# SmartDefend - IOT Security using Machine Learning

1st Dr Saumya Y M
*Associate Professor dept. of CSE*
*St Joseph Engineering College*
Vamanjoor, India

2nd Joyline Rencita Dsouza
*dept. of CSE*
*St Joseph Engineering College*
Vamanjoor, India

3rd Melanie Crystal Miranda
*dept. of CSE*
*St Joseph Engineering College*
Vamanjoor, India

4th Natasha Elizabeth Correia
*dept. of CSE*
*St Joseph Engineering College*
Vamanjoor, India

***Abstract*—In recent years, technology has advanced to the fourth industrial revolution, where the Internet of things (IoTs), fog computing, computer security, and cyberattacks have evolved exponentially on a large scale. Machine Learning(ML) is considered one of the most promising methods for addressing cybersecurity threats and providing security. Several studies have proposed smart intrusion detection systems (IDS) with intelligent architectural frameworks using AI to overcome the existing security and privacy challenges. In this study, we present a systematic literature review (SLR) that categorize, map and survey the existing literature on ML methods used to detect cybersecurity attacks in the IoT environment. It is found that support vector machines (SVM) and random forest (RF) are among the most used methods, due to high accuracy detection another reason may be efficient memory.**

***Index Terms*—cybersecurity, intrusion detection system, support vector machines, random forest**

## I. Introduction

In an increasingly interconnected world, the Internet of Things (IoT) has revolutionized the way we interact with technology, from smart homes and wearable devices to industrial automation and infrastructure management[1]. However, the proliferation of IoT devices has led to a corresponding increase in security vulnerabilities, exposing these interconnected networks to a wide array of cyber threats[4]. From distributed denial-of-service (DDoS) attacks to data breaches and unauthorized access, the diverse range of potential security risks within the IoT landscape demands a sophisticated and proactive security approach. Traditional security measures often fall short in addressing the complexity and scale of these emerging threats, necessitating the integration of advanced technologies to fortify the resilience of IoT networks. In response to this imperative, the integration of machine learning algorithms has emerged as a powerful tool in identifying, preventing, and predicting potential cyber attacks within the IoT ecosystem. By harnessing the capabilities of machine learning, IoT systems can proactively detect anomalies, analyze patterns of suspicious behavior, and predict potential vulnerabilities, thereby mitigating the risks associated with various cyber threats. This research seeks to explore the vulnerabilities inherent in IoT networks and the potential consequences of cyber attacks on these interconnected systems[2]. By analyzing case studies and real-world examples, this study aims to elucidate the multifaceted nature of cyber threats facing the IoT landscape, emphasizing the pressing need for robust and intelligent security measures. Through the application of machine learning algorithms, this research endeavors to propose innovative approaches for preventing and predicting cyber attacks, thereby enhancing the overall security posture of IoT networks and ensuring the protection of sensitive data and critical infrastructure in the digital age.

## II. Literature Survey

### *A. Detecting Cybersecurity Attacks in Internet of Things Using Artificial Intelligence Methods*

In article[1] thoroughly compares Machine Learning (ML) and Deep Learning (DL) approaches within the realm of IoT cybersecurity. It strives to identify the most effective AI methods for detecting threats, emphasizing the importance of secure and intelligent IoT infrastructures. The study explores a spectrum of ML algorithms, including Naïve Bayes, Decision Tree, Random Forest, Support Vector Machine, and DL algorithms like Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM). Notably, 15 studies focus on DL, highlighting its pivotal role.

Most studies rely on the KDD dataset, but the survey urges the need for diverse datasets, emphasizing real-world, real-time IoT systems. Identified threats in IoT span DoS, DDoS, malicious attacks, ransomware, and more, categorized into Probe, U2R, R2L, and DoS. Power grid disturbances, fog-based attack. Recommendations include the widespread use of Support Vector Machines (SVM) and Random Forest (RF) for their accuracy. DL methods, particularly based on Artificial Neural Networks, RNN, and LSTM, prove efficient in detecting IoT malware and attacks. Hybrid approaches and XGBoost, a boosting algorithm, are also highlighted.

The survey recognizes limitations in proposed frameworks, such as methodology and data analysis weaknesses, resulting in low accuracy. Some studies lack reporting on predictive features, and an emphasis on accuracy metrics may limit exploration of specific threats. Future research should focus on next-gen AI algorithms in IIoT, medical IoT, energy IoT, and CPS. Exploring diverse databases and enhancing intelligent frameworks is crucial. Researchers are encouraged to identify vulnerabilities beyond categorized attacks, anticipating the development of novel IDS/AI models for securing the IoT against evolving cybersecurity threats. Additionally, one paper recommends using two datasets to overcome training time challenges and enrich feature sets, showcasing a practical approach to optimize AI models in IoT security.

### *B. A Novel Feature Selection Approach to Classify Intrusion Attacks in Network Communications*

The literature survey [3] extensively compares Machine Learning (ML) and Deep Learning (DL) approaches in the context of intrusion detection[3]. Various feature selection techniques and classification methods have been explored to determine the effectiveness of these approaches.

The study investigates the performance of traditional ML techniques and modern DL algorithms in addressing the challenges posed by cyber threats. The comparison aims to classify which approach is more suitable for different situations in intrusion detection.

Several datasets are utilized in the literature survey to evaluate the proposed methods. Notable datasets include KDD '99, UNSW-NB15, CIC-IDS2017, NSL-KDD, AWID, and IoTID20. These datasets serve as the basis for training and testing intrusion detection models[2].

The literature covers a range of cyber attacks, including remote-to-user (R2L), user-to-remote (U2R), denial of service (DoS), distributed DDoS, and probing. These attacks are classified and detected using various approaches, including signature-based, anomaly-based, and hybrid models like SABADT (Signature- and Anomaly-Based Attack Detection Technique)[1]. Intrusion Detection Systems (IDSs) play a crucial role in monitoring network traffic for indications of malicious activities.

This paper introduces a novel method to improve intrusion detection systems (IDS) against modern cyber attacks[3]. The method involves a unique feature selection technique and a hybrid classification approach for faster and more accurate attack detection. Testing on KDD '99 and UNSW-NB15 datasets showed superior performance compared to traditional machine learning methods, achieving high accuracy in identifying attack types. The methodology also outperformed other studies in terms of feature usage and accuracy, making significant contributions to the field.

The proposed method exhibits superior speed and accuracy in classifying attacks but acknowledges room for improvement, intending to address lower detection rates for certain attack types, test on new unseen attacks, extend evaluation to diverse datasets, enhance algorithmic accuracy, and focus on real-world application considerations for scalability and practical deployment in future research endeavors.

### *C. Cyberattacks Detection in IoT-Based Smart City Applications Using Machine Learning Techniques*

The literature review focuses on smart city frameworks and the application of Intrusion Detection Systems (IDS) in enhancing cybersecurity within the Internet of Things (IoT) infrastructure[4]. Various datasets, such as the UNSW-NB15 and CICIDS2017, are explored to evaluate and identify normal and attack network traffic. These datasets offer a comprehensive understanding of contemporary cyber threats, including DoS, DDoS, PortScan, SQL injection, Infiltration, Brute Force, and Bot attacks[3].

In the context of IDS techniques, the review delves into the selection of anomaly-based Network Intrusion Detection Systems (NIDS) over host-based IDS (HIDS) and signature-based NIDS. The rationale behind choosing anomaly-based NIDS lies in its adaptability to resource-constrained IoT devices and its capability to handle new attacks efficiently. The proposed model emphasizes the tracking of network traffic through fog nodes in close proximity to IoT sensors, thereby facilitating rapid cyber-attack detection.

Data pre-processing involves feature selection using information gain ratio, with the top 25 relevant features chosen for prediction. The selected features are then encoded, and a threshold is applied for optimal filtering. Machine learning techniques, including LR, SVM, DT, RF, KNN, and ANN, are employed to build the IDS scheme[3]. The literature emphasizes the significance of ensemble methods (Bagging, Boosting, Stacking) to improve the accuracy of cyber-attack detection[5].

The review underlines the importance of smart city frameworks in delivering sustainable services, highlighting ongoing projects such as those in Hong Kong and Masdar City. Challenges related to vulnerable urban development plans and the need for sustainability in smart city services are discussed[6]. IoT and related cyber-physical systems play a crucial role in managing sustainability programs, including intelligent transportation systems, smart buildings, and resource usage.

In conclusion, the literature review provides a comprehensive exploration of smart city frameworks, IDS techniques, machine learning algorithms, and ensemble methods in the context of IoT-based cyber-attack detection. The datasets, attacks, classifiers, and frameworks discussed collectively contribute to a holistic understanding of the challenges and opportunities in securing smart city infrastructures against evolving cyber threats.

### *D. IOT-based cyber security identification model through machine learning technique*

The document discusses the development of a cognitive cybersecurity methodology using machine learning (ML) techniques, particularly focusing on threat classification and vulnerability rating severity[5]. The research involves setting up a database that connects various internet vulnerability

repositories and building a pipeline of ML methods to apply them to different groups of information categories.

Natural Language Processing (NLP) techniques are used to transform cybersecurity alerts into numerical representations, and ML algorithms are trained using the collected characteristics synchronously[1]. The study also emphasizes the importance of model selection and ensemble construction to address cybersecurity challenges effectively.

Furthermore, the research evaluates the performance of the proposed ML pipeline, demonstrating the effectiveness of the ensemble of LSTM, NBSVM, and MLP in threat classification. The study also evaluates the model's performance in predicting Common Vulnerability Scoring System (CVSS) scores, showing promising results[3]. The experimental analysis underscores the potential of the ensemble methods in risk categorization and intensity rating in the cybersecurity scenario[2]. In summary, the research contributes to the development of a cognitive cybersecurity methodology that leverages ML techniques to enhance threat classification and vulnerability rating severity, addressing the complexities of modern security repositories and providing valuable insights into cybersecurity analysis[1].

Additionally, the research emphasizes the positive prospects of the ensemble methods in risk categorization and intensity rating in the cybersecurity scenario[3]. The study also discusses the potential of the proposed cognitive cybersecurity methodology in modifying investment choices at various managerial levels. Overall, the conclusion highlights the potential of ML methodologies in enhancing cybersecurity monitoring and evaluation, as well as addressing the complexities of modern security repositories.

## III. System Design

### A. Architecture design

Architecture design serves as the backbone of a project, outlining the systematic arrangement of components and their interconnections. In the realm of Intrusion Detection on IoT, this design is pivotal. It delineates the modular structure, specifying how different components interact and the flow of data within the system. With a focus on scalability, it ensures that the Intrusion Detection System can effectively adapt to the dynamic nature of IoT environments. By fostering modularity and clarity, the architecture design not only fortifies the system against potential threats but also facilitates streamlined development and future enhancements.

### B. Proposed framework for attack classification

The Fig.1 illustrates framework for attack classification. The framework for attack classification in this study leverages machine learning techniques to identify fraudulent network traffic, utilizing the UNSW-NB15 dataset. Initial challenges with the dataset, such as unbalanced data collection, mismatched data types, and missing values, were addressed through thorough data preprocessing. This involved cleansing the data by removing redundant features, introducing derived features, and employing linear regression to impute missing

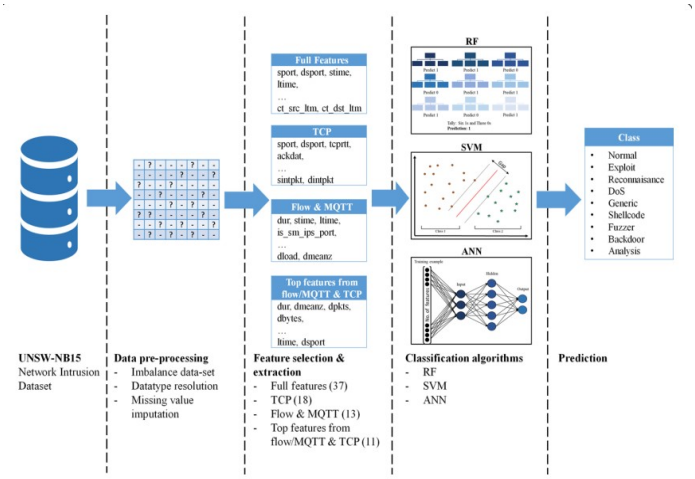


Fig. 1. Proposed framework for detecting cyber-attacks in IoT networks

values. Feature extraction, a fundamental aspect of machine learning, was utilized to improve prediction accuracy, decrease overfitting, and enhance training efficiency. The feature importance approach with Random Forest was employed for feature selection. Additionally, data type resolution led to the elimination of five features, leaving only the binary and multi-class labels for classification. Overall, these preprocessing steps are crucial for optimizing model performance and ensuring the reliability of the classification results in the research.

## Implementation

The basic idea behind machine learning techniques is that they primarily need learning datasets before applying the developed models t s to actual data. The data sets are regarded as being very significant and outdated for identifying current attack types. In this research, the dataset UNSW- NBIS is used to create a customised integrated classification-based model for identifying fraudulent network activity.

The UNSW-NB15 dataset has been introduced as the benchmark dataset in the proposed system. This chapter suggests a variety of machine learning and deep learning methods to improve the network's attack detection rate. The IDS model begins with an analysis of the UNSW-NB15 dataset, followed by data preparation, integrated model proposal, and testing on a test dataset. The performance of the proposed model is then assessed using the test dataset.

The dataset "UNSW NB15.csv" encompasses 49 features alongside a class label, constituting a portion of the larger original dataset known as UNSW NB1S. Descriptions of these characteristics are provided within the "UNSW-NB15 features.csv" file. For binary classification tasks, the dataset "bin data.csv" is utilized, while for multi-class classification tasks, "multidata.csv" serves as the dataset for computation. The UNSW-NB 15 dataset contains 175,341 number of instances in the training Dataset, 82,332 number of instances in testing.

he dataset comprises 49 features representing 9 attack categories, including Analysis, Fuzzers, Backdoors, DoS, Exploits, Reconnaissance, Generic, Shellcode, and Worms.

TABLE I
DESCRIPTIONS OF ATTACKS

| Attack | Description |
|---|---|
| Analysis | Used to penetrate web applications through emails, web scripts, and port scans. Allows unauthorized access to issue commands remotely. |
| Backdoor | Attempts to bring down network services or make resources unavailable for authorized users. |
| DoS | Tries to disrupt network services or resources for authorized users. |
| Exploit | Takes advantage of vulnerabilities in a system or network. |
| Fuzzers | Floods systems with random data to find vulnerabilities and crash them. |
| Generic | Attack disregarding cryptographic implementations. |
| Reconnaissance | Gathers information to evade security controls. |
| Shellcode | Injects code into applications to control compromised machines. |
| Worm | Replicates copies and uses system vulnerabilities to enter. |

Table I provides concise descriptions of nine attack categories present in the UNSW_NB-15 dataset. It outlines various tactics such as analysis for web application penetration, backdoors for network disruption, and denial-of-service (DoS) attacks for resource unavailability. Additionally, it covers exploit techniques targeting system vulnerabilities, reconnaissance methods for evading security controls, and shellcode injections to control compromised systems.

## IV. DATA PREPROCESSING

Pre-processing corresponds to cleaning the data. It involves removing redundant features, features that do not render a high IG and adding derived features—features derived from other features in data. Keeping in mind that certain ML models require details in a given format, i.e., no null values allowed in RF algorithms therefore records with null values must be removed or replaced with substitute values. This issue can be resolved using imputation. Moreover, some ML algorithms cannot process data types other than integers and foats. Tis compatibility issue can be overcome by typecasting the values or removing the features, that do not comply, altogether. Another important dimension of pre-processing of data is that the data should be compatible with more than one algorithms for consistency and for reducing computation complexity. Te pre-processing worked out on the UNSW-NB15 dataset is as follows:

Imbalance refers to an unfair class allocation within the data-set. Data imbalance causes the classifcation to be biased. In UNSW-NB15 data-set, this problem is apparent. Class distribution percentages are shown in Table 2. Normal packets comprise of above 87% of total trafc in the data-set. We use a technique similar to undersampling of imbalanced data-set to overcome this problem. We reduced number of normal packets by 50% but kept the original number of packets of other classes. The remaining data are now 60% of the actual data.

### A. Datatype resolution

Among 49 features, there are 5 features in UNSW-NB15 data-set whose data type is nominal (other than integer / foat) as shown in Table II. We removed these features from the original data-set and are left with 44 features. Second last and last features (43rd and 44th feature) are the binary and multi-class labels, respectively. Te multiclass label is also of nominal type, but during algorithm execution it is converted to integer type using factorization. The Table II shows the distribution

TABLE II
CLASS DISTRIBUTION IN FULL DATASET

| Class | Percentage (%) |
|---|---|
| Normal | 87.35 |
| Exploits | 1.75 |
| Reconnaissance | 0.55 |
| DoS | 0.64 |
| Generic | 8.48 |
| Shellcode | 0.06 |
| Fuzzers | 0.95 |
| Analysis | 0.11 |
| Backdoor | 0.10 |
| Worms | 0.01 |

of different attack classes in the full dataset. Most instances are labeled as 'Normal', while other attack classes have much lower percentages, indicating a highly imbalanced dataset. This distribution is crucial for understanding the prevalence of different types of attacks in the dataset.

### B. Exploratory data analysis

Exploratory Data Analysis (EDA) is crucial as it allows us to delve into our dataset and uncover patterns in the distribution of the 9 distinct attack categories. By examining the percentage composition of each attack type, we gain valuable insights into the relative prevalence of different security threats.

Fig. 2 presents a visual representation in the form of a pie chart, depicting the distribution of labels in the dataset. It reveals that a significant majority, comprising 75.99% of cases, fall under the category of normal instances. On the other hand, approximately a quarter of the cases, accounting for 24.01%, are classified as instances of attacks. This visualization provides a clear overview of the proportion of normal and abnormal labels within the dataset, highlighting the prevalence of normal instances compared to attack instances.

Fig. 3 illustrates the distribution of attack categories in the multiclass dataset. The largest proportion, approximately 48.66%, consists of normal instances, followed by generic attacks at 24.21%. Exploits represent about 15.25% of the dataset, while fuzzers and DoS attacks contribute 8.30% and 5.60%, respectively. Reconnaissance, analysis, backdoor,

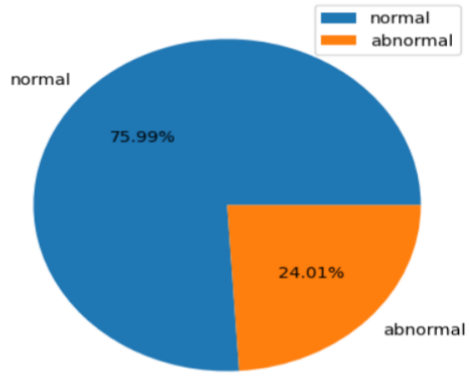


Fig. 2. Pie chart distribution of normal and abnormal labels

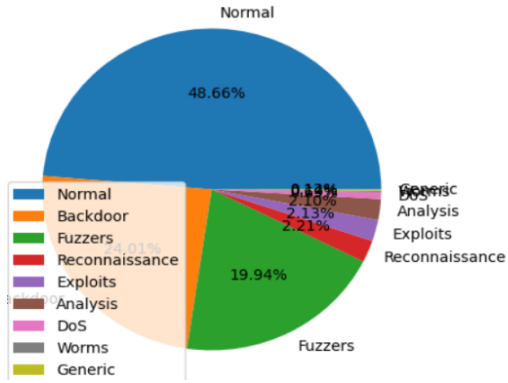


Fig. 3. Pie chart distribution of multi-class labels

shellcode, and worms comprise smaller percentages, ranging from 4.79% to 0.06%.

### C. Label encoding

The dataset initially consisted of 175,341 rows and 49 characteristic features. To ensure data integrity and model efficacy, null values were removed, resulting in a dataset with 82,333 rows and 49 features. Subsequently, to facilitate machine learning algorithms' compatibility with categorical data, a label encoding technique was applied. Under this process, categorical labels representing the main attack category were converted into numerical values, where 'Normal' was mapped to 0 and 'Attack' to 1. Similarly, for subcategories of attacks such as 'Analysis', 'Backdoor', 'DoS', 'Exploits', 'Fuzzers', 'Generic', 'Reconnaissance', 'Shellcode', and 'Worms', unique numerical mappings were assigned ranging from 1 to 9, respectively. This standardized numerical representation aids in algorithmic interpretation and classification accuracy.

### D. Machine learning models used

For our IoT security framework, we employed various machine learning models for classification. These included logistic regression, linear support vector machine (SVM), K-nearest neighbor (KNN), random forest (RF), extra-tree (ET), and XGBoost (XGB) classifiers. Logistic regression provided a baseline analysis due to its simplicity, while SVM effectively delineated classes. KNN classified instances based on proximity, while ensemble techniques like RF, ET, and XGB leveraged collective intelligence for improved accuracy. By integrating these diverse models, our framework aims to detect and mitigate potential threats in IoT environments effectively.

## V. Results

In our study, both binary and multi-class classifications were conducted using reduced datasets, specifically focusing on flow and MQTT features, TCP features, and top features derived from flow and TCP clusters. These experiments were performed on data with three imputations, as previously discussed. Through meticulous parameter tuning, optimal parameters were determined for both binary and multi-class classification tasks. Binary classification was exclusively applied to the full dataset, whereas multi-class classification was performed on both the complete dataset and layer clusters. This approach allowed for a comprehensive assessment of model performance across different subsets of the data, enabling a nuanced understanding of the classification task's intricacies.

### A. Binary classifcation

We employed a suite of machine learning algorithms, including Logistic Regression (LR), Random Forest (RF), Extra Trees (ET), XGBoost (XGB), Support Vector Machine (SVM), Decision Tree (DT), and K-Nearest Neighbors (KNN), for binary classification tasks. The attained accuracies were LR: 76.52%, RF: 92.73%, ET: 92.51%, XGB: 91.73%, SVM: 91.85%, DT: 91.85%, and KNN: 85.85%. Within the binary classification realm, Random Forest, XGB, and Extra Trees classifier were leveraged with optimized hyperparameters. Notably, the Random Forest model achieved the highest accuracy of 92.73% through mean imputation, followed closely by ET at 92.51%, and XGB at 91.73%. Despite comparable accuracies across various imputation techniques, instances of misclassification were evident, attributed to model simplicity and inherent noise in certain samples. The incorporation of confusion matrices provides a visually insightful representation of achieved accuracies and areas warranting further refinement.

TABLE III
ACCURACY OF ALL MODELS IMPLEMENTED FOR BINARY CLASS

| Algorithm | Accuracy (%) |
|---|---|
| Logistic Regression | 77.22 |
| Support Vector Machine | 87.37 |
| K-Nearest Neighbors | 85.85 |
| Random Forest Classifier | 92.73 |
| Decision Tree Classifier | 91.85 |
| Extra Trees Classifier | 92.51 |
| XGBoost Classifier | 91.73 |

The table III presents the accuracy percentages of various machine learning algorithms applied to binary classification tasks. Random Forest Classifier achieved the highest accuracy

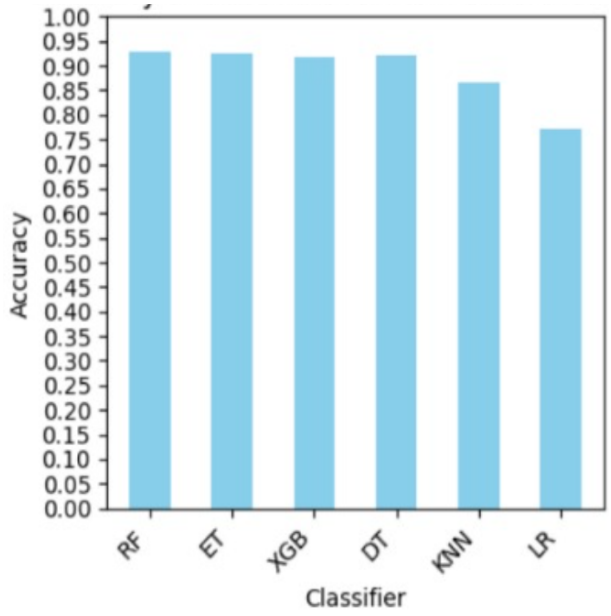


Fig. 4. Accuracy of classifiers for main label prediction

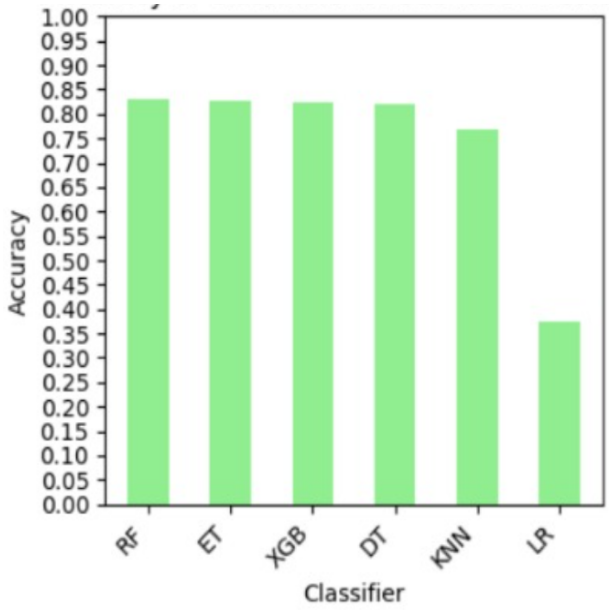


Fig. 5. Accuracy of classifiers for sub label prediction

of 92.73%, followed closely by Extra Trees Classifier at 92.51%. Logistic Regression exhibited comparatively lower accuracy at 76.52%.

Fig. 4 illustrates the accuracies of all models implemented for binary classification.

### *B. Multi-class classifcation*

We applied a range of machine learning algorithms, including Logistic Regression (LR), Random Forest (RF), Extra Trees (ET), XGBoost (XGB), Support Vector Machine (SVM), Decision Tree (DT), and K-Nearest Neighbors (KNN), for multiclass classification. The achieved accuracies were LR: 36.91%, RF: 82.67%, ET: 82.52%, XGB: 82.38%, DT: 81.48%, SVM: (accuracy not provided), and KNN: 76.57%. Notably, Random Forest demonstrated the highest accuracy at 82.67%, closely followed by Extra Trees at 82.52%, and XGBoost at 82.38%. While the models exhibited comparable accuracies, instances of misclassification were observed, likely due to model simplicity and noise in certain samples. The inclusion of confusion matrices offers a visual representation of achieved accuracies and areas for potential improvement.

TABLE IV
ACCURACY OF ALL MODELS IMPLEMENTED FOR MULTI-CLASS

| **Algorithm** | **Accuracy (%)** |
|---|---|
| Logistic Regression | 36.91 |
| Support Vector Machine | 63.03 |
| K-Nearest Neighbors | 76.88 |
| Random Forest Classifier | 82.67 |
| Decision Tree Classifier | 81.48 |
| Extra Trees Classifier | 82.52 |
| XGBoost Classifier | 82.38 |

The Table IV illustrates the accuracy percentages of various machine learning algorithms utilized for multi-class classification tasks. Notably, Random Forest Classifier achieved the highest accuracy of 82.67%, closely followed by Extra Trees Classifier at 82.52%. Conversely, Logistic Regression exhibited the lowest accuracy at 36.91%.

Figure 5 illustrates the accuracies of all models implemented for multi-class classification.

## ACKNOWLEDGMENT

We gratefully acknowledge the invaluable contributions of Dr. Saumya Y M, Dr. Sridevi Saralaya, and the entire faculty of the Department of Computer Science and Engineering for their guidance and support throughout the project. We also extend our appreciation to Dr. Rio D'Souza, Rev. Fr. Wilfred Prakash D'Souza, Rev. Fr. Kenneth Rayner Crasta, and our friends and family for their unwavering encouragement. Their collective efforts have played a pivotal role in the successful completion of this project.

## CONCLUSION AND FUTURE WORK

In this study, we built a robust model to analyze the UNSW-NB15 dataset related to IoT, covering various attacks. We applied both traditional machine learning algorithms and a Deep Learning approach to compare their effectiveness. The evaluation used practical metrics like accuracy, precision, recall, F-measure, and support for both binary and multi-class classifications, the latter involving nine parameters. Our focus is on detecting malicious attacks in network traffic using these techniques, demonstrating adaptability to novel records even during training. While our models showed results comparable to traditional methods, the Multi-Layer Perceptron model achieved superior performance, especially with balanced data. Additionally, the Random Forest classifier performed well

with binary data.

For future work, we propose a comparison of data mining techniques with selective feedback approaches to better capture the nuanced behavior of intrusions and normal activities. The vision includes developing a user-friendly interface for individuals and businesses. Further improvements may involve exploring different types of neural networks, such as LSTM, and diverse architectures within the neural network landscape.

## REFERENCES


[1] Mujaheed Abdullahi, Yahia Baashar, Hitham Alhussian, Ayed Alwadain, Norshakirah Aziz, Luiz Fernando Capretz and Said Jadid Abdulkadir, "Detecting Cybersecurity Attacks in Internet of Things Using Artificial Intelligence Methods: A Systematic Literature Review," Electronics, vol. 11, no. 2, p. 198, Jan. 2022.

[2] HasanAlkahtani and TheyaznHHAldhyani., "Botnet Attack Detection by Using CNN-LSTM Modelfor Internet of Things Applications", Security and Communication Networks 2021 (2021).

[3] Moustafa, Nour, et al, "An Ensemble Intrusion Detection Technique based on proposed Statistical Flow Features for Protecting Network Traffic of Internet of Things" IEEE Internet of Things Journal (2018).

[4] Kelton A. P. da Costa, João Paulo Papa, Roberto Munoz, Celso O. Lisboa, Victor Hugo C. de Albuquerque, "Internet of Things: A survey on machine learning-based intrusion detection approaches," Computer Networks 151 (2019): 147-157.

[5] Moustafa, Nour, et al., "A New Threat Intelligence Scheme for Safeguarding Industry 4.0 Systems," IEEE Access (2018).

[6] Zhiyan Chen, Jinxin Liu, Yu Shen, Murat Simsek, Burak Kantarci, Hussein T. Mouftah, and Petar Djukic, "Machine Learning-Enabled IoT Security: Open Issues and Challenges Under Advanced Persistent Threats," ACM Computing Surveys, vol. 55, no. 5, pp. 1–37, Dec. 2022.