Acknowledgements

We would sincerely like to thank our project guide Dr. Prachi Gharpure for giving us an opportunity to work under her guidance. Not only has she lent her support in her area of expertise, she has taken up the daunted task of going out of her area of expertise, thus making it an enriching and learning experience for all of us.

It is only right to express our sincere gratitude to the BE Computer batch of 2014, our classmates, who have helped and supported us in the development and successful completion of our project. The non-teaching staff of our department, Mr. Nandkumar Kinnare and Mr. Nalawade, also deserves a special mention for helping us with any difficulty we have faced with respect to the college resources. We would also like to thank a few seniors who have been nice enough to guide us in spite of their watertight working schedules.

Last but not the least, we would like to thank the open source community whose efforts and works have inspired us to working harder and achieving better.

Abstract

The Indian stock market has a highly stochastic nature which requires a highly sophisticated model for forecasting stock prices. Neural networks, as artificial intelligence (AI) methods, have application in many fields and can be applied to predict the future of the stock prices based on the past data. This report explains the proposed model to predict the share price using Artificial Neural Network with given input parameters of share markets. The proposed artificial neural network model is not designed to predict the exact values of the stocks rather it is designed to predict whether the stock price for any particular stock will increase or decrease for the following day.

In order to predict this rise or fall in stock prices we rely on the data provided by NSE India which is then processed. Previous research about Stock Market prediction using various methods has been studied before designing the model. Our aim is to design a model to predict stock prices with a better accuracy than the available algorithms.

Introduction

1.1 Motivation

Indian stock market has a highly stochastic nature which requires a highly sophisticated model for forecasting stock prices. Neural networks, as artificial intelligence (AI) methods, have application in many fields and can be applied to predict the future of the stock prices based on the past data. This report explains the proposed model to predict the share price using Artificial Neural Network with given input parameters of share markets. The proposed artificial neural network model is not designed to predict the exact values of the stocks rather it is designed to predict whether the stock price for any particular stock will increase or decrease for the following day.

The prediction algorithms coupled with its application in the field of Stock Market are our area of interests. The fluctuation and trends in stock market could be determined using these prediction techniques. The knowledge that Artificial Neural Networks, studied in academic course 'Soft Computing', can assist in predicting the market trends developed an interest amongst us to develop a particular model that would help predict market trends of National Stock Exchange (NSE) of India. Thus, an attempt is made to contribute to the area of Stock Market Prediction.

1.2 Aims and Objectives

The aims and objectives of this project can be stated as follows:

1. To create an application which allows the user to select a stock and a date and the application will predict whether the stock price for that date will rise or fall.

2. To understand the concept of Artificial Neural Network and Data mining Classification algorithms.

3. To understand the working and subtleties of Stock market of India.

4. To research and study existing Stock market prediction models.

5. Devise and develop a model to predict the trends in NSE (National Stock Exchange) of India.

1.3 Problem Definition

A stock market facilitates exchange of stocks between the buyers and the sellers. A stock, simply being a share in ownership of a company, represents the owner's claim on the company's assets, earnings, profits as well as loss. There has been a great amount of research done to predict the stock prices of the companies listed on different stock exchanges. However, the stock market trend depends on various factors, namely, political events, economic condition, financial reports and general news. The mentality, attitude and reaction to these factors are different for investors in different nations. Hence, a successful prediction model for a stock exchange in a particular nation might not be able to successfully predict for a stock exchange of another nation. Hence, it becomes important to develop a prediction model that will accurately predict the stock prices of the companies listed on the National Stock Exchange (NSE) of India. The Indian stock market has a highly stochastic nature which requires a sophisticated model for forecasting stock prices. We study various existing prediction models and their consistencies with the stocks listed on NSE of India. Also, understanding the fact that prediction of exact values of stocks is hypothetical, we confine our studies to predict if the stock price for any particular stock listed on NSE of India will increase or decrease for the following day. For the investors investing in any of the company listed on the National Stock Exchange of India, such model can prove to be a very helpful tool to take the right decision regarding their stocks.

1.4 Topic Organization //to be updated at last

The Thesis is organized as follows chapter 2 describes the literature surveyed from various papers, previous works and some websites. It explains Stock Marketand various classification algorithms studied.

Chapter 2

Literature Review

There has been a great amount of research done to predict the stock prices of the companies listed on different stock exchanges. However, the stock market trends depends on various factors, namely, political events, economic condition, financial reports and general news. The mentality, attitude and reaction to these factors are different for investors in different nations. Hence, a successful prediction model for a stock exchange in a particular nation might not be able to successfully predict for a stock exchange of another nation. Hence, it becomes important to develop a prediction model that works for a particular stock exchange and so we propose a model that will accurately predict the stock prices of the companies listed on the National Stock Exchange (NSE) of India.

The research by [1] used data mining methods and neural networks for predicting stock market trends. An attempt has been made in this study to investigate the predictive power of financial and economic variables by adopting the variable relevance analysis technique in machine learning for data mining. The authors examined the effectiveness of the neural network models used for level estimation and classification. It was observed from the results that the stock buying or selling strategies recommended by the neural network classification models had higher accuracy over those recommendations that were suggested by other strategies.

The research by [2] aimed to aid the investors in the stock market to decide the best timing for buying or selling stocks. Id3 classification algorithm was used for making the decision from the knowledge extracted from the historical prices of stocks listed in Amman Stock Exchange (ASE).The research was able to provide moderate accuracy in predicting the stock prices. It was observed from the results that the proposed model had less efficiency as it did not take into account the political events, economic conditions and general news that have a great influence on the stock market.

As explained in [4], classification learning operates under supervision. Learning is done by providing the actual outcome for each of the training examples. The success of classification learning can be judged by trying out the concept description that is learned on an independent set of test data, as stated in [4].

J48 is an open source Java implementation of the C4.5 decision tree algorithm. That being said the algorithm used in the previously surveyed literature is C4.5 but this model uses J48 for sake of convenience.

2.1 Stock Market

A stock market is the aggregation of buyers and sellers of stocks. These may include the securities listed on a stock exchange as well as those only traded privately. A share of stock is literally a share in the ownership of a company.
A company wants to share its assets and earnings with the general public because it needs the money to function. Companies only have two ways to raise money to cover start-up costs or expand the business: It can either borrow money (a process known as debt financing) or sell stock (also known as equity financing). If a company issue a lot of shares, that would lower the price of each individual share, perhaps making the stock more attractive to lone investors. Each person who buys a share of stock essentially owns a piece of the company and has a say in how the company is run.

A business that wants to sell shares of stock to private or public investors needs to become a corporation first. The legal process of turning a business into a corporation is called incorporation. A corporation is registered with the government, has its own Social Security number, can own property, and make contracts. By definition, a corporation has stock that can be bought and sold; all of the owners of the corporation hold shares of stock in the corporation to represent their ownership.

The buying and selling of stocks is concentrated in one place, and since it's all done electronically, we can track the constantly fluctuating price of a stock in real time. Investors can watch, for example, how a stock's price reacts to news from the company, media reports, national economic news and lots of other factors. A stock market index is a measurement of the value of a section of the stock market. It is computed from the prices of selected stocks . It is a tool used by investors and financial managers to describe the market, and to compare the return on specific investments.

Indian Stock Exchange refers to:

1. National Stock Exchange of India
    a. It is located at Mumbai, Maharashtra, India and was established in the mid-1990s.
    b. CNX Nifty is the stock market index of NSE of India
2. Bombay Stock Exchange

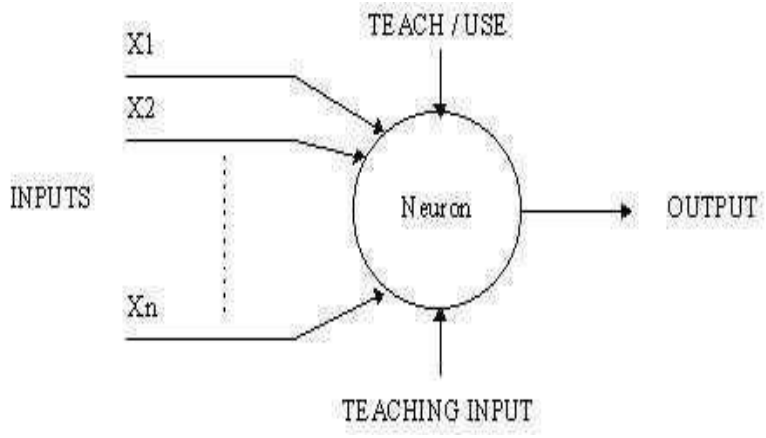
a. It was established in 1875.

S&P BSE SENSEX is the stock market index of BSE.

2.2 Artificial Neural Network

An Artificial Neural Network (ANN) is an information processing model that is inspired by the way biological nervous systems, such as the brain, process information. It is composed of a large number of highly interconnected processing elements (neurons) working in union to solve specific problems. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process.

An artificial neuron is a device with many inputs and one output. The neuron has two modes of operation; the training mode and the testing mode. In the training mode, the neuron can be trained to fire (or not), for particular input patterns. In the testing mode, when a taught input pattern is detected at the input, its associated output becomes the current output. If the input pattern does not belong in the taught list of input patterns, the firing rule is used to determine whether to fire or not.



The features of ANN are as follows:
1) Adaptive learning: An ability to learn how to do tasks based on the data given for training or initial experience.
2) Self-Organization: An ANN can create its own organization or representation of the information it receives during learning time.
3)Real Time Operation: ANN computations may be carried out in parallel, and special hardware devices are being designed and manufactured which take advantage of this capability.
4) Fault Tolerance via Redundant Information Coding: Partial destruction of a network leads to the corresponding degradation of performance. However, some network capabilities may be retained even with major network damage.

Multi-Perceptron Neural Network

A multilayer perceptron (MLP) is a feedforward artificial neural network model that maps sets of input data onto a set of appropriate outputs. A MLP consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one. Except for the input nodes, each node is a neuron (or processing element) with a nonlinear activation function. MLP utilizes a supervised learning technique called backpropagation for training the network. MLP is a modification of the standard linear perceptron and can distinguish data that are not linearly separable.

If a multilayer perceptron has a linear activation function in all neurons, that is, a linear function that maps the weighted inputs to the output of each neuron, then it is easily proved with linear algebra that any number of layers can be reduced to the standard two-layer input-output model (see perceptron). What makes a multilayer perceptron different is that some neurons use a *nonlinear* activation function which was developed to model the frequency of action potentials, or firing, of biological neurons in the brain. This function is modeled in several ways.

The two main activation functions used in current applications are both sigmoids, and are described by

$$y(v_i) = \tanh(v_i) \quad \text{and} \quad y(v_i) = (1 + e^{-v_i})^{-1},$$

in which the former function is a hyperbolic tangent which ranges from -1 to 1, and the latter, the logistic function, is similar in shape but ranges from 0 to 1. Here y_i is the output of the ith node (neuron) and v_i is the weighted sum of the input synapses. Alternative activation functions have been proposed, including the rectifier and softplus functions. More specialized activation functions includeradial basis functions which are used in another class of supervised neural network models.

2.3 Classification Algorithms

Classification problem is designing the mapping function m in such a way that each tuple of Database D belongs to a class C. A classification problem predicts categorical class labels and classifies data based on the training set and values in a classifying attribute and uses it in classifying new data. The typical applications of Classification algorithms include medical diagnosis, treatment effective analysis and credit approval.
There are three different Classification techniques:

1. Regression

   Regression analysis is a statistical process for modeling and analyzing several variables, when the focus is on the relationship between a dependent variable and one or more independent variables. Thus, regression analysis explains how the typical value of the dependent variable changes when any one of the independent variables is varied, while the other independent variables are held fixed.

2. Decision trees

   A decision tree is a tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. A decision tree is a flowchart-like structure in which internal node represents a "test" on an attribute, each branch represents the outcome of the test and each leaf node represents a class label. The paths from root to leaf represents classification rules.

3. Rules

The decision tree can be linearized into decision rules, where the outcome is the contents of the leaf node, and the conditions along the path form a conjunction in the if clause. In general, the rules have the form: if condition1 and condition2 and condition3 then outcome.

Classification algorithms for Stock Market Prediction

Data mining refers to extracting or mining knowledge from large data stores or sets. Stock market prediction refers to identifying the patterns in the stock prices of a particular company. Thus, data mining could be effectively used in stock market prediction. Data classification can be done in many different methods; one of those methods is the classification by using Decision Tree and Rules. Decision trees and artificial neural networks can be trained by using an appropriate learning algorithm. Thus, data mining classification algorithms could be effectively used in forecasting the trends in stock market.

2.3.a Adaboost

AdaBoost, short for "Adaptive Boosting", is a machine learning meta-algorithm The output of the other learning algorithms ('weak learners') is combined into a weighted sum that represents the final output of the boosted classifier. AdaBoost is adaptive in the sense that subsequent weak learners are tweaked in favor of those instances misclassified by previous classifiers. AdaBoost is sensitive to noisy data and outliers. While every learning algorithm will tend to suit some problem types better than others, and will typically have many different parameters and configurations to be adjusted before achieving optimal performance on a dataset, AdaBoost (with decision trees as the weak learners) is often referred to as the best out-of-the-box classifier. When used with decision tree learning, information gathered at each stage of the AdaBoost algorithm about the relative 'hardness' of each training sample is fed into the tree growing algorithm such that later trees tend to focus on harder to classify examples.

AdaBoost refers to a particular method of training a boosted classifier. A boost classifier is a classifier in the form

$$F_T(x) = \sum_{t=1}^{T} f_t(x)$$

where each f_t is a weak learner that takes an object x as input and returns a real valued result indicating the class of the object. The sign of the weak learner output identifies the predicted object class and the absolute value gives the confidence in that classification.

At each iteration of the training process, a weight is assigned to each sample in the training set equal to the current error $E(F_{t-1}(x_i))$ on that sample. These weights can be used to inform the training of the weak learner, for instance, decision trees can be grown that favor splitting sets of samples with high weights.

Applicability of AdaBoost for Stock Market Prediction:

AdaBoost algorithm is favorable for stock market prediction as it can handle both continuous and categorical variables. Stock markets attributes namely Closing price, opening price, volume traded, maximum price and minimum price are continuous changing values. Also, AdaBoost algorithm provides a clear indication of which fields are most important for prediction and classification. Thus, it suits best in predicting stock market prediction.

2.3.b Bayesian Algorithm

Bayes' Theorem is a means of quantifying uncertainty. Based on probability theory, the theorem defines a rule for refining a hypothesis by factoring in additional evidence and background information, and leads to a number representing the degree of probability that the hypothesis is true. Bayesian logic is a branch of logic applied to decision making and inferential statistics that deals with probability inference: i.e. using the knowledge of prior events to predict future events. Bayesian theorem is stated mathematically as the following simple form:

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}.$$

Applicability of Bayesian algorithm for Stock Market Prediction:

Bayesian algorithm combines prior knowledge and observed data by multiplying prior probability of a hypothesis with the probability of the hypothesis given with the training. Thus, the fluctuation in the stock prices are effectively accounted for, as the prior knowledge of stock price are crucial for predicting the increase or decrease in the stock price for the following day.

2.3.c J48 Algorithm:

J48 is a Java implementation of C4.5 algorithm. C4.5 is a well-known algorithm used to generate a decision trees. It is an extension of the ID3 algorithm used to overcome its disadvantages. The decision trees generated by the C4.5 algorithm can be used for classification, and for this reason, C4.5 is also referred to as a statistical classifier.

The C4.5 algorithm made a number of changes to improve ID3 algorithm. Some of these are:

- Handling training data with missing values of attributes
- Handling differing cost attributes
- Pruning the decision tree after its creation
- Handling attributes with discrete and continuous values

Applicability of J48 for Stock Market Prediction:

J48 provides an important feature by handling attributes with differing weights. The different stock market attributes namely, opening price, closing price, maximum price and minimum price have varying importance in prediction of stock market prices for the following day. Thus, J48 is an important classification algorithm to predict the stock markets trend. Also, J48 can handle training data with missing attributes values, aiding the prediction model when all the attributes of stock data is not available.

Chapter 3

Proposed Neural Network Classification Model

The classification algorithms discussed above alone have been found to predict stock market with about 40-45% certainty [1]. While this result might be better than some other proposed methods is still not good enough to predict the market. The AdaBoost, J48, Bayesian algorithm being based on sheer probability or learning of fixed data set it still cannot guess the fundamental changes in a company due to sudden changes in trends or and thus it fails to improve itself accordingly.

Neural networks on other had have been found to predict the stock markets with better probability and are able to learn from some changes in global economy which cannot be understood by classification algorithms. Neural networks can be helpful for fundamental analysis of the company that underlines the stock itself. This in turn helps in predicting the results after sudden changes in stock market.

Thus we propose the combination of above techniques. Here neural networks and classification algorithms are not in competition but complement each other. There are tasks are more suited to an algorithmic approach like arithmetic operations and tasks that are more suited to neural networks. This task requires systems that use a combination of the two approaches viz. classification algorithms and neural networks in order to perform at maximum efficiency.

The results from classification algorithms can be fed to the neurons. The neurons with their learning capabilities can be expected to weigh each input after learning from the past trends of stock market and thus assign right weight to the inputs or the results from classification algorithms in this case.

A. The collection of the data

The website www.nseindia.com contains the historical prices of companies listed on the National Stock Exchange of India from the year 1994.The website provides a detailed data on the stocks providing the values of the Date, Previous Close price, Open Price, High Price, Low Price, Last Price, Close Price, Average Price, Total Traded Quantity, Turnover Rupees in Lakhs, Deliverable Quantity and percentage Delivery Quantity to Traded Quantity.

The collected data had 11 attributes out of which the following 7 attributes were selected: Previous Close price, Open Price, High Price, Low Price, Close Price, and Total Traded Quantity as other factors were not significant in the prediction. The class attribute signifies if the opening price of the stock increased or decreased for the current day with respect to the closing price of the previous day and it is named as "Result".

Table 1: Attribute Description

Attribute	Description	Possible Values
Previous	Previous day close price of the stock	Positive, Negative, Equal
Open	Current day open price of the stock	Positive, Negative, Equal
Low	Current day minimum price of the stock	Positive, Negative, Equal
High	Current day maximum price of the stock	Positive, Negative, Equal
Last	Current day close price of the stock	Positive, Negative, Equal
Volume	Current day quantity of stocks traded	Positive, Negative, Equal
Result	Increase or decrease	Up, Down

A. Preparing the data

Initially, when the data were collected, the values of the attributes selected were continuous numeric values. Hence, data transformation was applied by generalizing data to a higher-level concept.

• If the values of the attributes open, high, low, last were greater than the value of attribute previous for the same trading day, the numeric values of the attributes were replaced by the value Positive.

• If the values of the attributes open, high, low, last were less than the value of attribute previous for the same trading day, the numeric values of the attributes were replaced by the value Negative.

• If the values of the attributes open, high, low, last were equal to the value of attribute previous for the same trading day, the numeric values of the attributes were replaced by the value Equal.

• If the number of the shares traded on the current day was greater than the number of the shares traded on the previous day, the numeric values of the attribute Volume was replaced by the value Positive.

• If the number of the shares traded on the current day was less than the number of the shares traded on the previous day, the numeric values of the attribute Volume was replaced by the value Negative.

If the number of the shares traded on the current day was equal to the number of the shares traded on the previous day, the numeric values of the attribute Volume was replaced by the value Equal.

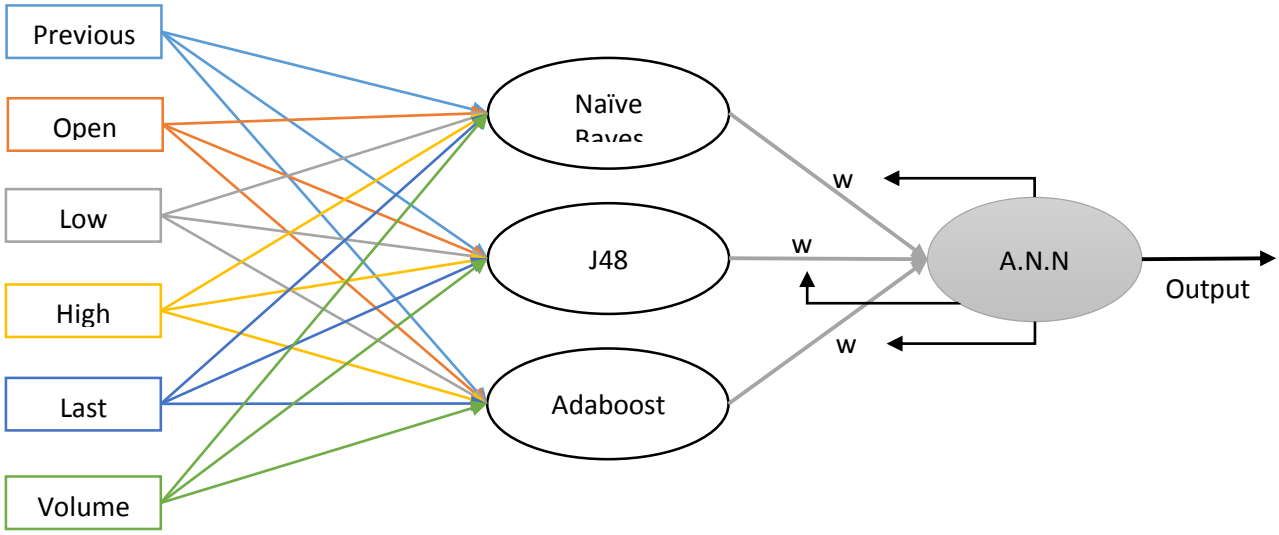
The value of the attribute Result was determined by comparing the values of the closing price of the previous day and the opening price of current day.

•If the opening price of current day was less than the closing price of the previous day, then the numeric value of the attribute Result was replaced by Down.

•If the opening price of current day was greater than or equal to the closing price of the previous day, then the numeric value of the attribute Result was replaced by Up.

B. Explanation of The Model

The following Neural network model is designed for effective prediction:



The three data mining classification algorithms used as neurons in the hidden layer are AdaBoost, J48 and Bayesian algorithm. Each of these algorithms predicts whether the opening price of a particular company will increase or decrease for the following day by providing the output as Up or Down. The weights given to outputs of each of these algorithms are updated using the Back Propagation Weight update rule. Final output is calculated by the weights attached to each of the algorithms and the answer predicted by the algorithms. The final output serves as the prediction whether the stock price will increase or decrease. Then, the actual result of the opening price of the company is compared to the individual outputs of the three algorithms. Based on the accuracy of these algorithms, the rules and the weights attached to the output of these algorithms are updated.

The designed neural network model for stock prediction aims to determine whether the stock price of a company will increase or decrease for the following day. The model is not designed to predict the exact value of the opening price of a stock for the following day. Hence, the proposed model has to identify if the stock price of a company will increase or decrease. For this reason, three data mining classification algorithms are used to predict the outcome of the intended result. The three data mining classification algorithms are Bayesian, AdaBoost and J48 decision trees. These classification algorithms are chosen for two prime reasons. Firstly, they provide the decision output in a format that matches the intended output, which means that the output of each of these algorithm will be either increase or decrease. Secondly, they have a high accuracy rate and take comparatively less computation time.

C. Training of the model

The transformed values of the seven attributes, namely previous, open, low, high, last, volume and result serve as the input parameters to the three classification algorithms for the training process. Initially, the training data that is provided to each of these algorithm is the historical data that is fetched from the database of National Stock Exchange Board of India. The algorithms are trained using this data and the corresponding rules are generated by each of these algorithms. The six attributes that are considered for generating rules are namely previous, open, low, high, last and volume and each of these attributes has three possible values namely positive, negative and equal. Hence, the maximum number of rules generated are 3^6. Thus, by this training process each classification algorithm generates its own set of rules. Depending on the rules generated, the classification algorithms determine the output result.

D. Working of the model

The transformed values of the six attributes, namely previous, open, low, high, last and volume serve as the input parameters to the three classification algorithms for the prediction process. The attributes' value of the current day determines whether the opening price of stock for the following day will increase or decrease relative to the closing price of the current day. Each of the three classification algorithms, namely AdaBoost, Bayesian and J48 generate an output result based on the rules generated by them. These output results serve as the input to the neural network. The results of each of the three classification algorithm are weighted and the weights are updated using the Back Propagation Weight update rule. The output and weight attached to each of the output are averaged and based on this calculation final decision is taken and provided to the user. Thus, the model predicts whether the opening price of a stock will increase or decrease for the following day.

On the following day, the actual result of the opening price of the company is compared with the predicted output of each of the three algorithms and depending on the accuracy of the algorithm the weights are updated using the Back Propagation Weight update rule. Not only the weights, but also the rules generated by the three algorithms are also updated i.e. if the existing rule generated by the algorithm proves false for the recent prediction then that rule is modified.

Chapter 4

Implementation

Chapter 5

Results and Discussions

The proposed neural network model was trained and tested using the stock data of the period 1st January 201 to 31st December 2014 for three companies viz. TCS, Bank of India, Infosys.

The comparison criteria used for the results are:

- Correctly Classified Instances: Abbreviated as 'CCI (%)'. It is a percentage of instances that have been predicted correctly by the algorithm.
- RMS Error: It is the Root Mean Squared Error for the algorithm.
- F-Measure: It is s measure of a test's accuracy.
- Precision: Precision is the fraction of retrieved instances that are acceptable or correctly predicted instances in this case.
- Recall: Recall is the fraction of relevant instances that are retrieved.

TABLE I. RESULTS FOR 'TCS'

ALGORITHMS	CCI (%)	RMS Error	F-Measure	Precision	Recall
Naïve Bayes	63.7584	0.3915	0.637	0.637	0.638
J48	68.4564	0.4596	0.682	0.719	0.685
Adaboost M1	74.4966	0.4172	0.745	0.759	0.745
Hybrid Model	81.2500	0.3853	0.816	0.825	0.813

TABLE II. RESULTS FOR 'BANK OF INDIA'

ALGORITHMS	CCI (%)	RMS Error	F-Measure	Precision	Recall
Naïve Bayes	75.1678	0.4122	0.741	0.778	0.752
J48	74.4966	0.4172	0.745	0.759	0.745
Adaboost M1	68.4564	0.4596	0.682	0.719	0.685
Hybrid Model	76.2500	0.4753	0.616	0.725	0.613

TABLE III. RESULTS FOR 'INFOSYS'

ALGORITHMS	**CCI (%)**	**RMS Error**	**F-Measure**	**Precision**	**Recall**
Naïve Bayes	70.3682	0.3124	0.659	0.689	0.734
J48	68.8889	0.4172	0.689	0.759	0.745
Adaboost M1	68.4564	0.4596	0.682	0.719	0.685
Hybrid Model	72.2500	0.3753	0.593	0.689	0.589

Thus the Hybrid model performs better than the individual classification algorithms.

//Expand the results

Chapter 10

Conclusions

This study presents a proposal to apply neural network and data mining and classification algorithms on the historical prices of the stocks to predict whether the stock price of any particular company will increase or decrease for the following day. For the investors investing in any of the company listed on the National Stock Exchange of India, it can prove to be a very helpful tool to take the right decision regarding their stocks.

The proposed model shows better results compared to that of other algorithms studied viz. Naïve Bayes, J48, and Adaboost. However, the results are not perfect. The reason being that many factors like economic conditions, political events, financial reports and general news that influence the stock market were not considered. Such factors can cause random and otherwise unpredictable changes in prices of stocks. These factors can be evaluated by adding a separate attribute "News" and providing inputs in integer form.

The model only focuses on the rise or fall of stock prices and ignores the magnitude of change. It can be upgraded to predict the magnitude of change during further work. The stochastic nature of stock market makes exact change prediction an unsurmountable problem. To simplify this task the magnitude could be broken up into blocks of percentages. Such prediction could be more useful in decision making.

References

1. Enke, D., Thawornwong, S. (2005) "The use of data mining and neural networks for forecasting stock market returns", Expert Systems with Applications, 29, pp. 927-940.

2. Qasem A. Al-Radaideh, Adel Abu, Assaf Eman Alnagi (2013) "Predicting Stock Prices using Data Mining Techiniques".

3. http://www.nseindia.com/products/content/equities/equities/eq_security.htm.

4. Witten, Ian H., and Eibe Frank. Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, 2005.

Acknowledgements

We would sincerely like to thank our project guide Dr. Prachi Gharpure for giving us an opportunity to work under her guidance. Not only has she lent her support in her area of expertise, she has taken up the daunted task of going out of her area of expertise, thus making it an enriching and learning experience for all of us.

It is only right to express our sincere gratitude to the BE Computer batch of 2014, our classmates, who have helped and supported us in the development and successful completion of our project. The non-teaching staff of our department, Mr. Nandkumar Kinnare and Mr. Nalawade, also deserves a special mention for helping us with any difficulty we have faced with respect to the college resources. We would also like to thank a few seniors who have been nice enough to guide us in spite of their watertight working schedules.

Last but not the least, we would like to thank the open source community whose efforts and works have inspired us to working harder and achieving better.