IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received May 24, 2020, accepted June 7, 2020, date of publication June 30, 2020, date of current version July 20, 2020.

*Digital Object Identifier 10.1109/ACCESS.2020.3005994*

# Discovery and Prediction of Stock Index Pattern via Three-Stage Architecture of TICC, TPA-LSTM and Multivariate LSTM-FCNs

**HONGBING OUYANG[1], XIAOLU WEI[2], AND QIUFENG WU[3]**

[1]Department of Economics, Huazhong University of Science and Technology, Wuhan 430074, China
[2]Business School, Hubei University, Wuhan 430062, China
[3]Department of Science, Northeast Agricultural University, Harbin 150030, China

Corresponding author: Xiaolu Wei (xiaoluees@hust.edu.cn)

This work was supported by the Ministry of Education (MOE) in China Project of Humanities and Social Sciences under Project 19YJA790067.

**ABSTRACT** In this study, we attempt to discover and predict stock index patterns through analysis of multivariate time series. Our motivation is based on the notion that financial planning guided by pattern discovery and prediction of stock index prices maybe more realistic and effective than traditional approaches, such as Autoregressive Integrated Moving Average (ARIMA) model. A three-stage architecture constructed by combining Toeplitz Inverse Covariance-Based Clustering (TICC), Temporal Pattern Attention and Long-Short-Term Memory (TPA-LSTM) and Multivariate LSTM-FCNs (MLSTM-FCN and MALSTM-FCN) is applied for pattern discovery and prediction of stock index. In the first stage, we use TICC to discover repeated patterns of stock index. Then, in the second stage, TPA-LSTM that considers weak periodic patterns and long short-term information is used to predict multivariate stock indices. Finally, in the third stage, MALSTM-FCN is applied to predict stock index price pattern. The Hangseng Stock Index and eleven industrial sub-indices are used in the experiment. Empirical results show that the three-stage architecture achieves satisfactory and better performance than traditional methods, such as Naive Bayes Classifier (NB), Support Vector Machine Classifier (SVM), Random Forest (RF), etc. Moreover, we construct equal proportion portfolios based on the bullish trading rules to further analyze the feasibility of the proposed three-stage architecture. Seven comprehensive stock indices are used in the experiment. Empirical results show that the portfolio based on the proposed three-stage architecture presents better performance than the market-based portfolio. These findings may provide new direction for the portfolio construction and risk aversion.

**INDEX TERMS** Stock index pattern, pattern discovery, pattern prediction, multivariate time series.

## I. INTRODUCTION

Stock index prediction is one of the most important subjects in financial time series forecasting. Investors can invest passively through a stock index or compare active investment performance with stock index. Therefore, developing a more realistic model to predict stock index is of great importance for investors and professional analysts. However, stock index characteristics, including ''noisy'' and ''non-stationary'', make prediction face challenges. ''Noisy'' implies that there is insufficient information for investors to observe past behaviors of stock index. 'Nonstationary' means that stock index may change dramatically in different periods. These characteristics lead to poor stock index prediction results as predicted by traditional econometric models such as linear model, Auto-Regressive Integrated Moving Average (ARIMA), and Vector AutoRegression (VAR) [1]–[3]. The aforementioned methods belong to short-term predictions in time series, which are seriously affected by ''noisy'' and ''non-stationary''. However, if stock index prediction only focuses on forecasting the trend over a certain period, the effects of ''noisy'' and ''non-stationary'' on the prediction results will be eliminated. One of methods of stock index trend prediction over a certain period is to decompose stock

The associate editor coordinating the review of this manuscript and approving it for publication was Junxiu Liu.

index over long-period into many stock index fragments over short-period, and then stock index fragments over short-period are classified into some pattern in pattern sets including "W" shape, "head and shoulders", etc. The core of the process above is to find some inherent patterns in the existing stock index sequences and make appropriate prediction, which is referred to as "pattern discovery" [4] and "pattern prediction". According to the repeated patterns which were discovered in previous work, we can predict repeated patterns of stock index over a certain period in the future, and then make appropriate actions to take gains and avoid losses more effectively. Therefore, this paper focuses on discovering and forecasting stock index patterns.

In this study, we attempt to discover and predict stock index pattern through a three-stage architecture that consists of Toeplitz Inverse Covariance-Based Clustering (TICC), Temporal Pattern Attention and Long- Short-Term Memory (TPA-LSTM) and Multivariate LSTM-FCNs (MLSTM-FCN, MALSTM-FCN), which are developed by Hallac *et al.* [5], Shih *et al.* [6] and Karim *et al.* [7], respectively. Taking Hangseng Composite Stock Index (HSCI) and 11 industry stock indices in HSCI as an example, this paper investigates the feasibility of proposed three-stage architecture in financial time series. In the first stage, this paper applies TICC algorithm to cluster prices of industrial indices in HSCI, including consumer good manufacturing, consumer service, energy, finance, industry, information technology, integrated industry, raw material, real estate, utilities. Then, this paper maps the clustering results of industry indices to HSCI and discover repeated patterns of HSCI. In the second stage, TPA-LSTM is used to predict industry indices. In the third stage, this paper applies Multivariate LSTM-FCNs to classify industry indices and predict the pattern of HSCI in the future. Based on the idea that industry indicators are predominant factors in explaining stock market co-movements [8], the proposed three-stage architecture with TICC, TPA-LSTM and Multivariate LSTM-FCNs might be more effective in discovery and prediction of stock index patterns. Moreover, we could conduct early warnings of stock index and make corresponding measures more efficiently. The integral three-stage architecture of TICC, TPA-LSTM, Multivariate LSTM-FCNs that is proposed in this paper is shown in Fig. 1.

The remainder of this paper is organized as follows. Section 2 summarizes the literature review of stock index pattern prediction in machine learning. Section 3 describes the mathematical model for the discovery and prediction of the stock index. Section 3 provides details of proposed three-stage architecture, including TICC, TPA-LSTM, Multivariate LSTM-FCNs. Section 4 describes the empirical preliminaries, which contain empirical dataset and stationarity test. Section 5 presents selection of evaluation criteria and empirical results. Section 6 constructs a bullish trading rule and presents the portfolio performance based on this trading rule and the proposed three-stage architechture. Finally, the conclusion is drawn in section 7.

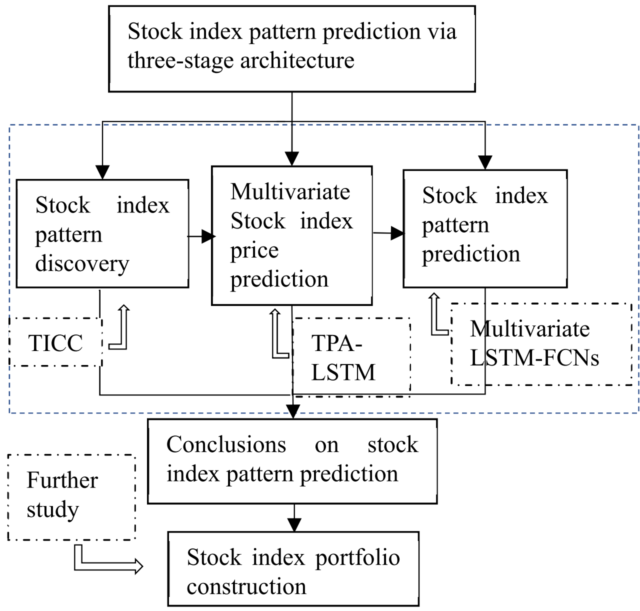


**FIGURE 1.** Three-stage architecture of TICC, TPA-LSTM, Multivariate LSTM-FCNs.

## II. LITERATURE REVIEW

The methods used to predict stock index have two categories, including traditional econometric methods and neural networks. However, there mainly exist three problems in the literature related to stock index prediction through traditional econometric methods (e.g. ARIMA, VAR). One is that prediction of multivariate time series through traditional econometric methods is rarely explored because of model capacity and their high computational cost, especially when the data is high-dimensional [2]. The second one is that these models are generally based on the assumption that variables are independent, normal distribution, which is impractical in the real market [3]. The third one is that these studies rarely explore pattern prediction rather than point prediction [9], [10]. Recently, deep neural networks provide another promising tool in time series forecasting [11], [12] due to its ability to model nonlinear patterns, realize complex causal relationships and learn from huge history dataset. In the field of pattern prediction in time series, there exist various approaches, such as recurrent neural network (RNN) [13] and convolutional neural network (CNN) [14]. The studies related to pattern prediction of financial time series through deep neural network mainly have three categories. One is to identify significant events or patterns through some templates, such as stock market bulletin [15], [16]. The second one is to seek and predict the inherent structure in the time series [17], [18]. The third one is to predict simple up-down trend in time series [19], [20], [21].

However, the performance of neural network techniques in stock index scenario is relatively less explored [22]–[24]. Moreover, there exist five problems in stock index pattern prediction through neural networks. One is that existing studies only study pattern discovery of stock index instead of forming a complete structure and investigating pattern prediction further [25]–[27]. The second one is that these studies mainly focus on up-down prediction, while fail to recognize and predict various patterns of stock index further [28], [29]. The third one is that these researches mainly focus on pattern prediction of a single stock index without considering differences of various industries [19], [24]. The fourth one is that these models are mainly designed for multivariate time series with strong repeated patterns and fixed time periods, which are unadaptable to datasets with non-periodic or flexible periodic patterns [30], [31]. The fifth one is that these studies fall short in distinguishing a mixture of short-term and long-term repeating patterns explicitly [22], [23].

The contributions of this paper are, first, to construct a comprehensive framework to discover and predict repeated patterns of stock index through a proposed three-stage architecture that consists of TICC, TPA-LSTM and Multivariate LSTM-FCNs, which could fill the gap between stock index pattern prediction and machine learning techniques, and secondly, to predict stock index patterns with full consideration of interdependencies among eleven industry stock index prices in HSCI. These industry stock index prices include consumer good manufacturing, consumer service, energy, finance, industry, information technology, integrated industry, raw material, real estate, utilities, and thirdly, to discover repeated patterns of stock index price with flexible time period through TICC technique, and fourthly, to learn and predict multivariate time series of industry indices with weak periodic long and short-term patterns more effectively through TPA-LSTM method. Therefore, we could discover and predict stock index repeated patterns more completely and effectively through the proposed three-stage architecture.

## III. STOCK INDEX PRICE PATTERN IDENTIFICATION AND PREDICTION BASED ON TICC, TPA-LSTM, MULTIVARIATE LSTM-FCNs

### A. MATHEMATICAL MODEL ON STOCK INDEX PRICE PATTERN IDENTIFICATION AND PREDICTION

Identification and forecasting of stock index patterns are important topics in the financial field, which could help investors make appropriate investments and avoid huge losses. In recent years, neural networks are widely used in financial time series. Unlike traditional models, deep neural networks have several distinct advantages as non-parametric, self-learning, non-assumption, and noise-tolerant, which are unavailable in traditional models [32], [33], [34]. However, there are four problems that need to be considered in the discovery and prediction process. The first one is the length of different patterns. Many studies prefer pattern templates of fixed length for representing repeated patterns rather than adjustable length, which is used by our method. The second one is the ignorance of weak periodic patterns in financial multivariate time series, which usually yields unsatisfactory outcomes. The third one is that many studies prefer univariate time series prediction rather than multivariate time series prediction, which is considered by the method applied in this paper. The last one is the efficiency of discovery and prediction process.

This paper aims at discovering and predicting repeated patterns of HSCI through a proposed three-stage framework that consists of TICC, TPA-LSTM and Multivariate LSTM-FCNs. Given datasets $\{x_t\}_{t=1}^{T}$, wherein $x_t \in R^n$ represents the observed value at time $t$ and $n$ is the variable dimension, the process of pattern discovery and prediction contains three stages.

In the first stage, this paper focuses on clustering industry indices based on the complex relationship among them, and mapping the clustering results to discover repeated patterns of HSCI. Instead of looking at $x_t$ only, a short subsequence of size $W$ ($W<T$, wherein $T$ is the length of the training sample) is clustered. This subsequence includes observations going from time $t\text{-}w+1$ to $t$, which is called $X_t$. We call this new sequence $X$, going from $X_1$ to $X_t$. Given new datasets $\{X_t\}_{t=1}^{n}$, TICC method is applied to get cluster results which is denoted by $\{Clu_i\}_{i=1}^{a}$.

In the second stage, this paper aims at predicting multivariate time series of industry stock index prices in a rolling forecasting step. Instead of looking at a single independent variable $y_t$, this stage predicts $x_{T+h}$, wherein h is the desirable horizon ahead of the current time stamp and $\{x_t\}_{t=1}^{T}$ are available. Moreover, this stage uses only $\{x_t\}_{T-w+1}^{T}$ to predict stock index prices $x_{T+h}$, wherein $w$ is the window size. This is based on the assumption that there is no useful information before the window $w$ which is set to be 30 in this paper [31]. Therefore, given datasets $\{x_t\}_{T-w+1}^{T} \in R^{n\times T}$, TPA-LSTM is applied to get the point prediction results of industry stock indices, which is denoted by $\{x_t\}_{T+h}^{T'} \in R^n$.

In the third stage, this paper is interested in classifying the predicted multivariate time series $\{x_t\}_{T+h}^{T'}$ that produced in the second stage, wherein $T'$ is the length of test sample. Depending on the needs of Multivariate LSTM-FCNs, the output of the second stage, $\{x_t\}_{T+h}^{T'} \in R^n$, is defined as a tensor of shape $(M, T', N)$, where $M$ is the number of samples in the dataset, $T'$ is the length of test time steps, $N$ is the number of variables in the dataset. Given a new shaped dataset, Multivariate LSTM-FCNs are applied to get classification results of industry stock indices and mapping results of HSCI, which is denoted by $\{Cla_i\}_{i=1}^{a}$.

### B. TOEPLITZ INVERSE COVARIANCE-BASED CLUSTERING METHOD ON STOCK INDEX PATTERN DISCOVERY

TICC was proposed by David Hallac *et al.* in 2017 [5], which was runner up in the research track of Knowledge Discovery and Data Mining (KDD). Compared to other clustering methods, TICC is the first method to cluster multivariate time series based on graphical dependency structure. Based on the graphical dependency structure, TICC could explore complex relationships among different variables in multivariate time series, and provide a basis for interpreting the clustering results of pattern discovery. For example, raw sensor data from an automobile can be interpreted as a sequential timeline of states, such as running, slowing down, stopping, etc. In each state, sensors present different synchronic and intertemporal relationship. TICC algorithm could analyze the relationship between different sensors, and interpret the states of the automobile more efficiently. In this section, we describe the details of the TICC algorithm applied in this paper.

TICC is a clustering technique that cluster short subsequences of size $w$ ($w < T$, wherein $T$ is the length of time series), going from time $t-w+1$ to $t$. In this paper, window size $w$ and cluster number $k$ of stock index are chosen to be 3 and 5, respectively. Each cluster is defined based on a Gaussian inverse covariance $\Theta_i \in R^{n\times n}(i = 1, 2, 3, 4, 5)$, which describes the structural representation of cluster i. The objective of the TICC method applied in this paper is to solve the covariance of each cluster $\Theta \in \{\Theta_1, \Theta_2, \Theta_3, \Theta_4, \Theta_5\}$ and the assignment results in $P \in \{P_1, P_2, P_3, P_4, P_5\}$, where $P \subset \{1, 2, \cdots, T\}$. The objective is achieved through two steps: cluster assignment and Toeplitz graphical lasso. In the Cluster assignment, TICC solves the subproblem through a dynamic programming algorithm. In the Toeplitz graphical lasso, TICC updates the cluster parameters based on the alternating direction method of multipliers (ADMM). This combinatorial algorithm is equivalent to expectation maximization (EM). Based on the assignment results, we could discover repeated clusters of industry stock indices and HSCI. Then, the optimization problem in TICC is written as follows,

$$argmin_{\Theta\in\Gamma,P} \sum_{i=1}^{5}\left[\overbrace{\|\lambda\circ\theta_i\|_1}^{sparsity} + \sum_{X_i\in P_i}\left(\overbrace{-\ell\ell(X_t,\theta_i)}^{log\ likelihood} + \overbrace{\beta I\{X_{t-1}\notin P_i\}}^{temporal\ consistency}\right)\right] \quad (1)$$

Here, $\Gamma$ is Toeplitz matrices, $\lambda(\lambda \in R^{n\times n})$ is a parameter that determines the sparsity level in the MRFs. In other words, we will minimize the negative log likelihood and make sure $\Theta_i(i = 1, 2, 3, 4, 5)$ is sparse based on regularization parameter $\lambda$. $\beta$ is another parameter that incentivize neighboring points to be assigned to one cluster, the neighboring subsequences are more likely to belong to one cluster as $\beta$ gets larger. In this paper, $\lambda$ and $\beta$ are chosen to be 11e-3 and 2, respectively. $X_t$ contains 11 industry indices in HSCI, including consumer good manufacturing, consumer service, energy, finance, industry, information technology, integrated industry, raw material, real estate, utilities. $\|\lambda\circ\Theta_i\|_l$ is an $\ell_1$-norm penalty to encourage a sparse $\Theta_i$ and prevent overfitting. $-\ell\ell(X_t, \Theta_i)$ is the negative log likelihood of $X_t$ comes from cluster $i$. $I\{X_{t-1} \notin P_i\}$ is a function that analyze whether adjacent subsequences are assigned to one cluster.

#### 1) CLUSTER ASSIGNMENT

Given the initial inverse covariances $\Theta_i(i = 1, 2, 3, 4, 5)$ and one of regularization parameters $\beta$, this section solves the following subproblem for $P \in \{P_1, P_2, P_3, P_4, P_5\}$ in this section.

$$minimize \sum_{i=1}^{6}\sum_{X_i\in P_i}\left(-\ell\ell\left(X_t, \Theta_i\right) + \beta I\left\{X_{t-1}\notin P_i\right\}\right) \quad (2)$$

The objective of this subproblem is to assign the $T$ subsequences to these 5 clusters based on the tradeoff of maximizing the log likelihood of the data and minimizing the volatility of cluster assignment. We can get the initial results of points assignment $P_i$ from $3^{\mathrm{T}}$ potential assignments of points to 5 clusters through dynamic programming.

#### 2) TOEPLITZ GRAPHICAL LASSO

Given the initial cluster assignments $P \in \{P_1, P_2, P_3, P_4, P_5\}$ and inverse convariances $\Theta \in \{\Theta_1, \Theta_2, \Theta_3, \Theta_4, \Theta_5\}$, this section solves the TICC's optimization problem and update the parameters $\Theta_i$ and $P_i(i = 1, 2, 3, 4, 5)$. Based on a convergence criterion ($\varepsilon = 2\text{e-}5$), We are able to get the global optimum of cluster assignments $P_i$ through ADMM.

### C. TEMPORAL PATTERN ATTENTION AND LONG SHORT-TERM MEMORY ON STOCK INDEX PRICES PREDICTION

TPA-LSTM was proposed by Shun-Yao Shih *et al.* in 2018 [6], which was submitted to the journal track of European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2019). Compared to other forecasting methods, TPA-LSTM is the first method to predict $n$ dimensional time series with a mixture of weak long- and short-term patterns [35]. Based on TPA-LSTM method, we could fully consider the mixed structure of weak long- and short-term repeated patterns inherent in financial time series, and predict multivariate stock indices more accurately. In this section, we describe the details of TPA-LSTM algorithm applied in this paper.

TPA-LSTM consists of a non-linear part and a linear part. The non-linear part is a temporal attention mechanism which includes a recurrent layer, convolutional layer and a temporal pattern attention layer, while the linear part uses an autoregressive model (AR) to forecast the result.

#### 1) RECURRENT LAYER

The first layer of TPA-LSTM is a long short-term memory network (LSTM). Given the input matrix $X = \{x_1, x_2, \cdots, x_t$, wherein $x_t \in R^n(n = 11)$, this recurrent layer aims at capturing long-term information. The outputs of recurrent layer are the hidden states at each time stamp. The hidden sates of recurrent layer's units at time t can be formulated as

$$h_t, c_t = F(h_{t-1}, c_{t-1}, x_t) \quad (3)$$

Which is defined by the following equations,

$$i_t = \sigma(x_t W_{xi} + h_{t-1} W_{hi} + b_i) \quad (4)$$
$$f_t = \sigma(x_t W_{xf} + h_{t-1} W_{hf} + b_f) \quad (5)$$
$$o_t = \sigma(x_t W_{xo} + h_{t-1} W_{ho} + b_o) \quad (6)$$
$$c_t = f_t \odot c_{t-1} + i_t \odot tanh(x_t W_{xg} + h_{t-1} W_{hg})) \quad (7)$$
$$h_t = o_t \odot tanh \odot (c_t) \quad (8)$$

where $i_t, f_t, o_t \in R^m$, $W_{xi}, W_{xf}, W_{xo}, W_{xg} \in R^{m\times n}(n = 8)$, $h_t \in R^{m\times(t-1)}$, $\odot$ is the element-wise product, and $\sigma$ is the sigmoid function.

#### 2) CONVOLUTIONAL LAYER

Given previous LSTM hidden states $H = \{h_1, h_2, \cdots, h_{t-1}\}$ and initial input matrix $X = \{x_1, x_2, \cdots, x_t\}$, this section extracts short-term signal patterns and interdependencies among eleven variables. The output in this section can be expressed as

$$H^C_{i,j} = \sum\nolimits_{l=1}^{w} H_{i,(t-w-1+l)} \times C_{j,(T-w+l)} \quad (9)$$

where $H^C_{i,j}$ represents the convolutional value of the $i$-th row vector and the $j$-th filter, $C_i$ denotes the $k$ filters we have, $T$ is the maximum length this paper is paying attention which is set to be 30.

#### 3) TEMPORAL PATTERN ATTENTION LAYER

Traditional attention mechanism selects relevant information relative to current time step, which may lead to a failure to ignore noisy variables and detect temporal useful patterns in multivariate time series forecasting. To alleviate this problem, TPA-LSTM develops a new temporal pattern attention mechanism which could select useful variables and capture temporal information for forecasting. Given previous convolutional value $H^C_t$, recurrent value $H$, and initial input matrix $X$, the output of this temporal pattern attention layer is a non-linear projection part, which is computed as

$$h^D_t = W_{h'} (W_h h_t + W_V v_t) + b \quad (10)$$

where $v_t = H^C_i \alpha_t$ is the weighted context of hidden states of the convolutional matrix, $\alpha_t$ is the attention weights which can be expressed as

$$\alpha_t = \sigma(AttnScore(H^C_t, h_t)) \quad (11)$$

#### 4) AUTOREGRESSIVE LAYER

Due to non-linear property of the proposed attention mechanism, TPA-LSTM method decomposes the prediction into a non-linear part and a linear part. The prediction of the non-linear part is captured by a recurrent layer, a convolutional layer, and a temporal pattern attention layer. In contrast, the prediction of the linear part is solved by the Autoregressive (AR) model in this section. Given the initial input $X$, we can get the forecasting result of the non-linear part through AR Layer, which is formulated as follows,

$$h^L_{t,i} = \sum\nolimits_{k=0}^{q^{ar-1}} W^{ar}_k x_{k-1,i} + b^{ar} \quad (12)$$

Then the forecasting result of TPA-LSTM can be expressed as follows,

$$\hat{Y}_t = h^D_t + h^L_t \quad (13)$$

### D. MULTIVARIATE LSTM-FCNs ON STOCK INDEX PATTERN CLASSIFICATION

Multivariate LSTM-FCNs were proposed by Fazle Karim *et al.* in 2019 [7]. Compared to other classification methods, Multivariate LSTM-FCNs improve the classification performance by combining a Fully Convolutional Network (FCN) block with Long Short-Term Memory (LSTM) block, which considers the complex structure of multivariate time series and could classify multivariate time series without heavy preprocessing on the data or feature engineering. Based on Multivariate LSTM-FCNs, we could classify and predict the future pattern of financial time series more accurately and quickly with the consideration of interdependences among different variables and the complex structure inherent in financial time series. In this section, we introduce the details of Multivariate LSTM-FCNs in this paper.

#### 1) FULLY CONVOLUTIONAL BLOCK

The fully convolutional block consists of three temporal convolutional blocks, which is used as a feature extractor. Each temporal convolutional block contains a convolutional layer with filter sizes of 128, 256, 128. Each temporal convolutional block is followed by a batch normalization layer [36] with an epsilon of 0.001 and a Rectified Linear Unit (RELU) activation function [37]. Moreover, the first two temporal convolutional blocks are succeeded by a squeeze-and-excite block with reduction ratio r of 16, which is used to recalibrate the input feature maps adaptively. Finally, a global average pooling layer is applied after the final convolution block.

Temporal Convolutional Network (TCN) is proposed by Colin Lea *et al.* in 2016 [38], which is used as a feature extraction module in this paper. The input to our temporal convolutional blocks applied in this paper, $X_t \in \hat{E}R^{N_0}$, is the output of the TPA-LSTM, which is defined as a multivariate time series with $T'$ time steps and N variables. The true pattern label for each time step $t$ is given by $y_t = \{1, 2, \cdots, C\}$, wherein $C$ is the number of classifying classes.

Specifically, this paper applies a set of one-dimensional filters on each of $L$ convolutional layers, which are used to capture how variables evolve in the whole time period $T'$. The number of convolutional layers $L$ is set to be 3 in this paper. Moreover, we parametrize the one-dimensional convolutional filters for each layer $l$ using tensor $W^{(l)} \in R^{N_l \times d \times N_{t-1}}$ and biases $b^{(l)} \in R^{N_l}$, where $l \in \{1, \ldots, L\}$ is layer index and d is filter duration. The i-th part of the unnormalized activation $\hat{E}^{(l)}_t \in R^{N_l}$ in the $l$-th layer then can be expressed as

$$\hat{E}^{(l)}_{i,t} = f(b^{(l)}_i \sum\nolimits_{t'=1}^{d} \langle W^{(l)}_{i,t'}, E^{(l-1)}_{t+d-t'} \rangle \quad (14)$$

where f(·) is a Leaky Rectified Linear Unit (RELU).

Squeeze-and-excitation block was proposed by Hu *et al.* [39], which acts as a unit of calculation for transforming: $F_{tr} : X \rightarrow U, X \in R^{W' \times H' \times C'}, U \in R^{W \times H \times C}$. The output of $F_{tr}$ is expressed as

$$U = [u_1, u_2, \ldots, u_C] \tag{15}$$

where $u_c = v_c * X = \sum_{s=1}^{C'} v_c^s * x^s$. Moreover, we use $*$, $v_c^8$ to represent convolutional operation and two-dimensional spatial kernel, respectively. In order to increase the network's sensitivity to information so that subsequent transformations can take advantage of them, the transformation process is divided into two steps: squeeze and excitation.

The squeeze operation exploits information outside the local receptive field through a global average pool, which could generate channel-wise statistics. For spatial data with dimensions $W \times H$, the $c$-th element of channel-wise statistics z (z $\in R^C$) is calculated by

$$z_c = F_{sq}(u_c) = \frac{1}{W \times H} \sum_{i=1}^{W} \sum_{j=1}^{H} u_c(i, j) \tag{16}$$

For time series data with temporal dimension $T$, the $c$-th element of channel-wise statistics $z(z \in R^C)$ is calculated by

$$z_c = F_{sq}(u_c) = \frac{1}{T} \sum_{t=1}^{T} u_c(t) \tag{17}$$

In order to analyze the channel-wise dependencies in the following step, an excitation operation processes the summary information produced by squeeze operation through a simple gating mechanism. The mechanism is expressed as

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)) \tag{18}$$

where $\sigma$, $\delta$ are sigmoid activation function and ReLU activation function, respectively. $W_1 \in R^{\frac{C}{r} \times C}$, $W_2 \in R^{\frac{C}{r} \times C}$ are used to limit model complexity, wherein $r$ is reduction ratio.

Finally, the output of the block is rescaled as

$$\tilde{x}_c = F_{scale}(u_c, s_c) = s_c \cdot u_c \tag{19}$$

where $\tilde{X} = [\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_C]$, $F_{scale}(u_c, s_c)$ refers to channel-wise multiplication between the feature map $u_c \in R^T$ and the scale $s_c$.

#### 2) LSTM BLOCK

Besides fully convolutional block, the transposed multivariate time series $X \in R^{N \times T'}$, which is achieved through a dimension shuffle layer, is conveyed into the LSTM block. The LSTM block contains either a LSTM layer or an Attention LSTM layer, which is followed by a dropout layer. Finally, a SoftMax classification layer uses output obtained from the fully convolutional block and LSTM block to produce the final classification result.

LSTM layer can capture long-term information and temporal dependencies in the multivariate time series, as described in previous section. However, it fails to learn long-term dependencies, which could be solved by an attention mechanism proposed by Bahdanau *et al.* [40].

The attention mechanism is widely used in financial time series analysis in recent years [41], [42]. According to the attention mechanism, context vector $c_i$ depends on a sequence of annotation $(h_1, \ldots, h_{T_x})$. Each annotation $h_i$ contains information about the entire input, while pays more attention on the surroundings of the $i$-th word. The context vector $c_i$ is calculated as

$$c_i = \sum_{j=1}^{T_x} \alpha_{ij} h_j \tag{20}$$

$$\alpha_{ij} = \frac{exp(e_{ij})}{\sum_{k=1}^{T_x} exp(e_{ik})} \tag{21}$$

where $\alpha_{ij}$ is the weight of each annotation $h_i$, $e_{ij} = a(s_{i-1}, h_j)$ is an alignment model. The alignment model is a feedforward neural network that measures the degree of matching between the input around $j$ and the output $i$.

## IV. DATA

### A. DATA SET AND DATA PREPROCESSING

The data of Hangseng Stock Composite Index (HSCI) and eleven industry stock indices are obtained from Wind platform. Eleven industry stock indices include consumer good manufacturing, consumer service, energy, finance, industry, information technology, integrated industry, raw material, real estate, utilities. The dataset covers time period from 01/09/2006 up to 01/02/2019.

More specifically, closing prices on daily bases are used as datasets, which are illustrated in Fig. 2. In order to have a better understanding of historical variation, (a), (b), (c) show prices of HSCI and eleven industry indices in daily, monthly and yearly scale, respectively. Moreover, figure (d) shows the daily closing prices of these stock indices during 2006. The short-term and long-term repeated patterns are not clear due to non-stational time series or patterns with a flexible time period. Each sample data of stock index prices is split into a training set (60%), validation set (20%), test set (20%) in chronological order. Based on the training set and validation set, we discover repeated patterns through TICC and improve the prediction power of TPA-LSTM through tuning hyperparameters. Moreover, we use a test set to get prediction and classification results of TPA-LSTM and Multivariate LSTM-FCNs. Also, the Null values are immediately dropped due to its little scale.

### B. STATIONARITY TEST

The stock index has characteristics of noisy and non-stationary, which may make prediction results unsatisfied. Therefore, this section implements a stationarity test of HSCI and eleven industry stock indices, which include the Levin-Lin-Chu test (LLC), the Im-Pesaran-Shin test (IPS) and the Phillips-Perron test (PP). The results of the statistical value are found in Table 1.

Table 1 shows the stationarity test results of HSCI and eleven industry indices. With the null hypothesis that assumes unit root process, these stock indices are nonstationary in LLC test while stationary in IPS test and PP test. Moreover,

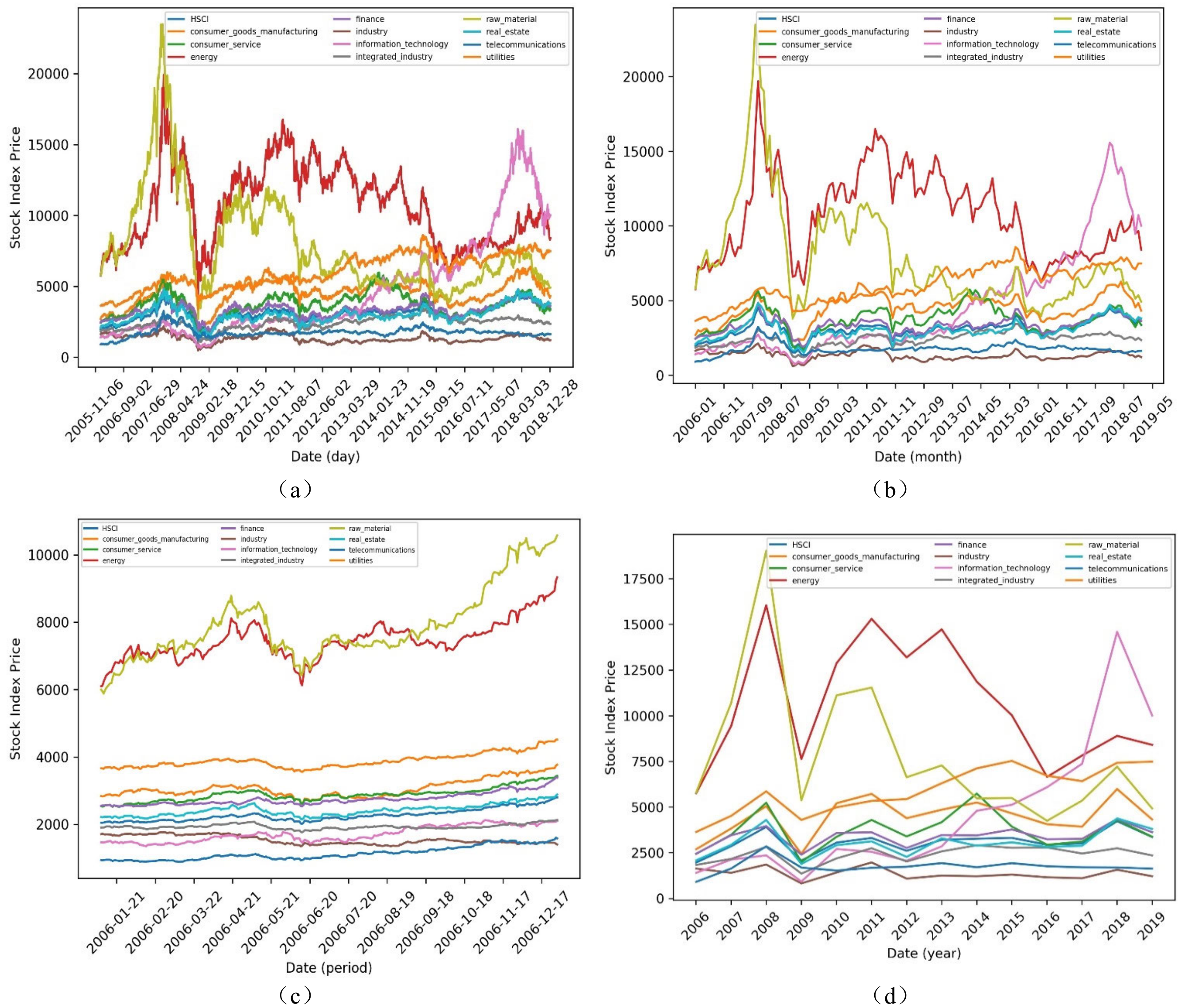


**FIGURE 2.** Closing prices of HSCI and eleven industry indices.

**TABLE 1.** Stationarity test for HSCI and eleven industry indices.

| Method | Statistic | Prob.** |
|---|---|---|
| No difference | | |
| LLC | -0.67131 | 0.0.2510 |
| IPS | -3.04907 | 0.0011 |
| PP | 4504599 | 0.0023 |
| First order difference | | |
| LLC | -237.992 | 0.0000 |
| IPS | -201.280 | 0.0000 |
| PP | 202.627 | 0.0000 |

the first-order difference of these stock indices is stationary in all stationarity tests. Therefore, in order to compare the predicted performance of TPA-LSTM and other techniques (such as CNN and RNN) that are applied in the second stage of the proposed architecture, we will apply these techniques to datasets with no difference and first-order difference.

## V. RESULTS AND DISCUSSION

This paper proposes a three-stage architecture that consists of TICC, TPA-LSTM, Multivariate LSTM-FCNs to discover and predict repeated patterns of stock index. In this section, this paper takes HSCI and eleven industry stock indices as an empirical example to analyze the feasibility of the proposed three-stage architecture in financial stock indices. In the first stage, this paper applies TICC to discover repeated patterns of HSCI. In the second stage, this paper predicts multivariate time series of industry stock indices with consideration of short-term and long-term repetitive and weak patterns. In the third stage, this paper predicts repeated patterns of HSCI through Multivariate LSTM-FCNs. The empirical results and discussion are described in the following section.

### A. PERFORMANCE CRITERIA

The performance of the proposed three-stage architecture that comprises of TICC, TPA-LSTM, Multivariate LSTM-FCNs is evaluated in the last two stages.

In the second stage, this paper predicts multivariate time series of eleven industry stock indices in HSCI through TPA-LSTM. The prediction performance of the TPA-LSTM method is compared with four methods, including Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), VAR, ARIMA. These first three methods are able to analyze multivariate variables, while ARIMA is a univariate method in which we will analyze each variable independently. In order to evaluate the performance of TPA-LSTM more comprehensively, we choose three performance evaluation metrics and two tests for analyzing performance difference. Three performance evaluation metrics include Root Relative Squared Error (RSE), Relative Absolute Error (RAE), and Empirical Correlation Coefficient (CORR). The RSE and RAE are in scaled versions, which are designed to make comparisons more efficient and valid. Two tests for analyzing performance difference include a multistep conditional predictive ability test [43] and a fluctuation test [44]. The multistep conditional predictive ability test is a performance test on average, while the fluctuation test is a performance test to evaluate the time-varying relative forecast performance of the models. In these two tests, we sum the RSE and RAE at each time step to analyze the relative performance of TPA-LSTM.

In the third stage, this paper classifies the repeated patterns of HSCI and industry stock indices that are discovered in the first stage. The classification performance of Multivariate LSTM-FCNs is compared with Naive Bayes Classifier (NB), Support Vector Machine Classifier (SVM), Random Forest (RF), XGBoost (XGB), Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN). Additionally, we evaluate classification performance of a model through four classification metrics, including precision, recall, F1-score and accuracy. The definitions of these criteria are found in Table 2.

Here, $Y, \hat{Y} \in R^{n\times T}$ are true values and predicted values of HSCI and industry indices, respectively. $n$ is the number of test prediction, $\tau$ is prediction horizon, $T$ is total sample size, $m$ is maximum estimation of window size, $h_t$ is a test function, $\Delta L_i$ is loss differences of two methods, TP is true positive, FP is false positive, FN is false negative. Moreover, $Z_{m,t+\tau} = h_t \Delta L_{m,t+\tau}$, $\tilde{\Omega}_n = n^{-1} Z_{m,t+\tau} Z'_{m,t+\tau} + n^{-1} \sum_{j=1}^{\tau-1} w_{n,j} \times \sum_{t=m+j}^{T-\tau} [Z_{m,t+\tau} Z'_{m,t+\tau-j} + Z_{m,t+\tau-j} Z'_{m,t+\tau}]$, where $w_{n,j}$ is a weight function [44].

### B. EXPERIMENT ON STOCK INDEX PATTERN DISCOVERY VIA TICC

In this section, we conduct clustering on multivariate time series of industry stock indices in HSCI and mapping the

**TABLE 2. Performance criteria and their calculations.**

| Criteria | Calculation |
|---|---|
| RSE | $\mathrm{RSE} = \dfrac{\sqrt{\sum_{(i,t)\in\Omega_{Test}}(Y_{it}-\hat{Y}_{it})^2}}{\sqrt{\sum_{(i,t)\in\Omega_{Test}}(Y_{it}-\mathrm{mean}(Y))^2}}$ |
| RAE | $\mathrm{RAE} = \dfrac{\sum_{(i,t)\in\Omega_{Test}}\lvert Y_{it}-\hat{Y}_{it}\rvert}{\sum_{(i,t)\in\Omega_{Test}}\lvert Y_{it}-\mathrm{mean}(Y)\rvert}$ |
| CORR | $\mathrm{CORR} = \dfrac{1}{n}\sum_{i=1}^{n}\dfrac{\sum_t(Y_{it}-\mathrm{mean}(Y_i))(\hat{Y}_{it}-\mathrm{mean}(\hat{Y}_i))}{\sqrt{\sum_t(Y_{it}-\mathrm{mean}(Y_i))^2(\hat{Y}_{it}-\mathrm{mean}(\hat{Y}_i))^2}}$ |
| Multistep Conditional Predictive Ability Test | $T^h_{m,n,\tau} = n\left(n^{-1}\sum_{t=m}^{T-\tau} h_t\Delta L_{m,t+\tau}\right)\tilde{\Omega}_n^{-1}\left(n^{-1}\sum_{t=m}^{T-\tau} h_t\Delta L_{m,t+\tau}\right) = n\bar{Z}'_{m,n}\tilde{\Omega}_n^{-1}\bar{Z}_{m,n}$ |
| Fluctuation test | $F^{OOS}_{t,m} = \hat{\sigma}^{-1} m^{-1/2} \sum_{j=t-m/2}^{t+\frac{m}{2}-1} \Delta L_j$ |
| Precision | $\mathrm{Precision} = \dfrac{TP}{TP+FP}$ |
| Recall | $\mathrm{Recall} = \dfrac{TP}{TP+FN}$ |
| F1-score | $F1 = \dfrac{2\times \mathrm{Precision}\times \mathrm{Recall}}{\mathrm{Precision}+\mathrm{Recall}}$ |
| Accuracy | $\mathrm{Accuracy} = \dfrac{\text{True Predictions}}{\text{All Predictions}}$ |

clustering results to HSCI by using the TICC method. Through this process, we could discover repeated patterns of HSCI. The dataset of 80% is used to train the TICC method. A cluster number of 5 is found to produce the best possible results. $W$, $\lambda$, $\beta$ and $\varepsilon$ are arbitrarily chosen to be 3, $11e^{-3}$, 600 and $2e^{-5}$, respectively. The program is constructed using python 3 language. Fig. 3 and Fig. 4 show the clustering and mapping results of HSCI, respectively.

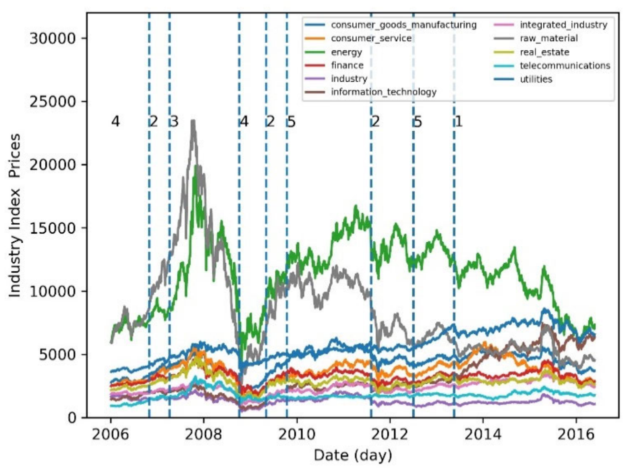


**FIGURE 3.** Clustering results of industry indices in HSCI.

Fig. 3 shows the clustering results of industry indices in HSCI, in which the blue dotted line divides the entire time period from 2007 to 2016 into nine sub-time periods, and the numbers above the line chart represent the cluster each time period belongs to. Specifically, 1 represents cluster 1, 2 represents cluster 2, and etc. As we can see from Fig. 3, in cluster 1, the energy industry index declines with considerate fluctuations, while the information industry index rises year by year. Moreover, other industry indices change synergistically and smoothly. Cluster 2 repeats three times throughout the entire time period, which is more complicated than other clusters. Specifically, the overall performance of cluster 2 is that all industry indices rise relative steadily, while the energy industry index and the raw material industry index rise by a large margin. The pattern characteristics of cluster 3 and cluster 4 are more obvious. In cluster 3, all industry indices show an ''inverted V''-shaped peak pattern. Specifically, all industry indices rise rapidly, and later, after the inflection point around 2008, all industries indices fall sharply. This pattern coincides with the financial crisis of 2008-2009. Cluster 3 is similar to cluster 1, but the difference is that cluster 3 has more volatility and greater risk. In cluster 4, all industry indices exhibit a ''W''-shaped valley pattern. Specifically, at first, all industry indices fall in concert. And then, these indices continue to fluctuate in a small range at the bottom of the valley. Finally, after a period of troughs in the valley, all industry indices climb up. Cluster 5 repeats twice during the entire time period and generally exhibits a ''M'' or ''inverted V'' dynamic pattern with one to two peaks. Specifically, the energy industry index and the raw material industry index fluctuate significantly, and other industry indices fluctuate in synergy.

According to the dynamic trend of industry indices in HSCI, which is presented in Fig. 3, this paper combines macroeconomic theory to make relevant assumptions about the repeated patterns of the HSCI and industry indices. This paper assumes that cluster 1 represents a ''head and shoulder'' pattern, in which prosperity of stock indices is followed by a smoother decline. Cluster 2 represents a steady recovery mode. cluster 3 represents a repetitive pattern of ''inverted V'', in which prosperity of stock index is followed by a volatile recession. Cluster 4 represents a ''W''-shaped pattern with recession at the bottom. Cluster 5 represents an ''M-type'' pattern with large fluctuations.

In order to confirm the relevant assumptions of repeated pattern and discover repeated patterns of HSCI, this paper then maps the clustering results of industry stock indices to HSCI in Fig. 4. Cluster 1 with brown marking is similar to the technical analysis of ''head and shoulders''. In the ''head and shoulders'', stock price presents a ''mountain'' shape, which is a typical signal of dramatic falls. A peak of stock price is followed by a power of resistance, which leads to a failure of continuing to a new high. When the neckline of the ''head and shoulders is broken, a real sell signal occurs. Although compared with the highest point, the stock price has fallen back to a considerable extent, the decline trend is only at the beginning and the unsold investors should continue to sell. Cluster 2 with orange marking is more complexed

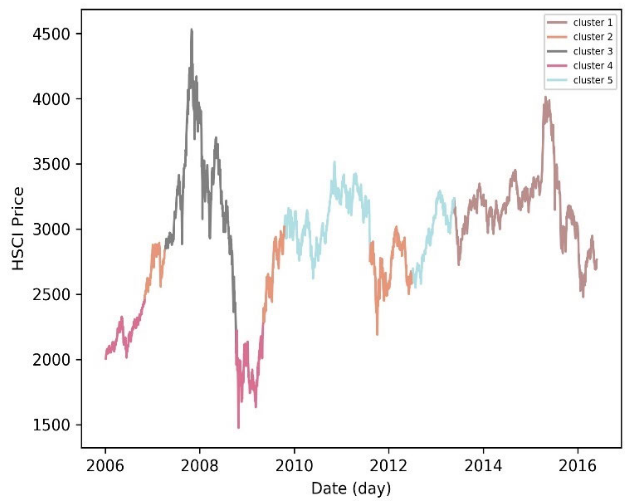


**FIGURE 4.** Mapping results of HSCI.

in shape, which denotes ''recovery and stable'' in mapping results. In Cluster 2, we recommend that investors explore in depth to ensure that stock price rise steadily and purchase the corresponding stock index in a timely manner. Cluster 3 with grey marking is similar to a type of spine, which is consistent with the economic crisis in 2008. In the shape of ''spine'', the price will go straight up to the highest point and then plummet. Speculators need to be aware of the stock market and sell stocks in time to avoid suffering losses. Cluster 4 with red marking is similar to the technical analysis of a ''W'' shape. In the ''W'' shape, the stock market rebounds before the end of the downward trend and then fell again slightly. The stock price will stop at the previous low point and begin to rise. When the stock price fell back to the last low, investors could immediately take long position. In cluster 4, we suggest that investors immediately buy the relevant stock index and hold a long position when the stock index price falls to the last low. Moreover, when more and more investors take a long position, the demand for this stock will drive the stock price to rise, and the stock price will also break through the last high point. If investors can't accurately determine the last low of stock index, they can hold the relevant stock index at the beginning of the ''W'' type of the stock index price and wait patiently for the stock index price to rise. Cluster 5 with brown marking which is similar to the technical analysis of ''rising flag''. In the ''rising flag'', investors are often rational and expect stock prices to increase. Stock price experiences a short-time soar, a slight downward trend with frequent falling and rebound, and a final increase. However, as we can see from Fig. 4 above, cluster 5 has a larger volatility and greater risk is greater. Investors should be more cautions when stock index is in this pattern.

The clustering and mapping results of industry stock indices and HSCI are roughly consistent with clustering assumptions, economic theories, and technical analysis. By applying TICC algorithm to pattern discovery of HSCI,

**TABLE 3.** Performance comparison of different methods.

| Methods | Metrics | Horizon 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 |
|---|---|---|---|---|---|---|---|---|---|
| TPA-LSTM | RSE | **0.0659*** | **0.0940*** | **0.1122*** | **0.1203*** | **0.1686*** | **0.1552*** | **0.1656*** | **0.1859*** |
| | RAE | **0.0562*** | **0.0832*** | **0.0972*** | **0.1029*** | **0.1433*** | **0.1368*** | **0.1422*** | **0.1678*** |
| | CORR | **0.9750*** | **0.9558*** | **0.9400*** | **0.9240*** | **0.9088*** | **0.8937*** | **0.8757*** | **0.8595*** |
| RNN | RSE | 0.1556 | 0.1930 | 0.2142 | 0.2346 | 0.2970 | 0.2784 | 0.3438 | 0.3636 |
| | RAE | 0.1397 | 0.1722 | 0.1879 | 0.2081 | 0.2729 | 0.2492 | 0.3174 | 0.3424 |
| | CORR | 0.9637 | 0.9396 | 0.9217 | 0.9060 | 0.8679 | 0.8575 | 0.8256 | 0.8081 |
| CNN | RSE | 0.0908 | 0.1203 | 0.1375 | 0.1740 | 0.1745 | 0.1935 | 0.1965 | 0.2474 |
| | RAE | 0.0791 | 0.1047 | 0.1196 | 0.1512 | 0.1541 | 0.1726 | 0.1760 | 0.2234 |
| | CORR | 0.9701 | 0.9473 | 0.9332 | 0.9133 | 0.8940 | 0.8809 | 0.8570 | 0.8025 |
| ARIMA-stationary | RSE | 1.0036 | 1.0022 | 1.0028 | 1.0033 | 1.0027 | 1.0028 | 1.0033 | 1.0018 |
| | RAE | 1.0020 | 1.0009 | 1.0011 | 1.0019 | 1.0018 | 1.0015 | 1.0025 | 1.0013 |
| | CORR | -0.0580 | -0.0472 | -0.0488 | -0.0519 | -0.0468 | -0.0437 | -0.0556 | -0.0393 |
| VAR-stationary | RSE | 1.0003 | 1.0001 | 1.0001 | 1.0001 | 1.0001 | 1.0001 | 1.0001 | 1.0001 |
| | RAE | 1.0002 | 0.9999 | 0.9999 | 1.0001 | 1.0001 | 1.0000 | 1.0001 | 0.9999 |
| | CORR | -0.0558 | -0.0559 | -0.0555 | -0.0553 | -0.0582 | -0.0566 | -0.0614 | -0.0464 |
| TPA-LSTM - stationary | RSE | 0.1984 | 0.1976 | 0.1996 | 0.1981 | 0.1987 | 0.1984 | 0.2032 | 0.2096 |
| | RAE | 0.1463 | 0.1464 | 0.1484 | 0.1473 | 0.1484 | 0.1478 | 0.1532 | 0.1620 |
| | CORR | -0.0166 | -0.0158 | -0.0484 | -0.0201 | -0.0294 | -0.0374 | -0.0385 | -0.0497 |
| RNN-stationary | RSE | 1.0178 | 1.0076 | 1.0087 | 1.0094 | 1.0060 | 1.0148 | 1.0130 | 1.0087 |
| | RAE | 1.0202 | 1.0081 | 1.0094 | 1.0110 | 1.0084 | 1.0213 | 1.0165 | 1.0111 |
| | CORR | -0.0397 | 0.0174 | 0.0020 | -0.0180 | 0.0177 | -0.0329 | -0.0531 | -0.0170 |
| CNN-stationary | RSE | 1.0451 | 1.0366 | 1.0611 | 1.0381 | 1.0523 | 1.0151 | 1.0590 | 1.0455 |
| | RAE | 1.0541 | 1.0404 | 1.0714 | 1.0490 | 1.0649 | 1.0252 | 1.0716 | 1.0508 |
| | CORR | -0.0127 | -0.0031 | 0.0127 | -0.0095 | -0.0029 | 0.0510 | -0.0448 | -0.0352 |

we can find different types of repeated patterns that other clustering techniques can't identify and provide interpretable results for the third section to predict repeated patterns of HSCI.

### C. EXPERIMENT ON STOCK INDEX PREDICTION VIA TPA-LSTM

In the second section, we apply the TPA-LSTM method to multivariate time series prediction of stock index prices in different industries, which are included in HSCI. TPA-LSTM uses four components to extract short-term and long-term repetitive and weak patterns with consideration of linear and non-linear structure of time series. These four components are recurrent component, convolutional component, temporal pattern attention component, and autoregressive component. We repeat the algorithm until we find the lowest validation loss value. The results of the experiment are described Table 3.

In the prices prediction of industry stock indices based on TPA-LSTM, an attention length of 30 is found to produce the best possible results. Moreover, the dropout rate, the learning rate, the horizon $h$ and the optimization algorithm are arbitrarily chosen to be 0.2, $3\times10^{-3}$, 24, and the Adam algorithm, respectively. The program is constructed using python 3 language. Table 3 shows the performance comparison of industry index price prediction on average, which include the evaluation metrics of RSE, RAE, CORR, and multistep conditional predictive ability test. Moreover, the main diagnostic tests of ARIMA and VAR are reported in the Appendix.

Table 3 summarizes the prediction performances of all methods on all test sets (20%) in three performance evaluation metrics, including RSE, RAE, CORR of TPA-LSTM, RNN, CNN, ARIMA-stationary, VAR-stationary, TPA-LSTM-stationary, RNN-stationary, CNN-stationary. The data for the last five models is industry index prices after first-order difference. Moreover, we set horizon = {3, 6, 9, 12, 15, 18, 21, 24}, respectively. The larger the horizons, the worse the prediction results. The best results for five methods and three metrics are highlighted in bold face in Table 3. The total count of the bold-faced results is 24 for TPA-LSTM and 0 for other methods, which indicates the relative predictive power of TPA-LSTM. Moreover, an asterisk sign (∗) indicates that the test rejects equal conditional predictive ability at the 1% level and that the TPA-LSTM method outperforms other methods through conditional predictive ability tests.

Clearly, even though the periodic patterns of industry stock index prices are not clear, the methods proposed by TPA-LSTM still perform much better than other neural network methods (RNN, CNN) and traditional econometric methods (ARIMA, VAR) on average. Specifically, when horizon is 12, TPA-LSTM outperforms the neural baseline methods RNN, CNN, ARIMA, VAR by 11.43%, 5.37%, 91.78%, 87.98% in RSE metric, 10.52%, 4.83%, 94.61%, 89.72% in RAE metric, and 1.81%, 1.07%, 93.35%, 97.93% in CORR metric, respectively. When the horizon is 24, TPA-LSTM outperforms the neural baseline methods RNN, CNN, ARIMA, VAR by 17.77%, 6.15%, 85.96%, 81.41% in RSE metric, 17.47%, 5.56%, 88.31%, 83.21% in RAE metric, and 5.14%, 5.70%, 89.47%, 90.59% in CORR metric, respectively. The TPA-LSTM method has robust performance in different metrics, partly due to its consideration of interdependencies among multiple variables, long-term information, linear structure, and weak periodic patterns.

However, the results on average seem too good to be true. In order to confirm the prediction ability of TPA-LSTM, we further introduce a fluctuation test [43] to evaluate the relative performance of TPA-LSTM in a dynamic environment. The relative performances based on RSE and RAE at each point with a horizon of 12 are reported in Fig. 5 and Fig. 6, respectively.

Fig. 5 and Fig. 6 report the dynamic prediction performance of TPA-LSTM based on GW Fluctuation test, in which one-sided critical value is 5%. The results show that in the horizon of 12, the results of fluctuation test based on RSE are basically the same as those based on RAE. Specifically, in addition to CNN and the TPA-LSTM_diff model, TPA-LSTM has a relatively higher predictive ability overall in the dynamic environment. Moreover, although CNN and the TPA-LSTM_diff model perform better in some periods, TPA-LSTM is significantly more predictive during most of time.

Through this section, we obtain predicted results of industry index prices through TPA-LSTM, which has a better predictive power during most of time. Based on the mapping cluster results of HSCI through TICC and prediction results of industry index prices through TPA-LSTM, we could classify and predict repeated patterns of HSCI through MALSTM in the next section.

### D. EXPERIMENT ON STOCK INDEX PATTERN PREDICTION VIA MULTIVARIATE LSTM-FCNs

In the third section, we conduct two steps to predict repeated patterns of HSCI. Firstly, we use the clustering results and predicted results of industry stock indices in HSCI to classify the repeated patterns of industry stock indices through Multivariate LSTM-FCNs. Secondly, we map the classification results of industry stock indices to HSCI and get predicted patterns of HSCI. The dataset of the first 80% is used to train the Multivariate LSTM-FCNs method and compare the classification performance of different methods, while dataset of the last 20% is used to produce predicted results. During the training phase, number of training epochs, dropout rate, batch size and optimizer are arbitrarily chosen to be 200, 80%, 128, Adam optimizer, respectively. The program is constructed using python 3 language. Table 4 shows the performance of HSCI classification through Multivariate LSTM-FCNs and other six classification methods in four metrics.

Table 4 summarizes the classification performance of Multivariate LSTM-FCNs and baseline methods in four metrics, including precision, recall, F1-score and accuracy. The methods we use to compare the classification performance include MLSTM-FCN, MALSTM-FCN, Naive Bayes Classifier (NB), Support Vector Machine Classifier (SVM), Random Forest (RF), XGBoost (XGB), Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN). The best results for the six methods are highlighted in bold face in Table 4. The total count of the bold-faced results is 4 for MALSTM-FCN, which indicates that proposed methods outperform other baseline methods (NB, SVM, RF, XGB, CNN, RNN) in classification power on average. More specifically, MALSTM-FCN outperforms the baseline methods NB, SVM, RF, XGB, CNN, RNN by 38.74%, 29.74%, 1.95%, 3.57%, 14.36%, 5.69% in precision metric, 26.39%, 22.77%, 5.07%, 5.47%, 9.53%, 3.88% in recall metric, 33.53%, 28.49%, 7.75%, 9.50%, 11.28%, 4.01% in F1-score metric, 18.55%, 20.55%, 0.75%, 1.00%, 11.28%, 4.01% in accuracy metric, respectively.

**TABLE 4.** **Performance comparison of HSCI classification via different methods.**

| Methods | Precision | Recall | F1-score | Accuracy |
|---|---|---|---|---|
| MLSTM-FCN | 0.7421 | 0.6927 | 0.7379 | 0.7379 |
| MALSTM-FCN | **0.8048** | **0.7352** | **0.7748** | **0.7748** |
| NB | 0.4930 | 0.5412 | 0.5150 | 0.6311 |
| SVM | 0.5655 | 0.5678 | 0.5541 | 0.6155 |
| RF | 0.7891 | 0.6979 | 0.7147 | 0.7689 |
| XGB | 0.7761 | 0.6950 | 0.7012 | 0.7670 |
| CNN | 0.6893 | 0.6651 | 0.6874 | 0.6874 |
| RNN | 0.7590 | 0.7066 | 0.7437 | 0.7437 |

Therefore, we obtain much better classification results of HSCI through MLSTM-FCN and MALSTM-FCN. Based on the classification results, we could accurately determine which market-level HSCI would be in and make corresponding countermeasures in time. Moreover, we could reduce risks in the financial market and construct a more effective portfolio through predicting repeated patterns of industry stock indices and HSCI.


**FIGURE 5.** **Fluctuation test with RSE.**


**FIGURE 6.** **Fluctuation test with RAE.**

## VI. ROBUSTNESS TEST BASED ON STOCK INDEX TRADING RULES

In the previous section, this paper takes HSCI as an example, and analyzes the performance of the proposed three-stage architecture through various metrics. In order to evaluate the prediction performance of the proposed three-stage architecture in financial time series more comprehensively, this section will further take seven comprehensive stock indices as examples, construct simple stock index trading rules, and evaluate the prediction performance of the proposed three-stage architecture in practice.

### A. BULLISH TRADING RULES

This section takes seven comprehensive stock indices as examples to analyze the practical application of the proposed three-stage architecture in financial time series. Here, the comprehensive stock index refers to a comprehensive index covering multiple industry stock indices. The seven comprehensive stock indices selected in this section include HSCI, SZZZ, SH380, GZ1000, ZZ80, S&P500 (SPX), and Dow Jones Industrial Average Index (DJI). Based on these comprehensive stock indices and the literature of Wang and Chan [46], we define the bullish signal trading rules as shown below.

$$B_{i,k} = \begin{cases} 1 & Clu_{i,a} = bullishsignal \\ -1 & Clu_{i,a} = othersignal \end{cases} \tag{22}$$

Here, $i$ is the stock index number, k is the trading day, and $B_{i,k}$ is the investor's buying behavior on stock index $i$. When $B_{i,k}$ is equal to 1, investors will buy the corresponding stock index investor's buying behavior on stock index $i$. When $B_{i,k}$ is equal to -1, the investor will sell the corresponding stock index investor's buying behavior on stock index $i$.

Under the proposed trading rules, if the repeating pattern of HSCI is 2, 4, 5, which suggests a bullish signal in the previous section, investors should buy HSCI on that trading day, and hold it for a certain number of days ($q$). If the repeating pattern of the HSCI presents another transaction signal, investors should take short position on HSCI, and close the position after a certain number of days ($q$). Similarly, when the repeating pattern of other comprehensive stock index $i$ recommends a buying transaction, investors should buy the corresponding index $i$ on that trading day, and hold it for a certain number of days ($q$). If the repeating pattern of the comprehensive stock index $i$ recommends a sell transaction, investors should take short position on the corresponding stock index, and close the position after a certain number of days ($q$).

Under the bullish trading rules defined above, this paper compares the equal proportion investment portfolio constructed based on the proposed three-stage architecture and the market through three performance metrics, including average return, excess return, and market timing. The definitions and calculations of these performance metrics are found in Table 5.

**TABLE 5. Performance metrics and their calculations.**

| Metrics | Calculation |
|---|---|
| Average $\text{Return}_{\text{p}}$ | $\text{Average Return}_{\text{p}} = \frac{\sum_{i=1}^{N}\sum_{\text{k=m}}^{\text{n}}\left[\left(\text{X}_{\text{i,k+q}} - \text{X}_{\text{i,k}}\right) \times \text{B}_{\text{i,k}}/\text{X}_{\text{i,k}}\right]}{\sum_{\text{k=m}}^{\text{n}} \text{B}_{\text{i,k}} \times \text{N}}$ |
| Average $\text{Return}_{M}$ | $\text{Average Return}_{M} = \frac{\sum_{i=1}^{N}\sum_{\text{k=m}}^{\text{n}}\left[\left(\text{X}_{\text{i,k+q}} - \text{X}_{\text{i,k}}\right)/\text{X}_{\text{i,k}}\right]}{\sum_{\text{k=m}}^{\text{n}} \text{B}_{\text{i,k}} \times \text{N} \times (\text{n} - \text{m} + 1)}$ |
| Excess Return | $\text{Excess Return} = \text{Average Return}_{\text{p}} - \text{Average Return}_{M}$ |
| Market Timing | $\text{Market Timing} = \frac{\text{FTTM N}(\text{r} > 0)}{\text{FTTM N(buys)}}$ |

Here, $X_{i,k}$ is the price of stock index $i$ on trading day $k$, m is the first trading day, n is the last trading day, $N$ ($N = 7$) is the number of stock indices, $q$ is the holding period which is set to be 100, 150, 200, 250, 300.

### B. ROBUSTNESS TEST

In this section, we conduct three steps to evaluate the prediction performance of the proposed three-stage architecture further. In the first step, we select seven comprehensive stock indices to construct our stock pool, and establish a bullish trading rule to provide suggestion for our invest behavior. In the second step, we select eligible stock indices based on the prediction results and the bullish trading rule. In the third step, we construct two equal proportion portfolios based on the proposed three-stage architecture and the market, and compare their performance based on three performance metrics, including average return, excess return, and market timing. Table 6 shows the performance of two portfolios in three metrics.

Table 6 summarizes the performance results of two equal proportion portfolios based on the proposed three-stage architecture and the market. As shown in Table 6, the forecast time span based on the bullish trading rule is referred to horizon in this section, which is set to be {3, 6, 9, 12, 15, 18, 21, 24}. In each time span, the holding period of the portfolio is set to be q = {100, 150, 200, 250, 300}.

Empirical results show that the portfolio returns based on the proposed three-stage architecture are generally higher than the portfolio returns based on the market. In addition, when the prediction horizon is fixed, the portfolio returns based on the proposed three-stage architecture increase with the length of holding time. Specifically, when the prediction horizon is 24, as the length of holding time increases, the average return of the portfolio based on the proposed three-stage architecture is 6.1090%, 13.5129%, 22.8862%, 31.5248%, and 38.5600%, respectively. The average return of the portfolio based on the market is 0.4320%, 3.6887%, 7.1686%, 10.4040%, and 12.8272%, respectively. The excess return of the portfolio based on the proposed three-stage architecture is 14.4194%, 16.6355%, 19.9613%, 21.4220%, and 21.7870%, respectively. These excess returns have passed the t-test at a significance level of 1%. The results reject the null hypothesis, which indicate that the excess returns are significant at a significance level of 1%.

**TABLE 6.** Performance metrics and their calculations.

| Horizon | Holding period | Average $\text{Return}_p$ (%) | Average $\text{Return}_M$ (%) | Excess Return | Market Timing (%) | Horizon | Holding period | Average $\text{Return}_p$ (%) | Average $\text{Return}_M$ (%) | Excess Return | Market Timing (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 100 | 7.0939 | 0.4320 | 16.9210 (0.0000)*** | 42.5360 | 15 | 100 | 6.1090 | 0.4320 | 14.4194 (0.0000)*** | 42.5360 |
| | 150 | 15.0419 | 3.6887 | 19.2248 (0.0000)*** | 47.4689 | | 150 | 13.5129 | 3.6887 | 16.6355 (0.0000)*** | 47.4689 |
| | 200 | 25.0617 | 7.1686 | 22.7242 (0.0000)*** | 48.5527 | | 200 | 22.8862 | 7.1686 | 19.9613 (0.0000)*** | 48.5527 |
| | 250 | 34.0727 | 10.4402 | 24.0107 (0.0000)*** | 50.5731 | | 250 | 31.5248 | 10.4402 | 21.4220 (0.0000)*** | 50.5731 |
| | 300 | 41.2145 | 12.8272 | 24.0345 (0.0000)*** | 49.5741 | | 300 | 38.5600 | 12.8272 | 21.7870 (0.0000)*** | 49.5741 |
| 6 | 100 | 6.1090 | 0.4320 | 14.4194 (0.0000)*** | 42.5360 | 18 | 100 | 6.1090 | 0.4320 | 14.4194 (0.0000)*** | 42.5360 |
| | 150 | 13.5129 | 3.6887 | 16.6355 (0.0000)*** | 47.4689 | | 150 | 13.5129 | 3.6887 | 16.6355 (0.0000)*** | 47.4689 |
| | 200 | 22.8862 | 7.1686 | 19.9613 (0.0000)*** | 48.5527 | | 200 | 22.8862 | 7.1686 | 19.9613 (0.0000)*** | 48.5527 |
| | 250 | 31.5248 | 10.4402 | 21.4220 (0.0000)*** | 50.5731 | | 250 | 31.5248 | 10.4402 | 21.4220 (0.0000)*** | 50.5731 |
| | 300 | 38.5600 | 12.8272 | 21.7870 (0.0000)*** | 49.5741 | | 300 | 38.5600 | 12.8272 | 21.7870 (0.0000)*** | 49.5741 |
| 9 | 100 | 6.1090 | 0.4320 | 14.4194 (0.0000)*** | 42.5360 | 21 | 100 | 6.1090 | 0.4320 | 14.4194 (0.0000)*** | 42.5360 |
| | 150 | 13.5129 | 3.6887 | 16.6355 (0.0000)*** | 47.4689 | | 150 | 13.5129 | 3.6887 | 16.6355 (0.0000)*** | 47.4689 |
| | 200 | 22.8862 | 7.1686 | 19.9613 (0.0000)*** | 48.5527 | | 200 | 22.8862 | 7.1686 | 19.9613 (0.0000)*** | 48.5527 |
| | 250 | 31.5248 | 10.4402 | 21.4220 (0.0000)*** | 50.5731 | | 250 | 31.5248 | 10.4402 | 21.4220 (0.0000)*** | 50.5731 |
| | 300 | 38.5600 | 12.8272 | 21.7870 (0.0000)*** | 49.5741 | | 300 | 38.5600 | 12.8272 | 21.7870 (0.0000)*** | 49.5741 |
| 12 | 100 | 7.0211 | 0.4320 | 16.7362 (0.0000)*** | 42.5360 | 24 | 100 | 6.1090 | 0.4320 | 14.4194 (0.0000)*** | 42.5360 |
| | 150 | 14.9455 | 3.6887 | 19.0615 (0.0000)*** | 47.4689 | | 150 | 13.5129 | 3.6887 | 16.6355 (0.0000)*** | 47.4689 |
| | 200 | 24.9413 | 7.1686 | 22.5712 (0.0000)*** | 48.5527 | | 200 | 22.8862 | 7.1686 | 19.9613 (0.0000)*** | 48.5527 |
| | 250 | 33.9279 | 10.4402 | 23.8636 (0.0000)*** | 50.5731 | | 250 | 31.5248 | 10.4402 | 21.4220 (0.0000)*** | 50.5731 |
| | 300 | 41.0386 | 12.8272 | 23.8856 (0.0000)*** | 49.5741 | | 300 | 38.5600 | 12.8272 | 21.7870 (0.0000)*** | 49.5741 |

The empirical results also show that the market timing ability of the proposed portfolio is generally stronger than that of the market-based portfolio. In addition, when the forecast time span is fixed, as the length of the holding period increases, the market timing ability of the proposed portfolio and the market-based portfolio both increase as well. Specifically, when the prediction horizon is 24, as the holding period increases, the positive return ratio of the proposed portfolio is 55.9114%, 64.0945%, 69.3575%, 76.8017%, and 79.2338%, respectively. The positive return ratio of the market-based portfolio is 42.5360%, 47.4689%, 48.5527%, 50.5731%, and 49.5741%, respectively.

Through the performance metrics and the bullish trading rules that are mentioned above, we can conclude that compared to the market-based portfolio, the portfolio based on the proposed three-stage architecture performs better in the practice. Moreover, the proposed three-stage architecture shows certain feasibility and application potential in the finance market.

## VII. CONCLUSIONS

Discovery and prediction of stock index pattern are of great importance to reduce uncertainty and risks in financial markets and, more specifically, is crucial in constructing a financial portfolio. In the literature of stock index pattern discovery and prediction through neural networks, previous studies mainly focus on pattern discovery and up-down prediction of stock index with strong repeated patterns and fixed time periods. This paper makes up for the shortcomings of previous research, which forms a complete structure of stock index pattern discovery and prediction through a proposed three-stage architecture of TICC, TPA-LSTM, and Multivariate LSTM-FCNs. Through proposed three-stage architecture, this paper could analyze and predict stock index prices with weak periodic and flexible patterns.

The proposed three-stage architecture contains three stages. In the first stage, we apply TICC to cluster industry stock indices in the comprehensive stock index and map cluster results to that stock index. Based on the mapping results, we could discover repeated patterns of the comprehensive stock index on the training dataset. In the second stage, we predict multivariate time series of industry stock indices simultaneously through TPA-LSTM. In the third section, we predict repeated patterns of the comprehensive stock index on the test dataset through Multivariate LSTM-FCNs. HSCI and eleven industry indices that are included in the HSCI are used in the experiment. The empirical results show that the proposed three-stage architecture, including TICC, TPA-LSTM, and Multivariate LSTM-FCNs significantly improves the state-of-the-art results in pattern discovery and prediction of HSCI. Moreover, we propose a bullish trading rule and construct an equal proportion portfolio based on this trading rule and the prediction results of the proposed three-stage architecture. Seven comprehensive stock indices are used in the experiment. The empirical results show that, the constructed portfolio based on the bullish trading rule and the proposed three-stage architecture performs significantly better than the market-based portfolio. Therefore, the proposed three-stage architecture is a feasible and promising method to discover and predict repeated patterns of stock index in financial markets.

There are two promising extensions of pattern discovery and prediction in stock index prices. One possible extension of stock index pattern prediction is to conduct proactive index-tracking or construct other trading strategies with predicted patterns of stock index. The other extension is to search for more suitable and effective price prediction and pattern matching methods to improve the performance of the proposed structure.

## APPENDIX

See Table 7 and 8.

**TABLE 7.** Diagnostic test of ARIMA*.

| | Horizon | | | |
|---|---|---|---|---|
| Diagnostic Test | 3 | 6 | 9 | 12 |
| Durbin Watson | 1.934016 | 1.93815 | 1.940566 | 1.94198 |
| Normality_Jarque Bera | 374.6563 | 370.9624 | 372.0882 | 371.2866 |
| | (4.41E-82) | (2.80E-81) | (1.59E-81) | (2.38E-81) |
| Normality_Omni | 107.8993 | 107.4276 | 107.5246 | 107.3897 |
| | (3.72E-24) | (4.70E-24) | (4.48E-24) | (4.79E-24) |
| | Horizon | | | |
| Diagnostic Test | 15 | 18 | 21 | 24 |
| Durbin Watson | 1.947629 | 1.943887 | 1.944517 | 1.95355 |
| Normality_Jarque Bera | 370.5208 | 364.4936 | 365.0941 | 367.0618 |
| | (3.49E-81) | (7.10E-80) | (5.26E-80 | (1.97E-80) |
| Normality_Omni | 107.4352 | 106.4886 | 106.3259 | 105.8894 |
| | (4.69E-24) | (7.52E-24) | (8.16E-24 | (1.01E-23) |

* The numbers within parentheses are p-value.

**TABLE 8.** Diagnostic test of VAR*.

| | Horizon | | | |
|---|---|---|---|---|
| Diagnostic Test | 3 | 6 | 9 | 12 |
| Durbin Watson | 1.9469 | 1.947634 | 1.95051 | 1.953028 |
| Normality_Jarque Bera | 350.2051 | 343.4979 | 344.9102 | 350.2408 |
| | (8.99E-77) | (2.57E-75) | (1.27E-75) | (8.83E-77) |
| Normality_Omni | 102.3902 | 101.0115 | 101.1529 | 101.6643 |
| | (5.84E-23) | (1.16E-22) | (1.08E-22) | (8.39E-23) |
| | Horizon | | | |
| Diagnostic Test | 15 | 18 | 21 | 24 |
| Durbin Watson | 1.956045 | 1.952713 | 1.955437 | 1.959986 |
| Normality_Jarque Bera | 351.764 | 345.4471 | 349.1734 | 350.4943 |
| | (4.12E-77) | (9.71E-76) | (1.51E-76) | (7.78E-77) |
| Normality_Omni | 101.5642 | 100.2622 | 100.7499 | 100.1289 |
| | (8.82E-23) | (1.69E-22) | (1.33E-22) | (1.81E-22) |

* The numbers within parentheses are p-value.

## REFERENCES


[1] H. Takayasu, "Practical fruits of econophysics," in *Proc. 3rd Nikkei Econophys. Symp.*, Tokyo, Japan, 2006.

[2] Y. Zuo and E. Kita, "Stock price forecast using Bayesian network," *Expert Syst. Appl.*, vol. 39, no. 8, pp. 6729–6737, Jun. 2012, doi: 10.1016/j.eswa.2011.12.035.

[3] A. Kazem, E. Sharifi, F. K. Hussain, M. Saberi, and O. K. Hussain, "Support vector regression with chaos-based firefly algorithm for stock market price forecasting," *Appl. Soft Comput.*, vol. 13, no. 2, pp. 947–958, Feb. 2013, doi: 10.1016/j.asoc.2012.09.024.

[4] G. B Fang, 'The features of volatility clustering parametric and nonparametric analysis about China's stock marcket," *Technol. Econ.*, no. 10, 2007.

[5] D. Hallac, S. Vare, S. Boyd, and J. Leskovec, "Toeplitz inverse covariance-based clustering of multivariate time series data," in *Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Halifax, NS, Canada, Aug. 2017, pp. 215–223.

[6] S.-Y. Shih, F.-K. Sun, and H.-Y. Lee, "Temporal pattern attention for multivariate time series forecasting," *Mach. Learn.*, vol. 108, nos. 8–9, pp. 1421–1441, Sep. 2019, doi: 10.1007/s10994-019-05815-0.

[7] F. Karim, S. Majumdar, H. Darabi, and S. Harford, "Multivariate LSTM-FCNs for time series classification," *Neural Netw.*, vol. 116, pp. 237–245, Aug. 2019, doi: 10.1016/j.neunet.2019.04.014.

[8] D. Lamponi, "Is industry classification useful to predict US Stock-Price co-movements?" *J. Wealth Manage.*, vol. 17, no. 1, pp. 71–77, Apr. 2014.

[9] C. Narendra Babu and B. Eswara Reddy, "Prediction of selected Indian stock using a partitioning–interpolation based ARIMA–GARCH model," *Appl. Comput. Informat.*, vol. 11, no. 2, pp. 130–143, Jul. 2015.

[10] R. Stoean *et al.*, "Evolutionary regressor selection in ARIMA model for stock price time series forecasting," in *Proc. Int. Conf. Intell. Decis. Technol.*, Sorrento, Italy, 2017.

[11] Z. Tang, C. de Almeida, and P. A. Fishwick, "Time series forecasting using neural networks vs. Box-Jenkins methodology," *Simulation*, vol. 57, no. 5, pp. 303–310, Nov. 1991, doi: 10.1177/003754979105700508.

[12] M. Adya and F. Collopy, "How effective are neural networks at forecasting and prediction? A review and evaluation," *J. Forecasting*, vol. 17, nos. 5–6, pp. 481–495, Sep. 1998.

[13] R. Achkar, F. Elias-Sleiman, H. Ezzidine, and N. Haidar, "Comparison of BPA-MLP and LSTM-RNN for stocks prediction," in *Proc. 6th Int. Symp. Comput. Bus. Intell. (ISCBI)*, Basel, Switzerland, Aug. 2018, pp. 48–51.

[14] L. Sayavong, Z. Wu, and S. Chalita, "Research on stock price prediction method based on convolutional neural network," in *Proc. Int. Conf. Virtual Reality Intell. Syst. (ICVRIS)*, Jishou, China, Sep. 2019, pp. 173–176.

[15] Y. Zhou, "Failure trend analysis using time series model," in *Proc. 29th Chin. Control Decis. Conf. (CCDC)*, Chongqing, China, May 2017, pp. 859–862.

[16] A. Kaushal and P. Chaudhary, "News and events aware stock price forecasting technique," in *Proc. Int. Conf. Big Data, IoT Data Sci. (BID)*, Pune, India, Dec. 2017, pp. 8–13.

[17] X. Gao, Q. Du, L. Tian, Z. Mao, and Y. Wang, "The data mining technique of time-series trending structure series," in *Proc. 6th World Congr. Intell. Control Autom.*, Dalian, China, 2006, pp. 6019–6023.

[18] S. Zhong and W. Gang, "Study on algorithm of dependent pattern discovery of multiple time series data stream," in *Proc. Int. Conf. Comput. Sci. Service Syst. (CSSS)*, Nanjing, China, Jun. 2011, pp. 767–769.

[19] S. S. Ratakonda and S. Sasi, "Seasonal trend analysis on multi-variate time series data," in *Proc. Int. Conf. Data Sci. Eng. (ICDSE)*, Kochi, India, Aug. 2018, pp. 1–6.

[20] J. Zhang, S. Cui, Y. Xu, Q. Li, and T. Li, "A novel data-driven stock price trend prediction system," *Expert Syst. Appl.*, vol. 97, pp. 60–69, May 2018, doi: 10.1016/j.eswa.2017.12.026.

[21] M. Wen, P. Li, L. Zhang, and Y. Chen, "Stock market trend prediction using high-order information of time series," *IEEE Access*, vol. 7, pp. 28299–28308, 2019, doi: 10.1109/ACCESS.2019.2901842.

[22] P. K. H. Phua, X. Zhu, and C. Haur Koh, "Forecasting stock index increments using neural networks with trust region methods," in *Proc. Int. Joint Conf. Neural Netw.*, Portland, OR, USA, 2003, pp. 260–265.

[23] W. Huang, Y. Nakamori, and S.-Y. Wang, "Forecasting stock market movement direction with support vector machine," *Comput. Oper. Res.*, vol. 32, no. 10, pp. 2513–2522, Oct. 2005, doi: 10.1016/j.cor.2004.03.016.

[24] A. Rifa'i and D. Mahdiana, "Optimized fuzzy backpropagation neural network using genetic algorithm for predicting Indonesian stock exchange composite index," in *Proc. 6th Int. Conf. Electr. Eng., Comput. Sci. Informat. (EECSI)*, Bandung, Indonesia, Sep. 2019, pp. 195–199.

[25] Y. Du, "Application and analysis of forecasting stock price index based on combination of ARIMA model and BP neural network," in *Proc. Chin. Control Decis. Conf. (CCDC)*, Shenyang, China, Jun. 2018, pp. 2854–2857.

[26] S. Jeon, B. Hong, and V. Chang, "Pattern graph tracking-based stock price prediction using big data," *Future Gener. Comput. Syst.*, vol. 80, pp. 171–187, Mar. 2018, doi: 10.1016/j.future.2017.02.010.

[27] T.-L. Chen and F.-Y. Chen, "Examining stock index return with pattern recognition model based on cumulative probability-based granulating method by expert knowledge," *Granular Comput.*, vol. 4, no. 4, pp. 671–685, Oct. 2019.

[28] Y. Liu and X. Liu, "A trend-based stock index forecasting model with gated recurrent neural network," in *Proc. IEEE Int. Conf. Prog. Informat. Comput. (PIC)*, Suzhou, China, Dec. 2018, pp. 425–429.

[29] T. Xia, Q. Sun, A. Zhou, S. Wang, S. Xiong, S. Gao, J. Li, and Q. Yuan, "Improving the performance of stock trend prediction by applying GA to feature selection," in *Proc. IEEE 8th Int. Symp. Cloud Service Comput. (SC2)*, Paris, France, Nov. 2018, pp. 122–126.

[30] X.-X. Cen and J.-T. Qin, "Application of improved k-means clustering in stock price fluctuation trend," *Sci. Technol. Ind.*, vol. 1, p. 28, Jan. 2016.

[31] G. Lai, W.-C. Chang, Y. Yang, and H. Liu, "Modeling long-and short-term temporal patterns with deep neural networks," in *Proc. 41st Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Ann Arbor, MI, USA, Jun. 2018, pp. 95–104.

[32] R. Sharda and R. B. Patil, "Connectionist approach to time series prediction: An empirical test," *J. Intell. Manuf.*, vol. 3, no. 5, pp. 317–323, Oct. 1992, doi: 10.1007/BF01577272.

[33] R. J. Van Eyden, *The Application of Neural Networks in the Forecasting of Share Prices*, no. 5. Haymarket, VA, USA: Finance & Technology Publishing, 1995.

[34] S. Haykin and N. Network, "A comprehensive foundation," *Neural Netw.*, vol. 2, p. 41, Feb. 2004.

[35] H. Ouyang, X. Wei, and Q. Wu, "Agricultural commodity futures prices prediction via long-and short-term time series network," *J. Appl. Econ.*, vol. 22, no. 1, pp. 468–483, Jan. 2019, doi: 10.1080/15140326.2019.1668664.

[36] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. 32nd Int. Conf. Mach. Learn. (ICML)*, Lille, France, 2015, pp. 448–456.

[37] L. Trottier, P. Giguere, and B. Chaib-draa, "Parametric exponential linear unit for deep convolutional neural networks," in *Proc. 16th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Cancún, Mexico, Dec. 2017, pp. 207–214.

[38] C. Lea, R. Vidal, A. Reiter, and G. D. Hager, "Temporal convolutional networks: A unified approach to action segmentation," in *Proc. Eur. Conf. Comput. Vis., Comput. Vis. (ECCV)*, Amsterdam, The Netherlands, 2016, pp. 47–54.

[39] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Salt Lake City, UT, USA, Jun. 2018, pp. 7132–7141.

[40] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, San Diego, CA, USA, 2015, pp. 1–15.

[41] D. T. Tran, A. Iosifidis, J. Kanniainen, and M. Gabbouj, "Temporal attention-augmented bilinear network for financial time-series data analysis," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 5, pp. 1407–1418, May 2019, doi: 10.1109/TNNLS.2018.2869225.

[42] Y. Mäkinen, J. Kanniainen, M. Gabbouj, and A. Iosifidis, "Forecasting jump arrivals in stock prices: New attention-based network architecture using limit order book data," *Quant. Finance*, vol. 19, no. 12, pp. 2033–2050, Dec. 2019.

[43] R. Giacomini and H. White, "Tests of conditional predictive ability," *Econometrica*, vol. 74, no. 6, pp. 1545–1578, Nov. 2006.

[44] R. Giacomini and B. Rossi, "Forecast comparisons in unstable environments," *J. Appl. Econometrics*, vol. 25, no. 4, pp. 595–620, Jun. 2010.

[45] W. K. Newey and K. D. West, "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix," Nat. Bur. Econ. Res., Cambridge, MA, USA, Tech. Rep. 55, 1986.

[46] J.-L. Wang and S.-H. Chan, "Stock market trading rule discovery using pattern recognition and technical analysis," *Expert Syst. Appl.*, vol. 33, no. 2, pp. 304–315, Aug. 2007.

**HONGBING OUYANG** was born in Hubei, China, in 1968. He received the Ph.D. degree in finance from the City University of Hong Kong, in November 2004. He is currently a Professor with the School of Economics, Huazhong University of Science and Technology. His research interests include financial economics, financial econometrics, investments, and financial regulation.

**XIAOLU WEI** was born in Hubei, China, in 1991. She received the master's degree in mathematical economics from Northeastern Agricultural University, Harbin, China, in June 2016. She is currently pursuing the Ph.D. degree in finance with the Huazhong University of Science and Technology, Wuhan, China. She has passed Level III examination of CFA, in 2016. Her current research interests include financial application, economic prediction, and corporate finance.

**QIUFENG WU** was born in Heilongjiang, China, in 1979. He received the Ph.D. degree in computer application technology from the Harbin Institute of Technology, Harbin, China, in January 2014. He is currently working as an Associate Professor with the College of Science, Northeast Agricultural University. His current research interests include machine learning, computer vision, and smart agriculture. He is a CCF member.