ISSN: 2249-7196

IJMRR/April. 2025/ Volume 15/Issue 2s/284-288

Vendra Devendra Kumar/ International Journal of Management Research & Review

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

Vendra Devendra Kumar
PG scholar, Department of MCA, DNR College, Bhimavaram, Andhra Pradesh.
K.Sridevi
(Assistant Professor), Master of Computer Applications, DNR college, Bhimavaram, Andhra Pradesh.

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 LongShort-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.

I. INTRODUCTION

1.1 AIM OF THE PROJECT

Discover Patterns in Stock Indices: Identify recurring patterns and trends in historical stock index data. Use Temporal Interval Clustering and Classification (TICC) to segment and analyze time series data. Predict Stock Index Movements: Forecast future stock index values based on identified patterns. Utilize Time-series Pattern Attention Long Short-Term Memory (TPA-LSTM) for enhanced prediction accuracy by capturing temporal dependencies and pattern attention mechanisms.

1.2 SCOPE OF THE PROJECT

The project aims to develop a sophisticated framework for discovering and predicting patterns in stock index data using a three-stage architecture that combines Temporal Interval Clustering and Classification (TICC), Time-series Pattern Attention Long Short-Term Memory (TPA-LSTM), and Multivariate Long

Short-Term Memory Fully Convolutional Networks (LSTM-FCNs).

1.3 OBJECT OF THE PROJECT

To develop a comprehensive and accurate framework for discovering patterns and predicting movements in stock indices using a novel three-stage architecture that integrates Temporal Interval Clustering and Classification (TICC), Time-series Pattern Attention Long Short-Term Memory (TPA-LSTM), and Multivariate Long Short-Term Memory Fully Convolutional Networks (LSTM-FCNs).

1.4 Introduction

Integrate Multivariate Data for Enhanced Forecasting: Combine multiple data sources (e.g., economic indicators, trading volumes, other indices) for robust prediction. Implement Multivariate Long Short-Term Memory Fully Convolutional Networks (LSTM-FCNs) to handle and process multivariate time series data effectively. Methodology Stage 1: Temporal Interval Clustering and Classification (TICC) Segment the time series data into intervals with similar statistical properties.

Identify and classify different temporal segments that capture the underlying structure and patterns of the stock index data. Stage 2: Time-series Pattern Attention Long Short-Term Memory (TPA-LSTM) Apply TPA-LSTM to the segmented data for capturing long-term dependencies and applying attention mechanisms to identify significant patterns. Enhance the model's ability to focus on critical time periods that influence future movements.

Stage 3: Multivariate Long Short-Term Memory Fully Convolutional Networks (LSTM-FCNs) Integrate various features and indicators to construct a multivariate forecasting model. Use LSTM-FCNs to leverage both LSTM's sequential learning capabilities and FCN's feature extraction strengths for improved prediction performance. Expected Outcomes Enhanced Pattern Recognition: Improved understanding of the underlying patterns and temporal segments within stock index data.

Identification of key periods that significantly impact future index movements. Accurate Stock Index Predictions: Development of a robust predictive model with high accuracy in forecasting stock index values. Application of advanced neural network architectures to capture complex temporal dynamics and multivariate relationships. Comprehensive Analysis Framework: Creation of a versatile framework that can be applied to various stock indices and financial markets.

II. LITERATURE SURVEY

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

One challenge of economics is that the systems treated by these sciences have no perfect metronome in time and no perfect spatial architecture—crystalline or otherwise. Nonetheless, as if by magic, out of nothing but randomness one finds remarkably fine-tuned processes in time. To understand this "miracle," one might consider placing aside the human tendency to see the universe as a machine.

Instead, one might address the challenge of uncovering how, through randomness (albeit, as we shall see, strongly correlated randomness), one can arrive at many temporal patterns in economics. Inspired by principles developed by statistical physics over the past 50 years—scale invariance and universality—we review some recent applications of correlated randomness to economics.

[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.

Bayesian network is a probabilistic graphical model that represents a set of random variables and their conditional dependencies via a directed acyclic graph. This paper describes the price earnings ratio (P/E ratio) forecast by using Bayesian network. Firstly, the use of clustering algorithm transforms the continuous P/E ratio to the set of digitized values. The Bayesian network for the P/E ratio forecast is determined from the set of the digitized values.

NIKKEI stock average (NIKKEI225) and Toyota motor corporation stock price are considered as numerical examples. The results show that the forecast accuracy of the present algorithm is better than that of the traditional time-series forecast algorithms in comparison of their correlation coefficient and the root mean square error.

[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.

Due to the inherent non-linearity and non-stationary characteristics of financial stock market price time series, conventional modeling techniques such as the Box–Jenkins autoregressive integrated moving average (ARIMA) are not adequate for stock market price forecasting. In this paper, a forecasting model based on chaotic mapping, firefly algorithm, and support vector regression (SVR) is proposed to predict stock market price. The forecasting model has three stages. In the first stage, a delay coordinate embedding method is used to reconstruct unseen phase space dynamics.

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

Most procedures for modeling and forecasting financial asset return volatilities rely on restrictive and complicated parametric GARCH or stochastic volatility models. The method of realized

ISSN: 2249-7196

IJMRR/April. 2025/ Volume 15/Issue 2s/284-288

volatility constructed from high-frequency intraday returns is an alternative choice for volatility measurement.

In this paper we make an empirical analysis on Chinese stock index data by using the method of nonparametric realized volatility. We find that the realized volatility can describe the Chinese stock index volatility very well. The original Chinese stock index return series show obvious leptokurtic, fat-tailed relative to the Gaussian distribution.

[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.

Subsequence clustering of multivariate time series is a useful tool for discovering repeated patterns in temporal data. Once these patterns have been discovered, seemingly complicated datasets can be interpreted as a temporal sequence of only a small number of states, or clusters. For example, raw sensor data from a fitness-tracking application can be expressed as a timeline of a select few actions (i.e., walking, sitting, running). However, discovering these patterns is challenging because it requires simultaneous segmentation and clustering of the time series.

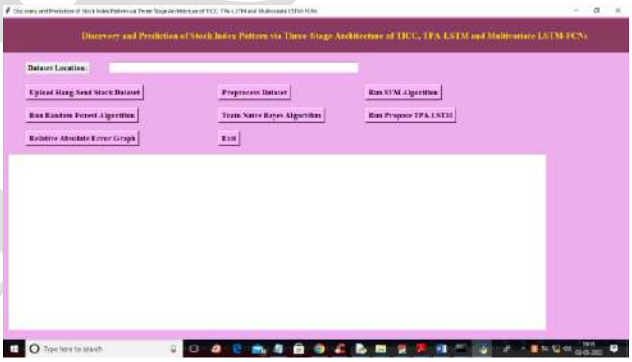
III. PROPOSED METHOD

Identification and forecasting of stock index patterns are crucial in finance to guide investment decisions and minimize losses. Recently, deep neural networks have outperformed traditional models due to their self-learning, non-parametric, and noise-tolerant nature. However, challenges like variable pattern lengths and weak periodic patterns in multivariate time series remain. TICC, a clustering method based on graphical dependency structures, helps uncover complex inter-variable relationships. TPA-LSTM, combining temporal attention and LSTM layers with an autoregressive model, effectively captures both short- and long-term patterns in financial data. Multivariate LSTM-FCNs enhance classification by integrating LSTM and convolutional blocks, capturing deep temporal features without heavy preprocessing. The dataset includes daily closing prices of the Hangseng Stock Composite Index and eleven industry indices from 2006 to 2019, preprocessed by dropping nulls and split chronologically for training, validation, and testing. TICC discovers patterns, while TPA-LSTM and LSTM-FCNs handle prediction and classification. A stationarity test using LLC, IPS, and PP methods was applied to ensure data suitability for modeling.

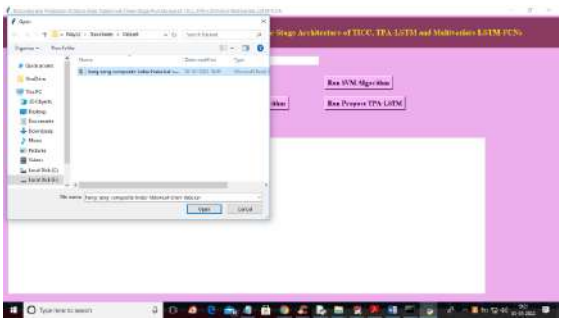
IV. RESULT

To implement this project author has used Hang-Sang dataset and I am also using same dataset

To run project double click on 'run.bat' file to get below screen



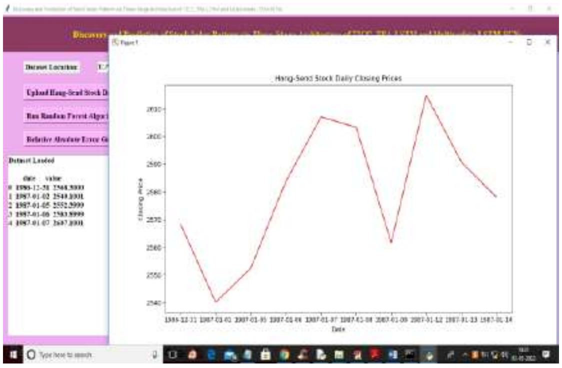
In above screen click on 'Upload Hang Send Stock Dataset' button to upload dataset and to get below output



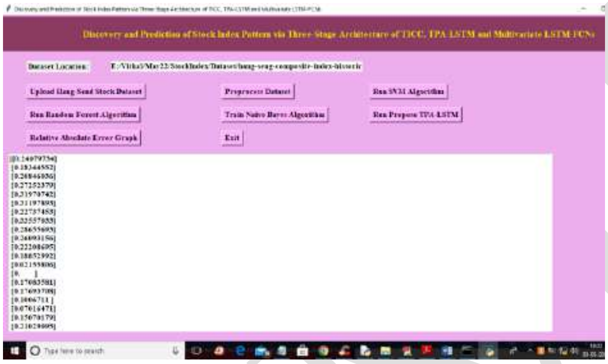
In above screen selecting and uploading 'Hang Seng' dataset file and then click on 'Open' button to load dataset and to get below output

ISSN: 2249-7196

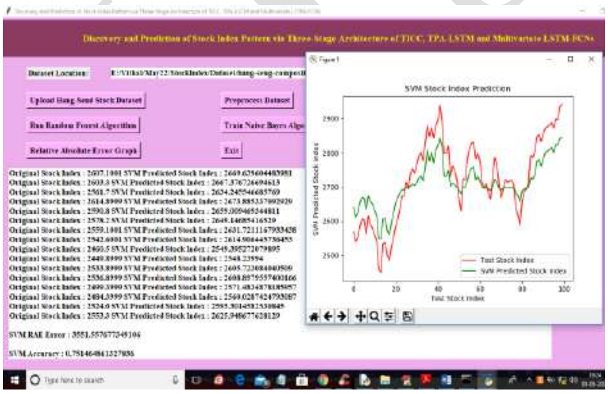
IJMRR/April. 2025/ Volume 15/Issue 2s/284-288



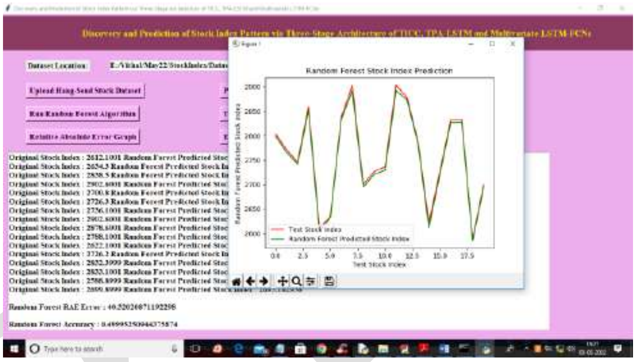
In above screen in text area we can see dataset loaded and in graph x-axis represents DATE and y-axis represent stock value on that date and now close above graph and then click on 'Preprocess Dataset' button to read all values and then normalize them and get below output



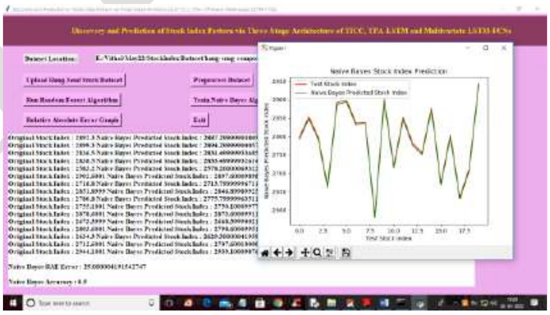
In above screen all stock values are normalize between 0 and 1 and now click on 'Run SVM Algorithm' button to get below output



In above screen text area we can see original test value and predicted value from SVM and then calculate difference between original test value and predicted value as RAE and we got SVM RAE as 3551 and accuracy as 75% and in graph x-axis represents days and y-axis represents stock values and RED line represents original test value and green line represents predicted value and we can see there is huge gap between red and green line so prediction is not accurate and if prediction is accurate then both lines get overlap and now close above graph and then click on 'Run Random Forest' button to get below output



In above screen with random forest both lines are overlapping so its RAE error reduce to 40 and its prediction is little accurate and now close above graph and then click on 'Run Naïve Bayes Algorithm' button to get below output

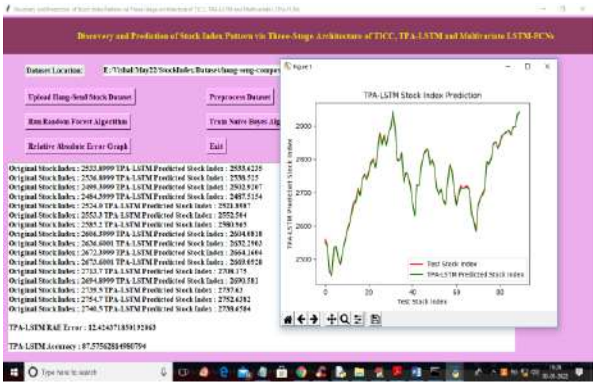


In above screen with Naïve Bayes we got 25% error rate and both lines are overlapping so its prediction is also little accurate and now close above graph and then click on 'Run Propose TPA-LSTM' button to train propose algorithm and get below output

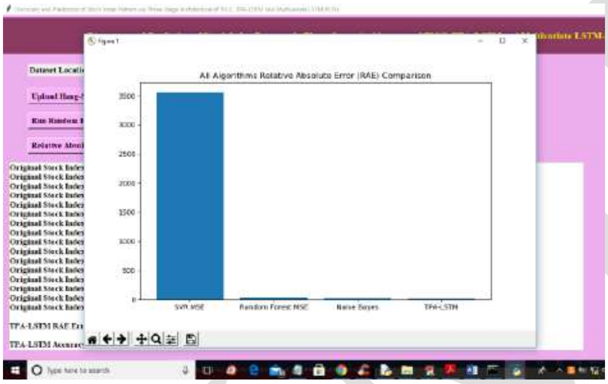
ISSN: 2249-7196

IJMRR/April. 2025/ Volume 15/Issue 2s/284-288

Vendra Devendra Kumar/ International Journal of Management Research & Review



In above screen with propose LSTM-TPA we got RAE as 12% and accuracy as 87% and we can see both lines are overlapping without any gap so propose algorithm prediction is accurate. Now close above graph and then click on 'Relative Absolute Error Graph' button to get below graph



In above screen x-axis represents algorithm names and y-axis represents RAE error and in all algorithms propose LSTM-TPA got less error rate so its performance is good

V. CONCLUSION

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.

REFERENCE

[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 LSTMFCNs 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.