Applied Soft Computing 19 (2014) 86–92

Contents lists available at ScienceDirect

Applied Soft Computing

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

A hybrid ANFIS based on *n*-period moving average model to forecast TAIEX stock

Liang-Ying Wei[a], Ching-Hsue Cheng[b,*], Hsin-Hung Wu[c]

[a] Department of Information Management, Yuanpei University, 306 Yuanpei Street, Hsin Chu 30015, Taiwan, ROC
[b] Department of Information Management, National Yunlin University of Science and Technology, 123, Section 3, University Road, Touliu, Yunlin 640, Taiwan, ROC
[c] Department of Business Administration, National Changhua University of Education, No. 2 Shida Road, Changhua City, Changhua 500, Taiwan, ROC

ARTICLE INFO

Article history:
Received 3 June 2010
Received in revised form 4 April 2012
Accepted 18 January 2014
Available online 7 February 2014

Keywords:
Autocorrelation
Subtractive clustering
Adaptive-network-based fuzzy inference system
Moving average
Adaptive learning

ABSTRACT

Linear model is a general forecasting model and moving average technical index (MATI) is one of useful forecasting methods to predict the future stock prices in stock markets. Therefore, individual investors, stock fund managers, and financial analysts attempt to predict price fluctuation in stock markets by either linear model or MATI. From literatures, three major drawbacks are found in many existing forecasting models. First, forecasting rules mined from some AI algorithms, such as neural networks, could be very difficult to understand. Second, statistic assumptions about variables are required for time series to generate forecasting models, which are not easily understandable by stock investors. Third, stock market investors usually make short-term decisions based on recent price fluctuations, i.e., the last one or two periods, but most time series models use only the last period of stock price. In order to overcome these drawbacks, this study proposes a hybrid forecasting model using linear model and MATI to predict stock price trends with the following four steps: (1) test the lag period of Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX) and calculate the last *n*-period moving average; (2) use subtractive clustering to partition technical indicator values into linguistic values based on data discretization method objectively; (3) employ fuzzy inference system (FIS) to build linguistic rules from the linguistic technical indicator dataset, and optimize the FIS parameters by adaptive network; and (4) refine the proposed model by adaptive expectation models. The proposed model is then verified by root mean squared error (RMSE), and a ten-year period of TAIEX is selected as experiment datasets. The results show that the proposed model is superior to the other forecasting models, namely Chen's model and Yu's model in terms of RMSE.

1. Introduction

Stock markets are influenced by complex factors and nonlinear relationships among factors existed in different periods such that forecasting the future stock trends for investors are extremely difficult. Therefore, many forecasting methods have been employed in predicting stock prices since the stock market begins.

Many conventional numerical forecasting models by financial researchers have been proposed, such as Engle's [13] autoregressive conditional heteroscedasticity (ARCH) model, Bollerslev's [2] generalized ARCH (GARCH) model, Box and Jenkins' [3] autoregressive moving average (ARMA) model, and the autoregressive integrated moving average model (ARIMA). Autocorrelation (AR) is the correlation (relationship) between members of time series of observations, such as weekly share prices or interest rates and, then, the same values at a fixed time interval. More technically, autocorrelation occurs when residual error terms from observations of the same variable at different time periods are correlated (related). AR, a popular and important method in conventional time-series models, has been applied to time-series forecasting problems [6]. However, AR technique has limited capabilities for modeling time series data, and more advanced nonlinear methods including neural networks have been frequently applied with success [6]. Further, fuzzy logic-based modeling techniques are appealing because of their interpretability and potential to address a broad spectrum of problems.

Many researchers have been focusing on technical analysis to improve the investment return [1,11,26], because technical analysis methods are one of major analysis approaches for investors to make investment decisions with the ability to forecast the future price direction by studying past market data, primarily stock price, and volume. The technical analysis method has assumed that stock price and volume are the two most relevant factors in determining

* Corresponding author.
E-mail addresses: lywei@mail.ypu.edu.tw (L.-Y. Wei), chcheng@yuntech.edu.tw (C.-H. Cheng), hhwu@cc.ncue.edu.tw (H.-H. Wu).

1568-4946/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.asoc.2014.01.022

the future direction and behavior of a particular stock or market. The technical indicators, which come from the mathematical formula based on stock price and volume, can be applied to predict the future price fluctuation and also be provided for investors to determine the timing of buying or selling the stocks [11]. One of the most widely known trading rules used by financial traders is the moving average. Specifically, if the current stock price is higher or lower than the previous moving average of a specific period, then this gives a buy or sell signal, respectively. Moving average is one of technical analysis methods and has gained a significant increase of interest in the academia. Additionally, stock market investors usually make their short-term decisions based on recent stock information such as the latest price fluctuations. Therefore, in this paper we calculate the latest 3-period moving average values by our proposed model.

In the evolution of time series models, many studies have applied data mining techniques in financial analysis. In 1990, Kinoto et al. [18] have developed a prediction system for stock market by using neural network. Nikolopoulos and Fellrath [22] have combined genetic algorithms (GA) and neural network to develop a hybrid expert system for investment decisions. Kim and Han [17] have proposed genetic algorithms approach to feature discretization and the determination of connection weights for artificial neural networks (ANNs) to predict the stock price index. Huarng and Yu [15] have applied back-propagation neural network to establish fuzzy relationships in fuzzy time series for forecasting stock price. Moreover, Roh [23] has integrated neural network and time series model for forecasting the volatility of stock price index.

From the above literature reviews, there are three major drawbacks found in these forecasting methods/models. First, for most statistical methods, variables are assumed to follow a particular distribution. If the underlying assumption does not exist, these statistical methods cannot be applied to these datasets for analysis. Second, ANN is a black-box method, and the rules mined from the methods are not easily understandable. Third, stock market investors usually make short-term decisions based on recent price fluctuations (the last one or two periods), but most time series models use only the last period of stock prices in forecasting.

To improve the existing forecasting models, a thoughtful model should be able to overcome these drawbacks and be ease-of-use for investors. In addition, stock market investors usually make their short-term decisions based on recent stock information such as the latest market news or price fluctuations. Chen et al. [10] have pointed out that the price patterns in the Taiwan and Hong Kong stock markets are short-term. Besides, the linear relationships between recent periods of stock prices in forecasting models are typically assumed to forecast stock prices [9]. To consider the linear relationships between recent periods, we apply one-period and two-period adaptive expectation models to the proposed model for enhanced forecasting performance.

Under such circumstances, this paper proposes a hybrid forecasting model to refine the existing models in forecasting stock prices with the following processes: (1) test the lag period of Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX) by linear model and calculate the last n-period moving average; (2) use subtractive clustering to discretize condition features (technical indicators); (3) apply fuzzy inference system (FIS) to produce rules from the linguistic values of technical indicators and employ adaptive network to optimize FIS parameters to improve forecasting accuracy; and (4) refine forecasting accuracy of the proposed model by adaptive expectation models. Further, this paper employs stock index as experimental datasets for model verification by showing the effectiveness of the proposed processes in improving forecasting accuracy. More importantly, stock analysts or investors can employ the proposed refined processes to improve their forecasting tools or models.

The rest of this paper is organized as follows. Section 2 introduces the related works including moving average, subtractive clustering, adaptive-network-based fuzzy inference system (ANFIS), and multi-period adaptive model. The proposed model and algorithm are demonstrated in Section 3. Model verification is described in Section 4. Findings and conclusions are drawn in Section 5.

2. Related works

This section reviews related works of different forecasting models on stock market, moving average, subtractive clustering, ANFIS, and multi-period adaptive model.

2.1. Different forecasting models on stock market

Stock market is one of the most exciting and challenging monetary activities. The market climates are dramatically changed in a second and the gain-loss is decided in a twinkling decision. Fuzzy time-series models have been applied to handle economic forecasting, such as stock index forecasting, and various models have been proposed. The fuzzy time series model constructs the fuzzy relation R and uses a max–min composition operator to calculate forecasting values. Chen [7] proposed another model which uses equal interval lengths to partition the universe of discourse and generate forecasting rules with a simplified calculation process. In the process of establishing fuzzy relationships and forecasting, Yu [28] argued that recurrent fuzzy relationships should be considered in forecasting and recommended that different weights should be assigned to various fuzzy relationships. Therefore, Yu [28] proposed a weighted fuzzy time-series method to forecasting the TAIEX.

In summarizing, this paragraph provides some arguments: (1) in fuzzy time-series process, the concept of fuzzy logic was introduced to cope with the ambiguity and uncertainly of most of the real-world problems; (2) the fuzzy time-series are appropriately applied to linguistic values datasets to generate high accuracy forecasting rules.

2.2. Simple moving average trading rule

A common moving average trading rule is presented in Eq. (1)

$$ma_t = \frac{1}{L}\sum_{i=0}^{L-1} P_{t-i} \tag{1}$$

where ma_t represent t-period moving average value, L represent the period of moving average, and P_{t-i} represent stock price, by Eq. (1), a buy signal is generated when the current price is higher or above the moving average, while a sell signal is generated when it is below. There are variations in Eq. (1). When the buy–sell signals are incurred in future currencies and exchange rates with different current price above or below the moving average by a certain percentage or other modifications discussed in [24], the results of applying these procedures do not depart significantly. Moreover, the proposed model performs much better under the same scenario. The length of the moving average is selected by the technician, and the most popular length is 1–200, where the short period is 1 day and the long period is 200 days. There are other popular and much-used trading rules such as 1–50, 1–150, and others [4].

2.3. Autoregressive model

In time series forecast, predictions are practically obtained by forecasting a value at the next time period based on a specific prediction algorithm. In addition, forecasting non-periodic short-term time series is much more difficult than that for long-term time