Classification and Prediction of Stock Market Index Based on Fuzzy Metagraph

Thirunavukarasu Anbalagan a, S. Uma Maheswari b

Abstract

Stock market price forecasting is one of the challenging tasks due to the difficulty in predicting the non-linear and non-stationary time series data. In this paper a Fuzzy Metagraph (FM) based stock market decision making, classification and prediction are proposed for short term investors of Indian stock market. Simple Moving Average (SMA), Exponential Moving Average (EMA), Moving Average Convergence Divergence (MACD) and Relative Strength Index (RSI) are some of the Technical Indicators which are used as input to train the system which is integrated with Fuzzy Metagraph. This approach of incorporating FM with SMA, MACD and RSI would be a new attempt in classification and prediction on share market investment. Stocks listed in Bombay Stock Exchange (BSE) in India are used to evaluate the performance of the system. The results obtained from the proposed FM based model are found to be satisfactory with very low risk error.

Keywords: FIS; Fuzzy Metagraph; Stock Market Classification; Stock Market Prediction; RSI; MACD; SVM;

*Corresponding author. Tel.: +0-000-000-0000; fax: +0-000-000-0000. E-mail address: thsa07@gmail.com, thirunavukarasuphd@gmail.com.
Introduction

Over the past decades, a number of fuzzy graphs have been used to represent uncertain relationships between fuzzy elements or sets of fuzzy elements. However, existing fuzzy graphs are not capable of effectively modelling the directed relationships between sets of fuzzy elements. Fuzzy Metagraph (FM) is an emerging technique used in the design of many information processing systems like transaction processing systems, decision support systems, and workflow Systems [13]. When a FM said to be dynamic a constantly changing stream of data.

Financial markets are highly volatile and generate huge amount of data on a day to day basis. Stock market prediction has always been one of the hottest topics in research, as well as a great challenge due to its complex and volatile nature. The Timing Problem to buy low and then sell high is a non-trivial problem and it is considered a dream of each investor. To solve this problem, the prediction of stock prices trend is important to be considered.

A trader decides about what share to trade, when to trade, using either fundamental or technical analysis. Fundamental analysis use information about economic condition of the nation and world, financial state and management of the company’s involved, current political and weather information. Technical analysis is a method of analyzing stock price movement by using historical price information to predict future price of a stock. There are many way of analyzing stock movement. Technicians also look for relationships between price/volume indices and market indicators. Technical analysis involves predicting stock price using technical indicators like SMA, EMA, RSI and MACD [4, 11].

Soft computing techniques are widely applied to stock market problems. They offer useful tools in forecasting noisy environments like stock markets, capturing their non-linear behaviour. Utilizing intelligent systems such as neural networks, fuzzy systems and genetic algorithms for the purpose of prediction in the field of finance has extensive applications. Lately, artificial neural networks (ANNs) and support vector machines (SVMs) have been successfully applied to solve the problems of predicting financial time series, including financial stock market prediction. Although Adaptive Neuro-Fuzzy Inference System (ANFIS) has been applied in several studies, few of these have contributed to research in the financial area. The method of Support Vector Classification can be extended to solve regression problems. This method is called Support Vector Regression (SVR). SVR has been applied to stock market forecasting problems.

1. Technical Indicators of Stock market and Fuzzy Metagraph

Technical indicators of stock market

Technical indicators are used to develop the quality of the data, improve the efficiency and simplicity the stock prediction process. The two most popular types of moving averages are the Simple Moving Average (SMA) and the Exponential Moving Average (EMA).

A SMA is calculated by adding the security's prices for the most recent "n" time periods and then dividing by "n" .This calculation is done for each period in the chart.

$$EMA = Price (t) \times k + EMA (y) \times (1-k)$$

Where $t$ represents today, $y$ denotes yesterday, $N$ is the number of days in EMA and $k = 2 / (N+1)$.

MACD indicator tries to forecast market trends by comparing short and long-term tendencies. It is the difference between a security's 26-day and 12-day EMAs.

$$MACD = 26 \text{days EMA} - 12 \text{days EMA}$$

A 9-day exponential moving average, called the "signal" (or "trigger") line is plotted on top of the MACD to show buy/sell opportunities. If MACD is above the signal line then buy. If MACD is below the signal then sell.

The Relative Strength Index (RSI) considers whether an asset is over bought or oversold. RSI is describes in following equation

$$RSI = \frac{100-(100/1+RS)}{Average \ gain/Average \ loss}$$

Where, $RS = Average \ gain/Average \ loss$

Simple 14 period averages are used as initial value for average gain and average loss. For subsequent values the following formula is used

Average Gain = [(previous Average Gain) x 13 + current Gain] / 14.

Average Loss = [(previous Average Loss) x 13 + current Loss] / 14.
If RSI increases to above 70 (implies overbought) then sell. If RSI is between 30 and 70 (implies normal) then hold. If RSI decreases to below 30 (implies oversold) then buy.

1.1. Fuzzy Inference System (FIS) and Fuzzy Metagraph (FM)

The main components of Fuzzy Inference System are a fuzzification interface, a fuzzy rule base (knowledge base), an inference engine (decision-making logic), and a defuzzification interface. Fuzzy if-then rules and fuzzy reasoning are the backbone of fuzzy inference systems, which are the most important modeling tools based on fuzzy set theory. The basic FIS can take either fuzzy inputs or crisp inputs, but the outputs it produces are always fuzzy sets. In fuzzification stage crisp values of input variables are converted into linguistic fuzzy variables. In defuzzification stage fuzzy variables are converted to crisp variables. These crisp values are used in decision making process [2, 10].

A Fuzzy Metagraph \( \tilde{S} \) can be defined as a triple \( \{ X, \tilde{X}, \tilde{E} \} \), here \( X \) is its generating set where \( \tilde{X} \) is a fuzzy set on \( X \) and \( \tilde{E} \) is a fuzzy relation on \( X \times X \). A fuzzy set \( X \) on \( X \) is completely characterized by its membership function \( \mu: X \rightarrow [0, 1] \) for each \( x \in X \), \( \mu(x) \) is the truth value of the statement of “\( x \) belongs to \( \tilde{X} \)”. \( \tilde{E} \) is a fuzzy edge set \( \{ \tilde{e}_m, m=1, 2, 3, \ldots, m \} \). Figure 1 shows the FM whose element set is \( X = \{ \tilde{X}_1, \tilde{X}_2, \tilde{X}_3, \tilde{X}_4, \tilde{X}_5 \} \) is known as fuzzy Meta Node and whose edge set consists of: \( \tilde{e}_1 = \{ \tilde{X}_1, \tilde{X}_2 \}, \{ \tilde{X}_3 \} \}, \tilde{e}_2 = \{ \tilde{X}_2 \}, \{ \tilde{X}_4 \} \} \}, \{ \tilde{X}_5 \} \}. \) The knowledge of problem will be stored in the computer memory or knowledge database of Expert system. The Expert System will use inference rule-based techniques that are known as goal driven and data driven reasoning methods. FM is an effective modelling and analysis tool for rule based systems. Data management and model management are major functions of a DSS. Fuzzy metagraphs provide a better foundation for decision analysis, and also facilitate more effective design of DSS [13].

2. FM Based Classification and Prediction of Stock Market Data

The overall framework of the proposed model is illustrated as Fig. 2 and three major phases are provided. To detail the proposed model, each process of the proposed model is described as follows.

Step 1: Collect experimental datasets
Step 2: Data transformation and select essential technical indicators
Step 3: Classification and Prediction of Stock Market based on Fuzzy Metagraph

Stock data are collected from the website and the dataset has been collected from popular Indian companies like Tata consultancy service (TCS) and Reliance Industry Limited (RIL) from January 2011 to December 2012. The year 2011-2012 had been very challenging year for Indian share market. In this study stock information for that period is taken to analyze the performance of the system at hard times.
Stocks listed in Bombay Stock Exchange (BSE) are used to evaluate the system. For experimentation, the stock market datasets are divided into two sets such as: (1) training dataset and (2) testing dataset. 80% of the daily data were used for the training dataset and the remainder of the daily data was used for the testing dataset [10, 11].

3. **Experimental results and performance comparison**

In the FM module of the proposed model, a four layer has been chosen, which is completely linked architecture that has been shown in Fig. 3. The value of every variable is described by one of the feasible five fuzzy membership sets (VL L M H VH) which means very low (VL), low (L), medium (M), high (H), and very high (VH). Fuzzy membership functions may be crisp data observations about the membership degree of these fuzzy sets. Finally, resultant output has a single floating-point value from the set {0.0, 0.25,...,1.0}. As a result, the total number of probable if–then rules for fuzzy system. The output of the FM system has five unique fuzzy sets i.e. Strong Buy, Buy, Hold, Sell, Strong Sell that are defined. A Strong Buy signal is created when the output is almost to 1.0 and a Strong Sell signal is created when the output is almost to 0.0 from the above-mentioned case.

The preliminary fuzzy rules are measured by the prediction system with technical indicators generating the following a few fuzzy metagraph rules.

1. If \( \tilde{X}_4 \) and \( \tilde{X}_5 \) and \( \tilde{X}_{10} \) and \( \tilde{X}_{15} \) then \( \tilde{X}_{20} \) (Rating \( \tilde{e}_1 = 0.00 \))
2. If \( \tilde{X}_4 \) and \( \tilde{X}_6 \) and \( \tilde{X}_{14} \) and \( \tilde{X}_{16} \) then \( \tilde{X}_{21} \) (Rating \( \tilde{e}_2 = 0.25 \))
3. If \( \tilde{X}_4 \) and \( \tilde{X}_7 \) and \( \tilde{X}_{12} \) and \( \tilde{X}_{17} \) then \( \tilde{X}_{22} \) (Rating \( \tilde{e}_3 = 0.50 \))
4. If \( \tilde{X}_4 \) and \( \tilde{X}_8 \) and \( \tilde{X}_{13} \) and \( \tilde{X}_{18} \) then \( \tilde{X}_{23} \) (Rating \( \tilde{e}_4 = 0.75 \))
5. If \( \tilde{X}_4 \) and \( \tilde{X}_9 \) and \( \tilde{X}_{14} \) and \( \tilde{X}_{15} \) then \( \tilde{X}_{24} \) (Rating \( \varepsilon_5 = 1.0 \))

Where \( \tilde{X}_1 = \text{SMA is Very High} \), \( \tilde{X}_2 = \text{SMA is High} \), \( \tilde{X}_3 = \text{SMA is Medium} \), \( \tilde{X}_4 = \text{SMA is Low} \).
\( \tilde{X}_5 = \text{EMA is Very High} \), \( \tilde{X}_6 = \text{EMA is High} \), \( \tilde{X}_7 = \text{EMA is Medium} \), \( \tilde{X}_8 = \text{EMA is Low} \), \( \tilde{X}_9 = \text{EMA is Very Low} \).
\( \tilde{X}_{10} = \text{MACD is Very High} \), \( \tilde{X}_{11} = \text{MACD is High} \), \( \tilde{X}_{12} = \text{MACD is Medium} \), \( \tilde{X}_{13} = \text{MACD is Low} \), \( \tilde{X}_{14} = \text{MACD is Very Low} \).
\( \tilde{X}_{15} = \text{RSI is Very High} \), \( \tilde{X}_{16} = \text{RSI is High} \), \( \tilde{X}_{17} = \text{RSI is Medium} \), \( \tilde{X}_{18} = \text{RSI is Low} \), \( \tilde{X}_{19} = \text{RSI is Very Low} \).
\( \tilde{X}_{20} = \text{Strong Sell} \), \( \tilde{X}_{21} = \text{Sell} \), \( \tilde{X}_{22} = \text{Hold} \), \( \tilde{X}_{23} = \text{Buy} \), \( \tilde{X}_{24} = \text{Strong Buy} \).

\( \varepsilon_1 \) represents the rule \( \tilde{X}_1 \tilde{X}_5 \tilde{X}_{10} \tilde{X}_{15} \rightarrow \tilde{X}_{20} \), where “\( \wedge \)" denotes conjunction and “\( \rightarrow \)" denotes implication.

---

Fuzzy Metagraph based Decision and Prediction for dynamic stock exchange databases have been constructed by using weighted fuzzy production rules. In these rules, a weight parameter can be assigned to each proposition in the antecedent of a fuzzy production rule and Certainty Factor (CF) can be assigned to
each rule. CF can be calculated by using some important variables in dynamic stock market. In Proposed System, FM learning algorithm is used to predict the future prices of stock. Figure 3 shows that Proposed System Architecture for stock market prediction and classification technique.

![Figure 4: SVM based Stock market Classification result.](image1)

Figure 4: SVM based Stock market Classification result.

![Figure 5: Stock market Classifier tree result.](image2)

Figure 5: Stock market Classifier tree result.

The main objective of the paper is to build a prediction system to predict the future occurrence of an event. If the prediction is based on past prices only, then the strategy for buy, sell and hold follows five simple rules:

1. If the predicted trend is positive and the share has been bought, then hold.
2. If the prediction is negative and the share hasn’t been bought, then do nothing.
3. If the prediction is positive and the share hasn’t been bought, then buy.
4. If the prediction is negative and the share has been bought, then sell.
5. All transaction take place at the end of each trading day, thus the closing price is assumed to be the trading price.

News contents are one of the most important factors that have influence on market. If the prediction is based on news only, then the above plans are no longer applicable. This is due to the fact that news articles for each day can only be obtained in the morning before stock market opens. Therefore, changes are made to accommodate for the late arrival of news input and it is summarized as follows:

Assuming the overall trend of the stock is rising. News is assumed to have valid influence on the stock only on the same day it is published.

1. If the news prediction is positive and the share hasn’t been bought, then buy at opening price.
2. If the prediction is positive and the share has been bought, then hold.
3. If the prediction is negative and share hasn’t been bought, buy at closing price.
4. If the prediction is negative and the share has been bought, sale at opening price.
5. If the news prediction is absent, then buy at closing price.

A buy signal from the expert system is used to buy stock and sell signal is used to sell that stock. Only short or long type investment is used to assess the performance of the system. Forecasting performance of different classification methods is shown in table 1. Comparison of classification rate for two machine learning model is shown in table 2.

Table 1. Forecasting performance of different classification methods

<table>
<thead>
<tr>
<th>Classification Method</th>
<th>Hit Ratio (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Walk (RW) Model</td>
<td>50</td>
</tr>
<tr>
<td>Neural Networks</td>
<td>69</td>
</tr>
<tr>
<td>SVM</td>
<td>73</td>
</tr>
<tr>
<td>FM</td>
<td>75</td>
</tr>
</tbody>
</table>

This study developed a novel hybrid model based on a FM and FIS for stock market price forecasting. The proposed model performs best based on two error measures, namely Root Mean Square Error (RMSE) and Mean Magnitude Relative Error (MMRE). The RMSE and MMRE error comparison with a different classification model is shown in Table 3. Figure 4 show that SVM based Stock market Classification technique. Classifier tree is shown in figure 5. FM based Prediction result has been shown in figure 6. Graphs in Figure 7 show that FM is able to follow the trend of the target prices.

![Graph showing stock market prediction](image)

Figure 6: FM based Stock market prediction result.

Table 2. Comparison of classification rate for machine learning model.

<table>
<thead>
<tr>
<th>Machine Learning Model Algorithm</th>
<th>Misclassification Rate</th>
<th>Correct classification Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>FM</td>
<td>0.5611</td>
<td>0.4389</td>
</tr>
<tr>
<td>SVM</td>
<td>0.7426</td>
<td>0.2574</td>
</tr>
</tbody>
</table>

Table 3. Comparison of prediction techniques for data of RIL.

<table>
<thead>
<tr>
<th>Prediction Technique</th>
<th>RMSE</th>
<th>MMRE</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>0.0171</td>
<td>0.0228</td>
<td>97.50</td>
</tr>
<tr>
<td>FM</td>
<td>0.0139</td>
<td>0.0176</td>
<td>98.97</td>
</tr>
</tbody>
</table>
6. Conclusion

Stock market prediction and classification are important and of great interest because successful prediction of stock prices may promise attractive benefits. These tasks are highly complicated and very difficult. We have proposed a FM approach for stock market price forecasting. Experimental results, obtained by running on datasets taken from BSE India, show that our method performs better than other methods. This research is just a beginning and the long term goal is to predict the trend of the price variation by including various influential factors such as technical analysis and fundamental analysis. As a result, the system can be further applied for the daily trading purpose. In the future works, the model can be verified by other stock market such as China, Japan and Hong Kong.

References