Gradient Boosting and LSTM Based Hybrid Ensemble Learning for Two Step Prediction of Stock Market

Pratyush Ranjan Mohapatra [1], Ajaya Kumar Parida [1,*], Santosh Kumar Swain [1], and Santi Swarup Basa [2]

[1] School of Computer Engineering, Kalinga Institute of Industrial Technology (Deemed to be University), Bhubaneswar, Odisha, India; Email: pratyush.mohapatra@gmail.com (P.R.M.), sswainfcs@kiit.ac.in (S.K.S.)
[2] Department of Computer Science, Maharaja Sriram Chandra Bhanjadeo University, Baripada, Odisha, India; Email: santiswarup.basa@gmail.com (S.S.B.)
*Correspondence: ajaya.paridafcs@kiit.ac.in (A.K.P.)

***Abstract*—Prediction of stock market price using different artificial intelligent techniques have become an efficient and effective method for stock market prediction with higher prediction accuracy. In this present work, thus we provide an ensemble technique that comprises of two base models namely extreme gradient boosting method and long short term memory method for short term prediction of stock market. Previously the prediction of stock price was confined to all the data available, irrespective of its significance in prediction accuracy. This study investigates different issues for predicting the closing price of the stock market. Based on the two step ensemble method (including a feature selection and combination of two different intelligent techniques). Convolutional Neural Network (CNN) method is used for feature selection purpose based on the correlation coefficient of different technical indicators for predicting the closing price. Additionally, ensemble learning is applied for increasing the prediction accuracy. The subset of selected input features enhances the model's accuracy. The performance evaluation of the proposed model is performed by comparing it with different other models like Support Vector Machine (SVM), Long Short Term Memory (LSTM), Kernel Extreme Learning Machine (KELM), etc. As a new addition to the previous literature the proposed combined method extracts the features that mainly influences the accuracy of the predicted price hence better result in less time is observed. The proposed ensemble learning technique exhibited the best predicted output as compared with other methods discussed in this study.**

***Keywords*—artificial intelligence, technical indicators, ensemble learning, Extreme Gradient Boosting (XGB)**

I. Introduction

The stock market is expressed as one of the most challenging real-world problems, it depends upon various factors. Main aim of stock price prediction is to recommend certain countermeasures to reduce the risk management of the stock market. In order to perform the prediction, the market needs a thorough hypothesis [1–3]. The stock markets movement is affected by various factors making it more complicated for analysing purpose [4, 5]. The universal data of the stock price makes the analysis difficult when different markets are considered. In general, two methods are implemented for stock market prediction, technical analysis and fundamental analysis. Technical analysis is mainly based on the statistical analysis, this includes the historical movement of the prices. Different technical indicators may include moving average, stochastic %K, etc. Whereas the fundamental analysis mainly focuses on different economic aspects by the company and their political conditions. Irrespective of different existing model stock price analysis is still a challenging job and have scope for improvement. To overcome these challenges in stock market prediction, the modern researchers have introduced different artificial intelligent techniques for prediction and analysis purposes. Various techniques primarily used for this purpose, traditional approaches extensively used previously are Auto Regressive Moving Average (ARMA), Auto Regressive Integrated Moving Average (ARIMA), etc. [6–8]. The main drawback of these models is, that it is not capable of handling nonlinear data, the computational complexity is more, and the problem of local minima therefore slow convergence rate. These limitations led to the introduction of other new models for more accurate predictions.

The main contribution of this paper can be given as a new contribution to the previously present study is that this paper focuses on Convolutional Neural Network (CNN) based feature selection, where different technical indicators (like Simple moving average, Momentum stochastic %K, Accumulator/Distributor, historical closing price and Open-High-Low-Close (OHLC)) are considered as the features and the best features are selected based on the Correlation coefficient with the closing price. This paper shows a one day ahead prediction of stock prices. Conclusively the CNN algorithm is hybridized with the XGB (Extreme Gradient Boosting) and Long Short Term Memory (LSTM) algorithm for more accurate prediction of stock market price. The closing price is not only

Manuscript received March 2, 2023; revised March 21, 2023; accepted May 30, 2023; published November 22, 2023.

dependent on the previous closing price value but also depends upon various indicators and emotional status of the market. Hence a very accurate prediction technique is required. The proposed model gives an accurate prediction by considering the important features only that mostly influences the next day closing price. The experimental data analysis shows better performance of the proposed model.

The remaining paper is organized as following: Section II provides a brief description of different methods used in this study for prediction purpose, Section III demonstrates the proposed methodology. Section IV shows the experimental verification and result analysis showing different comparative study between different prediction techniques, and finally Section V concludes the study.

II. Literature Review

The new research includes Back Propagation (BP), Artificial Neural Network (ANN), Support Vector Machine (SVM), Recurrent Neural Network (RNN), and Extreme Learning Machine (ELM) [9–14]. The Single layered feed-forward neural network gained more importance because of less computational complexity and hence more accurate prediction. ELM depends on the propitious selection of the hidden neurons and the activation function for stability and generalization. The limitation of ELM is the selection of the number of neurons in the hidden layer and weight matrix. To solve this problem the kernel fu stochastic %Knction came into existence. This improved the stability and generalization of the method. The kernel function enhances the system stability and when applied in combination with the ELM technique it is termed Kernel-based Extreme Learning Machine (KELM) [15]. The basic concept of KELM is such that the kernel matrix is formulated to describe the feature mapping where no random weights are used for its processing. Later the EL also came into use for the purpose of stock price prediction. EL comprises of mainly three types bagging boosting and stacking. The most commonly used method being the bagging and boosting method. The recent researches mainly focus on combination of statistical and AI models. Methods used like clustering [16, 17], multiple kernel-based prediction [18 –21], ensemble-based methods included Neural Network Regression Ensemble (NNRE), Support Vector Regression Ensemble (SVRE), Boosted Regression Tree (BRT), and Random Forest Regression they mainly used the process of boosting. Very less researches have been performed for stacking based Ensemble Learning (EL) process. In recent studies different tree-based ensemble methods have found its importance. Methods like Random Forest (RF), Gradient Boosting (GB), have owned importance because of their stability, simplicity, and robustness in prediction [22–25].

CNN and LSTM were previously used as a general procedure for prediction of short-term stock price, and showed high accuracy in result that's why this method is used in our model in a different way.

III. Background Theory and Algorithm

This section briefly describes different methods used in this study for prediction purposes.

A. Convolutional Neural Network

CNN possess the characteristic of giving importance to the most obvious features. This characteristic helps in feature selection process. CNN is a type of feed forward neural network; CNN comprises of two parts namely the convolution layer and the pooling layer. It is mostly used for time series prediction as the weight sharing process greatly improves the accuracy of the model by reducing the number of parameters. The convolution layer comprises of convolution kernels, after the working functioning of the convolutional layer, the features are extracted. The dimension of the extracted feature tends to be very high and hence a pooling layer is added for reduction of the features dimension. Eq. (1) shows the comprehensive calculation of the kernel present in the CNN for reduction in features. Fig. 1 shows the architecture of the LSTM.

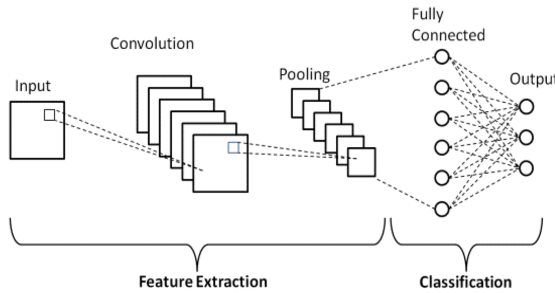

Figure 1. Basic concept of CNN technique.

$$O_p = tanh(x_t \times k_t + b_t) \quad (1)$$

O_p shows the output result after performing the convolution, x_t, k_t, and b_t represents the input, weight vector and bias value respectively and tan h being the activation function. Fig. 1 demonstrates the process of CNN. Different activation function can be used but for this particular study tanh function is used as it is observed to give better result.

B. Long Short Term Memory

LSTM is a type of the RNN algorithm and belongs to the same family, the prime conception of this RNN technique being the feedback layer. The model's output depends upon both the input data as well as the data of the preceding result or hidden layer. The major drawback of this technique is the back propagated data that gets increased or decreased at every step initiating to a vanishing data followed after noticing separate steps. Thus, the LSTM technique resolved the vanishing data problem [22, 23]. Different from the conventional neural networks LSTM comprises of memory blocks alternatively to the neurons connected between the layers. Independent blocks include different gates operating the blocks output. It is mainly composed of three gates that functions for solving the network, namely the input gate, the forget gate and the output gate. The activation function acts as the decision-making term. Primary steps of the gates are given as follows:

1) Depending on different conditions the activation process is controlled and thus enhancing the up gradation of the memory. The function occurs in the memory gate.
2) As per the name given the forget gate discard the inadequate information. By forgetting the previous cell information, this phenomenon takes place in the forget gate.
3) Depending upon the previous information from the forget gate the main output is obtained. This happens in the output gate.

Mathematically the nodes can be represented as:

$$i_g = \delta(w_{xi} \times X_{tt} + W_{hi} \times h_{tt-1} + W_{ci} \times C_{tt-1} + b_i) \quad (2)$$

$$f_g = \delta\left(W_{xf} \times X_{tt} + W_{hf} \times h_{tt-1} + W_{cf} \times C_{tt-1} + b_f\right) \quad (3)$$

$$C_g = f_t c_{t-1} + i_t \, tanh(W_{xc} X_{tt} + W_{hc} h_{tt-1} + b_c) \quad (4)$$

$$O_g = \delta(W_{xo} X_{tt} + W_{ho} h_{tt-1} + W_{co} c_t + b_o) \quad (5)$$

$$h_t = o_g \, tanh(c_t) \quad (6)$$

Eqs. (2)–(6) shows the formation of various nodes, here i_g, f_g, and o_g being the input gate, forget gate and output gate respectively. The input of the LSTM is twofold process, where the current sample X_{tt} is the previous hidden layer sample, h_{tt-1}, being the cell state and c_{t-1} gives the internal source of each gate. While the activation function is given as δ and tan h. W_{xf} is the weight matrix for signal *a* and *f*.

The activation of f_g further multiplies the cell state and it gets updated, in this process tan h function is used to process the cell and the final LSTM output is obtained as shown in Eq. (6). The primary motive of LSTM is preserving the internal cell memory in its entire life process. In this technique C_t determines the future of the elements, i.e., which needs up gradation and which can be erased. This being the basic difference as compared to the traditional RNN. Fig. 2 shows the framework of the LSTM. Fig. 3 shows the network connection between the input and output layer of LSTM.

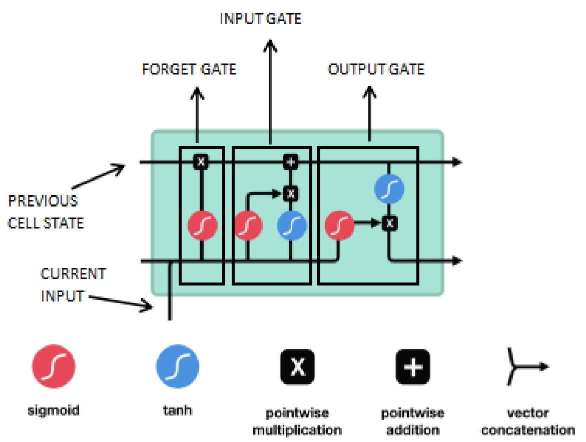

Figure 2. Detailed framework of the LSTM technique.

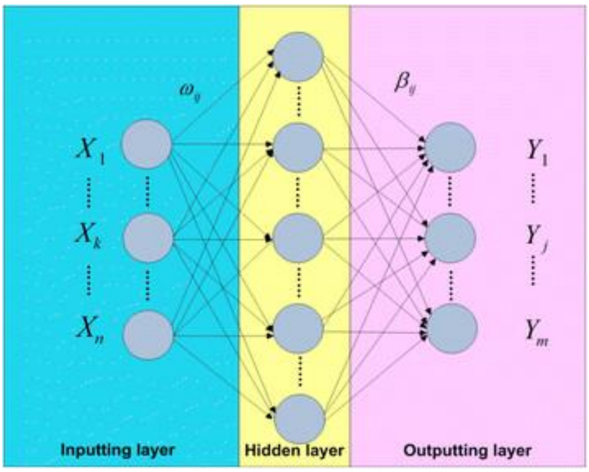

Figure 3. Structure of the LSTM technique.

C. Proposed Method

The proposed model consists of XGB model along with the LSTM model. the base models comprises of the independent input features and therefore provides the prediction result individually. These data functions as the metadata and fed as input to the regression which provides us the final predicted output. Thus, the output obtained from the ridge regressor consists of the weighted average of each output of all the base models.

As given in the input data set

$$I_{NGB} = \lfloor x_1, x_2, \ldots, x_{ngb} \rfloor \quad (7)$$

X_i denotes the input vector comprising of different technical indicators of the input features

$$x_i < x_1, x_2, \ldots, x_n > \quad (8)$$

The main aim is defining a function for predicting stock price at the time *t*+1 based on the technical indicator of the previous day.

$$\hat{y}_{t+1} = f(x)_t = f(x_1, x_2, \ldots x_n)_t \quad (9)$$

$\hat{y}_{t+1}$ is the predicted price and *f*(.) is the mathematical function of the proposed model (XGB and LSTM)

$$f(x) = [f_1(x), f_2(x), \ldots, f_m(x)] \quad (10)$$

where $f_1(x), \ldots, f_{m-1}(x)$ denotes the $m-1$ XGB model and $f_m(x)$ represents the LSTM model.

The basic model is trained for predicting the closing price as y_i is the individual base model's predicted closing price, ridge regression's input is the output of the base model. this provides the weights to the individual base model for final closing price prediction.

$$\begin{aligned} \hat{y}_1 &= f_1(x) \\ \hat{y}_2 &= f_2(x) \\ \hat{y}m &= f_m(x) \end{aligned} \quad (11)$$

$$\hat{y} = w_0 + w_1 \times \hat{y}_1 + \ldots\ldots + w_m \times \hat{y}_m \quad (12)$$

W_i represents the weight assigned to the base elements, the cost function optimizes the weights with the help of the cost function:

$$c_f(w) = \frac{1}{2n}\sum_{i=1}^{n}\left(\hat{y}_i - \hat{y}_j\right)^2 + X\sum_{j=1}^{m} w_j^2 \quad (13)$$

$C_f(w)$ is the cost function which is utilized to optimize the weight of the ridge regressor. And y_i is the actual closing price.

In this paper, CNN is applied in the two-stage hybrid prediction models, i.e., XGB-LSTM in this to propose. This proposed method comprises of CNN based features. Different technical indicators used in this paper for the purpose of input parameters. Table I. shows the different technical editors included for the purpose of prediction using the available data. The proposed model structure in shown in Fig. 4.

TABLE I. SHOWS THE DIFFERENT TECHNICAL EDITORS

Indicator Names	Formulae
Simple moving average	$(C_p+C_{p-1}+\ldots+C_{p-n})/z$
Momentum	C_p-C_{p-n}
Stochastic% K	$[(C_p-L_{p-(n-1)})/(L_{p(n-1)}-L_{p-(n-1)})]\times 100$
Accumulator/Distributor accumulator	$(H_p-C_{p-1})/(H_p-L_p)$
Historical Closing Price	NA
OHLC	NA

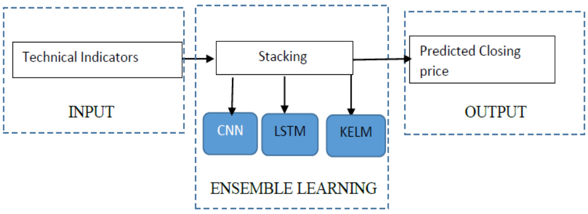

Figure 4. Basic architecture of the proposed technique.

IV. EXPERIMENTAL SETUP AND RESULT ANALYSIS

This section demonstrates a detailed explanation of the experimental setup and performance of the hybrid model. And that how the proposed model gives a better performance accuracy as compared to other previously implemented models. Direct prediction of the closing price helps in tracking the stock's trend. Based on recent studies we can say that the extraction of important features can help in more accurate prediction.

A. Feature Selection

As shown in Table I, a number of technical features are there for predicting the closing price of a particular stock, but all the features does not hold equal amount of importance while predicting the closing price. The primary procedure followed is selecting the features those are best suited for prediction of stock market with maximum accuracy. CNN technique is used in this study for performing the feature selection and based on the correlation between the closing price the best features are considered. Table II shows the relation between the feature and the closing price, the more the correlation coefficient the better is the feature for predicting the closing price

TABLE II. CORRELATION COEFFICIENT BETWEEN DIFFERENT TECHNICAL FEATURES AND CLOSING PRICE

Technical Indicators	CCR
Simple moving average	0.96
Momentum	0.90
Stochastic D%	0.85
Stochastic %K	0.89
Accumulator/ Distributor accumulator	0.97
Historical closing Price	0.98
OHLC	0.98

Based on Table II, the best CCR is obtained for OHLC, Simple moving average, historical closing price and lastly OHLC.

B. Data Pre-Processing

The data obtained is highly non-linear and also processes some overlapping and missing data. To maintain a proper synchronisation of the data present the whole data set is categorised under a single format [0 1]. This helps in maintaining the uniformity of the stock price data. The process of data normalization is performed using Eq. (14) this also improves the training speed and minimises the calculation overflow.

$$N_{norm} = \frac{x_N - x_{N\min}}{x_{N\max} - x_{N\min}} \quad (14)$$

N_{norm} is the normalized value, X_N being the total number of data, X_{Nmin} being the minimum data of the data set and X_{Nmax} shows the maximum of the data set.

C. Data Collection

The stock market data is obtained from the yahoo finance [26] latest verified on 1st September 2022. The opening price, high, low, and closing price is collected for Yes bank and HDFC bank for performing the proposed model. The experiment is performed on MATLAB 2017.

D. Result Analysis

This work predicts the closing price of the two banks in advance for 1 day ahead. For solving the prediction problem. The entire data set which is expressed as $D=(X,Y)$, where X represents the sample input and Y represents the corresponding labels of the input sample. X can be defined as; $X=(X_1, X_{2,} \ldots, X_{n-1}, X_n)$ in a similar manner Y is also represented in terms of total number of sample (n). n represents the technical indicators considered for predicting the closing price. X_i can be expressed in terms of $X_{i0}, X_{i1}, X_{i2}, X_{i3}, X_{i4}$. The flow chart of the proposed model in given in Fig. 5. This section shows the process for ascertaining the best structure of XGB model. Initially, the XGB model is executed for validation purpose and obtaining the best hyper parameter with the help of grid search technique, the following observations are obtained. Algorithm 1 is followed for XGB-LSTM performance.

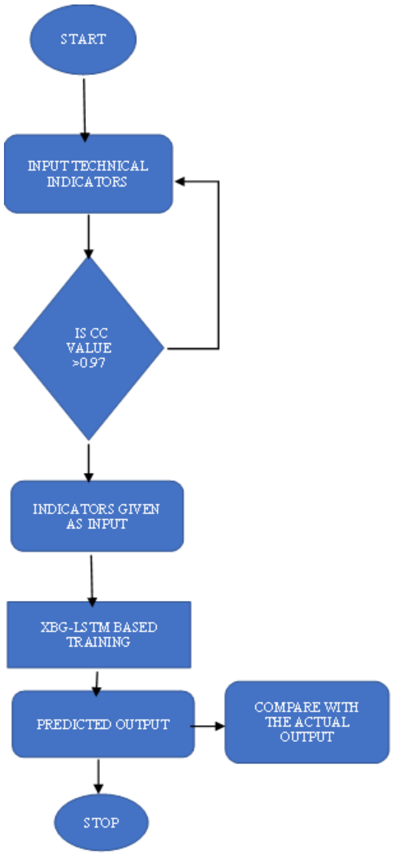

Figure 5. Flow chart of the proposed processes.

Algorithm 1. Ensemble proposed technique (XGB-LSTM with ridge regression)

Initialise
Begin
Different features of the CNN, let us consider the features be s $=\{s_1, s_2,\ldots s_n\}$
Now generate the data for training and cross validation, s= $(SD_{tr}\ SD_{cv})$
dM denotes maximum acceptable accuracy rate of drop
N denotes the network
N is trained for loss minimisation with (s) such that the accuracy rate is in inacceptable region
Calculate accuracy rate of SD_{tra} and SD_{cv}
Now rank the networks
If I <N, i++, NEXT
Else stop
End

The important hyper parameters are given below:

First being the regression tree number (n_e estimators, k) and, secondly the maximum depth of a regression tree (d_p).

Different XGB techniques are implemented for building the XGB forest. Where Nm being the optimal number of XGB trees among XGB forest. It depends upon the accuracy of the model and validation data. The models multiplicity is enhanced by varying the values of the hyper parameters. The optimal hyper-parameters used are given below:

$$\text{n\_estimators}, k = [25,50,100,150,200]$$

$$\text{max\_depth}, d_p = [2,4,6,8,10]$$

$$\text{learning\_rate } L_r = [0.1,0.3,0.5]$$

The algorithm responsible for feature selection is given in Algorithm 2. Fig. 6 shows the original closing price data of YES bank and it is found that the source data is highly nonlinear based on this observation it can be well understood the complexity of accurate prediction of the stock market. Fig. 7 shows the comparative analysis of different models used for predicting stock market in this paper and it can be well observed that the proposed model outperforms the other models considered in this paper based on their implementation. The corresponding data showing the corresponding values are shown in Table III. The table clearly shows that the proposed model performs better than other previously used techniques. The Mean Absolute Percentage Error (MAPE) value of the proposed ensemble technique is 1.172% which is less when compared to the SVM technique (MAPE value of 2.53%). As it is the matter of a financial market every percentage is of greater importance. The data represented in the graph is for 35 days only for proper visualization and distinguishing between different techniques. Table IV shows the comparative analysis of the second data that is for HDFC bank. Although there is a little bit of increase in the computational time while showing better result for both the case studies, but it can be ignored as the fluctuation of time in comparison to the prediction accuracy is very less. Algorithm 2 shows the framework of the proposed technique. This study did not consider the emotional background of the market which is also an important aspect in predicting the future closing price which can be considered as a backlog of the proposed method that can be considered in future studies. So taking the markets emotional status as constant the prediction is performed which can be considered hereafter for more accurate closing price prediction.

Previously different prediction techniques were used for price prediction some of them are used for comparison purposes in the proposed study also. The MAPE value of XGB, ELM, and LSTM (1.91, 2.53 and 1.54, respectively) is observed to be greater than the proposed hybrid model (1.17). CNN-LSTM was already implemented for predicting the closing price of the stock market but it is observed that the proposed method performed better not only in terms of accuracy but also in terms of computational time. This is of great use when big data analysis is performed.

Algorithm 2. CNN based feature selection

Begin
Different features of the CNN, let us consider the features be s $=\{s_1, s_2,\ldots s_n\}$
Now generate the data for training and cross validation, s= $(SD_{tr}\ SD_{cv})$
dM denotes maximum acceptable accuracy rate of drop
N denotes the network
N is trained for loss minimisation with (s) such that the accuracy rate is in inacceptable region
Calculate accuracy rate of SD_{tra} and SD_{cv}
Now rank the networks
If I <N, i++, NEXT
Else stop
End

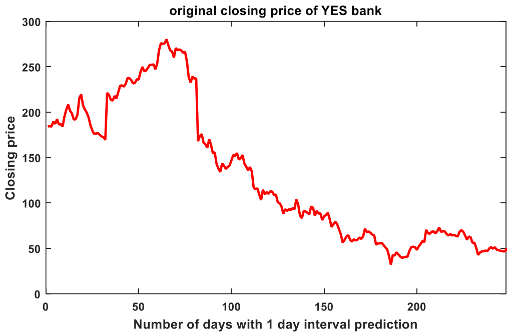

Figure 6. Original closing price.

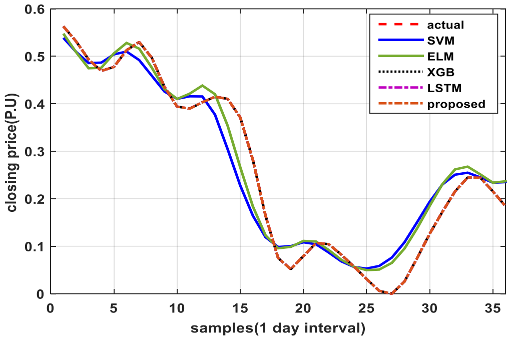

Figure 7. Predicted closing price.

TABLE III. COMPARISON ANALYSIS OF DIFFERENT TECHNIQUE FOR CLOSING PRICE PREDICTION OF YES BANK

Technique	MAPE (%)	MAE (P.U)	RMSE (P.U)	Tr (sec)
SVM	2.95	0.382	0.051	10.22
ELM	2.53	0.033	0.418	9.14
XGB	1.91	0.025	0.033	9.33
LSTM	1.54	0.020	0.027	10.23
proposed	**1.172**	**0.015**	**0.021**	**11.12**

TABLE IV. COMPARATIVE ANALYSIS BETWEEN DIFFERENT PREDICTION MODELS FOR HDFC BANK

Technique	MAPE (%)	MAE (P.U)	RMSE (P.U)
SVM	2.82	0.062	0.077
ELM	1.68	0.036	0.046
XGB	1.41	0.028	0.043
LSTM	1.23	0.025	0.053
proposed	**0.87**	**0.028**	**0.025**

V. CONCLUSION

From this paper it can be concluded that the proposed hybrid ensemble-based learning method performs better than the other models discussed in this study. The prediction thus can help in selecting the increase or decrease or rather buy or sell criteria of the stock market. The main drawback of the previously included the input variable selection which is solved by implementing the CNN based feature selection that segregates the best input that gives an accurate prediction result. It is also observed that the computational timing is low as compared to the traditional methods. Thus, making the proposed CNN based XGB-LSTM better in predicting the closing price. In future all the three methods of ensemble learning can be combined together to obtain the output. The features can be optimized using different optimization technique. Table V shows the comparative study of different prediction models that were previously followed.

TABLE V. COMPARATIVE STUDY OF PREVIOUSLY EMPLOYED TECHNIQUE AND PROPOSED

Stock market	Result obtained from literature review	Ref.
S&P 500 index	Percentag of correct directional predictions of; ANN: µANN>51.78 at 5% significance level	[29]
ISE National 100 index	Technical Analysis (10): Simple Moving Average, Weighted Moving Average Percentage of correct directional Percentage of correct directional predictions (average) of; ANN: 75. Polynomial SVM: 71. (at α=0.05 significance level, (at α=0.05 significance level, difference mean difference is performances.	[30]
S&P 500, DAX, and FTSE	Input variables: 1 (for each time series) Basic Price Data (1): Lagged values of S&P 500 Returns obtained (%) at 0.5% transaction costs by; **ANN** / **B&H** / **AR(1)** S&P 500 29.52 21.02 0.43 DAX 32.52 23.88 2.65 TOPIX 35.59 −6.69 2.93 FTSE 28.38 13.45 4.25	[27]
Dow Jones Industrial Average index	Percentage of correct directional predictions of; ANN: 55.1	[28]
Proposed method	Data for YES BANK, 1 Day ahead prediction of closing price with a MAPE error of	

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHORS CONTRIBUTION

Pratyush Ranjan Mohapatra developed the concept and did the literature survey along with Ajaya Kumar Parida. Santosh Kumar Swain and Ajaya Kumar Parida collected the data and performed the analysis and implemented the code. Santi Swarup Basa did the paper writing along with Ajaya Kumar Parida. All authors had approved the final version.

REFERENCE

[1] E. F. Fama, "Efficient capital markets: A review of theory and empirical work," *The Journal of Finance*, vol. 25, no. 2, pp. 383–417, May 1970.

[2] B. G. Malkiel, "The efficient market hypothesis and its critics," *Journal of Economic Perspectives*, vol. 17, no. 1, pp. 59–82, 2003.

[3] A. Timmermann and C. W. J. Granger, "Efficient market hypothesis and forecasting," *International Journal of Forecasting*, vol. 20, pp. 15–27, Jan.–Mar. 2004.

[4] K. Nti, D. F. Adekoya, and B. A. Weyori, "A systematic review of fundamental and technical analysis of stock market predictions," *Artificial Intelligence Review*, vol. 53, no. 4, pp. 3007–3057, Aug. 2019.

[5] K. O. Nti, A. Adekoya, and B. Weyori, "Random forest based feature selection of macroeconomic variables for stock market prediction," *American Journal of Applied Sciences*, vol. 16, no. 7, pp. 200–212, Sep. 2019.

[6] I. Rojas *et al.*, "Soft-computing techniques and ARMA model for time series prediction," *Neurocomputing*, vol. 71, no. 4–6, pp. 519–537, Jan. 2008.

[7] Q. Cao, K. B. Leggio, and M. J. Schniederjans, "A comparison between Fama and French's model and artificial neural networks in predicting the Chinese stock market," *Computers & Operations Research,* vol. 32, no. 10, pp. 2499–2512, Oct. 2005.

[8] Z. Zhongbao *et al.*, "Big data and portfolio optimization: A novel approach integrating DEA with multiple data sources," *Omega*, vol. 104, 102479, 2021.

[9] F. Ghashami, K. Kamyar, and S. A. Riazi, "Prediction of stock market index using a hybrid technique of artificial neural networks and particle swarm optimization," *Applied Economics and Finance*, vol. 8, no. 1, Dec. 2021.

[10] L. He *et al.*, "Supportive emergency decision-making model towards sustainable development with fuzzy expert system," *Neural Computing and Applications*, vol. 33, no. 22, pp. 15619–15637, Jun. 2021.

[11] Z. Tingting *et al.*, "Multi-step-ahead crude oil price forecasting based on two-layer decomposition technique and extreme learning machine optimized by the particle swarm optimization algorithm," *Energy*, vol. 229, 120797, 2021.

[12] X. Jingming *et al.*, "Financial time series prediction using ℓ2, 1RF-ELM," *Neurocomputing*, vol. 277, pp. 176–186, 2018.

[13] I. Majumder, P. K. Dash, and R. Bisoi, "Variational mode decomposition based low rank robust kernel extreme learning machine for solar irradiation forecasting," *Energy Conversion and Management*, vol. 171, pp. 787–806, 2018.

[14] W. Zhaoyue, H. Jinsong, and W. Yongjie, "A bimodel algorithm with data-divider to predict stock index," *Mathematical Problems in Engineering,* vol. 2018, 2018.

[15] F. Zhou *et al.*, "EMD2FNN: A strategy combining empirical mode decomposition and factorization machine based neural network for stock market trend prediction," *Expert Systems with Applications*, vol. 115, pp. 136–151, 2019.

[16] Y. Xu *et al.*, "A hybrid two-stage financial stock forecasting algorithm based on clustering and ensemble learning," *Applied Intelligence,* vol. 50, no. 11, pp. 3852–3867, 2020.

[17] W. Bin *et al.*, "Predicting short-term stock prices using ensemble methods and online data sources," *Expert Systems with Applications*, vol. 112, pp. 258–273, 2018.

[18] W. Chengzhang and B. Xiaoming, "Boosting learning algorithm for stock price forecasting," *IOP Conference Series: Materials Science and Engineering,* vol. 322, no. 5, 2018.

[19] C. Weiling *et al.*, "Leveraging social media news to predict stock index movement using RNN-boost," *Data & Knowledge Engineering*, vol. 118, pp. 14–24, 2018.

[20] M. A. Hassan *et al.*, "Exploring the potential of tree-based ensemble methods in solar radiation modelling," *Applied Energy*, vol. 203, pp. 897–916, 2017.

[21] A. Mosavi *et al.*, "State of the art of machine learning models in energy systems, a systematic review," *Energies*, vol. 12, no. 7, 1301, Feb. 2019.

[22] C. Tianqi and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proc. the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining,* New York USA, 2016, pp. 785–794.

[23] B. Suryoday *et al.*, "Predicting the direction of stock market prices using tree-based classifiers," *The North American Journal of Economics and Finance*, vol. 47, pp. 552–567, 2019.

[24] M. A. I. Sunny, M. M. S. Maswood, and A. G. Abdullah, "Deep learning-based stock price prediction using LSTM and bi-directional LSTM model," in *Proc. the 2020 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES)*, Giza, Egypt, 2020.

[25] I. Majumder, M. K. Behera, and N. Nayak, "Solar power forecasting using a hybrid EMD-ELM method," in *Proc. the 2017 International Conference on Circuit, Power and Computing Technologies (ICCPCT)*, Kollam, India, 2017.

[26] Yahoo Finance. [Online]. Available: https://finance.yahoo.com/

[27] S. T. A. Niaki and S. Hoseinzade, "Forecasting S&P 500 index using artificial neural networks and design of experiments," *Journal of Industrial Engineering International*, vol. 9, no. 1, pp. 1–9, Feb. 2013.

[28] K. Yakup, M. A. Boyacioglu, and Ö. K. Baykan, "Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange," *Expert systems with Applications*, vol. 38, no.5 pp. 5311–5319, 2011.

[29] J. Teo and D. Wood, "The profitability of daily stock market indices trades based on neural network predictions: Case study for the S&P 500, the DAX, the TOPIX and the FTSE in the period 1965–1999," *Applied Financial Economics*, vol. 14, no.4, pp. 285–297, 2004.

[30] R. R. Lisa *et al.*, "Meeting preceptor expectations to facilitate optimal nurse practitioner student clinical rotations," *Journal of the American Association of Nurse Practitioners*, vol. 32, no. 5, pp. 400–407, 2020.