IRJCS: International Research Journal of Computer Science
Volume 12, Issue 08, October 2025
ISSN: 2393-9842
https://www.irjcs.com/archives
https://doi.org/10.26562/irjcs.2025.v1208.11

IntelliGuard: Intelligence on Cyber Risk for Commercial Protection

Ms.Sowmya J,
Assistant professor, Dept. of Computer Science & Application,
The Oxford College of Science, Bangalore, India
Krishna P
PG Student [MCA], Dept. of Computer Applications and Science,
The Oxford College of Science, Bangalore, India

Publication History
Manuscript Reference: IRJCS/RS/Vol.12/Issue08/OCCS10090
Research Article | Open Access | Double-Blind Peer Reviewed Article ID: IRJCS/RS/Vol.12/Issue08/OCCS10090
Received:23,October 2025, Revised: 09, October 2025, Accepted: 31October 2025 Published Online: 21November 2025
https://www.irjcs.com/volumes/Vol12/iss-08/11.OCCS10090.pdf
Article Citation:Sowmya,Krishna(2025), IntelliGuard: Intelligence on Cyber Risk for Commercial Protection**,**IRJCS: International Research Journal of Computer Science, Volume 12, Issue 08 of 2025 pages 368-373
doi:> https://doi.org/10.26562/irjcs.2025.v1208.11
BibTeX Sowmya@2025IntelliGuard
IRJCS papers should be cited as IRJCS (International Research Journal of Computer Science, AM Publications, India 2025, ISSN 2393-9842, https://doi.org/10.26562/ijrcs.2025.v1208.11 The journal's official abbreviation is IRJCS.
Orcid: https://orcid.org/0009-0004-9398-7488

Abstract: The Internet of Things paradigm has transformed connectivity, but it has also introduced previously unheard-of security risks. The classification of assaults using a unfathomable erudition approach and other machine learning approaches is the prime focal point of this study. Using machine learning (ML) technique such a shaphazard Forest (RF),conclusion tree(DT), Extra Tree Trainer (ETC), Support Vector Machine(SVM),and k-Nearest Neighbor (KNN) and Deep Learning (DL) design concepts, including Deep Neural Network (DNN), the study analyzes various attack types in IoT environments and proposes a binary and multiclass classification framework. Benchmark datasets with real-world IoT attack scenarios, such as Edge-IIoT set, are used for experiments. The dataset is preprocessed using Principal Componenet Analysis (PCA) for choosing features, Synthetic Minority Over sampling Technique to solve class imbalance, and Standard Scaling for feature scaling. These approaches' comparative performance and efficacy are examined. The outcomes demonstrate how well machine learning strategies generalize across various attack classes and how well the DL model manages intricate assault patterns. The DNN model yields the most excellent results with 100% accuracy for binary classification, 96.15% accuracy for 6-class classification, and 94.68% effectiveness for 15-class classification. A ten-fold cross validation approach has also been used to check for over fitting in the model. This work helps with the improvement of security mechanisms by offering insights into the selection of appropriate strategies for the binary and multiclass breakdown of threats.
Keywords: Support Vector Machine, Deep Learning, Haphazard Forest, DNN model, Synthetic Minority Oversampling Technique

I. INTRODUCTION

The hurried growth of the Internet of equipment has revolutionized industries by enabling seamless device interconnectivity and automation. However, this proliferation of connected devices has introduced unprecedented safekeeping vulnerabilities and attack surfaces. conventional sanctuary mechanism are often ill-equipped to knob the dynamic and heterogeneous nature ofenvironments. This necessitates the deployment of intelligent threat detection models that can efficiently identify and classify various forms of cyber attacks. In this context, apparatus learning (ML) and deep education (DL) techniques offer a promising alternative to rule-based detection, providing the ability to gain knowledge of from facts and adapt to evolving threats. The present work focuses on developing a robust IoT threat detection framework, "Smart Sentry," leveraging both ML classifiers such as accidental Forest (RF), Decision Tree (DT), Extra Tree Classifier (ETC), carry Vector apparatus(SVM), and k-Nearest Neighbor (KNN) and a Deep Neural Network (DNN). This hybrid framework is evaluated on the Edge-I IoT set dataset, encompassing real-world attack scenarios, through both double and multiclass organization approaches. By integrating dimensionality reduction, resampling strategies, and scaling techniques, the system aims to enhance performance, minimize latency, and ensure generalizability.

II. RELATED WORK

1.M.A.Farrago, O.Friar, D.Hammond, L.Magyars, and H.Janacek, "Edge- IIoTset: A new all-inclusive down-to-earth cyber safety measures dataset of IoT and Riot function for federal and federate learning," IEEE Access, vol. 10, pp. 40281–40306, 2022.

In this broad sheet, we recommend a new A wide-ranging and realistic cyber security dataset for IoT and IIoT applications, named Edge-IIoTset, designed to support machine learning based infringement uncovering system in two operational modes: centralized learning and federated learning.

Dataset has been generated using a purpose-built IoT/ IIoT test bed with a large representative set of devices, sensors, protocols and cloud/edge configurations. The IoT data were collected from more than ten different types of devices, including glow-cost digital sensors for warmth and clamminess, ultrasonic sensors, hose down point detectors, pH meters, soil moisture sensors, heart rate monitors, and flame sensors. In addition, fourteen types of attacks targeting IoT and IIoT connectivity protocols were identified and analyzed, which were grouped into five major threat categories: DoS / DDoS, information gathering, man-in-the-middle, injection, and malware attacks.Moreover,featureswereextractedfromdiversesourcessuchas alerts, system resources, logs, and network traffic, leading to the proposal of 61 new highly correlated features.1176 found features.

2.K.Gupta, D.K.Sharma's, K.Data Gupta, and A.Kumar,"A tree classifier based network intrusion detection model for Internet of medical things," Compute. Electra.Eng.,vol.102, Sep. 2022

Healthcare is a prominent domain where the Internet of equipment (IoT) has shown significant potential, giving rise to the Internet of health check Things (IoMT) for improved medical services. While IoMT offers substantial benefits, it also introduces cyber security risks that may jeopardize patient privacy and even endanger lives. The rapid adoption of IoMT strategy to provide efficient health care services at scale highlights the requirement of developing secure frameworks to safeguard patients. Nevertheless, creating robust security models for IoMT environments remains a considerable challenge. In this work, a tree classifier–based imposition uncovering classification is proposed for IoMT networks. The model reduces the dimensionalities' of input data to accelerate anomaly detection, while still achieving a high detection accuracy of 94.23%.

3. G.Bhandari,A.Lyth,A.Shalaginov,andT.-M.Grønli,"Distributed deep neural-network-based middleware for cyber-attacks detection in smart IoT ecosystem: A novel frame work and performance evaluation approach," Electronics, vol. 14, no. 2, p. 2908, Jan. 2023.

Cyber attacks continue to be one of the most significant challenges in today's digital era. The rapid expansion of Internet of Things (IoT) devices has introduced several security concerns, including in sufficient encryption, malware, ransom ware, and IoT bot nets, which make these devicessusceptible to exploitation. Attackers can compromise sensitive data, disrupt systems, or demand ransom. Past incidents of cyberattacks emphasize the need for establishing strong cyber security benchmarks, particularly within modern Smart Environments. This study presents a framework that leverages artificial intelligence (AI) techniques to detect malware attacks across diverse and distributed scenarios. The proposed method enables proactive monitoring of network traffic to identify malicious activity in the IoT ecosystem, thereby enhancing the security and resilience of Smart Environments against emerging threats.

4. F.EusufzaiMI.AA.RedWan,S.R.Sabuj,M.Elsharief,andM.M.MohamedRashid,S.U.Khan,and M."A federated learning-based approach for on the road to recovery intrusion uncovering in industrial Internet of possessions networks," Network, vol. 03, no. 1, pp. 1508–1719, Jan. 2023.

The Internet of Things (IoT) refers to a network of smart electronic devices that communicate wirelessly through the Internet. These interconnected devices continuously generate massive volumes of user-related data, making the system highly vulnerable to security threats and malicious intrusions. As the number of IoT devices grows rapidly, relying on centralized deep learning models raises serious privacy concerns. Traditional centralized machine learning (ML)-based frameworks are also challenging to implement, as they demand large datasets to be collected and stored in a single location. Since IoT data is widely distributed across multiple networks, there is a pressing need for decentralized learning approaches. To address this, our study introduces a Federated Learning (FL) framework for intrusion detection, aimed at strengthening IoT network security. By enabling asynchronous training on local IoT device data, this method preserves user confidentiality and enhances privacy protection.

III. EXISTING SYSTEM

The study in [1] introduces a novel dataset, **Edge-I IoT set**, designed to evaluate machine learning–based intrusion detection systems. This dataset was developed using a customized test bed comprising a diverse set of devices, sensors, protocols, and cloud/edge configurations. It incorporates multiple attack scenarios, such as malware, man-in-the-middle, and denial-of-service, while covering a broad spectrum of IoT and IIoT applications in a realistic and comprehensive manner. The authors further benchmark the performance of an assortment of appliance learning algorithms on Edge-I IoT set under both centralized and federated learning settings.

In [2], the authors propose a new intrusion detection system (IDS) framework for Internet of Medical Things (IoMT) environments. The approach is based on a decision tree classifier, where the model first preprocesses network traffic data to extract relevant features and then uses them totrain the classifier. Once trained, the mock-up can tell apart sandwiched between benevolent and malicioustravel.Theeffectiveness of the proposed IDS is evaluated on the publicly available UNSW-NB15 dataset of IoMT set-up traffic, achieving superior performance compared to several state-of-the-art IDS models, with an truth of 94.23%.

The study examines IDS for IIoT networks as they presently exist [7]. When discussing detection of breaches in IIoT networks, the authors cover a number of topics, such as the diversity of devices, their resource constraints, and the requirement for real-time functioning in many IIoT applications. The authors also discuss the various types of intrusion detection technologies (IDSs) that can be used on networks, including those that rely on signatures, anomalies, or machine learning. They identify which IDS types are best suited for certain applications by weighing all the benefits along with the drawbacks of each type. They argue that automatic learning-based intrusion detectionsystems are the most effective tactic.

IRJCS: International Research Journal of Computer Science ISSN: 2393-9842
Volume 12, Issue 08, October 2025 https://www.irjcs.com/archives
https://doi.org/10.26562/irjcs.2025.v1208.11

IV.PROPOSED SYSTEM

1. This study presents a comprehensive review of prior research focused on classifying anddetecting attacks within the Edge-IIoTset dataset.
2. The methodology adopted for classification is described in detail, outlining the steps andalgorithms used.
3. Techniques such as PCA for dimensionality reduction, SMOTE for handling class imbalance, and standardscalingforfeaturenormalizationsignificantlyenhancetheoverallperformance.
4. A thorough performance evaluation is conducted for binary, 6-class, and 15-class attack classifications using the Edge-IIoT set dataset.
5. **Smart Sentry** employs multiple machine learning algorithms, including Random Forest (RF), Decision Tree (DT), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Extra Trees Classifier (ETC), along with a deep learning model (DNN). For binary classification, most models achieve 100% accuracy, while DNN records the best results for 6-class (96.15%) and 15-class (94.86%) tasks.
6. Smart Sentry demonstrates efficient performance with in minimal execution time and negligible data loss. Specifically, the DNN model achieves testing times of 20.05s,18.11s, and 27.99s for 2-class,6-class, and 15-class classifications, respectively. Its efficiency systems from fewer parameters, reduced computational demand, and lower model complexity.
7. Strong results from 10-fold cross-validation confirm that the model avoids over fitting and generalizes effectively to unseen data.

Advantages:-

Decision Tree Classifier (DT): A decision tree is a widely used machine learning method for both classification and regression tasks. It resembles a flowchart where internal nodes represent features, branchescorrespondtodecision rules, and leaf nodes denote outcomes. The model recursively split the dataset base on the most informative features to construct the tree structure.

Extra Tree Classifier (ETC): An ensemble machine learning technique that is closely connected to the random forest concept is the Extra Tree Classifier, sometimes referred to as Extremelyrandomly generated Trees. It uses random subsets of the data to create several decision trees, each with its own twist. By choosing random split points for features, the additional trees increase the degree of randomization without offering even less bias and variation than conventional Random Forests.

Random Forest Classifier (RF) By mixing numerous decision trees, the Random Forest Classifieris an exceptionally efficient ensemble machine learning model that lowers excess fitting and raises prediction accuracy. The way it works is by building a forest of decision trees, each one of which has been trained using a different set of qualities and data.

IV. SYSTEM ARCHITECTURE

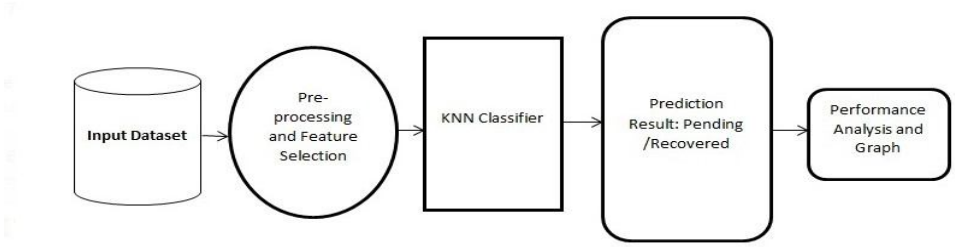

Fig.1System Architecture Diagram

Methodology

1. Machine Learning Algorithms:

Result Tree Classifier (DT):-A Decision Tree is solitary of the simple stand most intuitive models. It creates a tree-like organization everywhere each knob represents a mark (attribute), and each branch represent a decision rule base on that feature.

Working: The model splits the dataset a teach decision node based on the best feature that minimize simpurity (e.g.,Gini index, Information Gain).This process continues until the tree reaches a stopping condition, such as a maximum depth or minimum data at a node.

Pros: Easy to interpret and visualize. Good for both categorization and waning tasks.

Cons: Can overfit the data, especially with noisy data or in sufficient pruning.

Random Forest Classifier (RF):- The casual of forest model is an assembly scholarship system that builds numerous decision trees and combines their results to improve accuracy.

Working: Random Forest creates a compilation of result trees, each trained on a casual subset of the data. For classification, each tree votes on the class, and the class with the majority votes becomes the prediction. It uses bootstrap aggregation (bagging) to reduce variance and put off more than correct.

Pros: Less horizontal to over right compared to a on its own decision tree. Handles large datasets with higher accurateness.handle missing values better than individual decision trees. Cons: More computationally expensive than decision trees due to multiple trees.

IRJCS: International Research Journal of Computer Science ISSN: 2393-9842
Volume 12, Issue 08, October 2025 https://www.irjcs.com/archives
https://doi.org/10.26562/irjcs.2025.v1208.11

Less interpretable than a single decision tree. Extra Tree Classifier (ETC):- The Extra Trees classifier is analogous to Random Forest, but it introduces additional randomness in the way trees are created.

Working: Like Random Forest, it builds many decision trees. However, it selects the split point of each node randomly, instead of searching for the unsurpassed possible split. This leads to more randomness in the representation. The model uses extremely randomized trees and works well with noisy datasets.

Pros: Typically faster than Random Forest because it doesn't search for the best split. More robust with smaller datasets.

Cons: May sacrifice some accuracy due to the randomness in splits.

Support Vector appliance (SVM):- The SVM algorithm is new for classification household tasks, specially when the statistics is not linearly separable.

Working: SVM find the hyper aircraft that best separate the statistics hooked on two classes. The margin between this hyperplane and the closest data points (support vectors) is maximized to improve generalization. Kernel trick is used when the data is not linearly separable. It maps the datato a higher-dimensional space, where a hyperplane can be used for separation.

Pros: Effective in high-dimensional spaces. Works well for both linear and non-linear data.

Cons: Computationally expensive for large datasets. Difficult to interpret the model, especially with non-linear kernels.

k-Nearest Neighbors (KNN):- The KNN algorithm is a non-parametric, instance-based learning algorithm that works by comparing the query point with its k nearest neighbors in the feature space.

Working: For classification, KNN finds the k-nearest neighbors of the analysis sample and assignsthe class base on the popular class of those neighbors. It uses distance metrics such as Euclidean distance to calculate similarity.

Pros: Simple and trouble-free to understand. No training phase, since it stores the entire dataset for comparison.

Cons: Computationally expensive during prediction, as it requires calculating the distance to all data points. Performance can degrade with high-dimensional data.

2. Deep Learning Algorithm:

Deep Neural Network (DNN):- Deep Neural network (DNNs) are a group of pupils of deep knowledge models that consist of multiple unseen layers between the input and output layers, enabling the reproduction to become skilled at complex, abstract representations of data.

Working: DNNs learn a set of weights and biases for each node in each layer, with the goal of reducing the error (loss function) through back propagation and optimization (gradient descent).Each layer extracts features at varying levels of abstraction. The deeper the network, themore abstract the features it can learn.

Pros: Can model complex, non-linear relationships. Capable of learning from raw data.

Cons: Requires large amount of facts and computational resources. Less interpretable compare to former models like decision trees.

3. SupportingTechniques:

Principal module Analysis (PCA):- PCA is a dimensionality reduction technique used to diminish the amount of character in a dataset while retaining as much variance as possible.

Working: PCA transform the figures into a new set of orthogonal axes (principal components), ordered by the amount of variance they capture. It helps in attribute mixture, improving computational efficiency and removing noise.

Synthetic Minority Oversampling Technique (SMOTE):- SMOTE is new to lecture to class unevenness in datasets by generating synthetic instances for the minority class.

Working: It creates synthetic examples by taking samples beginning the marginal course group and generating new samples by interpolating between them. Helps to balance the dataset and improve model performance on underrepresented classes.

Standard Scaling:- Standard scale is a method of feature scaling that standardize the description by removing the mean and scaling to unit variance.

Working: The formula is $x$scaled = $x - \mu\sigma$ xscaled = σx − μ, where $\mu\sigma$ is the mean and σx is the standard deviation. Ensures that all description make a payment similarly to the model's learning process.

4. SYSTEM ARCHITECTURE AND WORK FLOW:

Data Preprocessing PCA for trait selection with dimensionality reduction. SMOTE to handle class imbalance.Standard Scaling for normalization.

Representation tutoring and Evaluation:-Train and evaluate each model (DT, RF, ETC, SVM, KNN) and deep learning model (DNN).Use cross-validation (e.g.,10-foldcross-validation)to ensure generalization.

Attack Classification

Binary Classification: Distinguishing stuck between benevolent and male volentinter change. Multi-Class Classification: Classifying multiple types of attacks (e.g.,malware, denial-of-service). DNN shows the best performance, especially for multiclass classification tasks.

Performance Metrics:- accurateness, exactitude, bring to mind, and F1-score are second-hand to appraise the presentation of the models.

InferenceTimes: The DNN model' slow computational requirements allow it to achieve faster predictions

The methodology adopted in this research follows a structured data pipeline that begins with data preprocessing and culminates in model evaluation. The Edge-IIoT set dataset is selected due to its comprehensive coverage of diverse IoT attack vectors, ranging from DoS and reconnaissance to data injection and information theft.

IRJCS: International Research Journal of Computer Science ISSN: 2393-9842
Volume 12, Issue 08, October 2025 https://www.irjcs.com/archives
https://doi.org/10.26562/irjcs.2025.v1208.11

Initially, most important section investigation (PCA) is employed for attribute reduction to eliminate redundant and non-informative attributes, thus minimizing computational overhead. To address the class imbalance often found in cyber security datasets, the Synthetic Minority Oversampling Technique (SMOTE) is applied, ensuring a balanced distribution of classes and improving model sensitivity to minority attacks. Following this, Standard Scaler is used for feature normalization to ensure uniformity in the data distribution across all features. The classification framework is structured into three distinct levels: binary classification (normal vs. attack), 6-class classification (broad attack categories), and 15-class classification (fine-grained attack labels). A variety of ML classifiers are trained and evaluated, alongside a DNN that consists of multiple hidden layers capable of learning abstract patterns. Each model undergoes rigorous evaluation using accuracy, precision, recall, F1-score, and testing time metrics. Furthermore, 10-fold cross-validation is used to confirm the model's robustness and resistance to over fitting.

V. MODULE DESCRIPTION

1) Data Preprocessing Module: This module handles the initial stage of data transformation. It implements PCA to diminish feature dimensionality while retain variance. SMOTE is then applied to synthetically over sample minority classes, followed by standard scaling to normalize the statistics. These steps collectively enhance model good organization and diminish computational cost.

2) Binary and Multiclass Classifier Module: The core classification tasks are handled in this module. Binary classification distinguishes stuck between caring and malevolent transfer, while 6-class and 15-class classifiers delve deeper into attack types. Machine education algorithms such as RF,DT,SVM,KNN,and ET Care employed, each bringing unique advantages like interpretability or ensemble robustness.

3) Unfathomable Neural Network (DNN) Module: This module incorporates a deep learning model designed with multiple hidden layers. The DNN is trained using back propagation and dropout techniques to prevent overfitting. Its architecture enables it to learn complex data patterns and relationships that traditional ML models might miss, especially in high-dimensional spaces.

4) Evaluation and Cross-Validation Module: This module calculates performance metrics including accuracy, precision, recall, F1-score, and execution time for each classification type. It also performs 10-fold cross-validation to test the model's ability to generalize across diverse subsetsof the information, thereby preventing over fitting and ensuring stability.

VI.RESULTS

The experimental evaluation demonstrates that the DNN model outperforms other classifiers across all classification levels. For binary classification, the DNN achieves an outstanding 100% accuracy, indicating flawless separation between normal and attack data. In the 6-class scenario, the DNN achieves 96.15% accuracy, showcasing its capability in handling grouped attack types. Even under the more complex 15-class classification, which requires distinguishing between fine-grained attack labels, the DNN still performs impressively with a 94.68% accuracy. The ML models also show competent performance, with RF and ETC closely following DNN in most evaluations. Notably, Smart Sentry's performance is achieved with relatively low computational time test durations are 20.05s for2-class,18.11s for 6-class, and 27.99s for 15-class classification using DNN. This suggests the system's viability in near-real-time environments. Furthermore, high cross-validation scores validate that the models are not over fitting and can effectively generalize to unseen attack types. These results collectively confirm the robustness and effectiveness of the projected hybrid classification approach.

Fig.2 Accuracy of the Results

VII. CONCLUSION

This study presents a comprehensive ML and DL-based IoT attack classification system, "Smart Sentry," that addresses the growing need for intelligent security in connected ecosystems. By integrating powerful classifier like accidental Forest and DNN with advanced preprocessing technique such as PCA, SMOTE, and normalization, the system successfully classifies binary, broad, and granular categories of IoT threats. The use of the Edge-I IoT set dataset ensures the validity of the come within reach of in real-worlds scenario. The deep learning model demonstrates superior performance, especially in complex multiclass classification, due to its ability to capture non-linear relationships in high-dimensional data. The inclusion of 10-fold cross-validation further enhances the trustworthiness of the results by proving the model's generalization capabilities. In conclusion, this work significantly contributes to the domain of IoT cyber security by donation an accurate, fast, and scalable solution for detect and classify a wide array of cyber threats. hope effort possiblywillexplorefederateeducationfordistributedenvironmentsandtheincorporationof real-time anomaly detection for enhanced security in live IoT networks.

ISSN: 2393-9842
https://www.irjcs.com/archives
https://doi.org/10.26562/irjcs.2025.v1208.11

REFERENCES

1. M.A.Ferrag, O.Friha, D.Hamouda, L. Maglaras, and H. Janicke, "Edge- IIoTset: A new comprehensive realistic cyber security dataset of IoT and IIoT applications for centralized and federated learning," IEEE Access, vol. 10, pp. 40281–40306, 2022.
2. K.Gupta, D.K.Sharma, K.Datta Gupta, and A.Kumar, "A tree classifier based network intrusion detection model for Internet of medical things,"Comput.Electr.Eng.,vol.102,Sep. 2022,
3. G.Bhandari, A. Lyth, A. Shalaginov, and T.-M. Grønli, "Distributed deep neural-network-based middleware for cyber-attacks detection in smart IoT eco system: A novel framework and performance evaluation approach," Electronics, vol. 12, no. 2, p. 298, Jan. 2023.
4. G.P.Bhandari, A.Lyth, A.Shalaginov, and T.M. Grønli, "Artificial intelligence enabled middleware for distributed cyber attacks detection in IoT-based smart environments," in Proc. IEEE Int. Conf. Big Data, Dec. 2022
5. M.M.Rashid, S.U.Khan, F.Eusufzai, M.A.Redwan, S.R.Sabuj, and M.Elsharief, "A federated learning-based approach for improving intrusion detection in industrial Internets of equipment networks," Network, vol. 30, no. 11, pp. 1585–1279, Jan. 2023.