



A Survey on Stock Forecasting Model Based on Combined Fuzzy-Genetic Algorithm

Ajeeta Sachdeva, Vivek Sharma

Department of Information Technology,
SATI, Vidisha, India

Abstract: *This paper surveys recent literature in the area of Data Mining, fuzzy time series, genetic algorithm and Fuzzy-Genetic system used to predict the stock market fluctuation. Fuzzy time series and Fuzzy-Genetic systems are identified to be the leading machine learning techniques in stock market index forecasting area. The conventional techniques are not cover all the possible relation of the stock price fluctuations. There are new approaches to acknowledged in detail of an analysis of stock price variations. Fuzzy time series models and Genetic algorithm can be used exclusively in the finance markets and forecasting of stock price.*

Keywords: *Data Mining, Stock Market forecasting, fuzzy time series, Fuzzy-Genetic Systems, Forecasting.*

I. INTRODUCTION

One of the important tools for development of any country is empowerment financial markets. Economic growth needs efficient capital market to allocate optimal capital resources. Forecasting future of market conditions and adopting proper methods to obtain more profits from the market is one of the main needs for industries and people. One of the tools for forecasting is time series and fuzzy time series.

A time series is a set or sequence of observed data arranged in chronological order and in an equally spaced time intervals such as daily or hourly air temperature. Time series occurs in many fields and the analysis of time series has got a wide application in areas like process control, economic forecasting models, marketing strategy, various population studies, biomedical science and many more areas. Time series analysis uses systematic approaches to extract information and understand the characteristics of a physical system that creates the time series. There are a number of different approaches to deal with time series analysis including dynamic model building and performing correlations.

II. STOCK MARKET

The stock markets of all countries need some sort of prediction for buyers. Moreover investing in stock markets is an exciting challenging and monetary activity that has influence on the economy of individuals and countries. For forecasting stock prices initially traditional statistical techniques were used. Several regression models were also used by researchers for forecasting stock prices. These techniques can predict merits and demerits of the stock prices. However, these techniques call for historical data and with normal distribution of data for accurate prediction. The traditional forecasting methods fail in forecasting linguistic data. In order to deal with the linguistic historical data the conventional time series methods has failed, therefore an alternative forecasting method known as fuzzy time-series is used.

III. FUZZY TIME-SERIES

Fuzzy time-series methods are efficient as they can also work with linguistic values with maximum accurateness. Though it works well, many other models came into existence. They are explored in enrollment applications stock index, tourism demand road accidents and weather. Many researchers developed their fuzzy time series models that can effectively deal with stock market forecasting. Some researchers used a element known as length of linguistic intervals. These models are good for forecasting. However, theory was linguistic in nature as Millers said. The current models comprise first order models, high order models, genetic algorithms, hybrid models bivariate models, adaptive expectation models, and neural network models. Out of them the first order fuzzy time-series models are highly useful and suitable for analyzing stock markets.

Fuzzy theory was originally developed to deal with problems involving linguistic terms (L.A. Zadeh,1976;). Time-series models had failed to consider the application of this theory until fuzzy time-series was defined by Song and Chissom (Q. Song, B.S. Chissom, 1993; 1994). They proposed the definitions of fuzzy time-series and methods to model fuzzy relationships among observations. The framework of Song and Chissom's model includes six steps: (1)Define and partition the universe of discourse; (2) define fuzzy sets for the observations;(3)fuzzification of observation, (4) establish the fuzzy relationship, R, (5) forecast, and (6) defuzzification of forecasting results (Q. Song, B.S. Chissom, 1993; 1994).

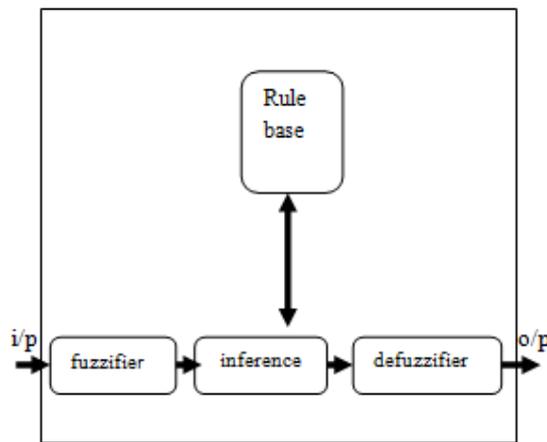


Figure 1. Structure of fuzzy logic system

IV. LITERATURE REVIEW

The fuzzy time series models were pioneered by Zadeh, and further improved by Song and Chissom, used fuzzy relational equations, defined different order of fuzzy time series models, defined important properties of fuzzy time series, and presented a step-by-step procedure for the implementations of a fuzzy time series with linguistic variables.

Song and Chissom [1] subsequently published an application of their fuzzy time series model, a 1993 study in which they forecasted enrollments for the university of Alabama, and developed a first order model from historical data of enrolment. From which they proposed a bit by bit forecasting procedure. In comparison with other similar studies in the literature, Song and Chissom verified that their fuzzy time series model had smaller average forecasting error than that of other approaches.

In the evolution of fuzzy time-series models, Chen, (1994) presented a model in which he applied simplified arithmetic operations in forecasting algorithms whereas Song and Chissom presented complicated maximum–minimum composition operations, in their model. [2] In later study, Chen presented various methods, such as genetic algorithms and high-order fuzzy time series models to improve his first model (S.M. Chen, C.C. Hsu, 2004; S.M. Chen, N.Y. Chung, 2006; J.R. Hwang, et al.,1998).

Hwang [3] et al. (1998) used the differences of the available historical data as fuzzy time series instead of direct usage of raw numeric values. Sah and Degtiarev also used a similar approach in (Sah & Degtiarev, 2005). Huarng improved the forecasting accuracy based on the determination of effective length of intervals in time series (Huarng, 2001a) and heuristic approaches.

Shyi-Ming and Jeng-Ren Hwang [4] proposed two algorithms for temperature prediction using two – factors time variant fuzzy time series. The daily average temperature was taken as the first factor and the daily cloud density as the second factor. They considered the historical data of the daily average temperature from June 1996 to September 1996 in Taipei as main factor and historical data of daily cloud density as second factor to predict the temperature in future.

QiSen Cai, Defu Zhang, Bo Wua, Stephen C.H. Leung [5] proposed a paper in which, a hybrid model FTSGA based on genetic algorithm and fuzzy time series was proposed FTSGA improved the performance by applying the various genetic algorithm operations which is known as selection, crossover and mutation applied iteratively to search a good discourse partition. They selected the TAIEX data set for the experiment. And the results of the experiment showed that comparing with other models based on fuzzy time series FTSGA can greatly reduce the root mean square error and improve accuracy.

The partition of the universe of discourse plays an important role in improving the forecasting results of fuzzy time series model. In order to find better partition of universe. To improve the results of the fuzzy time series models, they found better partition of universe with the help of a hybrid method combined with genetic algorithm was proposed in this paper. Genetic algorithm is a well-known search heuristic that mimics the process of natural evolution. This heuristic approach is widely used to generate useful solutions to optimization and search problems including the partition problem in fuzzy time series. Their model FTSGA used genetic algorithm to achieve certain better results compared with existing conventional fuzzy time series models by obtaining the suitable partition of universe. They used the daily TAIEX closing prices covering the period from 1990 to 1999 as the experimental dataset, and its performance was compared with the existing traditional fuzzy time series models. They used the data of each year from January to October to perform the estimation while those for November and December are used for forecasting. The root of mean squared error (RMSE) is employed as evaluation criterion for the forecasting performance of the planned model and the similar models.

Mustafa demetgul, Osmin Ulkir, Tayyab Waqar proposed a paper in which, fuzzy logic control for a very familiar household, i.e. washing machine, has been developed [6]. The washing machine has been modeled based on the parameters regarding washing of cloths. To model a system using fuzzy logic, we have to first determine the inputs and output parameter of it. The important duty of washing machine is to clean the cloths without damaging them. In order to achieve it, the output parameters of fuzzy logic, which are related to washing, must be paid importance.

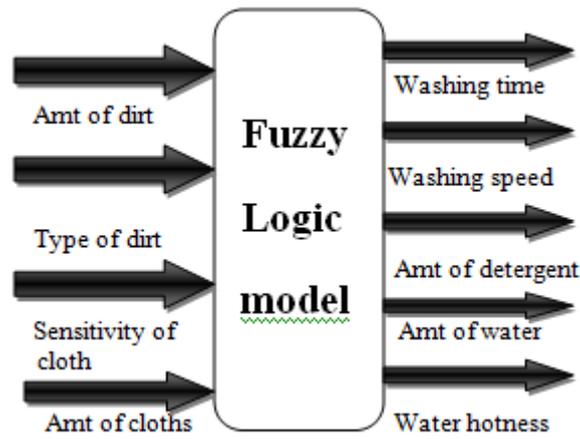


Figure 2. Inputs and outputs of the system

Shashikumar G.Totad, G.suresh [7] developed a method to forecast stock prices considering different parameters. In this article they considered the problem of forecasting stock prices. To forecast, they implemented a hybrid algorithm to improve the prediction performance of the forecasting procedure. The process adopted by the algorithm was equal frequency division. The name of algorithm was fast Fourier transform algorithm. In order to improved the forecasting accuracy on predicting stock price the paper presented a new methodology and adopts equal frequency division to achieved reliable linguistic parameter definition which was lacked by the earlier models which adopted equal distance division method.

Samira M. Boaisa, and Saleh M. Amaitik presented a forecasting model [8]. The main purpose of this paper has been to use fuzzy time series technique for forecasting problems based on average lengths of intervals. It has been successfully implemented to the forecasting of the average calendar day. In this paper the results obtained shows the effectiveness of the proposed model compared to previous works in accuracy and effortlessness. Using and analyzing of fuzzy time series that uses linguistic values enhances the power of decision making systems, thus introduces more accuracy than traditional methods.

Li and Cheng proposed a deterministic forecasting model to achieve the issues of controlling uncertainty in forecasting, effectively partition the intervals, and constantly achieving forecasting accurateness with different interval lengths. Moreover their model followed the consistency principle that a shorter interval leads to more perfect results.

Kunhuang Huarng, Member, IEEE, and Tiffany Hui-Kuang Yu [9] proposed a paper in which objective of study is to explore ways of determining the useful lengths of intervals in fuzzy time series. The Ratio based lengths of intervals are, therefore, expected to get better fuzzy time series forecasting. The experimental analysis suggests that the ratio-based lengths of intervals can also be used to improve fuzzy time series forecasting. In terms of the RMSE, the ratio-based lengths of intervals were found to provide the smallest rank in the enrollment forecasting, the smallest total rank in the stock index forecasting, and the smallest rank in the inventory forecasting as well.

Eren Bas, Vedid Rezlan Uslu, Ufuk Yolsu proposed a method contributes to the fuzzification stage by using a modified genetic algorithm (MGA) [10]. Fuzzy time series methods are used when observations of time series holds uncertainty. Furthermore, these techniques do not require the assumptions needed for conventional time series approaches. In general fuzzy time series methods consist of three stages, namely first one is fuzzify the crisp input, second one is determination of fuzzy relations, and the final one is defuzzify the output into crisp. Artificial intelligence algorithms are normally used in these stages with genetic algorithms being the most popular of these algorithms owing to their rich operators and good performance. However, the mutation operation of a GA may cause some negative results in the solution set. Thus, they proposed a modified genetic algorithm to find optimal interval lengths and control the effects of the mutation operator. The consequences of applying our new approach to real datasets show superior forecasting performance when compared with those obtained by other techniques. The proposed method contributes to the fuzzification stage by using a modified genetic algorithm (MGA). In particular, their aim to check the negative effects of the mutation operation so as to obtain better forecasts furthermore, by using MGA they avoid subjective judgments in determining the intervals, and to obtain more sensible results, they find dynamic instead of fixed length intervals.

Prof. Sharadindu Roy & Uttam Kumar Panja, [11] presented a method to solve real world problem. In this paper genetic algorithm was used to solve the travelling salesman problem (TSP). They applied genetic algorithm operator to this problem and find the solution. The solution obtained of the traveling salesman problem is global optimum. In this problem number of cities and distances are given between them (cities). A Traveling salesman has to trip all of them. The salesman will begin from city and after visiting the remaining cities (exactly once) he will reach to his starting position.

Wu Xiao-qin proposed a prediction model. [12] In this paper, the financial time series analysis and forecasting as the main case study to the theory of soft computing technology framework that focuses on the fuzzy theory and genetic algorithms (FGA) as a method of combination. The economic time series forecasting model based on genetic algorithm and fuzzy time series was built. They used index cards of Sheng Zheng as an example. The experimental results show that FGA performs much better than BP neural network not only in the precision but also in the searching speed. The hybrid algorithm had a strong feasibility and superiority.

Financial time series data contains a wealth of system information. The time series analysis is to discover the hidden knowledge (relationships, rules, trends, etc.). The causal relationship is one of the most important one that find out market trends based on historical sequence. Fuzzy logic construct the fuzzy domain of the main division of the initial interval. The initial interval find out the division of fuzzy functions and forecast the prospect data. The finest length of the past range was designed based on genetic algorithm. The suitable fitness function was chosen to estimate the most favorable range of the best fitness division. So as to achieve a specific value of the fuzzy. The route of the algorithm is

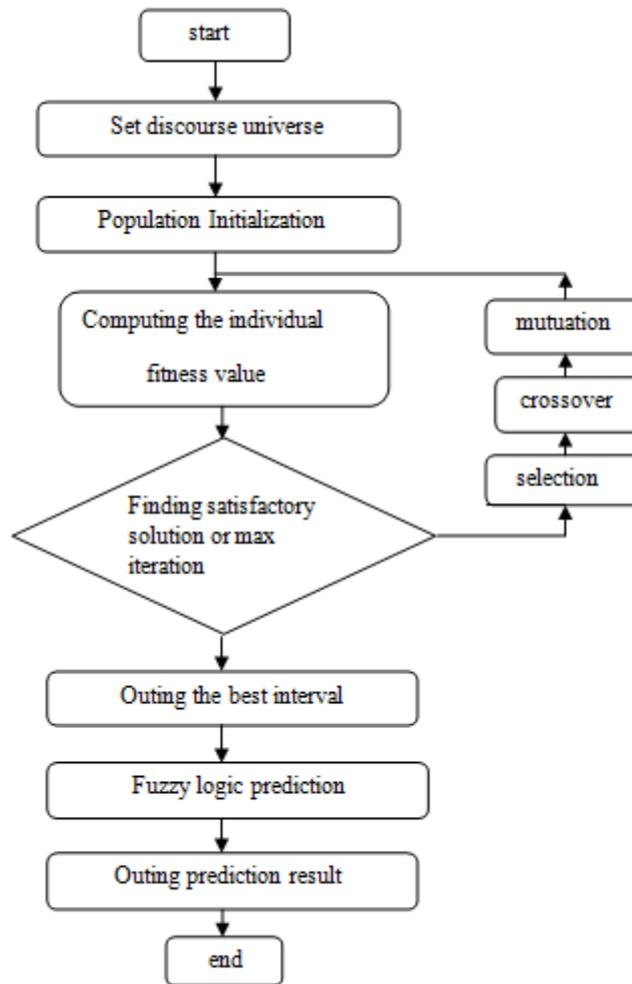


Figure 3. Flowchart of algorithm

In this paper, they combined the two data mining algorithms, they make use of genetic algorithms to achieve optimal solution and fuzzy logic algorithm to solve things in the direction of the development of the application of the rules of analysis of financial time series forecast problem, and contrast with other algorithms. A predictive model was developed, predictive accuracy was enhanced, and more stable results were achieved.

Manjul Saini, A.K.Singh, used a soft computing technology in the forecasting model. The key purpose of this article was to study the use of soft computing in the field of Stock Exchange Market and meteorology. Neural Network with well developed back - propagation algorithm was used to forecast the daily stock market returns and weather condition in dehradun (U.K.). In this paper they had composed stock market exchange records sets and weather records set of Dehardun. In Dehradun, Uttrakhand, the months of April to November are identified as the rainfall season with May, June, July, August, and October as the main monsoon seasons. In this paper the study used the data of these months from 1975 to 2002. The average Humidity and the average Wind Speed, Average Temperature etc are used as input parameters. The data taken was retrieved from www.Indiastat.com and the IMD website. The unidentified values were randomized keeping in mind the average value of the data.

With the advancement of technology the Neural Network tools were designed with new and efficient algorithm which in turn increases its competence, scalability, helpfulness to predict, classify the unseen data. This study focuses on the highly developed Backpropagation algorithm learning method .The Advanced Back propagation algorithm seeks to minimize the error term between the output of the neural net and the actual desired output value. The inaccuracy term was considered by comparing the net output to the desired output and was then feedback through the network causing the synaptic weights to be changed in an effort to reduce error. The process was continual until the inaccuracy reaches a lowest value.

GaneshV.Kumbhar, RajeshV.Argiddi developed a model which works in two phase [14]. In first phase they created artificial neural network & be trained it with stock market historical data. To efficiently learn the pattern of stock market

Genetic algorithm was used to optimize the weights of neural network. In the second phase they used the learned neural network to predict the future nature of stock market that is whether it goes up, down or it remains steady for next few days.

To initialize the population of Genetic algorithm the weights of neural network were sent to algorithm as an input. After that Genetic algorithm applied the selection, crossover & mutation on population & returns the finest population with weights of NN. In order to achieve the desired output the optimized weights were set to Neural network & then it was tested. If desired output has found then it stops Genetic Algorithm otherwise GA continues until termination state was occurred. The second phase of projected methodology was the prediction of future with the help of neural network which was saved when we got the desired output. The goal of the paper was to raise the capitulate from the investment. They had proposed the hybrid neuro-genetic technique to predict stock market price. Genetic algorithm was used for optimization of weights of neural network & NN was used for prediction.

V. CONCLUSION

In this paper our aim is to forecast the stock market with the help of hybrid algorithm based on Fuzzy time series and Genetic algorithm. It was found that genetic algorithm and fuzzy logic have performed well in forecasting process. We can consider the different parameters in terms of error as mean square error, Root mean square error or average forecasting error to calculate the forecasting accuracy of the model. Lower the error, better the forecasting method is.

REFERENCES

- [1] Song Q and Chissom BS. "Fuzzy time series and its models". Fuzzy Sets and Systems, 1993, (54), 269-277.
- [2] Chen, S.M., C.C. Hsu, "A new method to forecast enrollments using fuzzy time series" Appl. Sci Eng., 2: 234-244, 1994.
- [3] Huarng, K. Effective lengths of intervals to improve forecasting in fuzzy time series. Fuzzy Sets and Systems , 123,387-394, 2001.
- [4] Chen SM and Hwang JR. "Temperature Prediction using fuzzy time series". IEEE Transactions of Systems, Man, Cybernetics Part : Cybernetics, 2000, vol. 30, no. 2,263-275.
- [5] QiSen Caia, Defu Zhanga, Bo Wua, Stephen C.H. Leung " A novel stock forecasting model based on fuzzy time series and genetic algorithm". International Conference on Computational Science. Procedia Computer Science18 (2013) 1155 – 1162. Elsevier Sciencedirect.
- [6] Mustafa Demetgul, Osman Ulkir, Tayyab Waqar "Washing machine using fuzzy logic Automation, Control and Intelligent Systems; 2(3):27- 32, 2014.
- [7] Shashikumar G.Totad, G.Suresh "A novel Fuzzy time-series based forecasting of Stock price". International Journal of Soft Computing and, Artificial Intelligence, ISSN: 2321 404X, Volume-1, Issue-2 Nov-2013.
- [8] Samira M.Boaisha and Saleh M.Amatik "Forecasting Model based on fuzzy time Series".
- [9] Huarng K and Yu TH. "Ratio-Based lengths of intervals to improve fuzzy time series forecasting". IEEE Transactions on Systems, Man, Cybernetics Part B: Cybernetics, 36(2), 328-340, 2006.
- [10] Eren Bas, Vedide Rezan Uslu, Ufuk Yolcu, Erol Egrioglu "A modified genetic algorithm for forecasting fuzzy time series". Journal Applied Intelligence, Volume 41 Issue 2, September 2014.
- [11] Prof. Sharadindu Roy, Uttam Kumar Panja, Nikhil Kumar Sardar "Efficient technique to solve Travelling Salesman Problem using Genetic Algorithm". International Journal of Computer Engineering and Technology, Volume 5, Issue 3, March (2014), pp. 132-137
- [12] Wu Xiao-qin "Research on Prediction Model of Time Series Based on Fuzzy Theory and Genetic Algorithm" International Conference on Medical Physics and Biomedical Engineering. Physics Procedia 33 (2012) 1241-1247. Elsevier sciencedirect
- [13] Manjul Saini, A.K.Singh "Forecasting Stock Exchange Market and Weather Using Soft Computing". International Journal of Advanced Research in Computer Science and Software Engineering. Volume 4, Issue 5, May 2014.
- [14] Ganesh V. Kumbhar, Rajesh V. Argiddi "Hybrid Neuro-Genetic Data Mining Technique for Market Forecasting" Indian Journal of Applied Research Volume : 5 | Issue : 2 | Feb 2015 ISSN 2249-555X.