Expert Systems with Applications 37 (2010) 4742–4748

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Adapted Neuro-Fuzzy Inference System on indirect approach TSK fuzzy rule base for stock market analysis

Akbar Esfahanipour, Werya Aghamiri *

Department of Industrial Engineering, Amirkabir University of Technology (Polytechnic of Tehran), Tehran, Iran

ARTICLE INFO

Keywords:
FCM clustering
Fuzzy Rule Based Systems
Forecasting
Stock market
Forecasting accuracy
ANFIS

ABSTRACT

Nowadays because of the complicated nature of making decision in stock market and making real-time strategy for buying and selling stock via portfolio selection and maintenance, many research papers has involved stock price prediction issue. Low accuracy resulted by models may increase trade cost such as commission cost in more sequenced buy and sell signals because of insignificant alarms and otherwise bad diagnosis in price trend do not satisfy trader's expectation and may involved him/her in irrecoverable cost. Therefore, in this paper, Neuro-Fuzzy Inference System adopted on a Takagi–Sugeno–Kang (TSK) type Fuzzy Rule Based System is developed for stock price prediction. The TSK fuzzy model applies the technical index as the input variables and the consequent part is a linear combination of the input variables. Fuzzy C-Mean clustering implemented for identifying number of rules. Initial membership function of the premise part approximately defined as Gaussian function. TSK parameters tuned by Adaptive Nero-Fuzzy Inference System (ANFIS). Proposed model is tested on the Tehran Stock Exchange Indexes (TEPIX). This index with high accuracy near by 97.8% has successfully forecasted with several experimental tests from different sectors.

1. Introduction

The price variation of stock market is a dynamic system and the chaotic behavior of the stock price movement duplicates complication of the price prediction; however, the highly non-linear, dynamic complicated domain knowledge inherent in the stock market makes it very difficult for investors to make the right investment decisions promptly. It is necessary to develop an intelligent system to get real-time pricing information, reduce one obsession of investors and help them to maximize their profits.

There are two common approaches has used for market analysis which are fundamental analysis and technical analysis. A fundamental analysis involved some statistics of the macroeconomics data as well as the basic financial status of the company. After taking all these factors into account, the analyst will then make a decision of selling or buying a stock. Another approach is based on the historical financial time series data called technical analysis. However, financial time series show quite complicated patterns (for example, trends, abrupt changes, and volatility clustering) and such series are often non-stationary, whereby a variable has no clear tendency to move to a fixed value or a linear Trend (Chang & Liu, 2008).

Stock price prediction has always been a subject of interest for most investors and professional analysts. Nevertheless, finding out the best time to buy or to sell has remained a very difficult task because commission (remuneration for services rendered) is commonly overlooked when doing research relating to stock market prediction; however, if any model is actually implemented it is going to incur fees which could greatly affect the profit predicted by the model. Chen and Linkens (2004) considers three different levels of commissions and how it would affect the best buying strategy used by Investors. Tehran Stock Exchange has evolved into an exciting and growing marketplace where individual and institutional investor trade securities of over 420 companies. TEPIX is a weighted market value all share prices appearing on the TSE Price Board. TEPIX calculation method is as follows:

$$TEPIX = \frac{\sum P_{it} C_{it}}{\sum P_{ib} C_{it}} (Base - Value),$$

where P_{it} and P_{ib} represent share price of company i, respectively at time t and at the close of trading on March 21st, 1990; and C shows the total number of shares. And other important index in this market contain Industrial Index, Financial Index, Top 50 Companies Index. The remaining sections of this paper are organized as follows. Section 2 provides the literature review. Section 3 introduces the essential concepts of a type of fuzzy rule based systems called TSK and in its sub section issue of input variable and performance measurement has been considered. Section 4 describes ANFIS struc-

* Corresponding author. Tel.: +98 21 64542782.
E-mail addresses: Esfahaa@aut.ac.ir (A. Esfahanipour), Werya_Aghamiri@aut.ac.ir (W. Aghamiri).

0957-4174/$ - see front matter © 2009 Published by Elsevier Ltd.
doi:10.1016/j.eswa.2009.11.020

ture. Section 5 discusses the Fuzzy c-mean clustering algorithms. Section 6 presents the underlying research methodology. In Section 7 we can see experimentation design and results and Comparisons of different forecasting models. Finally, concluding remarks are given in Section 8.

2. Literature survey

It is too difficult to forecast stock price variation and the price fluctuation behaves more like a random walk and time varying. In a stock market, how to forecast stock prices accurately and find the right time to trade are obviously of great interest to investors. To reach this goal, two approaches can be employed: statistical and Artificial Intelligence (AI).

The statistics school includes autoregressive integrated moving average (ARIMA), generalized autoregressive conditional heteroskedasticity (GARCH) volatility (Franses & Ghijsels, 1999), and smooth transition autoregressive (STAR) (Sarantis, 2001). These models depend on the assumption of linearity among variables and normal distribution.

However, the assumption of linearity and normal distribution may not hold even though it has been shown successful in dealing with stock price movement in past decades. On the other hand, with greater in the stock market and the increasing need for more efficient forecasting models, the AI school, operating without the limitation of such an assumption, outperforms the conventional statistical methods experimentally (Enke & Thawornwong, 2005; Hansen & Nelson, 2002; Ture & Kurt, 2006; Zhang, 2003). During the last decade, stocks and future traders have come to rely upon various types of intelligent systems to make trading decisions. Recently, artificial neural networks (ANNS) have been applied to this area (Aiken & Bsat, 1999; Chang, Wang, & Yang, 2004; Chi, Chen, & Cheng, 1999; Kimoto & Asakawa, 1990; Lee, 2001; Yao & Poh, 1995; Yoon & Swales, 1991). These models, however, have their limitations because of the great noise and complex dimensionality of stock price data and besides, the quantity of data itself and the input variables may also middle in each other. Therefore, the result may not be very satisfactory.

Other soft computing methods are also applied in the prediction of stock price and these Soft Computing (SC) approaches are supposed to use quantitative inputs, like technical indices, and qualitative factors, like political effects, to automate stock market forecasting and trend analysis. Kuo, Chen, and Hwang (2001) uses a genetic algorithm base fuzzy neural network to measure the qualitative effects on the stock price. Variable selection is critical to the success of any network for the financial viability of a company. They applied their system to the Taiwan stock market. Aiken and Bsat (1999) use a FFNN (Feed Forward Neural Network) trained by a genetic algorithm (GA) to forecast three-month US Treasury Bill rates. They come to this conclusion that an NN can accurately predict these rates. Thammano (1999) used a neuro-fuzzy model to predict future values of Thailand's largest government-owned bank. He concluded that the neuro-fuzzy architecture was able to recognize the general characteristics of the stock market faster and more accurately than the basic back-propagation algorithm. Baba, Inoue, and Asakawa (2000) used NNs and GAs to construct an intelligent decision support system (DSS) for analyzing the Tokyo Stock Exchange Prices Indexes (TOPIX). The essential feature of their DSS was that it projected the high and low TOPIX values four weeks into the future and suggested buy and sell decisions based on the average projected value and the then-current value of the TOPIX. Kim and Han (2000) used a NN modified by a GA to predict the stock price index. In this instance, the GA was used to reduce the complexity of the feature space, by optimizing the thresholds for feature discretization, and to optimize the connection weights between layers. They concluded that the GA approach outperformed the conventional models. Abraham, Philip, and Saratchandran (2003) investigate how the seemingly chaotic behavior of stock markets could be well represented using several connectionist paradigms and soft computing techniques. To demonstrate the proposed technique, they analyzed the 7 year's Nasdaq-100 main index and 4 year's NIFTY index values. They concluded that all the connectionist paradigms considered could represent the stock indices behavior very accurately. Recently, therefore regarding to neural network's ability to model any given function and in other wise, fuzzy rule based systems mimics the crucial ability of the human mind to summarize data and focus on decision-relevant information, retaining good results by the proposed model is our expectation in this area.

3. A TSK type fuzzy rule based system

This research is aimed to predict the future price of a stock by the technical index input to the adapted TSK fuzzy model. For a first order Takagi–Sugeno model, a common rule is represented as follows: If x_1 is A_1, and x_2 is A_2, then $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ where x_1 and x_2 are linguistic variables and A_1 and A_2 are corresponding fuzzy sets and β_0, β_1, β_2 are linear parameters. Usually determining the linear parameters by using the least-squares algorithm and the membership function parameters are fine tuned using a neural network learning method. Initial rules are generated using the grid partitioning method. The various rule parameters during the TSK adaption process into ANFIS will be fine tuned. This research mainly studies the fluctuations of short-term stock prices and tries to develop a forecasting model using TSK type fuzzy rule based approach. The TSE index. are selected for studying purposes, and for more evaluation of the proposed model we test it on Tehran stock market index.

3.1. Input variables

There are so many factors that might effect stock prices. Many papers have dealt with input selection to mapping financial indexes and stocks. We can divide Inputs into two different types of inputs, financial and political (which tend to be qualitative). Quah and Srinivasan (1999) specified 5 key factors which will influence the stock price movement, that is yield, liquidity, risk, growth, and momentum. Izumi and Ueda (1999) stated that stock returns can be affected directly by macroeconomic factors such as inflation and short-term interest rate. As for the measure of system performance, Yao and Poh (1995) indicated an example that a model with a low normalized mean square error (NMSE) had a lower return than a model with a higher NMSE. In order to enhance the understanding of inexperienced traders and other people, Brownstone (1996) suggested using percentages to measure performance. Chen, Leung, and Daouk (2003) used a sliding window to forecast the next day's price of the index. The network was retrained with the most recent 68 days of input attempting to forecast the next day. In other wise to show the current tendency of the stock price fluctuations, technical indexes are calculated from the stock price's time series, trading volumes and time which are following a set of formula. We can apply these indexes for decision making to evaluate oversold or overbought in the stock market. Basically, the technical index can be classified as index for TSE movement or particular stock price variations, such as KD, RSI, MACD, MA, BIAS. As Chang et al. (2004) says, seven technical indexes are illustrated as shown in Table 3.1. Because ease of compute and further more for fair comparison with other model has used these factors as input variable for the model.