Expert Systems With Applications 117 (2019) 228–242

Contents lists available at ScienceDirect

Expert Systems With Applications

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

Global stock market investment strategies based on financial network indicators using machine learning techniques

Tae Kyun Lee[a,b], Joon Hyung Cho[b], Deuk Sin Kwon[b], So Young Sohn[b,*]

[a] Samsung Asset Management Co., Ltd., Seoul, Republic of Korea
[b] Department of Information and Industrial Engineering, Yonsei University, Seoul, Republic of Korea

ARTICLE INFO

Article history:
Received 29 May 2018
Revised 2 September 2018
Accepted 2 September 2018
Available online 19 September 2018

Keywords:
Financial network
Global stock market
Machine learning
Investment strategy

ABSTRACT

This study presents financial network indicators that can be applied to global stock market investment strategies. We propose to design both undirected and directed volatility networks of global stock market based on simple pair-wise correlation and system-wide connectedness of national stock indices using a vector auto-regressive model. We examine the effect and usefulness of network indicators by applying them as inputs for determining strategies via several machine learning approaches (logistic regression, support vector machine, and random forest). Two strategies are constructed considering stock price indices: (1) global stock market prediction strategy and (2) regional allocation strategy for developed market/emerging market. According to the results of the performance analysis, network indicators were proven to be important supplementary indicators in predicting global stock market and regional relative directions (up/down). In particular, these indicators were more effective during market crisis periods. This study is the first attempt to construct strategies for global portfolio management using financial network indicators and to suggest how network indicators can be used in practical fields.

1. Introduction

In the management of global investment portfolios, four levels of portfolio strategies are typically considered. From macro to micro perspectives, they are asset allocation, regional allocation, country selection, and security selection. Among these strategies, the first two strategies employed for global stock investment portfolios are mainly affected by macro financial issues such as market regime cycles and global money flows, which have significant influences on profits. Therefore, almost all investment banks study these strategies. Allocating the amount of risk assets is the key factor for tactical asset allocation strategy. As the stock market accounts for most of the risk assets, prediction of stock markets is crucial. However, betting on the direction of stock markets is regarded as a high-risk strategy because there are too many external factors affecting it. In addition, this strategy could significantly influence portfolio returns with only a slight change in proportion of allocation of assets owing to high volatility. Therefore, establishing a strategy for assigning weights to stocks is crucial for portfolio returns in asset management fields. As for the regional allocation strategy, it is necessary to consider the relative attractiveness of specific regions in terms of investment returns. Portfolio managers adjust the proportions over regions based on a pre-assigned benchmark regional weight. As it is less risky than stock market directional strategy, it is used more frequently. Thus, regional allocation strategy contributes as much as stock market directional strategy to global portfolio returns (Morgan, 2012,2014).

In order to design an optimal investment strategy as described above, it is important to accurately predict the market by understanding its characteristics. In academic fields, many studies on market prediction and forecasting methods have been carried out. There are studies on market valuation based on price or financial statement information (Piotroski, 2000; Kang et al., 2002; Brown, 2005; Kim & Sohn, 2012; Na & Sohn, 2011; Sohn & Lim, 2007). Many studies have focused on bubble signals to capture market collapses (Werner, 1997; Abreu & Brunnermeier, 2003; Bakaert et al., 2014; Jarrow et al., 2011). Recently, there have been efforts to apply machine learning techniques as they have shown relative success in predicting financial time series (Lee, 2009; Dai et al., 2012; Barak & Modarres, 2015). Patel, Shah, Thakkar, and Kotecha (2015a) suggested the use of price technical parameters and compared the performances of several machine le models for stock prediction. Inthachot, Boonjing, and Intakosum (2016) applied a genetic algorithm for selecting machine learning model input variables. In addition, multiple machine learning models have been combined for constructing forex portfolio trading strate-

* Corresponding author.
E-mail addresses: tkaaron.lee@samsung.com (T.K. Lee), paul619@naver.com (J.H. Cho), kds0281@gmail.com (D.S. Kwon), sohns@yonsei.ac.kr (S.Y. Sohn).

https://doi.org/10.1016/j.eswa.2018.09.005
0957-4174/

gies (Petropoulos, Chatzis, & Vlachogiannakis, 2017). A study has been published which uses online data sources including google trends, Wiki and financial news with an ensemble learning model (Weng, Lu, Wang, Megahed, & Martinez, 2018). Volatility of markets is also considered an important factor to describe market environments; there has been research on forecasting market volatility with machine learning (Hajizadeh et al., 2012; Kristjanpoller & Minutolo, 2016; Kristjanpoller, Fadic, & Minutolo, 2014), and asset allocation strategies using volatilities (Kim & Enke, 2018).

One of the most important characteristics for global investment is that the stock markets of countries across the world are closely linked (Allen & Gale, 2000; Castiglionesi, 2009). The importance of this connectivity has become more prominent since the 2008 financial crisis (Haldane, 2009). The financial crisis has changed the general concept of risk diversification and financial connections have been recognized important globally even in separated markets. Therefore, it is necessary to examine the global linkage effect, which can be obtained from the complex connection among various markets (Hong & Sohn, 2013). Considering the connected financial structure, Nier, Yang, Yorulmazer, and Alentorn (2007) described the vulnerability of the financial system in terms of potential risk contagion owing to inter-bank connectedness. By constructing several banking systems with various key parameters of the structure of the financial system, the authors showed that concentrated banking systems tend to be prone to systemic risk. Aloui, Aïssa, and Nguyen (2011) showed the effect of global linkage between emerging and US markets on the financial crisis. Glasserman and Young (2015) suggested a model to measure systemic risk and estimated the risk contagion with financial connectivity.

The studies have been expanded to financial networks to describe complex linkage structures over global regions or markets more effectively. The phenomenon of financial connectivity including risk contagion and market collapses has been explained. Billio, Getmansky, Lo, and Pelizzon (2012) measured the ex-ante systemic risk possibility based on integrated financial network analysis of divided financial markets such as banks, insurances, hedge funds, and other financial markets. Networks were constructed using the Granger causality method with rolling period from 1994 to 2008. This method proposes network measures using principal-component analysis, which identifies and quantifies financial crisis periods. In addition, Diebold and Yilmaz (2016) constructed the equity return volatility networks of the financial institutions of United States (US) and Europe based on a vector auto-regressive (VAR) model over 2004–2014. The directed connections of each company with other companies were extracted from the error variance decomposition weight in the VAR model. This model analyzed the changes in the financial network during the 2008 global financial crisis and the 2012 European fiscal crisis. It also showed the risk transfer between the two systemic risks that occurred on different continents. Peralta and Zareei (2016) presented a portfolio strategy that improves the risk-adjusted performance through network analysis. They constructed a network based on the correlation between stock returns and presented a portfolio strategy by reflecting the centrality of each individual stock. Despite the several network indicators used for financial market analysis, their applications have limitations for developing practical portfolio strategies that can provide enhanced performances. Chuluun (2017) examines the impact of the financial network by country's centralities with international stock market co-movement. This shows that linkages derived from total portfolio holdings are important for portfolio diversification. Barunik, Kocenda, and Vacha (2017) show that volatility propagates though the forex market, providing evidence of asymmetric connectedness with traded currency datasets. This shows that currencies do not display similar patterns in their asymmetric directional spillovers. In any event such as the GFC period, negative directional spillovers transmitted from one currency impact the volatility of other currencies more than positive spillovers.

This study aims to present a practical methodology to design an optimal portfolio strategy of the global stock market with financial network indicators. Although there have been many studies on the importance of network analysis (Chi et al., 2010; Barigozzi & Hallin, 2016), especially in high volatility states (Billio et al., 2012; Diebold & Yilmaz, 2014 & 2016), there have been only a few attempts that construct investment strategies through these network indicators (Peralta & Zareei, 2016). Recent studies of network indicators for investment (Barunik et al., 2017; Chuluun, 2017) provide the evidence of the network effects of investment strategies. However, these papers are also not expanded to practical investment strategy and effects in portfolios. Moreover, there have been no attempts to utilize them in the most important and difficult global financial market outlook and portfolio strategy.

Especially, this paper clearly shows the impacts of the financial network on financial market among market turn-moil periods. Even though it is a common fact in portfolio managements, we proved it with more objective and quantitative method by data simulation. This study is the first to demonstrate the applicability of network indicators to global investment. It suggests steps for the implementation of the strategy and shows the effect of network indicators. In order to establish the investment strategy, we propose to construct volatility networks of the global stock market based on a simple pair-wise correlation and system-wide connectedness of national representative financial indices. By extending the study by Diebold and Yilmaz (2016), which only analyzed the US & Europe banking networks, we broadly construct a system-wide volatility network based on global stock indices. In addition, we examine the effect and usefulness of network indicators in terms of the global investment strategy. Two global investment strategies are designed considering stock price indices: 1) stock market prediction strategy (forecasting total stock market direction) and 2) regional allocation strategy (forecasting the relative direction of developed market (DM)/emerging market (EM)). These two strategies are the most dominant strategies employed for global stock market portfolios in practice. In particular, we first built an investment strategy for regional allocation utilization using financial network indicators and validated the usefulness of the global stock market strategies.

As previously stated, machine learning applications have recently been advanced in the forms of deep running and graphical models. Since deep learning can take so many forms, it has been proposed in various ways depending on the financial market or viewpoint. However, it is difficult to apply a deep learning model in this paper owing to the lack of a commonly used deep learning structure in financial fields.

Accordingly, we apply three machine learning approaches that have been applied in numerous studies (Huang et al., 2005; Dutta et al., 2012; Cavalcante et al., 2016) and used for prediction in the financial industry: logistic regression (LR), support vector machine (SVM), random forest (RF) and Deep Neural Network (DNN) (Kim & Sohn, 2010; Sohn, Moon, & Kim, 2005). We predict the directions of the total global stock market indices and regional relative attractiveness using simple index prices and network indicators. We also conduct sensitivity analyses by changing the forecasting and volatility periods to determine which period is the most appropriate for volatility construction to predict market directions. Finally, we recommend a suitable investment strategy and provide guidance regarding global portfolio management strategies.

This paper is organized as follows. The data and methodology are explained in Section 2. Section 3 presents the results. Section 4 presents the conclusions and contributions.

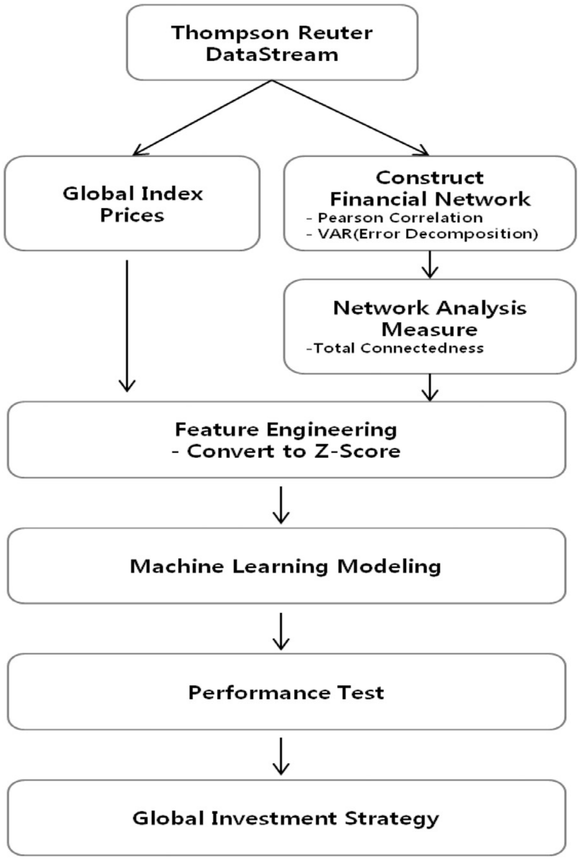

Fig. 1. Steps for global stock market investment strategy.

2. Methodology

2.1. Framework

The framework proposed for a global stock market investment strategy is described in Fig. 1. We first collect the stock indices data of 10 individual countries and construct two global volatility networks by using their Pearson correlation and VAR model. The financial network consists of a network based on an undirected pair-wise Pearson correlation and system-wide VAR model presented by Diebold and Yilmaz (2016). The Pearson correlation network can identify undirected relationships between countries. It can describe network linkage density, whereas the system-wise VAR model network can capture directional impacts across countries. Subsequently, network indicators representing connectedness are extracted from these two kinds of networks. Connectedness measures are the edge densities among global stock indices. The edge densities based on the average of edge strength are calculated from dynamic financial networks over rolling time windows, which construct time-series data. Subsequently, a time-series dataset including stock indices and connectedness indicators is converted to z-score based on historical time-series data of 52 weeks as they are observed in different scales over countries and observation periods. Subsequently, we predict stock market directions with z-score input data using classification models such as LR, SVM, and RF techniques. Finally, we evaluate the performance of our proposed approach and propose global investment portfolios based on the global investment strategy, which constitutes different stock volatility and prediction periods. Table 1 shows several time periods used in this study for stock volatility, prediction, and z-score conversion.

2.2. Research data

In order to establish an investment strategy for the global indices, 10 countries were selected as displayed in Table 2. The countries were selected by considering the Morgan Stanley Capital International (MSCI) Index countries and the regional representation. Five of the 10 countries were classified as developed market (DM) and the remaining countries were classified as emerging market (EM) based on the MSCI world index market classification. The data were collected over a period of 22 years from January 1995 to December 2016, and all data were sourced from Thompson Reuters DataStream Database. The summary statistics of the period, stock index, annualized return, and volatility of each selected country are described in Table 2.

2.3. Financial volatility network

The volatility of a stock index is defined as the standard deviation of the daily log returns. The volatility was calculated for each period (1 week, 4 weeks, 8 weeks, and 12 weeks). Based on this value, the annualized volatility of the stock index can be obtained as follows:

$$v_t = \left[252 \sum_{k=1}^{L} (R_{t-k} - \mu)^2 / (L-1) \right]^{1/2},$$

where R_t is the daily log return of stock index, μ is the sample mean over the volatility calculation period, and L is the number of days (Figlewski, 1997).

First, we construct a correlation-based network by using an edge with the volatility correlation between the stock indices of 10 countries. Second, we construct a financial volatility network from the variance decomposition matrix of the VAR model. The p-lag vector autoregressive (VAR(p)) model is defined as follows:

$$V_t = \sum_{i=1}^{p} \phi_i V_{t-i} + \varepsilon_t, \text{ where } \varepsilon_t \sim (0, \Sigma)$$

where $V_t = (v_{1t}, v_{2t}, \cdots, v_{Nt})'$, volatility of a stock index is an $N \times 1$ vector of jointly determined dependent variables, φ_i represents $N \times N$ coefficient matrices, and ε_t is an $N \times 1$ unobservable zero mean white noise vector. The moving average representation is $V_t = \sum_{i=0}^{\infty} A_i \varepsilon_{t-i}$, where the $N \times N$ coefficient matrices, A_i, obey the recursion $A_i = \Phi_1 A_{i-1} + \Phi_2 A_{i-2} + \cdots + \Phi_p A_{i-p}$, with A_0 an $N \times N$ identity matrix and $A_i = 0$ for $i < 0$.

The variance decompositions are the transformations of the moving-average coefficients. They allow splitting of the H-step-ahead forecast error variances of each variable (Seymen, 2008). This provides information on the extent to which each variable contributed to changes in the post-H periods of the linked international financial markets. In this study, we construct a 3-lag VAR model and H-step was used for 10 days (2 weeks).

We use the generalized variance decomposition framework of Koop, Pesaran, and Potter (1996) and Pesaran and Shin (1998). The contribution of variable j to the H-step-ahead generalized forecast error variance of variable i is

$$\theta_{ij}^{\delta}(H) = \frac{\sigma_{jj}^{-1} \sum_{h=0}^{H-1} (e'_i A_h \sum e_j)^2}{\sum_{h=0}^{H-1} (e'_i A_h \sum A'_h e_i)},$$

where Σ is the variance matrix for the error vector ε, σ_{jj} is the standard deviation of the error term for the j-th equation, and e_i is the selection vector with unity as the i-th element and zeros elsewhere.

The explanatory power of future forecast error variance of individual variables derived using variance decomposition indicates

Table 1.
Time periods used in global stock market investment process.

Period 1 (Stock volatility period)	Period 2 (Prediction period)	Period 3 (Z-score conversion)
1, 4, 8, 12 weeks	1, 4, 8, 12 weeks	Past 52 weeks

Table 2.
Summary statistics of global stock indices.

Country	Stock index	Region	# of Day	Annualized average return	Volatility
Germany	DAX 30 Performance	DM	5480	14.9%	23.7%
United Kingdom	FTSE 100	DM	5480	6.9%	18.5%
Hong Kong	HANG SENG	DM	5480	9.9%	25.4%
Japan	NIKKEI 225	DM	5480	3.8%	23.6%
United States	S&P 500 Composite	DM	5480	11.2%	18.9%
Brazil	BOVESPA	EM	5480	25.0%	31.7%
China	SHANGHAI SE Composite	EM	5480	16.5%	26.3%
India	S&P BSE 30 SENSITIVE	EM	5480	18.5%	24.8%
Korea	KOREA SE Composite	EM	5480	11.0%	27.4%
Taiwan	TAIWAN SE Weighted	EM	5480	7.5%	22.1%

the degree of effectiveness of stock index of each country in the financial network. The countries have in and out constituents for effectiveness. The net values of effectiveness are the edges between countries. Based on this, we constituted a VAR network.

In both models (Pearson correlation and VAR models), the connectedness measure was calculated as follows, which is the average of all connections of the network:

$$\text{Connectedness measure} = \frac{\sum_{i,j=1,\ i\neq j}^{N} E_{ij}}{M},$$

where E_{ij} = value of edge between countries i and j, M = number of connections (edges; undirected: correlation, directed: Θ of VAR model).

In order to determine the most effective period of volatility for market prediction, financial volatility networks of each volatility period (1 week, 4 weeks, 8 weeks, and 12 weeks) are constructed. Fig. 2 shows the illustration of the volatility financial network derived from the above methods. To identity the structural changes in financial networks, we compare three periods (1996: stable period, 1997/2008: turmoil period). The Pearson correlation network shows that the inter-state volatility network was strongly built during the emerging market crisis and the 2008 global financial crisis. In the VAR network model, the soaring influence of Asian countries during the emerging market crisis was depicted and the US influence surged during the 2008 financial crisis. The VAR network provided further detailed information as it is a directed method (Fig. 3).

2.4. Machine learning techniques

In this study, we used LR, RF, and SVM in a financial market forecasting model to predict the direction of movements of individual stocks and stock price indices (Patel, Shah, Thakkar, & Kotecha, 2015b; Zhang et al., 2014). The LR, which is a multivariate analysis model, is an appropriate model for predicting the presence of a characteristic with a set of predictor variables (Lee, 2004). As LR is useful when the dependent variable is binary or a multinomial categorical variable, it has commonly been used in the field of investments for the prediction of directions or classification of companies.

RF belongs to the category of ensemble algorithms. It is a classification learning technique developed by Breiman (2001). This learning method is the extension of traditional decision tree techniques, and has a meta-learning form of decision trees instead of one. The RF is composed of multiple decision trees, which are randomly selected variables. The combined prediction trees are expected to increase the accuracy and stability of the model performances as compared with those of a single classification model. In the case of RF, the larger the size of the forest (the number of trees), the more the convergence of the generalization error to a specific value, and thus, the over-fitting can be avoided. RF uses randomly extracted data from the total training dataset and is not significantly affected by noise or outliers. In addition, RF has been presented as the best predictive classification algorithm for imbalanced data sets according to Brown and Mues (2012). In this study, we use the Gini index to calculate the criterion of decision trees.

SVM is also a widely used machine learning technique for classification in the industry. The learning processes for determining decision boundaries should maximize the shortest distance (margin) to the boundary in order to determine the best classifier in the given data. SVM is well suited to many classification problems owing to the minimization of over-fitting, high accuracy, easy modification, and ability to handle high-dimensional data (Cortes & Vapnik, 1995). However, it requires a long time because it has a large computational complexity depending on the kernel function. In this study, we selected a radial basis kernel function.

$$K(x_i, x_j) = \exp\left(-\gamma \left\| x_i - x_j \right\|^2\right),$$

where γ is the constant of the radial basis function.

As for the LR model, the cut-off probability is used as a parameter. The number of trees and random features are the parameters of RF. The regularization parameter and the gamma in the kernel function are the parameters of the SVM model. Table 3 lists the parameters used in each machine learning technique.

In this study, the performance of the reference model, which only uses the stock index Z-score of each country as inputs without connectedness measures, is compared with that of the suggested model, which uses connectedness measures and stock index Z-score. In addition, we attempted to recommend the most suitable model for portfolio strategy by analyzing the change in forecasting period and volatility depending on the observation period and the machine learning model.

2.5. Investment framework

In this study, we suggested two approaches: (1) stock market prediction strategy (investing 1/10 each to 10 countries based on global stock market directional forecasting), (2) regional allocation strategy (investing 50% each to DM/EM regions as long/short position by comparing the attractiveness between DM/EM country stock indices). Investment simulations were conducted under the

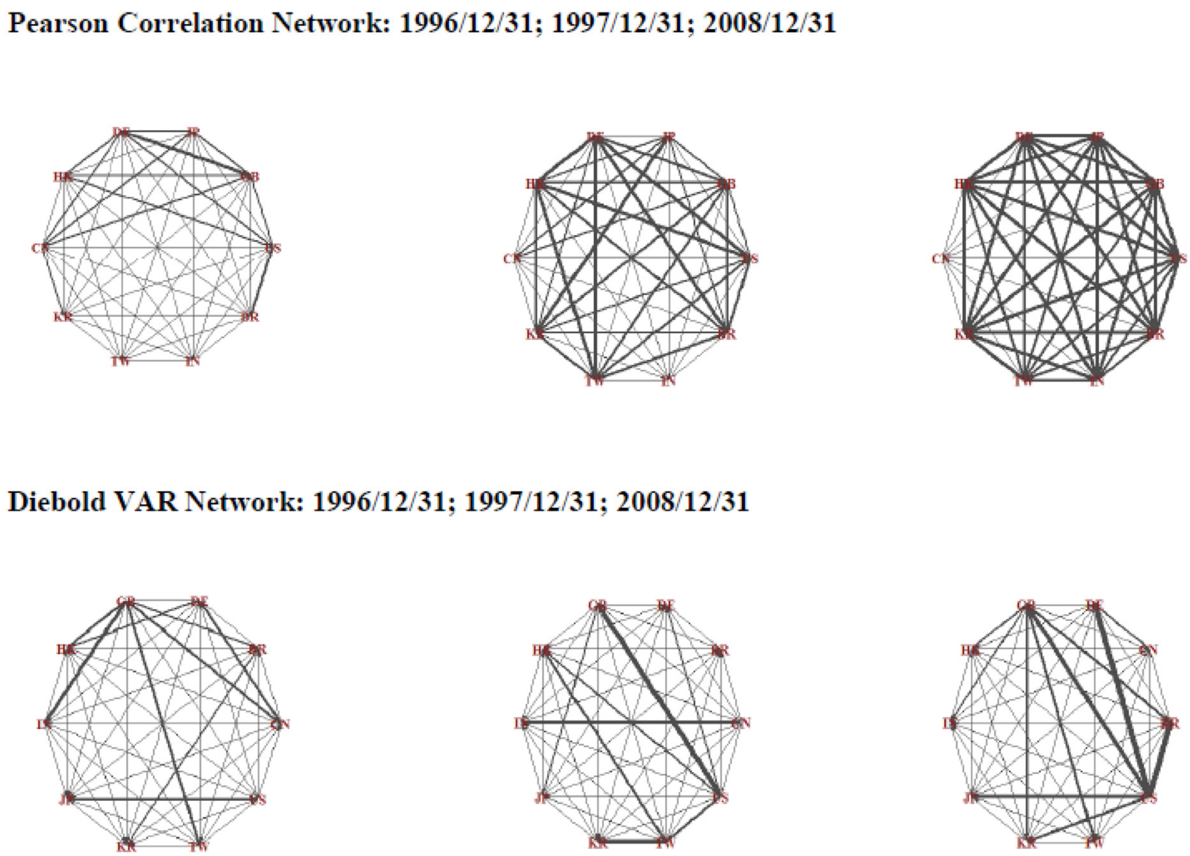

Fig. 2. Global stock market volatility network (week_vol = 4weeks).

Table 3.
Machine learning parameters and their levels tested in parameter setting.

Machine Models	Parameters	Levels
Logistic regression (LR)	Cut-off probability (p)	0.38–0.62 changed by 0.01
Random forest (RF)	Number of features (k)	2, 4, 6, 8, 10
	Number of trees (p)	100, 150, 200, 250, 300
Support vector machine (SVM)	Kernel function	Radial basis function
	Gamma in kernel function (γ)	0.25, 0.5, 1, 2, 4
	Regularization parameter (c)	2, 4, 6, 8, 10

assumption of daily investing with the holding periods same as the forecasting periods.

The investment period (forecasting period) and volatility period were divided into four periods (1 week, 4 weeks, 8 weeks, and 12 weeks). Fig. 4 shows the entire structure of strategies divided into two parts: total stock market prediction and DM/EM regional allocation strategies.

Fig. 5 shows the testing methodology for the strategies. We set up a training set that adds one year from the initial training set (1995–2004) by rolling the time series data, and in the testing set, we measured the accuracy of the suggested strategies based on the testing data for the subsequent year after the training period. The performances were measured as the average of the top five parameter results. The top 5 mixed parameter selection is chosen since it is appropriate for measuring the average performance of machine learning models. It could be applied to various situations and there are some papers in which the same method was used for selecting parameters (Patel et al., 2015b). In addition, in order to determine the actual profit according to the model used in the study, we constructed the model portfolio within the testing period and examined the profit rate realized via the investment simulation.

3. Results

3.1. Accuracy performance

In order to measure the performances of the strategies, directional accuracies of 16 portfolios consisting of different periods of forecasting and volatility for network construction, were tested by using the LR, RF and SVM model. The suggested model is learned with the training data by rolling dates and tests the accuracy during the testing period 2005 to 2016. In addition, to identify the effect of network indicators, we compared the performances obtained with and without using network indicators (using only price indicators) in the stock market prediction strategy and regional allocation strategy. The usefulness of network indicators was confirmed in terms of effective support ability for directional forecasting of the whole stock market and regional allocation strategy of DM and EM.

Table 4 shows the results of forecasting models for the overall stock market and the direction of the global stock indices. The directional forecasting of the global stock market using the network measures shows better accuracy than the strategy that does

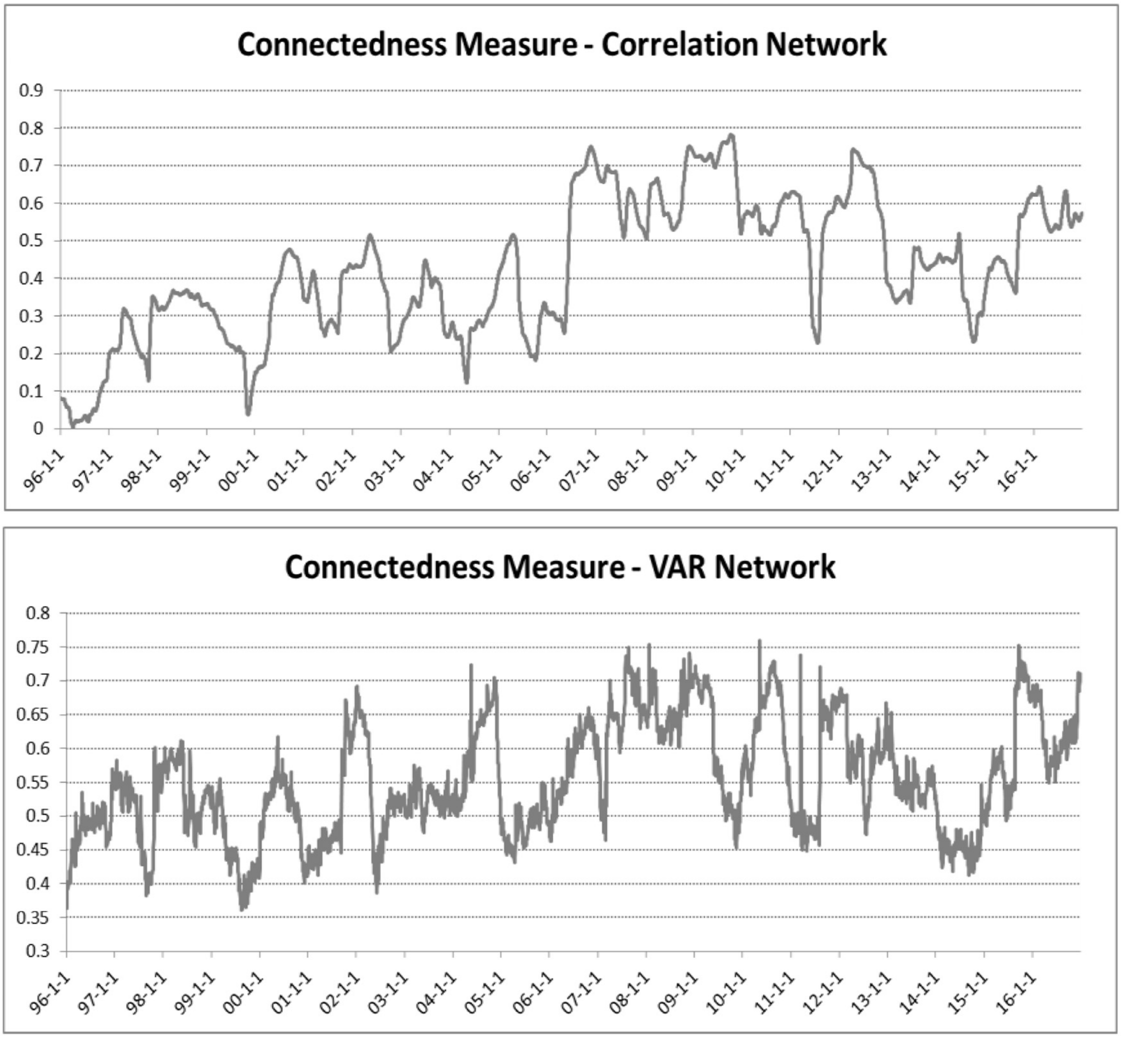

Fig. 3. Global stock volatility connectedness measure time series.

Table 4.
Result for each volatility and forecasting period in global stock market: Up/Down accuracy of average global stock index direction.

Forecasting period		With network				Without network			
Machine learning Model / volatility period		1 week	4 weeks	8 weeks	12 weeks	1 week	4 weeks	8 weeks	12 weeks
LR	1w	61.1%	65.2%	64.7%	68.6%	60.3%	65.8%	62.7%	64.1%
	4w	60.5%	63.6%	63.6%	65.2%	60.3%	65.9%	63.5%	64.1%
	8w	60.9%	63.9%	62.8%	64.0%	60.1%	65.6%	62.8%	64.2%
	12w	61.2%	65.3%	63.8%	65.3%	60.4%	65.9%	62.9%	64.4%
	Avg	60.9%	64.5%	63.7%	65.8%	60.3%	65.8%	63.0%	64.2%
RF	1w	56.9%	60.4%	63.3%	68.9%	56.3%	56.8%	62.6%	63.2%
	4w	55.5%	57.0%	62.9%	65.6%	55.8%	56.4%	62.3%	62.8%
	8w	56.0%	58.5%	64.4%	64.8%	55.9%	56.6%	62.0%	63.2%
	12w	56.8%	56.8%	61.9%	65.7%	56.2%	56.8%	62.2%	62.8%
	Avg	56.3%	58.2%	63.1%	66.2%	56.0%	56.6%	62.3%	63.0%
SVM	1w	59.8%	68.4%	70.4%	72.3%	59.6%	65.4%	67.1%	70.5%
	4w	58.8%	66.8%	66.6%	71.5%	59.5%	65.2%	67.2%	70.4%
	8w	59.2%	65.8%	67.6%	71.6%	59.1%	65.6%	67.5%	70.6%
	12w	59.3%	65.4%	67.3%	70.5%	59.3%	65.2%	66.6%	70.4%
	Avg	59.3%	66.6%	68.0%	71.5%	59.4%	65.4%	67.1%	70.5%

not use them, especially in 8- and 12-week forecasting periods. This suggests that the use of network indicators is more appropriate for medium-term rather than short-term investment. In addition, the accuracy results obtained for periods of volatility show that short-term volatility is more effective than mid-term volatility. The reason is that short-term volatility could further explain the market movement. In practical asset management, short-term volatility is used as an important risk indicator as well. Overall, the performance accuracy was satisfactory as it was over 60% or 70%, but it could be viewed as insufficient because of enhancement of only approximately 2–5% in accuracy after adding the network indicators. However, considering that the prediction enhancement of the upward downward of the overall stock market is very difficult owing to the influences of external factors such as monetary pol-

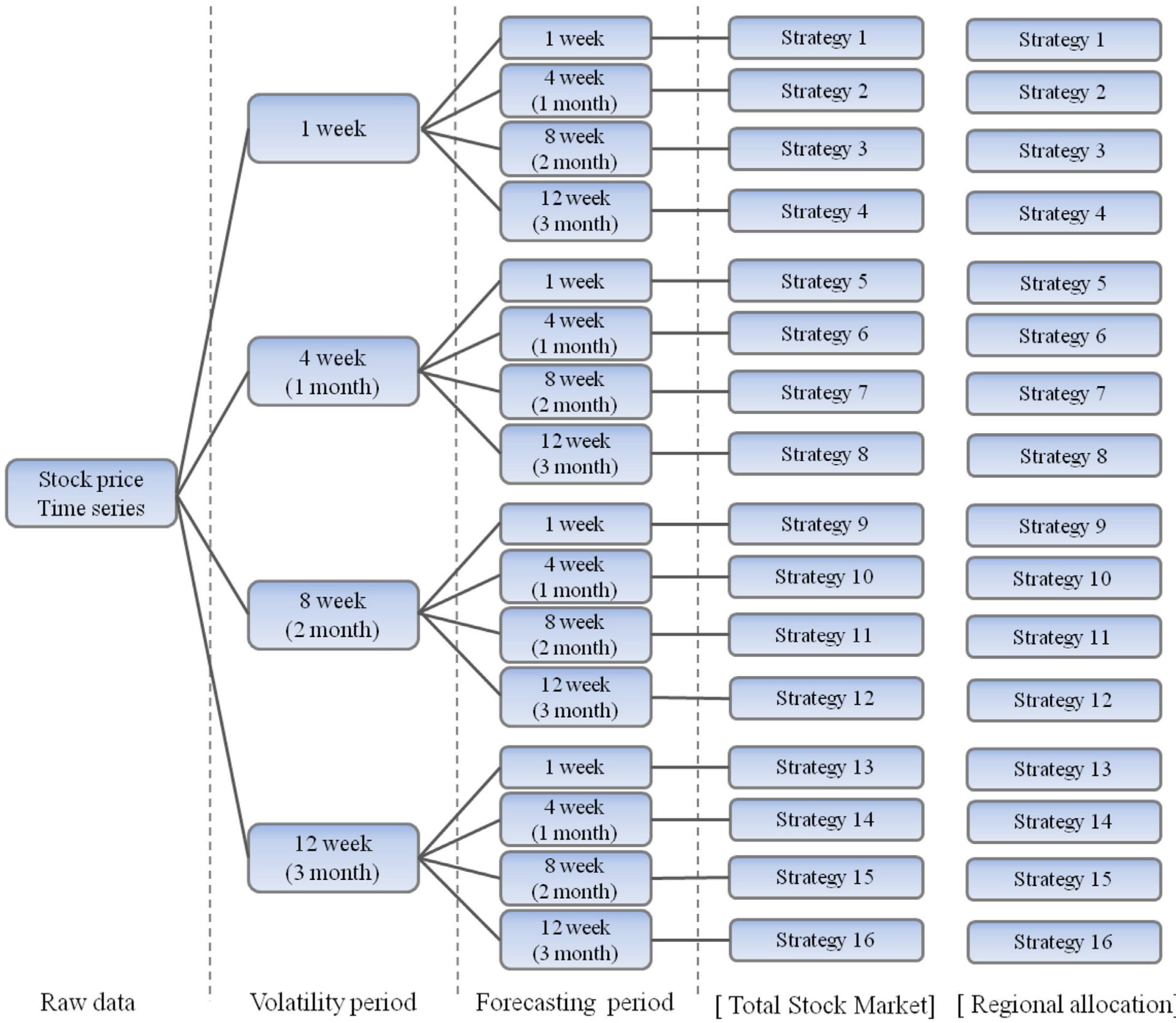

Fig. 4. Global stock market investment strategy structure.

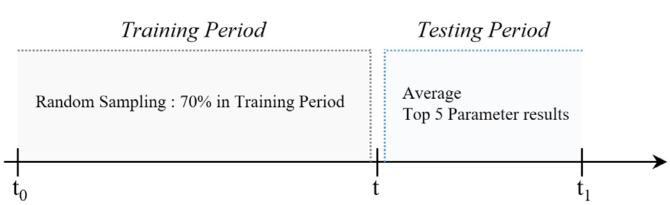

Fig. 5. Testing methodology of global market investment strategies.

icy and interest rate, this is meaningful improvement. More than almost 3% enhancement in the directional accuracy could result in significant enhancement in the return on investment. Among the machine learning techniques, SVM shows the best performance in forecasting directions whereas the accuracy enhancement with the addition of network indicators is similar for all the techniques.

Table 5 shows the results of testing accuracy for the relative directions of DM/EM. Although the accuracy of the relative directional forecasting is lower than that of the stock market directional forecasting in terms of absolute level, most of the effects of the network indicators on the forecasting period over 8 weeks are similar to those on the stock market forecasting. We can confirm that network indicators are good support indicators for relative directional forecasting in the mid-term forecasting period. In addition, the relative directional forecasting does not show improvement differences over the measurement period of volatility. The SVM model shows the best performance similar to a stock prediction strategy. However, the LR model does not show enhancement with the addition of network indicators. Tables 4 and 5 include results of daily forecasting for 6 years (almost 1,500 attempts). This means that there are considerable attempts being made to forecast the directions of markets.

Fig. 6 shows the model accuracy with the addition of network indicators according to the parameters. Fig. 7 shows the comparison of the accuracy enhancement with and without the addition of network indicators. The absolute accuracies are the best in the SVM model for almost all the parameters. Even though accuracies are affected by the parameters, accuracy enhancement is observed for all the models and parameters. This shows that the enhancement with the addition of network indicators is robust and that they are good support for market price indicators. Considering that model accuracies and enhancements vary for different parameters, we must exercise care when selecting parameters for machine learning techniques in practice.

Tables 6 and 7 show the changes in accuracy by year. It is apparent that the effectiveness of the network indicators in the forecasts of the overall market direction is high in the period of the financial crisis in 2008. This result is similar to those of other studies such as (Billio et al., 2012; Diebold and Yilmaz, 2016), which showed the effectiveness of network indicators for the prediction

Table 5.
Result for each volatility and forecasting period in DM/EM allocation: Relative up/down accuracy of average DM stock index over average EM stock index.

Forecasting period		With network				Without network			
Machine model / Volatility period		1 week	4 weeks	8 weeks	12 weeks	1 week	4 weeks	8 weeks	12 weeks
LR	1w	58.1%	57.6%	54.9%	49.4%	57.6%	57.3%	55.9%	51.2%
	4w	57.5%	57.1%	53.1%	50.3%	57.6%	57.5%	55.5%	51.1%
	8w	57.2%	57.7%	54.6%	50.3%	57.6%	57.2%	55.3%	51.3%
	12w	56.2%	57.4%	54.6%	50.2%	57.2%	57.0%	55.6%	51.4%
	Avg	57.2%	57.4%	54.3%	50.0%	57.5%	57.2%	55.5%	51.3%
RF	1w	52.6%	54.7%	52.2%	50.5%	52.2%	53.8%	48.9%	46.2%
	4w	52.7%	54.9%	50.2%	49.8%	51.9%	54.1%	48.7%	46.1%
	8w	52.7%	56.0%	51.1%	46.7%	51.8%	53.5%	48.9%	46.8%
	12w	53.0%	56.0%	53.1%	47.7%	52.4%	53.6%	48.7%	46.7%
	Avg	52.7%	55.4%	51.6%	48.7%	52.1%	53.8%	48.8%	46.4%
SVM	1w	55.9%	60.5%	58.6%	62.3%	56.5%	59.6%	57.8%	58.3%
	4w	55.8%	61.1%	59.9%	62.0%	57.3%	59.9%	57.7%	58.2%
	8w	56.3%	62.2%	60.6%	60.5%	57.9%	59.8%	58.1%	58.9%
	12w	56.6%	60.7%	59.0%	60.0%	57.7%	59.7%	57.7%	58.3%
	Avg	56.1%	61.1%	59.5%	61.2%	57.4%	59.7%	57.8%	58.4%

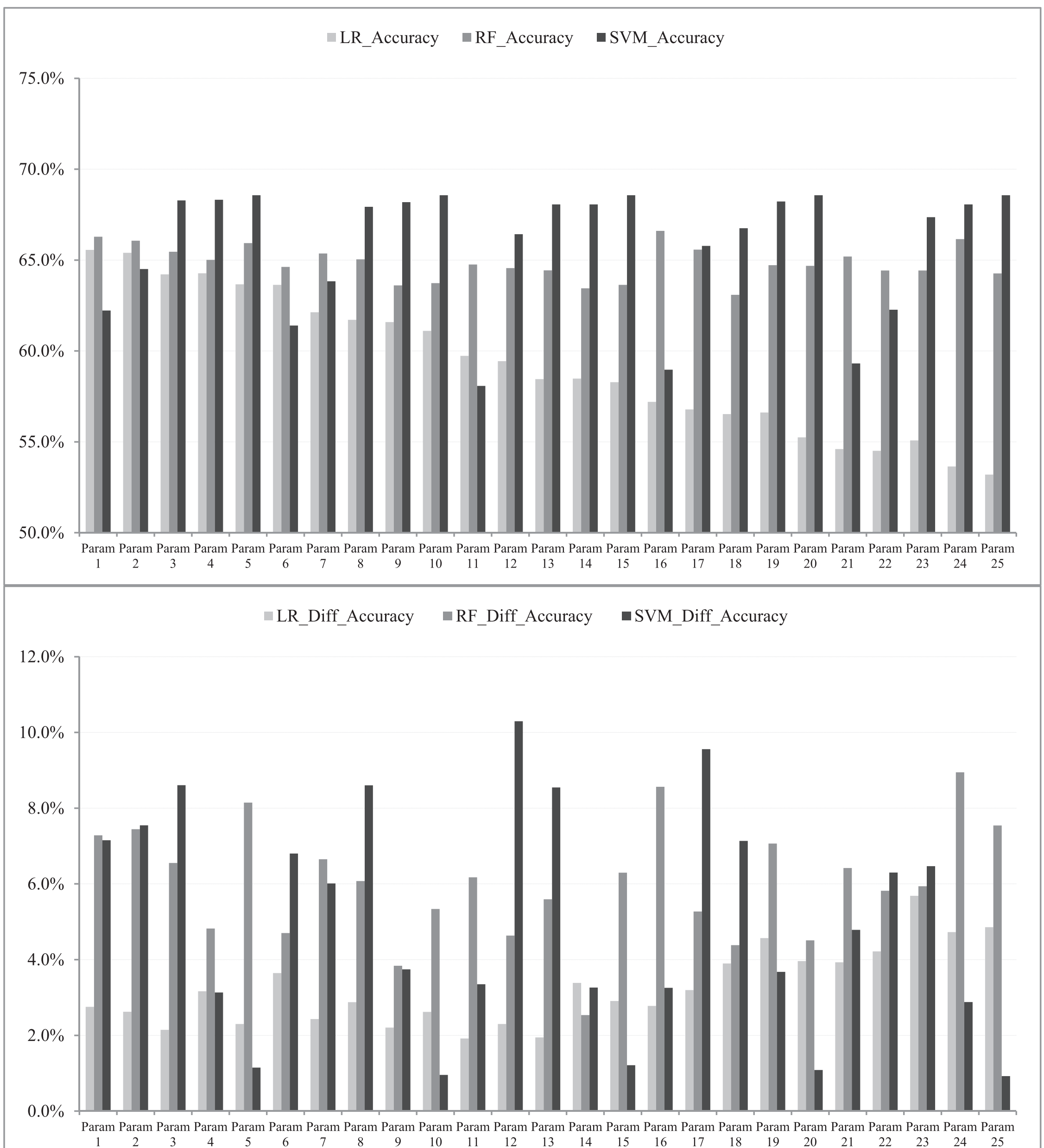

Fig. 6. Accuracy and network enhancement results for machine learning techniques for parameters in global stock market (week_vol = 1 w, week_forecasting = 12 w).

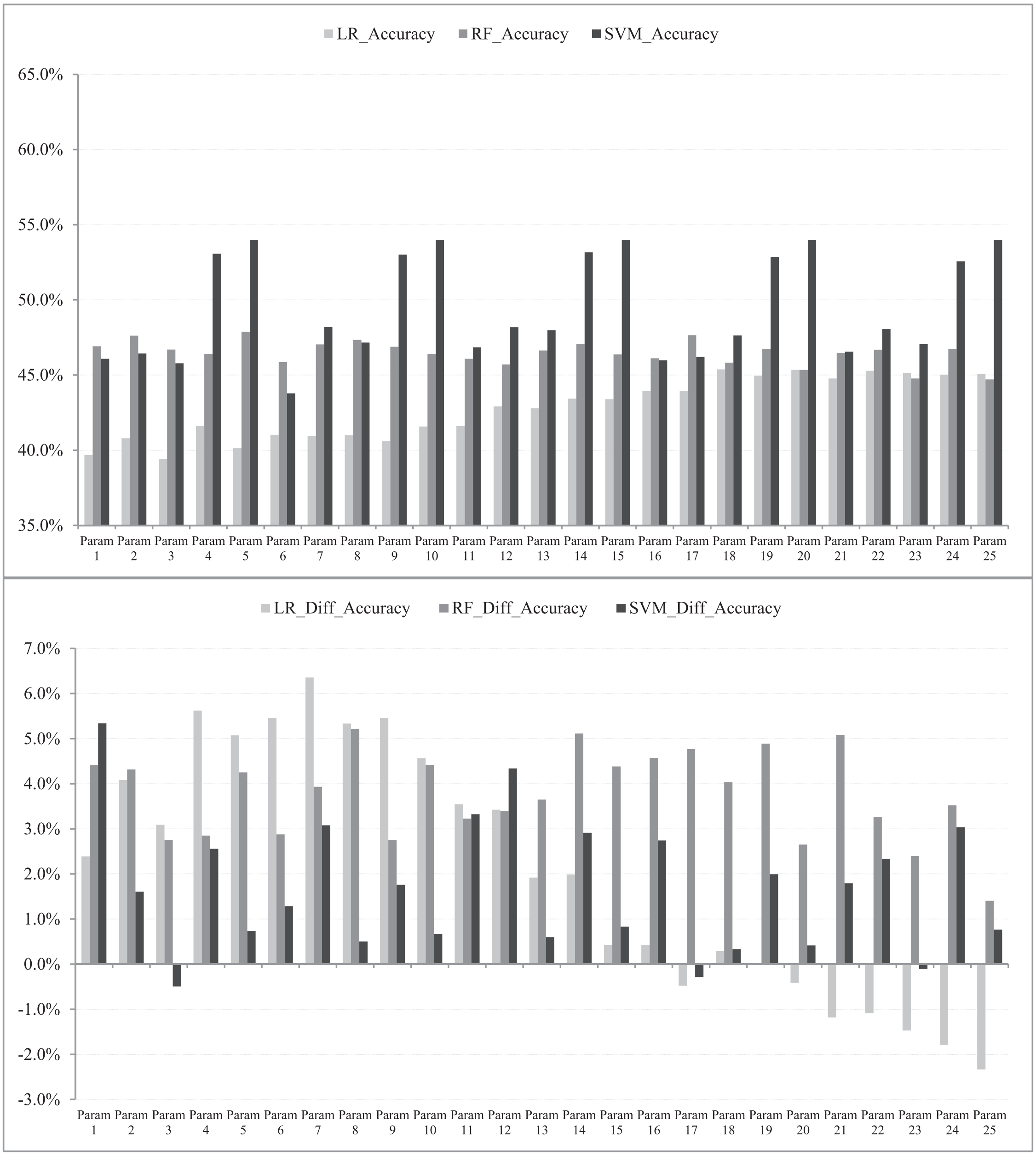

Fig. 7. Accuracy and network enhancement result for machine learning techniques for parameters in DM/EM allocation (week_vol = 1 w, week_forecasting = 12 w).

Table 6.
Result for each prediction year in global stock market: (week_vol = 1 w, week_forecasting = 12 w).

Year	Logistic		Random forest		SVM		Avg_acc	Avg_Diff
	acc	diff_acc	acc	diff_acc	acc	diff_acc		
2005	83.5%	0.0%	83.4%	−0.1%	83.5%	0.0%	83.5%	0.0%
2006	81.3%	11.3%	83.5%	22.1%	82.4%	0.9%	82.4%	11.4%
2007*	65.8%	11.0%	78.0%	8.8%	94.6%	0.5%	79.5%	6.8%
2008*	48.2%	10.9%	63.1%	21.8%	36.0%	18.8%	49.1%	17.2%
2009*	64.2%	−16.9%	64.5%	−2.0%	84.7%	6.5%	71.1%	−4.1%
2010	70.5%	1.3%	74.2%	0.6%	80.2%	7.1%	74.9%	3.0%
2011	72.6%	28.5%	52.5%	10.2%	56.2%	−15.6%	60.4%	7.7%
2012	68.8%	0.0%	68.5%	0.2%	69.7%	1.9%	69.0%	0.7%
2013	74.3%	0.0%	66.4%	−5.8%	74.9%	−3.3%	71.9%	−3.0%
2014	70.0%	−2.1%	71.5%	−0.5%	75.6%	2.1%	72.4%	−0.2%
2015	56.2%	20.8%	58.5%	16.2%	62.8%	0.0%	59.2%	12.4%
2016	67.6%	−10.8%	63.3%	−3.2%	67.6%	2.9%	66.1%	−3.7%
Total Avg	68.6%	4.5%	68.9%	5.7%	72.3%	1.8%	70.0%	4.0%

*Global financial crisis period.

Table 7.
Result for each prediction year in DM/EM allocation: (week_vol = 1 w, week_forecasting = 12 w).

Year	Logistic		Random forest		SVM		Avg_acc	Avg_Diff
	acc	diff_acc	acc	diff_acc	acc	diff_acc		
2005	75.5%	10.0%	62.8%	−0.1%	62.8%	0.0%	67.0%	3.3%
2006	62.7%	4.5%	60.0%	17.5%	88.1%	0.1%	70.3%	7.4%
2007*	46.2%	−11.8%	66.3%	9.0%	98.1%	4.3%	70.2%	0.5%
2008*	52.0%	14.4%	44.3%	−4.4%	61.5%	0.3%	52.6%	3.4%
2009*	89.0%	14.3%	78.7%	−3.3%	74.7%	−7.9%	80.8%	1.0%
2010	55.9%	−12.3%	60.7%	2.0%	68.7%	11.6%	61.8%	0.4%
2011	44.5%	−7.0%	49.3%	4.2%	54.3%	9.1%	49.4%	2.1%
2012	12.1%	−8.5%	17.8%	3.0%	20.8%	4.9%	16.9%	−0.2%
2013	13.8%	0.0%	28.4%	8.9%	40.6%	22.4%	27.6%	10.4%
2014	27.7%	−16.6%	38.2%	−6.9%	65.9%	−1.7%	43.9%	−8.4%
2015	50.3%	−4.8%	47.2%	14.1%	49.8%	0.0%	49.1%	3.1%
2016	63.1%	−4.4%	52.7%	8.1%	62.6%	5.6%	59.5%	3.1%
Total Avg	49.4%	−1.8%	50.5%	4.3%	62.3%	4.1%	54.1%	2.2%

*Global financial crisis period.

Table 8.
Simulation performances of testing period by volatility periods, forecasting periods, and machine learning models: stock market prediction strategy.

Forecasting period		With network				Without network			
Machine model / Volatility period		1 week	4 weeks	8 weeks	12 weeks	1 week	4 weeks	8 weeks	12 weeks
LR	1w	17.1%	13.4%	11.9%	**12.5%**	16.1%	14.8%	10.6%	10.5%
	4w	16.1%	12.8%	11.3%	**10.9%**	16.4%	14.9%	11.0%	10.5%
	8w	16.7%	13.0%	10.9%	10.2%	15.8%	14.6%	10.8%	10.6%
	12w	17.1%	14.4%	11.6%	**10.9%**	16.1%	14.8%	10.8%	10.7%
	Avg	16.8%	13.4%	11.4%	**11.1%**	16.1%	14.8%	10.8%	10.6%
RF	1w	10.1%	9.2%	10.0%	11.2%	9.9%	7.4%	9.4%	8.9%
	4w	9.9%	7.2%	9.9%	10.4%	9.1%	7.2%	9.2%	8.7%
	8w	11.3%	8.7%	10.3%	9.8%	9.0%	7.5%	9.2%	8.9%
	12w	12.1%	8.0%	9.9%	10.9%	9.5%	7.4%	9.2%	8.7%
	Avg	10.8%	8.2%	10.0%	10.5%	9.4%	7.4%	9.3%	8.8%
SVM	1w	14.6%	15.1%	14.2%	**13.9%**	14.5%	12.5%	12.5%	12.2%
	4w	12.8%	14.5%	12.2%	**13.2%**	14.1%	12.5%	12.7%	12.2%
	8w	13.8%	13.3%	12.8%	**13.2%**	13.4%	12.9%	12.9%	12.3%
	12w	14.3%	13.2%	12.5%	**13.1%**	13.9%	12.6%	12.1%	12.4%
	Avg	13.9%	14.0%	12.9%	**13.3%**	14.0%	12.6%	12.5%	12.3%
Long only Portfolio return		11.4%	11.6%	11.8%	12.1%	11.4%	11.6%	11.8%	12.1%

of financial risk. In addition, the network indicators are also good support indicators even after the period of financial crisis and the addition of network indicators shows poorer performance for only two years over twelve years as compared with the models without network indicators. The relative forecasting of DM/EM shows similar results. The periods of accuracy enhancement are different between the overall stock market and regional relative forecasting as the cycles of regional relative direction are different from those of the overall stock markets.

3.2. Profit performance with global investment strategy

The forecasting directions for stock markets are not equally linked to portfolio strategy simulation performances because some periods with high volatility are more important than the other periods. Even if the investment strategy shows high accuracy, the portfolio return can be low. Therefore, we validate the significance of network indicators for profit performance by simulating portfolio investment.

Tables 8 and 9 show the results of simulation of the actual investment portfolio according to the model period, investment period, and machine learning model. From the assessed accuracy, it can be observed that, in the prediction of the global stock market, the relative performance improvement of 1–3% per year is higher than the returns obtained using long-only portfolio strategy and without using network measures. In addition, the prediction of the regional allocation strategy of DM/EM shows similar results. It not only shows a 3–10% annual return improvement rate compared with the simple long DM-short EM strategy, but also shows a 1–3% improvement compared with the model without network indicators. In the regional allocation strategy, even though the LR model achieves the highest return in the short-term forecasting period, higher returns are achieved overall with the SVM model in terms of absolute and relative returns.

Thus, the use of network indicators is effective for global stock portfolio strategy. Without using network indicators, the performance of this model has difficulty in exceeding that of the reference portfolio. In the regional allocation strategy, the effectiveness of network indicators was more significant than expected. The reason is that accuracy may be high in periods with large volatile returns, which achieves high returns in simulation. Compared with other machine learning techniques, the performance enhancement of the SVM model is shown to be higher overall than those of the RF and LR models. Even though higher accuracies are not necessary for achieving better and more stable returns, the simulation returns increase with improving accuracies in this paper. Trading costs including taxes and slippages are not considered in simulation performances.

Table 9.
Simulation performances of testing period by volatility periods, forecasting periods, and machine learning models: DM/EM regional allocation strategy.

Forecasting period		With network				Without network			
Machine Model / Volatility period		1 week	4 weeks	8 weeks	12 weeks	1 week	4 weeks	8 weeks	12 weeks
LR	1w	14.0%	7.5%	4.9%	1.5%	12.7%	7.2%	3.8%	2.1%
	4w	12.7%	7.0%	3.6%	1.8%	12.7%	7.6%	3.7%	2.0%
	8w	12.5%	6.3%	3.5%	1.5%	12.7%	7.1%	3.6%	2.0%
	12w	10.4%	6.5%	4.1%	1.3%	11.9%	6.9%	3.8%	2.1%
	Avg	12.4%	6.8%	4.0%	1.5%	12.5%	7.2%	3.7%	2.0%
RF	1w	4.6%	3.8%	2.4%	1.6%	3.4%	3.9%	0.9%	0.4%
	4w	4.5%	4.3%	1.2%	1.0%	3.1%	3.7%	0.7%	0.3%
	8w	5.1%	5.3%	2.4%	0.3%	3.2%	3.6%	0.7%	0.6%
	12w	5.6%	5.8%	3.6%	1.0%	3.5%	3.5%	0.8%	0.6%
	Avg	4.9%	4.8%	2.4%	1.0%	3.3%	3.7%	0.8%	0.5%
SVM	1w	9.2%	8.7%	7.6%	9.2%	9.8%	7.9%	6.6%	6.8%
	4w	9.1%	9.3%	8.4%	8.8%	10.5%	7.9%	6.6%	6.6%
	8w	10.5%	10.6%	8.6%	8.3%	11.6%	7.8%	6.8%	6.9%
	12w	10.9%	9.0%	7.8%	8.3%	11.2%	8.0%	6.7%	6.9%
	Avg	9.9%	9.4%	8.1%	8.6%	10.8%	7.9%	6.7%	6.8%
DM-EM Portfolio Return		−1.9%	−2.2%	−2.5%	−2.8%	−1.9%	−2.2%	−2.5%	−2.8%

4. Discussion and conclusion

For network analysis in the field of finance, symptoms of market crisis and structural changes could be closely described. However, practical usefulness studies are scant in this regard. The purpose of this study was to examine the effectiveness and enhancement of a financial network based on global stock indices of 10 countries in building a global stock portfolio strategy. In addition, three machine learning techniques such as LR, RF, and SVM were used for the prediction of the global financial market. Through the detailed analysis according to forecasting periods, we investigated the effects of network indicators on the prediction of the global stock. Moreover, volatility period analysis could suggest an effective period for prediction markets.

The results of the study showed that the global stock market network indicators are effective for forecasting market direction (up and down) during the market turmoil period as expected. Moreover, mid-term investments with short-term volatility showed better performance than other strategies in terms of stock market prediction. The better performance obtained over one week is because the financial network measure does not immediately affect the market rise or fall but affects the implicit market risk, eventually influencing the market direction in the mid-term periods. The short-term volatility shows better performance among the other volatility periods as short-term is a more sensitive measure than others. It is regarded as important in practical fields and many practitioners focus on the short-term volatility measure for investment as well.

In addition, the network indicators can be performance enhancers for the regional allocation strategy, especially during market volatility period. Further, the information of network in the regional allocation level is stably effective similar to that in the total market direction level over the prediction and volatility periods. The network measures are not used in markets for regional allocation yet, but we observed that they are good indicators for regional direction forecasting and allocation strategies.

The usefulness of the network indicators can be confirmed based on the average accuracies of the entire parameters as shown in the appendix Tables A1 and A2 (results for all parameters). In addition, we observe the same results that performances are improved during the period of the financial crisis. However, in investment simulations, average performances are not good enough, with low returns for some parameter combinations. It is necessary to improve parameter selection techniques to obtain more robust simulation results. There were differences in the performance results among the three machine learning methodologies employed in this study. In the prediction model combining the price index and the network indicators, the effectiveness of SVM, a non-parametric approach, was higher than that of LR or RF. It can be concluded that the methodology of SVM is more appropriate for forecasting the market through a non-linear relationship among variables.

The limitations of this study are as follows. First, in this study, we have used only a connectedness measure but we must develop other network indicators for capturing diverse market dynamics. Second, the three machine learning techniques are not sufficient for measuring the effects of numerous latent factors in complex financial markets. Therefore, a new methodology such as model improvement or deep learning will be considered in the future. Finally, forecasting indicators such as price and network indicators have only been used whereas other indicators that are already considered to be important in the market such as valuation and economic indicators have not been used (e.g., P/E ratio, FX). These limitations will be addressed in future studies.

All persons who meet authorship criteria are listed as authors, and all authors certify that they have participated sufficiently in the work to take public responsibility for the content, including participation in the concept, design, analysis, writing, or revision of the manuscript. Furthermore, each author certifies that this material or similar material has not been and will not be submitted to or published in any other publication before its appearance in the *Expert Systems with Applications.*

Author's contributions

Category 1: Conception and design of study: T.K. Lee, J.H. Cho, D.S. Kwon; acquisition of data: J.H. Cho, D.S. Kwon; machine learning modeling: J.H. Cho, D.S. Kwon, T.K. Lee; analysis and interpretation of data: T.K. Lee, J.H. Cho, D.S. Kwon, S.Y. Sohn. project administration: S.Y. Sohn.

Category 2: Drafting the manuscript: T.K. Lee, J.H. Cho, S.Y. Sohn, D.S. Kwon; revising the manuscript critically for important intellectual content: T.K. Lee, D.S. Kwon, S.Y. Sohn.

Category 3: Approval of the version of the manuscript to be published: T.K. Lee, J.H. Cho, D.S. Kwon, S.Y. Sohn; funding acquisition: S.Y. Sohn.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (2016R1A2A1A05005270).

Table. A1.
Result for each volatility and forecasting period in global stock market: Up/Down accuracy of average global stock index direction: the averages of all parameters.

Forecasting period		With network				Without network			
Machine Learning Model / Volatility period		1 week	4 weeks	8 weeks	12 weeks	1 week	4 weeks	8 weeks	12 weeks
LR	1w	53.4%	57.2%	56.2%	59.2%	53.5%	56.0%	53.3%	56.0%
	4w	52.8%	55.5%	53.6%	56.4%	53.1%	55.9%	53.6%	55.9%
	8w	52.8%	55.3%	53.7%	55.9%	53.3%	56.1%	53.2%	56.0%
	12w	53.7%	57.1%	55.3%	57.3%	53.4%	56.0%	53.4%	56.1%
	Avg	53.1%	56.3%	54.7%	57.2%	53.3%	56.0%	53.4%	56.0%
RF	1w	52.5%	55.9%	58.9%	64.8%	52.0%	53.4%	57.7%	58.8%
	4w	51.7%	53.3%	57.8%	61.0%	52.0%	53.1%	57.7%	58.8%
	8w	51.9%	54.2%	59.7%	60.2%	52.0%	53.2%	57.7%	58.5%
	12w	53.1%	51.9%	56.7%	61.2%	52.3%	53.0%	57.7%	58.3%
	Avg	52.3%	53.9%	58.3%	61.8%	52.1%	53.2%	57.7%	58.6%
SVM	1w	54.0%	61.6%	62.9%	65.8%	53.5%	56.7%	61.0%	60.7%
	4w	55.2%	59.8%	62.2%	65.1%	53.3%	57.2%	61.0%	60.5%
	8w	54.3%	59.8%	63.2%	65.9%	53.4%	57.2%	61.2%	60.8%
	12w	54.1%	58.2%	62.6%	65.3%	53.5%	57.0%	60.8%	60.8%
	Avg	54.4%	59.9%	62.7%	65.5%	53.4%	57.0%	61.0%	60.7%

Table. A2.
Result for each volatility and forecasting period in DM/EM allocation: Relative up/down accuracy of average DM stock index over average EM stock index: the averages of all parameters.

Forecasting period		With network				Without network			
Machine Learning Model / Volatility period		1 week	4 weeks	8 weeks	12 weeks	1 week	4 weeks	8 weeks	12 weeks
LR	1w	49.2%	47.9%	44.8%	42.8%	49.0%	47.8%	43.4%	40.8%
	4w	49.0%	47.1%	43.8%	43.2%	48.9%	47.9%	43.2%	40.8%
	8w	48.8%	47.4%	44.6%	42.9%	49.0%	47.8%	43.3%	40.8%
	12w	48.9%	48.1%	45.1%	42.8%	48.9%	47.7%	43.5%	40.9%
	Avg	49.0%	47.6%	44.6%	42.9%	49.0%	47.8%	43.4%	40.8%
RF	1w	48.7%	50.9%	47.4%	46.5%	48.6%	49.5%	45.1%	42.7%
	4w	49.2%	50.6%	46.4%	45.5%	48.6%	49.7%	44.9%	42.6%
	8w	49.1%	52.0%	46.8%	42.3%	48.5%	49.5%	45.0%	43.0%
	12w	49.4%	50.9%	48.5%	43.1%	48.8%	49.8%	45.1%	42.8%
	Avg	49.1%	51.1%	47.3%	44.4%	48.6%	49.6%	45.0%	42.8%
SVM	1w	50.5%	51.1%	48.9%	49.5%	51.0%	51.3%	49.6%	47.8%
	4w	50.2%	51.6%	48.1%	51.1%	50.9%	50.7%	49.5%	47.5%
	8w	49.9%	53.7%	49.8%	49.8%	51.3%	50.8%	49.6%	47.8%
	12w	50.2%	51.2%	49.0%	49.5%	51.4%	51.0%	49.3%	47.7%
	Avg	50.2%	51.9%	49.0%	50.0%	51.1%	50.9%	49.5%	47.7%

Appendix A

Tables A3–A8.

Table. A3.
Result for each prediction year in global stock market: (week_vol = 1 w, week_forecasting = 12 w): the averages of all parameters.

Year	Logistic		Random forest		SVM		Avg_acc	Avg_Diff
	acc	diff_acc	acc	diff_acc	acc	diff_acc		
2005	83.2%	0.0%	81.6%	1.9%	82.3%	0.8%	82.3%	0.9%
2006	71.7%	12.8%	79.0%	22.7%	78.2%	7.9%	76.3%	14.5%
2007	34.7%	4.5%	73.9%	11.9%	76.3%	12.0%	61.6%	9.5%
2008	39.5%	8.2%	54.5%	18.6%	21.7%	6.5%	38.6%	11.1%
2009	54.3%	−3.2%	56.1%	−4.0%	80.4%	14.7%	63.6%	2.5%
2010	64.6%	7.4%	72.3%	2.8%	73.1%	7.5%	70.0%	5.9%
2011	53.2%	12.3%	46.9%	7.4%	47.0%	−7.2%	49.0%	4.2%
2012	63.2%	−3.8%	67.3%	2.4%	68.6%	12.5%	66.4%	3.7%
2013	70.7%	−3.2%	61.6%	−6.3%	70.6%	−4.5%	67.6%	−4.7%
2014	55.3%	−13.7%	69.6%	−0.7%	71.8%	3.9%	65.6%	−3.5%
2015	54.1%	22.9%	54.3%	18.3%	59.3%	3.1%	55.9%	14.7%
2016	66.5%	−5.3%	61.2%	−2.6%	60.4%	3.5%	62.7%	−1.5%
Total Avg	59.2%	3.2%	64.8%	6.0%	65.8%	5.1%	63.3%	4.8%

Table. A4.
Result for each prediction year in DM/EM allocation: (week_vol = 1 w, week_forecasting = 12 w): the averages of all parameters.

Year	Logistic		Random forest		SVM		Avg_acc	Avg_Diff
	acc	diff_acc	acc	diff_acc	acc	diff_acc		
2005	67.1%	3.6%	62.7%	−0.1%	62.3%	−0.2%	64.0%	1.1%
2006	52.5%	6.2%	54.2%	15.2%	69.1%	5.8%	58.6%	9.1%
2007	39.4%	1.9%	61.0%	10.0%	75.0%	18.2%	58.5%	10.1%
2008	39.0%	5.9%	42.4%	−1.5%	40.8%	−2.9%	40.7%	0.5%
2009	72.0%	13.3%	76.1%	−1.2%	62.4%	−15.3%	70.2%	−1.1%
2010	48.3%	−4.3%	55.4%	0.3%	53.8%	3.6%	52.5%	−0.1%
2011	43.9%	−2.5%	45.9%	2.1%	47.6%	3.6%	45.8%	1.0%
2012	9.8%	−1.8%	13.4%	1.6%	13.9%	2.2%	12.4%	0.6%
2013	13.8%	0.0%	20.3%	3.6%	23.9%	8.6%	19.3%	4.1%
2014	24.5%	−3.8%	31.9%	−8.9%	47.0%	−13.6%	34.5%	−8.7%
2015	48.8%	5.2%	43.2%	15.3%	45.9%	2.1%	45.9%	7.5%
2016	54.3%	0.5%	51.3%	8.8%	52.0%	8.2%	52.5%	5.8%
Total Avg	42.8%	2.0%	46.5%	3.8%	49.5%	1.7%	46.2%	2.5%

Table. A5.
Simulation performances of testing period by volatility periods, forecasting periods, and machine learning models: stock market prediction strategy: the averages of all parameters.

Forecasting period		With Network				Without Network			
Machine Model / Volatility period		1 week	4 weeks	8 weeks	12 weeks	1 week	4 weeks	8 weeks	12 weeks
LR	1w	8.0%	8.4%	7.4%	8.5%	7.8%	7.6%	5.0%	6.1%
	4w	7.5%	7.4%	5.9%	6.9%	7.6%	7.5%	5.1%	6.1%
	8w	7.5%	7.3%	5.8%	6.3%	7.6%	7.6%	5.0%	6.1%
	12w	8.6%	8.9%	7.1%	7.2%	7.7%	7.5%	5.1%	6.1%
	Avg	7.9%	8.0%	6.5%	7.2%	7.7%	7.6%	5.0%	6.1%
RF	1w	6.1%	6.9%	8.0%	9.7%	5.9%	5.6%	7.5%	7.0%
	4w	6.1%	5.3%	7.7%	8.6%	5.8%	5.4%	7.5%	7.0%
	8w	7.0%	6.0%	8.4%	7.9%	5.8%	5.5%	7.5%	6.8%
	12w	8.4%	5.4%	7.7%	9.0%	6.1%	5.4%	7.5%	6.7%
	Avg	6.9%	5.9%	7.9%	8.8%	5.9%	5.5%	7.5%	6.9%
SVM	1w	8.3%	10.5%	10.8%	10.8%	7.8%	7.2%	9.3%	7.9%
	4w	9.3%	9.8%	10.5%	11.0%	7.7%	7.5%	9.3%	7.8%
	8w	8.5%	9.4%	11.1%	10.7%	7.5%	7.5%	9.3%	7.9%
	12w	8.7%	9.1%	10.9%	10.9%	7.8%	7.3%	9.1%	7.9%
	Avg	8.7%	9.7%	10.8%	10.8%	7.7%	7.4%	9.2%	7.9%
Long only Portfolio Return		11.4%	11.6%	11.8%	12.1%	11.4%	11.6%	11.8%	12.1%

Table. A6.
Simulation performances of testing period by volatility periods, forecasting periods, and machine learning models: DM/EM regional allocation strategy: the averages of all parameters.

Forecasting period		With network				Without network			
Machine Model / Volatility period		1 week	4 weeks	8 weeks	12 weeks	1 week	4 weeks	8 weeks	12 weeks
LR	1w	−0.1%	−1.5%	−2.0%	−1.8%	−0.6%	−2.1%	−3.9%	−3.5%
	4w	−0.4%	−2.2%	−2.6%	−1.5%	−0.7%	−1.9%	−3.9%	−3.5%
	8w	−0.6%	−2.7%	−2.9%	−2.1%	−0.5%	−2.1%	−4.0%	−3.5%
	12w	−0.8%	−1.4%	−2.0%	−2.1%	−0.7%	−2.2%	−3.9%	−3.5%
	Avg	−0.5%	−2.0%	−2.4%	−1.9%	−0.6%	−2.1%	−3.9%	−3.5%
RF	1w	−0.6%	1.1%	−0.1%	0.2%	−1.2%	0.8%	−1.2%	−1.2%
	4w	0.3%	1.1%	−0.7%	−0.6%	−1.2%	0.8%	−1.3%	−1.3%
	8w	0.0%	2.2%	−0.1%	−1.8%	−1.2%	0.7%	−1.2%	−1.1%
	12w	0.7%	1.8%	1.0%	−1.1%	−0.9%	0.9%	−1.1%	−1.1%
	Avg	0.1%	1.5%	0.0%	−0.8%	−1.1%	0.8%	−1.2%	−1.2%
SVM	1w	1.4%	0.8%	0.8%	2.2%	1.7%	0.6%	2.0%	1.4%
	4w	1.0%	1.5%	0.6%	2.9%	1.4%	0.2%	2.1%	1.2%
	8w	1.1%	4.0%	1.8%	2.9%	2.0%	0.2%	2.0%	1.4%
	12w	0.7%	1.2%	1.4%	3.0%	2.0%	0.5%	1.9%	1.3%
	Avg	1.1%	1.9%	1.2%	2.7%	1.8%	0.4%	2.0%	1.3%
DM-EM Portfolio Return		−1.9%	−2.2%	−2.5%	−2.8%	−1.9%	−2.2%	−2.5%	−2.8%

Table. A7.
Simulation performances by models and prediction years.

Year	Logistic			Random Forest			SVM			Avg. Ret	Avg. Diff_A	Avg. Diff_B
	Ret	Diff_A	Diff_B	Ret	Diff_A	Diff_B	Ret	Diff_A	Diff_B			
2005	19.5%	0.0%	0.0%	19.3%	−0.1%	−0.1%	19.5%	0.0%	0.0%	19.4%	0.0%	0.0%
2006	30.2%	0.1%	5.9%	28.5%	−1.7%	7.8%	29.8%	−0.3%	−0.3%	29.5%	−0.6%	4.5%
2007	24.9%	−14.7%	3.6%	30.6%	−9.0%	3.9%	39.6%	0.0%	0.3%	31.7%	−7.9%	2.6%
2008	−18.5%	23.5%	13.3%	−22.0%	20.0%	11.1%	−25.9%	16.1%	16.1%	−22.2%	19.8%	13.5%
2009	40.6%	−13.5%	−11.5%	36.3%	−17.8%	1.8%	54.1%	0.0%	10.1%	43.7%	−10.4%	0.2%
2010	10.1%	0.0%	0.6%	11.8%	1.7%	0.0%	13.7%	3.6%	1.9%	11.9%	1.8%	0.8%
2011	0.7%	12.0%	12.0%	−10.3%	1.0%	0.4%	−9.7%	1.6%	−9.5%	−6.4%	4.9%	1.0%
2012	9.6%	0.0%	0.0%	9.3%	−0.2%	0.4%	9.9%	0.3%	0.7%	9.6%	0.0%	0.4%
2013	15.3%	0.0%	0.0%	12.4%	−2.9%	−0.9%	15.1%	−0.2%	−0.1%	14.3%	−1.0%	−0.3%
2014	8.1%	−1.0%	−0.6%	8.6%	−0.5%	0.0%	9.2%	0.1%	0.1%	8.6%	−0.4%	−0.2%
2015	3.8%	−1.6%	4.9%	5.3%	−0.1%	4.7%	5.4%	0.0%	0.0%	4.8%	−0.6%	3.2%
2016	5.8%	0.0%	−4.3%	4.5%	−1.4%	−2.4%	5.8%	0.0%	0.9%	5.4%	−0.5%	−1.9%
Total Avg.	12.5%	0.4%	2.0%	11.2%	−0.9%	2.2%	13.9%	1.8%	1.7%	12.5%	0.4%	2.0%
2005	7.8%	10.5%	4.1%	2.7%	5.4%	0.0%	2.7%	5.4%	0.0%	4.4%	7.1%	1.4%
2006	6.6%	22.7%	1.5%	5.4%	21.5%	10.4%	19.2%	35.3%	0.0%	10.4%	26.5%	4.0%
2007	−3.2%	27.0%	−8.7%	15.3%	45.4%	6.3%	43.1%	73.2%	5.3%	18.4%	48.5%	1.0%
2008	1.4%	−10.0%	7.8%	−7.3%	−18.7%	−5.3%	7.2%	−4.1%	−0.2%	0.4%	−10.9%	0.8%
2009	38.3%	64.1%	3.8%	25.9%	51.6%	−9.2%	34.5%	60.2%	2.7%	32.9%	58.6%	−0.9%
2010	2.9%	3.2%	−2.6%	3.4%	3.6%	0.6%	6.1%	6.3%	3.8%	4.2%	4.4%	0.6%
2011	−5.3%	−11.2%	−3.0%	−3.5%	−9.4%	1.7%	−1.3%	−7.1%	3.7%	−3.4%	−9.2%	0.8%
2012	−9.7%	−21.4%	−1.9%	−8.9%	−20.6%	0.5%	−7.0%	−18.7%	1.1%	−8.5%	−20.2%	−0.1%
2013	−17.5%	−38.7%	0.0%	−12.6%	−33.8%	3.6%	−4.1%	−25.3%	12.1%	−11.4%	−32.6%	5.2%
2014	−7.8%	−2.3%	−6.3%	−3.2%	2.3%	−2.6%	5.8%	11.3%	−2.2%	−1.7%	3.8%	−3.7%
2015	−0.3%	−0.7%	−1.5%	−1.0%	−1.4%	4.9%	−0.4%	−0.8%	0.0%	−0.6%	−1.0%	1.2%
2016	4.1%	8.0%	−0.8%	2.9%	6.8%	2.8%	4.0%	7.9%	2.4%	3.7%	7.6%	1.5%
Total Avg.	1.5%	4.3%	−0.6%	1.6%	4.4%	1.2%	9.2%	12.0%	2.4%	4.1%	6.9%	1.0%

Stock market prediction strategy (week_vol = 1 w, week_forecasting = 12 w).
Diff_A: Strategy Return - Return(long only portfolio).
Diff_B: Strategy Return - Return(without network indicators).
DM/EM regional allocation strategy (week_vol = 1 w, week_forecasting = 12 w).
Diff_A: Strategy Return - Return(DM-EM portfolio).
Diff_B: Strategy Return - Return(without network indicators).

Table. A8.
Parameters list used in Figs. 6 & 7 by models.

List of parameters	LR p	RF k	SVM p	γ	c
Param 1	0.38	2	100	2	0.25
Param 2	0.39	2	150	2	0.5
Param 3	0.4	2	200	2	1
Param 4	0.41	2	250	2	2
Param 5	0.42	2	300	2	4
Param 6	0.43	4	100	4	0.25
Param 7	0.44	4	150	4	0.5
Param 8	0.45	4	200	4	1
Param 9	0.46	4	250	4	2
Param 10	0.47	4	300	4	4
Param 11	0.48	6	100	6	0.25
Param 12	0.49	6	150	6	0.5
Param 13	0.5	6	200	6	1
Param 14	0.51	6	250	6	2
Param 15	0.52	6	300	6	4
Param 16	0.53	8	100	8	0.25
Param 17	0.54	8	150	8	0.5
Param 18	0.55	8	200	8	1
Param 19	0.56	8	250	8	2
Param 20	0.57	8	300	8	4
Param 21	0.58	10	100	10	0.25
Param 22	0.59	10	150	10	0.5
Param 23	0.6	10	200	10	1
Param 24	0.61	10	250	10	2
Param 25	0.62	10	300	10	4

References

Abreu, D., & Brunnermeier, M. K. (2003). Bubbles and crashes. *Econometrica*, 173–204.

Allen, F., & Gale, D. (2000). Financial contagion. *Journal of Political Economy, 108*, 1–33.

Aloui, R., Aïssa, M. S. B., & Nguyen, D. K. (2011). Global financial crisis, extreme interdependences, and contagion effects: The role of economic structure? *Journal of Banking & Finance, 35*, 130–141.

Barak, S., & Modarres, M. (2015). Developing an approach to evaluate stocks by forecasting effective features with data mining methods. *Expert Systems with Applications, 42*(3), 1325–1339.

Barunik, J., Kocenda, E., & Vacha, L. (2017). Asymmetric volatility connectedness on the forex market. *Journal of International Money and Finance, 77*, 39–56.

Barigozzi, M., & Hallin, M. (2016). A network analysis of the volatility of high dimensional financial series. *Journal of Royal Statistical Society: Series C, 66*, 581–605.

Bekaert, G., Ehrmann, M., Fratzscher, M., & Mehl, A. (2014). The global crisis and equity market contagion. *The Journal of Finance*, (November).

Billio, M., Getmansky, M., Lo, AW., & Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics, 104*, 535–559.

Brown, G. W. (2005). Investor sentiment and asset valuation. *The Journal of Business*.

Castiglionesi, F., Feriozzi, F., & Lorenzoni, G. (2009). *Financial integration, liquidity, and the depth of systemic crisis*. Massachuseetts Institute of Technology and University of Tilburg Unpublished working paper l.

Cavalcante, R. C., Brasileiro, R. C., Souza, V. S. F., Nobrega, J. P., & Oliveira, A. L. I. (2016). Computational Intelligence and Financial Markets: A survey with future directions. *Expert Systems with Applications, 55*, 194–211.

Chi, K. T., Liu, J., & Lau, F. C. M. (2010). A network perspective of the stock market. *Journal of Empirical Finance, 17*(4), 659–667.

Chuluun, T. (2017). Global portfolio investment network and stock market comovement. *Global Finance Journal, 33*, 51–68.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning, 20*(3), 273–297.

Dai, W., Wu, J. Y., & Lu, C. J. (2012). Combining nonlinear independent component analysis and neural network for the prediction of Asian stock market indexes. *Expert systems with applications, 39*(4), 4444–4452.

Diebold, F. X., & Yilmaz, K. (2014). On the network topology of variance decompositions: measuring the connectedness of financial firms. *Journal of Econometrics, 182*, 119–134.

Diebold, F. X., & Yilmaz, K. (2016). Trans-atlantic equity volatility connectedness: U.S. and European Financial Institutions, 2004-2014. *Journal of Financial Econometrics, 14*, 81–127.

Dutta, D. R. A., Bandopadhyay, G., & Sengupta, D. R. S. (2012). Prediction of stock performance in the Indian stock market using logistic regression. *International Journal of Business and Information, 7*(1).

Figlewski (1997). Forecasting Volatility. Financial Markets. *Institutions and Instruments 6*.

Glasserman, P., & Young, H. P. (2015). How likely is contagion in financial networks? *Journal of Banking & Finance, 50*, 383–399.

Hajizadeh, E., Seifi, A., Zarandi, M. F., & Turksen, I. (2012). A hybrid modeling approach for forecasting the volatility of S&P 500 index return. *Expert Systems with Applications, 39*(1), 431–436.

Haldane, AG. 2009. “Rethinking the Financial Network.” Speech presented at the Financial Student Association, Amsterdam. http://www.bankofengland.co.uk/archive/Documents/historicpubs/speeches/2009/speech386.pdf (accessed December 12, 2014).

Hong, S. J., & Sohn, S. Y. (2013). Peer group analysis for introducing weather derivatives for a city. *Expert Systems with Applications, 40*(14), 5680–5687.

Huang, W., Nakamori, Y., & Wang, S.-Y. (2005). Forecasting stock market movement direction with support vector machine. *Computers & Operations Research, 32*, 2513–2522.

Inthachot, M., Boonjing, V., & Intakosum, S. (2016). Artificial neural network and genetic algorithm hybrid intelligence for predicting Thai stock price index trend. *Computational Intelligence and Neuroscience* Nov 15..

Jarrow, R., Kchia, Y., & Protter, P. (2011). How to detect an asset bubble, society for industrial and applied mathematics, 839–865

Kang, J., Liu, M., & Ni, S. X. (2002). Contrarian and momentum strategies in the China stock market: 1993-2002. *Pacific-Basin Finance Journal, 10*, 243–265.

Kim, H. S., & Sohn, S. Y. (2010). Support Vector Machines for Default Prediction of SMEs based on Technology Credit. *European Journal of Operational Research, 201*(3), 838–846.

Kim, Y., & Sohn, S. Y. (2012). Stock fraud detection using peer group analysis. *Expert Systems with Applications, 39*(10), 8986–8992.

Kim, Y., & Enke, D. (2018). A dynamic target volatility strategy for asset allocation using artificial neural networks. *The Engineering Economist*.

Koop, G., Pesaran, M. H., & Potter, S. M. (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics, 74*, 119–147.

Kristjanpoller, W., Fadic, A., & Minutolo, M. (2014). Volatility forecast using hybrid neural network models. *Expert Systems with Applications, 41*(5), 2437–2442.

Kristjanpoller, W., & Minutolo, M. (2016). Forecasting volatility of oil price using an artificial neural network-GARCH model. *Expert System with Applications, 65*(15), 233–241.

Lee, M. C. (2009). Using support vector machine with a hybrid feature selection method to the stock trend prediction. *Expert Systems with Applications, 36*(8), 10896–10904.

Lee, S. (2004). Application of likelihood ratio and logistic regression models to landslide susceptibility mapping using GIS. *Environmental Management, 34*(2), 223–232.

Morgan, J. P.Europe Equity Research. (2012). Market Timing Model: Methodology. *Thinking Process and Strategy to time the market*.

Morgan, J. P.Global Quantitative and Derivatives Strategy. (2014). Framework for Regional Equity Allocation: Country Selection based on Fundamental. *Macro and Technical Signals*.

Na, S. H., & Sohn, S. Y. (2011). Forecasting changes in Korea composite stock price index (KOSPI) using association rules. *Expert Systems with Applications, 38*(7), 9046–9049.

Nier, E., Yang, J., Yorulmazer, T., & Alentorn, A. (2007). Network models and financial stability. *Journal of Economic Dynamics and Control, 31*(6), 2033–2060.

Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015a). Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques. *Expert Systems with Applications, 42*, 259–268.

Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015b). Predicting stock market index using fusion of machine learning techniques. *Expert Systems with Applications, 42*(4), 2162–2172.

Peralta, G., & Zareei, A. (2016). A network approach to portfolio selection. *Journal of Empirical Finance, 38*, 157–180.

Pesaran, H. H., & Shin, Y. (1998). Generalized impulse response analysis in linear multivariate models. *Economics Letters, 58*, 17–29.

Petropoulos, A., Chatzis, S., P. , Siakoulis,V., & Vlachogiannakis, N. (2017). A stacked generalization system for automated FOREX portfolio trading. *Expert Systems with Applications, 90*, 290–302.

Piotroski, J. D. (2000). Value Investing: The Use of Historical Financial statement Information to separate Winners from Losers. *Journal of Accounting research, 38*, 1–41.

Seymen, A critical note on the forecast error variance decomposition, ZEW, 2018, Discussion Papers No. 08-065, ZEW - Zentrum für Europäische Wirtschaftsforschung/Center for European Economic.

Sohn, S. Y., Moon, T. H., & Kim, S. H. (2005). Improved Technology Scoring Model for Credit Guarantee Fund. *Expert Systems with Applications, 28*(2), 327–331.

Sohn, S. Y., & Lim, M. (2007). Hierarchical forecasting based on AR-GARCH model in a coherent structure. *European Journal of Operational Research, 176*(2), 1033–1040.

Weng, B., Lu, L., Wang, X., Megahed, F. M., & Martinez, W. (2018). Predicting Short--Term Stock Prices using Ensemble Methods and Online Data Sources. *Expert Systems with Applications*.

Werner, J. (1997). Arbitrage, Bubbles, and Valuation. *International Economic Review*.

Zhang, X., Hu, Y., Xie, K., Wang, S., Ngai, E. W. T., & Liu, M. (2014). A causal feature selection algorithm for stock prediction modeling. *Neurocomputing, 142*, 48–59.