# Portfolio Risk Analysis Using Data Mining Techniques

[1] Mrs.M.Deivamani, [2]Dr.M.Kannan

[1] M.Phil Research Scholar, Department of Computer Science and Applications SCSVMV University, Enathur,Kanchipuram
[2] Assistant Professor, Department of Computer Science and Applications SCSVMV University, Enathur,Kanchipuram

*Abstract: : This research work aims to rate stocks based on the fundamentals of the company. Hence, it helps the investor in knowing the quality of the company in which he has invested. This work also analyses the risk of the user's portfolio based on three different parameters, which is different from the technique adopted by the modern portfolio theory. It is to be noted that the modern portfolio theory used only the covariance of stocks to analyse the rise and does not take the fundamentals of the said company into account. Thus, it does not help in capturing the holistic risk of the portfolio. The three parameters used in this work include the fundamental analysis of the company, technical analysis of the stock and association among stocks in the portfolio. Thus this report will be useful for individual investors and financial institutions in calculating the risk they are taking to achieve superior returns. The system essentially provides an interface for an investor to manage his/her portfolio. This system has also been extended to mutual funds.*

*Keywords- :* Portfolio, Stocks, Investors, Association

## I. INTRODUCTION

Data mining, *helps to identify and extract the useful information from large set of database also it helps companies to focus the most important aspects for their business.* With the help of data mining functionalities it is easier to predict the future trends and give scope to take necessary decisions. The automated, prospective analyses offered by data mining move beyond the analyses of past events provided by retrospective tools typical of decision support systems. Data mining tools can answer business questions that traditionally were too time consuming to resolve. They scour databases for hidden patterns, finding predictive information that experts may miss because it lies outside their expectations.

## II. Literature Review

Victoria Lemieux & Payam S. Rahmdel developed prototype visual analytics tools for portfolio risk analysis with the help of clustering techniques. B.Nagarathna & Dr.M.Moorthi examined different data mining approaches helpful in financial sector. Akinsola Adeniyi F, Sokunbi M.A, Lawal .O.N, Okikiola F.M, applied ID3 algorithm for Lagos company consisting of 17000 records with 57 fields of data to determine risk levels in insurance policies.

Karina Marvin & Swati Bhatt examined portfolio analysis for 229 stocks with the help of clustering techniques. Sakshi Singh, Harsh Mittal, and Archana Purwar, identified similar investment patterns over the consumers in USA. Fuzzy data mining technique and Frequent Pattern Growth algorithm are applied to interpret the result.

Adebimpe, L.A, Adedara, Olusola & Longe, O.B, analyze portfolio investment from six organizations for the benefits of customers to understand and make successful investments and returns. Sjors Otten, Marco Spruit and Remko Helms, proposed new classification algorithm which will classify and visualize portfolio sales data set.

Sivasree M S & Rekha Sunny T, proposed a model for loan predictability with the help of 4520 records consist of 17 attributes. Decision tree induction algorithm is deployed to interpret the results. The prediction algorithm reveals that whether the loan is approved or rejected for the customers is generated.

Stefan Klotz & Andreas Lindermeir, applied cluster analysis technique in portfolio data set through financial institution and identified risk and return patterns. Sara Saadati GHASEM SOLTANLO, Alireza Naser SADRABADI conducted a case study from 150 shares in Sari city to understand purchasing behavior of stake holders using classification technique and association rule. The obtained result will helpful for the stake holders to classify the risk levels such as high, medium and low.

Rajanish Dass, discussed different applications which relates to applications of data mining techniques. Abhijit A. Sawant and P. M. Chawan discussed different algorithms such as Bayesian, decision tree, boosting, bagging, back propagation, genetic, particle swamp optimization and support vector machine to predict loan risk analysis.

Suraj Baviskar, Nitin Namdev, explained fundamental concept of stock market and proposed data mining technique to predict stock market trends.

### III.     Problem Statement

This section contains the problem statement, its scope and the approach to solve the same.

The Problem

The existing risk analysis techniques proposed by Modern Portfolio Theory do not capture the holistic risk of a stock. MPT uses only volatility of the historic data to capture the risk of each portfolio. This however, is insufficient in capturing the holistic risk of the stock since there is no proven correlation between volatility and returns for any given period of time for historic stock prices. Also, volatility fails to remain constant over a period of time, as in the volatility in one period might have no correlation with the volatility of another period. Hence, MPT loses its predictive power.

The aim here is to develop a procedure for analysing the risk associated with a portfolio at any given time so that it can be managed in an effective way. This must, therefore take into account the association between the stocks as well as the technical and fundamental factors for analysis.

Scope of the Problem

Banks, NBFCs and other financial institutions have their own portfolio of investments. They also have facilities to provide credit on deposit of shares from clients. There are risk audit teams in each of these institutions which monitor the risk involved in each investment. There is hence a need to analyse the risk of a portfolio before extending credit or before investing in any shares of the prospective client. Thus a holistic risk evaluation mechanism is needed for such activities. The proposed solution, thus, helps in analysing the risk of the portfolio in an efficient way. It helps financial institutions in developing their own personalised risk assessment model for each of their customers.

### IV.     PROPOSED METHODOLOGY

After going through the earlier work done by Markowiztzetal, it becomes increasingly clear that the system is required to be designed in such a manner to avoid any of the pitfalls of the above referenced theories. Hence it is necessary to design of three problem statements and the desired solution is mandatory. The design of the system must follow a smooth pattern and mists provide all functionalities necessary.

The methodology adopted calculates risk holistically by taking the fundamentals, technical and the association of the stock into account. The seven fundamental parameters taken in to account capture the risk associated with the company and help in predicting the returns from the stock in the long term. The four technical parameters capture the price risk of the stock. Technical factors also take the valuation risk into account.

The risk of individual stocks changes when they are grouped in a portfolio. The overall risk of the portfolio is analysed by taking the association and grouping of the stocks in the portfolio.

### V.     RESULTS

Association, in rough terms, is a method that involves grouping or clustering together of similar kinds of data together. Association is a data mining technique that helps perform risk assessment process with smoothness and admirable accuracy. This work delves into the core data mining concepts of association, association networks, trend association measures, regression and moving windows are used to determine the level of risk that a portfolio has absorbed. This work introduces a technique called Moving Approximation Transform for the same purpose.

This work consists of a series of case studies conducted on the given system to a certain correct functioning of the association problem. After the user enters his/her choice, the system prompts the user to select the number of stocks in the portfolio and subsequently also prompts he/she to select the stocks from a list of pre-defined stocks. The following three case studies were performed on the association measure of the portfolio and the results are displayed appropriately. The case study consists of three portfolios of low risk, medium risk and high risk profile. We have taken 6 stocks in each portfolio for our convenience.

CASE 1-HIGH RISK PORTFOLIO:
Input: 6 stocks namely- Bharati-Airtel, Bajaj Auto, BHEL, CIPLA, DLF & Kavey Seed
Table 1: Association Matrix for Case 1

| C | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 1.000 | 0.391 | 0.370 | 0.489 | 0.490 | 0.930 |
| 2 | | 1.000 | 0.899 | 0.880 | 0.920 | 0.920 |
| 3 | | | 1.000 | 0.824 | 0.920 | 0.920 |
| 4 | | | | 1.000 | 0.870 | 0.870 |
| 5 | | | | | 1.000 | |
| 6 | | | | | | 1.000 |

C- Company, 1-Bharathi Airtel, 2- Bajaj Auto, 3- BHEL, 4-CIPLA, 5-DLF, 6-Kavery Seed

The group (for a threshold of 0.75) is:
BAJAJ-AUTO
BHEL
CIPLA
DLF
Kavery Seed
These stocks are said to "SWING TOGETHER".
So, the category of the given portfolio is HIGH RISK.
Inference: From the table 1 association matrix, we can see that 5 out of 6 stocks swing together. Hence, the portfolio is said to be of HIGH RISK nature.

CASE 2- MEDIUM RISK PORTFOLIO:
Input: The 5 input stocks are- TATA Power, TATA Steel, TCS, WIPRO and Infosys.
Table 2: Association Matrix for Case2

| C | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1 | 1.000 | -0.050 | -0.030 | -0.060 | 0.175 |
| 2 | | 1.000 | -0.202 | 0.775 | 0.875 |
| 3 | | | 1.000 | -0.113 | -0.200 |
| 4 | | | | 1.000 | 0.828 |
| 5 | | | | | 1.000 |

C-Company, 1 – TATA Power, 2 – TATA Steel, 3 – TCS, 4-WIPRO, 5- INFOSYS

The group (for a threshold of 0.75) is:
TATA STEEL
WIPRO
INFOSYS
These stock are said to "SWING TOGETHER"
So, the category of the given portfolio is MEDIUM RISK.

Inference: From the table 2, we can observe that out of the 5 stocks given as input, based on their fundamental and technical parameters and taking into account the association measure between each stock two at a time, Tata-Steel, Wipro and Infosys exhibit strong association (>0.75). These stocks are therefore said to be swinging together. Since 3 out 5 stocks swing together, this is a MEDIUM RISK portfolio.

CASE 3- LOW RISK PORTFOLIO:
Input: The 6 input stocks are- Crompton Greaves, M&M, Maruti, L&T, Wipro and RCom.

Table 3. Association Matrix for Case 3

| C | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 1.000 | | | | | |
| 2 | 0.40 | 1.000 | 0.580 | 0.302 | 0.488 | 0.524 |
| 3 | | | 1.000 | | | |
| 4 | | | | 1.000 | | |
| 5 | | | | | 1.000 | |
| 6 | | | | | | 1.000 |

C-Company, 1- Crompton Greaves, 2 – M & M, 3 – Maruthi, 4 – RCOM, 5 – WIPRO, 6 – L & T

The group (for a threshold of 0.75) is:
RCOM
WIPRO
These stocks are said to "SWING TOGETHER".
So, the category of the given portfolio is LOW RISK.
Inference: From the table 3, we can observe that out of the 6 stocks given as input, RCOM and WIPRO show strong association (< 0.75) among each other. These stocks are therefore said to be swinging together. Since only 2 out 6 stocks swing together, this is a LOW RISK portfolio.

## VI.     DISCUSSION

The technical and fundamental parameters go a long way in finding the worth of a stock and the level of risk it poses to the shareholder. By performing the technical and fundamental analyses, we have covered a very important facet of the risk assessment procedure.

In many cases, portfolios might contain stocks belonging to the same sector. So, in case a sector performs badly, the entire portfolio is at risk of falling flat on its face. This is because one badly performing stock belonging to a certain portfolio might influence another stock belonging to the same sector, as the association between them is found to be rather high. It is hence considered very prudent to choose stocks belonging to diverse sectors to avoid the risk of high association.

Also, the case studies highlight the importance of association measures and networks in the risk analysis procedure. They provide a graphical clear and highly effectively means of measuring the risk associated with stocks in a portfolio, which the Modern Portfolio Theory fails to achieve.

## VII.     CONCLUSION & FUTURE DISCUSSIONS

The process of analysing the risk of portfolios is of great importance owing to the fact that the user might be prone to errors in judgement or in some cases might not be able to spot an underperforming stock in his/her portfolio. The system not only ensures that the right mix of stocks is maintained in their portfolio, but also gives the user ample suggestions of potential stocks with better fundamental and technical characteristics that can boost the performance of the said portfolio. These suggestions, if implemented, will result in providing more stability to the portfolio or mutual fund. Practitioners often add additional constraints to improve diversification and further limit risk. Examples of such constraints are asset, sector, and region portfolio weight limits.

The system takes in to consideration the different kinds of risks in holding a portfolio like fundamental, technical, association and concentration risk and also helps the user to optimise the risk-return ratio of the portfolio.

The system has taken only the 6 large stocks in the market which are the constituents of Sensex. This can be extended to include several other stocks in the mid cap and small cap category. Shares of more than 6000 companies' trade in the market and the system can be extended to include all those stocks. It can also be implemented for developed markets with different classification thresholds according to the local market needs. The designed system not only analyses the holistic risk of the portfolio by taking fundamental, technical and association parameters into consideration but also helps the user reduce the risk of the portfolio and increase the expected returns from the portfolio. The association network which has been used in

this work can also be used to analyse the association between the currencies of various countries and thereby benefit the people who want to hedge their currency against a basket of currencies.

## VIII.     REFERENCES

1. Victoria Lemieux & Payam S. Rahmdel, "Clustering Techniques and their Effect on Portfolio Formation and Risk Analysis", Office of Financial Research Staff Discussion paper, No.2015-01, January 2015, pp.1-7
2. B.Nagarathna & Dr.M.Moorthi, "A Review on Data Mining and its Techniques in Finance Sector", International Journal of Advance Research in  Computer Science and Management Studies,  ISSN: 2327782, Volume 3, Issue 3, March 2015, pp.29-37
3. Akinsola Adeniyi F, Sokunbi M.A, Lawal .O.N, Okikiola F.M, "A Data Mining Approach to Insurance Risk Analysis", International Journal of Engineering And Computer Science ISSN:2319-7242,Volume 4 Issue 2 February 2015, Page No. 10255-10258
4. Karina Marvin  & Swati Bhatt, "Creating Diversified Portfolios Using Cluster Analysis", Independent Work Report Fall, 2015, pp.1-26
5. Sakshi Singh, Harsh Mittal, and Archana Purwar, "Prediction of Investment Patterns Using Data Mining Techniques", International Journal of Computer and Communication Engineering, Vol. 3, No. 2, March 2014, pp.145-148
6. Adebimpe, L.A, Adedara, Olusola & Longe, O.B, "Forecasting Portfolio Investment Using Data Mining", African Journal of Computing & ICT, ISSN 2006-1781, Vol 5. No. 4, June 2012, pp.101-106.
7. Sjors Otten, Marco Spruit and Remko Helms, " Towards decision analytics in product portfolio management", Decision Analytics, 2015, pp.1-25
8. Sivasree M S & Rekha Sunny T,  "Loan Credibility Prediction System", International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181, Vol. 4 Issue 09, September-2015, pp.825-830
9. Stefan Klotz & Andreas Lindermeir, "Multivariate credit portfolio management using cluster analysis", The Journal of Risk Finance, Vol. 16 Issue: 2, 2015, pp.145 - 163
10. Sara Saadati GHASEM SOLTANLO, Alireza Naser SADRABADI,  "The application of data mining techniques in analysis the stock portfolio in order to identify common patterns in the behavior of shareholders (Case study of selected brokers in Mazandaran province)",  Cumhuriyet University Faculty of Science Science Journal (CSJ), Vol. 36, No: 4 Special Issue (2015), ISSN: 1300-1949, pp.730-738.
11. Rajanish Dass, "Data Mining in Banking and Finance: A Note for Bankers", Indian Institute of Management Ahmedabad, pp.1-15
12. Abhijit A. Sawant and P. M. Chawan,  "Study of Data Mining Techniques used for Financial Data Analysis", International Journal of Engineering Science and Innovative Technology, ISSN: 2319-5967, Volume 2, Issue 3, May 2013, pp.503-509.
13. Suraj Baviskar, Nitin Namdev,  "Analyzing and Predicting Stock Market Using Data Mining Techniques – A Review",  IJIRT,  Volume 2, Issue 3  ISSN: 2349-6002, 2015,pp.130-135