NOVEL APPROACH FOR PREDICTING THE RISE AND FALL OF STOCK INDEX FOR A SPECIFIC COMPANY

Anirudh Prabhu¹, Aldrin Rodrigues¹, Chirag Chauhan¹, G.Anuradha²

¹Student, St. Francis Institute of Technology, Borivali (West), Mumbai – 400103
²Associate Professor, St. Francis Institute of Technology, Borivali (West), Mumbai – 400103

ABSTRACT

The financial market is highly fluctuating in nature. If investment strategies are not adequately planned and designed, it may lead to high monetary losses. There’s a high risk involved in stock market investment as the investment strategies are purely based on expert knowledge about the market conditions and some amounts of instincts.

In our proposed method using the insight of technical analysis and the concept of fuzzy inference system, we try to analyse the previous buy and sell scenarios and try to predict whether to go for buying or selling of stocks the next day. Since the investor obviously wants to make profits and minimize the losses, the motivation behind this project is to minimize the risk involved in an investment by suggesting a profitable investment plan and keeping the investors away from a non-profitable deal based on extensive analysis of the current market trends.

Keywords: Subtractive Clustering, Prediction, ANFIS, Stocks,

I. INTRODUCTION

Stock data analysis has remained one of the challenging time series analysis problems over the years. Because of the complexity and instructions, the problem has received attention from many researchers. In this paper we attempt to predict the percentage rise and fall of the price of a specific stock index using a linear implementation of subtractive clustering, neuro-fuzzy inference systems and a novel proposed algorithm to calculate the contribution factor of the different clusters formed. The main strength of neuro-fuzzy
inference systems is that they can combine the human-like reasoning style of fuzzy systems with the connectionist structure and learning ability of neural networks. We use subtractive clustering because it is an easy and efficient method for automatically generating fuzzy inference systems by detecting clusters in input-output training data. The ANFIS helps to predict the rise and fall using previous set patterns. Using the proposed method, one can predict the percentage rise and fall of a specific stock index. This method will be very useful for intra-day traders as they are more interested at making small profits on a day to day basis.

II. PREVIOUS WORK

Vaidehi , V , et. al[1] build a prediction system to predict the future occurrence of an event. It is a combination of a clustering algorithm and fuzzy system identification which proves effective in improving the efficiency of the prediction. To train the prediction system, historical data is obtained from the web. Data specific to any company stock is obtained and is recorded. This recorded information is studied and parameters containing only the necessary inputs to the prediction system. The subtractive clustering algorithm is used for its computational advantages and fuzzy rules are formed using system identification technique.


Ademola Olayemi (2007) mentioned that forecast performance improves when pre-processed data is used. He has used a general pre-processing method, based on multi scale wavelet decomposition to provide a local representation of time series data prior to the application of fuzzy models [2].

[7] presents an innovative approach for indicating stock market decisions that the investor should take for minimizing the risk involved in making investments. The system uses Adaptive Neuro-Fuzzy Inference System (ANFIS) for taking decisions based on the values of technical indicators. Among the various technical indicators available, the system uses weighted moving averages, divergence and RSI (Relative Strength Index).


III. SUBTRACTIVE CLUSTERING

Finding similarities in data and putting similar data into groups can be done using data clustering. Clustering partitions a data set into several groups such that the similarity within a group is larger than that among groups [8].

Subtractive clustering is a technique for automatically generating fuzzy inference systems by detecting clusters in input-output training data. The measure of potential for a data point is estimated based on the distance of this data point from all other data points. Therefore, a data point lying in a heap of other data points will have a high chance of being a cluster centre, while a data point which is located in an area of diffused and not concentrated data points will have a low chance of being a cluster centre.
After measuring the potential of every data point, the data point with the greatest potential value is selected as the first cluster centre. To find the next cluster centre, potentials of data points must be revised. For each data point, an amount proportional to its distance to the first cluster centre will be subtracted. This reduces the chance of a data point near the first cluster being selected as the next cluster centre. After revising the potential of all data points, the data point with the maximum potential will be selected as the next cluster centre. The potential of data points in the first step is measured as [9]:

$$p_i = \sum_{j=1}^{n} e^{-\alpha \|x_i - x_j\|^2}$$  \[\text{Equation 1}\]

Where,

$$\alpha = \frac{4}{r^2}$$

And $x_i$ is the $i^{th}$ data point and $r_a$ is a vector which consists of positive constants and represents the hyper sphere cluster radius in data space. The constant $r_a$ is effectively the radius defining a neighbourhood; data points outside this radius have little influence on the potential.

The potential which has been calculated through Equation 1 for a given point, is a function of that point's distance to all other points, and the data point which corresponds to maximum potential value is the first cluster centre. Let $p_1^*$ denotes the maximum potential, if $x_1^*$ denotes the first cluster centre corresponding to $p_1^*$.

$$p_1^* = \bigcup_{i=1}^{n} p_i$$  \[\text{Equation 2}\]

Where $\bigcup$ denotes the maximum of all $p_i$'s

To revise the potential values and select the next cluster, the following formula is suggested.

$$p_i = p_i - p_1^* e^{-\beta \|x_i - x_1^*\|^2}$$  \[\text{Equation 3}\]

Where,

$$\beta = \frac{4}{r_b^2}$$
And $r_b$ is a vector which consists of positive constants and is called the hyper sphere penalty radius. The constant $r_b$ is effectively the radius defining the neighborhood which will have measurable reductions in potential. To avoid cluster centres being close to each other, $r_b$ must be greater than $r_a$. A desirable relation is as follows [9]:

$$r_b = 1.5r_a \rightarrow Equation \ 4$$

IV.   OUR CONTRIBUTION

One of the problems with subtractive clustering is that it can perform clustering with only two factors at a time. So, in this paper we propose a simple function/algorithm to compute the effects and contribution more than two factors. Let us call this function as “CF Algorithm” (Contribution Factor Algorithm). The algorithm is as follows:

**CF Algorithm**

- Combine the outcomes from both graphs.
- Construct graph 1 and graph 2 as Closing price Vs Volume and High price Vs Low price respectively.
- Initialize cluster radius $CL_1$ and $CL_2$ from graphs 1 and 2 respectively and the corresponding distance of current data point as $d_1$ and $d_2$ respectively.
Calculate initial contribution factor $R_0$ and new contribution factor $R_n$ as follows:

If $CL_1 > CL_2$ then,

$$R_0 = \frac{CL_1}{CL_2}$$

$$R_n = R_0 + \left( \frac{d_1}{CL_1} - \frac{d_2}{CL_2} \right)$$

Else

$$R_0 = \frac{CL_2}{CL_1}$$

$$R_n = R_0 + \left( \frac{d_2}{CL_2} - \frac{d_1}{CL_1} \right)$$

V. PROPOSED APPROACH

The paper proposes that if the following algorithm is used in the exact sequence then an accurate prediction of the fall and rise of the stock index can be made. The algorithm is as follows:

Step 1: Fix initial parameters as Closing price, Volume, High Price and Low Price.

Step 2: Obtain the historical data of the above parameters for a specific period (e.g. 5 years) of a particular stock index.

Step 3: Map data points as Closing Price vs. Volume and High Price vs. Low Price on different graphs.

Step 4: Perform subtractive clustering on individual graphs to get clusters and cluster centers.

Step 5: Convert the cluster into initial rules.

Step 6: Predict outcomes from individual graphs using ANFIS Modelling.

Step 7: Implement CF Algorithm.

Step 8: Stop
Start

Fix initial parameters as Closing price, Volume, High Price and Low Price

Obtain the historical data of the above parameters for a specific period (e.g., 5 years) of a particular stock index

Map data points as Closing Price vs. Volume and High Price vs. Low Price on different graphs

Perform subtractive clustering on individual graphs to get clusters and cluster centers

Convert the cluster into initial rules

Predict outcomes from individual graphs using ANFIS Modelling

Combine the outcomes from both graphs

Initialize: cluster radius CR1 and CR2 & the corresponding distance of current data points d1 and d2

Calculate initial contribution factor and new contribution factor

CF Algorithm
VI. EXAMPLE

Above are the graph clusters formed by 4 years of Accenture’s stock data.
VII. CONCLUSION/FUTURE SCOPE

The proposed paper aims to help intraday traders efficiently deal with the rise and fall of specific stock index. Using the given algorithm, additional number of factors or parameters can be added to further accurately predict progress of stocks. The algorithm used can be improvised and improved for advance research.

VIII. REFERENCES

[10] K. V. Sujatha and S. Meenakshi Sundaram, “Regression, Theil’s and MLP Forecasting Models of Stock Index” International journal of Computer Engineering & Technology (IJCET), Volume 1, Issue 1, 2010, pp. 82 - 91, ISSN Print: 0976 – 6367, ISSN Online: 0976 – 6375, Published by IAEME.