PREDICTING STOCK TRENDS WITH LSTM-CNN

T.Nagarjuna Reddy[1], K.Likhith Reddy[2], A.Yashwanth Reddy[3], K.Praveen[4]

[1]Assistant Professor, [2,3,4]Students

[1,2,3,4]Department of Artificial Intelligence and Machine Learning

Malla Reddy Institute of Technology and Science, Hyderabad, India.

Email Id: nag5588@gmail.com, likhith1718@gmail.com, Anamyashwanth4@gmail.com,katrothpraveen404@gmail.com

ABSTRACT

The stock market is a complex and dynamic system, affected by numerous factors includingeconomic indicators, political events, and investor sentiment.Traditional time series analysistechniqueshavebeenusedfordecades,butrecentadvancesindeeplearninghaveshownpromisingresultsformodelingandpredictingcomplextimeseriesdata.Inthisstudy,weuselongshort-termmemory(LSTM)neuralnetworksandconvolutionalneuralnetworks(CNN)toanalyzeandpredictstock market data. We compare the performance of these models against traditional time seriesmodels such as ARIMA and exponential smoothing. Our results show that deep learning modelscan significantly outperform traditional models for stock market prediction, with LSTM modelsachieving the best performance. We also show that the inclusion of additional data sources such asnews articles and social media sentiment can further improve prediction accuracy. Our findingssuggest that deep learning techniques can be a valuable tool for investors and analysts seeking tomake informed decisions based on stock market data. However, we caution that the complexity andvolatility of the stock market make accurate prediction a challenging task, and recommend usingthesemodelsasonepartofacomprehensiveinvestmentstrategy.

Inourprojectwewillbeusingthealgorithmssuchas DecisionTree(DT) and ArtificialNeuralNetwork(ANN)andforanalysingourresults.

I. INTRODUCTION

Predicting stock prices and making investment decisions is a complex and challenging task in the world of finance. The stock market is influenced by a multitude of factors, including economic indicators, company performance, market sentiment, and geopolitical events. Machine learning, a subset of artificial intelligence, has gained significant attention for its potential to enhance stock price prediction and assist in making informed investment decisions, especially in the context of multi-stock prediction or portfolio management [1].

Multi-stock prediction using machine learning involves the development and deployment of computational models to forecast the future prices or movements of multiple stocks within a portfolio. This approach offers several advantages, including the ability to process vast amounts of historical and real-time data, identify hidden patterns and correlations, and provide insights that can aid investors and portfolio managers [1].

In this field, machine learning models are used to analyze various data sources, such as historical stock prices, trading volumes, news sentiment, financial reports, and macroeconomic indicators [2]. By applying algorithms, these models aim to discover predictive relationships and trends, which can help optimize portfolio performance, reduce risk, and potentially generate higher returns.

While multi-stock prediction with machine learning has the potential to provide valuable insights, it's important to acknowledge the inherent risks and uncertainties associated with financial markets. Past performance does not guarantee future results, and machine learning models are not foolproof [3]. Moreover, financial markets can be influenced by unpredictable events and sentiment-driven shifts, which may challenge the accuracy of predictions.

Time series analysis is a critical tool for understanding and predicting complex temporal data. Inthe financial domain, time series analysis is particularly important for predicting the behavior ofstock markets, which are influenced by a multitude of factors including economic indicators,political events, and investor sentiment.Accurate predictions of stock market trends can havesignificant financial implications, making it an area of intense interest for investors and analysts.Traditional time series analysis techniques, such as Arima and exponential smoothing, have beenused for decades to analyze stock market data.However, these models have limitations in modelingthecomplexdynamicsofstockmarkets,particularlyinthepresenceofnon-linearitiesandnon-stationarities.Deeplearningtechniqueshaveemergedasapromisingalternativeformodelingand predicting time series data in recent years deep learning techniques have emerged as apromisingalternativeformodelingandpredictingtimeseriesdata.Deeplearningtechniques,suchas long short-term memory (LSTM) neural networks and convolutional neural networks

(CNN),canlearncomplextemporalpatternsindataandcapturenon-linearrelationshipsbetweenvariables[16][17][18].Thesemodelshavebeensuccessfullyappliedtoavarietyoftimeseries predictiontasks,includingstockmarketprediction.

In this paper, we explore the use of deep learning techniques for time series analysis andprediction of stock market data.We compare the performance of deep learning models againsttraditional time series models and also investigate the impact of incorporating additional datasourcessuchasnewsarticlesandsocialmediasentimentonpredictionaccuracy. The rest of the paper is organized as follows: in the next section, we review the relevant literature on time series analysis and deep learning for stock market prediction. In the following section, we describe our methodology and experimental setup. We then present and analyze our results, and discuss the implications of our findings. Finally, we conclude the paper with a summary of our contributions and suggestions for future research.

Disadvantages of previous existing model

ARIMA models can be limited and challenging when forecasting time series. They are linear models, so they cannot handle nonlinear relationships or complex dynamics, such as sudden shocks or regime changes. Additionally, they are parametric models and rely on assumptions about the data, such as normality or homoscedasticity. Outliers and missing values can affect the model estimation and forecasting performance, so the data may need to be preprocessed before applying ARIMA models. Moreover, they are not suitable for very short or very long time series as they may not have enough information or become unstable over time.

II. LITERATURE SURVEY

Research in finance has explored how stock markets are affected by their multi- source and heterogeneous data on some scales. Multi-source heterogeneous detain the stock market means that the data of the stock market includes data from different sources such as the stock market, the foreign exchange market and even the weather system, as well as the structure of stock prices, trading volumes, and stock news, announcements and social networks. and other unstructured data. In particular, the efficient market hypothesis believes that information from various sources in the stock market will have an impact on the stock market, while behavioural finance believes that financial markets are explained, studied and predicted from the individual behaviours of traders and the motivations that produce such behaviors the trend and extent of price fluctuations. These studies point out that the internal mechanism of the stock market is very complex, similar to Brownian motion. Combining the multi-source heterogeneous data in the stock market can more accurately classify and predict the stock market state. With the vigorous development 7 of the stock market, it continues to generate a large number of multi-source heterogeneous data of various scales. The traditional idea of relying solely on experts to analyze and predict has been difficult to meet the needs of industry development [4]. In order to quickly analyze massive stock market data and assist or even completely replace investors in making stock market investment decisions, a large number of researches on stock market forecasting based on information technology have emerged. These studies have also contributed to the rapid development of quantitative funds that rely on automated computer analysis to execute and even make investment decisions entirely on their own [5]. Obtaining accurate stock price forecasts can more effectively avoid future risks for decision makers; for regulators, it can strengthen the control of the stock market, regulate and guide the stock market in a timely manner, and contribute to the sustainable development of the economy.

Progress of stock price prediction

The research on stock behaviour was first conducted by Bachelier in 1900. Heused random walks to express stock price trends. Fama tested that stock price changesare characterized by random walks. Malkiel and Fama studied valid market assumptions in 1970 and found that all new information will be reflected in asset prices immediately without delay [19]. Therefore, changes in future asset prices havenothing to do with past and present information. From their perspective, predictingfuture asset prices is considered impossible. On the other hand, many studies trytoprove effective market hypotheses experimentally, and empirical evidence shows that the stock market can be predictable in some ways. In traditional time series models, 9 parameter statistical models are used for forecasting, such as ARMA model, ARIMAmodel and vector autoregressive model, etc., to find the best estimate. VirtanenandYliolli used six explanatory variables to estimate the Finnish stock market index, including the lagging index and macroeconomic factors in an econometric model based on ARIMA. Work proposed a stock priceprediction system based on ARIMA in 2014, which has been tested in the listed stocksoriginated from the Stock Exchange in New York and the Stock Exchange runningfrom the country Nigeria[3]. Then the ARIMA model is regarded as a high potential model for forecasting short-term series[7].The textual data from thenewspaper at Nikkei as the input of the LSTM neural network, and combined with thetime series data in stock market to predict the opening price of 10 selected companies[1].

Time series model

- **Stationary time series:**
  Stationary time series are divided into strictly stationary time series and wide stationary time series. Below we introduce their definitions. Strictly stationary timeseries provide important theoretical significance, but it is difficult to obtain the joint distribution of random sequences in the actual research process. Therefore, in order to better use in practical applications, researchers have defined a relatively weak wide stationary time sequence. Researchers choose to use the characteristic statistics of the sequence to define wide stationarity, which can make the constraint conditions a little looser. By ensuring the stationarity of the low-order moments of the sequence to ensure that the sequence can be approximately stationary. Time series analysis also belongs to the field of statistics. It can also analyse the population through samples like statistics. And from the statistical theorems, we can know that the number of random variables is directly proportional to the complexity of the analysis, and the sample size is inversely proportional to the accuracy of obtaining the overall information (obviously the sample information obtained when the population is selected as the sample is Overall information, but such an operation is obviously unrealistic). But time series data has its peculiarities. For a timeseries $\{\cdots, X1, X2, \cdot\cdot\cdot, Xt, \cdot\cdot\cdot\}$, its value Xt at any time t is a random variable, and since time is one-way, it cannot be repeated, So we can only get one sample value in this way, which leads to too little sample information for statistical analysis. But if we have the concept of stationarity, this problem will be solved.
- **Principles of the ARIMA model:**
  Autoregressive moving average model (ARIMA model) is the most commonly 13 used time series model. According to different conditions, it can be divided into the following three types of models: autoregressive (AR) model, moving average (MA) model and ARIMA model. Most of the data in real life is not stable. We need to smooth the data. Box and Jenkins proved that the difference method is an effective smoothing method. Therefore, applying the difference method to the ARIMA model will result in the well- known ARIMA model. The following article will introduce the general modelling steps of the ARIMA model.

III. MATERIALS AND MODELS

Long Term Short Memory Network

LSTM was designed by Hochreiter and Schmidhuber.

LSTM can be implemented in python using the Keras Library.

LSTM has 3 gates

Input gate

Forget gate

Output gate

LTSMs are a type of Recurrent Neural Network for learning long-term dependencies. It is commonly used for processing and predicting time-series data.

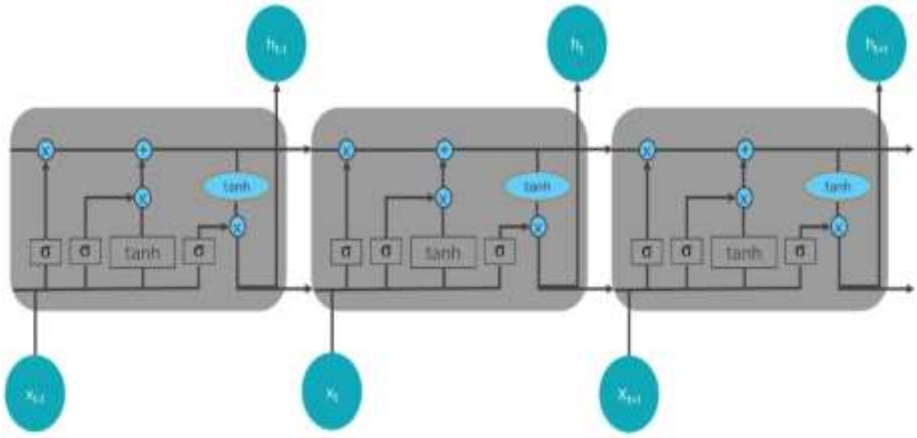

Figure 1:LSTM network

From the image on the top, you can see LSTMs have a chain-like structure. General RNNs have a

single neural network layer. LSTMs, on the other hand, have four interacting layers communicating extraordinarily.

LSTM process

In traditional neural networks, neurons in the same hidden layer are not connected to each other, and this structural defect directly leads to their poor performance in dealing with certain problems. This shortcoming becomes especially acute when dealing with time series and speech recognition problems where information is contextualized. The emergence of the recurrent neural network solves this problem very well. The neurons in the same hidden layer are connected to each other, which can effectively obtain the contextual information of the data. The output of the recurrent neural network is determined according to the input and the previous related information, so it can play its short-term memory when dealing with timeseries problems. Although the effect of recurrent neural network in dealing with time series problems is very good, there are still some problems. The more serious one is that gradient disappears or explodes easily in the processing of long-term span problems. Causes the phenomenon of small memory value. After the cyclic neural network is expanded, it can be regarded as a multi-layer feedforward neural network with each layer sharing the same weight parameters. Although it keeps trying to learn the long term dependencies of sequences, actual research finds this to be a difficult task indeed. Long-term reliance on signals tends to become very weak and highly susceptible to short-term signal fluctuations. There is a multiplier of the derivative of the activation function in time-based backpropagation, and the continuous accumulation will cause uncontrollable problems. Although it can be solved theoretically by adjusting the parameters, it is found that this problem is difficult to solve in practice, so it still needs to be optimized from the structure. This leads to its improved structure - the LSTM neural network.
Next, we can start the construction of the LSTM neural network. The first is thedetermination of several parameters. After n-fold cross-validation, we choose thehidden layer to have 10 neurons. The number of iterations is selected 50 times, andeach 72sample data is formed into a batch for training, that is, batchsize = 72, Adamalgorithm is used as the optimizer of the model, the learning rate is 0.001, andthetraining set data is randomly scrambled. Use the MSE indicator as the loss functionof the model for training. In the below figure it is a training diagram of the neural network. It canbe seen from this diagram that after iteration, the loss function of the model decreasesquickly and tends to converge. It can be seen that the prediction model is morereasonable.

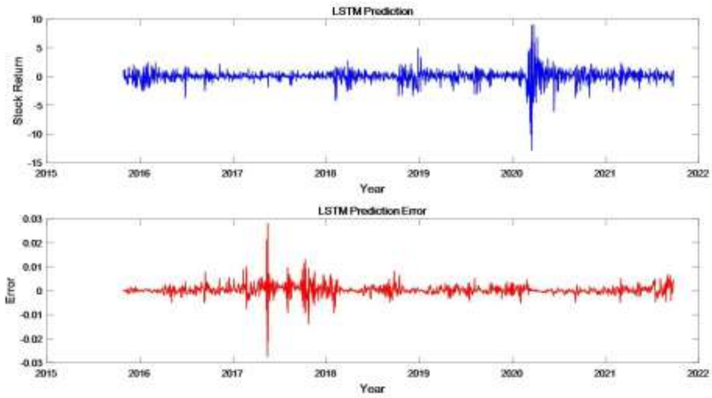

Figure 2: LSTM result prediction

Convolutional Neural Networks

Convolutional neural networks (convolutional neural network, CNN) are widelyused in image processing related tasks (such as image classification, target detection, object recognition, etc.), has also been applied to natural language speech processingand speech processing tasks [5]. The fully connected network requires correspondingparameters for each dimension of data. For image tasks, using a fully connectednetwork will cause a lot of parameters and a huge model, which is not conducivetotraining and deployment use. The convolutional neural network uses a smaller tensor as a parameter (called the convolution kernel) in the input. The input height and width dimensions are sliding processing, and the input at different positions shares thisparameter. This method is used to save province model parameters. Convolutional neural networks include convolution operations, nonlinear transformation and poolingoperations. Processing the image information is an example to illustrate thecalculation process of the convolution operation[11]. For the input picture I ∈RH ×W×C, where I is the picture, H is the height of the film, W is the width of the picture, Cisthe feature number of the picture, and its three primary colours (R, G, B) are generallyused. The colour value of as its characteristic, that is, C = 3, the whole picture is athree-dimensional tensor. Parameters of convolution operation, that is, the convolutionkernel is g ∈ R k×k×C×Count, where k is the size of the convolution kernel and Cout is thenumber of output features, also known as number is a four-dimensional tensor. Then, the convolution operation is

$$(I * g)(i, j)=\sum\sum I (i+ m, j + n)\ g (m +(k), n +(k))\ m=- k\ m=- k$$

The size of the convolution kernel and the number of output features need tobedesigned by the network designer, and there is also a step size (stride), voidrate 17 (dilation), filling method (padding) and other parameters can be designed/selected. The convolution kernel size is the size of the area that can be sensed bytheconvolution operation. When k = H, the convolution kernel sees the entire picture. It degenerates into a fully connected network. The step size indicates that theconvolution kernel is slipping. The step length of each sliding in the dynamiccalculation process. Filling means adding specific elements around the output imageto control the size of the output.

ToolsAndTechnologies

1. Python: Python is a popular programming language for data analysis and machine learning.YoucanusePythontoimplementyourstockmarketpredictionmodels.

2. Jupyter Notebook: Jupyter Notebook is an interactive environment that allows you to writeand execute Python code in a notebook-style interface.It’s great for experimentation andvisualization.

3. Pandas: Pandas is a Python library for data manipulation and analysis.You can use it tohandleandpreprocessyourstockmarketdata.

4. Numpy: Numpy is another Python library for numerical computing. It’s useful for

performingmathematicaloperationsonyourdata.

5. Matplotlib and Seaborn: These libraries can help you create visualizations to better understandyourdataandtheresultsofyourmodels.
6. Scikit-Learn: Scikit-Learn is a Python library that provides tools for machine learning anddatamining.Itincludesvariousalgorithmsforregressionandclassificationthatcanbeusedforstockpredictio n.
7. TensorFloworPyTorch:Ifyouwanttodiveintodeeplearningforstockprediction,youcanuseTensorFloworPyTorch, whicharepopulardeeplearningframeworks.
8. Finance Datasets: Look for publicly available financial datasets that you can use for yourproject.These datasets often include historical stock prices, company fundamentals, andeconomicindicators.

Block Diagram

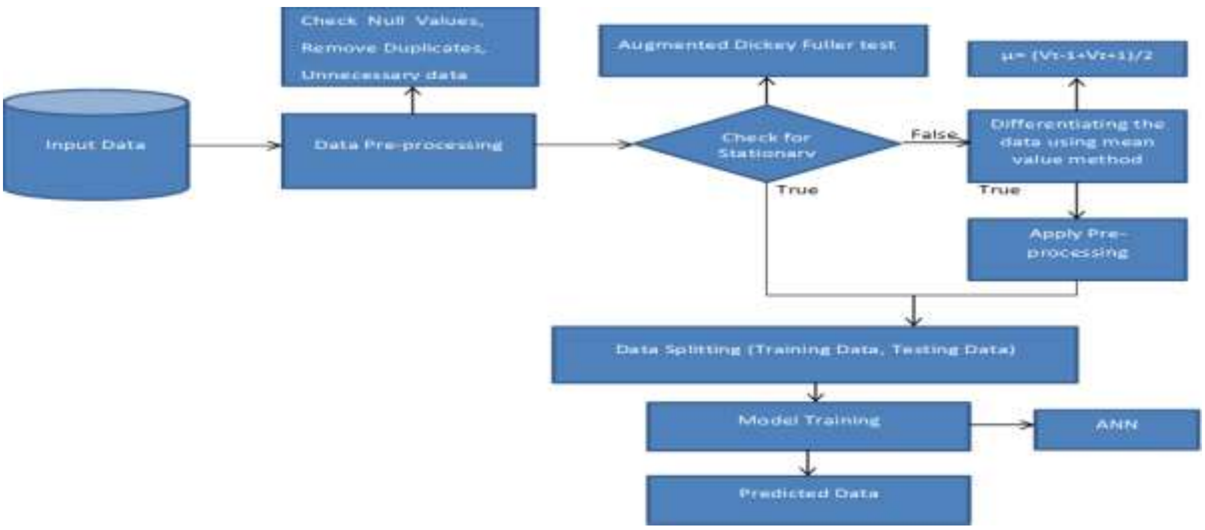

Figure 3: Block diagram

StepsOfImplementation

Inputdataset:

ThedatasetcanbetakenfromtheofficialNSEandESTorganizationsthatareopen-sourcedanddonot require any license. We have collected a set of stock market datasets which we are going toanalyze.Thenfortrainingthedataset,thecomparisonofthenon-stockdatasetsisbeentaken.

Analysisofdataset:

Heretheanalysisofthedatasettakesplace.Thesizeofthedataistakenintoconsiderationforthedataprocess.

Oversampling:

wehavecreatedadetailedhistoryofallstocksthathavebeendoneoveracertainamountoftimeanditissampledtofixathresholdva lue.

TrainingandTestingSubset:

As the dataset is imbalanced, many classifiers show bias for majority classes. The features of theminorityclassaretreatedasnoiseandareignored.Henceitisproposedtoselectasampledataset

IV. RESULT AND DISCUSSION

Predicting Stock Trends with Multi-Factor CNN-LSTM and Attention Mechanism is a title of a research paper that proposes a novel deep learning model for stock price prediction. The model uses three types of data: historical stock prices, technical indicators, and news articles. The model consists of three components: a convolutional neural network (CNN) for extracting features from stock prices, a long short-term memory (LSTM) network for capturing temporal dependencies, and an attention mechanism for weighting the importance of different news articles.

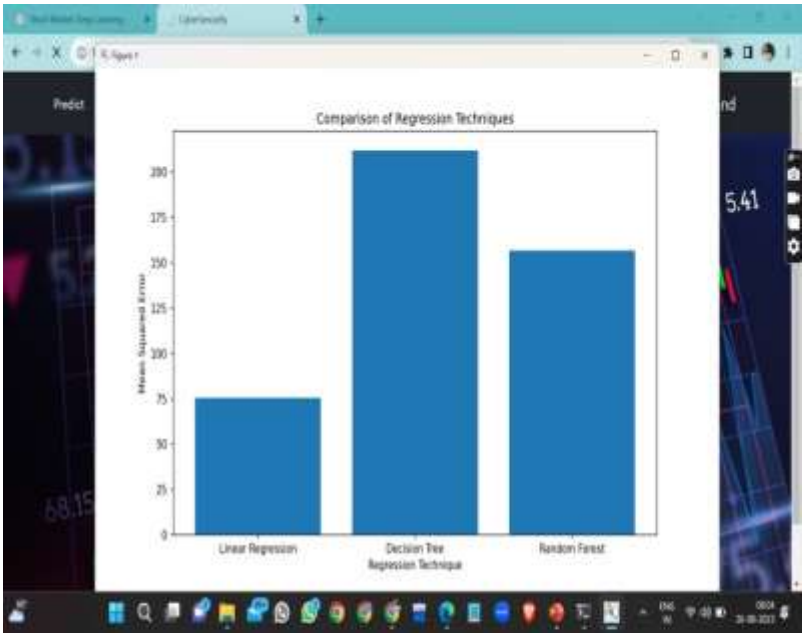

Figure 4: Accuracy analysis

V. CONCLUSION AND FUTURE SCOPE

In report, we will compare a machine learning models like LSTM model, the CNN model and also the hybrid approach of LSTM + CNN model. We have a tendency to train the model using the data of NSE listed companies to predict the stock future value. This shows the proposed method is capable to distinctive around interrelation with the data. Also, it is evident from the results that, Hybrid approach of LSTM+CNN model is capable to identify the changes in trends. For the projected method the Hybrid approach of LSTM+CNN is known as the best model. It uses the information given at a specific instant for prediction. Even if the other two models LSTM and CNN are utilized in a lot of other time-dependent data analysis, it is not outperforming over the Hybrid approach of LSTM+CNN architecture in this case. This is often because of quick changes occur in stock market. The changes in the stock market is not always be in a regular pattern or not always follow the continuous cycle. Based on the companies and sectors, the existence of the trends and the period of their existence will differ. The analysis of this type of cycles and trends can offer a more profit to the investors. In future work, we add more stock market data and compare more model to improve accuracy of predicted stock price. In the future, for better accuracy model can be trained with more varied and detailed data. Also, other algorithms along with proposed can be used to create a new hybrid model.

VI. ACKNOWLEDGEMENT

The team members of the research project want to sincerelythank our guide Assistant Professor T.Nagarjuna Reddy and the Department of Computing Science and Engineering, Malla Reddy Institute of Technology andScience, India for their encouragement and support for completion of this work.

VII. REFERENCES

[1]. Jonathan Frankle and Michael Carbin –"The Lottery Ticket Hypothesis Finding Sparse,Trainable Neural Networks", 2018, International Conference on Learning Representations(ICLR)

[2]. RickyT.Q.Chen–"NeuralOrdinaryDifferentialEquations",2018,ConferenceonNeuralInformationProcessingSystems(NeurIPS)

[3]. Clark, T. E., & West, K. D. (2007). Approximately normal tests for equal predictive

[4]. accuracy in nested models. Journal of Econometrics, 138, 291–311

[5]. Haven, E., Liu, X., & Shen, L. (2012). De-noising option prices with the wavelet

[6]. method. European Journal of Operational Research, 222(1).

[7]. Chen, J., Jiang, F., & Tong, G. (2017). Economic policy uncertainty in China and

[8]. stock market expected returns. Accounting and Finance, 57, 1265–1286.

[9]. Cochrane, J. H. (2007). The dog that did not bark: A defence of returns predictability. Review of Financial Studies, 21, 1533–1575.

[10]. Conrad, J., & Kaul, G. (1998). An anatomy of trading strategies. Review of FinancialStudies, 11, 489–515.

[11]. Cowles, A., 3rd (1933). Can stock market forecasters forecast? Econometrica. Journalof the Econometric Society, 309–324.

[12]. Dai, Z., Zhou, H., Wen, F., & He, S. (2020a). Efficient predictability of stock returnvolatility: The role of stock market implied volatility. The North American Journal ofEconomics and Finance, 52, 101174.

[13]. Dai, Z., & Zhu, H. (2020). Stock returns predictability from a mixed modelperspective. Pacific-Basin Finance Journal, 60, 101267.

[14]. Dai, Z. F., Dong, X. D., Kang, J., & Hong, L. (2020b). Forecasting stock marketreturns: New Technical indicators and two-step economic constraint method. TheNorth American Journal of Economics and Finance, 53, 101216.

[15]. HadovanHasselt–"DeepReinforcementLearningwithDoubleQ-learning",2016,ConferenceonNeuralInformationProcessingSystems(NeurIPS)

[16]. AlexKrizhevsky,IlyaSutskever,andGeoffreyHinton–"ImageNetClassificationwithDeepConvolutional Neural Networks", 2012, 25th International Conference on Neural InformationProcessingSystems(NIPS)

[17]. Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, SherjilOzair, Aaron Courville, and Yoshua Bengio –"Generative Adversarial Networks", 2014,NeuralInformationProcessingSystems(NIPS)

[18]. Wojciech Zaremba,Ilya Sutskever,and Oriol Vinyals–"Recurrent Neural NetworkRegularization",2014,32ndInternationalConferenceonMachineLearning(ICML)

[19]. Yann LeCun, Yoshua Bengio, and Geoffrey Hinton –"Deep Learning", 2015, published in thejournalNatureinMay2015

[20]. Volodymyr Mnih–"Playing Atari with Deep Reinforcement Learning", 2013, conferenceproceedingsoftheNeuralInformationProcessingSystems(NIPS)conference

[21]. Volodymyr Mnih–"Deep Reinforcement Learning", 2015, International Conference onMachineLearning(ICML)

[22]. XiaoDing–"DeepLearningforStockPredictionusingNumericalandTextualInformation",2015,Twenty-NinthAAAIConferenceonArtificialIntelligence.