

Stock Market Prediction and Investment Portfolio Selection Using Computational Approach

Keerti. S.Mahajan[#], Dr. R .V. Kulkarni*

[#]Department of Computer Studies, SIBER, Kolhapur. (MS), India-416004

* H.O.D Department of Computer Studies, SIBER, Kolhapur. (MS), India-416004

Abstract: Stock Market is considered as one of the fundamental pillars of national economy, as the important purpose of stock market in the economy is to raise capital and also to transfer funds to the most profitable opportunities, Therefore Predicting Stock Price, Predicting Timing and Selection of Portfolio are considered as several extremely complex situations in stock market. because, they are affected by many interrelated economic, social, political and even psychological factors at both local and global levels and all these factors are interrelated with each other in a very critical manner, These makes use of Mathematical Approaches & Soft Computing techniques combinely. These techniques can be used to select scripts in the portfolio and predict stock price or trading signal automatically with acceptable precision in order to retrieve information that could help to suggest investors on whether to Buy, Hold or Sell. This paper reviews the literature concerning different computational methods, such as linear programming, goal programming, data mining and soft computing strategies with an emphasis on the historical perspective and from many other resources such as, web sources, research papers, articles etc. The various fundamental, technical, economical and political factors of Indian Stock Market are considered.

Keywords: Stock Market, prediction, portfolio selection, technical indicators, financial indicators, economic and political factors, literature survey.

I. Introduction

Indian Stock Market has always been an interesting area of discussion in the financial sector, but it is very difficult to select a particular script. As Indian Stock Market is having more than 8900 scripts and it's difficult for investors to find the proper timing and the correct price of stocks. The main objective of investors is to earn profit, so, it is essential for traders and other investors to have knowledge about the financial area, which helps to take correct decisions about the investment in stock market. Therefore, computational techniques are essential for analysing the data and combine the different information, which helps to generate effective returns by the stock market. There is not one reliable method for assigning value to a company's script So, the artificial intelligence and its associated technologies are one of the methods are used to reduce the human efforts is been the interesting research area. Applications of machine learning and statistical techniques to automate and improve processes in finance domain have started more than a decade ago [19]. As there are more than 3000 listed scripts are traded, but all scripts are not profitable, hence, the selection of particular script in the investor's portfolio is important to get the high returns. Therefore for portfolio selection nearly all market participants considers following approaches of stock analysis such as, Fundamental analysis and Technical analysis and also many other social, economic and political factors are considered. And to get best results effective blending of all the factors is essential.

Fundamental analysis is one of the important starting points for picking up profitable companies. The most important part of fundamental analysis is looking into the financial statements. This involves looking at assets, revenue, liabilities, group selection etc. all the other financial aspects of a company. Fundamental analysts explore the certainties which affecting the company's value and growth by following a top down approach. There are huge numbers of investment strategies used in stock market to select scripts which are different from each other but all strategies use these fundamentals. So, this approach is suitable for portfolio selection.

Where as Technical analysis takes a completely different approach it is not related to intrinsic value of a company or a commodity. It is related to the price movements in the market. Technical analysis is a method of evaluating securities by analysing the statistics generated by market activity, such as past prices and volume. Technical analysts analyse different tools to identify patterns that can help to suggest future trend of the market. Which is represented using charts which shows the movements of price through time like stochastic oscillators or RSI(relative Stock Index.) and this technical approach is suitable for predicting stock price and investment timing in the stock market[18][19].

The Indian market is chosen as the economic domain to conduct the experiments as the Indian Stock market is the oldest stock exchange in Asia and claims to have the largest number of listed companies in the world is well placed on global economy.

1.1 About Indian Stock Market:

There are two stock exchanges for Equity Trading in India. They are the Bombay Stock Exchange (BSE) and the National Stock Exchange (NSE).

BSE (Bombay Stock Exchange): is the oldest and fastest stock exchange in Asia, with a speed of 200 microseconds which was established in 1875, and claims to have the largest number of listed companies in the world. There are 8900 listed scripts only about 3000 are traded every day. BSE secured the 10th position in the world by market capitalization at \$1.7 trillion as of 23 January 2015 [20].

BSE Indices:

BSE Sensex index was launched in 1986 to measure a overall performance of the stock exchange. And expanded to derivatives market and trading futures contracts in 2000 which helped to expand the trading in market which was transitioned to electronic trading system introduced by CMC Ltd in 1995 called BOLT(online trading)[20].

This BSE Sensex was followed by BSE National index which contains 100 listed scripts from the Mumbai, Calcutta, Delhi, Ahmadabad and Madras. And index called as BSE-100 and its calculations are based on the prices of particular stocks which are listed in BSE. and other indices of BSE are BSE-200, DOLLEX-200, BSE-500, BSE-PSU, DOLLEX-300 and BSE TECH. now except BSE-PSU all other indices are based on free float method. Generally BSE circulate information based on P/E Ratio, Dividend Yield and Price to Book value ration on day to day bases. [20]

NSE (National Stock Exchange): NSE is also plays a vital role in Indian market which was established in 1992 and considered as first fully automated Screen based electronic exchange in country which benefited to investor to trade from anywhere in the country. it has market capitalization of US \$ 1.65 trillion and had 12th place in world[1]. different index of NSE are CNX Nifty, flag Ships index the 50 stock index are heavily considered by the investors of India.

NSE provide many number of services such as trading in equities, indices, Mutual Funds, Exchange Trade Funds and Security Lending and Borrowing Scheme, an also in derivatives, debt and clearing to name few of them. NSE is mainly meant for transparency in the market by creating the National Securities Depository Ltd)in which investors can hold and transfer their shares and bonds electronically and ensured anyone who have basic financial knowledge can trade rather than being confined to broker.[1][20]

Indians are hugely equity averse. Only 1.2% of the Indian household financial savings is directly invested in shares (2010-2011) [1]. Because, according to Indian people's mentality, education about finance and considering to their earning, Investing in stock market is considered as high risk, But now a day's investments in stock market is regarded as high risk and high gains, so attracts Indian people to invest in stock market. Since the information available about the stock market is incomplete, uncertain, confusing and chaotic, making many researchers to work on the prediction of stock market which can helps investors to add high returns scripts to their portfolio.

As per the research so far in this area are done on prediction, selection and optimization of portfolio, recommendation for trading in market etc.

Statistical methods such as, linear regression, goal programming, technical analysis, fundamental analysis, data mining techniques and Soft Computing all are used to attempt to predict and optimize portfolio selection.

None of these approaches has proven to be always correct prediction tool. J.G. Agrawal & Dr.V. S Chourasia (2013) stated that many of these techniques are used to pre-process raw inputs, and their results are fed into neural networks as input. [3] But the motto behind this successful stock market prediction and selection of script is achieving best results with minimum input data and least complex model. So the soft computing approach is useful tool for analysing the stock market which helps to predict the stock price and to guide investors to add scripts in their portfolio. The soft computing field contains number of techniques that try to mimic the ability of human mind to effectively employ modes of reasoning that are approximate rather than exact.[4] the main techniques of soft computing are fuzzy logic, neural network theory, ANFIS, genetic algorithms and probabilistic learning.

[4]Panchal & Jayesh kumar patel (December 2012) defined the difference between the traditional computing and soft computing and they contributed Soft computing to solve complex problem such as stock market direction prediction in which they have concluded that soft computing methods are effective in real world by dealing with ill-structured problem. According to them Fuzzy logic, neural network and probabilistic

methods make a solid set either in their pure form or combined in hybrid solutions, can be used to tackle issues related to imprecision, learning, uncertainty, partial truth and approximation. We can say soft computing act like a human mind. Soft computing (SC) solutions are unpredictable, uncertain and between 0 and 1. [5]

Artificial neural networks are generally presented as systems of interconnected "neurons" which can compute values from inputs, and are capable of machine learning as well as pattern recognition.[5], fuzzy logic is mainly used to handle imprecision and uncertainty, and genetic algorithms for search and optimization.

The objective of this paper is to study the existing methods for prediction of Indian stock market and selection of portfolio and other related parameters of Indian stock market.

The paper is organized into 5 further sections. The section 2 is about the different impact factors which affects the stock market movement. Section 3 Reviews the significant work done in the field of stock market prediction and portfolio selection using different computational approaches. Section 4. Comparisons of different approaches. Section 5 Concludes with the look at the current and the future direction of the subject.

II. Impact Factors Of Stock Market Movement:

There are huge numbers of factors that make an impact on the stock market. Among those the most important are technical indicators, financial indicators and economic policies of stock market and political factors of the country.

2.1. Technical Indicators Of Stock Market:

[16] Technical analysis is defined as the study of price, volume, and technical indicators to make trading decisions. This type of analysis is based on historic data related to price and volume to predict, interpret and alert about the movement of market. There are many numbers of technical indicators present in stock market as a tool for technical analysis. And these indicators are used as input variables for their model to find out output. Mainly there are three types of indicators such as,

1. **Leading Indicators:** these indicators lead the price movement of stocks. This helps to find out overbought and oversold position.
2. **Lagging Indicators:** It helps to find trend continuation by following price movement
3. **Coincident Indicators:** Both price movement and an indicator occurs simultaneously which don't have much weight age in prediction purpose. Some Leading and Lagging indicators are,
 1. **Moving Averages**

1.1. Simple moving Average

1.2. Weighted Moving Average

Moving average indicators are lagging indicators which are used efficiently to predict the movement of stock price.

2. **Trading Index:** this indicator is used find whether to buy or sell a particular share. And it is for short term period.
3. **Tick Index:** It is a simple short term indicator and it is much suitable for intraday trading. To view the overall market movement. There are two types of tick index. One is uptick and the other is downtick. And difference between the two is the tick index.
4. **Relative Straight Index:** this indicator is used to find out the momentum of gains and losses to find overbought and oversold conditions of stocks. It is a leading indicator and used to find strength and weakness of a particular stock.

5. **Pivot Point:** This technical indicator is used to find the next day trend of the overall market on today's price action. Trading below the pivot point gives better results.

2.2. Financial Indicators Of Stock Market:

1. **Return on Capital Employed (ROCE):** A financial ratio that measures a company's profitability and the company's capital is ROCE. A higher ROCE indicates more efficient use of capital. Therefore ROCE should be higher indicates to add the script in your portfolio
2. **Earnings per Share:** this indicator is used to find the company's net income towards the total number of shares and also help to find P/E ratio.

3. **Liquidity Ratio:** Liquidity ratio expresses a company's ability to repay short-term creditors out of its total cash. The liquidity ratio is also known as quick ration and helps to find liquidity of assets and it should not be higher than 1.00.
4. **Return on Investment:** Return on investment allows an investor to evaluate the stock and investors prefer higher ROI. It is a general indication of the organizations efficiency.
5. **P/E Ratio:** Price-earnings ratio, is another important financial indicator and this should be higher which helps investors to take decision whether to buy or not. By comparing the market price with EPS. [18]

2.3. Macroeconomic Indicators Of Stock Market:

Analysis of historic data indicates that macroeconomic indicators and the stock market index are dependent on each other. Macroeconomic indicators make use of factor analysis technique to determine how much these variables affect stock returns. Macroeconomic indicators are liable for fluctuating performance of the stock market. These economic indicators are used to maintain the movement of stock market which are influenced by changes in factors like money supply, interest rate, inflation rate, exchange rate, international crude oil prices, external debt, and external reserve.

1. Monetary policy and repo rate hike:

- 1) Repo rate and stock market are very much correlated with each other, i.e. whenever the RBI hiked the repo rate, the market went down. This shows that the Indian Stock Market is more dominated by the RBI monetary policies than SEBI.

Therefore, investor should consider this factor which makes an impact on stock market movement.

2. International crude oil price and inflation:

India imports crude oil from international market, almost 80% oil is imported from other countries. Therefore if any consequent change occurs in international market prices makes an impact on inflation numbers and this inflation affects the stock market.

3. Policy announcements of the government:

The stock market and the government policies are very much correlated with each other stock market follows the rules and regulations of government. And government supports market by growth impetus with policy announcements. Almost each economic factor makes an impact on stock market but these three factors are positively correlated with market from last five years.

4. Gross Domestic Product (GDP):

This indicator plays a vital role in stock market prediction as it shows the economic progress made by the country. If GDP is high then capital inflow will be higher in the market.

5. Inflation:

Higher Inflation results negative impact on Stock Market as it generates higher interest rates in the economy. This is not recommended in the economic system of the country. A result into loss in corporate and cost of debt goes up. Recently RBI has increased the repo rate from 3.25% to 8.25%. Which had increased the funds from corporate resulting stock market displayed a negative outlook. [18]

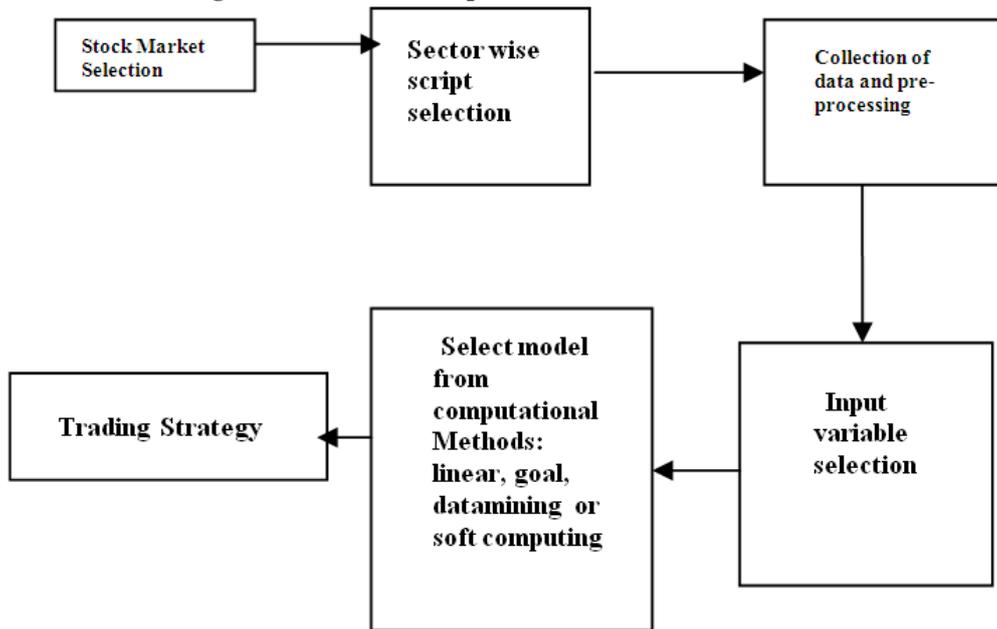
6. Foreign Institutional Investor Inflows

From last few years foreign investors are increased in India because investment made in Indian Stock Market has been profitable for foreign investors. Therefore FII is one of the important indicators for Indian stock market. As the flows are highly correlated with equity returns in India, and compared to local investors FIIs are more informatics.

III. Related Work:

Many models for stock market prediction and portfolio selection exists in the literature. The aim of this section is to highlight some of the more popular ones. Some common model will be assumed throughout this section and is given below.

Fig -1 Structure of Script Selection from Stock Market.



To select stock, a particular sector is to be chosen where the current study can be started because; there are 8900 listed scripts in Indian market. And it would become a difficult task to compare each and every stock from such a huge repository. To make easier to work on dataset particular sector has been chosen. Historical data and fundamental parameters are collected for respective sector. By data collecting and pre processing still the number of scripts gets reduced. Because there are so many extracted variables, so use some selected variables. Therefore, the suitable input variables such as technical indicators, financial indicators, and economic, social and psychological factors are considered.

A review on both prediction of stock market and selection of portfolio using different computational approach is presented.

Review on portfolio selection in Stock Market using Statistical and OR Approach:

Portfolio means a combination of assets. And it is said to efficient if it has a defined expected level of return for the given risk which is represented by standard deviation of the portfolio returns and said to be the efficient frontier. Therefore the efficient frontier is defined as the portion of the opportunity that gives the highest expected return for any given level of risk, and lies at the top of the opportunity set. Below in the related work many models considered this area as their research topic using operational research approach such as linear programming, goal programming etc.

3.1.1 Using Linear Programming Approach:-

[21] Abhijit Ravipati discussed in his theses about the Harry Markowitz portfolio theory and formulation of optimal portfolio selection is discussed with few examples. Harry Markowitz developed a theory which is to identify risk versus expected return. In March 1952 Markowitz developed the critical line algorithm using linear programming techniques for “Portfolio Selection”. And this model is based on mean-variance analysis, and it requires the knowledge of expected returns and standard deviation on each asset. And it was considered as the powerful tool in stock allocation for uncovering risk reduction using diversification.

According to author mean-variance is based on single period of investment. and in this one period model, by maximizing expected utility of ending the period wealth to choose portfolio weights is considered as a very complicated nonlinear programming problem.

According to him the investors should select such portfolio in which maximum return should be earned to the given variance for a given level of expected return.

In this paper authors he has given the optimal portfolio selection model as,

$$\text{Portfolio return: } R_p = \sum_i w_i E(R_i) \quad [21]$$

$$\text{Portfolio return variance: } \sigma_p^2 = \sum_i \sum_j w_i w_j \sigma_i \sigma_j \rho_{ij} \quad [21]$$

Where ρ_{ij} is the correlation coefficient between the returns on asset i and j .

By assuming portfolio have N assets with returns R_i

R_p is Return on portfolio

R_i is Return on asset i

W_i is Weight of component asset i

Σ_i is Standard deviation of asset i.

According to Markowitz the risk of stocks in portfolio are based on the asset weights and standard deviation of stock returns and correlation between them. The efficient frontier is one of the solutions to earn expected return of a portfolio. And this is used to define diversification. This diversification can increase returns without increasing risk.[23]

And also he has given many other important linear programming models such as Black- Litterman model which overcomes the problem of mean – variance optimization which consists of high concentrated portfolios, input- sensitivity estimation error maximization by using market equilibrium weights or capital asset pricing model. Regularization method, Bridge method. And used mat lab tool for mathematical calculation.[21]

Renata MANSINI et. al. (2003) presented on LP Solvable Models for Portfolio Selection

According to authors the Markowitz model used the standard deviation and variance as the risk measure. This has quadratic optimization problem. So the authors have shown that how to measure mean-risk in stock market using linear programming solvable model derived from the SSD shortfall criteria by using the simplest models which uses only the number of scenarios such as auxiliary variables and inequalities. And also used m-MAD model which is more detail and required more complex LP formulations. [8]

Mei Yu et.al discussed about Portfolio optimization problems with linear programming Models in which they compared four different models, such as, Markowitz's model, Konno's MAD model, Cai's model and Teo's model. They have considered Tokyo stock market and divided into two groups of data one is having 33 scripts for 72 months, and other data from 63 scripts for 120 months to compare them in the four models. They have considered efficient assets utility values for their model. Then they have constructed portfolios according to the defined models, compared the expected value with the true value. They have considered computational speed of each model and they found that, all the models have the same shape of frontier but the utility level may vary according to the risk-aversion factor. And authors concluded that, utility value is very high for Cai's model and lowest for Markowitz' model. Therefore it became difficult for the investor to select portfolio by considering the utility value of Cai's and Teo's models because, if the risk factor varies the utility value does not give proper result. They found that, the Markowitz's model and Konno's model have similar tendencies while Cai's and Teo' models seem to have similar when the expected returns are compared with true value for 10 months period. And also found that all the four models get higher final wealth compared with Nikkei 225 index and Nikkei 500 index respectively in most cases. and the Teo's model, is much influenced by the number of stocks and time periods.

According to them, the dynamic portfolio employing absolute deviation is considered for future work. And it is well known that a classical Mean-variance model has been extended to multiperiod cases using Mean Absolute Deviation (MAD) model is very interesting and challenging problem to be solved. [22]

Christos Papahristodoulou, discussed about the Optimal Portfolios using Linear Programming Models. According to him the classical Quadratic Programming formulation used to select portfolio is based on the expected return is multivariate which is distributed and on investors who are risk averter. He said that this model is bulky and time consuming so instead of this model he suggested the alternative models such as, maximin and minimization of absolute standard deviation, for these models he has constructed an optimal portfolio for customer for this he has taken only five OMX scripts from the Stockholm stock exchange A, B, C, D, and E for one year period and he found that the **maximin** model has the high risk and return, while the **min s** has the low risk and return, with the **min |s|** formulation being closed to **min s** formulation. For future work he has suggested to include more shares to examine.[24]

Mehmet Horasanli (2008), in his paper he extends the Optimal Portfolios using Linear Programming Models proposed by C Papahristodoulou to his new model which is based on a multi-asset setting to handle portfolio options and underlying assets using linear programming approach and is implemented to Novartis, Sanofi and AstraZeneca's call and put options. In his model portfolio and their related scripts is constructed using hedging strategy considering Greek letters such as delta, gamma, theta, rho and Vega, and found that, the impact of each Greek constraint on portfolio return by using duality analysis to get shadow prices. According to him shadow prices are very helpful to the investors to quantify the profit change based on whether the particular constraint is included or excluded to the model. And also the shadow prices are considered to identify, whether the optimum value increases or decreases, depending on Greek risk is increases or decreases.

Author suggested adding transaction cost to the presented model for further work to get optimum results. [24, 25]

Finally to conclude that using linear programming models with computer is very easy to setup and solve the problem but it lacks in dealing with options, as return on options are based on non-linear. Linear programming approach is not suitable for non-linearity. And one more problem is all the above models use constant historical volatility data taken from the past performance of particular Stock Market but volatility changes over time. Another limitation of the above presented models is that, an author has ignored transaction costs. As transaction costs play a vital role in dynamic hedging. Many researchers involved in this linear programming approach followed the mean-variance analysis of Markowitz (1952) for Portfolio Selection have only considered PS as risk adjusted return with little importance to other essential factors. Therefore, it is assumed that the investors are keen with only returns related to specific levels of risk when selecting scripts in their portfolios. But a portfolio analysis model includes many other important factors. To handle such several factors Goal Programming model is used by the researchers for portfolio selection.

3.1.2. Using Goal Programming Approach:-

In spite of number of research already has been done in portfolio selection, there has always been an intimation of multiple conflicting objectives in portfolio selection, such as liquidity, rate of risk, rate of return, macro economics, market dynamics etc.

The concept of goal is different from that of objective. An objective simply defines each direction like profit maximization and goal defines a target against which the attained solutions are compared. In this regard, GP optimizes the deviations from the prespecified targets, rather than performance of the solutions. The general form of GP model is the following: Kumar et al. (1978).

Max or Min $h(d_i^+, d_i^-)$

Subject to: $g_i(x) d_i^+, d_i^- = t_i$

$x \in S$

$d_i^+, d_i^- \geq 0$

Where,

- g_i is goal i defined as a function which may be linear or non-linear of decision variables x .
- t_i is the target value for goal g_i
- d_i^+, d_i^- are the deviations from the target value t_i . Represents underachievement and over achievement of the specified goal.
- h is a linear function of the deviational variables.

The right-hand side of these constraints defines the target values of the goal, which can be defined as some values of goals or their optimal values.

Goal Programming (GP) is widely used model in the field of multiple criteria decision making which allows decision maker to use number of variations of constraints and goals, particularly in the field of Portfolio Selection.

Therefore the Goal Programming is known as Multi-objective programming technique used to select the portfolio to satisfy the decision maker's prediction. Goal Programming is one of the important approaches for decision making problems such as to minimize the difference between the attainment of goals and their desire levels. [26]

These multi-objectives signals a stable financial environment, as the future is uncertain and investor is certain of the future, so the optimal Portfolio Selection problem is depend is reduced the structure of portfolio which will maximize the investors return and depends on the following facts. (Callin 2008):

1. A set of possible future course of events for the world.
2. Linking these set of possible future course of events to the returns of Individual securities.
3. A probabilities function of the possibility of each of the future course of events.
4. A method to determine whether one portfolio is preferable to another portfolio.

These facts are considered under different assumptions based on investor's mentality. Therefore GP approach is used to achieve using different strategies for analysis. Many researches are undergone using GP for PS problem Such as,

Kumar et al. (1978) discussed their work about the facts that Portfolio Selection techniques are characterized by assumptions of unified goals. Considering the real world situations with the presence of many conflicting goals is at best limited. They addressed the problem of goal conflicts in the portfolio selection of

Dual-Purpose Funds, and they have suggested extending the standard methodology, in terms of the development of a GP model in conceptual form, which can be applied for the resolution of inherent clash of interests.

Goal Programming used for portfolio selection is an analytical approach devised to address financial decision making problems where targets have been assigned to the attributes of a portfolio and where the decision maker is interested in minimising the non-achievement of the corresponding goals. [26]

Rania Azmi (2010), discussed in his paper about goal programming for Portfolio Selection to incorporating multiple goals such as portfolio's return, risk, liquidity, expense ratio, amongst other factors.

According to him GP model can be used for incorporating multi-period, extended factors and different risk measures into the Portfolio Selection analysis. Also, the decision maker can establish target values not only for the goals but also for relevant achievement functions.

To achieve this, a Meta-GP model could be formulated, which allows the decision maker to establish requirements on different achievement functions, rather than limiting their opinions to the requirements of a single variant. In this sense, this approach could be used as a second stage after GP problem for PS is being solved (Uria et al. 2002). Future research is warrant in the area of GP applications to PS, particularly for Mutual Funds as the need for incorporating extended factors is greatly manifest. [26]

Hameed Al-Qaheri et.al (2010) explored an perception about an end user DSS (Decision Support System) for portfolio selection based on Mean-Variance (M-V) Model of portfolio selection by Markowitz [1952, 1991]. The Decision Support System using Goal Programming model to achieve investor's objectives such as, asset allocation, diversification, rate of return in order to find optimum solution.

According to the authors the decision support system used in this model is flexible and could be continued through goal extension by adding more number of objectives and using other computational approaches such as Fuzzy, Neural Network, Rough Sets, Genetic Algorithm, Linear Programming etc. they have implemented the proposed GLP model using VB.NET and extended large scale LP/QP Solver within Solver Platform SDK from Frontline Systems, for portfolio selection.

Authors have generated 10 different scenarios by observing stock movements of 60 listed companies for the period 2002-2008 based on the degree of objective fulfilment and the level of return on investment. They suggested for future research on developing and utilizing other models using computation intelligence paradigm and a hybrid model consisting of OR and CI techniques. [27]

M. Arenas Parra et.al (2001) presented on portfolio selection using fuzzy goal programming for private investors considering three criterion such as, risk, return and liquidity to deal with optimum portfolio. They have formulated the problem using goal programming. Where the constraints and goals are fuzzified. Because the liquidity uses a high degree of subjectivity by quantifying the knowledge of experts. Their model suggests an efficient portfolio that verifies according to the investors goal. According to them if the investor is risk adverse, then the target value of the variance should be small. If the investors are ready to take higher risk, then there are high profitable targets and lower liquidity targets. An author also suggests which targets are not feasible. [28]

3.1.2 Using Data Mining Techniques :-

Data mining, is defined as the science and technology of investigating data in order to find previously unknown patterns, it is a part of knowledge discovery in database. In today's computer-driven world, these databases contain huge quantities of information. The accessibility and plentiful information makes data mining important and necessity. Financial institutions such as stock markets produce huge datasets that force to use data mining tools such as clustering, decision tree, neural network, association rules, factor analysis etc to predict the stock market movement. Many researches had been done using this approach such as, [29] [30]

K. Senthamarai Kannan, (2010), has implemented BSRCTB algorithm (Combinatorial Technique) .this algorithm was able to predict if the following day's closing price would increase or decrease better than chance (50%) with a high level of significance. Furthermore, their results shows that there is some validity to

technical analysis of stocks such as, Five methods of analysing stocks were combined to predict these methods were Typical Price (TP), Bollinger Bands, Relative Strength Index (RSI), CMI and Moving Average (MA). The algorithm BSRCTB has a profitable signal of 58.25% and from the five technical indicators the SMI and Bollinger Bands could produce more profitable signals. [31]

Mohammad Ali Afshar Kazemi (2014) explored fuzzy c-means method for Tehran stock exchange data for the period April 2010 to April 2014. Risk and Return as indicators to offer the best portfolio that has better performance than the market. In their study they have taken eight variables based on diversification Such as, prices, risk, return, P/E ratio, β , number of customers, number of traders and trading volume. To optimize suggested portfolio Markowitz model was used to cluster stock and to select portfolio clustering model is used and unsupervised data has been used by the authors. Three methods of k-means, fuzzy c-means and self-organizing maps (SOM) were selected for clustering that would be evaluated by the Silhouette index. [32]

Ekarin Varutbangkul (2013) explored a new stock trading which combines trading signals from six traditional trading systems: MACD, RSI, SO, MAE, BB, and CCI. Decision trees are used as filters in order to improve the quality of the combined trading signals. The result of his study shows that filtering with decision trees can improve the quality of the combined trading signals and this approach can overcome all individual trading systems, B&H strategy, and UA.[33]

As the economic globalization has been increased tremendously with increasing information technology, stock market data are generated and with the increase of economic globalization and evolution of information technology, financial data are being generated and collected at a remarkable pace. Therefore, there is a need for automated approaches to effective and efficient utilization of huge amount of financial data to support companies and individuals in strategic planning and investment decisions making. Therefore Data mining have been used to uncover hidden patterns and predict future trends and behaviours in stock market. [29][30]

3.1.3 Using Soft Computing Approach:

Soft Computing approach is used for constructing systems which can posse human like expertise in particular domain and can adapt to the changing environment to explain their decision by using various models such as, Fuzzy Logic (FL), Genetic Algorithm, Swarm Intelligence, Ant Colony Optimizations, Neural Network (NN), Machine Learning (ML). These soft computing models are applied to stock market data because the information they handle is uncertain and diffuse, so it is obvious to prefer realistic representations rather than just models assumed to be exact. The various related work has been conducted in this area are,

2.1. S.S.Lam (2001) defined in his paper an expandable genetic fuzzy system for stock market timing .according to him common trading rules are used in the system. In his model fuzzy rules are selected from the fuzzy knowledge base using genetic algorithm. And also he has used incremental approach to re-train the system in fix interval. And the results show that the system is more reliable as it can generate more profitable but-sell signals. And also been pointed out that the duration between re-training affects the performance of the system [7].

Abraham et.al. (2001) explored hybridization in soft computing techniques for automated stock market forecasting and trend analysis. Authors presented main component analysis to pre-process the input data, a Neural Network for 1 day ahead stock market forecasting and neuro-fuzzy system for analysing the trend of the predicted stock values .according to them forecasting and predicting stock market trends using hybridization model gives better results.[9]

Alberto Fernandez et.al. (2005) looked at heuristic method using artificial neural network to find out well organized portfolio selection problem .They have used Markowitz model based generalization which includes cardinality and bounding constraints and to deal with mixed quadratic and integer programming problems they used heuristic method of Hopfield neural network and used to solve the above problem.. And compared the results with genetic algorithms, tabu search, and simulated annealing. And they have concluded that neural network model has given better results than the other heuristic methods. [11]

Jin-Cherng Linet. al. (2007) shown that a time series forecasting model by independent component analysis mechanism. They have extracted some underlying factors using time series and used those factors as a forecasting base. They have concluded that within component ambiguity, correlation approximation and mean difference problems, independent component analysis mechanism had intrinsic limitations for time series forecasting. Under the linear time complexity, the component ambiguity and mean difference problem was solved by the proposed evaluation. The empirical data showed that their model exactly revealed the exhibity and accuracy in time series forecasting domain [6].

Ahmed A. Gamil et.al (2007) proposed a multi agent and fuzzy logic model for stock market decision making based on technical analysis using genetic algorithms. Authors have integrated DSS into an agent based frame work to enhance the information retrieval about stocks. And accessible in internet. They used real data from NASDAQ stocks and they suggested using more technical indicators for better results. [10].

Cheng Dong and Feng Wan (2009) presented on the techniques of Livermore as a set of fuzzy IF-THEN rules and its membership functions to find out stock market timing that when to buy and sell stocks of china's A stock market using different financial indicators such as ROCE, MACD and RSI and suggested for further research based no rule base optimization and auto adaptation of membership function for stock price forecasting. [12]

Esmail Hadavandi et.al (2010) defined about the genetic fuzzy expert model for stock price forecasting to improve the accuracy by using technical indicators as input variables to the genetic algorithm to extract rules of fuzzy expert system (GFS). And data base tuning for next day stock price prediction to extract use full patterns using rule induction method. They have used this GFS model for IBM stock price forecasting for short term and compared with the outcomes of with previous one using mean absolute percentage error. And achieved better prediction accuracy. [13]

Masoud Mansoury et.al (2012) reported some results of an hybrid approach as a decision support system is used to suggest investors to find out the best portfolio of stocks using financial indicators such as Chaikin Money flow, Relative Strength index, Random Fluctuations, Commodity channel index and Money Flow Index etc. is presented to evaluate a stock in the form of a decision support system. and this DSS will help to predict the Stock price. Authors have collected data from Iran Stock Market from 2001 to 2011. And the outcome is that the use indices and their combinations shows the DSS gives high accuracy precisions and they suggested divergence analysis can be used as a future research. The use of divergence analysis with examination of the saturation of buy / sell stocks can be very effective. Other work that can be done in this regard is the use of the proper indicators that weighs methods to select a stock in the group that relates to that stock. It means that assign high weight to the important indicators output rather than the same weight to them. [14]

Pravin S. Metkewar et. al.(2013) in this paper author used mathematical formulation of the Portfolio problem related to banking sector of the Indian stock market based on case 1..n. they have examined various input factors and correlating linear and non-linear membership functions and created fuzzy rule base. They have used Back propagation Neural Network to handle uncertainty of input parameters which helps for decision making. They used Adaptive Neuro Fuzzy Inference System (ANFIS) of type 'Sugeno' in the Matlab 7.5 and, outputs have been compared using triangular, trapezoidal and gbell membership functions. And authors conclusion is that non-linear membership function i.e., gbell works more prominently than linear functions because the nature of input data is non-linear and input is also depends upon three different sub-categories. [15].

IV. Conclusion:

In this paper, authors have attempted to review on the different techniques employed for stock prediction, portfolio selection and other related stock market parameters. The techniques reviewed are Linear programming, Goal Programming, Data Mining and soft computing models and it has been found that the above mentioned techniques and inputs are solely not much accurate for prediction of stock market trends as well as portfolio selection. There exist a gap between technologies and user requirements. If different impact factors such as social, economic, political and psychological factors which affect the stock market index are also taken into consideration as input variables, which helps to get better results. Also incorporating specific stock market sector knowledge will help in achieving better results.

References

- [1]. [thereformedbroker.com/.../10-things-you-need-to-know-about-indias-stock Market](http://thereformedbroker.com/.../10-things-you-need-to-know-about-indias-stock-Market)
- [2]. http://en.wikipedia.org/wiki/National_Stock_Exchange_of_India
- [3]. State of the Art in Stock Prediction Techniques by J.G.Agrawal , Dr. V.S. Chourasia, Dr.A.K.Mitra ,April 2013
- [4]. Role of soft computing techniques in predicting stock market direction by panchal Amitkumar Mansukhbhai and Patel , December 2012
- [5]. http://en.wikipedia.org/wiki/Soft_computing
- [6]. Jin-Cherng Linet. al.
- [7]. S.S.Lam , “ A Genetic Fuzzy Expert System for Stock Market Timing”, IEEE 2001,
- [8]. Renata MANSINI, Wlodzimierz OGRYCZAK, M. Grazia SPERANZA, “On LP
- [9]. Solvable Models for Portfolio Selection”, INFORMATICA,2003 “Institute of Mathematics and informatics, Vilnius”, 04,37-62.
- [10]. Abraham, A.B.Nath and P.K. Mahanti.(2001), “ Hybrid Intelligent Systems for Stock Market Analysis, “Computational Science, Springer-Verlag Germany, Vassil N. Alexandrov et.al.(Eds),ISBN 3-540-42233-1,San rancisco,USA,pp.337-345.
- [11]. Ahmed A.Gamil, Raafat S.El-Fouly, and Nevin M.Darwish, Cairo University,(2007), “Stock Technical Analysis using Multi Agent and Fuzzy Logic”. “World Congress on Engineering” -01 pp 2-4.
- [12]. Alberto Fernandez,Sergio Gomez,”Portfolioselection using neural networks”,
- [13]. ELSEVIER, “”Computer and Operations Research” 34 pp 1177-1191.
- [14]. Cheng Dong and Feng Wan (2009),”A Fuzzy Approach to Stock Market Timing”, IEEE
- [15]. Esmaeil Hadavandi,Hassan Shavandi, Arash Ghanbari,(2010), “A Genetic Fuzzy Expert System or Stock Price Forecasting”, IEEE.
- [16]. Masoud Mansoury, Bijan Mansoury, Seyyed Alireza Hashemi Golpayegan (2012), “Enhanced Decision Support System For Portfolio Management Using Financial Indicators”,”International Journal of Business Information Systems Strategies (IJBIS)” , Volume 1, Number 1,November 2012.
- [17]. Pravin S. Metkewar, Anuja M.Bokhare,Sanjay S. Karamपुरi, (2013) ”Application of Portfolio Model viewed as Adaptive Neuro Fuzzy Inference System (ANFIS)”,
- [18]. “International Journal of Emerging Technologies in Computational and Applied sciences (IJETCAS)”,pp 13-507.
- [19]. www.moneycontrol.com
- [20]. www.stockmangroup.com
- [21]. www.investopedia.com
- [22]. Carol Anne Hargreaves, Prateek Dixit, and Ankit Solanki,” Stock Portfolio Selection using Data Mining Approach”, “IOSR Journal of Engineering (IOSRJEN)”Vol. 3, Issue 11 (November. 2013), PP 42-48.
- [23]. "BSEIndia". BSEIndia. Retrieved 2010-08-26
- [24]. Abhijit Ravipati, “Markowitz’s Portfolio Selection Model and Related Problems”,
- [25]. Thesis, “Graduate school-new Brunswick Rutgers, the state university of new jersey”, January 2012.