International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 03 | March – 2026 DOI:10.55041/ISJEM.ACME010
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Design and Implementation of Hybrid Stock Market Analysis and Prediction using Dynamic Dataset

K M Thanuja[1], Dr. N.Deepak Kumar[2], Dr.M.Giri[3]

[1]P.G Scholar, Department of CSE,Sree Rama Engineering College, Tirupati, Andhra Pradesh, India, naiduthanuja4@gmail.com

[2] Professor, Department of Computer Science & Engineering, Sree Rama Engineering College, Tirupati, Andhra Pradesh, India, deepakkumarsvuphd@gmail.com

[3]Professor, Department of CSE (AI), Mother Theresa Institute of Engineering and Technology, Palamaner, Andhra Pradesh, India, dr.m.giri.cse@gmail.com

---

Abstract - Stock market data is in nonlinear fashion, heavy fluctuated data, time series data, present heavy noisy data, and handling such a data is a one of the crucial tasks. Traditional stock prediction approaches using ML approaches were exhibits moderate results and not appropriate to handle heavy fluctuated and noisy data samples related to stock market. Many of these techniques are depends on short term fluctuations in the stocks and many be representing wrong stock predictions. In this research we developed Hybrid Stock Market Analysis and Prediction approach, capable to handle heavy fluctuated stock market dynamic dataset to enhance prediction, feature encryption techniques are used to convert continuous data samples into consistent representation, and the model can be easily predicting the changes in the market trend. All the data samples are processed using hybrid ML approach for modelling based on temporal dependencies and the segments of time are prioritize with help of attention technique. Performance of proposed model is evaluated and compared with LSTM, MLP, RF, SVM, and LR approaches are evaluated in terms of accuracy, precision, recall, F1-score, FPR, FNR, MCC, and AUC values. Proposed model exhibits greater accuracy, exhibits greater precision, exhibits greater recall, exhibits greater F1-score, exhibits low FPR, exhibits low FNR, exhibits greater MCC, and exhibits greater AUC compare with other similar techniques.

Key Words: Machine Learning, Deep Learning, LSTM, Random Forest, Support Vector Machine, MLP, FPR, FNR, Logistic Regression

1.INTRODUCTION

In general stock prices are heavy noisy, continuous, dynamic, timely changing data samples. Predicting stock movements is a one of the crucial and challenging tasks due to heavy fluctuations, uncertainty exists, and many parameters influence the stock prices that are users trust on a particular industry, politics, and news about stock market [1]. Now a days predicting stock prices is gaining greater popularity due to successful prediction of stock prices may guide many investors to invest money on appropriate stocks [2]. The prediction probability is more then only the investor may get profits, if the model predicts consistently movement of stock price that will be helpful to investor to take decisions on investment [3].

A stock price in the market is the data gathered from buying and selling of stocks over a period of time. Prediction of stocks may help to analyze the current market of stock price and selection of appropriate techniques to study market is place equal role [4]. Many financial organizations and people who are performing trading has generated many trading applications to analyze the stock price and suggest to their customers, but unfortunately all these applications suggest only mean returns [5]. A stock price of any industry is just a value calculated by determining the expectation of cash transactions of a company in future and not considers current position of the industry in the market [6].

The present stock value of a company in the market can be deviated from original company stock value due to many unwanted causes on the basic operations of industry and completely it is decided based on opinion about the company on public [7]. suppose heavy fluctuation in the market due to many reasons, experts and knowledge

engineers only can able to analyze the correct stock value and these are helpful to predict stock of a company in future [8]. many of the investors depends purely on their luck, the advanced stock trading applications facilitate many services to analyze future stock price of industry and automatically helps stock analysts with more parameter values to perform better stock predictions [9].

Authors in [10] described how prices of stock impact on financial position of industry, prediction of any stock is completely based depends on information that are available from various sources, all these information is classified into two categories, one if qualitative (posts, news articles, announcements on stocks, annual reports on industry financial positions), and second one is quantitative (turnover, stock prices). Stock prices are affected on many parameters like sentiments of investors, news articles, and volume of stock trade [11]. In order to predict stocks depending on single source of information is not enough and before going to decision about stocks it is mandatory to take information from multiple sources (social media blogs, stock related information from newspapers, public talk about company stocks, and so on) [12].

2. Related Work

Authors in [13] developed a new model based on RF approach to analyze stock prices are heavily fluctuated data, users invest their money into a particular company stock based on some predictive decisions, in order to take such a decisions user may follow some apps or trading tools to gain more profits. Authors in [14] introduced NN approach to analyze local parameters, previous stock details are recollected with help of neurons, their model is capable to analyze continuous data samples, in addition to NN authors used PSO technique to optimize weights of NN, and provide suggestions how to predict opening price of a stock price. Authors in [15] determined a new approach by combining LSboost with RF, experiments are conducted with one-decade historical stock prices, the results are also compared with SVM, and explained how to predict closing value of stock price.

Authors in [16] focused social media information related to stock prices, studied how the people attracted towards a particular company stocks, authors used data mining methodologies to determine decisions on stock prices. Authors in [17] used a new ANN approach to predict closing value of a stock in the market, their model achieved less computing time, occupied limited memory, and also studied peak value of stocks, number of volumes sold or buy per day, lowest value of stock, and their model achieved acceptable performance. Authors in [18] conducted detailed analysis on stock prices, they mainly concentrated on stock market in India, authors used naive Bayes classification to decide the tweets on stock is either positive or negative, and their model shows moderate robustness while predicting stock prices.

Authors in [19] implemented a new stock prediction hybrid approach by combining PCA, DNN, and auto encoding approaches. their model accepting large volume of stock related data samples to analyze present stock trends, and they are tried to predict stock prices under heavy fluctuated stock data. Authors in [20] discussed LSTM approach and achieved best results than RF, DNN, and LR classification, stock price data were fully studied under three different phases, stock fluctuations are observed with long term stock market prices, the model predicts up and down in the stock prices, and shows acceptable accuracy while predicting stock prices. Authors in [21] elaborated CNN approach, their model applied to historical stock price data samples, the news spread by many experts are also taken into consideration in this process, their model predicts peak value of stock and average price of stock, and achieved better performance than other similar kinds of approaches.

Authors in [22] used ML based approach for predicting stock prices with different types of features by combining RF and decision trees, showed enhanced accuracy between different stock prices but robustness of their model varied based on market conditions. Authors in [23] discussed a new approach to analyses stock prices with available historical data samples, authors decided that prediction of stocks changes continuously is a one of the challenging and difficult tasks, and shows less robustness to predict stock prices under heavy fluctuations. Authors in [24] introduced a new method how prices of stock impact on financial position of industry and prediction of any stock is completely based depends on information that are available from various sources.

3. Hybrid Stock Prediction Approach

In this research we developed hybrid Stock Market Analysis and Prediction approach, capable to handle heavy fluctuated stock market dynamic dataset to enhance prediction, feature encryption techniques are used to convert continuous data samples into consistent representation, and the model can be easily predicting the changes in the market trend. All the data samples are processed using hybrid DL approach by combining CNN

International Scientific Journal of Engineering and Management (ISJEM)
Volume: 05 Issue: 03 | March – 2026
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata
ISSN: 2583-6129
DOI:10.55041/ISJEM.ACME010

for identifying local level patterns with LSTM for modeling based on temporal dependencies and the segments of time are prioritize with help of attention technique. The proposed Stock prediction approach architecture is visualized in figure 1.

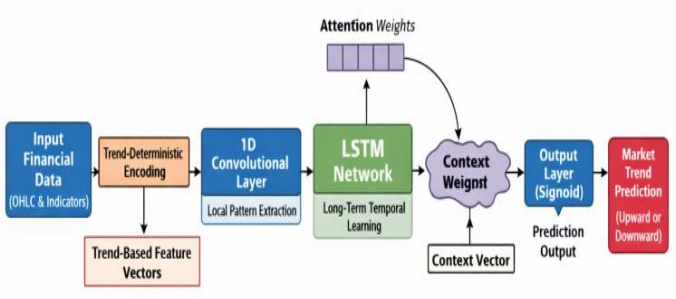

Fig -1: Architecture of proposed Stock Prediction Approach

In our proposed stock prediction approach used historical stock price data, it is a continuous time sample data about stock prices, dataset is generated from couple of decades, and dataset is collected under normal and heavy fluctuated market. The name the dataset used in our approach is OHLC dataset, it is an open source dataset available in Kaggle information repository, and sample in the dataset are recorded with time and price values. From the OHLC dataset discovered more important features like opening price of stock, closing price of stock, highest price of stock, lowest price of stock, and the volume of stock traded. The dataset samples are applied to feature engineering process, eliminate noise from data samples, and all data samples are applied to scaling. Once preprocessing dataset is completed apply split technique to partition dataset into seventy percent for training, fifteen percent for testing, and fifteen percent for validation. Hyper parameters in proposed stock prediction approach are tuned with {Size of temporal window =20}, {filters=64}, {number of_ parameters=12}, {Pool size=2}, {kernel size=3}, {units hidden=128}, {Layers in LSTM=1}, {dimension in attention =64}, {dropout=0.2}, {learning rate=0.005}, {Epochs =50}.

4. Performance Evaluation

Performance of proposed, CNN (Convolution Neural Network), LSTM (Long Short-Term Memory), MLP (Multilayer Perceptron), XGBoost (Extreme Gradient Boosting), RF (Random Forest), SVM (Support Vector Machine), LR (Logistic Regression) are evaluated in terms of accuracy, precision, recall, F1-score, FPR (False Positive Rate), FNR (False Negative Rate), MCC (Matthews Correlation Coefficient), and AUC (Area Under ROC Curve). Accuracy of proposed approach is calculated and then compared with CNN, LSTM, MLP, XGBoost, RF, SVM, LR methods, SVM and LR shows moderate accuracy because not able to handle dynamic temporal data samples, ensemble technique were exhibits acceptable accuracy, our model exhibits greater accuracy (95.6 percent), the accuracy values of different approaches are visualized in figure 2.

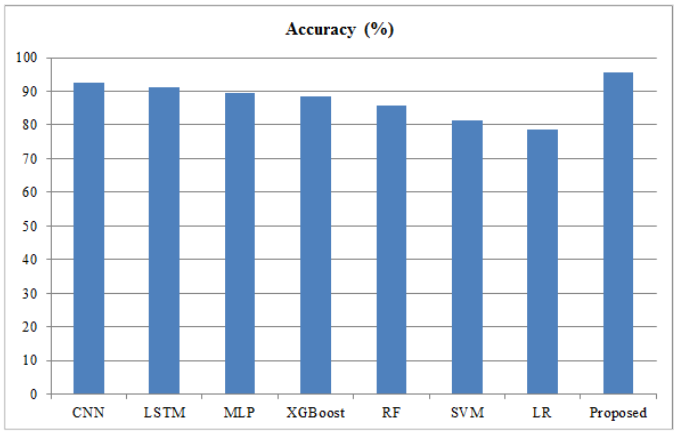

Fig. 1. Accuracy values of different Prediction Models

Precision of proposed approach is calculated and then compared with CNN, LSTM, MLP, XGBoost, RF, SVM, LR methods, traditional approaches exhibits low precision values because not able handle heavy fluctuated data samples, ML and DL approaches optimize false detections and shows moderate precision values, our model exhibits greater precision (94.8 percent) compare with other similar techniques, the precision values of different approaches are visualized in figure 3.

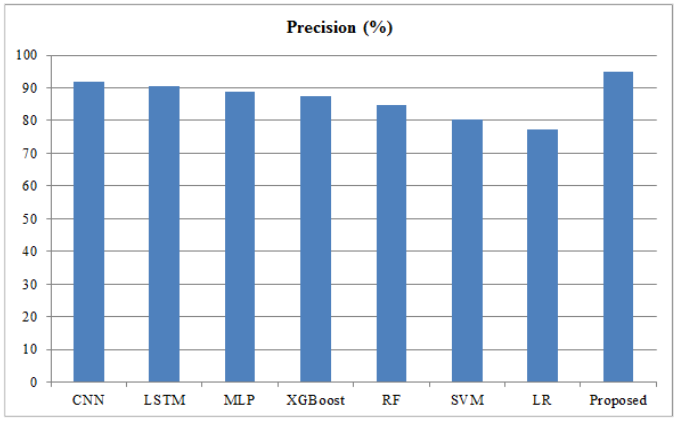

Fig. 2. Precision values of different Prediction Models

Recall of proposed approach is calculated and then compared with CNN, LSTM, MLP, XGBoost, RF, SVM, LR methods, traditional approaches exhibits low recall values because not able handle heavy fluctuated data samples, ML and DL approaches optimize false detections and shows moderate recall values, our model exhibits greater recall (94.1 percent) compare with other similar techniques, the recall values of different approaches are visualized in figure 4.

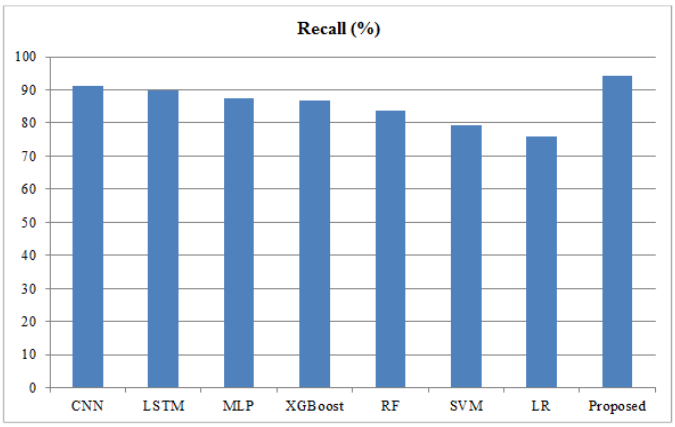

Fig. 3. Recall values of different Prediction Models

F1-score of proposed approach is calculated and then compared with CNN, LSTM, MLP, XGBoost, RF, SVM, LR methods, and traditional approaches exhibits low F1-score values because not able handle heavy fluctuated data samples.

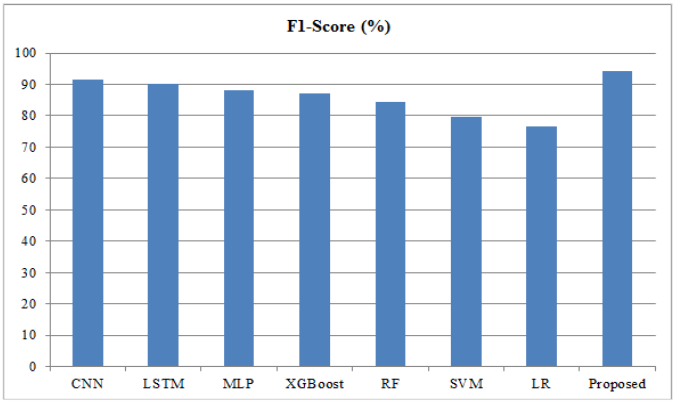

Fig. 4. F1-Score values of different Prediction Models

ML and DL approaches optimize false detections and shows moderate F1-score values, our model exhibits greater F1-score (94.1 percent) compare with other similar techniques, the F1-score values of different approaches are visualized in figure 5. FPR represents count of wrong positive predictions. FPR of proposed approach is calculated and then compared with CNN, LSTM, MLP, XGBoost, RF, SVM, LR methods, traditional approaches exhibits greater FPR values because not able handle heavy fluctuated data samples with noise, ML and DL approaches optimize false detections and shows moderate FPR values, our model exhibits low FPR (5.7 percent) compare with other similar techniques, the FPR values of different approaches are visualized in figure 6.

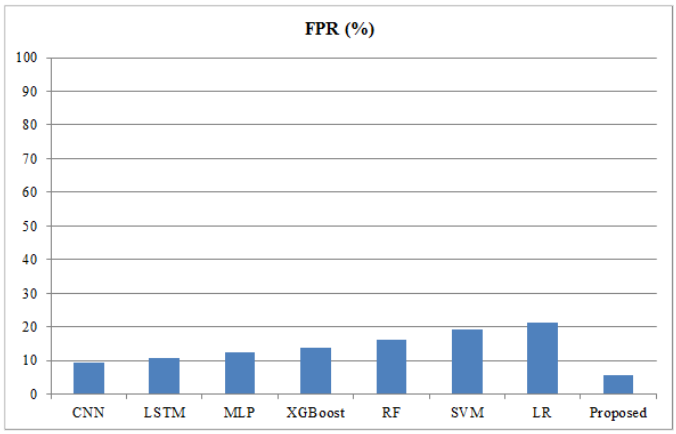

Fig. 5. FPR values of different Prediction Models

FNR represents count of missing correct predictions. FNR of proposed approach is calculated and then compared with CNN, LSTM, MLP, XGBoost, RF, SVM, LR methods, traditional approaches exhibits greater FNR values because not able handle heavy fluctuated data samples with noise, ML and DL approaches optimize false detections and shows moderate FNR values but facing certain issues while handling failure transactions, our model exhibits low FNR (5.9 percent) compare with other similar techniques, the FNR values of different approaches are visualized in figure 7.

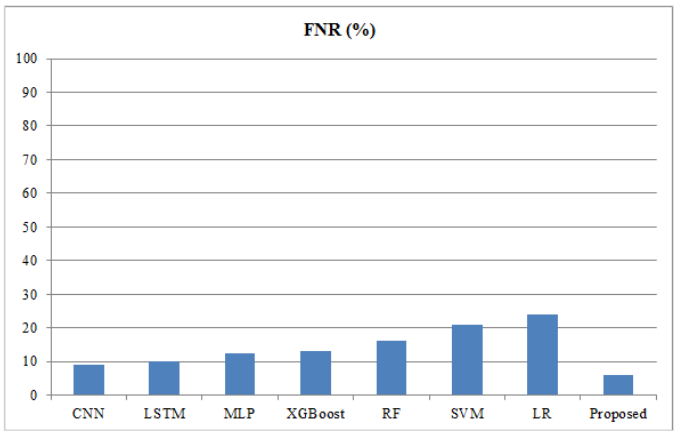

Fig. 6. FNR values of different Prediction Models

MCC represents calculation of class labels under imbalanced distribution of classes. MCC of proposed approach is calculated and then compared with CNN, LSTM, MLP, XGBoost, RF, SVM, LR methods, traditional approaches exhibits low MCC values because not able

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 03 | March – 2026 DOI:10.55041/ISJEM.ACME010
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

handle heavy fluctuated data samples with noise, ML and DL approaches optimize false detections and shows moderate MCC values but facing certain issues while handling failure transactions, our model exhibits greater MCC (0.89) compare with other similar techniques, the MCC values of different approaches are visualized in figure 8.

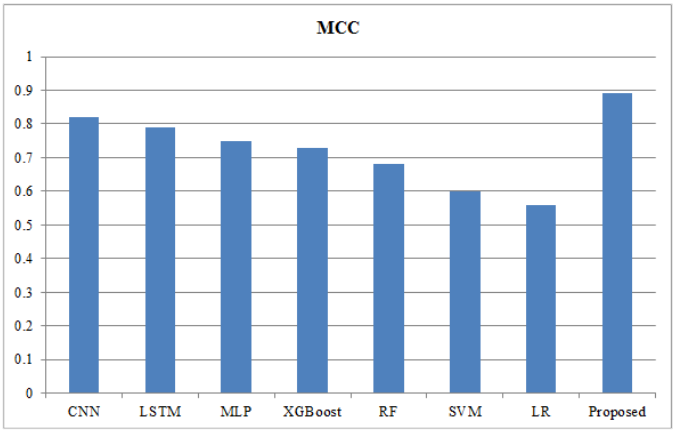

Fig. 7. MCC values of different Prediction Models

AUC represents calculation of class labels under different types of predefined threshold values. AUC of proposed approach is calculated and then compared with CNN, LSTM, MLP, XGBoost, RF, SVM, LR methods, traditional approaches exhibits low AUC values because not able handle heavy fluctuated data samples with noise, ML and DL approaches optimize false detections and shows moderate AUC values but facing certain issues while handling failure transactions, our model exhibits greater AUC (97 percent) compare with other similar techniques, the AUC values of different approaches are visualized in figure 9.

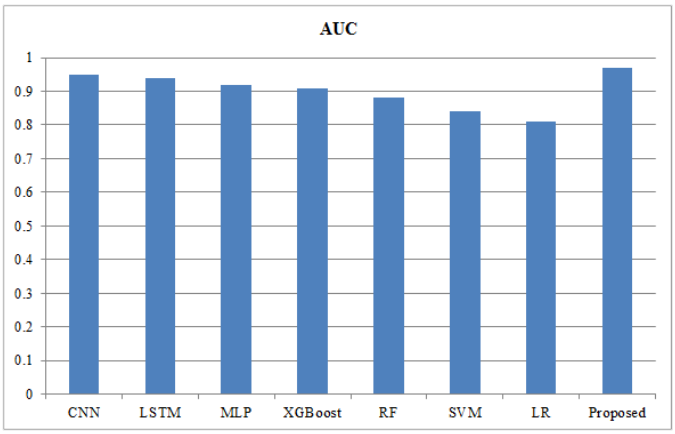

Fig. 8. AUC values of different Prediction Models

Performance of proposed approach is calculated and then compared with CNN, LSTM, MLP, XGBoost, RF, SVM, LR methods, traditional approaches exhibits low performance values because not able handle heavy fluctuated data samples, ML and DL approaches optimize false detections and shows moderate performance values, our model exhibits greater performance values compare with other similar techniques, the performance values of different approaches are visualized in figure 10.

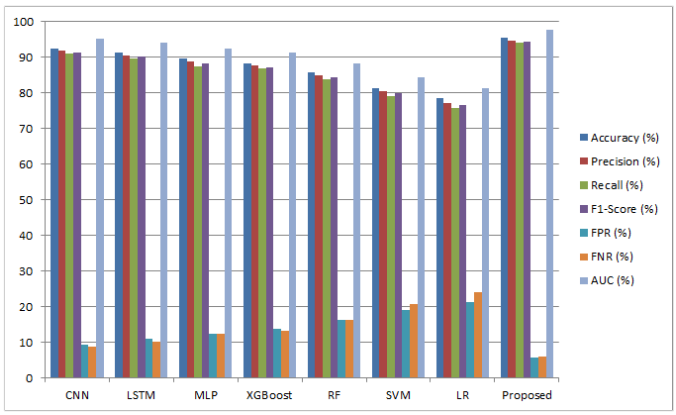

Fig. 9. Performance values of different Prediction Models

5. CONCLUSIONS

Conventional stock prediction approaches using ML and DL approaches were exhibits moderate results and not appropriate to handle heavy fluctuated and noisy data samples related to stock market. In this research we developed Hybrid Stock Market Analysis and Prediction approach, capable to handle heavy fluctuated stock market dynamic dataset to enhance prediction, feature encryption techniques are used to convert continuous data samples into consistent representation, and the model can be easily predicting the changes in the market trend. All the data samples are processed using hybrid DL approach by combining CNN for identifying local level patterns with LSTM for modeling based on temporal dependencies and the segments of time are prioritize with help of attention technique. Performance of proposed method is evaluated and compared with CNN, LSTM, MLP, XGBoost, RF, SVM, and LR approaches are evaluated in terms of accuracy, precision, recall, F1-score, FPR, FNR, MCC, and AUC values. Our model exhibits greater accuracy (95.6 percent), exhibits greater precision (95.6 percent), exhibits greater recall (94.4 percent), exhibits greater F1-score (94.1 percent), exhibits low FPR (5.7 percent), exhibits low FNR (5.9 percent), exhibits greater MCC (0.89), and exhibits greater AUC (97 percent) compare with other similar techniques.

REFERENCES

1. G. Singh, P. Chaturvedi, A. Shrivastava and S. Vikram Singh, "Breast Cancer Screening Using Machine Learning Models," 2022 3rd International Conference on Intelligent Engineering and Management (ICIEM), London, United Kingdom, 2022, pp. 961–967.

2. Ozen, Fatih; Souri, Ali. "Cloud-based disaster management architecture using hybrid machine learning approach in IoT." Multimedia Tools and Applications, 83, 72357–72370, 2024.

3. M. Mallam, M. K. L. Murthy, T. S. Devi, J. V. Suman and S. Rao Polamuri, "Stock Market Price Prediction Using Machine Learning," 2024 Second International Conference on Advances in Information Technology (ICAIT), Chikkamagaluru, Karnataka, India, 2024, pp. 1-4.

4. R. Surekha, M. Giri, M. Kusuma, T. Naveena, E. R. Kalyan Gowd and N. Pavan Tej, "Prediction: Landslides and Floods Detection using Deep Learning Techniques," 2025 2nd International Conference on Circuits, Power and Intelligent Systems (CCPIS), Bhubaneswar, India, 2025, pp. 1-6.

5. L. Mathanprasad and M. Gunasekaran, "Analysing the Trend of Stock Marketand Evaluate the performance of Market Prediction using Machine Learning Approach," 2022 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), Chennai, India, 2022, pp. 1–9.

6. Ajay Kumar Sahu, Pankaj Kumar Gupta, Amit Kumar Singh, Kushal Singh, Ujjwal Kumar, "A Convolutional Neural Network Approach for Detecting the Distracted Drivers ", International Journal of Scientific Research in Computer Science and Engineering, Vol. 11, Issue.l, pp. 1 - 6, 2023.

7. M. Giri, N. Malavika, R. L. Devi, E. Ranjith Kumar, K. Rohith and P. Rushikesh, "Prediction of Accidents by Surveillance and Continuous Monitoring of Road Conditions Using Deep Learning Techniques," 2025 2nd International Conference on Circuits, Power and Intelligent Systems (CCPIS), Bhubaneswar, India, 2025, pp. 1-6.

8. K. Kumari, S. Das, A. Bhowmick, A. Saha, A. B. Puja and K. Karmakar, "Stock Price Prediction Using Machine Learning," 2024 International Conference on Big Data Analytics in Bioinformatics (DABCon), Kolkata, India, 2024, pp. 1-6.

9. V. Tiwari, B. P. Lohani, A. Rana, U. P. Pandey and B. Dhariwal, "Stock Market Prediction using different Machine Learning Algorithms," 2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), Gautam Buddha Nagar, India, 2023, pp. 147-151.

10. B. Math, O. Bevinakatti, S. Hosamani and P. Patil, "Stock Market Prediction Using Machine Learning," 2024 5th International Conference for Emerging Technology (INCET), Belgaum, India, 2024, pp. 1-6.

11. Z. U. R. K and M. Giri, "Enhancing Hotel Services in the Industry 6.0 Era: Optimizing Customer Satisfaction through Sentiment Analysis," 2025 6th International Conference for Emerging Technology (INCET), BELGAUM, India, 2025, pp. 1-6.

12. Indra; Hamdani, Agus Umar; Setiawati, Suci; Mentari, Zena Dwi; Purnomo, Mauridhy Hery. "Comparison of K-NN, SVM, and Random Forest Algorithm for Detecting Hoax on Indonesian Election 2024." Jurnal Nasional Pendidikan Teknik Informatika: JANAPATI, 13(1), 166-179, 2024.

13. Khan, Fahim Ullah; Shah, Ibrar Ali; Jan, Sadaqat; Ahmad, Shabir; Whangbo, Taegkeun. "Machine Learning-Based Resource Management in Fog Computing: A Systematic Literature Review." Sensors, 25(3), 687, 2025.

14. Polamuri, S.R., Srinnivas, K. & Mohan, A.K. Prediction of stock price growth for novel greedy heuristic optimized multi-instances quantitative (NGHOMQ). Int J Syst Assur Eng Manag 14, 353–366, 2023.

15. P. Sukanya, M. Giri, B. H. Krishna, G. Devendra, G. Abhinay and G. H. Kiran, "Prediction: Edge Industrial IoT (EIIoT) Dataset Analysis and Determination of threats by using DNN Technique," 2025 4th International Conference on Advances in Computing, Communication, Embedded and Secure Systems (ACCESS), Ernakulam, India, 2025, pp. 771-776.

16. V. Deswal, D. Kumar and Suman, "Stock Market Price Prediction using Machine Learning Techniques: A Review," 2023 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES), Greater Noida, India, 2023, pp. 693-698.

17. Yash Thorat, Vishal Kumar, Shivani Vichare. "Comparative Study on Machine Learning Algorithms for Stock Price Prediction." International Journal of Financial Studies, 11 (2), 2023.

18. M. Giri, Z. U. Rahaman K, C. B. Giridharacharyulu, K. Ganesh, M. Darshan and P. Bhageeradh, "Implementation

of New Efficient Method to Analyze Texture of Plant Based Alternative Meat Using Machine Learning Techniques," 2025 International Conference on Emerging Technologies in Engineering Applications (ICETEA), Puducherry, India, 2025, pp. 1-5.

19. V. Goyal and R. R. Raj, "Using LSTM, DNN, and ARIMA approaches to predict the trends in the Stock Market," 2022 8th In-ternational Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2022, pp. 918–922.

20. Ganjirad, M.; Delavar, M.R. "Flood Risk Mapping Using Random Forest and Support Vector Machine." ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, X-4/W1-2022, pp. 201-208, 2023.

21. J. Sridhar, M. Giri, K. Gayathri, K. Yamini, K. Yaswanth and B. Syam, "Repairable Resilient Data Search: Forward Secure Attribute Encryption for IoT cloud security," 2025 4th International Conference on Advances in Computing, Communication, Embedded and Secure Systems (ACCESS), Ernakulam, India, 2025, pp. 876-881.

22. Bhadani, Rishita. "Cloud-Based IoT Systems for Disaster Management in Urban Areas." Risk Assessment and Management Decisions, 1(2), 50, 2024.

23. A. K. Sahu, "Stock Market Prediction Using Machine Learning," 2023 International Conference on Computational Intelligence, Communication Technology and Networking (CICTN), Ghaziabad, India, 2023, pp. 512-515.

24. N. Suriya, R. Pandimeena, K. Nisha and V. N. Dhenishaa, "Stock Market Prediction Analysis Using Machine Learning Transformer Model," 2023 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI), Chennai, India, 2023, pp. 1-6.