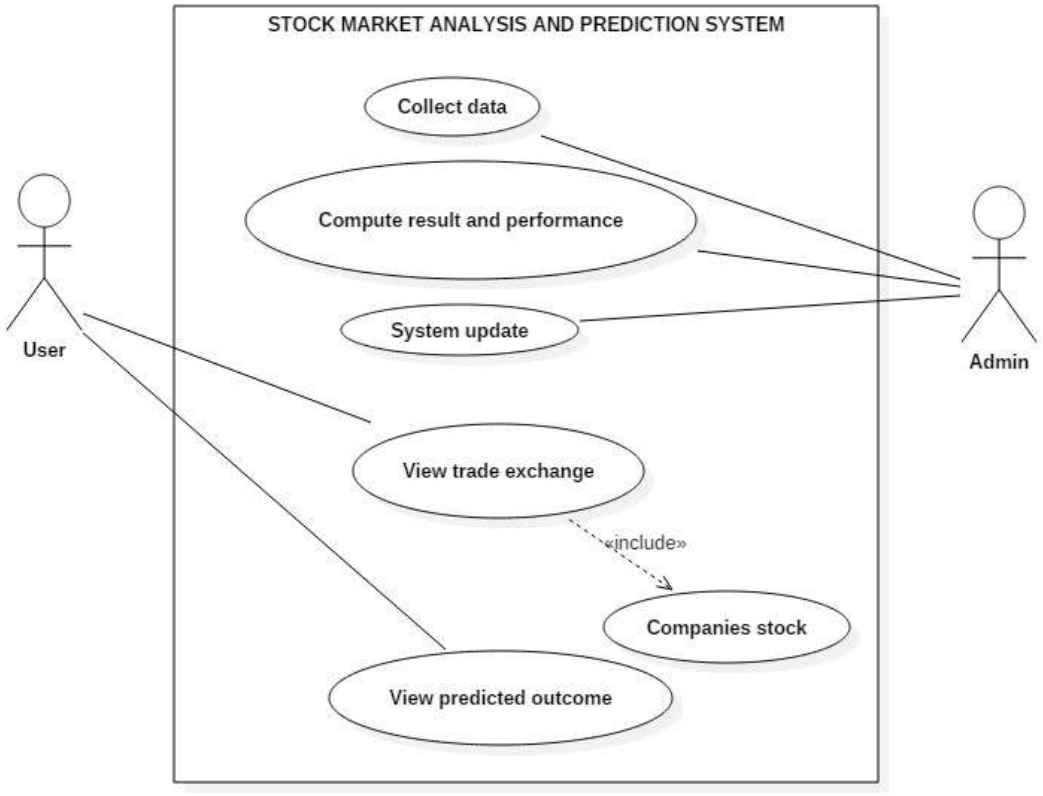


Figure 4.1: Use Case Diagram for the system

Use case index

Use case ID	Use case name	Primary actor	scope	complexity	priority
1	Collect data	admin	in	high	1
2	Compute result and prepare	admin	in	high	1
3	System update	admin	in	high	1
4	View trade exchange	user	in	medium	2
5	Company stock	user	in	medium	2
6	View predicted outcome	user	in	high	1

Use case description:

Use case ID:1

Use case name: Collect data

Description: Every required data will be available in Nepal stock exchange. 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 there the predict result of stock exchange and their actual price will be updated by admin in regular basis.

Use case ID: 4

Use case name: View traded exchange

Description: Company trading which is held at NEPSE can be viewed by user.

Use Case ID: 5

Use Case Name: Company Stock

Description: It is extended feature of view traded exchange. This includes the stock value of particular company.

Use Case ID: 6

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 hydropower companies. Thus, this will be available in user interface and viewer can observe them.

4.2. System Flow diagram

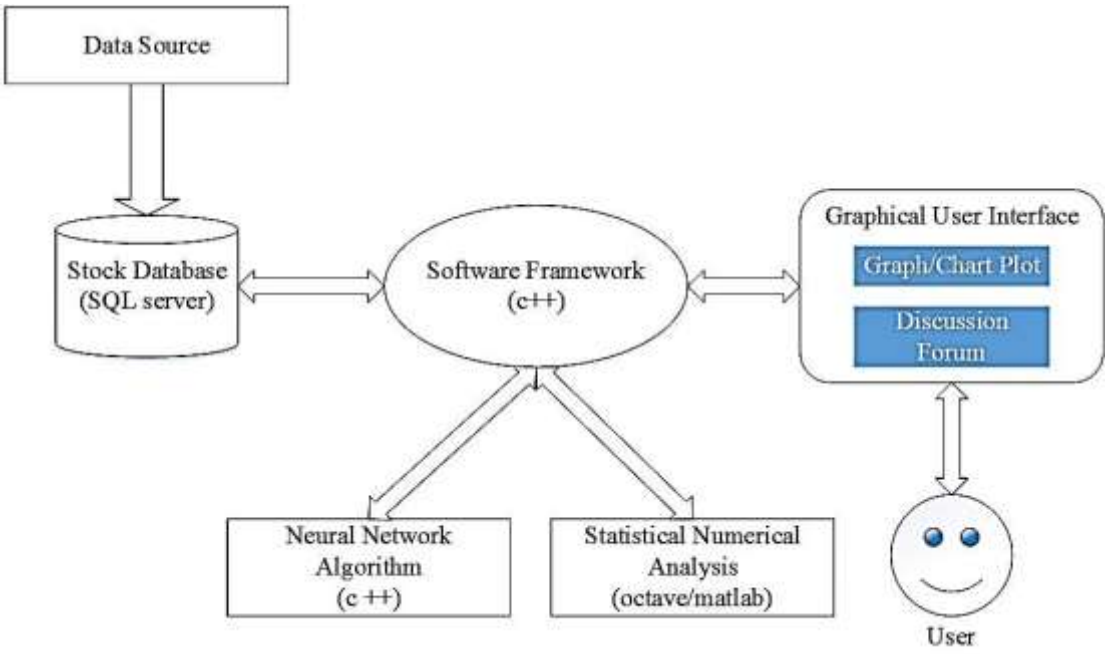


Figure 4.2: System Flow Diagram

Figure 4.3: Entity-Relation Diagram

5. METHODOLOGY

The purposed method for developing the system consists of mainly three main steps. Firstly, data is collected and sorted for relevancy from various sources. Secondly, analysis is carried out on the collected data by examining the current market direction, tracking the industry group and specific companies after which the data is represented and scored accordingly. At last, an ANN is designed and a suitable algorithm yielding best accuracy is chosen to predict the stock value.

5.1. Data Sources

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 Nepal Stock Exchange website (http://www.nepalstock.com.np) 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, increase or decrease in stock values for each financial companies. It also provides the overall performance of Nepal Stock Exchange and performance of companies of different categories. The site is updated on daily basis and it is also a repository for years of stock market data for Nepal.

There is no API provided by the website for providing data. We have written scripts to scrape all the required data from NEPSE website. Cron-jobs will be used to scrape data on regular basis.

Other sources include significant data from Open Data Nepal and relevant news related to finance and share market which might reflect the value and status of the listed companies.

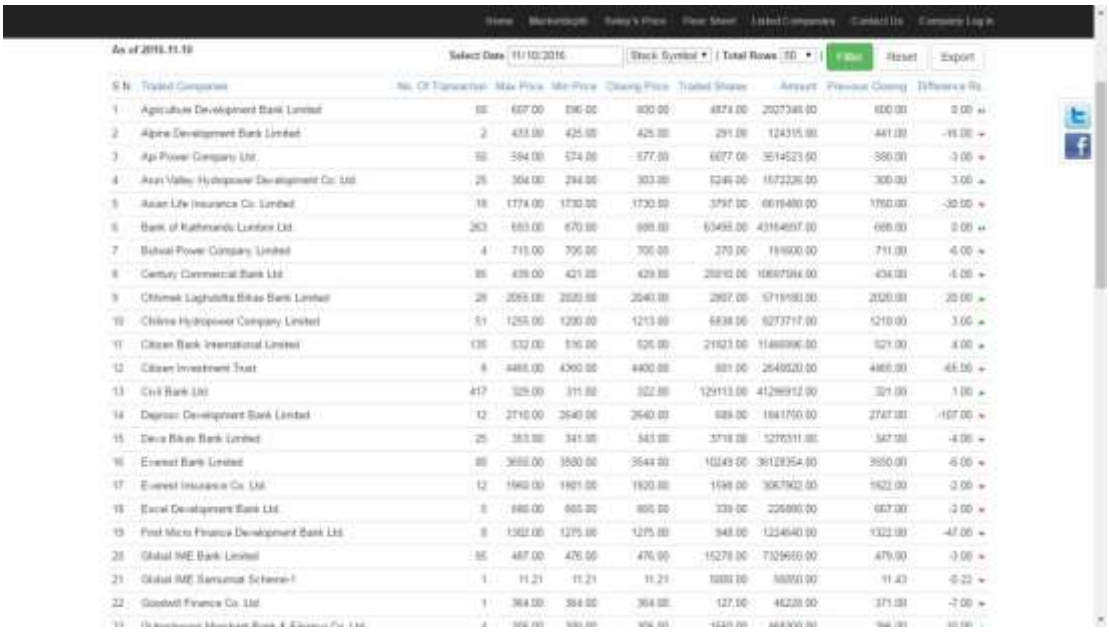


Figure 3.1: Nepal Stock Exchange website as of 2016-11-10

Other sources of data are banking and financial statistics published by Nepal Rastra Bank, annual report of different sample banks, supervision report of Nepal Rastra Bank and annual reports of concerned banks. In addition to these, different published articles, report, book, journal, and graduate research project will also analyzed.

5.2. 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 in NEPSE. So, this project performs analysis and prediction on only the companies that fall in the hydropower sector.

5.3. ANN 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 Backpropagation Algorithm is used for this training phase. These weights are used in prediction phase using same equations which are used in training phase.

5.3.1. Dataset Creation

First of all, a dataset is created for training the Artificial neural network. The collected data are arranged according to the format for the library we use for training. The dataset should be of exact format that FANN specifies. It includes number of training pair, number of input and number of output in the first line of the dataset file and data from the second line.

5.3.2. Data Normalization

The data is normalized before being input to the ANN. 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 minmax 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.

$$z = \frac{x - min(x)}{max(x) - min(x)}$$

5.3.3. Activation Function

We use log sigmoid function as the activation function at both hidden layer and output layer. A sigmoid function is a mathematical function having a characteristic "S"-shaped curve or sigmoid curve which is given by:

$$S(x) = \frac{1}{1 + e^{-x}} = \frac{e^x}{e^x + 1}.$$

It transforms linear inputs to nonlinear outputs. It bounds output to between 0 and 1 so that it can be interpreted as a probability and it also makes computation easier than arbitrary activation functions. The reason we choose this function as activation function is because it gives logistic neurons real-valued output that is a smooth and bounded function of their total input. It also has the added benefit of having nice derivatives which make learning the weights of a neural network easier.

5.3.4. Backpropagation with Feedforword 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

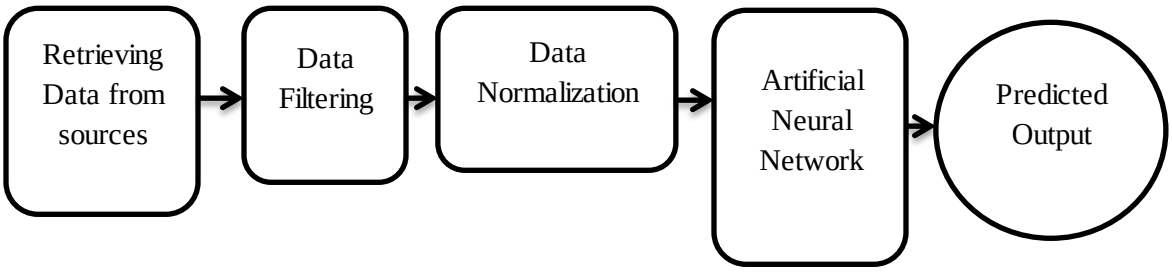


Figure 3.2: Prediction Process

5.4. 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 : 0.30

Learning rate : 0.30

Total Layers : 3

Input Neurons : 6

Hidden Layer Neurons : 4

Output Neurons : 1

Activation Function : Log Sigmoid

Limit of epochs : 50000

Minimum error : 0.000001

5.5. 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.

An input set like the one given below is given to the network.

```
0.2972972972973 0.0016617528812651 0.44117647058824 0.638720078125 0.28 0.292
0.2972972972973
```

The feed forward network then calculates an output according to the weights and bias of the individual neurons. The given input set generated an output as `0.6153274`

The output normalized data is then converted into normal form by using reverse minmax formula.

5.6. Model Design

We use feedforwad neural network which has a input layer with 6 neurons, a single hidden layer which has 4 neural neurons and a output layer with single neuron. The Backpropagagtion algorithm has been used for training the network.

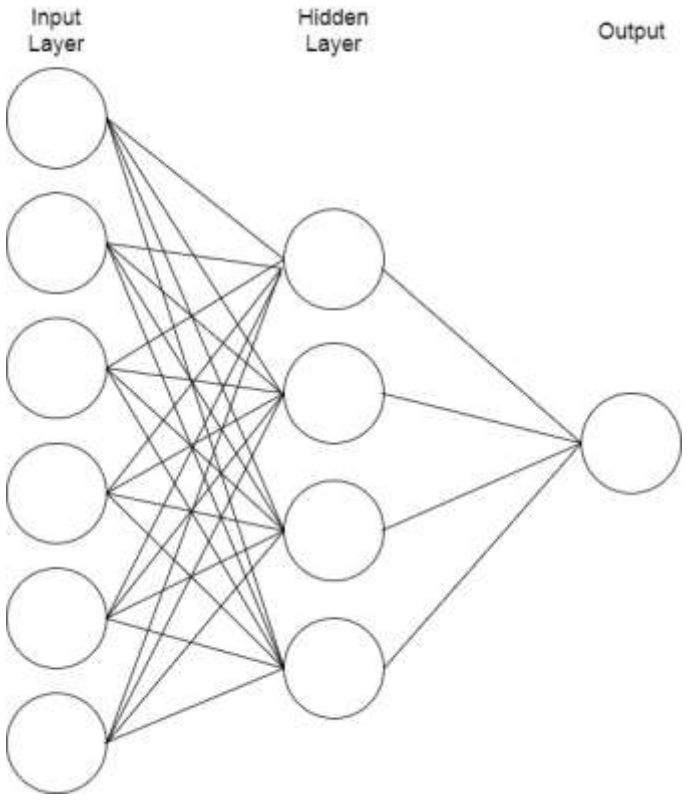


Figure 3.3: Neural Network

5.7. Input Data

Here is a brief description about the inputs that are fed to the neural network.

5.5.1. General Index (GI)

General index is a number that measure the relative value of a section of share market. It reflects the total economic condition of the market. If the general index goes down then it means the economic condition of that particular market is relatively in poor condition.

5.5.2. Price Difference

It is the difference between the current share value and previous share value of a specific company. The difference specifies whether the share price is increasing decreasing or equal throughout the day.

5.5.3. Highest Value

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

5.5.4. Lowest Value

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

5.5.5. 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.

5.5.6 Closing price

Closing price generally refers to the last price at which a stock trades during a regular trading session. For NEPSE , regular trading sessions run from Sunday - Thursday (11:00 AM . to 3:00 PM) and Friday(12:00 pm to 1:00 pm).

5.8. Programming tools and External Dependencies

5.8.1. SQL

Structured Query Language is a special-purpose programming language designed for managing data held in a relational database management system (RDBMS), or for stream processing in a relational data stream management system (RDBMS).

5.8.2. C++

We intend to use C++ language to design Neural Network since it has very good performance. C++ is a general-purpose programming language. It has imperative, object-oriented and generic programming features, while also providing facilities for low-level memory manipulation.

5.8.3. Fast Artificial Neural Network

Fast Artificial Neural Network Library is a free open source neural network library, which implements multilayer artificial neural networks in C with support for both fully connected and sparsely connected networks. Cross-platform execution in both fixed and floating point are supported. It includes a framework for easy handling of training data sets. It is easy to use, versatile, well documented, and fast. Several graphical user interfaces are also available for the library.

5.8.4. FusionChart

It is a plugin containing javascript and CSS files that assists in rendering line charts and other graphs in the analysis part of the project.

5.8.5. HTML

We use HTML for rendering the analyzed and forecasted data in the web page. HTML is the standard markup language for creating Web pages. HTML describes the structure of Web pages using markup. HTML elements are the building blocks of HTML pages.

5.8.6. JavaScript

We make the use of JavaScript to incorporate different plugins in our web page. JavaScript is the programming language of HTML and the Web. JavaScript resides inside HTML documents, and can provide levels of interactivity to web pages that are not achievable with simple HTML.

5.8.7. PHP

We use PHP to interact with the database since our data are stored in the database. PHP is a server-side scripting language designed primarily for web development but also used as a general-purpose programming language. It allows web developers to create dynamic content that interacts with databases.

5.6.9. NeuroXL

The NeuroXL is a neural network tool that can be used for simulation of Neural Network. Simulation using this tool can help us in selecting the right parameters for neural network such as maximum epoch, minimum error, initial weights and activation function.

6. TESTING

6.1. 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.

6.2. 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, behavior 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.

6.3. 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.

7. ANALYSIS AND RESULTS

7.1. Analysis

In this project, the factors that are taken into account for change in the closing price of a particular company are: General Index, Price difference, 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.

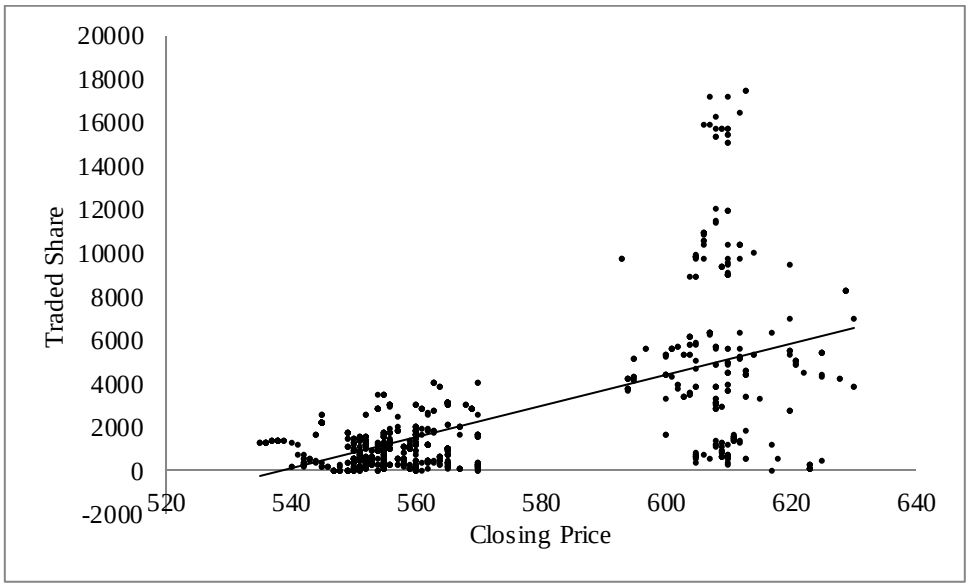


Figure 7.1: Scatter Plot – Closing Price and Traded Share

Figure 7.2: Scatter Plot – Closing Price and Difference

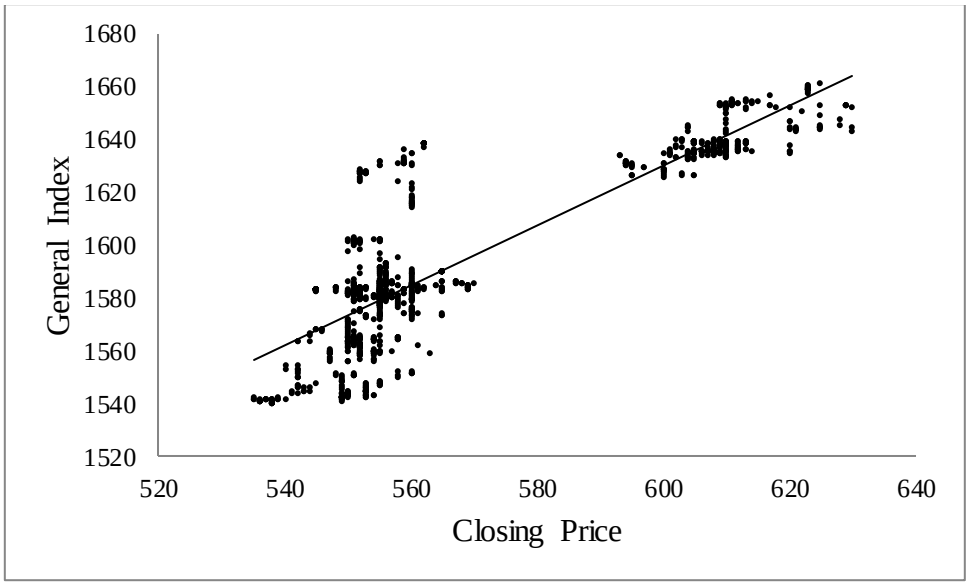


Figure 7.3: Scatter Plot – Closing Price and General Index

The above figure shows the data of API Hydropower Company. A total of 1520 data were used for calculation of correlation between Closing Price and Traded Share and Closing Price and Difference. Also a total of 1319 data were used for calculation of correlation between Closing Price and General Index. The coefficient of correlation was found to be 0.615437 for Closing Price and Traded Share and 0.145055 for Closing Price and Difference. Similarly, the coefficient of correlation was found to be 0.86695 for Closing Price and General Index.

7.2. Result

After collection of data the future share price is predicted using neural network. The value is then compared the next day with the actual value. The results and deviations of three random companies namely API, BARUN and KKHC are illustrated in graphical form below.

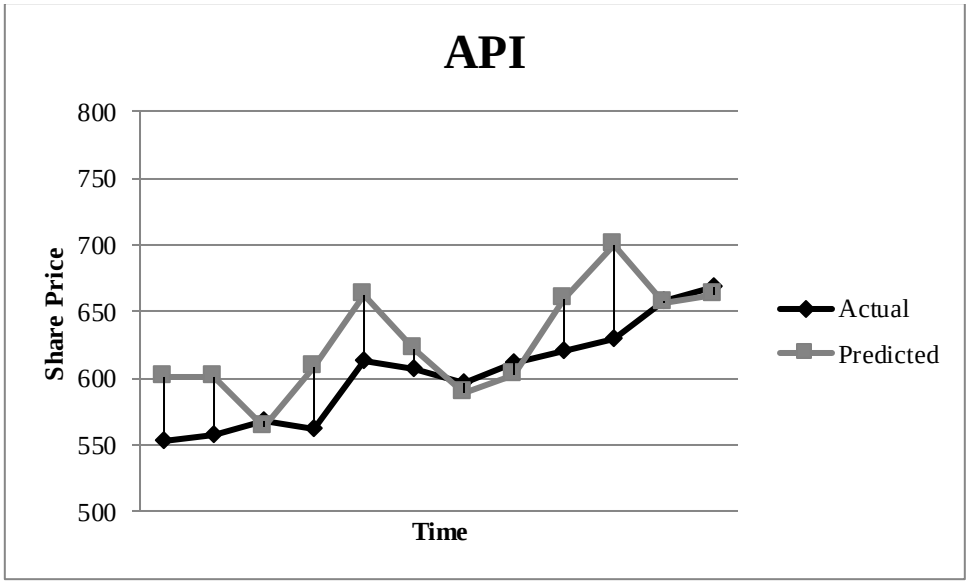


Figure 7.4: Actual vs Predicted – API

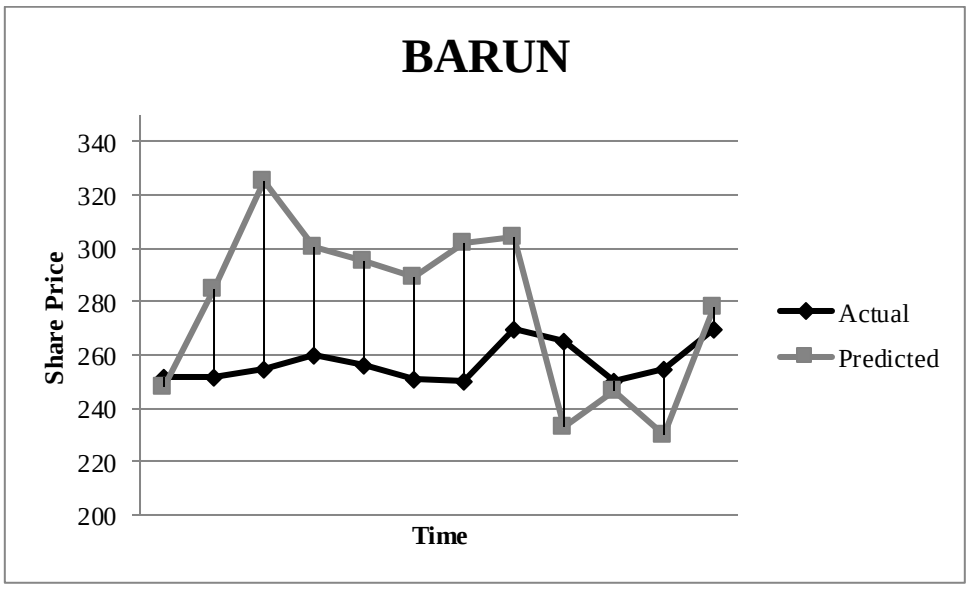


Figure 7.5: Actual vs Predicted – BARUN

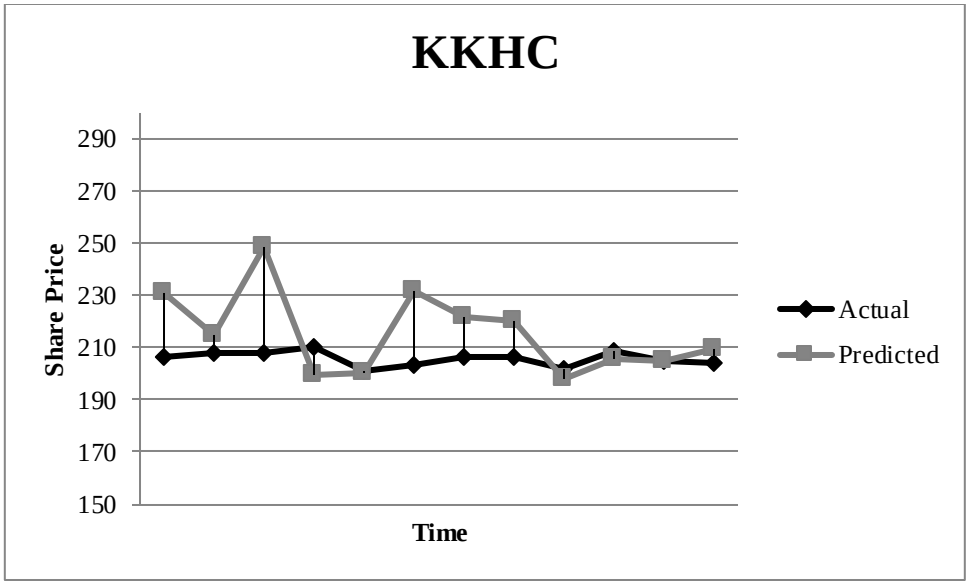


Figure 7.6: Actual vs Predicted – KKHC

8. LIMITATION AND FUTURE ENHANCEMENT

We implement the application of Artificial Neural Network to the task of stock market prediction and ANN model and salient feature. Our initial analysis show significant correlation between different input parameter.

The result obtained in both the cases was fairly accurate. As is indent from fig 7.1, 7.2, 7.3, the prediction is fairly accurate unless there is huge and sudden variation in the actual data. On other hand, this also proves the hypotheses that stock market are actually unpredictable.

After the phase of prediction and analysis, the result will be displayed to users in the form web pages.

8.1. Limitation

The main aim of this system is to provide a general idea of where the stock market is headed. It is only limited to a very basic prediction model. Thus, it cannot be used as a critical decision making tool. By incorporating only limited number of parameters, there is certain degree of accuracy. Since, there are many indeterminate parameters that directly affect stock market, each and every one of them cannot be taken into account. So, our model only depends on the relationship of our selected parameters with the share price.

At present, this system only performs analysis and prediction of hydropower sector companies of Nepal This system is limited to only certain users that have knowledge of stock market.

8.2. Future scope of improvement

- Potential improvement can be made to our data collection and analysis method.

- Future research can be done with possible improvement such as more refined data and more accurate algorithm.

- Implementation of discussion forums and economic news portal including other sector apart from hydropower and going in national level.

CONCLUSION

We implement the application of Artificial Neural Network to the task of stock market prediction and ANN model and salient feature. Our initial analysis show significant correlation between different input parameter.

The result obtained in both the cases was fairly accurate. As is indent from fig 7.1, 7.2, 7.3, the prediction is fairly accurate unless there is huge and sudden variation in the actual data. On other hand, this also proves the hypotheses that stock market are actually unpredictable.

After the phase of prediction and analysis, the result will be displayed to users in the form web pages.

REFERENCES

[1]R. Wilson and R. Sharda, "Bankruptcy prediction using neural networks", *Decision Support Systems*, vol. 11, no. 5, pp. 545-557, 1994.

[2]Atsalakis, G. S., Dimitrakakis, E. M., & Zopounidis, C. D. (2011). Elliott wave theory and neuro-fuzzy systems, in stock market prediction: The WASP system. *Expert Systems with Applications, 38*, 9196–9206

[3]K. Tsai and J. Wang, "External technology sourcing and innovation performance in LMT sectors", *Research Policy*, vol. 38, no. 3, pp. 518-526, 2009.

[4]K. Han and J. Kim, "Genetic quantum algorithm and its application to combinatorial optimization problem", *Evolutionary Computation, 2000*, vol. 2, pp. 1354-1360, 2000.

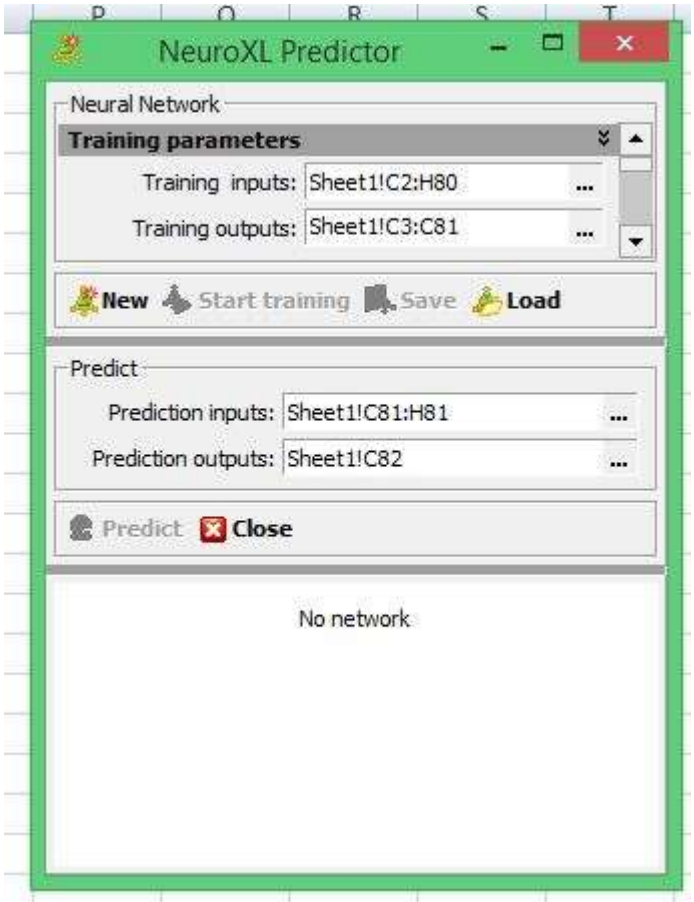
[5]"Market Capitalisation Definition from Financial Times Lexicon", *Lexicon.ft.com*, 2016. [Online]. Available: http://lexicon.ft.com/Term?term=market-capitalisation. [Accessed: 10- Nov- 2016].

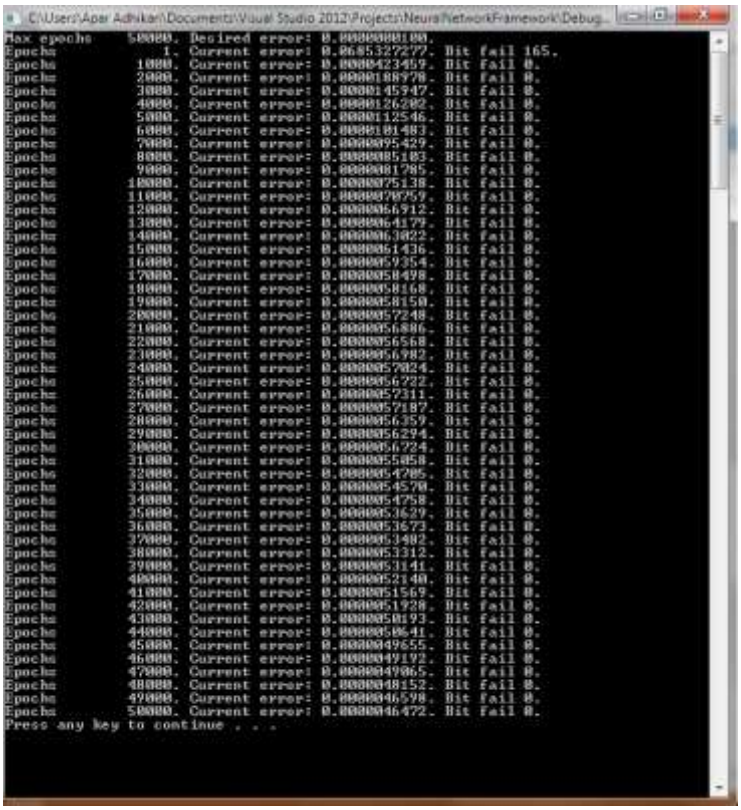
[6]"MarketTrak's Forecast Model Overview", *Markettrak.com*, 2016. [Online]. Available: http://www.markettrak.com/about.html. [Accessed: 25- Nov- 2016].

[7]S. Hannon, "5 Rules For Prediciting Stock Market Trends - StockTrader.com", *StockTrader.com*, 2016. [Online]. Available:https://www.stocktrader.com/2009/05/04/predicting-stock-market-trends-rules/. [Accessed: 16- Nov- 2016].

APPENDIX


STOCK MARKET
Home
Analysis
Forecasting
Today's Market
Company Price Change
AHPC 274 -1
AKPL 410 -1
API 655 -14
BPCL 666 -2
CHCL 805 -6
DHPL 200 -1
KKHC 204 0
NGPL 205 1
NHPC 145 0
RHPC 205 -3
160
100
40
0
What Is The Cost Of An
Investment Mistake
CORRECT MISTAKE
NEPAL STOCK EXCHANGE LTD
NEPSE


NeuroXL Predictor
Neural Network
Training parameters
Training inputs: Sheet1!C2:H80
Training outputs: Sheet1!C3:C81
New
Start training
Save
Load
Predict
Prediction inputs: Sheet1!C81:H81
Prediction outputs: Sheet1!C82
Predict
Close
No network


C:\Users\Apar Adhikari\Documents\Visual Studio 2012\Projects\NeuralNetworkFramework\Debug...
Max epochs 50000. Desired error: 0.0000000100.
Epochs 1. Current error: 0.8655322272. Bit fail 165.
Epochs 1000. Current error: 0.0000423459. Bit fail 0.
Epochs 2000. Current error: 0.0000188978. Bit fail 0.
Epochs 3000. Current error: 0.0000145947. Bit fail 0.
Epochs 4000. Current error: 0.0000126202. Bit fail 0.
Epochs 5000. Current error: 0.0000112546. Bit fail 0.
Epochs 6000. Current error: 0.0000101483. Bit fail 0.
Epochs 7000. Current error: 0.0000095429. Bit fail 0.
Epochs 8000. Current error: 0.0000085103. Bit fail 0.
Epochs 9000. Current error: 0.0000081785. Bit fail 0.
Epochs 10000. Current error: 0.0000075138. Bit fail 0.
Epochs 11000. Current error: 0.0000070959. Bit fail 0.
Epochs 12000. Current error: 0.0000066912. Bit fail 0.
Epochs 13000. Current error: 0.0000064179. Bit fail 0.
Epochs 14000. Current error: 0.0000063022. Bit fail 0.
Epochs 15000. Current error: 0.0000061436. Bit fail 0.
Epochs 16000. Current error: 0.0000059354. Bit fail 0.
Epochs 17000. Current error: 0.0000058498. Bit fail 0.
Epochs 18000. Current error: 0.0000058168. Bit fail 0.
Epochs 19000. Current error: 0.0000058150. Bit fail 0.
Epochs 20000. Current error: 0.0000057248. Bit fail 0.
Epochs 21000. Current error: 0.0000056886. Bit fail 0.
Epochs 22000. Current error: 0.0000056568. Bit fail 0.
Epochs 23000. Current error: 0.0000056982. Bit fail 0.
Epochs 24000. Current error: 0.0000057024. Bit fail 0.
Epochs 25000. Current error: 0.0000056722. Bit fail 0.
Epochs 26000. Current error: 0.0000057311. Bit fail 0.
Epochs 27000. Current error: 0.0000057187. Bit fail 0.
Epochs 28000. Current error: 0.0000056359. Bit fail 0.
Epochs 29000. Current error: 0.0000056294. Bit fail 0.
Epochs 30000. Current error: 0.0000056224. Bit fail 0.
Epochs 31000. Current error: 0.0000055058. Bit fail 0.
Epochs 32000. Current error: 0.0000054705. Bit fail 0.
Epochs 33000. Current error: 0.0000054570. Bit fail 0.
Epochs 34000. Current error: 0.0000054758. Bit fail 0.
Epochs 35000. Current error: 0.0000053629. Bit fail 0.
Epochs 36000. Current error: 0.0000053673. Bit fail 0.
Epochs 37000. Current error: 0.0000053402. Bit fail 0.
Epochs 38000. Current error: 0.0000053312. Bit fail 0.
Epochs 39000. Current error: 0.0000053141. Bit fail 0.
Epochs 40000. Current error: 0.0000052140. Bit fail 0.
Epochs 41000. Current error: 0.0000051569. Bit fail 0.
Epochs 42000. Current error: 0.0000051928. Bit fail 0.
Epochs 43000. Current error: 0.0000050193. Bit fail 0.
Epochs 44000. Current error: 0.0000050641. Bit fail 0.
Epochs 45000. Current error: 0.0000049655. Bit fail 0.
Epochs 46000. Current error: 0.0000049192. Bit fail 0.
Epochs 47000. Current error: 0.0000049065. Bit fail 0.
Epochs 48000. Current error: 0.0000048152. Bit fail 0.
Epochs 49000. Current error: 0.0000046598. Bit fail 0.
Epochs 50000. Current error: 0.0000046472. Bit fail 0.
Press any key to continue . . .


STOCK MARKET
Home
Analysis
Forecasting
Today's Market
Company Price Change
AHPC 274 -1
AKPL 410 -1
API 655 -14
BPCL 666 -2
CHCL 805 -8
DHPL 200 -1
KKHC 204 0
NGPL 285 1
NHPC 145 0
RHPC 205 -3
SPDL 245 0
UMHL 349 8
SELECT REQUIRED HYDROPOWER COMPANY
Api Power company Ltd
Show-Chart
API
2017-08-03
Current-Price
Date

Today's Market

Company	Price	Change
AHPC	274	-1 ▼
AKPL	410	-1 ▼
API	655	-14 ▼
BPCL	668	-2 ▼
CHCL	805	-6 ▼
DHPL	200	-1 ▼
KKHC	204	0 –
NGPL	285	1 ▲
NHPC	145	0 –
RHPC	205	-3 ▼
SPDL	245	0 –
UMHL	349	8 ▲

SEARCH PREDICTED RESULT : yyyy-mm-dd Show result

S.N	Company	Predicted Closing (2017-08-02)	Actual
1	API	663.03	669
2	AKPL	401.33	411
3	AHPC	270.1	275
4	BARUN	278.24	270
5	BPCL	720.39	668
6	CHCL	846.18	811
7	DHPL	231.21	201
8	KKHC	209.54	204
9	NGPL	251.18	284
10	NHPC	139.11	145
11	SPDL	280.17	245
12	RHPC	216.93	208
13	UMHL	328.34	341