# PORTFOLIO FORMATION AND ITS PERFORMANCE CALCULATION WITH THE STOCKS SELECTED BY SOM USING TECHNICAL INDICATORS

Arif Ullah Khan[1], Bhupesh Gour[2], Asif Ullah Khan[3]
[1]M.Tech Scholar Department of Computer Science & Engineering, TIT, Bhopal, India
[2,3]Professor Department of Computer Science & Engineering, TIT, Bhopal, India

***ABSTRACT***

*Identification of useful patterns in price movement of a stock in stock market needs tremendous analytical skills and effort. Careful analysis of the available technical indicators will help finding the right timing of trading of a stock to maximize the gains. To help investors manage their portfolios, we propose a tool for clustering and classification of stock market data using an unsupervised learning algorithm, Self-Organizing Map (SOM).Our research is intended to assist users in identifying stocks and guide them with the correct timings of buying or selling of these stocks. It will help the investors to maximize their gains. We found that the Self-Organizing Map algorithm can analyze and cluster the stock data reasonably well. This paper proposes an unsupervised clustering approach based method for stocks selection for our portfolio. The portfolio formed with selected stocks gives a maximum of 40.08% and an average of 20.6%.*

***KEY WORDS:*** *Neural Network, Stocks Classification, Technical Analysis, Fundamental Analysis, Self-Organizing Map (SOM).*

## I. INTRODUCTION

Selection of stocks and timing of investment is a challenging task. The stock price of a company is dependent on various factors, mainly known as fundamental technical parameters. The stock market contains a huge amount of variable data. For effective management of portfolio, investors must analyze stock market data regularly to identify overvalued / undervalued stocks. Finding useful information in such complex data, however, requires high analytical skills and effort. The complexity of stock market data and the tasks involved in analyzing this data call for a tool that can amplify human cognition. To help people gain insight into stock market data, we propose an analytical tool that analyzes the historical price movements of companies traded at NSE that clusters similar data together, and that visualizes the data onto a two-dimensional space using a learning algorithm called Self-Organizing Map (SOM). The purpose of this paper is to develop a method of classification of selected fundamentally strong stocks in to fixed number of classes by Self Organizing map using technical indicators. Each of the class is having its own properties; stocks having properties closer to a least class get selected for investment. Technical Analysis using this approach has short-term investment horizons, and access to price and exchange data. We combined a learning algorithm and information visualization to exploit human perceptual ability to recognize patterns and derive a lot of information from visualization with little effort. From a practical perspective, our technique can assist people in identifying stocks and the timing of trading.

## II. APPLICATION OF NEURAL NETWORKS IN STOCKS

### 2.1 Overview

The ability of neural networks to discover nonlinear relationships [3] in input data makes them ideal for modeling nonlinear dynamic systems such as the stock market. Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A neural network method can enhance an investor's forecasting ability [4]. Neural networks are also gaining popularity in forecasting market variables [5]. A trained neural network can be thought of as an expert in the category of information it has been given to analyze. This expert can then be used to provide projections given new situations of interest and answer "what if" questions. Traditionally forecasting research and practice had been dominated by statistical methods but results were insufficient in prediction accuracy [6]. Monica et al's work [7] supported the potential of NNs for forecasting and prediction. Asif Ullah Khan et al. [8] used the back propagation neural networks with different number of hidden layers to analyze the prediction of the buy/sell. Neural networks using back propagation algorithms having one hidden layer give more accurate results in comparison to two, three, four and five hidden layers.

## 2.2 Kohonen self-organizing map

Self-organizing maps (SOM) belong to a the class of neural network methods, which are nonlinear regression techniques that can be applied to find relationships between inputs and outputs or organize data so as to disclose so far unknown patterns or structures. It is an excellent tool in exploratory phase of data mining [9]. It is widely used in application to the analysis of financial information [10]. The results of the study indicate that self-organizing maps can be feasible tools for classification of large amounts of financial data [11]. The Self-Organizing Map, SOM, has established its position as a widely applied tool in data-analysis and visualization of high-dimensional data. Within other statistical methods the SOM has no close counterpart, and thus it provides a complementary view to the data. The SOM is, however, the most widely used method in this category, because it provides some notable advantages over the alternatives. These include, ease of use, especially for inexperienced users, and very intuitive display of the data projected on to a regular two-dimensional slab, as on a sheet of a paper. The main potential of the SOM is in exploratory data analysis, which differs from standard statistical data analysis in that there are no presumed set of hypotheses that are validated in the analysis. Instead, the hypotheses are generated from the data in the data-driven exploratory phase and validated in the confirmatory phase. In practical data analysis problems the most common task is to search for dependencies between variables. In such a problem, SOM can be used for getting insight to the data and for the initial search of potential dependencies. In general the findings need to be validated with more classical methods, in order to assess the confidence of the conclusions and to reject those that are not statistically significant. In this contribution we discuss the use of the SOM in searching for dependencies in the data. First we normalize the selected parameters and then we initialize the SOM network. We then train SOM to give the maximum likelihood estimate, so that we can associate a particular technical indicator with a particular node in the classification layer. The self-organizing networks assume a topological structure among the cluster units [2]. There are m cluster units, arranged in a one or two dimensional array: the input signals are n-dimensional. Fig. 1 shows architecture of self-organizing network (SOM), which consists of input layer, and Kohonen or clustering layer.

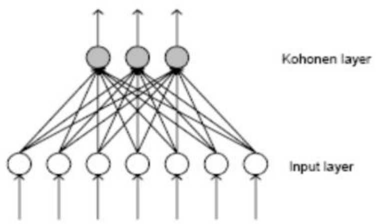


**Figure1:** Architecture of Kohonen self-organizing map

The shadowed units in the Fig. 1 are processing units. SOM network may cluster the data into N number of classes. When a self-organizing network is used, an input vector is presented at each step. These vectors constitute the "environment" of the network. Each new input produces an adaptation of the parameters. If such modifications are correctly controlled, the network can build a kind of internal representation of the environment.

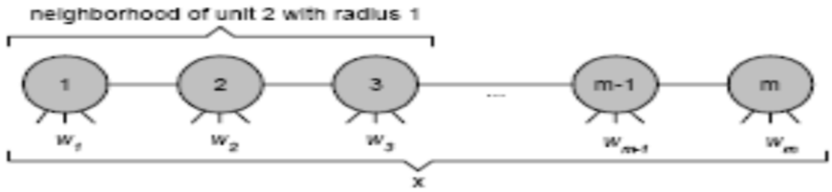


**Figure 2.** A one-dimensional lattice of computing units.

The n-dimensional weight vectors $w_1$, $w_2$, ...,$w_m$ are used for the computation. The objective of the clustering for each unit is to learn the specialized pattern present on different regions of input space as shown in Fig. 2. When an input from such a region is fed into the network, the corresponding unit should compute the maximum excitation [17]. SOM may distinctly reduce misclassification errors [12]. Kohonen's learning algorithm is used to guarantee that this effect is achieved. A Kohonen unit computes the Euclidian distance between an input $x$ and its weight vector $w$. The complete description of Kohonen learning algorithm can be found in [2] and [3].

## III. TECHNICAL ANALYSIS

Technical analysis is a method of analyzing statistics of securities, based on market activity like past prices and volume. In technical analysis, we looks for peaks, bottoms, trends, patterns, and other factors affecting a stock's price movement. Future trend of stock prices depends on their current & past values and volumes. Technical analysis looks for patterns and indicators on stock charts that will determine a stocks future performance. Technical analysts do not attempt to measure a security's intrinsic value, but instead use charts and other tools to identify patterns that can suggest future activity. In any case, technical analysts' exclusive use of historical price and volume data is what separates them from their fundamental counterparts. Security's past trading data is most important information. The field of technical analysis is based on three assumptions:

1. The market discount everything.
2. There is a trend in price movement.
3. History tends to repeat itself.

Technical analysis relies mainly on supply and demand of a particular script in the market in an attempt to determine what direction, or trend, will continue in the future. Technical analysis attempts to understand the emotions in the market. Williams's %R, Bollinger Bands, On Balance Volume, TRIX, CCI and ROC are mostly used technical indicators.

Recently, neural networks have been successfully applied in time-series problems to improve multivariate prediction ability. Neural networks have good generalization capabilities by mapping input values and output values of given patterns. Neural networks are usually robust against noisy or missing data, all of which are highly desirable properties in time series prediction problems. Various neural network models have already been developed for the stock market analysis.

### 3.1 Williams %R

William %R shows the relationship of the close relative to the high-low range over a set period of time. A value of %R close to zero indicates that the closing price is near the top of the range over N-days, where as a value close to -100 indicates that the closing price is near the bottom of the range. It is a momentum indicator especially popular for measuring overbought and oversold levels. It is used for identification of overbought/oversold levels. Low values (usually below -80) are considered as oversold levels where as values above -20 are considered overbought levels.

### 3.2 Bollinger Bands (BB)

Bollinger Bands are volatility bands placed above and below a moving average. Volatility is based on the standard deviation, which changes as volatility increases and decreases. The bands automatically widen when volatility increases and narrow when volatility decreases. This dynamic nature of

Bollinger Bands also means they can be used on different securities with the standard settings. For signals, Bollinger Bands can be used to identify M-Tops and W-Bottoms or to determine the strength of the trend.

### 3.3 On Balance Volume (OBV)

The OBV is calculated by adding the current period's volume to a cumulative OBV total when the trading instrument's price closes up and by subtracting the current period's volume from the cumulative OBV total when the trading instrument's price closes down.

### 3.4 Ease of Movement (EMV)

Ease of Movement is a volume-based oscillator that fluctuates above and below the zero line. As its name implies, it is designed to measure the "ease" of price movement. Ease of Movement takes Equivolume to the next level by quantifying the price/volume relationship and showing the results as an oscillator. In general, prices are advancing with relative ease when the oscillator is in positive territory. Conversely, prices are declining with relative ease when the oscillator is in negative territory.

### 3.5 Accumulation Distribution (AD)

The Accumulation Distribution Line is a volume-based indicator designed to measure the cumulative flow of money into and out of a security. Chaikin originally referred to the indicator as the Cumulative Money Flow Line. As with cumulative indicators, the Accumulation Distribution Line is a running total of each period's Money Flow Volume. First, a multiplier is calculated based on the relationship of the close to the high-low range. Second, the Money Flow Multiplier is multiplied by the period's volume to come up with a Money Flow Volume. A running total of the Money Flow Volume forms the Accumulation Distribution Line. Ultimate Oscillator is a momentum oscillator designed to capture momentum across

### 3.6 Chaikin Money Flow (CHMF)

Chaikin Money Flow measures the amount of Money Flow Volume over a specific period. Money Flow Volume forms the basis for the Accumulation Distribution Line. Instead of a cumulative total of Money Flow Volume, Chaikin Money Flow simply sums Money Flow Volume for a specific look-back period, typically 20 or 21 days. The resulting indicator fluctuates above/below the zero line just like an oscillator. Chartists weigh the balance of buying or selling pressure with the absolute level of Chaikin Money Flow. Chartists can also look for crosses above or below the zero line to identify changes on money flow.

## IV. EXPERIMENTAL RESULTS

The SOM algorithm has been implemented on Matlab platform, which run on Windows XP operating system. Stock data was sourced from National Stock Exchange (NSE) India. Historical data of one year, has been used to calculate the values of different technical indicators values. The input parameters of a particular stock over the period of one year are used with SOM algorithm are:

**Table 1:** List of Input Parameters

| Sl. No. | Input Parameter | Short Name |
|---|---|---|
| 1 | Open Rate | Open |
| 2 | High Rate | High |
| 3 | Low Rate | Low |
| 4 | Close Rate | Close |
| 5 | Volume | Vol |
| 6 | William % R | W%R |
| 7 | Bollinger Band | BB |
| 8 | On-Balance Volume | OBV |
| 9 | Ease of Movement | EMV |
| 10 | Accumulation Distribution | AD |
| 11 | Chaikin Money Flow | ChMF |

The historical data has been collected for one year (form 1/4/2013 to 13/03/2014) on the above specified parameters for Bajaj Auto Ltd., Bharti Airtel Ltd., BHEL Ltd., Infosys Ltd., NTPC Ltd., With these inputs SOM divides them into four classes[17][18]. As the SOM are more relevant to the problem where stocks of different companies are to be compared on some common parameters and arranges in the form of four different classes (Class1 to Class4). Out of these classes a very important pattern of stock price movement comes into picture, if the stock moves from lower class to higher class the price will move up and price will go down if the stock moves from higher class to lower class. We can purchase a stock on the second day after it enters in a lower class (Class1 or Class2) and sell the stock in reverse move. We can sell the stock on the second day as it enters in the upper class (Class3 or Class4).

**Table 2:** Bajaj Auto Stock giving sell signal

| Date | Rate | Class | Signal |
|---|---|---|---|
| 09-05-13 | 1826.5 | **4** | Sell |
| 10-05-13 | 1858.05 | **4** | Sell |
| 11-05-13 | 1862 | **4** | Sell |
| 13-05-13 | 1819.3 | **2** | Buy |
| 14-05-13 | 1787.7 | **2** | Buy |
| 15-05-13 | 1822.05 | **2** | Buy |

**Table 3:** Bajaj Auto Stock giving Buy signal

| Date | Rate | Class | Signal |
|---|---|---|---|
| 15/05/2013 | 1822.05 | **2** | Buy |
| 16/05/2013 | 1806.5 | **2** | Buy |
| 17/05/2013 | 1832.05 | **2** | Buy |
| 20/05/2013 | 1863.8 | **4** | Sell |
| 21/05/2013 | 1853.85 | **4** | Sell |
| 22/05/2013 | 1849.35 | **4** | Sell |

We have found that our selected stocks gives a maximum of 40.08% return and 20.6% average return on portfolio as shown in table 3.

**Table 4:** Total profit earned on Selected Stocks over a period of one year.

| Period = One Year | | | | | |
|---|---|---|---|---|---|
| Stock | Qty | Avg Rate | Invest Amt | Profit | Return % |
| Bajaj Auto | 100 | 1,898 | 189,798 | 29,445 | 15.5 |
| Bharti | 500 | 314 | 157,125 | 64,125 | 40.8 |
| BHEL | 500 | 157 | 78,704 | 20,125 | 25.6 |
| NTPC | 1000 | 142 | 142,100 | 600 | 0.4 |
| **Total** | | | **567,726** | **114295.0** | **20.6** |

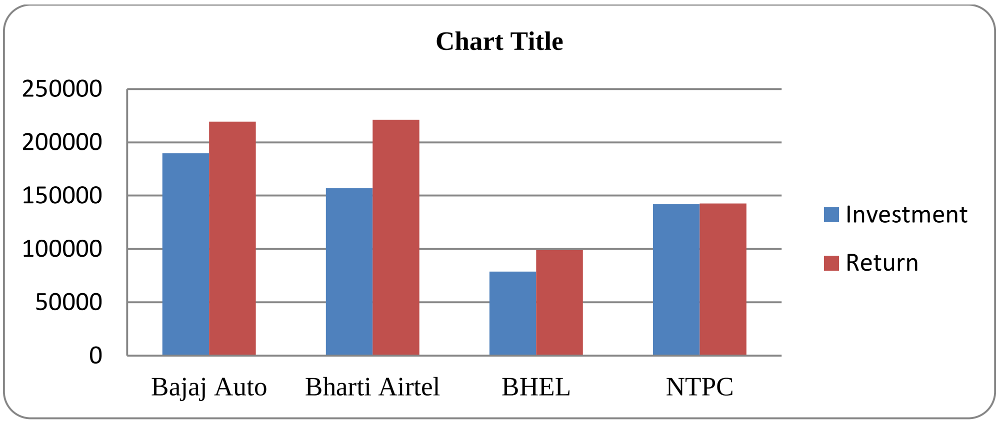


**Figure 3.** Return on investment.

## V. Conclusion

This paper presents the performances of the stock selected by using hybrid model of Self-Organizing Maps and technical indicators. The stocks selected by SOM and technical indicators help the investor not only in selecting stocks but also in identifying the timing of purchasing the particular stock. The result shows that the performance of stocks selected using self-organizing maps and technical indicators gives better returns on investment. Stock selected using SOM and Technical Indicators gives a maximum of 40.08% return and 20.6% average returns on our portfolio.

## VI. Acknowledgement

I sincerely thank TIT College, Bhopal for providing the opportunity and guidance for research work.

## References

[1] Mizuno, H., Kosaka, M., Yajima, H. and Komoda N., “Application of Neural Network to Technical Analysis of Stock Market Prediction”, *Studies in Informatic and Control*, 1998, Vol.7, No.3, pp.111-120.
[2] Haykin, Simon, “Neural Networks: A Comprehensive Foundation”, *Macmillian College Publishing Company*, New York, 1994.
[3] Phillip D. Wasserman, Van Nostrand, "Neural Computing: Theory and Practice", *Van Nostrand Reinhold*, New York, 1989.
[4] Youngohc yoon and George swales, “Predicting stock price performance: a neural network approach”, *IEEE publishing*, 1991.
[5] Shaikh A. Hamid, “Primer on using neural networks for forecasting market variables”, in *proceedings of the conference at school of business*, Southern New Hampshire university, 2004.
[6] Ramon Lawrence, “Using Neural Networks to Forecast Stock Market Prices”, *Course Project*, University of Manitoba Dec. 12, 1997.
[7] Monica Adya and Fred Collopy, “How Effective are Neural Networks at Forecasting and Prediction? A Review and Evaluation”, *Journal of Forecasting*, 1998.

[8] Asif Ullah Khan et al., "Stock Rate Prediction Using Back Propagation Algorithm: Analyzing the prediction accuracy with different number of hidden layers", *Glow gift*, Bhopal, 2005.
[9] Juha Vesanto and Esa Alhoniemi, "Clustering of the Self-Organizing Map", *IEEE Transactions on Neural Networks*, Vol. 11, No. 3, May 2000.
[10] Serrano, C., "Self Organizing Neural Networks for Financial Diagnosis", *Decision Support Systems Elsevier Science*, 1996, Vol 17, July, pp. 227-238.
[11] Tomas Eklund, "Assesing the feasibility of self organizing maps for data mining financial information", *ECIS*, June 6–8, 2002, Gdansk, Poland.
[12] Egidijus Merkevicius, Gintautas Garsva, "Forecasting of credit classes with the self organizing maps", *Informacines Technologies (ISSN 1392 – 124X)* Ir Valdymas, 2004, Nr.4(33).
[13] D. E. Goldberg, "Genetic Algorithms in Search, Optimization and Machine Learning." *New York: Addison-Wesley*, 1989.
[14] K. Bergerson and D. Wunsch, "A commodity trading model based on a neural network- expert system hybrid", *IJCNN-91- Seattle International Joint Conference*, Volume I, Issue 8-14 Jul 1991, Page(s): 289 – 293.
[15] Asif Ullah Khan et al., " Comparisons of Stock Rates Prediction Accuracy using Different Technical Indicators with Backpropagation Neural Network and Genetic Algorithm Based Backpropagation Neural Network", pp. 575-580, 978-0-7695-3267-7/08 $25.00 © *2008 IEEE DOI 10.1109/ICETET.2008.59.*
[16] H. White, "Economic prediction using neural networks: The case of IBM daily stock returns", in *Neural Networks in Finance and Investing*, chapter18, pages 315–328, 1993.
[17] Bhupesh Gour, et al., "Fingerprint Clustering Quality Analysis by using Self Organizing Map and K-Means Clustering Algorithms", Accepted in *International Conference on Pervasive Computing RGPV, Bhopal Dec*, 2008.

## AUTHORS BIOGRAPHY

**Arif Ullah Khan** holds a Bachelor degree in Engineering in Electrical from AMU, Aligarh, Masters Diploma in Business Administration from Symbiosis Institute of Management, Pune and Post Graduate Diploma in Human Rights from Indian Institute of Human Rights, New Delhi. He is also a Fellow of Institute of Electrical & Electronics Engineers (IETE). Arif Ullah Khan possesses more than 29 years of experience in field of Telecommunication. During this period, he has held various prestigious assignments with Government as well as with the corporate world. During his corporate tenure, he held post of Director - Regulatory Affairs & Industry Relations with Ericsson India Pvt Ltd and Director - Govt Affairs India & South Asia with Qualcomm India Pvt Ltd. Prior to his engagement with Qualcomm, Arif served Indian Air Force for 20 years as commissioned officer. He has also held the positions of Vice President of ITU-APT Foundation of India, Board Member of PTC India and Executive Member of TEMA.

**Bhupesh Gour** has received the B.E. degree in Computer Science & Engineering from Government Engineering College, Jabalpur in 2000, M-Tech in Computer Technology & Application from Rajiv Gandhi Technological University in 2005 and he received Ph.D. in Computer Science & Engineering from RGPV, Bhopal in year 2011. He is having total twelve years of industrial and academics experience. Presently he is working in Technocrats Institute of Technology, Bhopal.

**Asif Ullah Khan** is renowned educationist and administrator. He did his doctorate in Computer Science and Engineering from Rajiv Gandhi Technological University. He was the Gold Medalist in M.Tech. Degree in Computer Technology and Application from Rajiv Gandhi Technological University. He received the B.E. degree in Electronics and Instrumentation Engineering in 1990 from Samrat Ashok Technological Institute, Vidisha. He has organized various National Conferences and one IEEE international conference. He has been instrumental in getting MOU signed between TIT, Bhopal, India and Cincinnati University, USA. He was awarded the prestigious Bharat Shiksha Ratan Award in August, 2013. He has visited and delivered expert talk in Dubai, Malaysia, Thailand and Singapore. He is a senior member of Institute of Electrical and Electronics Engineers. He is also senior member of International Association of Computer Science and Information Technology.