Certain Trend Analysis on Stock Price Prediction using Machine Learning Algorithm-LSTM

R. Sivasankar[1], S. Nithyabharathi[2]
Research Student[1], Assistant Professor[2]
Department of CSE[1, 2]
IFET College of Engineering[1, 2]
Villupuram, India
Shivasan26001@gmail.com, Iniyanthigazh@gmail.com

Abstract--This paper exhibits a machine learning based frame work for stock price prediction. Stocks are a fundamental financial instrument that can be traded in the financial markets. Owning a company's stock means you have the opportunity to share in its profits and, in certain cases, participate in decision-making processes through voting rights. we proposed a nine machine learning models which consists Decision Tree, Random Forest, Adaptive Boosting (Adaboost), eXtreme Gradient Boosting (XGBoost), Support Vector Classifier (SVC), Naïve Bayes, K-Nearest Neighbors (KNN), Logistic Regression and Artificial Neural Network (ANN) and two powerful deep learning methods (Recurrent Neural Network (RNN) and Long short-term memory (LSTM).In this research, we concentrate on comparing prediction performance of different machine learning models and deep learning methods to predict stock market movement. Many technical indicators are utilized as inputs to our models. Our study includes two different approaches for inputs, continuous data and binary data. The future up or down trend is identified and when binary data is given as the input values to the predictors, we enter data with a recognized trend based on each feature's property.

Keywords-stock price prediction, machine learning models, ANN, RNN, LSTM;

I. Introduction

The prediction of stock prices has long been a subject of intense interest, not only among investors but also within the fields of finance and technology. As financial markets continue to evolve and become increasingly complex, the need for accurate and timely forecasts has grown exponentially. To address this challenge, the integration of machine learning and deep learning techniques has become a crucial focus of research and application. In an era where vast volumes of financial data are readily accessible and computing power is at an all-time high, machine learning and deep learning offer unique capabilities to analyze, model, and predict stock prices. These techniques enable us to uncover intricate patterns, relationships, and insights within historical market data, transcending the limitations of traditional approaches. Stock price prediction holds considerable significance as it can empower investors, traders, and financial institutions with informed decisions regarding asset allocation, risk management, and trading strategies. Traditional methods, relying on fundamental analysis and technical indicators, often struggle to adapt to the rapidly changing dynamics of financial markets. Machine learning and deep learning, on the other hand, provide the tools to construct adaptive models that can assimilate diverse data sources, including historical price movements, trading volumes, news sentiment, macroeconomic indicators, and more. This research endeavors to explore and apply the capabilities of machine learning and deep learning in the realm of stock price prediction. Leveraging historical market data, we aim to develop predictive models that outperform conventional techniques, enhancing our understanding of the potential of artificial intelligence in addressing real-world financial challenges.

The section of this paper is organized as follows, and section II narrates related works on Machine Learning utilizing different tools over the network. Section III describes the processing flow of modules with consensus rules and roles assigned to the users, section IV reviews tool analysis, and requirements for proposed system implementation. Section V settles the paper with impending investigations.

II. RELATED WORK

This section reviews related works of Machine Learning employed in stock price prediction. W. Long [4] proposed a novel end-to-end model named the multi-filters neural network (MFNN) specifically for the purpose of feature extraction from financial time series data and for the task of predicting price movements. M. R. Hassan [7] recommended a fusion model by combining the Hidden Markov Model (HMM), Artificial Neural Networks (ANN) and Genetic Algorithms (GA) to predict the financial market behaviour. Y. Nakamori [8] proposed a Support Vector Machine (SVM) is a unique machine learning algorithm distinguished by its ability to control the capacity of the decision function, its utilization of kernel functions, and the sparse nature of its solution. M. Ballings [14] aims to compare and evaluate ensemble methods such as Random Forest, AdaBoost, and Kernel Factory in contrast to single classifier models like Neural Networks, Logistic Regression, Support Vector Machines, and K-Nearest Neighbor to improve the performance of prediction rate. D. Bhanu Sravanthi [18] proposed to compare the two distinct categories of models utilized for stock prediction which is linear models and non-linear models. Specifically, the study delves into non-linear modeling techniques, employing Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), and subsequently compares the outcomes of these models with the actual results. J. Patel [25] recommended a two-stage fusion methodology that harnesses the power of Support Vector Regression (SVR) as its primary component in the initial stage. Subsequently, in the second stage of the fusion process, an ensemble of diverse machine learning models, including Artificial Neural Network (ANN), Random Forest (RF), and SVR, is employed.

The existing system focused only on macroeconomic or technical features with recent machine learning methods to detect stock index or values. The accuracy rate of existing system is comparatively low.

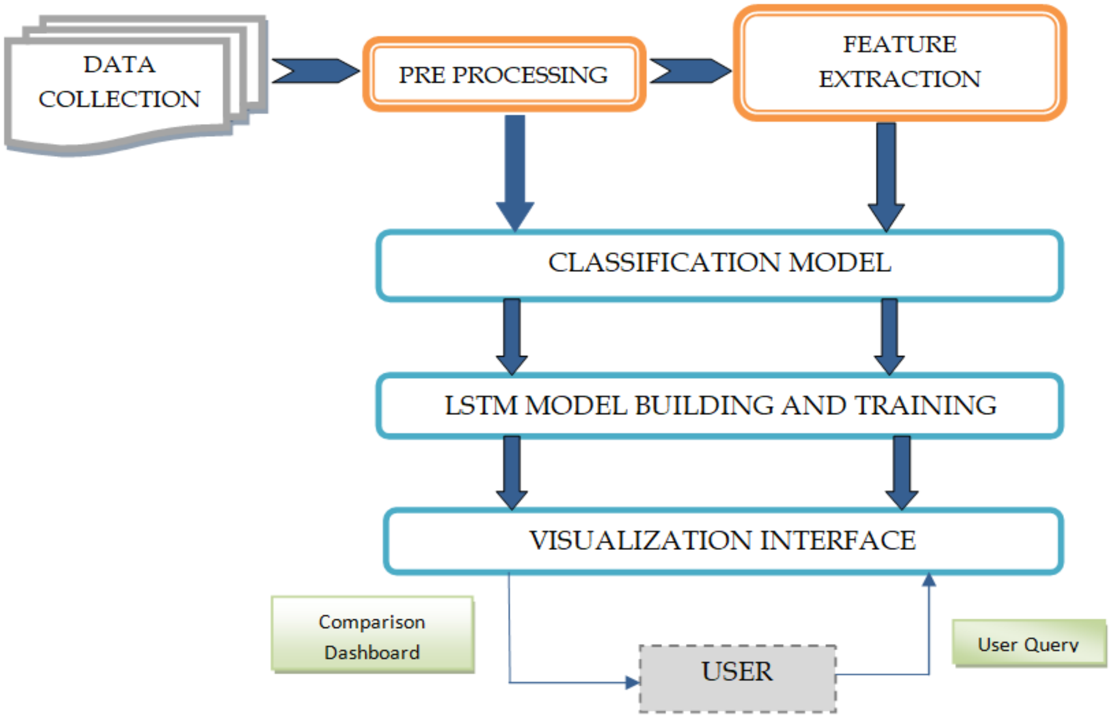


Fig.1.An overview of proposed framework Architecture

III. PROPOSED METHODOLOGY

This paper presents an enhanced stock price prediction system that leverages machine learning algorithms to provide accurate and reliable stock price predictions. The proposed system is a web application built using the Flask framework, allowing users to select a stock dataset, choose from various machine learning models, and visualize predictions. Additionally, it offers the functionality for users to upload custom datasets for prediction, demonstrating the versatility and adaptability of the system. Certainly! Here are more advantages of the proposed content and system. The system provides a user-friendly web interface that makes it accessible to a wide range of users, including investors, financial analysts, and individuals interested in stock price predictions. Users can easily interact with the system, select stocks, and choose prediction models. It supports multiple machine learning models for stock price prediction. This versatility allows users to experiment with different algorithms and choose the one that best suits their needs and preferences. Users have the option to upload their own datasets for prediction. This feature is particularly beneficial for those who have proprietary or specific data sources, making the system adaptable to various data inputs. The integration of machine learning.

IV. MACHINE LEARNING ALGORITHMS

A. RANDOM FOREST ALGORITHM (RF)

A random forest is an ensemble machine learning technique used for classification and regression tasks. It combines multiple decision trees to make more accurate predictions. Each decision tree is trained on a random subset of the data and a random subset of the features, which helps reduce overfitting and improve generalization. The final prediction in a random forest is determined by averaging or taking a majority vote from the predictions of individual trees, resulting in a robust and versatile predictive model.

A. LINEAR REGRESSION MODEL (LR)

Linear regression is a simple yet widely used statistical and machine learning technique for predicting stock prices. It is important to note that while linear regression can provide insights and basic predictions, it may not capture the complexities and dynamics of financial markets. Here's how linear regression can be applied to stock price prediction

$$Y=aX+b$$

Where Y is the predicted stock price,X represents the chosen features, such as historical prices, trading volume, and economic indicators, a is the coefficient that represents the impact of the features on the stock price,b is the intercept, indicating the baseline value.

B. LONG SHORT-TERM MEMORY (LSTM)

Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) architecture used for price prediction in financial markets. It is designed to capture and model complex, sequential patterns in time-series data. Unlike traditional feedforward neural networks, LSTMs are equipped to handle the temporal dependencies and memory retention essential for analysing and predicting stock prices.

Input Gate (i_t): $i_t = \text{sigmoid}(W_i * [h_{t-1}, x_t] + b_i)$

Output Gate (o_t): $o_t = \text{sigmoid}(W_o * [h_{t-1}, x_t] + b_o)$

In essence, LSTM models are a powerful tool for capturing and learning from historical price data, enabling them to provide more sophisticated and accurate predictions compared to simpler methods like linear regression. These models are commonly used in the field of quantitative finance and have contributed to improved decision-making in stock trading and investment strategies.

into financial decision-making. This demonstrates the potential of advanced technologies to assist the user.

C. RECURRENT NEURAL NETWORK (RNN)

A Recurrent Neural Network (RNN) is a type of artificial neural network designed to process sequential data, making it a valuable tool in stock price prediction. Unlike traditional feedforward neural networks, RNNs have a unique architecture that allows them to maintain internal memory or state, which is crucial for modelling time-dependent data like historical stock prices.

$$h_t = \text{activation\_function}(W_h * h_{t-1} + W_x * x_t + b)$$

where h_t represents the hidden state at time step t, x_t represents the input at time step t. In stock price prediction, this could be, for example, the price at a given time, W_h is the weight matrix for the previous hidden state, W_x is the weight matrix for the current input, b is the bias term for the activation.

D. DECISION TREE MODEL

In a decision tree, the process of predicting the class for a given dataset begins at the tree's root node. This algorithm carefully assesses the values of the root attribute, comparing them with the attributes of the dataset record in question. Based on this comparison, the algorithm takes a decisive step, following a specific branch that leads to the next node in the tree. At this new node, the algorithm once again engages in attribute value comparisons with the available sub-nodes. It proceeds further into the tree structure, and this process continues iteratively until it arrives at the ultimate destination: a leaf node. This step-by-step journey culminates when the algorithm reaches the leaf node of the tree. It is at this leaf node that the final class prediction is made, thus completing the classification process. This intuitive algorithm allows for a clear understanding of how a decision tree navigates its structure to make predictions.

E. XTREME GRADIENT BOOSTING

Extreme Gradient Boosting (XGBoost) is a cutting-edge machine learning algorithm that belongs to the ensemble learning family. It is widely recognized for its exceptional performance in a variety of machine learning tasks, making it a top choice among data scientists and practitioners. XGBoost's strength lies in its ability to combine the predictive power of multiple weak learners (typically decision trees) into a strong ensemble model. It does this through a process known as boosting, which sequentially trains new models to correct the errors made by the previous ones.

V. DATA PREDICTION ANALYSIS

We experimented our dataset by training and feature extraction. Price prediction is based on the extracted model generated by the algorithmic expression on different volumes of data streams. Datasets are used to validate and ensure the quality of data. They help identify inconsistencies, errors, and outliers, allowing for data cleaning and enhancement. In summary, our proposed workflow is experimented as shown below.

TABLE 1. Trades of Nifty Metal

Open	High	Low	Close	Volume
2877.75	2877.75	2798.5	2804.7	44,481,268
2782.1	2808.3	2726.6	2745.3	52,547,557
2789	2835.4	2726.6	2760.45	44,854,372
3050.75	2998.3	2943.4	2987.9	44,775,685

TABLE 2. Trades of NIFTY_FIN_SERVICE

Open	High	Low	Close	Volume
5239.4	5378.6	5224.6	5368.65	44253381
5371.85	5419.35	5353.7	5386.3	52087747
5340.95	5367.25	5313	5345.95	38253261
5338.35	5338.35	5213.3	5226.85	33273714

NIFTY METAL is influenced by various factors, including economic conditions, industry-specific news, and global market trends. Investors and traders closely monitor these price movements and trading volumes to make informed decisions about buying or selling shares within this index. the table encapsulates essential data points that are crucial for analysing the performance and trends of the NIFTY METAL index during the specified time frame, aiding market participants in making informed investment decisions. NIFTY referred as the Nifty Financial Services Index, is a prominent sectoral index within the National Stock Exchange of India (NSE). This index is designed to track the performance of companies operating in the financial services sector. The financial services sector includes a wide range of businesses, such as banks, non-banking financial companies (NBFCs), insurance companies, asset management firms, and other financial institutions. The companies included in this index are subject to periodic reviews and may change over time to ensure that it accurately represents the sector's composition. These reviews are carried out by the index provider to maintain the index's relevance and accuracy.

TABLE 3. Trades of NIFTY_IT

Open	High	Low	Close	Volume
4352.2	4397.35	4320.2	4388	10420573
4402.7	4430.65	4370.5	4386.95	13183048
4392.95	4407.15	4335.8	4351.95	12869250
4336.1	4385.1	4307.1	4376.15	11253911

ML models rely on features or input variables to make predictions. NIFTY IT data, including historical price trends, trading volumes, and sector-specific news, can be used for feature engineering, enhancing the predictive capabilities of models. Investors and traders use NIFTY IT to diversify their portfolios. ML models can help identify optimal diversification strategies by considering correlations and historical performance of NIFTY IT alongside individual stocks. stock prediction often involves time-series data, and NIFTY IT's historical data can be used for time-series analysis. Time-series analysis is crucial in stock prediction, as stock prices follow a time-dependent sequence. Techniques like Autoregressive Integrated Moving Average graph models are applied to analyse and forecast time-series data. ML models are used to predict future movements of the index based on historical price data, trading volumes, and other relevant indicators. These models employ techniques such as time series analysis and regression to make short-term and long-term predictions. Quantitative traders employ ML algorithms to make high-frequency trading. We used many algorithms like Svm, Linear Regression, Random Forest regression, K Nearest Neighbours (KNN), Decision Trees, LSTM model, SVR (Linear).

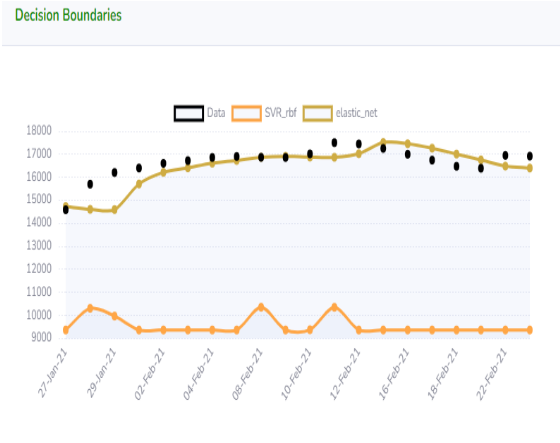
VI. EXPERIMENT RESULTS

This section shows the results of various plot values based on the algorithm used to predict the stock price. Machine Learning algorithms Plays a major role in classification and feature extraction to analyse the past behaviour and preferences as a model. Accuracy comparison has been tested by implementing the algorithms on data values to verify nifty based on the trained model using machine learning algorithms Random Forest regression, K nearest neighbours and etc.

a. Model comparison Linear Regression and Random Forest algorithm

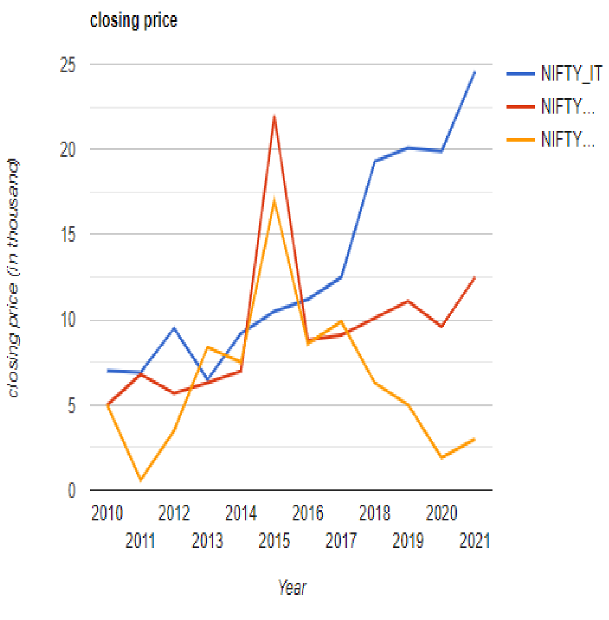
b. Model comparison K Nearest Neighbour and Decision Tree

c. Model comparison between Linear Regression and Decision Tree Algorithm.



d. Model comparison between Svr and LSTM algorithm.

The Difference in the performance were predicted from the above graphs using various ML algorithms on stock price prediction. Dataset collection involves various process from the live stock community to extract feature data to classify and training the data values for a long-term process. The result of the stock and goods were classified based on the cluster using K Nearest neighbour algorithm and predicts stock values without involving external factors to affect the priority of plot values. Here is the experimented stock prediction result based on the ml algorithms.



e. Predicted price on various Indices.

VII. CONCLUSION AND FUTURE WORK

In conclusion, the presented stock price prediction system represents a valuable tool for investors and financial analysts. The system leverages a user-friendly web interface built on the Flask framework, providing users with the capability to select from multiple machine learning algorithms, visualize predictions, and assess model performance. The integration of machine learning models, including Decision Trees, facilitates more data-driven investment decisions. The system's ability to adapt to custom datasets further enhances its versatility, accommodating diverse data sources and financial contexts. The comparative analysis of different machine learning models enables users to make informed choices, aligning their strategies with models that demonstrate superior predictive capabilities. The practical implications of the system are evident in its support for investors and financial analysts in portfolio management, risk assessment, and investment decision-making. It bridges the gap between technology and finance, showcasing the potential of machine learning in the financial domain. Will Implement most advanced algorithm to predict real time datasets by extracting values from stock community.

REFERENCES

[1] J. J. Murphy, Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications. Penguin, 1999.

[2] T. Turner, A Beginner's Guide To Day Trading Online, 2nd ed. New York, NY, USA: Simon and Schuster, 2007.

[3] H. Maqsood, I. Mehmood, M. Maqsood, M. Yasir, S. Afzal, F. Aadil, M. M. Selim, and K. Muhammad, ``A local and global event sentiment based efficient stock exchange forecasting using deep learning," Int. J. Inf. Manage., vol. 50, pp. 432451, Feb. 2020.

[4] W. Long, Z. Lu, and L. Cui, ``Deep learning-based feature engineering for stock price movement prediction," Knowl.-Based Syst., vol. 164, pp. 163173, Jan. 2019.

[5] J. B. Duarte Duarte, L. H. Talero Sarmiento, and K. J. Sierra Juárez, ``Evaluation of the effect of investor psychology on an artificial stock market through its degree of efciency," Contaduría y Administración, vol. 62, no. 4, pp. 13611376, Oct. 2017.

[6] Lu, Ning, A Machine Learning Approach to Automated Trading. Boston, MA, USA: Boston College Computer Science Senior, 2016.

[7] M. R. Hassan, B. Nath, and M. Kirley, ``A fusion model of HMM, ANN and GA for stock market forecasting," Expert Syst. Appl., vol. 33, no. 1, pp. 171180, Jul. 2007.

[8] 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. 25132522, Oct. 2005.

[9] J. Sun and H. Li, ``Financial distress prediction using support vector machines: Ensemble vs. Individual," Appl. Soft Comput., vol. 12, no. 8, pp. 22542265, Aug. 2012.

[10] P. Ou and H. Wang, ``Prediction of stock market index movement by ten data mining techniques," Modern Appl. Sci., vol. 3, no. 12, pp. 2842, Nov. 2009.

[11] F. Liu and J.Wang, ``Fluctuation prediction of stock market index by legendre neural network with random time strength function," Neurocomputing, vol. 83, pp. 1221, Apr. 2012.

[12] C.-F. Tsai, Y.-C. Lin, D. C. Yen, and Y.-M. Chen, ``Predicting stock returns by classier ensembles," Appl. Soft Comput., vol. 11, no. 2, pp. 24522459, Mar. 2011.

[13] R. D. A. Araäjo and T. A. E. Ferreira, ``A Morphological-Rank-Linear evolutionary method for stock market prediction," Inf. Sci., vol. 237, pp. 317, Jul. 2013.

[14] M. Ballings, D. Van den Poel, N. Hespeels, and R. Gryp, ``Evaluating multiple classiers for stock price direction prediction," Expert Syst. Appl., vol. 42, no. 20, pp. 70467056, Nov. 2015.

[15] S. Basak, S. Kar, S. Saha, L. Khaidem, and S. R. Dey, ``Predicting the direction of stock market prices using tree-based classiers," North Amer. J. Econ. Finance, vol. 47, pp. 552567, Jan. 2019.

[16] B. Weng, W. Martinez, Y.-T. Tsai, C. Li, L. Lu, J. R. Barth, and F. M. Megahed, ``Macroeconomic indicators alone can predict the monthly closing price of major U.S. indices: Insights from artificial intelligence, time-series analysis and hybrid models," Appl. Soft Comput., vol. 71, pp. 685697, Oct. 2018.

[17] J. Long, Z. Chen, W. He, T. Wu, and J. Ren, ``An integrated framework of deep learning and knowledge graph for prediction of stock price trend: An application in chinese stock exchange market," Appl. Soft Comput., vol. 91, Jun. 2020, Art. no. 106205.

[18] G. Rekha, D. Bhanu Sravanthi, S. Ramasubbareddy, and K. Govinda, ``Prediction of stock market using neural network strategies," J. Comput. Theor. Nanoscience, vol. 16, no. 5, pp. 23332336, May 2019.

[19] X. Pang, Y. Zhou, P. Wang, W. Lin, and V. Chang, ``An innovative neural network approach for stock market prediction," J. Supercomput., vol. 76, no. 3, pp. 20982118, Mar. 2020.

[20] Kelotra, A. and P. Pandey, ``Stock market prediction using optimized deep-convLSTM model," Big Data, vol. 8, no. 1, pp. 524, 2020.

[21] Y. Baek and H. Y. Kim, ``ModAugNet: A new forecasting framework for stock market index value with an overtting prevention LSTM module and a prediction LSTM module," Expert Syst. Appl., vol. 113, pp. 457480, Dec. 2018.

[22] H. Chung and K.-S. Shin, ``Genetic algorithm-optimized long short-term memory network for stock market prediction," Sustainability, vol. 10, no. 10, p. 3765, 2018.

[23] M. Nabipour, P. Nayyeri, H. Jabani, A. Mosavi, and E. Salwana, ``Deep learning for Stock Market Prediction," Entropy, vol. 22, no. 8, p. 840, Aug. 2020.

[24] Y. Kara, M. Acar Boyacioglu, and Ö. K. Baykan, ``Predicting direction of stock price index movement using articial neural networks and support vector machines: The sample of the Istanbul stock exchange," Expert Syst. Appl., vol. 38, no. 5, pp. 53115319, May 2011.

[25] J. Patel, S. Shah, P. Thakkar, and K. Kotecha, ``Predicting stock market index using fusion of machine learning techniques," Expert Syst. Appl., vol. 42, no. 4, pp. 21622172, Mar. 2015.

[26] J. Patel, S. Shah, P. Thakkar, and K. Kotecha, ``Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques," Expert Syst. Appl., vol. 42, no. 1, pp. 259268, Jan. 2015.

[27] R. Majhi, G. Panda, B. Majhi, and G. Sahoo, ``Efcient prediction of stock market indices using adaptive bacterial foraging optimization (ABFO) and BFO based techniques," Expert Syst. Appl., vol. 36, no. 6, pp. 1009710104, Aug. 2009.

[28] Y. Chen and Y. Hao, ``A feature weighted support vector machine and Knearest neighbor algorithm for stock market indices prediction," Expert Syst. Appl., vol. 80, pp. 340355, Sep. 2017.