International Journal of Advanced Trends in Multidisciplinary Research
ISSN: 3139-0986, Volume 2 Issue 7
July 2026, PP 1-6

A Machine Learning Approach for Real-Time Anomaly Detection in 5G Networks

S.Venkateswara Rao[1], B. Sindhuja[2]
[1]Assistant Professor, Department of MCA, Megha Institute of Engineering and Technology for Women, Edulabad (Village), Ghatkesar (Mandal), Medchal District, Telangana, India.
[2]MCA Student, Department of MCA, Megha Institute of Engineering and Technology for Women, Edulabad (Village), Ghatkesar (Mandal), Medchal District, Telangana, India.

***Abstract* – With its ultra-high data speeds, low latency, enormous device connectivity, and better service dependability, fifth-generation (5G) cellular networks have greatly improved wireless communication. These networks have been rapidly deployed. But, the attack surface has grown with the ever-increasing complexity of 5G infrastructures and the ever-increasing volume of network traffic, leaving these networks open to cyber attacks, illegal access, and strange communication patterns. The ever-changing nature of 5G settings makes it difficult for traditional security measures to detect complex assaults as they happen. In light of these difficulties, this research introduces a framework for anomaly identification in 5G cellular networks that is based on machine learning. Various machine learning techniques are used by the suggested system. These techniques include Autoencoders, Random Forest, One-Class Support Vector Machine (One-Class SVM), and two ensemble learning methods. Both ensemble models include AdaBoost, Decision Tree, and Gradient Boosting into a Voting Classifier; however, the second model enhances anomaly detection performance by integrating AdaBoost, Decision Tree, and ExtraTree. When creating the models, we used a Cellular Dataset that included both typical and unusual 5G network data. Data cleansing, normalisation, and label encoding are some of the preprocessing activities that the dataset goes through before training in order to make the data better and the model work better. The suggested ensemble strategies are shown to outperform individual machine learning models in extensive experimental assessment. In complicated 5G communication scenarios, Ensemble 2 outperforms all other techniques with 100% Accuracy, Precision, Recall, and F1-score, proving its usefulness for real-time anomaly identification. To improve network security, reduce cyber threats, and enable intelligent monitoring of next-gen cellular communication systems, this study proposes a framework that is both scalable and dependable, all while maintaining the original machine learning methodology.**

***Keywords* – 5G Cellular Networks, Network Anomaly Detection, Machine Learning, Autoencoder, One-Class SVM, Random Forest, Ensemble Learning, Voting Classifier, Cybersecurity, Network Traffic Analysis.**

I. INTRODUCTION

With its lightning-fast data transfer rates, decreased latency, increased device connectivity, and rock-solid service dependability, fifth-generation (5G) cellular networks have completely changed the game when it comes to contemporary wireless communication. Smart healthcare, the Internet of Things (IoT), cloud computing, intelligent urban infrastructure, industrial automation, and autonomous mobility have all been expedited in their rollout because to these characteristics. A major concern for telecom companies is ensuring the continued availability of their networks and services in the face of the massive amounts of data being sent by 5G communication systems involving millions of diverse devices. As a result, safeguarding 5G networks from ever-changing cyber threats is now a must for guaranteeing reliable and secure communication services.

5G networks, in contrast to earlier iterations of mobile communication, function in virtualised settings that perpetually produce complicated traffic patterns. Even if it makes things more adaptable and scalable, the increasing complexity also increases the surface area that bad actors may exploit. Unauthorised access, unusual traffic patterns, Distributed Denial-of-Service (DDoS) attacks, infiltration into networks, and interruption of services are all examples of cyber risks that might expose critical user information and drastically diminish the quality of communication. Thus, protecting next-generation communication infrastructures is becoming more and more dependent on intelligent systems that can detect anomalous network behaviour in real-time.

If you want to find suspicious activity on your network, you may use traditional methods that depend on established rules, signatures, or thresholds that you set yourself. Despite their efficacy against established attack patterns, these methods often fail to spot emergent or hidden abnormalities in 5G settings on a big scale. In addition, real-time security monitoring becomes impractical due to computationally costly human analysis caused by the ever-increasing amount of network data. Research on Machine Learning (ML) systems that may automatically learn typical network behaviour and spot deviations that signal possible security vulnerabilities has been prompted by these limitations.

An adaptive method for anomaly detection is provided by machine learning, which directly extracts significant patterns from past network traffic data. Machine learning algorithms constantly examine communication behaviour and categorise network events according to learnt features, rather than depending only on predetermined attack signatures. While there are many machine learning algorithms that can detect anomalies, some of the most promising include ensemble learning models, Autoencoders, One-Class Support Vector Machine (One-Class SVM), and Random Forest. These algorithms have shown great promise in detecting unusual behaviour in intricate communication networks. In instance, ensemble learning integrates a number of prediction models to

strengthen classification accuracy, decrease the number of false positives, and boost total detection precision.

Protecting 5G cellular networks from malicious traffic requires an anomaly detection system, and this research provides one that is based on machine learning. An example of 5G communication environment network traffic is included in the Cellular Dataset, which is used by the suggested system. Data cleaning, normalisation, label encoding, and training-testing partitioning are some of the preprocessing activities that the dataset goes through before model creation. This ensures that the machine learning algorithms are given high-quality input. Then, five different prediction methods are examined: Autoencoders, One-Class SVM, Random Forest, Ensemble 1 (Voting Classifier with AdaBoost, Decision Tree, and Gradient Boosting), and Ensemble 2 (Voting Classifier with AdaBoost, Decision Tree, and ExtraTree). The usefulness of Ensemble 2 for intelligent anomaly detection in 5G communication networks is highlighted by its greatest performance in comparison evaluations, which include 100% Accuracy, Precision, Recall, and F1-score.

Maintaining dependable communication performance while bolstering network security via intelligent anomaly detection is made easy with the suggested architecture. The research helps strengthen cybersecurity, enables real-time network monitoring, and makes current 5G cellular infrastructures more resistant to changing cyber threats by combining several machine learning models into an all-encompassing anomaly detection system.

II. RESEARCH BACKGROUND AND EXISTING DEVELOPMENTS

Introducing ultra-low latency, huge device connection, and ultra-high-speed data transfer, 5G cellular communication has revolutionised contemporary communications. Despite these technical improvements, new security concerns have emerged due to the complexity of 5G network designs, which are inefficiently handled by traditional rule-based monitoring systems. In order to keep networks available, secure sensitive data, and stop harmful cyber operations, it is crucial to recognise unusual communication patterns in real time, especially while network traffic keeps growing. Therefore, detecting anomalies in networks has become an essential part of contemporary telecommunication security research.

Statistical analysis, predetermined signatures, and threshold values that had to be manually specified were the mainstays of early anomaly detection methods. These methods were used to identify unusual network behaviour. When faced with new or changing cyber threats, these systems were often not very adaptable, even if they worked well for detecting existing attack patterns. Researchers are looking on intelligent data-driven methods that can automatically learn network behaviour since traditional intrusion detection techniques are becoming less successful in dealing with the ever-changing communication patterns and dynamic nature of 5G traffic.

By studying massive amounts of communication data and learning the traits of typical network behaviour, machine learning offers an effective method for detecting suspicious network activity, according to recent research. For the purpose of detecting anomalies in communication networks, supervised, unsupervised, and ensemble learning methods have all been explored. A few examples of these methods include Autoencoders, which are very good at detecting outliers via reconstruction error analysis, and One-Class Support Vector Machines, which can differentiate between typical and abnormal traffic patterns with just normal training data. The use of ensemble learning techniques has further enhanced the dependability of predictions by enhancing detection accuracy and decreasing the number of false positives.

In order to improve network monitoring and service assurance, some studies have concentrated on using machine learning in 5G network settings, which combine intelligent data analytics features. Scientific studies have shown that machine learning models can detect and prevent cyber threats like Distributed Denial-of-Service (DDoS) attacks, unauthorised access attempts, and deteriorating network performance by continuously analysing traffic behaviour and identifying unexpected communication patterns. In comparison to more traditional forms of protection, these smart methods greatly enhance both scalability and flexibility.

Despite a lot of success, researchers still have a long way to go before they can confidently build anomaly detection algorithms for massive 5G networks. We need detection models that can keep up with changing attack tactics, intricate communication settings, and network traffic variations while keeping computing cost to a minimum. In response to these difficulties, this research looks into a variety of machine learning methods for comparative anomaly detection in 5G network traffic using a Cellular Dataset. These methods include Autoencoders, One-Class SVM, Random Forest, and two ensemble models based on voting classifiers. The experimental study finds the best approach for improving network security while keeping the capacity to detect in real-time.

III. MACHINE LEARNING FRAMEWORK FOR 5G ANOMALY DETECTION

By consolidating many machine learning methods into one comprehensive anomaly detection system, the suggested framework may detect suspicious behaviour within 5G cellular networks. Traditional security measures often fail to identify complex threats in real time because of the growing complexity of contemporary communication infrastructures. This shortcoming is remedied by the suggested framework, which learns the features of valid communication and detects anomalies that can suggest malevolent activity by automatically analysing network

International Journal of Advanced Trends in Multidisciplinary Research
ISSN: 3139-0986, Volume 2 Issue 7
July 2026, PP 1-6

traffic patterns. An all-inclusive framework for intelligent network monitoring is provided by the following steps: dataset capture; data preprocessing; feature transformation; model training; ensemble learning; and anomaly prediction. Gathered from various communication contexts, the Cellular Dataset comprises typical 5G network traffic statistics, which is the first step in the anomaly identification process. In order to conduct a thorough study of communication behaviour, the dataset contains many network-related parameters. These elements include timestamps, network IDs, device types, data use, latency, and connection status. All things considered, these characteristics define typical and out-of-the-ordinary network activity and provide the groundwork for building machine learning models that can differentiate between normal and aberrant data.

In order to guarantee data compliance with machine learning algorithms and to increase data quality, the acquired data goes through a thorough preparation step prior to model building. To ensure the integrity of the dataset, we begin by removing entries with inconsistent or missing values. In order to ensure that all variables may make a meaningful contribution during model training, numerical features are normalised to decrease scale variations. Label encoding is used to transform category data into numerical representations that learning algorithms can handle more effectively, as many network properties are of a categorical kind. The suggested anomaly detection framework's learning power is improved and feature consistency is increased by these preprocessing activities.
To enable machine learning models to learn typical communication behaviour while keeping unseen samples for performance assessment, the produced dataset is split into training and testing subsets after preprocessing. Twenty to thirty percent of the dataset is kept aside for testing, whereas seventy to eighty percent is used for training the models. With this data partitioning method, we can objectively evaluate the generalisability of the model and reduce the likelihood of overfitting when detecting anomalies.

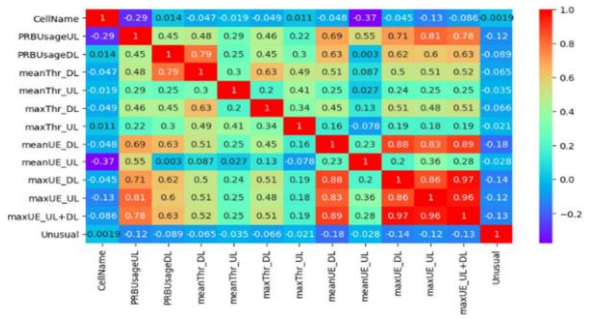

Fig. 1.Correlation Matrix

The suggested architecture explores many ML techniques to find the best one for 5G cellular network anomaly detection. Autoencoders first engage in unsupervised learning by attempting to recreate typical network traffic patterns and then spotting outliers by analysing the resulting reconstruction mistakes. After that, One-Class Support Vector Machine (One-Class SVM) models the distribution of reasonable communication behaviour and finds outliers. Furthermore, Random Forest enhances classification robustness via supervised classification, which it does by building numerous decision trees and pooling their predictions. We build two ensemble models that use Voting Classifiers to improve prediction accuracy even more. Both Ensemble 1 and Ensemble 2 use several classifiers to more reliably identify complex network abnormalities; Ensemble 1 uses AdaBoost, Decision Tree, and Gradient Boosting, while Ensemble 2 uses AdaBoost, Decision Tree, and ExtraTree.

Following model training, the testing dataset is used to assess each classifier using industry-standard performance measures such as Accuracy, Precision, Recall, and F1-score. By maintaining consistent preprocessing methods and dataset properties across trials, comparative analysis may be used to find the best anomaly detection model. The experimental analysis shows that the second ensemble model has the best classification performance, proving that intelligent anomaly detection in 5G cellular communication networks is possible by integrating various machine learning methods.

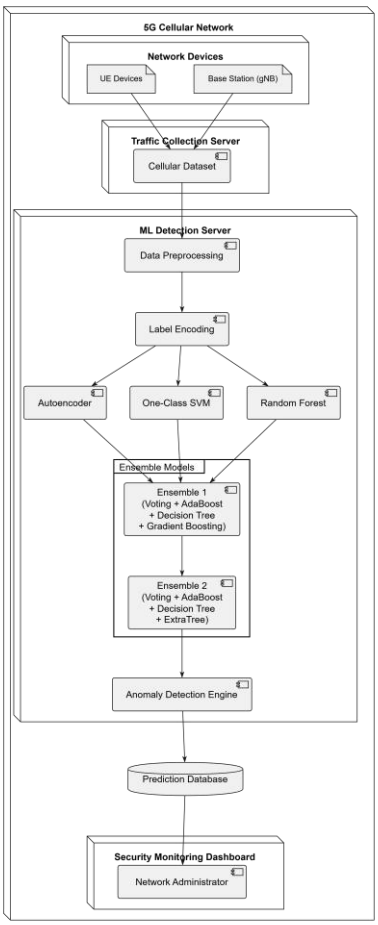

Fig. 2. UML Deployment Diagram for the Machine Learning-Based 5G Network Anomaly Detection Framework

IV. DATA PREPARATION AND MODEL DEVELOPMENT

The input data and training procedure used to build a model are the two most important factors in determining how well an anomaly detection system based on machine learning performs. The Cellular Dataset, which includes both typical and unusual network behaviours seen in a 5G cellular context, is the main source of network traffic statistics in the suggested architecture. Time stamps, network identifiers, device types, data use, latency, and connection status are some of the descriptive features included in the dataset that together depict the communication network's operational behaviour. Irregular traffic patterns may be an indication of cyber attacks or abnormal network conditions, and these qualities provide enough information to detect them.

A structured preprocessing pipeline is applied to the acquired dataset prior to machine learning model training in order to improve data quality and learning performance. In order to keep the dataset consistent, we first look for incomplete entries and eliminate them. We also look for missing attribute values. To facilitate effective model convergence during training and to minimise scale fluctuations, the remaining numerical characteristics are normalised. Label encoding is used to convert textual information into numerical representations that machine learning algorithms can understand, as many network properties are categorical. By standardising and enhancing the quality of the input data, these preprocessing activities increase the accuracy of anomaly detection for all models.
The dataset is split into training and testing subsets after preprocessing. Model training uses around 70–80% of the available samples, whereas performance assessment uses about 20–30%. The models may learn typical communication patterns and test their ability to categorise unseen network traffic using this data partitioning technique. To avoid the problem of overfitting and to evaluate the generalisability of the model in a realistic setting, it is best to keep the training and testing datasets separate.

In order to find the best way to spot suspicious activity in 5G networks, the anomaly detection framework looks at several machine learning techniques. Autoencoders first undertake unsupervised representation learning by attempting to recreate typical traffic patterns and then, using the reconstruction error to detect outliers. When it comes to detecting observations that drastically differ from taught patterns, One-Class Support Vector Machine (One-Class SVM) models the border of valid network behaviour. To further enhance classification reliability and robustness, Random Forest conducts supervised classification by building numerous decision trees and pooling their predictions.

The suggested system incorporates two ensemble learning algorithms to further increase anomaly detection capacity. In Ensemble 1, a Voting Classifier is used to analyse network traffic characteristics jointly. This classifier includes AdaBoost, Decision Tree, and Gradient Boosting. Also using a different configuration of Voting Classifiers, Ensemble 2 brings together AdaBoost, Decision Tree, and ExtraTree. The prediction stability, false detection rate, and overall anomaly detection performance may all be enhanced by using these ensemble approaches, which use the capabilities of several machine learning models. Finding the best method for protecting 5G communication environments from malicious network activity requires comparing and contrasting all five prediction models.

V. COMPARATIVE PERFORMANCE ASSESSMENT

By comparing five different anomaly detection models with the Cellular Dataset that has been produced, we may assess how successful the suggested machine learning framework is. The goal of the test is to find out how well each algorithm works in a 5G communication setting to differentiate between typical network activity and suspicious ones. A fair evaluation of their anomaly detection ability may be achieved by the comparison analysis, given all models are trained using the same preprocessing pipeline, dataset splitting technique, and evaluation criteria.

Autoencoder, One-Class Support Vector Machine (One-Class SVM), and Random Forest are the three machine learning models that are first tested separately. In order to detect suspicious network activity, each model uses a unique learning technique. Unsupervised learning situations are well-suited to the Autoencoder's anomaly identification capabilities, which it achieves via reconstruction error analysis. An anomaly is defined as an observation that falls outside of the regular range of network behaviour, and here is where One-Class SVM comes in. To make better predictions, Random Forest aggregates the outputs of individual decision trees and uses them to conduct supervised categorisation. Results from the experiments show that compared to the individual unsupervised models, Random Forest offers far superior classification performance.

We look at two ensemble learning methods to see if we can boost our prediction capabilities any further. Several classifiers may assess network traffic patterns simultaneously in Ensemble 1, which combines AdaBoost, Decision Tree, and Gradient Boosting within a Voting Classifier. Ensemble 2 follows the same voting procedure as the first, combining AdaBoost, Decision Tree, and ExtraTree. Ensemble approaches enhance classification resilience, decrease prediction uncertainty, and minimise false anomaly detections by integrating the capabilities of numerous predictive models. The framework is able to detect complicated traffic anomalies more efficiently than separate machine learning algorithms because to its collaborative decision-making process.

There are noticeable performance disparities across the models that were compared. Even while it achieves

excellent precision under some situations, One-Class SVM only delivers an accuracy of 26.7%, in contrast to the Autoencoder's 68.2%. Anomaly detection capability is greatly enhanced by Random Forest, achieving an impressive accuracy of 92.7%. Adding ensemble learning improves categorisation much more; Ensemble 1 has a 96.9% Accuracy. With perfect scores in all four metrics—Accuracy, Precision, Recall, and F1-score—Essential 2 outperforms all other ensembles when it comes to identifying outliers in 5G cellular network data. These results show that compared to machine learning models that work alone, the reliability of anomaly detection is much improved when numerous classifiers are integrated using ensemble learning.

VI. CONCLUSION

This research introduces a framework for 5G cellular network security and reliability enhancement using anomaly detection based on machine learning. By examining communication traffic obtained from a Cellular Dataset, the suggested framework explores several machine learning techniques for detecting anomalous network behaviour. An all-inclusive anomaly detection system that can differentiate between typical and unusual network activity is created by combining data pretreatment, label encoding, dataset segmentation, individual ML models, and ensemble learning approaches according to the whole methodology. To maintain consistency in experimental analysis, the research relies on maintaining the original machine learning methods, dataset features, and assessment methodology.

Autoencoder, One-Class Support Vector Machine (One-Class SVM), Random Forest, Ensemble 1 (Voting Classifier with AdaBoost, Decision Tree, and Gradient Boosting), and Ensemble 2 (Voting Classifier with AdaBoost, Decision Tree, and ExtraTree) are the five predictive models that are measured and compared. Ensemble learning outperforms standalone machine learning models in anomaly detection, according to the comparison. Ensemble 2 outperforms all other techniques by earning perfect scores in all four performance metrics—Accuracy, Precision, Recall, and F1-score—showcasing its capacity to reliably detect anomalous communication patterns in intricate 5G network settings. The experimental results show that using a combination of machine learning classifiers yields better results for dependable anomaly identification in real time.

An intelligent method to boost cybersecurity inside current cellular communication networks is provided by the suggested anomaly detection framework. The framework helps administrators keep communication services safe and uninterrupted by automatically monitoring network traffic and detecting suspicious behaviours. It also facilitates quick threat identification and decreases the chance of network intrusion. The suggested method is well-suited for implementation in next-generation telecommunications infrastructures because it incorporates ensemble learning, which further improves prediction stability and reduces false detections.

To further prepare for large-scale deployment, future studies may look at more varied and extensive 5G traffic datasets, test the framework in different network environments, and study optimisation methods. It is possible that future research may evaluate how well the suggested machine learning framework can adapt to changing mobile communication settings without compromising the core anomaly detection techniques used here.

REFERENCES

1. S. Sevgican, M. Turan, K. Gökarslan, H. B. Yilmaz, and T. Tugcu, “Intelligent network data analytics function in 5G cellular networks using machine learning,” Journal of Communications and Networks, vol. 22, no. 3, pp. 269–280, Jun. 2020.
2. U. Haider, M. Waqas, M. Hanif, H. Alasmary, and S. M. Qaisar, “Network load prediction and anomaly detection using ensemble learning in 5G cellular networks,” Computer Communications, vol. 197, pp. 141–150, 2023.
3. J. Lam and R. Abbas, “Machine learning based anomaly detection for 5G networks,” arXiv preprint arXiv:2003.03474, 2020.
4. A. U. Rahman, M. Mahmud, T. Iqbal, L. Saraireh, H. Kholidy, M. Gollapalli, and M. I. B. Ahmed, “Network anomaly detection in 5G networks,” Mathematical Modelling of Engineering Problems, vol. 9, no. 2, pp. 357–364, 2022.
5. L. F. Maimó, F. J. G. Clemente, M. G. Pérez, and G. M. Pérez, “On the performance of a deep learning-based anomaly detection system for 5G networks,” in Proc. IEEE SmartWorld Congress, 2017, pp. 1–8.
6. H. Ekström, “Non-Standalone and Standalone: Two Standards-Based Paths to 5G,” Ericsson Technology Review, Jul. 2019.
7. 3GPP, “Architecture Enhancements for the 5G System (5GS) to Support Network Data Analytics Services,” TS 23.288, Version 16.1.0, Sept. 2019.
8. 3GPP, “System Architecture for the 5G System (5GS),” TS 23.501, Version 16.2.0, Sept. 2019.
9. 3GPP, “5G System; Network Data Analytics Services; Stage 3,” TS 29.520, Version 16.1.0, Sept. 2019.
10. 3GPP, “5G System; Unified Data Management Services; Stage 3,” TS 29.503, Version 16.1.0, Sept. 2019.
11. C. Hernandez-Chulde and C. Cervello-Pastor, “Intelligent optimization and machine learning for 5G network control and management,” in Proc. Int. Conf. Practical Applications of Agents and Multi-Agent Systems (PAAMS), 2019, pp. 339–342.
12. F. Boccardi, R. W. Heath Jr., A. Lozano, T. L. Marzetta, and P. Popovski, “Five disruptive technology

directions for 5G," IEEE Communications Magazine, vol. 52, no. 2, pp. 74–80, Feb. 2014.
13. 3GPP, "Study on Latency Reduction Techniques for LTE," TR 36.881, Jul. 2016.
14. M. J. Khan, R. C. S. Chauhan, and I. Singh, "Outage probability and throughput of cooperative non-orthogonal multiple access with moving relay in heterogeneous network," Transactions on Emerging Telecommunications Technologies, vol. 33, no. 12, 2022.
15. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in Proc. Advances in Neural Information Processing Systems (NeurIPS), 2014, pp. 2672–2680.
16. S. Sevgican, M. Turan, K. Gökarslan, H. B. Yilmaz, and T. Tugcu, "Synthetic 5G cellular network data for NWDAF," 2019.
17. M. Ring, S. Wunderlich, D. Scheuring, D. Landes, and A. Hotho, "A survey of network-based intrusion detection data sets," Computers & Security, vol. 86, pp. 147–167, 2019.
18. T. Ahmed, B. Oreshkin, and M. Coates, "Machine learning approaches to network anomaly detection," in Proc. USENIX Workshop on Tackling Computer Systems Problems with Machine Learning Techniques, 2007.
19. D. E. Denning, "An intrusion-detection model," IEEE Transactions on Software Engineering, vol. SE-13, no. 2, pp. 222–232, Feb. 1987.
20. M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 dataset," in Proc. IEEE Symposium on Computational Intelligence for Security and Defense Applications, 2009, pp. 1–6.
21. N. Moustafa and J. Slay, "UNSW-NB15: A comprehensive data set for network intrusion detection systems," in Proc. Military Communications and Information Systems Conference (MilCIS), 2015, pp. 1–6.
22. G. Apruzzese, M. Colajanni, L. Ferretti, A. Guido, and M. Marchetti, "On the effectiveness of machine and deep learning for cyber security," in Proc. IEEE International Conference on Cyber Conflict, 2018, pp. 1–8.
23. P. Ramesh Babu, "Scalable VLSI Design for Edge AI Acceleration in Real-Time Audio-Visual Processing Systems," in Proc. IEEE Int. Conf., 2026.
24. P. Ramesh Babu, "Design of Flexible Polygon Shape Compact Patch Antenna with Slit for Biomedical Application," in Proc. IEEE Int. Conf., Feb. 24, 2026.
25. S. Venkateswara Rao, "Use Machine Learning to Predict the Probability of Being Admitted to a University," International Journal of Computer Networks and Wireless Communications (IJCNWC), vol. 15, no. 2, 2025.
26. S. Venkateswara Rao, "An AI-based Method for Analyzing the Relationships Between Financial Variables and Stock Price Changes," International Journal of Biology, Agriculture and Healthcare Research (IJBAR), vol. 15, no. 2, 2025.