Using Dempster Shafer Theory to Evaluate the Relative Performance of Securities and Sector Wise Indices

A Dissertation proposal

Submitted By

KOMALWALIA

11000042

to

Department of Computer Science and Engineering

In partial fulfilment of the Requirement for the

Award of the Degree

of

Master of Technology in Computer Science

Under the guidance of

Mr. Virrat Devaser

UID: 14591

(Assistant Professor)

(May-2015)

ABSTRACT

The investment of money these days is more prevailing in stock markets, currency and commodity markets rather than in conventional banks. Though banking system is from many last years but the emergence of stock market has taken over its influence. People have started investing their money in stock market to get the maximum profits. Stock market is very uncertain, unpredictable and volatile in nature. Thus, this study relates to finding the probability whether we will receive better returns or not. As stock market is uncertain so number of techniques have been proposed in order to handle its uncertainty such as Neural Networks, Fuzzy Logic, Dempster Shafer Theory, Support Vector Machine, Regression analysis etc. with the help of which the behaviour of stock market can be predicted up to certain extent and also by making some assumptions. In this proposal, Dempster Shafer Theory will be applied to determine the uncertain behaviour of the Stock Market by using the parameters factors i.e. relative price to earnings ratio, market capitalization, profit factor and will find out the certainties in the prices, resulting in the better prediction of the individual stock.

CERTIFICATE

This is to certify that **Komal Walia**, registration number **11000042** has completed dissertation proposal titled "**Using Dempster Shafer Theory for Evaluating Relative Performance of Securities and Sector Wise Indices**" under my guidance and supervision. To the best of my knowledge, the present work is the result of her original investigation and study. No part of the dissertation proposal has ever been submitted for any other degree or diploma. The dissertation proposal is fit for the submission and the partial fulfilment of the conditions for the award of **Master of Technology in Computer Science & Engineering.**

Date:

Signature of the advisor

Virrat Devaser

Assistant Professor

UID: 14591

Lovely Professional University

ACKNOWLEDGEMENT

DECLARATION

I hereby declare that the dissertation entitled, "**Using Dempster Shafer Theory for Evaluating Relative Performance of Securities and Sector Wise Indices**" submitted for the Degree of **Master of Technology in Computer Science & Engineering** is entirely my original work and all ideas and references have been duly acknowledged. It does not contain any work for the award of any other degree or diploma.

_____ **Komal Walia**

Date: _____ **Reg. No: 11000042**

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Today data is increasing day by day. Large amount of data is available which if not converted into the useful information is of no use. One needs to extract the useful patterns and information from such a huge amount of data. This is done with the help of Data Mining. Data mining is the mining of useful knowledge from data. It extracts useful information from data sets and transforms it into the structure which is understandable for further use .It uses the Extraction, Transformation, Loading (ETL) process in order to mine the data and help companies to focus on the information that is very important in their data warehouses. It is used in stock market, fraud detection, telecommunication industry, biological data analysis etc. Here we are considering the Stock Market. Stock market is the place where the shares are traded. It is the collection of buyers and sellers of stock which include either privately traded shares or the securities listed on a stock exchange list. It is also known as the equity market. Stock market is very uncertain and unpredictable. It is very hard to predict the nature of the stock market. Many investors who invested in the stock market are worried about seeing their investment value in red for months. This market is volatile in nature as price changes rapidly and it is difficult to predict the rise and fall in the prices in the stock market. This uncertainty has made the selection and the prediction of the stock very critical as because of this uncertainty, sometimes the company faces loss instead of profit. One cannot predict the expected return and risk of a company as various factors affects it either directly or indirectly making movement of stock prices unpredictable and uncertain. By applying the data mining techniques, one can predict the behaviour of a stock market. There are many techniques like, Neural Networks, Fuzzy Logic, Dempster Shafer Theory, Support Vector Machine, Regression analysis etc. with the help of which the behaviour of stock market can be predicted up to certain extend.

1.1 FUZZY LOGIC

Fuzzy logic is the mathematical theory that is an extension of conventional or Boolean logic and that handles the concept of partial truth i.e. the value that lies between the

complete truth and complete false. It is based on the Fuzzy set theory that depicts that the logical variables can take the values between 0 and 1. This theory is based upon the degree of truth. Imprecise and incomplete data can be handled by it. Fuzzy logic is also used for the prediction of stock market as it has the ability to discover patterns in nonlinear system. It has been applied by the advance trading system as it can respond to the changing market. It is also used where the uncertainty occurs but it can't be scaled well to large rule sets. Due to this limitation instead of fuzzy logic, Dempster Shafer Theory has been applied for the prediction of the stock.

1.2 DEMPSTER SHAFER THEORY

Dempster Shafer Theory was first introduced by Arthur P. Dempster and was later developed by Glenn Shafer. This theory deals with the uncertainty which is the lack of detailed knowledge to make a proper decision. It allows the combination of evidences from different sources and arrives at the degree of belief that consider all the available evidences. It is also used in the Artificial Intelligence as it deals with the uncertainty in the expert system. This theory is an alternative for traditional probability theory in which evidence is associated with one hypothesis while in DS theory evidence is associated with the set of hypothesis. This theory can deal with the union of hypothesis. This theory is the generalization of probability theory.

Let X is the set representing all possible states of a system under consideration or the universal set. Then the power set will be

2^x (Set of all possible subsets of X including the empty set φ)

For example- if X={1,2}, then

2^x= {{ φ },{1},{2},X}

This Theory consist of three main components-
- Basic Probability assignment (BPA or Degree of Belief or mass value)
- Belief function (Bel)
- Plausibility function (Pl)

Basic Probability Assignment-

BPA is the mapping of power set of hypothesis to the interval ranging from 0 to 1. It is denoted by m and is considered to be the first step in the Dempster Shafer Evidence Theory.

m: 2^x →[0,1]

So, BPA has two properties-

- Mass of empty set is zero

    i.e. m(φ) =0.

- Summation of masses of all the remaining elements of power set is 1.

    i.e. Σ m(A)=1 where A belongs to 2^x

The mass of A i.e m(A), a given member of power set, depicts that the actual state belongs to A and not to a particular subset of A and expresses the proportion of all relevant and available evidence.

Belief and Plausibility Measures- These measures are derived from the mass function (m) called Basic Probability Assignment Function.

Belief- Belief usually denoted by bel (A)It is the sum of all the masses of subset of set of interest and is written as-

Bel (A) = Σm(B)

Where, B is the subset of A

Plausibility- Plausibility usually denoted by pl(A) is the sum of masses of set B that intersect the set of interest A. So,

Pl (A) = Σ m (B)

Where, B intersection A is not equal to φ

Belief and Plausibity are related to each other as-

Pl (A) =1-bel (not A)

Dempster Combination Rule-

This rule helps us to combine the evidences coming from the different independent sources. Its main purpose is to combine the mass function on the basis of degree of beliefs.

Let's assume we are given two mass functions m1 and m2.

Then the combination rule is-

$m1,2(A) = \sum m1(B) * \frac{m2(c)}{1-K}$

Where B intersection C is A,A is not equal to φ

$m1,2(φ) = 0$

$K = \sum m1(B) * m2(C)$

Where, B intersection C is φ

1.3 NATIONAL STOCK EXCHANGE (NSE)

Established in 1992 and in order to bring the transparency to the Indian capital market, the group of leading Indian Financial institutions at the behest of government of India, set up the NSE. It is one of the prominent stock exchange of India, located in Mumbai. It is the first demutualized electronic exchange in the country which offered easy trading facility to the investors spread across the length and breadth of the country and provided a modern, fully automated screen-based electronic trading system. On 23 January 2015, it was declared as the world's 12th largest stock exchange with the market capitalization of more than US\$1.65 trillion. It comprised both domestic and global investors. It is the first exchange in India to connect together the investor's base of the entire country. It ensured that anyone who has experience and is qualified and who met the minimum financial requirements was allowed to trade. In order to allow investors to securely hold and transfer their shares and bonds electronically, National Securities Depository Limited (NSDL) is created by NSE.

Companies included under NSE are-

1. HDFC Bank Ltd

2. ICICI Bank Ltd

3. Axis Bank Ltd

4. State Bank of India

5. Punjab National Bank

6. Bank of Baroda

7. Tata Consultancy Services Limited

8. Infosys Ltd

9. Wipro Ltd

10. HCL Technologies Ltd

11. Vedanta Ltd

12. Hindalco Industries Ltd Preferred Shares

13. Tata Steel Limited Preferred Shares

14. Steel Authority of India Limited

15. Reliance Industries Limited

16. Cairn India Limited

17. Oil & Natural Gas Corporation Limited

18. GAIL (India) Limited

Generally people follow two approaches in analysing and predicting the stock prices. These are-

- Fundamental Analysis- It analyse the fundamental factors that influence the economy. Various fundamental factors are there whose values affect the stock prices. The analysis of these factors help us in determining the future trends.

- Technical Analysis- It analyses the movement of prices, volumes etc. It uses stock charts in order to identify the future trends that will help us in predicting how the stock will perform in future.

There are number of fundamental factors, some of them are-

- Price to Book Value (P/B)-P/B also known as the price-equity ratio is the ratio to compare company market value to its book value. Stock is undervalued if its P/B is lower. It also signify that something wrong is there with the company.

- Earnings per Share (EPS)-EPS serves as an indicator of company's profitability and is the portion of a company's profit that is allocated to each outstanding share of common stock.

- Debt Ratio- Company's or consumer's leverage is measured by this factor. Financial risk is more if its value is higher.

- Price to Sales Ratio (PSR) – PSR is the ratio that compares company's stock price to its revenue. To calculate it divide company's market capitalization to its total sale. In order to compare the companies in the same sector, this ratio is used. Low ratio indicates undervaluation while ratio more than average indicates over valuation.

- Dividend Yield-It tells us about the total dividend paid by the company to the shareholders per year.

In this research, I have done the fundamental analysis by including some critical factors that affects the economy of the stock. The fundamental factors included in this research are relative price to earnings ratio, profit factor and market capitalization.

- Relative price to earnings ratio or relative P/E –Relative P/E is the ratio of the company's P/E to the market P/E.
- Profit Factor-Profit Factor is the ratio of profit that a profitable trade generates to the loss that are generated by the losing trade. Risk will be less if the profit factor is high.
- Market Capitalization (MCAP)-MCap is the aggregation of the value of the company's stock. It is acquired by the multiplication of the number of outstanding shares with the current price per share.

Once these factors are obtained, Dempster Shafer Theory is applied on them in order to forecast the performance of the stock market.

CHAPTER 2

REVIEW OF LITERATURE

Ypke Hiemstra,(1994),[12] introduced a fuzzy logic forecasting model's architecture to support easy modification of knowledge base, simulation and facilities to trace problem solving so that users can meet their needs and beliefs. Uncertain, vague information have been used by the experts to predict the stock market. In this paper, a general approach to stock market prediction and an architecture of fuzzy logic forecasting support system has been presented. As stock market is very uncertain and involves imprecise reasoning and imprecise concepts, so, fuzzy logic is used as it deals with the uncertain and vague information. To represent the forecast of stock market, general model has been described below. The advantages of using this model are knowledge management and knowledge accumulation.
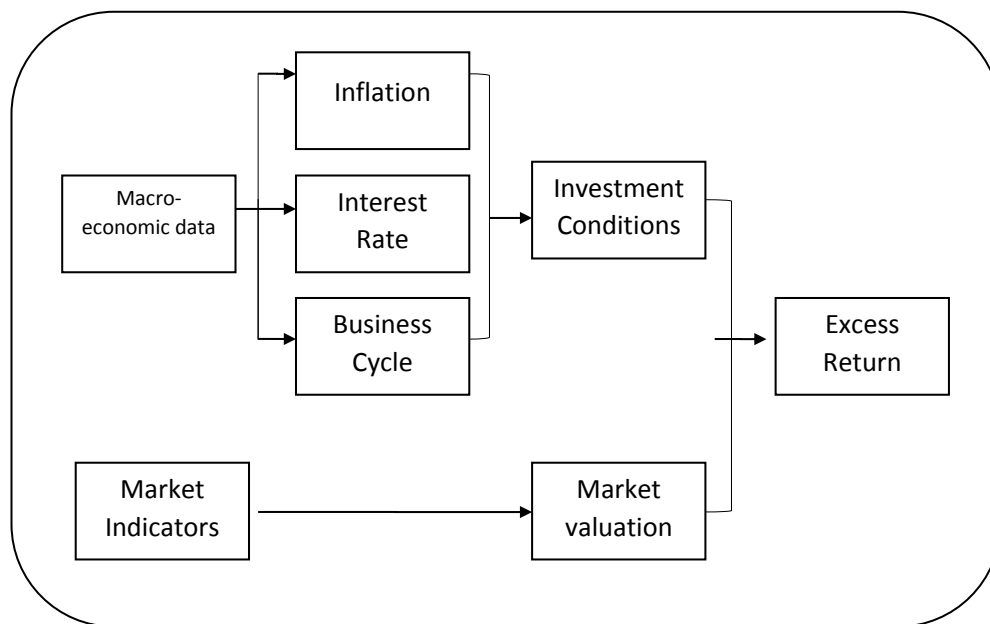

Fig.2.1: Basic approach to stock price prediction [12]

This model supports two components i.e. fuzzy logic model and user interface. User interface further support development interface and consultation interface and both of these interfaces accesses tracing and simulation facilities which are used for debugging

7

and explanation. General model has to be provided by the stock market forecasting model which user can process and offer interactive problem solving. Basic model is also discussed in which correlation coefficient of actual and predicted values have been used to evaluate and compare the forecasting problems. This model shows the great performance and user can improve it by tuning it.

Chiping Zhang et al., (2006), [3] proposed a new algorithm in which Dempster Shafer and Neural Network has been combined and presented to solve the problems of bad stabilization, low accurate identification and solution of uncertainty for solving the multi sensor system.

Information that the multi sensor system obtain is divided into groups, and each group is assigned a neural network. Also the unknown probability's concept is introduced. Neural networks and Dempster Shafer Theory has been applied to the problem of uncertainty. Neural networks have the tendency to recognize the targets and classify them whereas DS theory has been used to deal with the problem of uncertainty. In this paper, researchers have discussed a method in which both DS and BP Neural Network has been combined and is named as BPDS method. In BP Neural Network, weights are adjusted by back propagating if an error arises. This network consists of more than one layer. On the other hand, DS evidence Theory discusses Belief function and DS combination rule. In BPDS method, concept of unknown probability has been introduced which describes the unknown degree of targets. The unknown probability of targets described by the recognized targets $T_1,\ldots,T_n$ is:

$$\theta = 1 - \sum_{i=1}^{n} m(Ti)$$

As Neural network has low recognition rate and can sometimes produce false decisions to the system of great noise, Dempster Shafer Theory has been used to draw conclusions by combining the measurements of many times. But it also requires the proofs of every measurement. So the output of the network is regarded as the BPA function of every target and is considered to be the proof of DS theory. The result of Dempster Shafer evidential theory is the last decision. Simulation of robot system is also performed.

Chen Zhang et al.,(2007), [2] describes that in banks, operational risk (oprisk) is the primary management. Oprisks deals with uncertainty and the expert's knowledge plays an

important role in oprisk measurement. Unexpected oprisk has a lower frequency, but more serious loss consequences.
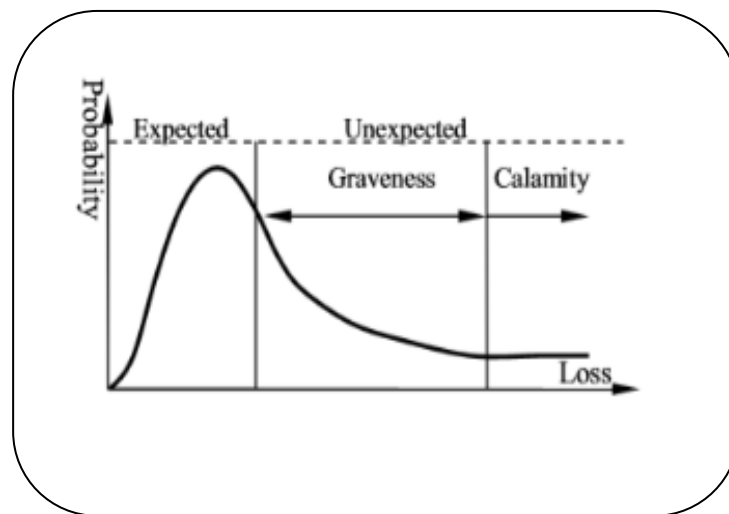
Fig.2.2: Distribution of the loss from oprisk [2]

Dempster Shafer Theory have been used in this paper in order to establish

- Frame of discernment
- Collection of expert's information
- Weight coefficients :(weight of the same group experts and weight between different groups) for the modification of Dempster's combining formula to find the final assessment of oprisk. The researchers confirmed the validity of these methods.

S. Chaigusin et al., (2008),[9] analysed the Stock Exchange of Thailand (SET) on the basis of the techniques used as it is very hard to determine the relevant factors that affect the stock market. Both at the local and global level, many uncertainties and interrelated economic and political factors affects the stock market. In this research, soft computing approach has been proposed on the SET forecasting that reveals the main driving indicators. SET index, SET50 index, SET100 index and the MAI index are the indices provided by the SET. One can calculate the SET index from the following formula:

$$SET\ Index = \frac{Current\ Market\ Value * 100}{BaseMarket\ Value}$$

Market value of all common stocks on April 30, 1975 is referred to as Base Market Value, which represents the base date.

Fundamental and Technical Analysis have also been discussed as these are the main concepts used in stock trading for economic strategy. Fundamental Analysis deals with the factors affecting the stock prices and expected returns from those prices. Soft computing simulates the human abilities and depicts the inexact solutions rather than the exact one. The exact solutions, certainty and precision is depicted by the hard computing. Fuzzy logic, neural networks and probabilistic reasoning are the areas where soft computing is used. Analysis on the basis of non-soft computing and soft computing has been done in determining the main factors influencing the stocks.

Xin Guan et al., (2008), [10] discusses about the Dempster Shafer Theory that deals with the uncertainty problems. It provides a method for combining the information from multiple sources and hence is applied in data fusion and pattern recognition. As the application of Dempster Factor is limited by the BPA function therefore it proposes three methods such as gray correlation analysis, fuzzy sets, and attributes measure in order to construct the BPA function. In the problem formulation, data from multiple sources has been obtained for pattern recognition. In this n classes $C_1, C_2 \ldots, C_n$ have been considered in which pattern x of k-dimensional have to be classified using training vectors with known classification. As multiple observations are available for each object, t measurements have been represented as a set$\{x^{(1)}, x^{(2)}, \ldots x^{(s)}, \ldots, x^{(t)}\}$, where $x^{(s)}$ is observation of source s. The frame of discernment is U={ C1,C2……,Cn } so, multi-feature and multi-classifier classifies the pattern in our observation, then results are combined from the observation using Dempster Shafer Theory. At last, decision is made on the basis of pattern recognition and classification on the basis of fusion result. After the problem formulation, the three methods are discussed. Also in order to demonstrate these methods, experiments to recognize emitter purpose are selected. To test the validity of these methods simulation has been performed by the researchers.

Yong Deng et al., (2008), [11] discussed about the modification of BPAs of hypotheses with the evidence sufficiency and evidence importance by the consensus index of evidence. For each hypothesis, different sets of BPAs are calculated by applying the Dempster's combination rule and new decision rules and the results for the same are obtained. In order to deal with the information from various sources, Dempster's combination rule is not efficient, so some new methods have been proposed. Fan used membership function and weights of features in order to modify BPAs of evidence and he

proposed the modification of Dempster's combination rule and decision rule. Further the Fan's method has been made easier with the introduction of consensus index of evidence γ on the basis of the distance function between evidences in order to increase the accuracy of the information fusion. After calculating the distance between evidences, its results are presented in distance matrix DM. After that similarity measure Sim_{ij} of two evidences is defined as: Sim(mi,mj) =1-d(mi,mj), and the similarity measure matrix (SMM) can obtained.

The support degree of the body of evidence is defined as

$Sup(m_i) = \sum_{j=1, j \neq i}^{m} Sim(mi, mj)$

Finally, consensus index(i = 1; 2; : : : ;m) of evidences is obtained-

$$\gamma = \frac{Sup(mi)}{\max Sup(mi)}$$

The new modification of BPAs considering the consensus index is :

$$m_{i,}.^*(A) = 1 - \sum_{B \subset \theta} \frac{\gamma imi,.(A), A \subset \theta}{\gamma imi,.(B), B \subset \theta, A = \theta}$$

We can find that the consensus index γ_i, when the conflict is bigger, the distance is far and the γ_i becomes smaller, and the γ_i will be more effect in reducing the conflict. After the consensus index, decision rule is formulated in which a threshold $\Delta = 0.6$ is set in order to qualify BPA. This implies that any given fault will have at least 60% of belief level. After this different sets of BPA are calculated using proposed method and then BPA of all evidences are combined and belief and plausibility function of each hypothesis are taken. Finally, both the maximum belief and maximum plausibility will be the final result. Fault diagnosis is taken as an example for the illustration of the proposed method and the comparison with the other methods. At last, by using appropriate modification of BPA of evidences and decision rules, conflicts among evidences are decreased and correct results are obtained.

Han Lock Siew et al., (2010),[5] examined the regression technique in order to predict the future trend for the stock. For handling currency values and financial ratios, the pre-transformed data source containing the data from different data types are used and the data format of these provides a process for computation of stock prices. Ranking of trends of stock prices are measured by the standardized ordinal data types in transformed data set and the outcomes of both the processes are examined. Weka machine learning

software has been used and its classifiers are used as an algorithm to produce an output. In this study, firstly the data in numerical values are converted into ordinal values, then regression based classifier is used to predict the ordinal data and at last the results are computed and compared.

Two models –random walk theory and efficient market hypothesis has been discussed in this paper. Along with this fundamental and technical analysis are also described. The approach called Machine Learning which is the branch of Artificial Intelligence is also depicted. This approach by finding the patterns from the data set and forming its own rules are used for making forecast in testing data sets. The Regression technique used here is the part of Machine Learning. This research starts by using fundamental analysis approach to gather numerical data in real valued format, after which the numerical values are converted into ordinal values using the transformation process. These ordinal values consist only the range of categorical enumerated values. Based on these enumerated values. The relationships between dependent and independent ordinal values are correlated. In this research data is collected from companies in Bursa Malaysia. Limitations of data are checked and after that methodology is proposed.
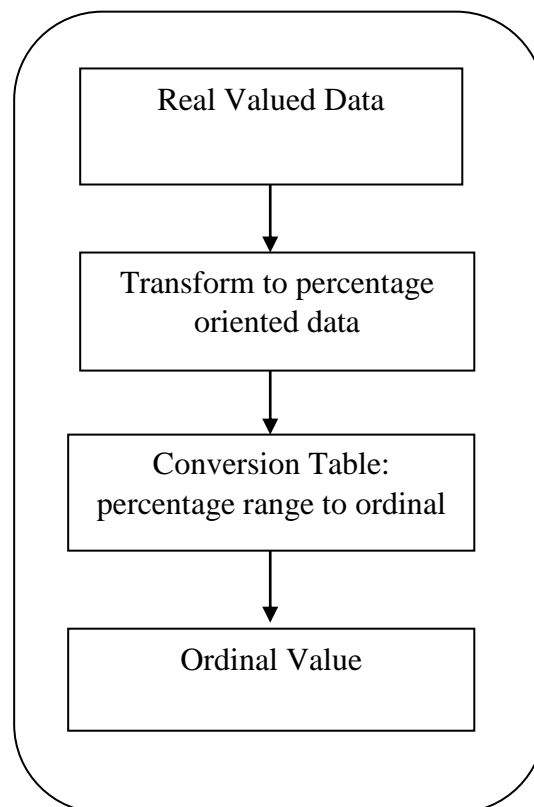
Fig.2.3: Data Transformation Model [5]

Data transformation model have also been described. Transformation objective is to make data more structured with the help of Regression technique.

The value of dependent variable based on the value of independent variable is thus predicted. Favourable outcomes are obtained when less structured data are transformed into more structured data in ordinal form. Since different types of data types exist, so further research can be conducted to compare its effects in regression techniques used for prediction of stock price trend.

Ryota Kato et al.,(2010) ,[7] proposed the prediction method that not only uses the predicting brands information but also the other brands information. The current patterns using the past incident patterns from the stock price are predicted by this method. This method firstly represents the stock price in the form of wave templates, and then it uses the classifier system to predict the incident patterns in the training period. At last predicts the current template in the test period. The method proposed is as-

- Simple Moving Average (SMA) is calculated in order to normalize the stock prices and smooth them.

  $$P_{SMA}(m) = \frac{p(m) + p(m-1) + \cdots \ldots + p(m-x+1)}{x}$$

  Where, p(t)= stock price

- Stock price p(t) is segmented into the regular window on N size. Each window has the difference of the stock price between a given time $P(N \times j + i + 1)$ and a one time later $P(N \times j + i)$ as a vector:

  $$V_j = \{v_i\}, i = 0, 1, \ldots, N-1, j = 0, 1, \ldots, M-1, \quad v_i = P(N \times j + i + 1) - P(N \times j + i),$$

  $$M = \frac{N_d}{N}$$

  Where N_d is the number of stock prices.

- K-means method is used to cluster each V_j and define the obtained representative vectors as "templates". Then these templates are clustered into three classes, up trend, down trend and flat trend.

- Each vector is approximated to the template that vector belongs to. Stock P is represented as-

P={T$_j$ }, j=0,1,…..M-1

- Incident patterns are discovered by the classifier system in the training period. Michigan approach is adopted here. One classifier represents one rule and function f is given in-

  f=N$_p$+w*N$_f$

  N$_p$= number of genes that matches the past states of referred brand, w=weight, N$_f$=number that all condition part genes match past genes of referred brand.

- Now check the general versatilities of all patterns. Strength S in introduced as the barometer of pattern's general versatilities. S is given the plus reward, if the condition part and the action part fits in otherwise it is given a minus reward.

  S=α*(N$_c$ - N$_w$)

  Where α=reward, N$_c$ =condition part and the action part fits in, N$_{w=}$only condition part fits in.

- Current template in the test period is predicted and reliability R is calculated on the basis of the value of S. If the highest S of some patterns are same then the pattern whose reliability R is highest is fired. R is calculated as-

  $$R=\frac{N_c}{N_f}$$

  N$_c$ = Number that the condition and action part fits in,

  N$_f$ = Number that the condition part fits in.

  R and S are revalued when every time the template is predicted and the correct value is renewed.

This proposal method is used in the buy sell strategy also. Stocks are bought at the beginning of the window if it predicts an uptrend template and are sold at the end of the window.

Mahbub Alam et al., (2011), [6] discussed about the two techniques Bin partitioning and Nearest Neighbour Algorithm. In bin partitioning, the real life data from Dhaka stock exchange as an input. Rule partitioning is the basis in order to normalize the data to bring

about the uniformity and correlation among attributes in order to suggest the trading pattern. Another technique is divided into the following phases in order to classify the stock market data-

- Pre-processing phase: Decision classes assignments, data cleansing, data normalization, attribute selection and discretization are included in this phase.

- Analysis and Rule Generating Phase: Preliminary knowledge, such as attribute selection based on correlation factor, derivation of rules based on attributes and discretization of values into bins, rule evaluation and prediction processes are included in this phase.

- Classification and Prediction phase: In this phase, rules generated from the previous phase to predict the stock price movement.

The nearest neighbour method is defined as the non-parametric class of regression in which series copy its own behaviour along the time. In Nearest neighbour algorithm a model has been discussed that depicts the behaviour of the stock data along the time span. Its main purpose is to locate the similar pieces of information in time, independently of the location. So, with the help of these algorithms, we are able to classify the series as buy or sell and predict the next day prices.

Chon Sern Tan et al.,(2012) [4] demonstrated the financial ratios of stocks using Bayesian networks. Financial ratios involves analysing company's financial statements to gain insight on its financial condition, future statements etc. and is a crucial part in fundamental analysis. Relationships among financial ratios and stock price changes are modelled in Bayesian network by using the local stock market data.

Two main features of BN are-

- Graphical Part- For enabling user to visualize the domain conveniently, it represents the structure of causal relationships between variables.

- Quantitative Part- this part is used for probabilistic inferencing and is attached to variables.

On the basis of the counts of correct predictions, network's reliability is tested on real data and also the comparison of portfolio mean return to that of a random portfolio has been made. As the resulting network consistently predicted the price changes more than

half of the times and also outperformed the random portfolio, it is considered to be reasonably reliable.

Ryota Kato et al., (2012), [8] proposed a method to predict stock market based on the interrelated time series. Interrelationships between the predicted stock and time series data have been fetched as it is difficult to predict which data is related to which stock. The daily up and daily down changes in the closing value are predicted with the help of discovered interrelationships. There are two ways to predict the stocks- by considering the data on predicted stock and the data aside from the predicted stocks. The first approach follows two methods (Fundamental Analysis and Technical Analysis). This method was unable to depict the interrelationships between the data and predicted stock. So, to solve this problem a method is proposed which fetches the interrelationships of change in price between the predicted stock and the time series.

In this method interrelations of changes in stock prices among predicted stock and time series are extracted and then these interrelationships are used to predict the stock.

This method is composed of two phases-

- Interrelation discovery phase

- Prediction phase

Evolution Strategy is used to calculate the variation pattern which represents how referenced time series have changed and then Likely Interrelated Data is fetched to predict the stock by using obtained interrelationships.

Aparna Anant Bhat et al., (2013),[1] presented an automatic trading system for Nifty in order to decide the buying and selling calls for intraday trading that combines various methods for improving the quality and precision of the prediction. In order to implement various technical indicators and to train the neural network, historical data has been used that predicts the movements for intraday Nifty. In this paper, technical indicators have been discussed that are available for technical analysis. Moving Average Convergence Divergence (MACD), Relative Strength Index (RSI) and Exponential Moving Average (EMA) are the various technical indicators.

- MACD- It has two lines- MACD line and Signal line. These two lines acts as a signal that depicts the trend changes with crossovers.

Fig.2.4: Moving Average Convergence and Divergence [1]

MACD Line: (12 day EMA-26 day EMA)

Signal Line: (9 days EMA of MACD line)

- EMA- It gives higher weight to recent prices.

Fig.2.5: Exponential Moving Average [1]

SMA: (10 period sum)/10

Multiplier: 2 / (Time periods+1) = (2 /(10+1)) = 0.1818(18.18%)

EMA: {Close− EMA (previous day)} multiplier + EMA (previous day)

- RSI- Overbought and over sold regions and hence change in momentum are indicated by this. It oscillates between 0-100. Above 70 is the oversold region and below 30 is the overbought region.

RS = (AverageGain /AverageLoss)

RSI = 100−(100 /(1+RS))

Backpropagation Neural Network and Sentiment Analysis has also been described here. In order to find the sentiment orientation, Sentiment analysis techniques are applied to popular blog articles written by domain experts and to user comments for achieving better prediction accuracy and to improve analysis. With the help of these methods, the system makes the prediction for every trading day that whether it will be a positive day or negative. Further it will decide the buy and sell calls thus achieving full automation in stock trading.

Zhen Hu et al., (2013), [13] used the Support Vector Machine (SVM) in order to predict the stock market. Firstly, for stock multivariate analysis, four company specific and six macroeconomic factors that may influence the stock market are selected. Secondly, in order to analyse the relationship of these factors and to predict the stock performance, SVM is used. It is a machine learning technique that can be used for the classification in which factors that may influence the performance of the stock are selected and then mapped to the class attributes which indicates the whole performance of stocks. It can be used for both classification and regression task. In this research, four company specific variables (net revenue, net income, price per earnings ratio of stock, diluted earnings per share) and six macroeconomic variables that direct the influence market (unemployment rate, consumer spending, consumer investment, federal funds rate, inflation rate and Dow Jones industrial average) have been used. Each stock is labelled in a specific year as either good or poor investment. If a price of a company rose over a given year, is it considered to be good otherwise poor investment? Companies are selected at random and then test and training samples are made. Then results are obtained that suggest SVM is a predictive tool for stock predictions in the financial market.

CHAPTER 3

PRESENT WORK

3.1 PROBLEM FORMULATION

Nowadays people invest most of their money in stock market rather than in banks to get more benefit. But due to fluctuations in stock market people do not have any idea from where will they get maximum profit. There are two factors reflecting the stock market, namely technical factors and fundamental factors. So proper prediction is required to make some assurance. Stock market which is the accumulation of buyers and sellers of stock have a very bright future and it is considered to be the method for a company to raise money. Stock market is very unpredictable and uncertain in nature. It is very hard to predict the movements in the stock market that what will happen on the next day, whether it will be a good day or the bad day i.e. the price of a stock will rise or will fall resulting in the bull market or the bear market. For the one who want to invest in the stock market needs to know what will be its performance in the future whether the investment should be done or not. So, one need to forecast the stock market performance.

In this research, I have used Dempster Shafer Theory to find out the result of stock market. This will help in finding the proper accuracy of various stock markets thus helping the people to get maximum return. Three factors namely, Relative Price to Earnings Ratio, Profit Factor and Market Capitalisation have been considered. These factors directly influence the stock market. They will act as the evidences of Dempster Shafer Theory and the mass values of returns good or poor or the combine degree of belief are taken as our hypothesis.

3.2 OBJECTIVE

- ▪ The objective of this research is to predict the performance of the securities based on the leader follower method. This leader follower method considers three factors. These factors are-
- ➢ Relative P/E ratio
- ➢ Profit Factor

- ➢ Market Capitalization
  - ▪ This research will help in finding the values for the leaders and followers for various companies whose result will forecast the performance of the stock.
  - ▪ To Determine Favourability of Stocks Using Dempster-Shafer Theory which indicates whether we will receive good or bad results for the stock analysis.

3.3 METHODOLOGY

The proposed methodology is based on the three factors. These factors are decided on the basis of the above papers. These factors are Relative Price to Earnings Ratio, Profit factor, Market Capitalization.

- ▪ Relative price to earnings ratio or relative P/E is the ratio of the company's P/E to the market P/E. It is compared over time and it helps in making the decision regarding the investment i.e. whether a person should invest or not.

- ▪ Profit Factor is the ratio of profit that a profitable trade generates to the losses that are generated by the losing trade. Risk will be less if the profit factor is high. Profit factor greater than 2 will indicates that the money that is made from the winning trades is twice than the one that is lost from the losing trade.

- ▪ Market Capitalization (MCAP) is the aggregation of the value of the company's stock. It is acquired by the multiplication of the number of outstanding shares with the current price per share. It is helpful for the investor in determining the return and risk in the share.

Once these factors are obtained, Dempster Shafer Theory is applied on them in order to forecast the performance. Historical data of different companies present in National Stock Exchange is collected from www.capitaline.com, www.nseindia.com. The factors discussed above are used to evaluate the performance of different stocks. These factors will act as the evidences of Dempster Shafer Theory and the mass values of returns good or poor or the combine degree of belief are taken as our hypothesis.

Once the data is collected, the basic probability is assigned to each factor in order to start the prediction process. Then, each hypothesis is assigned the degree of belief in the presence of these factors. The Leader Follower approach is followed here. In this approach the companies that are considered to be the leaders are followed by the companies that are considered to be their followers. The follower always follows the leader. If the leader is in

uptrend then the follower will also face the uptrend and vice versa with some time lag as follower will over a period of time approach the leader valuations.

3.3.1 RETURN PREDICTION PROCESS

In the prediction process, we start with assigning the Basic Probability of each factor. 6 years historical data for three factors of different companies present in National Stock Exchange (NSE) have been collected. Then degree of belief of each hypothesis is assigned on the presence of each factor.

Formula for each factor is:

- Relative Price to Earnings Ratio (P/E)- $\dfrac{Relative\ p\ e\ of\ follower}{Relative\ p\ e\ of\ Leader}$

- Profit Factor- $\dfrac{Profit\ Margin\ of\ Follower}{Profit\ Margin\ of\ Leader}$

- MCap= $\dfrac{\dfrac{Mcap\ of\ Follower}{Sales\ of\ Follower}}{\dfrac{Mcap\ of\ Leader}{Sales\ of\ Leader}}$

3.3.2 PROPOSED SYSTEM DESIGN
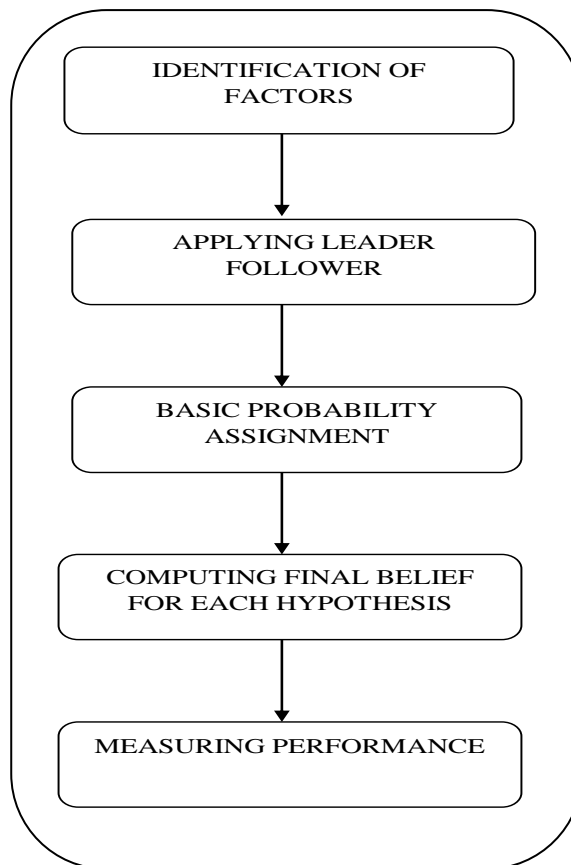
Fig.3.1: Design diagram of proposed system

21

3.3.3 BASIC PROBABILITY ASSIGNMENT

The first step in evidence theory is BPA. In this we assign the probabilities to each factor and on the basis of the result predicts whether the return will be good or bad in the future.

Leader-HDFCBANK

Followers-AXISBANK, ICICIBANK

HDFCBANK

S.No	Factors	Return	BPA
1.	PE	R_G	0.79
2.	Profit Factor	R_G	0.81
3.	MCap	R_P	0.61

FACTOR1 (P/E)

m1 (R_G) =0.79

m1 (Θ) =1-0.79=0.21

FACTOR2 (PROFIT FACTOR)

m2 (R_G) =0.81

m2 (Θ) =1-0.81=0.19

Combine factor 1 and factor 2

Combine factor 1 and factor 2	m2(R_G)=0.81	m2(Θ)=0.19
m1(R_G)=0.79	R_G=0.6399	R_G=0.1501
m1(Θ)=0.21	R_G=0.1701	Θ=0.0399

m3 (R_G) = 0.6399+0.1501+0.1701=0.9601

m3 (Ө) =0.0399

FACTOR3 (MCap)

m4 (R_P) = 0.61

m4 (Ө) =1-0.61=0.39

Combine factor 1 and factor 2

Combine factor 1 and factor 2	m4(R_P)= 0.61	m4(Ө)= 0.39
m3(R_G)= 0.9601	Φ=0.585661	R_G=0.374439
m3(Ө)=0.0399	R_P=0.024339	Ө=0.015561

m5 (R_G) = 0.374439/ (1-0.585661) = 0.374439/0.414339=0.90370204

m5 (R_P) = 0.024339/0.414339=0.05874175

m5 (Ө) =0.015561/0.414339=0.0375562

The value of m5 **(Return good)** is more than 0.5. So we can say that the performance or the return of HDFC Bank will be good in future.

AXISBANK

S.NO	FACTORS	RETURN	BPA
1.	PE	R_P	0.65
2.	Profit Factor	R_G	0.81
3.	Mcap	R_G	0.72

FACTOR1 (P/E)

m1 (R_P) =0.65

m1 (Ө) =1-0.65=0.35

FACTOR2 (PROFIT FACTOR)

m2 (R_G) =0.81

m2 (Ө) =1-0.81=0.19

Combine factor 1 and factor 2

Combine factor 1 and factor 2	m2(R_G)=0.81	m2(Ө)= 0.19
m1(R_P)=0.65	Φ=0.5265	R_P=0.1235
m1(Ө) =0.35	R_G=0.2835	Ө=0.0665

m3 (R_P) = 0.1235/ (1-0.5265) = 0.1235/0.4735=0.26082

m3 (R_G) = 0.2835/0.4735=0.59873

m3 (Ө) = 0.0665/0.4735=0.14044

FACTOR3 (MCap)

m4 (R_P, R_G) = 0.72

m4 (Ө) =1-0.72=0.28

Combine Factor 1, 2 and 3

Combine Factor 1, 2 and 3	m4(R_G)= 0.72	m4(Ө)= 0.28

m3(R_P)= 0.26082	Φ =0.18779	R_P=0.07302
m3(R_G)= 0.59873	R_G=0.43108	R_G=0.16764
m3(Ө)= 0.14044	R_G=0.10111	Ө=0.03932

m5 (R_P) = (0.18779+0.07302)/ (1-0.18779)=0.26081/0.81221=0.32111

m5 (R_G) = (0.43108+0.16764+0.10111)/0.81221=0.69983/0.81221=0.86163

m5 (Ө) = 0.03932/0.81221=0.21579

The value of m5 (**Return good**) is more than 0.5. So we can say that the performance or the return of AXIS bank will be good in future.

ICICIBANK

S.NO	FACTORS	RETURN	BPA
1.	PE	R_G	0.77
2.	Profit Factor	R_P	0.89
3.	Mcap	R_G	0.70

FACTOR1 (P/E)

m1 (R_G) =0.77

m1 (Ө) =1-0.77=0.23

FACTOR2 (PROFIT FACTOR)

m2 (R_P) =0.89

m2 (Ө) =1-0.89=0.11

Combining Factor 1 and 2

Combining Factor 1 and 2	m2(R_P)=0.89	m2(Θ)= 0.11
m1(R_G)=0.77	Φ =0.6853	R_G=0.0847
m1(Θ)= 0.23	R_P=0.2047	Θ=0.0253

m3 (R_G) = 0.0847/ (1-0.6853)=0.0847/0.3147=0.26914

m3 (R_P) = 0.2047/0.3147=0.65046

m3 (Θ) = 0.0253/0.3147=0.08039

FACTOR3 (MCap)

m4 (R_G) = 0.70

m4 (Θ) =1-0.70=0.30

Combining factors 1,2 and 3

Combining factors 1,2 and 3	m4(R_G)= 0.70	m4(Θ)=0.30
m3(R_G)= 0.26914	R_G=0.539	R_G=0.231
m3(R_P)= 0.65046	Φ =0.14329	R_P=0.06141
m3(Θ)= 0.08039	R_G=0.01771	Θ=0.00759

m5 (R_G) = (0.539+0.231+0.01771)/ (1-0.14329)=0.78771/0.85671=0.91945

m5 (R_P) = (0.06141)/ 0.85671=0.071681

m5 (Θ) = 0.00759/0.85671=0.00885

The value of m5 (**Return good**) is more than 0.5. So we can say that the performance or the return of ICICI BANK will be good in future.

Leader-SBIN

Followers-PNB, BANKBARODA

SBIN

S.NO	FACTORS	RETURN	BPA
1.	PE	R_G	0.81
2.	Profit Factor	R_P	0.62
3.	MCap	R_P	0.65

FACTOR1 (P/E)

m1 (R_G) =0.81

m1 (Θ) =1-0.81=0.19

FACTOR2 (PROFIT FACTOR)

m2 (R_P) =0.62

m2 (Θ) =1-0.62=0.38

Combining factor 1 and 2

Combining factor 1 and 2	m2(R_P)=0.62	m2(Θ)= 0.38
m1(R_G)=0.81	Φ=0.5022	R_G=0.3078
m1(Θ)=0.19	R_P=0.1178	Θ=0.0722

m3 (R_G) = 0.3078/ (1-0.5022) = 0.3078/0.4978=0.61832

m3 (R_P) = 0.1178/0.4978=0.23664

m3 (Θ) = 0.0722/0.4978=0.14503

FACTOR3 (MCap)

m4 (R_P) =0.65

m4 (Θ) =1-0.65=0.35

Combining factors 1, 2 and 3

Combining factors 1,2 and 3	m4(R_P)=0.65	m4(Θ) =0.35
m3(R_G)= 0.61832	Φ =0.40190	R_G=0.21641
m3(R_P)= 0.23664	R_P=0.15381	R_P=0.08282
m3(Θ)= 0.14503	R_P=0.09426	Θ=0.05076

m5 (R_G) = 0.21641/ (1-0.40190) = 0.21641/0.5981=0.36182

m5 (R_P) = (0.15381+0.08282+0.09426)/ 0.5981=0.33089/0.5981=0.55323

m5 (Θ) =0.05076/0.5981=0.08486

The value of m5 (**Return bad**) is more than 0.5. So we can say that the performance or the return of SBI BANK will be bad in future.

PNB

S.No.	FACTORS	RETURN	BPA
1.	PE	R_P	0.69
2.	Profit Factor	R_P	0.72
3.	MCap	R_G	0.61

FACTOR1 (P/E)

m1 (R_P) = 0.69

m1 (Θ) =1-0.69=0.31

FACTOR2 (PROFIT FACTOR)

m2 (R_P) = 0.72

m2 (Θ) =1-0.72=0.28

Combining factors 1 and 2

Combining factors 1 and 2	m2(R_P)= 0.72	m2(Θ)=0.28
m1(R_P)= 0.69	R_P=0.4968	R_P=0.1932
m1(Θ)=0.31	R_P=0.2232	Θ=0.0868

m3 (R_P) = (0.4968+0.1932+0.2232)/ (1-0) =0.9132

m3 (Θ) =0.0868

FACTOR 3 (MCap)

m4 (R_G) =0.61

m4 (Θ) =1-0.61=0.39

Combining factors 1, 2 and 3

Combining factors 1,2 and 3	m4(R_G)=0.61	m4(Θ) =0.39
m3(R_P) 0.9132	Φ =0.55705	R_P=0.35614
m3(Θ)=0.0868	R_G=0.05294	Θ=0.033852

m5 (R_P) = 0.35614/ (1-0.55705)= 0.35614/0.44295=0.80401

m5 (R_G) = 0.05294/0.44295=0.11951

m5 (Ө) =0.033852/0.44295=0.07642

The value of m5 (**Return bad**) is more than 0.5. So we can say that the performance or the return of PNB BANK will be bad in future.

BANKBARODA

S.NO	FACTORS	RETURN	BPA
1.	PE	R_G	0.61
2.	Profit Factor	R_P	0.87
3.	MCap	R_G	0.55

FACTOR 1 (P/E)

m1 (R_G) = 0.61

m1 (Ө) =1-0.61=0.39

FACTOR 2(PROFIT FACTOR)

m2 (R_P) = 0.87

m2 (Ө) =1-0.87=0.13

Combining factors 1 and 2

Combining factors 1 and 2	m2(R_P)= 0.87	m2(Ө)= 0.13
m1(R_G)= 0.61	Φ =0.5307	R_G=0.0793
m1(Ө)=0.39	R_P=0.3393	Ө=0.0507

m3 (R_G) = 0.0793/ (1-0.5307) = 0.0793/0.4693=0.16897

m3 (R_P) = 0.3393/0.4693=0.72299

m3 (Ө) = 0.0507/0.4693=0.10803

FACTOR 3 (MCap)

m4 (R_G) =0.55

m4 (Ө) =1-0.55=0.45

Combining factors 1, 2 and 3

Combining factors 1, 2 and 3	m4(R_G)=0.55	m4(Ө)=0.45
m3(R_G)= 0.16897	R_G =0.09293	R_G=0.07603
m3(R_P)= 0.72299	Φ =0.39764	R_P=0.32534
m3(Ө)= 0.10803	R_G=0.05941	Ө=0.04861

m5 (R_G) = (0.09293+0.07603+0.05941)/ (1-0.39764)=0.22837/0.60236=0.37912

m5 (R_P) = 0.32534/0.60236=0.54010

m5 (Ө) = 0.04861/0.60236=0.080699

The value of m5 (**Return bad**) is more than 0.5. So we can say that the performance or the return of BOB BANK will be bad in future.

CHAPTER 4
RESULTS AND DISCUSSIONS

4.1 RESULTS FROM DEMPSTER SHAFER THEORY

In this research, three factors have been considered and Dempster Shafer Theory have been applied on them in order to forecast the performance. Historical data from National Stock Exchange about different companies have been collected. These Factors will acts as an evidences of DS Theory and mass values of return good and poor are taken as our hypothesis. Leader Follower Approach is also followed in which the follower of a company will always follow the leader of a company. So, if a Leader is following an uptrend then all its followers will also follow the uptrend.

Firstly, 5 years data of each company is collected from NSE.

Secondly, factors are calculated for each company.

Stocks	P/E Ratio(2014)	P/E Ratio(2013)	P/E Ratio(2012)	P/E Ratio(2011)	P/E Ratio(2010)	P/E Ratio(2009)
HDFC	19.43826799	24.44932309	21.15150094	27.77104748	27.43552779	19.7582224
ICICI BANK	12.6612806	16.97759586	13.7566919	24.41703514	25.59630378	14.43981481
AXIS BANK	9.575252133	12.76001487	8.519444167	15.96621369	16.6974692	10.68781986
SBIN	11.70989569	12.29435504	9.815348427	20.59832106	16.3021685	8.872710361
PUNJAB NATIONAL BANK	6.646435665	6.959068835	5.802179022	8.57006673	7.676624782	5.160097221
BANK OF BARODA	6.060129691	8.571777105	5.899595687	8.061728395	6.615545701	4.621467026
TATA CONSULTANCY SERVICES	24.47199113	21.02370515	22.59587589	31.46062775	28.28186857	11.08719526
INFOSYS	20.64339882	15.64529943	20.26155935	32.93356131	26.05293025	12.15009208
WIPRO	19.23776224	18.60128617	22.50276549	25.31581741	20.87565405	12.46941896
HCL	14.97973054	12.27326036	15.84765178	28.77241806	24.17114094	8.053323594

Fig.4.1: Calculating PE for sample data

Stocks	Profit Factor(2014)	Profit Factor(2013)	Profit Factor(2012)	Profit Factor(2011)	Profit Factor(2010)	Profit Factor(2009)
HDFC	0.206108442	0.191823897	0.185371844	0.197027229	0.182325546	0.137454339
ICICI BANK	0.222066338	0.207744114	0.192747442	0.198327947	0.156571788	0.120869147
AXIS BANK	0.2029	0.1905	0.1928	0.2235	0.216	0.1675
SBIN	0.079876099	0.117880308	0.10990547	0.101536765	0.129110352	0.142991919
PUNJAB NATIONAL BANK	0.077332686	0.113347906	0.133901266	0.164285968	0.182305275	0.159932444
BANK OF BARODA	0.116618229	0.127305298	0.168733816	0.193808711	0.183151738	0.147578981
TATA CONSULTANCY SERVICES	0.285653058	0.264033505	0.288049594	0.25857606	0.243807724	0.209595342
INFOSYS	0.229900092	0.247953216	0.271005311	0.253811306	0.274503311	0.287159495
WIPRO	0.19056832	0.170035149	0.147314438	0.183465146	0.212898206	0.137594389
HCL	0.362762064	0.295955686	0.218970678	0.176360811	0.20803897	0.213324235

Fig.4.2: Calculating Profit Factor for sample data

Stocks	MCap(2014)	MCap(2013)	Mcap(2012)	Mcap(2011)	Mcap(2010)	Mcap(2009)
HDFC	4.36734112	4.23496092	4.37824861	5.46956901	5.4695966	2.52080259
ICICI BANK	3.25599017	3.00897177	3.04907915	4.9342789	4.1316021	1.19086494
AXIS BANK	2.23942069	2.24011172	2.15273205	3.80254525	4.0701446	1.37335409
SBIN	1.05033869	1.18492501	1.31976123	2.15938365	1.8591951	1.06152683
PUNJAB NATIONAL BANK	0.62331338	0.60549059	0.86110451	1.43240504	1.4916416	0.67037011
BANK OF BARODA	0.80054885	0.8164993	1.10690537	1.73513827	1.4049588	0.57038362
TATA CONSULTANCY SERVICES	6.44535553	6.35250902	5.99854925	7.9055004	6.6313405	2.35848484
INFOSYS	4.24628854	4.51379355	5.26386671	7.32065235	7.0983761	3.74294809
WIPRO	3.45213543	3.23998747	3.39394923	4.44653849	4.5106069	1.66336028
HCL	6.34979789	4.32044306	3.70653762	5.00169277	4.877059	2.66231078

Fig.4.3: Calculating MCap for sample data

Following Leader Follower Approach

Leader- HDFCBANK

Followers-ICICIBANK, AXISBANK

Leader-SBIN

Followers-PNB, BOB

Leader-TCS

Followers-INFOSYS, WIPRO, HCL

Stocks	P/E Ratio(2014)	P/E Ratio(2013)	P/E Ratio(2012)	P/E Ratio(2011)	P/E Ratio(2010)	P/E Ratio(2009)
relative pe=icici/hdfc	0.651358475	0.694399424	0.650388449	0.879226293	0.93296196	0.730825604
relative pe=axis/hdfc	0.49259801	0.521896448	0.402782015	0.574922991	0.608607544	0.540930234
relative pe=pnb/sbi	0.567591364	0.566037731	0.591133271	0.416056566	0.470895929	0.581569443
relative pe=bob/sbi	0.51752209	0.697212426	0.601058203	0.391377937	0.405807712	0.520863055
Relative pe=infy/tcs	0.843552072	0.744174222	0.896692806	1.046818314	0.921188435	1.095867061
Relative pe=wipro/tcs	0.786113485	0.884776781	0.995879319	0.804682526	0.738128529	1.124668473
Relative pe=hcl/tcs	0.612117357	0.583781987	0.701351514	0.914553209	0.854651484	0.726362566

Fig.4.4: Calculating PE after following Leader Follower approach for sample data

Stocks	Profit Factor(2014)	Profit Factor(2013)	Profit Factor(2012)	Profit Factor(2011)	Profit Factor(2010)	Profit Factor(2009)
icici/hdfc	1.07742476	1.08299392	1.03978813	1.00660171	0.8587485	0.879340355
axis/hdfc	0.98443323	0.99309837	1.04007165	1.13436098	1.18469411	1.218586488
pnb/sbi	0.96815802	0.96155082	1.21833123	1.61799491	1.41201129	1.118471906
bob/sbi	1.45998903	1.0799539	1.53526313	1.90875404	1.41856742	1.032079172
infy/tcs	0.8048228	0.93909754	0.94082865	0.9815731	1.1259008	1.370066205
wipro/tcs	0.66713209	0.6439908	0.5114204	0.709521	0.87322174	0.656476366
hcl/tcs	1.26993937	1.12090201	0.76018395	0.68204617	0.85329114	1.017790918

Fig.4.5: Calculating Profit Factor after following Leader Follower approach for sample data

Stocks	Mcap(2014)	Mcap(2013)	Mcap(2012)	Mcap(2011)	Mcap(2010)	Mcap(2009)
icici/hdfc	0.745531451	0.710507566	0.696415262	0.90213304	0.755376018	0.472414993
axis/hdfc	0.512765234	0.528956883	0.491687943	0.69521844	0.744139816	0.544808267
pnb/sbi	0.593440371	0.510994865	0.652469923	0.66333976	0.802305014	0.631514993
bob/sbi	0.762181627	0.689072552	0.838716385	0.80353404	0.755681207	0.537323788
infy/tcs	0.658813702	0.710552876	0.877523295	0.92602011	1.070428536	1.587013841
wipro/tcs	0.535600467	0.510032723	0.56579501	0.56246136	0.680195346	0.70526647
hcl/tcs	0.985174187	0.680116005	0.617905674	0.63268516	0.735455975	1.128822511

Fig.4.6: Calculating MCap after following Leader Follower approach for sample data

FINAL RESULT

Stocks	m5(R_G)	m5(R_P)	m5(Θ)
HDFC	0.903702	0.058742	0.037556
ICICI BANK	0.91945	0.071681	0.00885
AXIS BANK	0.32111	0.86163	0.21579
SBIN	0.36182	0.55323	0.08486
PUNJAB NATIONAL BANK	0.11951	0.80401	0.07642
BANK OF BARODA	0.37912	0.5401	0.080699
TATA CONSULTANCY SERVICES	0.57853	0.3425	0.07888
INFOSYS	0.9065	0.05605	0.03737
WIPRO	0.85503	0.0942	0.05072
HCL	0.91113	0.05774	0.03109

Fig.4.7: Final Mass Values of Sample Data

CHAPTER 5

CONCLUSION AND FUTURE SCOPE

The crucial goal of this research is to analyse the performance of stock market. Using Dempster-Shafer classification results are achieved that provide high accuracy with minimum error when applied to various stocks. These results will convey people a clear idea where to invest their money for maximum profits, thus arising their trust in stock markets. Although a number of challenges were encountered during prediction process but the proposed system provides an appropriate and efficient results despite of the biggest challenge of manually calculating the final mass values or degree of beliefs for different Stocks.

In future we can apply leader follower approach on the indexes corresponding to the large mcap, mid mcap and small mcap.

CHAPTER 6

REFERENCES

RESEARCH PAPERS

[1] Aparna Anant Bhat, Sowmya Kamath S., "Automated Stock Price Prediction and Trading Framework for Nifty Intraday Trading",2013

[2] Chen Zhang1, Weidong Zhu, Shanlin Yang, "Banking Operational Risk Management on DS Evidence Theory", 2007

[3] Chiping Zhang, Pingyuan Cui, Yingjun Zhang, "An Algorithm of Data Fusion Combined Neural Networks with DS Evidential Theory", 2006

[4] Chon Sern Tan, Chin Khian Yang, Yang Haur Tay, "Modeling Financial Ratios of Malaysian Plantation Stocks using Bayesian Networks",2012

[5] Han Lock Siew ,Md Jan Nordin, "Regression Techniques for the Prediction of Stock Price Trend", 2010

[6] Mahbub Alam, Asadullah AI Galib, Rashedur M Rahman , "Algorithms to Predict Opening Price and Trading Decision of Stocks in Dhaka Stock Exchange", 2011

[7] Ryota Kato, Noriko Yata and Tomoharu Nagao,"Stock Market Prediction Using Classifier System Based on Incident Pattern of Wave Template", 2010

[8] Ryota Kato and Tomoharu Nagao," Stock market prediction based on Interrelated Time Series Data",2012

[9] S. Chaigusin, C. Chirathamjaree, J. Clayden, "Soft computing in the forecasting of the Stock Exchange of Thailand (SET)", 2008

[10]   Xin Guan, Xiao Yi1, You He, "Study on Algorithms of determining basic probability assignment function in Dempster-Shafer Evidence Theory", 2008

[11]   Yong Deng,Dong Wang, Qi Li, "An improved combination rule in Fault diagnosis based on Dempster Shafer Theory", 2008

[12]   Ypke Hiemstra,"A Stock Market Forecasting Support System Based on Fuzzy   Logic",1994

[13]   Zhen Hu 1, Jie Zhu 2, and Ken Tse 3, "Stocks Market Prediction Using Support Vector Machine",2013

WEBSITES

[1] http://www.nd.com/define/profit.html

[2] http://www.investorwords.com/2969/market_capitalization.html

[3] http://business-fundas.com/2011/fuzzy-logic-and-fuzzy-systems/

[4] http://www.investopedia.com/terms/f/fuzzy-logic.asp

[5] http://en.wikipedia.org/wiki/Dempster%E2%80%93Shafer_theory

[6] http://capitaline.com/new/index.asp

[7]   http://www.nseindia.com/products/content/equities/equities/eq_security.htm

[8]   http://www.thearling.com/text/dmwhite/dmwhite.htm

[9]   http://en.wikipedia.org/wiki/National_Stock_Exchange_of_India

[10]   http://www.moneyworks4me.com/best-index/nse-stocks/top-nifty50-companies-list/

[11]   http://www.investopedia.com/

CHAPTER 7

APPENDIX

Table 7.1: Current price of the companies under NSE

Name of the company	Industry Name	Market Capitalization	Current Price (Change %)
ACC Ltd.	Cement & Construction Materials	Large Cap	1438.7 (-3.10%)
Ambuja Cements Ltd.	Cement & Construction Materials	Large Cap	234.2 (-1.95%)
Asian Paints Ltd.	Paints	Large Cap	762.9 -1.13%
Axis Bank Ltd.	Bank - Private	Large Cap	567.85 -2.70%
Bajaj Auto Ltd.	Automobile Two & Three Wheelers	Large Cap	1949.4 (-1.45%)
Bank Of Baroda	Bank - Public	Large Caps	168.95 (-1.92%)
Bharat Heavy Electricals Ltd.	Electric Equipment	Large Cap	237.2 -2.13%

Bharat Petroleum Corporation Ltd.	Refineries	Large Cap	765.75
			-1.34%
Bharti Airtel Ltd.	Telecommunication - Service Provider	Large Cap	381.65
			(-1.51%)
Cairn India Ltd.	Oil Exploration	Large Cap	213.6
			-1.16%
Cipla Ltd.	Pharmaceuticals & Drugs	Large Cap	637.05
			-1.22%
Coal India Ltd.	Mining & Minerals	Large Cap	363
			(-2.22%)
Dr. Reddys Laboratories Ltd.	Pharmaceuticals & Drugs	Large Cap	3308.2
			(-2.04%)
GAIL (India) Ltd.	Gas Transmission/Marketing	Large Cap	360.2
			(-1.60%)
Grasim Industries Ltd.	Diversified	Large Cap	3597.35
			(-0.29%)
HCL Technologies Ltd.	IT - Software	Large Cap	879.5
			(-0.90%)
HDFC Bank Ltd.	Bank - Private	Large Cap	989.2
			(-0.48%)
Hero MotoCorp Ltd.	Automobile Two & Three Wheelers	Large Cap	2329.45
			(-1.83%)
Hindalco Industries	Aluminium &	Large Cap	129.05

Ltd.	Aluminium Products		(-1.11%)
Hindustan Unilever Ltd.	Household & Personal Products	Large Cap	849.5
			(-0.97%)
Housing Development Finance Corporation Ltd.	Finance - Housing	Large Cap	1169.4
			(-2.70%)
ICICI Bank Ltd.	Bank - Private	Large Cap	331.25
			-0.33%
Idea Cellular Ltd.	Telecommunication - Service Provider	Large Cap	174.95
			(-3.74%)
IDFC Ltd.	Finance Term Lending	Large Cap	167.85
			-1.21%
IndusInd Bank Ltd.	Bank - Private	Large Cap	823.95
			(-1.07%)
Infosys Ltd.	IT - Software	Large Cap	1942.4
			(-0.98%)
ITC Ltd.	Cigarettes/Tobacco	Large Cap	322.25
			(-1.81%)
Kotak Mahindra Bank Ltd.	Bank - Private	Large Cap	1333.9
			-0.25%
Larsen & Toubro Ltd.	Engineering - Construction	Large Cap	1632.5
			(-1.35%)
Lupin Ltd.	Pharmaceuticals & Drugs	Large Cap	1772.8
			-1.55%

Mahindra & Mahindra Ltd.	Automobiles-Tractors	Large Cap	1145.15 (-2.25%)
Maruti Suzuki India Ltd.	Automobiles - Passenger Cars	Large Cap	3732.05 (-1.01%)
NMDC Ltd.	Mining & Minerals	Large Cap	127.85 0.00%
NTPC Ltd.	Power Generation/Distribution	Large Cap	150.3 (-0.73%)
Oil & Natural Gas Corporation Ltd.	Oil Exploration	Large Cap	304.6 (-0.52%)
Power Grid Corporation Of India Ltd.	Power Generation/Distribution	Large Cap	142.3 (-0.59%)
Punjab National Bank	Bank - Public	Large Cap	159.65 (-0.19%)
Reliance Industries Ltd.	Refineries	Large Cap	862.85 -1.36%
Sesa Sterlite Ltd.	Mining & Minerals	Large Cap	210 -1.38%
State Bank Of India	Bank - Public	Large Cap	270.05 (-0.64%)
Sun Pharmaceutical Industries Ltd.	Pharmaceuticals & Drugs	Large Cap	939.4 -0.02%
Tata Consultancy	IT - Software	Large Cap	2466.65

Services Ltd.			(-0.76%)
Tata Motors Ltd.	Automobiles-Trucks/Lcv	Large Cap	508.45
			(-2.21%)
Tata Power Company Ltd.	Power Generation/Distribution	Large Cap	75.85
			(-1.94%)
Tata Steel Ltd.	Steel/Sponge Iron/Pig Iron	Large Cap	360.5
			(-2.37%)
Tech Mahindra Ltd.	IT - Software	Large Cap	623.3
			-1.05%
Ultratech Cement Ltd.	Cement & Construction Materials	Large Cap	2666.1
			(-0.55%)
Wipro Ltd.	IT - Software	Large Cap	538.55
			(-0.53%)
Yes Bank Ltd.	Bank - Private	Large Cap	839.8
			-0.66%
Zee Entertainment Enterprises Ltd.	TV Broadcasting & Software Production	Large Cap	312.75
			(-2.42%)

7.1 QUESTIONNAIRE

Ques-1 What is Data mining?

Ans-1 Data mining is the mining of useful knowledge from data. It extracts useful information from data sets and transforms it into the structure which is understandable for further use .It uses the Extraction, Transformation, Loading (ETL) process in order to mine the data and help companies to focus on the information that is most important in their data warehouses.

Ques-2 What is stock?

Ans-2 A share in the ownership company is known as stock. It is a supply accumulated for future use.

Ques-3 What is a stock market?

Ans-3 Stock Market is a place where shares are traded. It is also known as the equity market.

Ques-4 what is a Fuzzy Logic?

Ans-4 Fuzzy logic is the mathematical theory that is an extension of conventional or Boolean logic and that handles the concept of partial truth i.e. the value that lies between the complete truth and complete false. It is based on the Fuzzy set theory that depicts that the logical variables can take the values between 0 and 1. This theory is based upon the degree of truth.

7.2 GLOSSARY OF TERMS

Dempster Shafer Theory - Dempster Shafer Theory deals with the uncertainty which is the lack of detailed knowledge to make a proper decision. It allows the combination of evidences from different sources and arrives at the degree of belief that consider all the available evidences.

Relative Price to Earnings Ratio- Relative price to earnings ratio or relative P/E is the ratio of the company's P/E to the market P/E.

Profit Factor- Profit Factor is the ratio of profit that a profitable trade generates to the losses that are generated by the losing trade.

Market Capitalization- Market Capitalization (MCAP) is the aggregation of the value of the company's stock. It is acquired by the multiplication of the number of outstanding shares with the current price per share.

Belief- Belief usually denoted by bel (A).It is the sum of all the masses of subset of set of interest.

Plausibility- Plausibility usually denoted by pl(A) is the sum of masses of set B that intersect the set of interest A.

Basic Probability Assignment (BPA)-BPA is the mapping of power set of hypothesis to the interval ranging from 0 to 1. It is denoted by m.

7.3 ABBREVIATIONS

Bel- Belief

BP- Back Propagation

BPA- Basic Probability Assignment

DS Theory-Dempster Shafer Theory

MCAP- Market Capitalization

Pl- Plausibility

Relative P/E-Relative Price to Earnings

SET- Stock Exchange of Thailand

BPDS- Back Propagation and Dempster Shafer

SVM-Support Vector Machine

NSDL- National Securities Depository Limited

PSR-Price to Sales Ratio

P/B-Price to Book Value

EPS-Earnings per Share

SBIN-State Bank of India

BANKBARODA-Bank of Baroda

PNB- Punjab National Bank

SMA-Simple Moving Average

MACD-Moving Average Convergence Divergence

RSI-Relative Strength Index

EMA-Exponential Moving Average

BN-Bayesian Network