Expert Systems with Applications 39 (2012) 4444–4452

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

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

Combining nonlinear independent component analysis and neural network for the prediction of Asian stock market indexes

Wensheng Dai [a,b], Jui-Yu Wu [c], Chi-Jie Lu [d,*]

[a] China Financial Policy Research Center, Renmin University of China, PR China
[b] Financial School, Renmin University of China, PR China
[c] Department of Business Administration, Lunghwa University of Science and Technology, Taiwan, ROC
[d] Department of Industrial Engineering and Management, Ching Yun University, Taiwan, ROC

A R T I C L E I N F O

Keywords:
Asian stock market
Stock index prediction
Feature extraction
Nonlinear independent component analysis
Neural network

A B S T R A C T

With the economic successes of several Asian economies and their increasingly important roles in the global financial market, the prediction of Asian stock markets has becoming a hot research area. As Asian stock markets are highly dynamic and exhibit wide variation, it may more realistic and practical that assumed the stock indexes of Asian stock markets are nonlinear mixture data. In this research, a time series prediction model by combining nonlinear independent component analysis (NLICA) and neural network is proposed to forecast Asian stock markets. NLICA is a novel feature extraction technique to find independent sources from observed nonlinear mixture data where no relevant data mixing mechanisms are available. In the proposed method, we first use NLICA to transform the input space composed of original time series data into the feature space consisting of independent components representing underlying information of the original data. Then, the ICs are served as the input variables of the neural network to build prediction model. Among the Asian stock markets, Japanese and China's stock markets are the biggest two in Asia and they respectively represent the two types of stock markets. Therefore, in order to evaluate the performance of the proposed approach, the Nikkei 225 closing index and Shanghai B-share closing index are used as illustrative examples. Experimental results show that the proposed forecasting model not only improves the prediction accuracy of the neural network approach but also outperforms the three comparison methods. The proposed stock index prediction model can be therefore a good alternative for Asian stock market indexes.

1. Introduction

With higher average growth rates and a significant consumer base that is fueling economic expansion, the economic successes of several Asian economies and their increasingly important roles in the global financial market have motivated several studies on Asian stock markets (Ané, Ureche-Rangau, Gambet, & Bouverot, 2008; Kim & Shamsuddin, 2008; Mazouz, Joseph, & Palliere, 2009). While Japanese and China's stock markets are the biggest two in Asia, they respectively represent the two types of stock markets. Japan's stock market represents the advanced financial market and has already attracted many researches (Atsalakis & Valavanis, 2009; Huang, Nakamori, & Wang, 2005; Lee & Chen, 2002; Lu, 2010; Lu, Lee, & Chiu, 2009). On the contrary, China's stock market is a typically emerging market. Few studies are reported on the prediction of China's stock market (Atsalakis & Valavanis, 2009; Cao, Leggio, & Schniederjans, 2005). But as one of the so-called BRICs, China's stock market is now the second largest in Asia following only Japan, and it is speculated with the potential to rank among the top four or five in the world within the coming decade. In 27th February 2007 which is called "Black Tuesday", with China's stock slumping near to 9%, Global stock marked suffered a sharp fall. It shows that China's stock market begin to play an important role for the world economy. So, China's stock market has recently attracted some interest in literatures. These includes the cross section of expected stock returns (Wang & Iorio, 2007), market efficiency (Fifield & Jetty, 2008), turning points of stock price (Yan, Powell, Shi, & Xu, 2007), and stock price forecasting (Cao et al., 2005).

A large number of successful applications have shown that neural networks can be a useful technique for stock price/index prediction since they do not require strong model assumptions and can map any nonlinear function without a priori assumption about the properties of the data (Atsalakis & Valavanis, 2009; Vellido, Lisboa, & Vaughan, 1999; Zhang, Patuwo, & Hu, 1998). In the most existing prediction methods using neural networks, the observed original values of prediction variables are usually directly used

* Corresponding author. Tel.: +886 3 4581196x6712; fax: +886 3 4683298.
E-mail address: jerrylu@cyu.edu.tw (C.-J. Lu).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.09.145

for building prediction models (Vellido et al., 1999; Zhang et al., 1998). However, a stock price may be affected by some underlying factors like seasonal variations or economic events (Back & Weigend, 1997; Górriz, Puntonet, Salmerón, & Lang, 2003; Kiviluoto & Oja, 1998; Malaroiu, Kiviluoto, & Oja, 2000; Mok, Lam, & Ng, 2004). For example, in Mok et al. (2004), the observed daily stock returns reflected the stock market reaction to a few factors such as release of various economic indicators, company news, psychological factors, government intervention, or political issues. Kiviluoto and Oja (1998) showed that the underlying factors affected the cash flow time series of several stores belonging to the same retail chain were trends, holiday effects, seasonal variation and competition. Back and Weigend (1997) used ICA to exact the features of the daily returns of the 28 largest Japanese stocks. The results showed that the dominant ICs can reveal more underlying structure and information of the stock prices than principal component analysis. Therefore, if the underlying/interesting information that can't be observed directly from the observed original data can be revealed through transforming the input space into the feature space with suitable feature extraction methods, the prediction performance will be improved by using the features to build prediction model.

Independent component analysis (ICA) is a novel feature extraction technique. It aims at recovering independent sources from their mixtures, without knowing the mixing procedure or any specific knowledge of the sources (Hyvärinen, Karhunen, & Oja, 2001). The independent sources, called independent components (ICs), are hidden information of the observable data. ICA model can be classified as linear ICA (LICA) and nonlinear ICA (NLICA) models. Linear ICA model is mainly based on the assumption that the observed mixture data are generated by a linear mixing from ICs (Hyvärinen et al., 2001). Conversely, the nonlinear ICA model assumes that the observations are a nonlinear combination of ICs (Almeida, 2003; Hyvärinen & Pajunen, 1999; Valpola, 2000).

Linear ICA has been employed successfully in various fields, such as signal processing, face recognition and time series forecasting (Górriz et al., 2003; Hyvärinen & Oja, 2000; Lu, 2010; Lu et al., 2009; Malaroiu et al., 2000; Popescu, 2003). Oja, Kiviluoto, and Malaroiu (2000) applied LICA in foreign exchange rate time series prediction. They firstly used LICA to estimate independent components and mixing matrix from the observed time series data and filtered the independent components (ICs) to reduce the effects of noise through the linear and nonlinear smoothing techniques. Then, the autoregression (AR) modeling was employed to predict the smoothed independent components. Finally, they combined the predictions of each smoothed IC by using mixing matrix and thus obtained the predications for the original observed time series. Cao and Chong (2002) employed LICA as a feature extraction tool in developing a support vector machine (SVM) forecaster. The ICs were considered as features of the forecasting data and used to build the SVM forecasting model. Lu et al. (2009) proposed a two-stage forecasting model by integrating LICA and support vector regression (SVR) for financial time series. They first used LICA to the forecasting variables for generating the ICs. After identifying and removing the ICs containing the noise, the rest of the ICs were then used to reconstruct the forecasting variables which contain less noise and served as the input variables of the SVR forecasting model. Lu (2010) utilized LICA to generate a denoising scheme for stock price data. The scheme was then used as preprocessing tool before building a stock price prediction model.

Linear ICA is suitable when it is reasonable to assume that the mixing mechanism is linear or the observations are linear mixture data. However, in many realistic cases/problems, the observed data (for example, stock index data) might be a nonlinear mixture of latent source signals. Asian stock markets are highly dynamic and exhibit wide variation since a lot of global investors are attracted to Asian stock markets. This makes Asian stock market indexes can be realistically and practically assumed to be the nonlinear mixture data. Therefore, NLICA has the potential to become rather powerful tools in the prediction of Asian stock market indexes. Moreover, unlike LICA model, NLICA remains an active field of research. Few applications of NLICA are proposed in literatures, such as image denoising (Haritopoulos, Yin, & Allinson, 2002) and feature extraction (Almeida, 2003; Zhang & Chan, 2007).

In this study, a stock index prediction model by integrating NLICA and neural network is proposed for two major Asian stock markets-China and Japan. The proposed approach first uses NLICA on the input space composed of original forecasting variables into the feature space consisting of independent components representing underlying/hidden information of the original data. The hidden information of the original data could be discovered in these ICs. The ICs are then used as the input variables of the neural network for building the prediction model. In order to evaluate the performance of the proposed approach, the Shanghai B-Share stock index closing price and Nikkei 225 closing cash index are used as the illustrative examples.

The rest of this paper is organized as follows. Section 2 gives brief overviews of nonlinear ICA and neural network. The proposed model is described in Section 3. Section 4 presents the experimental results and this paper is concluded in Section 5.

2. Methodology

2.1. Nonlinear independent component analysis

Linear ICA is becoming a well research area. In the linear ICA model (Hyvärinen et al., 2001), it is assumed that m measured variables, $\tilde{x} = [x_1, x_2, \ldots, x_m]^T$ can be expressed as linear combinations of n unknown latent source components $\tilde{s} = [s_1, s_2, \ldots, s_n]^T$:

$$\tilde{x} = \sum_{j=1}^{n} \tilde{a}_j s_j = \mathbf{A}\tilde{s} \tag{1}$$

where $\tilde{a}_j$ is the jth row of unknown mixing matrix $\mathbf{A}$. Here, we assume $m \geqslant n$ for $\mathbf{A}$ to be full rank matrix. The vector $\tilde{s}$ is the latent source data that cannot be directly observed from the observed mixture data $\tilde{x}$. The linear ICA aims to estimate the latent source components $\tilde{s}$ and unknown mixing matrix $\mathbf{A}$ from $\tilde{x}$ with appropriate assumptions on the statistical properties of the source distribution. Thus, linear ICA model intents to find a de-mixing matrix $\mathbf{W}$ such that

$$\tilde{y} = \sum_{j=1}^{n} \tilde{w}_j x_j = \mathbf{W}\tilde{x}, \tag{2}$$

where $\tilde{y} = [y_1, y_2, \cdots, y_n]^T$ is the independent component vector. The elements of $\tilde{y}$ must be statistically independent, and are called independent components (ICs). The ICs are used to estimate the source components s_j. The vector $\tilde{w}_j$ in Eq. (2) is the jth row of the de-mixing matrix $\mathbf{W}$.

Since the nonlinear ICA model assumes that the observations are a nonlinear combination of latent sources, it can be formulated as the estimation of the following model (Almeida, 2003)

$$\tilde{x} = F(\tilde{s}) \tag{3}$$

where $\tilde{x}$ and $\tilde{s}$ denote the data and source vector as before, and F is an unknown nonlinear transformation function. Assume now for simplicity that the number of independent components equals the number of mixtures. The nonlinear ICA problem then consists of finding a mapping $G : \Re^n \rightarrow \Re^n$ that yields components

$$\tilde{y} = G(\tilde{x}) \tag{4}$$

which are statistically independent.