International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802004

Machine Learning Based Intrusion Detection System using Supervised and Unsupervised Learning

Vangara Navaneetha, Prathi Bhargavi, Rayapalli Chandu, Vadi Bhavani, D. Bhagyaraj Yadav, Dr. Prasad Dharnasi

UG Student, Department of Computer Science and Engineering, Holy Mary Institute of Technology and Science, Telangana, India.

UG Student, Department of Computer Science and Engineering, Holy Mary Institute of Technology and Science, Telangana, India.

UG Student, Department of Computer Science and Engineering, Holy Mary Institute of Technology and Science, Telangana, India.

UG Student, Department of Computer Science and Engineering, Holy Mary Institute of Technology and Science, Telangana, India.

Assistant Professor, Department of Computer Science and Engineering, Holy Mary Institute of Technology and Science, Telangana, India.

Professor, Department of Computer Science and Engineering, Holy Mary Institute of Technology and Science, Telangana, India.

Publication History: Received: 28.02.2026; Revised: 07.03.2026; Accepted: 11.03. 2026; Published: 15.03.2026.

ABSTRACT: An Intrusion Detection System (IDS) can monitor and sift through data traffic to identify and protect the network from unauthorized or dangerous activities. In this paper, we propose a hybrid approach to Intrusion Detection Systems that utilizes both supervised and unsupervised machine learning models. Supervised learning models use labelled data to train the models to classify traffic, while unsupervised learning models identify anomalies and perform outlier detection without any labelled data. Examples of supervised learning models include Decision Tree, Random Forest, Support Vector Machine, and K-Nearest Neighbours. Examples of unsupervised learning models include K-Means Clustering, Autoencoders, and DBSCAN. Our work use the cleansed and pre-processed NSL-KDD and UNSW-NB15 to develop and evaluate machine learning models to achieve the greatest data accuracy. Our approach hybrid Intrusion Detection System improves the detection of both known and unknown intrusions, decreases the false positive and false negative rates and increases the level of protection offered by IoT and corporate networks. Yes, intrusive detection and preventative systems can learn and continue to evolve over time as new data enters the system, deepening the data accuracy as time goes by. This enhanced system calibres will guarantee a proactive data protection and threat counteraction system to net and wireless net environments.

KEYWORDS: Intrusion Detection System, Machine Learning, Supervised Learning, Unsupervised Learning, Anomaly Detection, Cybersecurity.

I. INTRODUCTION

The first target of cyber-attacks is generally weaknesses in a user's behaviour, systems, or networks. When it comes to exploitation, TCP (Transmission Control Protocol), UDP (User Datagram Protocol), and ICMP (Internet Control Message Protocol) offer great insight into the behaviour of traffic in a network at the level of protocols, which is used to develop features that further improve the efficiency of machine learning models used in intrusion detection systems. These models, after being trained on what is 'normal' behaviour during communications, are able to detect 'abnormal' or 'strange' behaviour that can be seen during transmissions, which can be indicative of a cyber-attack. Models are trained

International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802004

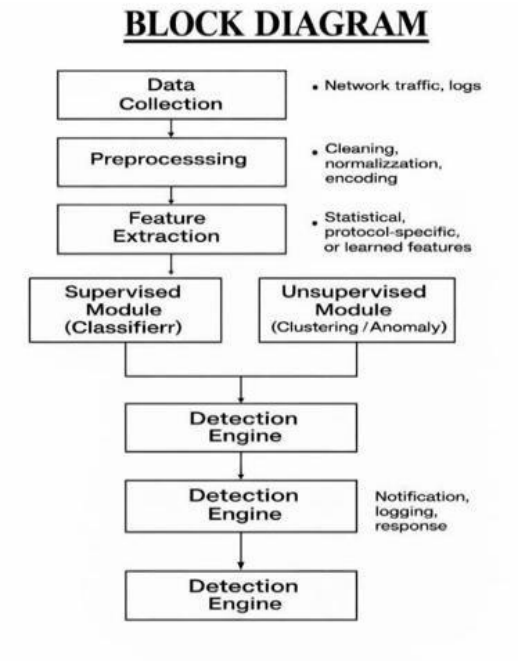
using the KDD Cup 99, NSL-KDD datasets, among many others. An Intrusion Detection System (IDS) is able to 'see' traffic that it is processing, therefore, allowing the traffic to flow, and providing 'real time' analysis of the flow, capturing and processing events to create alerts or logs to record the monitored behaviour. By monitoring network traffic, IDSs support the objective of protecting sensitive data in a computer and network system by establishing the range of data that is supposed to be maintained or utilized, therefore enabling the system to be performed and utilized.

II. LITERATURE REVIEW

The goal of this literature review is to gain insight into Machine Learning based Intrusion detection system using supervised and unsupervised learning. M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, in their paper entitled "A Detailed Analysis of the KDD CUP 99 Data Set," review the over-utilized KDD Cup 99 dataset and explain the problems of data redundancy and biased assessments of the intrusion detection field. They suggested the dataset NSL-KDD as a better alternative, which provides a more reliable and less imbalanced benchmark for the evaluation of machine-learning-based intrusion detection systems. Niyaz et al. demonstrated that deep feature learning yields better classification results than traditional methods, thereby illustrating the efficacy of deep learning to intrusion detection systems. They reviewed the NSL-KDD dataset and proposed a deep learning-based intrusion detection system using sparse autoencoders. However, their evaluation was obsolescent as it focused solely on old datasets and failed to consider practical real-world issues such as the costs of the computations and the speed of the system. N. Moustafa and J. Slay proposed the UNSW-NB15 dataset, which, with the incorporation of real network traffic and constructed contemporary attacks, gives a multifaceted benchmark for machine learning-based intrusion detection systems, which has received significant appreciation.

III. PROPOSED SYSTEM

This paper presents an intrusion detection system that incorporates both supervised and unsupervised machine learning approaches for the detection of nefarious actions in network traffic. It improves the system's accuracy and adaptability to changing protective security challenges, thus, offering an even better protection to computer networks. Now, we are now proposing the architecture to detect these attacks using ml models.



Data Processing:
The system here performs network data collection, which is a massive undertaking. The data includes network traffic, system logs, and user activity. It is the first of the intrusion detection system that needs to capture the entire network to facilitate a detailed examination.

International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802004

Preprocessing:
The initial data is typically raw, erroneous, and contains superfluous data. Preprocessing is the stage where data is cleaned, which includes the removal of noise and unnecessary information, the normalization to ensure values are on a common scale, and data is converted to a numerical format. This stage is aimed at preparing the data to be fed to machine learning algorithms.

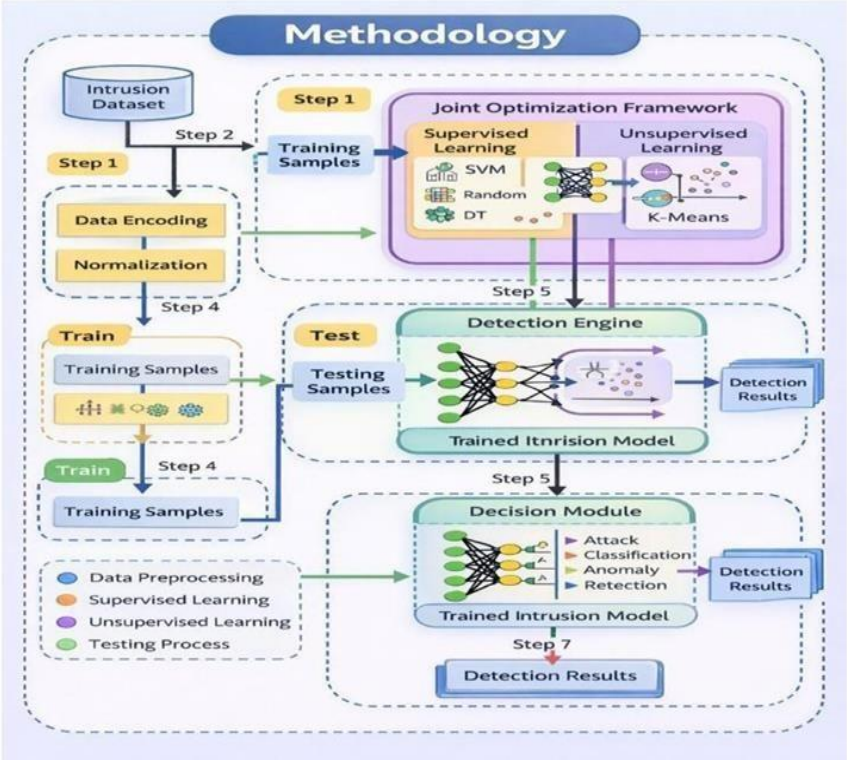
Feature Extraction:
This phase involves picking key elements from the raw data. Each type shows something different about the like how traffic shifts between small and big packets. Every feature ties back to protocol type, and connection duration. Feature Starting from messy datasets, extraction simplifies what gets gathered. This step often clears the way for clearer patterns later on detection accuracy.

Supervised Learning Module:
The supervised learning module uses labelled data to classify network traffic. It learns from examples that are So far, each entry shows whether it's been flagged as usual or suspicious. From that point on from time to time, the system updates what it knows. Instead of guessing, it looks at past events. When fresh data arrives, the model decides if there will be more. Sometimes it acts right, other times not.

Unsupervised Learning Module:
The unsupervised learning module works with unlabelled data. It groups similar data and detects Spotting oddities in how networks act. This part deals with what stands out in their behaviour. useful for spotting fresh or overlooked threats, including those not previously noted like nothing we've noticed before.

Detection Engine:
The detection engine gets output from each of the two processes supervised and unsupervised modules. It analyses these Outcomes shape how choices are made, leading to action on initiatives a sign shows whether a breach happened or not.

IV. METHODOLOGY



|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802004

This methodology, we enhance the performance of the intrusion detection process by distinguishing between data preprocessing, training, and testing phases and by integrating supervised and unsupervised learning in one framework. After encoding and normalizing the data set, we divide it into training samples and testing samples. We use training samples for building supervised models for learning known attack patterns and for building unsupervised models for learning normal network behaviour and for detecting anomalies. We optimize both models so that better feature representations are learned. We use test samples for testing and pass them through both models; then we process the output using a unified detection engine. Using this framework, we can accurately classify known attacks and detect unknown or new attacks as well. This proposed methodology enhances intrusion detection accuracy and robustness compared with traditional approaches that use only one type of learning.

V. IMPLEMENTATION

The current working document addresses the design of an intelligent IDS system that will use both supervised and unsupervised learning and how that will aid system robustness and reliability in detecting and delineating both known and unknown attack patterns.

TECHNIQUES TO DETECT ATTACKS IN THE FUTURE:

While building an IDS that incorporates the working principle of supervised and unsupervised learning and algorithms, we are facing some attacks that obstruct the system from acquiring the ability to distinguish in various ways. of cyber-attacks that are commonly found in networks and systems.

- **Denial of Service (DoS) Attacks:**
- Some service is attempted to be stopped from functioning by overuse of service, thus blocking a great deal of the space.
- **Probing attacks (scanning or monitoring):** these attacks are where an opening is exposed through a network.
- **Malware-based attacks:** these are attacks where the software that is the cause is the damaging one, and
- **Zero Day attack characterized attacks**: these are the attacks that are characterized by surprise, and there are no traces to be analyzation method for tackling these problems ahead of time is the adoption of certainIntrusion Prevention Systems analyze huge data streams from networks, system logs, and individuals. They do not rely on manually constructed detection systems. Rule-based system(s) have the ability to define and/or detect the patterns of incidents/ cases along with what is deemed to be the 'normal' to mask malicious behavior. This is akin to a Synaptic Neural Network. Once trained, the system is placed into operation and it becomes a continuous stream analytical engine. This system can analyze and categorize data streams to determine if the data is 'safe' or 'suspicious' and/or to detect the anomalies. The system can detect a sudden spike/problem concerning the volume of alerts. A condition relating to the volume of alerts is considered an anomalous state. Achieving all of the aforementioned is a product of applying state of the art techniques related to deep learning.

- **Convolutional Neural Networks (CNNs) are the premier state of the art systems for detecting patterns in organized and easily identifiable structures - specifically, "flows" of data packets.**
- **Autoencoders** are designed to learn acceptable (or 'standard') patterns in data. If the data fails to replicate the acceptable (standard) pattern, a state of 'reconstruction error' is condition - which presents an indication of potential traffic congestion.

EVALUATION METRICS

Different evaluation metrics were used to assess the proposed Random Forest, Decision Tree, K-means clustering, and DBSCAN models, enabling a thorough comprehension of the models, their dependability, and how they function in classification.

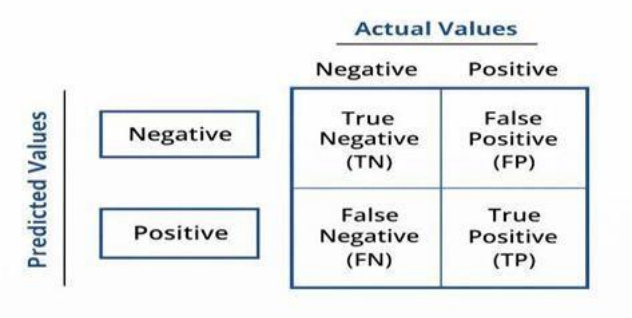
Confusion Matrix: A Confusion Matrix is a classification model evaluation tool that offers a complete overview of how a model makes predictions. Comparisons are made between the \text it{true} class labels and predicted class labels. The Confusion Matrix provides a model's correct predictions and the errors it makes (also called misclassifications).

International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802004



Accuracy:
The evaluation of accuracy is simply the ratio of correct predictions to the total number of observations (predictions).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Classification Report:
The effectiveness of a deep learning model in classification tasks is often assessed using a classification report. It is common for a classification report to consist of the following metrics in every class/label of the dataset.

Precision:
The percentage of true positive predictions, thereby measuring classification accuracy, is referred to as precision. This metric is highly applicable for use in scenarios where the occurrence of false positives is detrimental, such as in high-confidence prediction scenarios.

$$Precision = \frac{TP}{TP + FP}$$

Recall:
Recall is a measure of the proportion of true positive observations that a classification model is able to identify and is one of the most important metrics in classification.

$$Recall = \frac{TP}{TP + FN}$$

F1-Score:
Combining Precision and Recall into a single number for a classification model's performance evaluation has led to the widespread adoption of the F1 score.

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

VI. CONCLUSION

This project demonstrates how combining supervised and unsupervised learning can improve intrusion detection systems (IDS). Random Forest and Support Vector Machines (SVM) are examples of supervised learning models that are particularly good at recognizing certain types of attacks because they are trained with labelled data. Unsupervised learning models, such as K-Means and DBSCAN, identify different types of attacks and/or anomalous activities in a network absent of labelled data. Results indicate that the models are performing exceedingly well in recognizing and classifying different types of cyber-attacks. The integration of both models into a single system has increased the

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802004

efficiency and flexibility of the IDS while decreasing the incidence of false alarms. The combination of both models as the underpinning of the IDS system affords computer networks protection while providing an intelligent and reliable solution for intrusion detection.

VII. FUTURE SCOPE

To keep up with evolving threats, intrusion detection systems need to be more sophisticated and adaptable. Below are some future opportunities in the field of ML-based IDS:

1. Temporal & Spatial Learning: Slow-moving or concealed intrusions can be detected sooner by learning specific behaviours, and accomplished through models like LSTM, GRU, and Transformers.
2. Graph Neural Networks (GNNs): Model the network as a graph with devices as nodes and connection as edges. Find and address anomalous behaviours of the devices or edges.
3. Federated Learning: In terms of privacy and security, it allows several parties to collaboratively train a model while not sharing the data.
4. Transfer Learning: 'Data scarcity' implies learning new things in a given situation. Using the knowledge of one network for the other one is the right approach in these cases.
5. Hybrid Detection Models: Signature-based techniques to detect known attacks are used to enhance systems, therefore, fewer false positives are noted, thus improving the accuracy.

REFERENCES

1. Vaidya, S., Shah, N., Shah, N., & Shankarmani, R. (2020, May). Real-time object detection for visually challenged people. In *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 311–316). IEEE.
2. Dharnasi, P. (2025). A multi-domain AI framework for enterprise agility integrating retail analytics with SAP modernization and secure financial intelligence. *International Journal of Humanities and Information Technology, 7*(4), 61–66.
3. Krishna, G., Rajesh, B., Dinesh, B., Sravani, B., Rajesh, G., Dharnasi, P., & Sarvanan, M. (2026). Smart agriculture system using IoT with help of AI-techniques. *International Journal of Computer Technology and Electronics Communication, 9*(2), 479–487.
4. Patnaik, S. K., Sidhu, M. S., Gehlot, Y., Sharma, B., & Muthu, P. (2018). Automated skin disease identification using deep learning algorithm. *Biomedical & Pharmacology Journal, 11*(3), 1429.
5. Saravanan, M., & Sivakumaran, T. S. (2016). Three phase dual input direct matrix converter for integration of two AC sources from wind turbines. *Circuits and Systems, 7*, 3807–3817.
6. Akshaya, N., Balaji, Y., Chennarao, J., Sathwik, P., & Dharnasi, P. (2026). Diabetic retinopathy diagnosis with deep learning. *International Journal of Research Publications in Engineering, Technology and Management (IJRPETM), 9*(2), 506–512.
7. Lakshmi, A. J., Dasari, R., Chilukuri, M., Tirumani, Y., Praveena, H. D., & Kumar, A. P. (2023, May). Design and implementation of a smart electric fence built on solar with an automatic irrigation system. In *2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)* (pp. 1553–1558). IEEE.
8. Reddy, N. H. V., Reddy, N. T., Bharath, M., Hemanth, N., Dharnasi, D. P., Nirmala, B., & Jitendra, A. (2026). AI based learning assistant using machine learning. *International Journal of Engineering & Extended Technologies Research, 8*(2), 495–504.
9. Kumar, A. S., Saravanan, M., Joshna, N., & Seshadri, G. (2019). Contingency analysis of fault and minimization of power system outage using fuzzy controller. *International Journal of Innovative Technology and Exploring Engineering, 9*(1), 4111–4115.
10. Gopinathan, V. R. (2025). Intelligent workload scheduling for telecom cloud architecture using reinforcement learning. *International Journal of Research Publications in Engineering, Technology and Management, 8*(6), 13244–13255.
11. Bhagyasri, Y., Bhargavi, P., Akshaya, T., Pavansai, S., Dharnasi, P., & Jitendra, A. (2026). IoT based security & smart home intrusion prevention system. *International Journal of Computer Technology and Electronics Communication, 9*(2), 457–462.
12. Roy, S., & Saravana Kumar, S. (2021). Feature construction through inductive transfer learning in computer vision. In *Cybernetics, Cognition and Machine Learning Applications: Proceedings of ICCCMLA 2020* (pp. 95–107). Springer.
13. Reddy, V. N., Rao, P. H. S., Singh, N. S., Kumar, V. S. S., Reddy, Y. B., & Dharnasi, P. (2026). Face recognition using criminal identification system. *International Journal of Research Publications in Engineering, Technology and Management, 9*(2), 520–527.

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802004

14. Fazilath, M., & Umasankar, P. (2025, February). Comprehensive analysis of artificial intelligence applications for early detection of ovarian tumours: Current trends and future directions. In *2025 3rd International Conference on Integrated Circuits and Communication Systems (ICICACS)* (pp. 1–9). IEEE.
15. Vijayakumar, R., & Gireesh, G. (2013, July). Quantitative analysis and fracture detection of pelvic bone X-ray images. In *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)* (pp. 1–7). IEEE.
16. Chinthala, S., Erla, P. K., Dongari, A., Bantu, A., Chityala, S. G., & Saravanan, M. S. (2026). Food recognition and calorie estimation using machine learning. *International Journal of Engineering & Extended Technologies Research, 8*(2), 480–488.
17. Saravanan, M., Kumar, A. S., Devasaran, R., Seshadri, G., & Sivaganesan, S. (2019). Performance analysis of very sparse matrix converter using indirect space vector modulation. *International Journal of Innovative Technology and Exploring Engineering, 9*(1), 4756–4762.
18. Rupika, M., Nandini, G., Mythri, M., Vasu, K., Abhiram, M., Shivalingam, N., & Dharnasi, P. (2026). Electronic gadget addiction prediction using machine learning. *International Journal of Research Publications in Engineering, Technology and Management, 9*(2), 500–505.
19. Prasad, E. D., Sahithi, B., Jyoshnavi, C., Swathi, D., Arun Kumar, T., Dharnasi, P., & Saravanan, M. (2026). A technology driven solution for food and hunger management. *International Journal of Computer Technology and Electronics Communication, 9*(2), 440–448.
20. Chanamalla, B., Murali, V. N., Suresh, B., Deepak, M. S., Zakriya, M., Yadav, D. B., & Saravanan, M. (2026). AI-driven multi-agent shopping system through e-commerce system. *International Journal of Computer Technology and Electronics Communication, 9*(2), 463–470.
21. David, A. (2020). Air pollution control monitoring & delivery rate escalated by efficient use of Markov process in MANET networks: To measure quality of service parameters. *Test Engineering & Management*.
22. Thotla, S. B., Vyshnavi, S., Anusha, P., Vinisha, R., Mahesh, S., Yadav, D. B., & Dharnasi, P. (2026). Traffic congestion prediction using real time data by using deep learning techniques. *International Journal of Engineering & Extended Technologies Research, 8*(2), 489–494.
23. Nagamani, K., Laxmikala, K., Sreeram, K., Eshwar, K., Jitendra, A., & Dharnasi, P. (2026). Disaster management and earthquake prediction system using machine learning. *International Journal of Research Publications in Engineering, Technology and Management, 9*(2), 495–499.
24. Vimal Raja, G. (2024). Intelligent data transition in automotive manufacturing systems using machine learning. *International Journal of Multidisciplinary and Scientific Emerging Research, 12*(2), 515–518.
25. Rakesh, V., Vinay Kumar, M., Bharath Patel, P., Varun Raj, B., Saravanan, M., & Dharnasi, P. (2026). IoT-based gas leakage detector with SMS alert. *International Journal of Computer Technology and Electronics Communication, 9*(2), 449–456.
26. Vishwarup, S., et al. (2020). Automatic person count indication system using IoT in a hotel infrastructure. In *2020 International Conference on Computer Communication and Informatics (ICCCI)* (pp. 1–4). IEEE. https://doi.org/10.1109/ICCCI48352.2020.9104195
27. Poornima, G., & Anand, L. (2024, April). Effective strategies and techniques used for pulmonary carcinoma survival analysis. In *2024 1st International Conference on Trends in Engineering Systems and Technologies (ICTEST)* (pp. 1–6). IEEE.
28. Nandhini, T., Babu, M. R., Natarajan, B., Subramaniam, K., & Prasanna, D. (2024). A novel hybrid algorithm combining neural networks and genetic programming for cloud resource management. *Frontiers in Health Informatics, 13*(8).
29. Chinthamalla, N., Anumula, G., Banja, N., Chelluboina, L., Dangeti, S., Jitendra, A., & Saravanan, M. (2026). IoT-based vehicle tracking with accident alert system. *International Journal of Research Publications in Engineering, Technology and Management, 9*(2), 486–494.
30. Rachana, P., Kalyan, P. P., Kumar, T. S., Reddy, P. M., Rohan, P., Saravanan, M., & Dharnasi, P. (2026). Secure chat application with end-to-end encryption using deep learning. *International Journal of Computer Technology and Electronics Communication, 9*(2), 472–478.
31. Pavan Kumar, T., Abhishek Goud, T., Yogesh, S., Manikanta, V., Dinesh, P., Srinu, B., & Dharnasi, P. (2026). Smart attendance system using facial recognition for staff using AI/ML. *International Journal of Research Publications in Engineering, Technology and Management, 9*(2), 513–519. https://doi.org/10.15662/IJRPETM.2026.0902005
32. Charumathi, M. V., & Inbavalli, M. FAMILIARIZING THE PINE NUT OIL BY FUSING IT INTO DIFFERENT FOOD PRODUCTS.