IJCSE

Review Paper

Pattern Recognition and Machine Learning Approach for Stock Trading Decisions: A Review

Sonalika Nayak[1] ⓘ, Jibendu Kumar Mantri[2*] ⓘ, Prasanta Kumar Swain[3] ⓘ

[1,2,3]Department of Computer Application, MSCB University, Baripada, India

Corresponding Author: jkmantri@gmail.com

***Abstract*:** Stock Trading Decisions are important in selection of the right stock at the right time. There are traditional and regular methods for identifying superior stocks for investment but looking into volatility of current market scenario, new technologies must be incorporate to accomplish the target. Here, we presented a review on use of pattern recognition approach and machine learning techniques for Stock Trading Decisions. Usually common patterns are seen in the buying and selling data of stocks for a specific business house. Analysing these data patterns with the use of machine learning approach will produce a better result for Trading Decision. Different machine learning models has been built and applied by different authors to achieve better stock trading decisions.

***Keywords*:** Pattern, Candlestick, ANN, CNN, Open, Close, High, Low.

1. Introduction

Due to high degree of volatility and high noise [1], the stock trading decisions are need tricky approach. A feasible prediction helps the investor in appropriate trading decision and yield benefit with little risk. Therefore, prediction of trend in stock market by using different intelligent algorithms increases the profit ratio of investors. But it is an ever challenge to determine the stock market trend because market is strongly influenced by sentiment of trader, performance of share companies, news and social media.

There are many statistical and machine learning tools developed for prediction of trend followed in stock market. Examples are: Artificial Neural Network, Statistical regression model, genetic algorithm, Support vector machine etc.[2]. One of the most important tools for stock trading decision is candlestick pattern recognition. Candlestick patterns generated from the price factors of stock market with open, high, low and close parameters shows the difference in demand and supply along with investors sentiment [3, 4]. The profitability in trading by using candlestick method is further established [5, 6]. Complex candlestick patterns are also used for prediction of trends in latest researches. The predictive power of 2-way and 4-way patters are examined and studied [7, 8].

Current research shows the effective use of machine learning methods in the prediction of financial market. In the same series use of candlestick pattern for with machine learning algorithms are also established [9, 10]. The Convolutional Neural Network (CNN) model is an established model for image identification and prediction. The candlestick pattern images are made input to the model and trained to predict the future trend for stock trading decision [11].

Remaining of the paper is presented as follows: Section 2 discussed the literature review. The comparative study of various techniques used candlestick patterns is shown in section 3. Section 4 concludes the paper.

2. Related Work

In this review article 10 numbers of research papers were selected and all are based on a single platform of using candlestick pattern images with machine learning technique to predict future market trend.
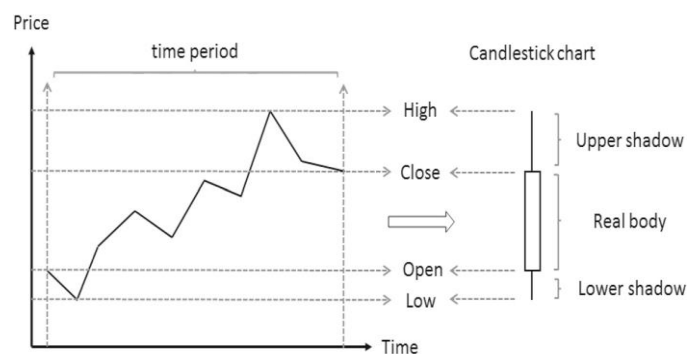

Figure.1. Candlestick patterns showing user information

Paper 1: In this paper the authors proposed a novel pattern recognition model for candlestick called PRML (pattern recognition using machine learning technology). Machine

learning technology is used to improve stock trading decisions. Here 11 types of features for daily stock market patterns and four types of machine learning algorithm are used to make all possible combinations to make the pattern recognition schedule [12].
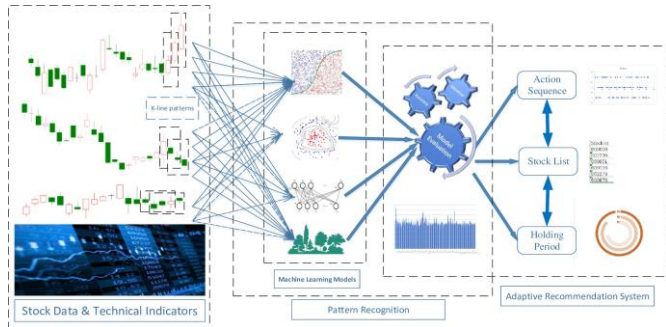

Figure 2. Overview of PRML model [12].

The overview of the methodology adapted in PRML is shown in figure 2. The prediction results of this model shows two-day and three-day patterns for one day in advance forecast and found to be profitable.

Paper 2: The performance evaluation study for CNN and LSTM is made in paper 2. It find the common chart patterns in a stack of historical data of stock market. Defining the methodology the LSTM and 1D CNN use same type of input. 2D CNN model with low recall rate cannot provide better detection rate than hard coded algorithm. The summery is shown in table-1[13].

Table-1: Recall rate.

Algorithm	Recall	Generalization
LSTM	0.97	0.3%
2D CNN	0.73	—
1D CNN	0.64	—

Paper 3: The patterns traced in high dimensional data are difficult to identify because it is not easy to visualize the data. Machine learning methods can be used for data classification and prediction from different data sets. In this paper tactics are made to identify patterns hidden in historical data and predict the long term values [14].
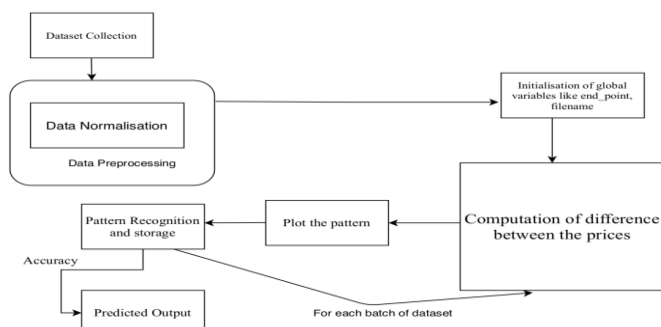

Figure. 3. Methodology followed [14].

Here, identification of patterns which are new is carried out by matching the new one with existing one and the similarity is checked with the fixation of threshold value.

Paper 4: This paper proposed a new deep learning model by integrating stock market data and candlestick chart patterns so that a optimal dynamic forecast may be made for investment [15]. A reward function is also mapped to handle investment risk.
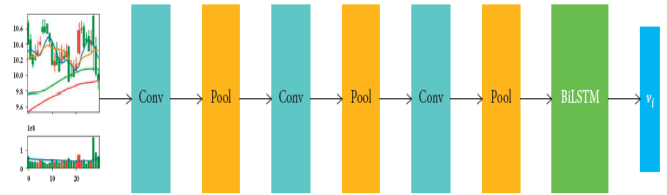

Figure.4. Network structure for extracting features of candlestick chart [15].

Paper 5: The presences of noise make it difficult for analysing and forecasting stock market and to take trading decision. Traders who take help of technical analysis to take decision for market investment face difficulty in identifying candlestick patterns quickly and minutely. Here the neural network model (ANN and CNN) is trained with candlestick images and find patterns [16].
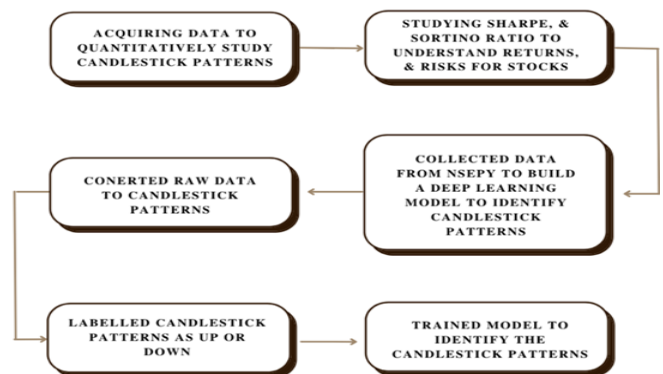

Figure .5. Methodology followed in paper-5 [16].

Paper 6: The primary focus of share traders is to place at the accurate time and direction by the help of appropriate future decision of the financial chaotic series which contain price description of stock such as open, close, high and low of a specific period. In this paper authors make a study and proposed an extensible architecture software framework using object-oriented approach and factory patterns for generating candlestick charts and use them to make a smart algorithm. [17].

Figure.6. Proposed approach.

The process proposed here consists of four stages. The first stage is about creation of object oriented framework for 18 candlestick patterns. The second stage coding was made to determine the candle type based on OHLC (open, high, low, and close). In third stage the data set is marked with type of candle pattern. In fourth stage the forecasting is done by developing algorithmic strategies with candlestick patterns.

Paper 7: This paper represents the use of candlestick pattern as a part of the tool for technical analysis and help in predicting market trend for investment using Artificial Neural Network (ANN).

Parameters used in ANN modelling uses a layer concept such as input, hidden and output layer, optimizer, loss function, metrics, epoch, batch size and learning rate. The accuracy of prediction is more than 70% and highest accuracy is 85.96% [18].

Paper 8: Use of single classifier in machine learning algorithm is not efficient in prediction of stock deflection. In this paper a new approach has been developed by combining sentiment features of stock and candlestick charts in prediction of price of a stock for a period of four, six and ten days. This paper proposed a joint network where classification for sentiment analysis is done in one branch of 1D CNN. Image classification for candlestick chart pattern is done using the 2D CN. The output of the two branches is joined and fed to dense layer in CNN and predict future trend of stock market in near future [19].
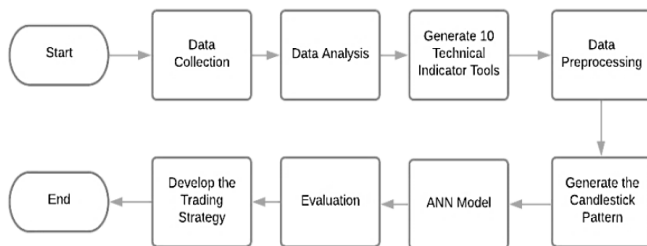
Paper 9: In this paper the candlestick charts are taken from moneycontrol.com. The authors of this paper use the old share prices taken from stock exchange of India to predict the future stock price trend and appropriate decision. Comparison with actual price decides the correct or wrong decisions. Here the sample sizes are considered for analysis is 30 companies and one year of data on share prices [20].

Paper 10: Here an automatic pattern recognition system for candlestick images is built in a two way approach. In the 1st step candlestick patterns are created form time series data using Gramian Angular Field (GAF). In second step eight important candlestick charts are determined by using CNN with GAF pictures. The proposed method GAF-CNN achieves an accuracy of 90.7% average [21].
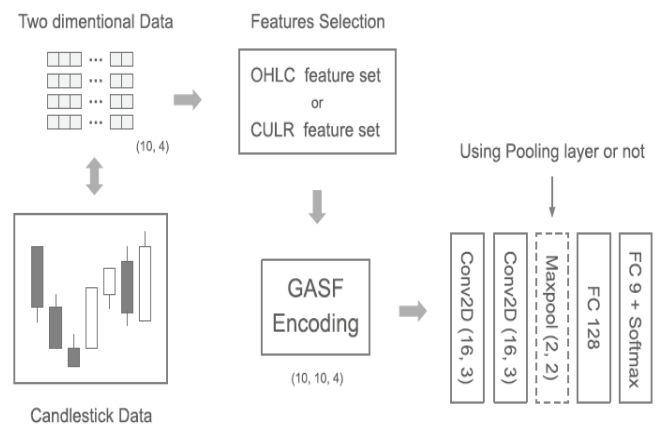

Figure.8. The workflow of the entire experiment [21].

Figure.7. Flowchart of the study [18].

3. Comparative Study

Sl. No	Techniques	Parameter Used	Advantages	Disadvantages	Future scope
1	1. Candlestick pattern recognition model 2. Machine learning Models LR KNN RF RBM	For LR (Logistic Regression): L2 termed used for regularized and warn termed as solver parameter. C=1, No. of Iteration=100 Stopping criterion = 0.0001. For KNN (k-Nearest Neighbors): Range(1,10) for neighbors. parameters for algorithm = auto, ball_tree, kd_tree, brute. size of leaf= range(1,2) CV=10 For RBMv(Restricted Boltzmann Machine): No. of iteration is 10,C=6000, Components is 100, Learning-rate set to 0.06. For RF (Random Forest): GridSearchCV optimizer, CV=10 Maximum depth range(1,10) Minimum sample leaf(2,4,6,50) Criterion(gini, entropy) estimators range(10,100,5)	Pattern recognition and machine learning method are used for Candlestick patterns. The predicted result is more accurate than the application of simple ML methods.	Complex candlestick patterns are not considered.	Deep learning models may be used to get better prediction performance
2	1. CNN	Old Candlestick patterns used by	Higher Detection rate is	Models of 1D and	Diverse patterns are to

	2. LSTM	traders for buy and sell. They used four factors like: Open, High, Low and Close.	obtained using LSTM model.	2D CNN could not perform for better accuracy.	be identified. Multi-objective learning may be deployed.
3	1. Pattern Recognition 2. Accuracy Computation using Backtest	Candlestick graphs with parameters like High, Low, Open, Close.	Proposed model detects the pattern found in the old stock market data and predict the future pattern.	The stock data are uncertain in nature, hence accuracy calculated is nearly 80%.	The prediction accuracy can be increased by including news, sentiment analysis.
4	In the proposed model candlestick charts and stock trading data are combinely analyzed and deep reinforcement learning model is applied to predict future stock trends.	Basic stock data : open, close, low, high price, trading volume. The parameters used as indicator are: KDJ, BIAS, RSI, WILLR, MACD, EMA, DIFF and DEA.	The proposed method helps the investor to generate more profit under same risk situation. The performance of the model is better when stocks showing different trends and the average SR value are found to be highest.		As a future scope of the model text information can also be included.
5	Candlestick charts are used to train ANN and find patterns.	As per NSE India and CMIE Prowess stock data have parameters: Open, Close, High, Low and total 16 years back data is collected.	The candle stick patterns obtained from stock market data are evaluated and analyzed using Candles canner software.		In Indian stock equities Deep learning models can be used to identify candle patterns.
6	In first stage, the software framework was created with object-based coding for the 18 wax patterns. In the second stage, one-hot encoding was done to determine the candle Type In the third stage, a data set labeled with candle types was created. In the fourth stage algorithmic buy/sell strategies based on candlestick patterns is developed and compared	opening (Open) and closing (Close) and the highest (High) and lowest (Low) price movements in a single visual, and it is called OHCL	Candlestick patterns for both the bist100 index and global financial assets, revealed the success of an algorithmic trading strategy based solely on candlestick charts.	Only 21 of the most popular patterns were used in the proposed study	To increase the originality number of candlestick patterns may increase.
7	Stock Market information has indicators that follow a pattern. Here the prediction of pattern as candlestick charts is done by using ANN.	Indicators like SMA, WMA, Momentum, Stochastic K%, Stochastic D%, RSI, MACD, Larry William's R%, (A/D) Oscillator, CCI are calculated using values of input parameters open, close, volume, high, low, adj close.	Score generated from confusion matrix and k-Fold Cross-Validation helps to evaluate the performance. Inclusion of ANN is found to be useful for prediction of candlestick pattern by using the indicators.		
8	Sentiment analysis data from social media like twitter are input into the Natural Language software and convert the old stock trading data in to candle stick patters. Also use CNN technique to predict the future stock trading pattern.	Date and stock wise open, close, high, low, volume data.	The result obtained from combined network model gives better result than single model using candlestick charts with accuracy 75.38%.		
9	Candlestick chart	opening and closing prices	Correct prediction possibility is worked out as 0.5	The prediction status is out of thirty companies, 15 companies share	Prediction using candle stick charts were reliable and the accuracy can be

18

| 10 | 1) GAF

2) CNN | Opening, high, low, closing, upper shadow, lower shadow, real-body, closing prices. | By using simulation data an accuracy of 92.42% is achieved. | Eight types of candlestick charts are generated out of the experiment conducted. | Apart from regular indicator new indicators like head and bottom included with GAF-CNN to have more candle stick patterns. |

(row above continues from previous page)

| prices are correctly predicted, 11 companies share prices are wrongly predicted, 3 companies will be rejected and 1 company status is unsolved | increased with increase in sample size |

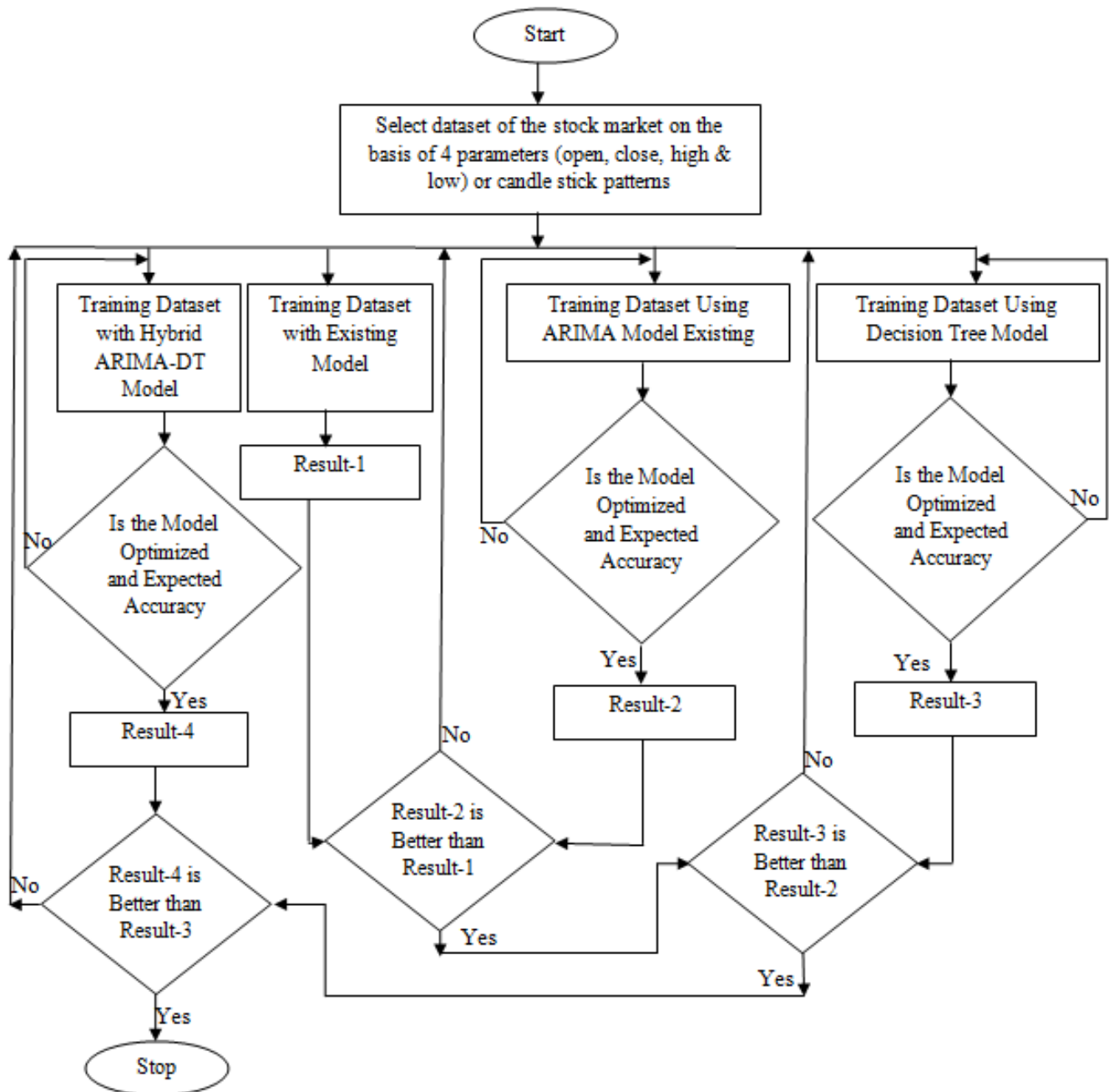## 4. Proposed Model for Prediction of Stock Price

Figure.9. The Flow Diagram for the Proposed Model.

19

However, with these developments of computational intelligent methods for stock market trading decisions, it is possible now to reduce most of the risk and the ability to forecast stock market (NSE and BSE indices) response using ARIMA, Decision Tree and Hybrid ARIMA-Decision Tree model. Hence, a proposed model follows track as shown in figure 9.

6. Conclusion

The stock trading decision according to daily alteration in stock trading and forecasting is an ever challenging task. The machine learning, deep learning and data mining technique shows their effectiveness in forecasting future trends in stock market. Candlestick pattern recognition based forecasting with machine learning and deep learning techniques are found to be more effective over traditional techniques for stock trading decision. In this paper we have reviewed ten research papers where authors use candlestick based pattern recognition technique to establish their effectiveness of their forecasting result. The comparative analysis of techniques and parameters used with advantages, disadvantages are presented in tabular form. In this age of artificial intelligence, this review with the proposed model will enlighten various authors for further analysis in developing new models with candlestick pattern recognition technique integrating with other optimization methods to design a better forecasting model for taking stock trading decision.

References

[1]. T. Fischer, C. Krauss, "Deep Learning With Long Short-Term Memory Networks for Financial Market Predictions", *European Journal of Operation Research*, Vol. **270**, No **2**, pp.**654-669**, **2018**.

[2]. Ding, Xiao, Yue Zhang, Ting Liu, Junwen Duan, "Deep learning for event-driven stock prediction." *In Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)"*, pp.2327-2333, 2015.

[3]. Hendrik. Bessembinder, Kalok Chan, "Market Efficiency And the Returns to Technical Analysis", *Financial Management,* pp.**5-17, 1998**.

[4]. Ben R. Marshall, Martin R. Young, Lawrence C. Rose. "Candlestick Technical Trading Strategies: Can They Create Value for Investors?." *Journal of Banking and Finance,* Vol. **30**, No **8**, pp.**2303-2323**, **2006**.

[5]. T. H. Lu, "The Profitability of Candlestick Charting in the Taiwan Stock Market." *Pacific-Basin Finance Journal,* Vol. **26**, pp. **65-78**, **2014**.

[6]. T. H. Lu, Yi-Chi Chen, Yu-Chin Hsu. "Trend Definition or Holding Strategy: What Determines The Profitability of Candlestick Charting?.", *Journal of Banking & Finance*, Vol. **61**, pp. **172-183**, **2015**.

[7]. Min. Zhu, Said Atri, Eyub Yegen, "Are Candlestick Trading Strategies Effective in Certain Stocks With Distinct Features?", Pacific-*Basin Finance Journal*, Vol.**37**, pp. **116-127**, **2016**.

[8]. Shi Chen, Si Bao, Yu Zhou, "The Predictive Power of Japanese Candlestick Charting In Chinese Stock Market", *Physica A: Statistical Mechanics and its Applications,* Vol. **457**, pp. **148-165, 2016**.

[9]. J. Fock, Henning, Christian Klein, Bernhard Zwergel. "Performance of Candlestick Analysis on Intraday Futures Data", The *Journal of Derivatives*, Vol. **13**, No 1, pp. **28-40**, **2005**.

[10]. Matthieu Duvinage, Paolo Mazza, Mikael Petitjean, "The Intra-Day Performance of Market Timing Strategies And Trading Systems Based on Japanese Candlesticks." *Quantitative Finance*, **Vol. 13**, No **7**, pp. **1059-1070, 2013**.

[11]. Marc'Aurelio Ranzato, Y-Lan Boureau, Yann Cun. "Sparse Feature Learning for Deep Belief Networks", *Advances In Neural Information Processing Systems*, Vol. **20**, **2007**.

[12]. Yaohu Lin, ,Shancun Liu, Haijun Yang, Harris Wu, Bingbing Jiang, "Improving Stock Trading Decisions Based on Pattern Recognition Using Machine Learning Technology", *PloS one*, Vol. **16**, No. **8, 2021**.

[13]. Marc Velay, Fabrice Daniel, "Stock Chart Pattern Recognition With Deep Learning." arXiv preprint arXiv:1808.00418 , **2018**.

[14]. D. Sorna Shanthi, T. Aarthi, A.K. Bhuvanesh, RA. Chooriya Prabha,"Pattern Recognition in Stock Market", *International Journal of Computer Science and Mobile Computing*, Vol. **9**, No **3**, pp.**106 – 111**, **2020**.

[15]. Li. Yawei, Liu. Peipei, Wang. Ze, "Stock Trading Strategies Based on Deep Reinforcement Learning", Scientific Programming, Vol. 2022, pp. 1-15, 2022. https://doi.org/10.1155/2022/4698656

[16]. A. HhUpreti, A. Agrawal, J. K. Joshi, S. Seniaray,(2022). "Quantitative Study of Candlestick Pattern & Identifying Candlestick Patterns Using Deep Learning For The Indian Stock Market", *International Journal of Health Sciences*, Vol. **6** No **S3**, pp. **5739–5749, 2022**.

[17]. Üzeyir AYCEL, Yunus SANTUR, "A New Algorithmic Trading Approach Based on Ensemble Learning And Candlestick Pattern Recognition in Financial Assets", *Turkish Journal of Science and Technology,* Vol. **17**, No **2**, pp. **167-184, 2022**.

[18]. Ni Putu Winda Ardiyanti, Irma Palupi, Indwiarti Indwiarti. "Trading Strategy On Market Stock By Analyzing Candlestick Pattern Using Artificial Neural Network (Ann) Method", *Journal Media Informatika Budidarma,*Vol. **5**, No **4**, pp. **1273-1282, 2021**.

[19]. T, T. Ho, Y. Huang, "Stock Price Movement Prediction Using Sentiment Analysis And Candlestick Chart Representation", *Sensors*, Vol. **21**, No **23**, pp. **7957, 2021.**

[20]. U. GOKUL, "Forecast Share Price Using Technical Analysis Tool", *Pacific International Journal*, Vol. **4**, No **1**, pp. **01-06, 2021.**

[21]. J. H. Chen, Y. C. Tsai, "Encoding Candlesticks As Images For Pattern Classification Using Convolutional Neural Networks", *Financial Innovation*, Vol. **6**, No **1**, pp. **1-19, 2020.**

AUTHORS PROFILE

Ms. Sonalika Nayak currently pursues Ph.D. in Computer Science and IT at Maharaja Sriram Chandra Bhanja Deo University, Baripada, Odisha. Her main research work focuses on Data analytics, Machine learning related to stock market data.

Dr. J.K. Mantri is working as an Associate Professor in the Department of Computer Application, Maharaja Sriram Chandra Bhanja Deo University, Baripada, Odisha, India. He has more than 28 years of experience of teaching and research. His area of specialization includes: AI, Business Process Re-Engineering, and Computer Security. He has authored 6 books and having chapters in edited books. He has also authored more than 87 papers published in refereed International / National Journals and Conferences. He is an active reviewer for peer-reviewed journals and has delivered invited talks pertaining to his research field.

Dr. P.K. Swain is working as Assistant Professor in Department of Computer Application, Department of Computer Application, Maharaja Sriram Chandra Bhanja Deo University, Baripada, Odisha, India.. He has published more than 15 research papers in reputed international journals and conferences. His main research work focuses on wireless sensor network, mobile computing and IoT and Data analytics. He has 16 years of teaching and research experience.