International Journal of Electrical and Computer Engineering (IJECE)
Vol. 15, No. 2, April 2025, pp. 2192∼2201
ISSN: 2088-8708, DOI: 10.11591/ijece.v15i2.pp2192-2201

A novel technique for selecting financial parameters and technical indicators to predict stock prices

Sneha S. Bagalkot[1,2], Dinesha H. A.[3], Nagaraj Naik[4]
[1]Nagarjuna College of Engineering and Technology, Bengaluru and Visvesvaraya Technological University, Belagavi, India
[2] Department of Computer and Science Engineering, B.M.S. College of Engineering, Bangalore, India
[3]Department of Computer and Science Engineering, Shridevi Institute of Engineering and Technology, Tumkuru, India
[4]Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education Manipal, Manipal, India

Article Info

Article history:

Received May 30, 2024
Revised Nov 13, 2024
Accepted Nov 20, 2024

Keywords:

Artificial neural network
Deep neural network
Financial parameters
Recursive feature elimination
XGBoost

ABSTRACT

Stock price predictions are crucial in financial markets due to their inherent volatility. Investors aim to forecast stock prices to maximize returns, but accurate predictions are challenging due to frequent price fluctuations. Most literature focuses on technical indicators, which rely on historical data. This study integrates both financial parameters and technical indicators to predict stock prices. It involves three main steps: identifying essential financial parameters using recursive feature elimination (RFE), selecting quality stocks with a decision tree (DT), and forecasting stock prices using artificial neural networks (ANN), deep neural networks (DNN), and extreme gradient boosting (XGBoost). The models' performance is evaluated with root mean square error (RMSE) and mean absolute error (MAE) scores. ANN and DNN models showed superior performance compared to the XGBoost model. The experiments utilized Indian stock data.

Corresponding Author:

Ngaaraj Naik
Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education
Manipal-576 104, India
Email:nagaraj.naik@manipal.edu

1. INTRODUCTION

The stock market attracts numerous investors and traders due to its potential for rapid and substantial gains [1], [2]. Both institutional and individual investors engage in stock trading [3], [4]. Institutional investors, being trained professionals, manage significant investment portfolios [5], [6] and typically focus on long-term investments. In contrast, retail investors often lack the trading expertise necessary for optimal buy or sell decisions [7]. Consequently, a robust stock forecasting model is essential for assisting retail investors.

To identify stock price trends, investors utilize two primary methods of analysis. The first is technical analysis, which assesses stock trends using historical data such as opening, closing, high, and low prices, along with trading volume. While technical analysis is applicable for both short-term and long-term investments, its limitation lies in its exclusive reliance on historical data, often overlooking fundamental factors crucial for future price predictions [8]. This study, therefore, integrates both financial parameters and historical stock data to enhance forecasting accuracy.

The fundamental approach involves analyzing financial statements [9] and evaluating metrics like earnings, assets, liabilities, and book value [10], [11]. The challenge arises from the plethora of financial parameters, making it difficult to determine which ones are most predictive. To address this, the study employs the recursive feature elimination (RFE) technique to identify the most relevant financial parameters.

Subsequently, the study uses a decision tree method to identify high-quality stocks. Decision trees are effective for classifying target variables [12], [13], as they partition data based on features, with nodes representing features and leaf nodes indicating outcomes based on entropy reduction or information gain. After identifying quality stocks, the study predicts their future prices using data mining methods, particularly machine learning [14],[15]. Machine learning approaches, including supervised and unsupervised methods [16], are employed, with artificial neural networks (ANNs) being used for their ability to handle non-linear and volatile data [17], [18]. Additionally, boosting algorithms, particularly XGBoost, are gaining traction for their gradient boosting ensemble approach [19], providing scalable solutions for end-to-end tree boosting. Although many studies focus solely on historical stock data for forecasting [20]–[22], integrating financial parameters with historical data is crucial for accurate predictions. This study, therefore, combines these elements to improve the estimation of future stock prices.

The contributions of this paper are summarized as follows:

a. We considered the RFE technique to identify the important financial parameters.
b. To identify the quality of the stock, we considered the decision tree.
c. Financial parameters and historical stock data are considered to forecast the future value of the stock price using ANN, deep neural networks (DNN), and XGBoost models.

Section 2 describes literature reviews, and section 3 provides methodology. Results are presented in section 4. Section 5 presents the conclusion of the work and future directions.

2. LITERATURE REVIEWS

Yun *et al.* [23] introduced the Genetic-XGBoost algorithm with a three-stage feature engineering method. Their model, incorporating 67 technical indicators, achieved an average prediction accuracy of 93.28% on the KOSPI dataset, significantly outperforming previous models. The study emphasizes the importance of selecting an optimal feature set to avoid overfitting and simplify interpretation. Gunduz *et al.* [24] used LightGBM and long short-term memory (LSTM) classifiers to predict banking stock movements on the Borsa Istanbul exchange. The LSTM models utilized 2D tensors, incorporating data from the previous eight hours for hourly predictions, demonstrating the effectiveness of temporal data integration. Peng *et al.* [25] considered 124 technical indicators for stock price prediction, employing feature selection techniques to eliminate irrelevant indicators. They used a deep neural network model and fine-tuned hyperparameters, such as hidden layers and dropout rates, to enhance prediction accuracy.

Li *et al.* [26] proposed a deep learning framework for gross domestic product (GDP) prediction involving three steps: feature extraction with feature crossing, relationship analysis with Boruta-RF and Q-learning, and model creation with TCN. Their method improved temporal convolutional network (TCN) prediction accuracy by 10%. Selecting the best subset of features is NP-hard, necessitating heuristic methods [27], [28]. Feature selection's scalability is crucial for large datasets, yet high-dimensional data can lead to local optima and costly parameter adjustments.

Handling non-linear and noisy data for feature selection is challenging [29]. Recent studies show that deep learning (DL) models, particularly LSTM, excel in financial forecasting by leveraging long-term dependencies [30], [31]. Time complexity can be reduced by eliminating variables before model integration. Filter feature selection scores each feature, while the wrapper technique uses subsets for model training. Combining these approaches, the genetic algorithm optimizes feature subsets via natural selection. Identifying market styles is vital for predicting stock prices [32]. The summary of literature reviews are described in Table 1.

Chun *et al.* [33] developed a stock forecasting model integrating investor emotions by analyzing microblogging data. The model employs part-of-speech (POS) tagging to extract emotional terms such as adjectives, nouns, adverbs, and interjections, which are classified into emotions like joy and sadness. While technical indicators are commonly used for stock price forecasting, they primarily rely on historical data, often neglecting fundamental company factors. Evaluating stock quality requires considering financial parameters, which provide insights into future stock value.

Table 1. Literature reviews on stock price prediction

Author	Data	Method	Target output	Gap
Yun *et al.* [23]	Technical indicators	XGBoost	Prediction	Stock historical data is considered.
Gunduz *et al.* [24]	Technical indicators	LSTM	Classification	Stock historical data is considered.
Peng *et al.* [25]	Technical indicators	Deep neural network	Prediction	Stock historical data is considered.
Li *et al.* [26]	GDP	Deep learning	Prediction	Stock historical data is considered.
Chen *et al.* [30]	Technical indicators	Deep learning	Prediction	Stock historical data is considered.
Kim *et al.* [1]	Technical indicators	Genetic algorithm	Prediction	Stock historical data is considered.
Kumar *et al.* [34]	Technical indicators	ENN	Prediction	Stock historical data is considered.
Chun *et al.* [33]	Micro-blogging data	Deep learning	Prediction	Stock historical data is considered.

3. PROPOSED METHODOLOGY

The overall work is described in Figure 1. The proposed work is classified into three tasks. The first is the identification of important financial parameters, the second is identifying quality stock, and the third is estimating the future value of stock price. The financial parameters are collected from NSE. However, it consists of many financial parameters such as price to earnings, book value, and returns on equity.. These financial parameters help identify the intrinsic value of stock prices. In this work, we considered 14 financial parameters. The aim is to identify important financial parameters. Therefore, we considered the RFE approach to remove the noisy financial parameters. The next step is to identify the quality stock; therefore decision tree is considered. In DT, the accuracy metric is used to identify stock quality. Finally, selected quality stocks extracted historical data and given input to the ANN, DNN, and XGBoost models to forecast the future stock price.

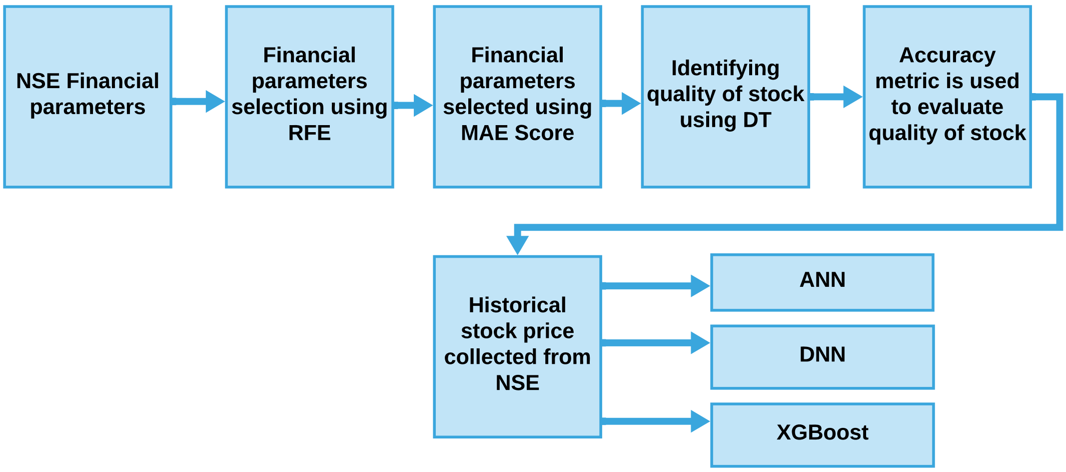


Figure 1. Financial parameters based stock price prediction

3.1. Recursive feature elimination (RFE)

The RFE method identifies significant stock financial parameters by sequentially eliminating features based on their MAE scores. This study evaluated 14 financial parameters using random forest regression to assess feature importance, with earning per share (EPS) as the dependent variable. The list of financial parameters are described in Table 2. Features with MAE scores exceeding 20% were deemed irrelevant; hence, other income, depreciation, and dividend payout were excluded due to their high MAE scores. The remaining parameters were used as inputs for the decision tree to determine stock quality. The step by step are described in algorithm 1.

3.2. Decision tree (DT)

DT is a common approach for solving regression and classification tasks [35], [36]. The relevant financial parameters selected by the RFE method are given to DT for classification. The 11 financial parameters are given input to the DT to identify quality stock. The supervised learning method is used in DT. Earning per share (EPS) is considered the target variable to classify the quality stock. Here more than 20% EPS returns are considered good quality stock; otherwise, it is classified as no category. The decision tree determines whether the stock has quality, i.e., yes or no, based on the financial parameters. A DT is used to classify the stock, either quality stock or not. These yes and no labels are maps in decision tree classifier vectors. The decision tree simple IF and ELSE conditions are mapped on classifier vectors. A decision tree is generated based on the input vector data, and the information gain metric is used to split the decision tree. In DT, 70% of data is considered training, and 30% of data is considered for testing. We have considered the accuracy, precision, and recall metric to evaluate the decision tree. Using a decision tree, we found that stock with more than 75% accuracies is considered good quality stock, shown in Table 3. Later these stocks are given input to the ANN, DNN, and XGBoost for stock price forecasting.

Algorithm 1. Recursive feature wlimination (RFE) algorithm

```
Input:
        X: Feature matrix (including financial parameters and technical indicators)
        y: Target variable vector
        estimator: Machine learning estimator
        n_features: Desired number of features to select

Output:
        Selected feature subset
Initialize a list features with indices of all the features in X
Initialize current_score to a very large number
while length of features ¿ n_features do
    Fit the estimator model on X[:, features] and y
    Calculate the importance of each feature using the estimator model
    Rank the features based on their importance scores in ascending order
    Remove the least important feature from features
    Fit the estimator model on the reduced feature set and calculate the Mean Absolute Error (MAE)
    if new MAE score is lower than current_score then
        Update current_score to the new MAE score
    else
        Break the loop
    end if
end while
return features
```

Table 2. Financial parameters MAE score using RFE

Serial No	Financial parameter feature	MAE Score
1	Sales	18.5
2	Expenses	19.6
3	Operating Profit	17.7
4	Other Income	**23.4**
5	Depreciation	**25.7**
6	Interest	16.7
7	Profit before tax	19.8
8	Tax	17.8
9	Net profit	18.7
10	EPS	17.6
11	Price to earning	16.3
12	Price	17.1
13	Dividend Payout	**26.2**
14	OPM	19.1

Table 3. Decision tree to evaluate quality stock

Stock	Accuracy	Precision	Recall
SunPharma	77.50	0.85	0.77
Cipla	76.30	0.81	0.70
Bharti-Airtel	78.50	0.86	0.77
Bajaj Auto	76.80	0.84	0.74
Federal Bank	77.50	0.82	0.79
RBL Bank	79.60	0.87	0.78
Bank of Baroda	76.50	0.81	0.79

3.3. Artificial neural network (ANN)

ANN was developed based on a biological neuron model that mimics human behavior [37]. ANN has neurons interconnected with different layers. Using a decision tree, we have identified the quality stock. These quality stock historical data have been collected. It consists of five variables open price, low price, high price ,volume, and close price. Here the close price is considered the target variable for the ANN model. In ANN, each neuron performs the computational task [38]. In this work, we considered neural networks, which have one hidden layer between the input and the output layer. The input data is multiplied with random weight along with bias. To deal with the nonlinearity in data sigmoid activation function is considered in the ANN model. The loss function goal is to reduce the error rate in the model. The back-propagation algorithm computes the gradient of the loss function by randomly assigning weight to neurons in each iteration.

3.4. Extreme gradient boosting (XGBoost)

XGBoost is one of the popular methods to deal with the non-linearity in data [39]. Using a decision tree, we have identified the quality stock. The next task is forecasting the future value of stock prices. For that, we considered historical stock data as input to the forecasting model, such as open price, low price, high price, volume, and close price. Here the close price is considered the target variable for the XGBoost model. Using the XGBoost method, computed the residual differences between actual and predicted [40]. Based on the residuals regression tree is constructed in the proposed work. The similarity score of each tree is computed based on the the (1) and (2).

$$Similarity - score = \frac{residualsum, square}{residual + \lambda} \tag{1}$$

$$Output - value = \frac{residualsum}{residual + \lambda} \tag{2}$$

The loss function goal is to reduce the residual error in the tree. We considered the square of residuals and adjusted the parameter value of λ . The output value will decrease if λ is more than zero. Because we are maximizing the value produced by the first tree, lambda is reset to zero.

3.5. Deep neural network (DNN)

The deep neural network is one of the popular methods for stock price forecasting. Therefore, this work considered DNN to forecast the future stock price. Using a decision tree, we have identified the quality stock. These quality stock historical data have been collected. It consists of five variables open price, low price, high price, volume, and close price. Here the close price is considered the target variable for the DNN model. The input layer reads input data and performs the computational task. The output layer produces the output [41], [42]. The proposed work constructed four hidden layers in DNN architecture. A DNN with four hidden layers enhances stock price prediction by progressively capturing complex, non-linear relationships within the data, with each layer refining the patterns to improve accuracy while reducing overfitting. The DNN Equation is defined in (3).

$$Y_i = (Bias + \sum_{k=1}^{n} w_k * A(x)) \tag{3}$$

Here w_k represents the weights, and $A(x)$ represents the input vector. To deal with the nonlinearity in the data, we considered the rectified linear units activation functions. Backpropagation is used to determine the appropriate value for weights and biases. We have fine-tuned the DNN parameters, such as the number of neurons and hidden layers, and adjusted the learning rate to get the best results.

4. RESULTS AND DISCUSSION

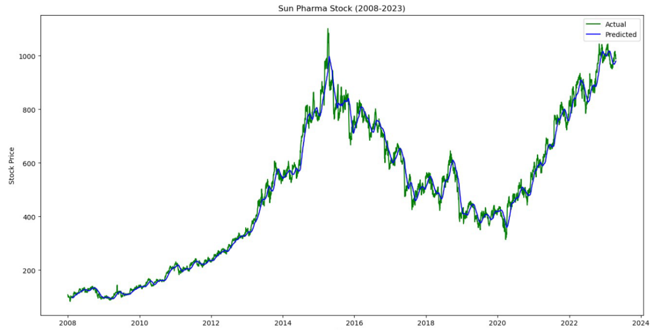
In the experiment setup, national stock exchange (NSE) data were used for investigation ($https : //www.nseindia.com/$). The proposed work goal is to identify the relevant financial parameters and forecast the stock price. The proposed work is divided into three tasks. The first is the recursive feature elimination (RFE) method to identify the relevant financial parameters. Using the RFE method, we identified the important financial parameters. Out of 14 features, three features are considered irrelevant features. The irrelevant features are other income, depression, and dividend payout. The financial parameters with the highest MAE are considered irrelevant indicators. The second is to identify the quality stock using the decision tree. The decision tree determines whether the stock has quality, i.e., yes or no, based on the accuracy metric. The accuracy metric is used to assess the stock quality.

Machine learning models are evaluated based on accuracy metrics. The third is stock price forecasting. The historical stock price is collected for further analysis. It includes stock data such as the open, high, low, and close prices and the stock price volume. Around 4910 trade sessions were considered for experiments from May 1, 2007, to October 11, 2022. To forecast the future value of stock price, we considered ANN, DNN, and XGBoost models. To get the best performance in prediction models, we have varied the number of neurons, hidden layer, and learning rate.

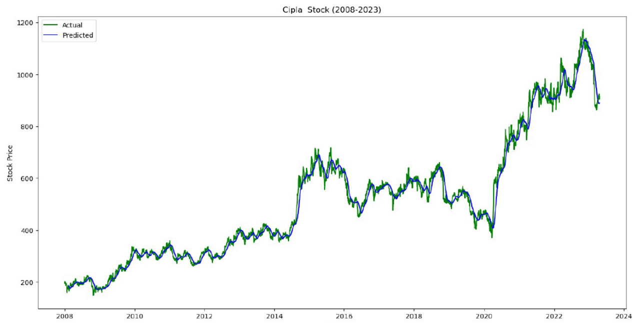
$$MAE = (\frac{1}{n})\sum_{i=1}^{n}|y_i - x_i| \tag{4}$$

$$RMSE = \sqrt{(\frac{1}{n})\sum_{i=1}^{n}(y_i - x_i)^2} \tag{5}$$

The MAE and RMSE metric is defined in (4) and (5). Figure 2 shows the stock prices predicted using deep neural network: Figure 2(a) Sun Pharma Stock and Figure 2(b) Cipla Stock. Figure 3 shows the stock prices predicted using deep neural network: Figure 3(a) Bharti Airtel Stock and Figure 3(b) Federal Bank Stock. Here red line indicates the predicted price, and the black line indicates the actual price. We found that the predicted price needs to fit volatile stock prices better. For that, we have fine-tuned the parameters of a model, like the number of neurons and learning rate. When setting up a DNN for stock prediction, we carefully tuned various hyperparameters to enhance model performance. We balanced fast convergence and training stability with a 0.01 learning rate. We balanced processing speed and consistent gradient updates using a 32-batch size. We trained the model over 100 epochs to prevent overfitting and ensure learning. We used the Adam optimizer to handle sparse gradients and automatically modify the learning rate during training. We reduced overfitting by randomly deactivating neurons during training with a dropout rate of 0.5. The rectified linear units (ReLU) activation function added non-linearity to the buried layers. The network had four hidden layers, with neuron counts decreasing from 128 to 64 to 32. As indicated in Table 4, the DNN outperformed both the ANN and XGBoost models in most scenarios. For example, in forecasting SunPharma stock, the DNN achieved an MAE of 3.210941, RMSE of 7.845115, and an R-squared value of 98.60. In contrast, the ANN had an MAE of 3.937627, RMSE of 8.452742, and an R-squared value of 98.20, while the XGBoost model recorded an MAE of 5.507795, RMSE of 8.986556, and an R-squared value of 97.80. Comparable results were observed for Cipla, Bharti-Airtel, Federal Bank, RBL Bank, and Bank of Baroda, with the DNN consistently delivering lower MAE and RMSE values along with higher R-squared values, signifying superior predictive accuracy and overall model performance. Additionally, these outcomes were benchmarked against existing literature by [43], where their multilayer perceptron (MLP) and convolutional neural network (CNN) models exhibited higher MAE and RMSE values, highlighting the effectiveness of our DNN configuration. For instance, their CNN model's performance for stock prediction resulted in an MAE of 25.66 and an RMSE of 36.87, showcasing the superior accuracy of our DNN approach.

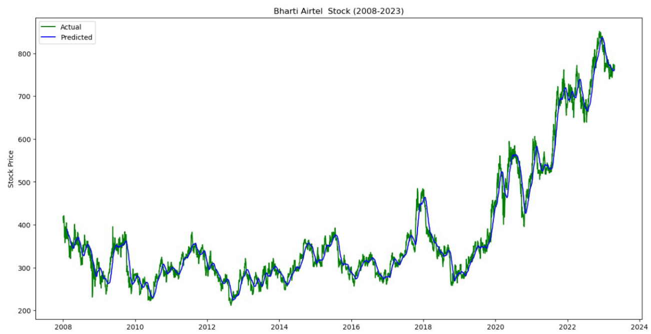


(a)



(b)

Figure 2. Stock price predicted using DNN (a) Sun Pharma Stock and (b) Cipla Stock



(a)

(b)

Figure 3. Stock price predicted using DNN (a) Bharti Airtel Stock and (b) Federal Bank Stock

Table 4. Result

Stock	Predictions Techniques	MAE	RMSE	R-squared
SunPharma	ANN	3.937627	8.452742	98.20
SunPharma	DNN	3.210941	7.845115	98.60
SunPharma	XGBoost	5.507795	8.986556	97.80
Cipla	ANN	3.174641	5.191448	98.95
Cipla	DNN	3.346246	5.445911	98.85
Cipla	XGBoost	5.251966	7.658853	98.50
Bharti-Airtel	ANN	6.334113	10.11714	97.30
Bharti-Airtel	DNN	7.088787	11.96207	97
Bharti-Airtel	XGBoost	5.490016	7.15828	98.30
Bajaj Auto	ANN	25.35	28.33	89.60
Bajaj Auto	DNN	28	30	88.60
Bajaj Auto	XGBoost	33	35	86.60
Federal Bank	ANN	18.17	18.98	93.20
Federal Bank	DNN	14.50	14.90	94.40
Federal Bank	XGBoost	17	17.60	94.10
RBL Bank	ANN	21	21.30	91.10
RBL Bank	DNN	17	17.60	92.40
RBL Bank	XGBoost	18.50	18.60	93.40
Bank of Baroda	ANN	22	22.70	91.60
Bank of Baroda	DNN	19	19.40	93.10
Bank of Baroda	XGBoost	20.10	20.80	92.40
Lu *et al.* [43]	MLP	31.49	39.26	96.99
Lu *et al.* [43]	CNN	25.66	36.87	97.35
Lu *et al.* [43]	BiLSTM	21.95	31.62	98

5. CONCLUSION

The proposed work encompasses three main tasks aimed at enhancing stock price prediction to assist investors in making profitable decisions. First, we identified relevant financial parameters using the RFE method. Our findings indicated that parameters such as other income, depreciation, and dividend payouts do not significantly contribute to stock price prediction, as evidenced by MAE scores greater than twenty, reflecting high data variance. Second, we utilized DT to identify quality stocks, providing a basis for robust stock selection. Lastly, we forecasted stock prices using ANN, DNN, and XGBoost methods. As demonstrated in experimental work ANN and DNN outperformed XGBoost in most cases due to their ability to handle nonlinearity in stock price data. Specifically, the DNN achieved lower MAE and RMSE values and higher R-squared values across various stocks, indicating superior predictive accuracy. In contrast, XGBoost was less effective in scenarios with high data volatility. Future research could explore integrating both technical and fundamental indicators to improve long-term stock price predictions. Additionally, leveraging options and derivatives data may enhance the accuracy of future stock value forecasts.

REFERENCES

[1] M. Kim, "A data mining framework for financial prediction," *Expert Systems with Applications*, vol. 173, p. 114651, 2021, doi: 10.1016/j.eswa.2021.114651.

[2] S. Gao, Q. Meng, and K. C. Chan, "IPO pricing: do institutional and retail investor sentiments differ?," *Economics Letters*, vol. 148, pp. 115–117, 2016, doi: 10.1016/j.econlet.2016.09.029.

[3] A. Kumar and C. M. C. Lee, "Retail investor sentiment and return comovements," *Journal of Finance*, vol. 61, no. 5, pp. 2451–2486, 2006, doi: 10.1111/j.1540-6261.2006.01063.x.

[4] W. Li and S. S. Wang, "Daily institutional trades and stock price volatility in a retail investor dominated emerging market," *Journal of Financial Markets*, vol. 13, no. 4, pp. 448–474, 2010, doi: 10.1016/j.finmar.2010.07.003.

[5] F. Wen, L. Xu, G. Ouyang, and G. Kou, "Retail investor attention and stock price crash risk: Evidence from China," *International Review of Financial Analysis*, vol. 65, p. 101376, 2019, doi: 10.1016/j.irfa.2019.101376.

[6] E. L. Von Thadden, "Long-term contracts, short-term investment and monitoring," *Review of Economic Studies*, vol. 62, no. 4, pp. 557–575, 1995, doi: 10.2307/2298077.

[7] T. Vo-Van, H. Che-Ngoc, N. Le-Dai, and T. Nguyen-Trang, "A new strategy for short-term stock investment using Bayesian approach," *Computational Economics*, vol. 59, pp. 887–911, 2022, doi: 10.1007/s10614-021-10115-8.

[8] B. Itri, Y. Mohamed, B. Omar, E. M. Latifa, M. Lahcen, and O. Adil, "Hybrid machine learning for stock price prediction in the Moroccan banking sector," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 14, no. 3, pp. 3197–3207, 2024, doi: 10.11591/ijece.v14i3.pp3197-3207.

[9] C. F. Lee, "Fundamental analysis, technical analysis, and mutual fund performance," in *Handbook of Financial Econometrics, Mathematics, Statistics, and Machine Learning, World Scientific*, 2021, pp. 3001–3058.

[10] G. D. Abuselidze and A. N. Slobodianyk, "Value assessment of shares of corporate issuers by applying the methods of fundamental analysis in the stock exchange market," in *Lecture Notes in Networks and Systems*, vol. 206, Springer, 2021, pp. 25–39. doi: 10.1007/978-3-030-72110-7_3.

[11] J. Guerard and A. Mark, "Earnings forecasts and revisions, price momentum, and fundamental data: Further explorations of financial anomalies," *Handbook of Financial Econometrics, Mathematics, Statistics, and Machine Learning (In 4 Volumes)*, vol. 5, pp. 1151–1209, 2020, doi: 10.1142/9789811202391_0030.

[12] M. S. Saeed et al., "An efficient boosted C5.0 decision-tree-based classification approach for detecting non-technical losses in power utilities," *Energies*, vol. 13, no. 12, p. 3242, 2020, doi: 10.3390/en13123242.

[13] M. C. Wu, S. Y. Lin, and C. H. Lin, "An effective application of decision tree to stock trading," *Expert Systems with Applications*, vol. 31, no. 2, pp. 270–274, 2006, doi: 10.1016/j.eswa.2005.09.026.

[14] A. Kurani, P. Doshi, A. Vakharia, and M. Shah, "A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting," *Annals of Data Science*, vol. 10, no. 1, pp. 183–208, 2023, doi: 10.1007/s40745-021-00344-x.

[15] A. M. Priyatno, W. F. R. Sudirman, and R. J. Musridho, "Feature selection using non-parametric correlations and important features on recursive feature elimination for stock price prediction," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 14, no. 2, pp. 1906–1915, 2024, doi: 10.11591/ijece.v14i2.pp1906-1915.

[16] M. Alloghani, D. Al-Jumeily, J. Mustafina, A. Hussain, and A. J. Aljaaf, "A systematic review on supervised and unsupervised machine learning algorithms for data science," in *Supervised and Unsupervised Learning for Data Science*, Springer, 2020, pp. 3–21. doi: 10.1007/978-3-030-22475-2_1.

[17] G. S. Atsalakis and K. P. Valavanis, "Surveying stock market forecasting techniques - Part II: Soft computing methods," *Expert Systems with Applications*, vol. 36, pp. 5932–5941, 2009, doi: 10.1016/j.eswa.2008.07.006.

[18] L. A. Laboissiere, R. A. S. Fernandes, and G. G. Lage, "Maximum and minimum stock price forecasting of Brazilian power distribution companies based on artificial neural networks," *Applied Soft Computing Journal*, vol. 35, pp. 66–74, 2015, doi: 10.1016/j.asoc.2015.06.005.

[19] J. Pesantez-Narvaez, M. Guillen, and M. Alcañiz, "Predicting motor insurance claims using telematics data—XGboost versus logistic regression," *Risks*, vol. 7, no. 2, 2019, doi: 10.3390/risks7020070.

[20] Z. Dai, X. Dong, J. Kang, and L. Hong, "Forecasting stock market returns: New technical indicators and two-step economic constraint method," *North American Journal of Economics and Finance*, vol. 53, p. 101216, 2020, doi: 10.1016/j.najef.2020.101216.

[21] E. H. Houssein, M. Dirar, L. Abualigah, and W. M. Mohamed, "An efficient equilibrium optimizer with support vector regression for stock market prediction," *Neural Computing and Applications*, vol. 34, no. 4, pp. 3165–3200, 2022, doi: 10.1007/s00521-021-06580-9.

[22] N. Naik and B. R. Mohan, "Optimal feature selection of technical indicator and stock prediction using machine learning technique," in *Communications in Computer and Information Science*, 2019, vol. 985, pp. 261–268. doi: 10.1007/978-981-13-8300-7_22.

[23] K. K. Yun, S. W. Yoon, and D. Won, "Prediction of stock price direction using a hybrid GA-XGBoost algorithm with a three-stage feature engineering process," *Expert Systems with Applications*, vol. 186, p. 115716, 2021, doi: 10.1016/j.eswa.2021.115716.

[24] H. Gunduz, "An efficient stock market prediction model using hybrid feature reduction method based on variational autoencoders and recursive feature elimination," *Financial Innovation*, vol. 7, no. 1, pp. 1–24, 2021, doi: 10.1186/s40854-021-00243-3.

[25] Y. Peng, P. H. M. Albuquerque, H. Kimura, and C. A. P. B. Saavedra, "Feature selection and deep neural networks for stock price direction forecasting using technical analysis indicators," *Machine Learning with Applications*, vol. 5, p. 100060, 2021, doi: 10.1016/j.mlwa.2021.100060.

[26] Q. Li, G. Yan, and C. Yu, "A novel multi-factor three-step feature selection and deep learning framework for regional GDP prediction: evidence from China," *Sustainability (Switzerland)*, vol. 14, no. 8, 2022, doi: 10.3390/su14084408.

[27] B. Xue, M. Zhang, W. N. Browne, and X. Yao, "A survey on evolutionary computation approaches to feature selection," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 4, 2016, doi: 10.1109/TEVC.2015.2504420.

[28] C. Zhang, Y. Xue, F. Neri, X. Cai, and A. Slowik, "Multi-objective self-adaptive particle swarm optimization for large-scale feature selection in classification," *International Journal of Neural Systems*, vol. 34, no. 3, pp. 1–27, 2024, doi: 10.1142/S012906572450014X.

[29] V. Bolón-Canedo, N. Sánchez-Maroño, and A. Alonso-Betanzos, "A review of feature selection methods on synthetic data," *Knowledge and Information Systems*, vol. 34, no. 3, pp. 483–519, 2013, doi: 10.1007/s10115-012-0487-8.

[30] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, "Deep feature extraction and classification of hyperspectral images based on convolutional neural networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 10, pp. 6232–6251, 2016, doi: 10.1109/TGRS.2016.2584107.

[31] H. I. Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P. A. Muller, "Deep learning for time series classification: a

review," *Data Mining and Knowledge Discovery*, vol. 33, pp. 917–963, 2019, doi: 10.1007/s10618-019-00619-1.
[32] X. Li and P. Wu, "Stock price prediction incorporating market style clustering," *Cognitive Computation*, vol. 14, no. 1, pp. 149–166, 2022, doi: 10.1007/s12559-021-09820-1.
[33] J. Chun, J. Ahn, Y. Kim, and S. Lee, "Using deep learning to develop a stock price prediction model based on individual investor emotions," *Journal of Behavioral Finance*, vol. 22, no. 4, pp. 480–489, 2021, doi: 10.1080/15427560.2020.1821686.
[34] S. Kumar Chandar, "Grey wolf optimization-Elman neural network model for stock price prediction," *Soft Computing*, vol. 25, no. 1, pp. 649–658, 2021, doi: 10.1007/s00500-020-05174-2.
[35] B. Charbuty and A. Abdulazeez, "Classification based on decision tree algorithm for machine learning," *Journal of Applied Science and Technology Trends*, vol. 2, no. 1, pp. 20–28, 2021, doi: 10.38094/jastt20165.
[36] J. Xianya, H. Mo, and L. Haifeng, "Stock classification prediction based on spark," *Procedia Computer Science*, vol. 162, pp. 243–250, 2019, doi: 10.1016/j.procs.2019.11.281.
[37] A. Safari and A. A. Ghavifekr, "International stock index prediction using artificial neural network (ANN) and Python programming," in *2021 7th International Conference on Control, Instrumentation and Automation*, 2021, pp. 1–7. doi: 10.1109/ICCIA52082.2021.9403580.
[38] D. Sharma, S. K. Singla, and A. K. Sohal, "Stock market prediction using ARIMA, ANN and SVR," in *Lecture Notes in Electrical Engineering*, vol. 698, Springer, 2021, pp. 1081–1092. doi: 10.1007/978-981-15-7961-5_100.
[39] Z. Peng, Q. Huang, and Y. Han, "Model research on forecast of second-hand house price in chengdu based on XGboost algorithm," in *2019 IEEE 11th International Conference on Advanced Infocomm Technology*, pp. 168–172, 2019, doi: 10.1109/ICAIT.2019.8935894.
[40] Y. Yang, Y. Wu, P. Wang, and X. Jiali, "Stock price prediction based on XGBoost and LightGBM," in *E3S Web of Conferences*, 2021, vol. 275, p. 1040. doi: 10.1051/e3sconf/202127501040.
[41] P. R. Srivastava, Z. Zhang, and P. Eachempati, "Deep neural network and time series approach for finance systems: predicting the movement of the Indian stock market," *Journal of Organizational and End User Computing*, vol. 33, no. 5, pp. 1–24, 2021, doi: 10.4018/JOEUC.20210901.oa10.
[42] A. Thakkar and K. Chaudhari, "A comprehensive survey on deep neural networks for stock market: The need, challenges, and future directions," *Expert Systems with Applications*, vol. 177, 2021, doi: 10.1016/j.eswa.2021.114800.
[43] W. Lu, J. Li, J. Wang, and L. Qin, "A CNN-BiLSTM-AM method for stock price prediction," *Neural Computing and Applications*, vol. 33, no. 10, pp. 4741–4753, 2021, doi: 10.1007/s00521-020-05532-z.

BIOGRAPHIES OF AUTHORS

Sneha S. Bagalkot currently works as assistant professor in the Department of Computer Science and Engineering, B.M.S. College of Engineering, Bengaluru and is pursuing Ph.D. in VTU, Belagavi in the Research Center, Department of Computer Science and Engineering, Nagarjuna College of Engineering and Technology, Bangalore, in the machine learning and data analytics domain. She has about 12 years of teaching experience. She can be contacted at email: snehasb.cse@bmsce.ac.in.

Dinesha H. A. is a serving as professor and dean (RD), at Shridevi Institute of Engineering and Technology. He is a fledgling idealistic entrepreneur indulged passionately in serving and uplifting society through advance research, innovation, and knowledge imparting in the arena of engineering and technology. Email id: sridini@gmail.com.

Nagaraj Naik earned his B.E. degree in Computer Science Engineering from VTU, India, in 2006, and his M.Tech. degree in 2012. He completed his Ph.D. in 2021 in the Department of Information Technology at the National Institute of Technology Karnataka, India. His research focuses on financial data analysis, stock price analysis, and statistical techniques. Email id: nagaraj.naik@manipal.edu.