International Journal of Management, IT & Engineering
Vol. 9 Issue 5, May 2019,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

Data Optimization for Price Variations Using PSO and Machine Learning Integration

Jayshree Bhangre
Research Scholar,
Veer Bahadur Singh Purvanchal University

Dr. Sanjeev Gangwar
Research Supervisor,
Veer Bahadur Singh Purvanchal University

Abstract

The stock market's primary draw is the rapid increase in economic value of stocks over a short period of time. Before making a market investment, the investor examines the organization's demonstration, expected value, and growth. Using the standard procedure or some of the already available methods recommended by various studies might not be sufficient for the analysis. With so many stocks currently on the market, it might be challenging to research each one using the few recommended foretelling techniques. We need some cutting-edge stock market prediction tool in order to know the projected stock value. This essay provides a sophisticated technique for planning, evaluating how various market organisers execute their stock, and identifying the best stock to buy by projecting the stock's closing price.The multilayer deep learning neural network optimised using Adam optimizer is the foundation of the proposed arrangement. Data from several organisations over the last six years (2010–2016) are added to the model to show the model's efficacy. The proposed framework is found to be the most appropriate for all distinct data sets from diverse sectors, according to the results. The outcome graph of the system clearly demonstrates how little prediction error there is.

Keywords: Deep learning, multilayer perception, forecasting, and soft computing in the stock market

International Journal of Management, IT & Engineering
Vol. 9 Issue 5, May 2019,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

Introduction

The value of the stock market and currency exchange rates have a significant impact in the current global economic scenario. The increase in a company's stock price reflects growth in the company's ability to conduct business, the improvement of its financial situation, and the value of its shareholders. Nowadays, people all around the world are more interested in investing money directly or indirectly in the stock market for faster fund growth than in other areas. Therefore, everyone wants to know which stock market offers the best return before making an investment. They are able to solve the problem thanks to modern technology. Today, it is nearly hard to estimate the stock price accurately using a chart.The efficient market hypothesis states that stock valuation provides recent data, which causes impulsive behaviour. The random walk suggested that it is impossible to accurately estimate stock price using historical data. Due to the high return on investment after all of this risk and uncertainty, individuals participate in the stock market. It has been difficult for researchers to create high-efficiency forecasting models because of this attraction. Although several mathematical models have been created, the results are still unsatisfactory.

For stock price prediction, ANN-based models have been utilised, and hybrid models have occasionally been used in conjunction with stock market forecasting. Because of how unpredictable the stock market is, index prediction is usually the most difficult endeavour. Many academics have embraced this challenge and created effective models that employ fuzzy logic to anticipate stock index prices using real data. In addition to these methods, which combine fuzzy with particle swarm optimization, neural network models are also being developed to improve predicting accuracy for stock market value.Using a few additional features in addition to historical data from the intended market as input to supervised models will improve the outcome of financial predictions. By focusing less on the historical data of the time series than was anticipated, the researchers were able to outperform their rival who only used historical data to achieve the same goal by using some stock market data from the South Korean stock market and product costs in an ANN-based semi-supervised model.

A model that is helpful in projecting composite relationships between unexpected changes in financial data was also proposed by the researchers. Time gaps between anticipating and talking about the importance of data regularity in sampling are another research dimension. It is difficult to determine the track for short period regularities due to the noisy nature of financial time series. In order to estimate close stock price, this article focuses on the deep learning method known as Multi Layered Perceptron (MLP). The simulation performance of the suggested model is found to be best suited to each stock market data set after this model has been trained and then applied to various stock data. For all data sets, the anticipated outcome is expressed in terms of Mean Absolute Percentage Error (MAPE).It has been found that the proposed model has a very low prediction error and is prospective enough to forecast stock market price.

The literature study section of the work exhibits information about reviews in relevant fields, and the materials and methods segment includes data sets, MLP, and optimization techniques. In the segment titled "Projected Methodology for Prediction," the projected foretelling framework, activity flow of the framework, and normalised real data were used to test the projected framework. In the segment titled "Empirical Result," there is no regeneration of the representation and empirical outcome, input option provision, statistical and financial conditions, MAPE and similitude with other data set results, as well as the projected foretelling framework.The part on "Summery and Emerging Directions" wraps up the proposed framework by describing its limitations and use.

The state of the art in stock price prediction research was conducted using MATLAB software and other statistical analytic techniques like ANOVA. Data from the BSE and NSE of the Indian stock market were used to validate the algorithm. the four to six-year data interval. The analysis of results is covered in chapter 5, along with the empirical assessment factors RMSE, MSE, and MI. The results of the investigation indicated that the proposed strategy to stock price prediction outperforms and is enhanced over the current approaches.Improved price prediction has a positive impact on national prosperity and helps investors and stock market buyers regain confidence.

Literature Review

Ecer et al. [2014] to predict the movement of the Borsa Istanbul (BIST) 100 index using a novel machine learning (ML) model. Tanh (x) and the standard Gaussian function were used as the output functions for modelling using multilayer perceptron-genetic algorithms (MLP-GA) and multilayer perceptron-particle swarm optimization (MLP-PSO), respectively. Nine technical indicators made up of historical financial time series data from 1996 to 2006 were used in this study. The accuracy and performance of the created models are compared using the Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and correlation coefficient values of the results. According to the findings, using Tanh (x) as the output function greatly increased model accuracy when compared to using the default Gaussian function.Higher testing accuracy was provided by MLP-PSO with a population size of 125, followed by MLP-GA with a population size of 50, with RMSE values of 0.732583 and 0.733063, MAPE values of 28.16% and 29.09%, and correlation coefficient values of 0.694 and 0.695, respectively. The findings show that the hybrid ML approach can successfully raise prediction accuracy.

Ronaghi et al. [2015] a system called the Noisy Deep Stock Movement Prediction Fusion (ND-SMPF) for forecasting stock price movement. To improve the accuracy of the stock movement prediction task, the suggested ND-SMPF predictive system leverages information fusion to mix twitter data with extended horizon market historical prices. In order to extract news level temporal information, Noisy Bi-directional Gated Recurrent Unit (NBGRU) and a Hybrid Attention Network (HAN) are specifically used. To complete the prediction task, historical price data is combined with relevant terms that have the highest correlation and influence on stock trends. This is done using a two level attention layer. The proposed ND-SMPF framework is evaluated using a real dataset, and the results show that it performs better than its recently created competitors.

Mehtab et al. [2016] a collection of regression models powered by deep learning that produce extremely accurate stock price predictions. This study used historical stock price data from a well-known firm listed on the National Stock Exchange (NSE) of India between December 31, 2012 and January 9, 2015 to create our forecasting models. Every working day of the week, the stock prices are noted at intervals of five minutes. The suggested system has developed four convolutional neural network (CNN) and five long- and short-term memory (LSTM)-based deep learning models for precise forecasting of the future stock prices using these incredibly detailed stock price data.Additionally, based on the execution times and root mean square errors (RMSE) values of each suggested model, this research effort presents detailed results on forecasting accuracies.

In order to create a prediction model for predicting stock price, Biswas et al. [2016] used the commonly used algorithms Long Short Term Memory, Extreme Gradient Boosting (XGBoost), Linear Regression, Moving Average, and Last Value Model using more than twelve months of historical stock data. The Mean Absolute Percentage Error (MAPE) measurement is used to compare the models, and it is shown that the LSTM technique outperforms them all with a MAPE of 0.635. Moving Average has the highest error rate among the five models for our situation as well.

Yu et al. [2017] Financial product pricing data is reconstructed using the time series phase-space reconstruction (PSR) approach as a one-dimensional series produced by the projection of a chaotic system made up of numerous components into the time dimension. To predict stock prices, a DNN-based prediction model is created based on the PSR technique using long- and short-term memory networks (LSTMs) for DL. Multiple stock indices for various time periods are predicted using the suggested prediction model as well as some other methods. The results of a comparison demonstrate that the proposed prediction model has a greater level of prediction accuracy.Through three process steps—time series data processing, network model development, and result assessment and analysis—a prediction model is proposed for financial price data, which are non-stationary and generally noisy time series. A DNN-based LSTM network model is integrated with the PSR time series analysis approach. At the pre-processing stage, data de-noising and normalisation are carried out. Data structure is additionally accomplished by partitioning using a temporal window.

Objective of the study

We defined the following goals for our research after determining the need for development in the area of machine learning-based option price prediction:

1. To analyse several methods for enhancing the accuracy of option price prediction.
2. To create and put into use an MLP network for prediction.

3. To create and put into practise a feature re-education classification algorithm.

Research Methodology

Firefly Algorithm

The meta-heuristic optimization technique known as the "firefly algorithm" is based on the environment-dependent flashing characteristics of fireflies. The Firefly algorithm manages the changing behaviours of data and solves NP-hard problems. It uses a random algorithm, or to put it another way, it does a random search to find a set of solutions. At its most basic level, the FA concentrates on developing solutions inside a search area and choosing the best surviving option. A random search prevents getting bogged down in regional optimums. In metaheuristic algorithms, exploration means finding several solutions inside the search space, but exploitation means narrowing the search to the best surrounding solutions.

The three fundamental components of the Firefly algorithm are

(1) The firefly changes into a All fireflies are of the same sex, and when they move randomly, they become brighter and more alluring.

(2) The brightness of the light and the distance from it affect the firefly' appeal. The light intensity reduction is calculated using the light absorption coefficient. The luminance of the firefly is likewise determined by the value of the goal function.

(3) Equation (1) is used to calculate the distance between fireflies, and Xi,k is the kth component of the spatial coordination and the ith firefly.

$$r_{i,j=}\sqrt{(xi-xj)^2+\ (yi-yj)^2 \ldots\ldots\ldots\ldots\ldots\ldots.(4.4.1)}$$

The movement of firefly and attracted fireflies measured as

$$X_i = X_i + B0^{erij\,2}(Xj - Xi) + a\left(rand - \frac{1}{2}\right) \ldots\ldots\ldots\ldots.(4.4.2)$$

A is a randomizer variable, rand is a random integer between [0, 1], and B is the attractiveness of the light source. The parameter is determined by variations in attraction.

The process of stock price optimization

The firefly algorithm (FA) employed in stock data as initial population and set the value of parameters as α, γ, β min $and\ t = 0$ and Fes=0; the factors of brightness of data is Ii at Xi is measured by f(Xi).

Data Analysis and Discussion

Table 1 compares the performance of the proposed technique with that of EM, ML, SVM, SOM, and ANN using the parameters RMSE, NMSE, MAE, and MI analysis of the SBI bank dataset using 0.1, 0.2, 0.3, 0.4, and 0.5 settle prices.

SETTLE PRICE	METHOD	RMSE	NMSE	MAE	MI
0.1	EM	6.84	7.36	5.46	4.36
	ML	6.16	7.46	5.64	4.56
	SVM	5.94	6.23	4.23	4.31
	SOM	6.23	6.78	4.38	5.01
	ANN	7.59	6.37	4.87	5.26
	PROPOSED	5.35	5.89	3.96	5.98
0.2	EM	6.63	7.25	5.31	5.37
	ML	6.54	7.13	5.12	5.47
	SVM	5.78	6.12	4.65	5.49
	SOM	6.24	6.01	4.51	5.79
	ANN	7.21	5.96	4.32	6.01
	PROPOSED	5.21	5.34	3.67	6.35
0.3	EM	6.34	7.14	5.16	5.65
	ML	6.21	6.98	5.01	5.72
	SVM	5.67	6.14	4.35	5.78

	SOM	6.01	6.03	4.21	5.98
	ANN	6.57	5.81	4.08	6.13
	PROPOSED	5.15	5.29	3.56	6.75
0.4	EM	6.11	6.87	5.01	5.49
	ML	6.19	6.71	4.96	5.56
	SVM	5.37	6.06	4.21	5.73
	SOM	5.98	5.91	4.65	5.91
	ANN	5.66	5.73	4.29	6.24
	PROPOSED	5.11	5.01	3.49	6.89
0.5	EM	5.99	6.89	5.11	5.24
	ML	5.68	6.74	4.93	5.32
	SVM	5.29	6.01	4.68	5.51
	SOM	5.31	5.70	4.31	5.68
	ANN	5.01	5.43	5.01	6.31
	PROPOSED	4.99	4.31	3.21	7.01

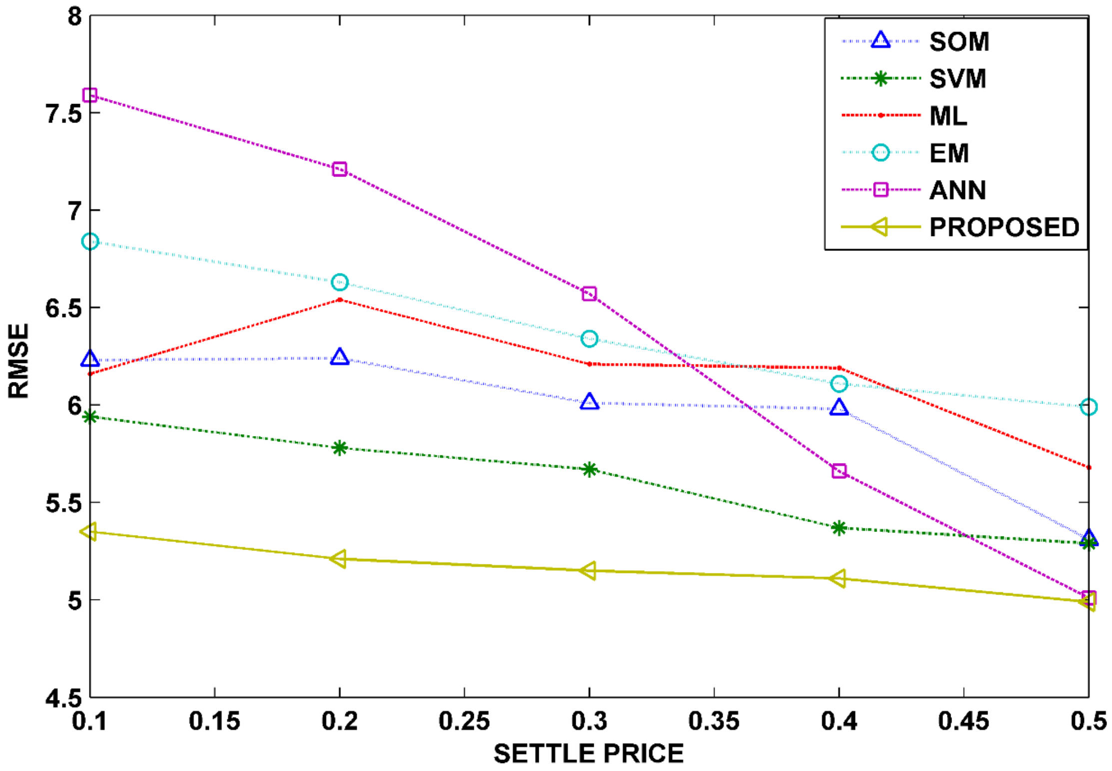

Fig 1: pricing performance of the option Settle price vs. RMSE

This graph compares the root mean square error for the SBI Bank dataset using SOM, SVM, ML, EM, ANN, and the suggested approach. The variation's distribution across different settle prices, such as 0.1, 0.2, 0.3, 0.4, and 0.5, shows that the suggested method's optimization process and improved prediction have improved the value of RMSE. Here, we see that the proposed method's root mean square error value is superior to that of previous approaches.

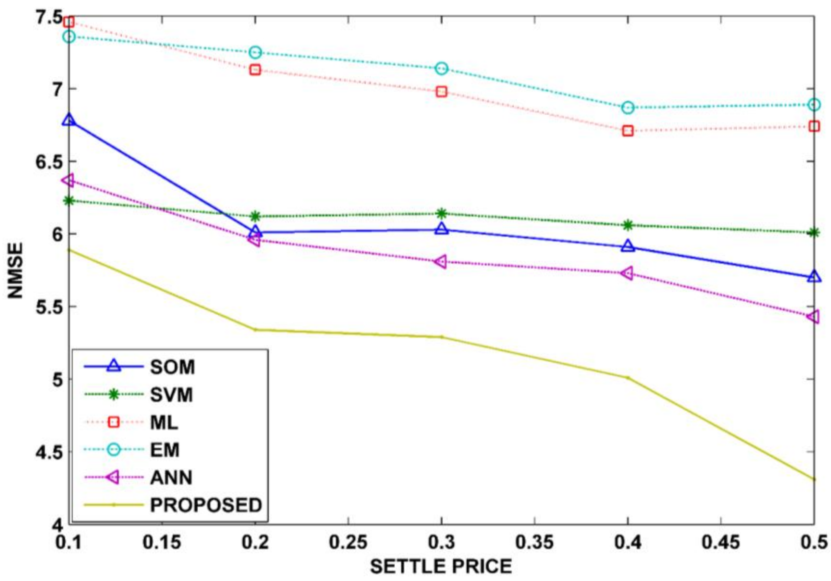

Fig 2: NMSE vs. Settle price.

This graph compares the normalised mean square error for the SBI Bank dataset using SOM, SVM, ML, EM, ANN, and the suggested approach. The variation's distribution across different settle prices, such as 0.1, 0.2, 0.3, 0.4, and 0.5, shows that the suggested method's optimization process has improved the value of NMSE. Here, we see that the proposed method's normalised mean square error value is superior to previous approaches.

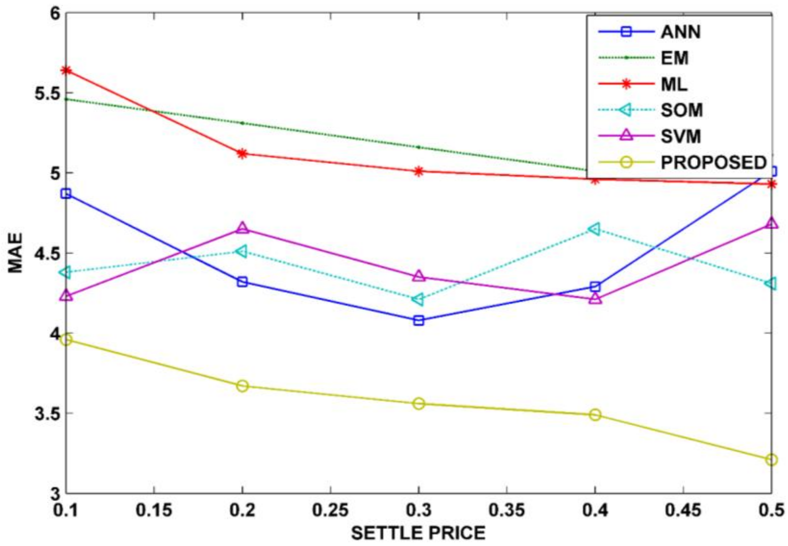

Fig 3: MAE vs. Settle price.

The mean absolute error variation for the SBI Bank dataset is shown in this graph for SOM, SVM, ML, EM, ANN, and the suggested approach. The variation's distribution across multiple settle prices, such as 0.1, 0.2, 0.3, 0.4, and 0.5, shows that the proposed approach's method is more accurate and efficient in optimising MAE value. Here, we see that the mean absolute error value of the suggested strategy is lower than that of other approaches.

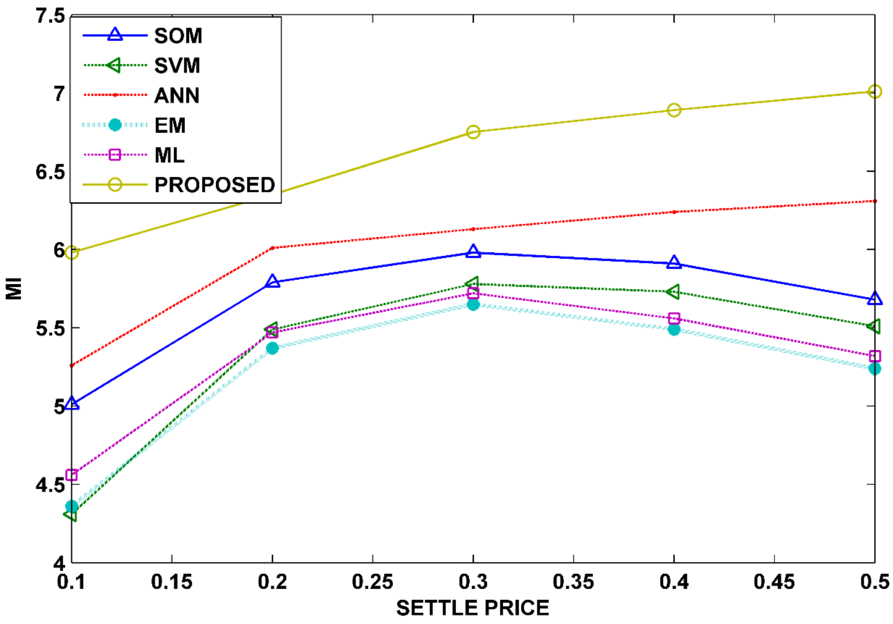

Fig 4 : MI vs. Settle price.

This graph shows how the suggested technique for the SBI Bank dataset compares against SOM, SVM, ML, EM, and SVM mutual information variations. The variation's distribution across different settle prices, such as 0.1, 0.2, 0.3, 0.4, and 0.5, shows that the suggested method's optimization process has improved the value of MI. Here, we see that the proposed method's mutual information value is superior than existing approaches.

Conclusion

Combining an optimization technique with MLP neural network-based categorization constitutes the used algorithms. The variation in attribute selection for the classification process is decreased by the optimization approach. Particle swarm optimization is the name of the optimization algorithm. And finally, evaluate the value of computation for the ultimate choice of the ideal price for the MLP classifier's input processing. Selected numbers of attributes were used in every data instance. Four standard parameters were employed for the examination of performance: NMSE, RMSE, MAE, and MI. The results show that MLP classifier performs better than classifiers based on deep neural networks. The selection of features properties of price variation causes the deep neural network-based classifier to fail.The value of option pricing is increased by the particle swarm optimization approach, which decreases attribute variation.

The stock data with the lowest variance is then used for detection and classification. Three machine learning methods were used in the classification procedure. Support vector machines, ensemble classifiers, and BP neural network models make up these algorithms. The sample size and degree of data fluctuation are modest. The OK prediction rate is carried out by the support vector machine. The rate of classification compromise is, however, rising as the sample size increases. Ensemble-based classifiers improve price prediction accuracy. Bagging techniques are used in the bag of two classifiers, support vector machines and decision trees, for the ensemble processing.The decision tree method maps the data with the strike price for the support vector machine and calculates the price entropy factor of the risk interest of the data. The support vector machine is compressed as a result of the performance of the ensemble of two classifiers. Another classification strategy was employed by the BP neural network model. The BP neural network model also yields effective outcomes. Four parameters, such as RMSE, MAE, MI, and MSE, should be calculated in order to validate the performance of the classifier. As a result, choosing the right qualities and features might affect stock price, therefore in the future, a feature selection algorithm and cascaded classifier design will be developed for better prediction.

References

[1]. Jin, Zhigang, Yang Yang, and Yuhong Liu. "Stock closing price prediction based on sentiment analysis and LSTM." Neural Computing and Applications 32, no. 13 (2010): 9713-9729.

[2]. Livieris, Ioannis E., Emmanuel Pintelas, and Panagiotis Pintelas. "A CNN–LSTM model for gold price time-series forecasting." Neural computing and applications 32, no. 23 (2010): 17351-17360.

[3]. Wu, Jimmy Ming-Tai, Zhongcui Li, Norbert Herencsar, Bay Vo, and Jerry Chun-Wei Lin. "A graph-based CNN-LSTM stock price prediction algorithm with leading indicators." Multimedia Systems (2011): 1-20.

[4]. Nti, Isaac Kofi, Adebayo Felix Adekoya, and Benjamin AsubamWeyori. "A systematic review of fundamental and technical analysis of stock market predictions." Artificial Intelligence Review 53, no. 4 (2010): 3007-3057.

[5]. Xu, Ying, Cuijuan Yang, Shaoliang Peng, and Yusuke Nojima. "A hybrid two-stage financial stock forecasting algorithm based on clustering and ensemble learning." Applied Intelligence 50, no. 11 (2015): 3852-3867.

[6]. Pang, Xiongwen, Yanqiang Zhou, Pan Wang, Weiwei Lin, and Victor Chang. "An innovative neural network approach for stock market prediction." The Journal of Supercomputing 76, no. 3 (2012): 2098-2118.

[7]. Huang, Ming-Hui, and Roland T. Rust. "A strategic framework for artificial intelligence in marketing." Journal of the Academy of Marketing Science 49, no. 1 (2014): 30-50.

[8]. Hao, Yan, and Chengshi Tian. "A hybrid framework for carbon trading price forecasting: the role of multiple influence factor." Journal of Cleaner Production 262 (2016): 120378.

[9]. Long, Wen, Zhichen Lu, and Lingxiao Cui. "Deep learning-based feature engineering for stock price movement prediction." Knowledge-Based Systems 164 (2009): 163-173.

[10]. Mondal, Saikat, A. Dutta, and P. Chatterjee. "Application of Deep Learning Techniques for Precise Stock Market Prediction." In National Conference on Machine Learning and Artificial Intelligence. 2013.

[11]. Sunny, MdArifIstiake, Mirza MohdShahriarMaswood, and Abdullah G. Alharbi. "Deep learning-based stock price prediction using LSTM and bi-directional LSTM model." In 2012 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES), pp. 87-92. IEEE, 2012.

[12]. Parray, Irfan Ramzan, Surinder Singh Khurana, Munish Kumar, and Ali A. Altalbe. "Time series data analysis of stock price movement using machine learning techniques." Soft Computing 24, no. 21 (2008): 16509-16517.

[13]. Wei, Xiangyu, ZhilongXie, Rui Cheng, Di Zhang, and Qing Li. "An Intelligent Learning and Ensembling Framework for Predicting Option Prices." *Emerging Markets Finance and Trade* 57, no. 15 (2008): 4237-4260.