# Designing Secure Model to Classify IoT Vulnerabilities Using Binary and Multi Class Deep Neural Network Classification

**RICHA SINGHAI [1], RAMA SUSHIL [2]**

*[1]Research scholar, Department of Computer Science and Engineering DIT UNIVERSITY, DEHRADUN*
*[2]Professor, Department of Computer Science and Engineering DIT UNIVERSITY, DEHRADUN*
*E-mail address: 1000013410@dit.edu.in*

**Abstract:** Excessive and exponential deployment of the Internet of Things (IoT) devices as wireless network makes it difficult to secure data. It is a high time to make IoT network vulnerable against various attacks. Since multiple vulnerabilities are simultaneously available in the IoT network, therefore, classifying the IoT vulnerabilities is multi-class problem. The paper contributed to design and develop a deep learning (DL) IoT security framework to classify and predict vulnerabilities. The principal component analysis (PCA) is used for feature set reduction and then rectified linear unit (ReLU) activation is used for eliminating vanishing gradient problem for binary classification. Further, the binary classification performance is compared with multi class classification models. Determining benignity or maliciousness of a sample is the main objective of detection models. The classification algorithms' objective is to assign each sample into one of the subsequent classes: beneficent, tsunami, Mirai, or Gafgyt. The classification of attacks is implemented using the deep neural network (DNN) based learning model. The N-BaIoT multiple vulnerabilities dataset is used which is of unbalanced nature. The paper suggested employing the oversampling approach Synthetic-Minority Oversampling Technique (SMOTE) for balancing the data set and compared performance with Random Oversampling and Undersampling (ROU) approach. Performance is compared with conventional decision tree, Random Forest (RF). The accuracy with proposed DNN approach is more than 99.99% in all cases.

**Keywords:** Internet of Things, Vulnerabilities, Binary Classification, ReLU, Multiclass classification, Vanishing Gradient Problem, Deep Neural Network, Oversampling, SMOTE

## 1. INTRODUCTION

The excessive as well as exponential deployment of IoT devices like wireless networks makes it vulnerable against various malicious attacks. Ever increasing deployment of IoT devices makes it challenging to safeguard data [1]. The IoT technology integrates heterogeneous technological devices using an immense amount of data generated through the billions of linked gadgets. As the quantity of Internet of Things, or IoT, gadgets linked to the network grows at a swift pace, so do network threats like floods and denial of service attacks (DoS) [2,3]. These attacks interrupt networks and deprive IoT devices and their services. Because type-imbalanced data in ML can be a common difficulty, there is a noticeable slant in the class distribution of data in binary [4] as well as multi-class classification issues [4, 5]. Ge, and Aldhaheri et al. (2020) have opted deep learning (DL) based intrusion and IoT attacks detection approaches are presented [6, 7]. Wang et al. [8] have proposed convolutional neural network (CNN) based binary classification problem but approach was image based. Various IoT vulnerabilities issues are addressed [9-14] for improving data security. The classification problem using most frequent Botnet data set for Bashlite and Mira data classification are presented in the [15-18]. The specific case of graph based learning is presented in [19], Accuracy improvement is still an open challenge due to data imbalance. Thus oversampling techniques by Random Oversampling and Undersampling (ROU) [20], Synthetic-Minority Oversampling Technique (SMOTE) [21, 22] and PF-SMOTE [23] used for the elimination of the class imbalance for classification problem. Thus, this paper aimed to compare DL based classifiers for various over sampling approaches.

Nevertheless, a lot of different devices used in the IoT ecosystem make it hard to identify IoT threats with conventional rule-based security solutions. Creating the best security models for every kind of device is difficult. Another approach that enables the development of ideal security models employing empirical data from individual devices is machine learning (ML). Some frequent security issues for IoT devices are poor authentication as IoT devices are renowned for using weak and default passwords, several big botnets, including Mirai, infected numerous gadgets simply by signing in with default as well as entered passwords. Today's cyber security is a significant difficulty since it takes a lot of work to prevent attacks when connecting devices through Internet or IoT devices within smart-cities, healthcare (HC), government organisations, and home appliances, among other places. The present study uses machine learning (ML) to detect IoT attacks or vulnerabilities. Research is carried out in two phases, initially binary classification is tested using neural network (NN) and in second pass multi-class classification problem is applied for comparison. Concept is to create ML-based models to represent every kind of device as well as concentrate on botnet assaults that target different IoT devices. Paper make use of the N-BaIoT dataset, which was created by inadvertently implanting botnet attacks (Bashlite or Mirai) into a variety of IoT devices, such as a webcam, doorbell, baby monitor, and security camera.

*Contributions of Work*

The prime concern of the paper is to classify and detect the various IoT vulnerabilities. Paper contributed to design and develop DL based IoT security model to classify and predict vulnerabilities. This study has been conducted on bot net traffic data, gathered from nine commercial IoT devices infected by authentic botnets from two classes (Bashlite or Mirai), however the study has been performed on an unbalanced sample dataset of size (706258, 116) from N-BaIoT dataset. Paper contributed in two pass, first vanilla deep neural network (DNN) is used for solving vanishing gradient problem in binary classification. The large features set is minimized using the principal component analysis (PCA) and then used Rectified Linear Unit (ReLU) in the hidden layer is to solve the problem of vanishing gradient problem and use sigmoid activation function at the output layer. But to explore the real scenario of multiple hybrid attacks in the second pass the binary classification case is extended to multiclass problem for data is gathered from 9 commercial IoT devices authentically infected by Mirai and Bashlite. Therefore in the paper this has been considered as a classification problem for the targets where malicious traffic count is 650666 and benign count 55592. This is an imbalanced dataset. Paper proposed to use oversampling technique random sampling ROU and SMOTE to balance the dataset and then used three algorithms to understand if these techniques are effective in classifying the targets labels or not. 141254 records chosen for the study extracted as 2% random sample from around 89 files of dataset. The accuracy of detection is compared for binary classification using CNN and three multiclass models as Decision tree, Random Forest (RF), and the proposed DNN.

## 2. CHALLENGES AND IOT VULNERABILITY

The IoT vulnerability refers to the susceptibility of IoT devices and networks to cyber-attacks and unauthorized access. This is a significant issue due to the increasing number of connected devices and the often weak security measures. This paper has focused on most frequently available botnet attacks classification and detection. The dataset is traffic data set and has many attack classes. And this makes the classification problem is most challenging. Selection of most prominent attacks is essential for accurate prediction. Basic problem is to detect malicious or not. Nowadays, cyber security is a huge challenge. If one wishes to link devices through the internet or connected devices across the smart cities, healthcare, government organisations, house hold appliances etc. it is a very laborious task for preventing assaults.

There are many security challenges for IoT networks as represented in Figure 1. The IoT vulnerability assessment: This is the process of identifying and analysing potential vulnerabilities in IoT systems. This is crucial for mitigating risks and protecting sensitive data. One challenge is the vanishing gradient issue. Another major challenge to deal with IoT networks is data imbalance. The presence of multi class vulnerabilities and several large botnets as in Figure 1 may cause imbalance. This paper has presented the methods to overcome and handle the both these issues as separate cases of binary and multiclass classification problems.

## 3. THE VANISHING GRADIENT PROBLEM (VGP)

The vanishing gradient problem occurs in artificial neural networks, particularly recurrent neural networks (RNNs) used for processing sequential data. In this paper for implementing the binary classification problem in part one vanishing gradient problem is used. During back-propagation, the gradients used to adjust the weights in the network can become very small as they flow backward through multiple layers. This makes it difficult for the network to learn long-term dependencies in the data and can hinder its performance. While the vanishing gradient problem itself isn't directly vulnerability, it can indirectly impact the security of IoT systems in several ways. The first contribution of the paper is to validate and test data to be vulnerable or not using binary classification using deep RNN. An example layered architecture representation of binary classification problem is shown in Figure 2. Where benign means not vulnerable (0), and Malware means vulnerable (1).

The actual Feature set example of Bot IoT data set is shown in the Figure 3. The dataset has large 115 features set and is therefore minimized using the principal component analysis (PCA) and then used ReLU activation in the hidden layer for eliminating the problem of vanishing gradient problem and use sigmoid activation function at the output layer. The binary classification problem proposed a mutual information-based feature selection and Pearson correlation coefficient (PCC) approach with machine learning methods to enhance the efficiency and performance of IoT botnet attack detection. A freely available benchmark dataset was used to show the benefit of the proposed feature selection method. The activation functions used for the proposed approach are shown in the Figure 4 and are mathematically defined in equations;

$$f_{sigmoid}(x) = 1/(1 + e^{-x}) \quad (1)$$

$$f_{ReLu}(x) = \begin{cases} 0 & if\ x < 0 \\ x & if\ x > 0 \end{cases} \quad (2)$$

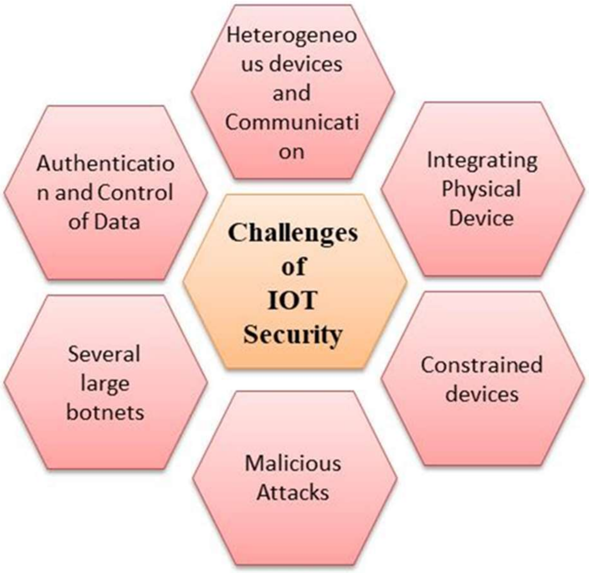
Heterogeneous devices and Communication
Authentication and Control of Data
Integrating Physical Device
Challenges of IOT Security
Several large botnets
Constrained devices
Malicious Attacks

Figure 1. Challenges of IoT networks and devices

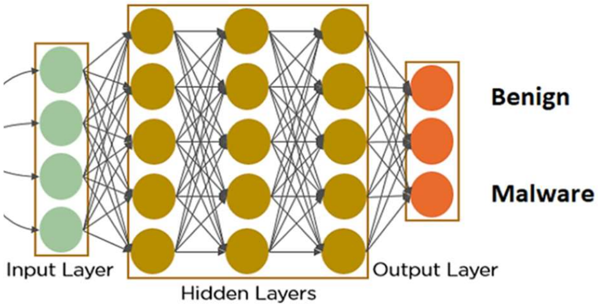

Figure 2 Basic layered architecture of the deep RNN binary classifier network

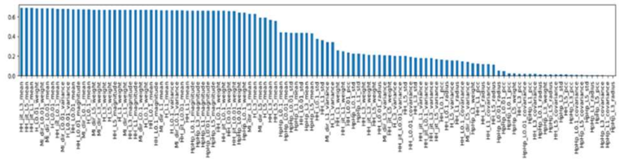

Figure 3. Actual available features in the dataset

The sigmoid is s shape exponential (nonlinear) function with x thus may have VGP issues. ReLU does-not saturated with positive input x. When x <= 0, ReLU's derivative/gradient is 0, while whenever x > 0, it is 1. Multiplying ReLU derivative yields a value of 0 or 1, and therefore there will be no vanishing gradients.

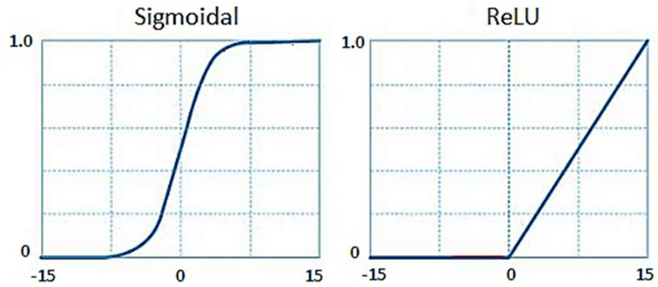

Figure 4. Sigmoidal and the ReLU activation functions example for NN

## 4. VULNERABILITIES CLASSIFICATION METHODS

There were many approaches designed for detections and classification of the IoT vulnerabilities. The major focus of this paper is to explore the ML based detection approaches. The most frequent methods are classified in the Figure 5.

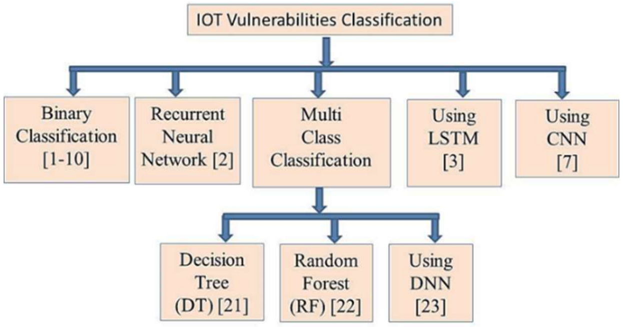
IOT Vulnerabilities Classification
Binary Classification [1-10]
Recurrent Neural Network [2]
Multi Class Classification
Using LSTM [3]
Using CNN [7]
Decision Tree (DT) [21]
Random Forest (RF) [22]
Using DNN [23]

Figure 5. Methodologies for classification of the IoT vulnerabilities

## 5. LITERATURE REVIEW

There are many types of IoT vulnerabilities or attacks are available and many databases are used for classifying them. Therefore this section has reviewed the most relevant research. Sarah Bin Hulayyil et al. [1] have proposed an approach to study devices employed in IoT contexts for vulnerability identification. They investigated ML algorithms on a variety of datasets, including IoT23. They proposed ML pipeline for identifying IoT challenges and most prevalent potential vulnerabilities were examined. The ML and DL methods are reviewed using low-performance, manufacturing, storage constraints, and use of energy, also security difficulties. Adat et al. [2] stated that using Worldwide Web and IoT devices prevalence of Internet has increased recently. As a result of all these advancements, cyber-attacks also rise. Due to the fact that there are more devices to corrupt and less secured locations to attack, distributed disruptions of service (DDoS) assaults have sharply grown. They have developed a DDoS mitigation system to protect IoT networks from DDoS attack in there study. Richa et al. [3] showed that various security and privacy issues generated by the Iot network in real time system. Ferrag et al. [4] introduced DeepCoin, a novel blockchain (BC)-powered deep neural energy framework for smarter grids. The DeepCoin structure employs two strategies: one based on DL and other on BC. They assess the proposed software effectiveness employing three separate datasets: the CICIDS2017 knowledge set, a power systems dataset, including an online robotic (BotIoT) dataset. Ge et al. [5] in (2019) has applied DL to provide a unique attacks detection for IoT traffic dataset attacks such as DDoS, surveillance, and data theft attempts, They uses multi-class classification via creating feed-forward artificial neural network (ANN) framework by embedding layers that encode highly dimensional category data. They achieve accuracy and F1 score of 0.99 for their approach. Ge et al. [6] letter in (2020) created two FNN algorithm by using general packet-level includes obtained by extracting header data from each packet. The outcome of the evaluation showed that the models performed well, with exceptional precision and low misunderstandings around additionally FN and FP. Sahar Aldhaheri et al. [7] had proposed DNN algorithm in combination to Dendritic Cells Arithmetic (DeepDCA) that detects intrusions using Self-Normalization NN. Algorithms that, when coupled with a digitized intrusion detection system (IDS), performs better for classification task using support vector machine (SVM), K-Nearest Neighbour (KNN), multi perceptron (MLP) NN and proposed IDS with respective accuracy of 96%, 91.69, 97.99% and 98.3% for Bot IoT dataset.

Wang et al. [8] have proposed intrusion detection by binary classification using Imaging Improved CNN. Have used over sampling method to enhance the volume of original information. Use of DL a 1D feature is transformed to 2D image then extracting features using the pooling and layers of convolution thus lowers the overall dimension of data. The Tan h value is included as an activation parameter along with Dropout approach enhances the prediction models. This contemporary algorithm employed KDDCup99 assortment of data. Their ID-IE-CNN approach has achieved accuracy of 97.1 % with F1 score of 0.9949. Lonzetta et al. [9] introduce the Bluetooth protocol for IoT and discuss its security, risks, limitations, including risk mitigation measures. The research focuses on understanding attack dangers along with mitigating techniques when using Bluetooth capabilities on our devices. Rajagopal et al. [10] proposed a meta-classification approach based on layered generalization. They investigated multiple datasets, including the UNSW NB-15 packet-based dataset and the UGR'16 flow-based dataset. A team model using a meta-classification strategy enabled by stacking generalization is shown to give better accurate projections than an emulated one. Ivan Cvitić et al. [11] stated that stringent hardware limits of existing devices and data transit technologies, this article looks at the risks and flaws in IoT environments with potential mitigation strategies. In light of research findings recommendations are made for identifying safety concerns of various layers of architectures and security threats determined by the type of application of the IoT concept. According to how the idea is applied, the risk rating will allow future research to focus on the building's highest danger levels and the implementation of appropriate security measures.

Three methods aimed on security concerns associated with IoT systems, according to Vesna Antoska Knights et al. [12]. Two of these methods rely on modeling based on anomalies and signatures to detect intrusions. The final section of this section includes a real-world example (entry alert system that uses Cobra at a poultry farm) that is currently in use. The PIR, (privacy, integrity and reliability) using alternative forms of cryptography and security. Shafiq et al. [13] evaluated the "transfer-ability" for a self-encoding model amongst identical and dissimilar IOY devices. Testing was carried out to reconstruct the algorithm which encodes the models' central layers employing a little bit of new input from another device. The first levels of the automated encoder indicate the high level basic properties shared by related devices, whilst the fundamental level of ANN catches more unique qualities. They also tested the effectiveness of the automatic encoder-oriented anomalous model on flow-based network activity data using the CIC-IDS2017 information set. 5.Richa et al. [14] investigated various vulnerabilities in IoT network sub system. Marzano et al. [15] investigated Mirai and Bashlite botnet assaults. Pay close attention to the ways the infection changes over time, including behavioral variations in botnet controllers. They use observation data collected over an eleven-month span from 47 honeypots that are our findings shed new light on those botnets and lend support to previous studies by demonstrating that spyware, network management, and malicious activity are becoming increasingly sophisticated. This study looks at Mirai, a successor to Bashlite botnets, and finds that it uses more durable host and management infrastructures, as well as more powerful attacks.

Kim et al. [16] investigated how assaults disrupt networks along with deprive IoT devices of their functionality. The IoT of many devices utilized within IoT ecosystem made it difficult to detect IoT threats using traditional based on rules security solutions. Developing the appropriate security models for all types of devices is difficult. ML and DL are two separate approaches that allow for building of optimal security frameworks using actual data from particular devices. Proposed technique searches for wormhole assaults using encryption. The maximum F1 score of 1 is achieved with a decision tree (DT) and a RF approach. Sarika Choudhary et al. [17] proposed model that use Linear Discriminant Analysis (LDA) to compress large datasets into low-dimensional space while retaining less significant characteristics. The particle component analysis (PCA) is then used to select valuable features. SVM and an ANN classifier were used in combination to achieve goal of reducing false alarm rates while boosting recognition rates. The NSL-KDD and UNSW-NB15 techniques are used. Richa et al. [18] investigated cyber security system with the help of machine learning algorithms based on the accuracy and recall parameter.

Alasmary et al. [19] provide a control flow graphs (CFG)-based IoT malware detection approach. They compared essential characteristics of IoT malware against other types of ransom ware. Using CFGs from over 6000 viral and benign IoT dismantled the specimens; they show 99.66% detection accuracy. According to Chaw Su Htwe et al. [20], attackers primarily target IoT devices. Once these devices become infected with spyware, criminals can instruct them to change into bots that attack organizations with the goal of not only stealing important data but additionally disrupting the network. The proposed detection architecture is based on machine learning approaches, including CART methodology and N-BaIoT accessible IDS data. Hawla et al. [21] describe their proposed strategy for creating classifiers from unbalanced data. A dataset is considered imbalanced if classification categories are not represented equally. Practical data sets comprise largely of "acceptable" cases, with small fraction of "abnormal" or "interesting" instances. Furthermore, it is typically far more expensive to accurately identify an uncommon (fascinating) specimen as a typical instance compared to it is to commit the opposite mistake. Rather than just underestimating the overwhelming class, our study shows that our method of oversampling the rare category while under-sampling the majority or typical category can improve classifier efficacy. Author have also proposed to smooth the data imbalance for building classifier from data. The method was Synthetic-Minority Oversampling Technique (SMOTE). Rather than just underestimating the overwhelming class, our study shows that our method of oversampling the rare category while under sampling the plurality and normal class can improve classifier efficacy. Joloudari et al. [22] have used the over sampling method SMOTE for decreasing the amount of majority classes to balance the data or creating novel information in the minority class artificially. In light of this, we examine the efficacy of techniques in this research that combine a number of well-known unbalanced data solutions—that is, oversampling and under sampling—with neural networks such as DNNs and CNNs. They have utilized the KEEL, Z-Alizadeh Sani, and breast cancer dataset to assess our approaches. Chen et al. [23] have proposed parameter free PF-SMOTE to avoid integrating noisy samples and generates enough relevant artificial instances. The results of the study show that suggested CNN approach with PF can compete with both the traditional SMOTE and its cutting-edge variations. Yi et al. [24] introduced ASN-SMOTE, an efficient and simple oversampling approach-based on the SMOTE oversampling and its k-near neighbors. Aadaptive qualified synthesizer selection ASN-SMOTE initially filters away noise form the minority class by determining if the closest neighbor of each minority instance belongs to the majority as well as minority class. Next, ASN-SMOTE uses the overwhelming example that is most similar to every minority scenario to effectively detect the choice border.

Damtew et al [25]. In the context of network intrusion detection structures, this study incorporates cooperative selection of the features (CoFS) approach with a Simulated Multi-minority (SMMO) Approach, which selects relevant features along with improves the class break down for the data set in an attempt to enhance the identification precision for the most significant minority classes. They evaluated structure with J48, AdaBoostM1, BayesNet, and arbitrary forests. Author also proposed the context of network intrusion detection structures. Research integrates cooperative selection of features (CoFS) technique with a Simulated Multi-minority (SMMO) Approach selects pertinent features and improves the class breakdown of the information set in an effort to improve identification precision of most severe minority classes (i.e., user-to-root and remote-to-local assaults). They tested structure using J48, AdaBoostM1, BayesNet, and arbitrary forests. Abbasi et al. [26] thoroughly explore earlier methods to using intelligent machines techniques for IOT intrusion detection, and we propose two algorithms for feature extraction and categorization. The first strategy extracts and classifies features using Logistic Regression (LR), whilst the second method categories using an ANN Six gadgets in the N-BaIoT information set. Results from simulations indicate that LR approach is superior and produces classification reliability of 90% compared to DL approaches in areas of precision, recall, and the F1- score achieved. Author thoroughly explore earlier methods to using intelligent machines techniques for IoT intrusion detection, and we propose two algorithms for feature extraction and categorization. The first strategy extracts and classifies features using Logistic Regression (LR), whilst the second method categorizes using an ANN Six gadgets in the N-

BaIoT information set. Results from simulations indicate that utilizing logistic regression as an approach is superior and produces classification reliability of above 90% when compared to various additional DL approaches in the areas of precision, recall, performance, and the F1- score achieved. McDermott et al. [27] describe a method for recognizing botnet behaviors on consumer IOT networks and devices. A detection system using the opposite direction short-term memory oriented neural network recurrent network (BLSTM-RNN) is built by applying a novel approach known as DL. Word embed is used to identify text and translate attack packets into tokenized integer formats. The database references are Koggel [28] and Doorbell [29].

Bagui et al. [30] proposed an impact of sampling on effectiveness of multi-class learners using ANN is examined. Comparison safety information sets KDD99, UNSW-NB15, UNSW-NB17, and UNSW-NB18 were subjected to a variety of replication techniques: at random under estimating, arbitrarily oversampling and arbitrarily reducing and arbitrarily the oversampling arbitrary under-sampling with artificial minority Oversampling approach and random reducing with ASM techniques. The outcome was assessed using macro F1-score, macro accuracy, and micro memory. Meidan Y et al. [31] provide a statistical study of a unique system-based anomaly detection technique that uses deep auto encoders to find anomalous network activity originating through hacked IoT devices by retrieving network-wide behavior samples. To evaluate their technique, and used Mirai and BASHLITE attacks, the two of the most recognized Connected- devices based botnets, to infect an aggregate of nine commercial IoT devices in our laboratory. Examination revealed that proposed method could swiftly and precisely identify attacks as soon as they launched from compromised IoT devices that were part of a botnet. Luque A et al. [32] proposed case study and use binary encoder simulation to establish an organized investigation of this influence. When handling data that is imbalanced, a number of performance measures based on the two-dimensional matrix of confusion can be compared thanks to a set of algorithms and numerical indications that are obtained. During the work, an improved method of measuring balance is identified, outperforming inequality ratio employed in earlier research.

Charlie Obimbo et al. [33] employed ML approaches to detect credit card-based fraud. As a result, utilizing a real-world dataset of customer credit card activity, this study used the machine learning approach LightGBM to detect illegal credit card use. The Artificial Minority The scheme of overs technique (SMOTE) is the data selection strategy used when there is an imbalance in information between the dishonest as well as non-fraudulent classes. The employed ML algorithm to detect credit card-based fraud. As a result, utilizing a real-world dataset of customer credit card activity, Leevy et al. [34] uses Bot-IoT to construct predictive algorithms for detecting assaults modeled by dataset examples belonging to the subcategories of recording keystrokes and knowledge theft, as well as stealing data. They focus work on assessing impact of group methods for selecting features (FSTs) on the effectiveness of categorization for these particular assault cases. They employ non-ensemble LR, DT, NB, and MLP based classification. Popoola et al. [35] demonstrated an efficient DL-based DDoS detection technique that can handle data from substantially skewed network traffic. To be more specific, the SMOTE generates more minority samples in order to achieve class balance, but the Deep Recurrent Neural Networks (DRNN) use equal network traffic data for biassed classification. Many deep training and Machine Learning (ML) approaches were developed to detect botnet attacks in IoT dev. R. Biswas et al. [36] have installed malevolent fake node(s) or machine(s) known as Botnet(s) to organise security breaches such DDoS, and Defeat of Services (DoS), The suggested technique uses cutting-edge ML including ANN, Gatted Repetitive Units (GRU), and Long or short- term memory (LSTM) models. Hassan et al. [37] suggested method for traffic category detection on severely unbalanced datasets is assessed in a multi-class issue environment. Chih-Chieh et al. [38] Two common DDoS attack methods were deployed. Chirag et al. [39] showed comparative study for multiclass classification.

A very new transfer learning auto-encoder model is introduced to detect noisy DDoS malware attacks such as Mirai and Bashlite in IoT devices. The most well-known oversampling approach is the SMOTE.

The major research shortcomings to be addressed are as follows;

- Most reported efforts utilise only single binary classifiers that distinguish amongst malicious and benign messages therefore fail to recognise the nature of IoT attacks. The multi-class classification of attacks is still open field of research.
- Binary classifiers are sensitive to the vanishing gradients problem and are to be addressed in recent designs.
- Many research works have not adopted the data balancing methods and this restricted the accuracy of classification.
- The Bot-IoT dataset is frequently being used in literature for DL based classification problem. But

large data features set has to be optimized for better results.

## 6. PROPOSED DL WITH OVERSAMPLING TECHNIQUE SMOTE

It's crucial to remember that the database mentioned above can cause some unforeseen issues. For example, un-resampling techniques could ignore important information needed for classifier training. For IoT unbalanced datasets, oversampling is suggested as a viable pre-processing method for creating new minority instances that will balance the dataset. On the other hand, algorithms that oversample may result in over fitting. It can be observed from the Figure 6 that paper proposed binary classification using RNN, and multi class classification using DNN as proposed approach. The feature set is reduced for classification. The oversampling method SMOTE is adopted to balance the dataset. By lowering the dimensionality of data, PCA allows us to better generalise ML models. Paper proposed to use Benign and the malicious as binary class and Mirai and Bashlite attacks and Benign as multi classes for classification using Bot–IoT attacks Doorbell dataset. Paper load data and create target classes as "benign" and "malicious" and thus produces an imbalanced dataset. Finally, before classification uses oversampling to create balanced target class in turn improves accuracy. The sequential design flow diagram for the proposed algorithm is shown in Figure 7.

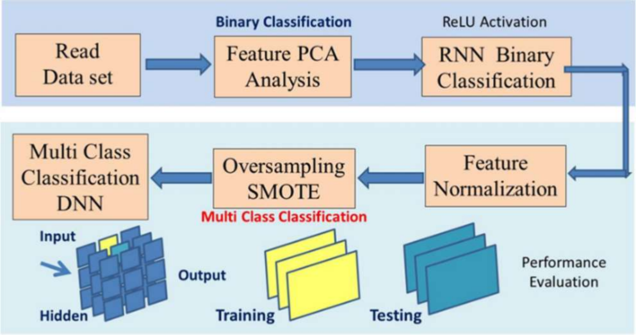

Figure 6 Proposed multi class classification systems

## 7. RESULT AND DISCUSSION

The paper has compared multi-class models of classification against binary classification efficiency. Determining an object's benign or malicious nature is the main goal of binary classification models. Assigning all samples to one of the subsequent categories—benevolent, tsunami, Mirai, or Gafgyt—was the goal of proposed classification approach. The proposed learning model is based on DNNs is used to implement attack (vulnerabilities) classifications. First details of the features available in database are described. Paper presented the results of three different experimentations sequentially. The first experiment is performed to represent the result of binary classifications using only single hidden layer. The performance is compared with the three different classifiers using two hidden layer architecture for better and fast convergence. The comparison of classification results with two oversampling methods as (SMOTE and ROU) are presented as an experiment. And the final experimentation has been presented for multiclass classification with three different attacks classes.

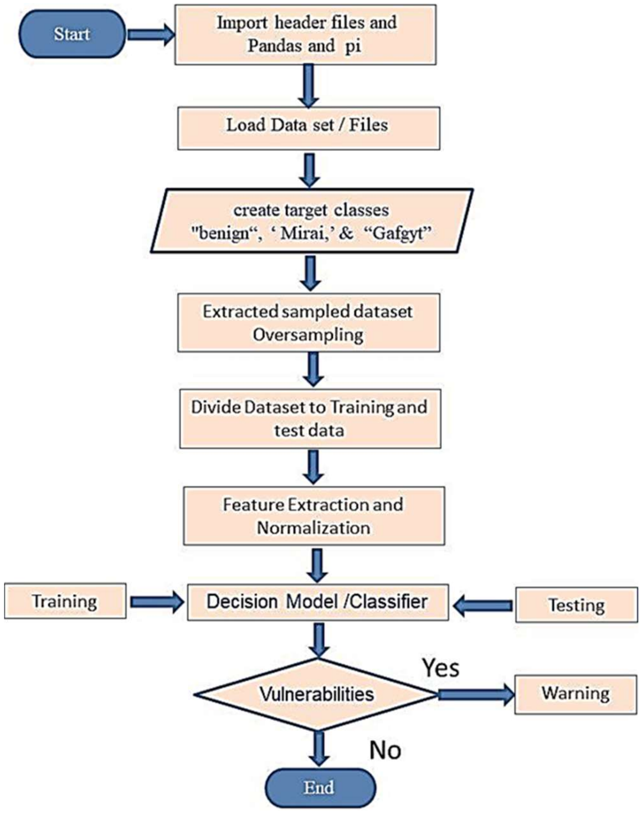
Start
Import header files and Pandas and pi
Load Data set / Files
create target classes "benign", 'Mirai,' & "Gafgyt"
Extracted sampled dataset Oversampling
Divide Dataset to Training and test data
Feature Extraction and Normalization
Training
Decision Model /Classifier
Testing
Vulnerabilities
Yes
Warning
No
End

Figure 7. Flow chart of proposed DL methodology for IoT vulnerabilities classifications

### 7.1 Database Description

The N-BaIoT standard traffic attack dataset is used for evaluation in the current research work. The feathers of the used data base are represented in the Table 1. This is designed to detect botnet attack types, using 9 commercial IoT devices that provided the real traffic data. The IoT devices were attacked by two botnet attack families, namely Bashlite and Mirai. In total, there are about five million items of data, grouped in separate files. Each file contains 115 features and a class labels and thus jave rich set of attacks data for 9 devices.

TABLE 1 STATISTICS OF N-BAIOT DATASET

| Description | Feature name | Number of Instances,% |
|---|---|---|
| IoT device type | Security cameras<br>Webcam<br>Smart baby monitor | 4<br>1<br>1 |

| | | |
|---|---|---|
| | Thermostat<br>Smart door-bell devices | 1<br>2 |
| General Features | Total number of Instances<br>% of features in dataset<br>Time windows | 6,273,053<br>115<br>100 ms, 500 ms, 1.5 s, 10 s and 1 min |
| Distribution of data (2 classes) | # of "Benign" records<br># of "attack" records | 555,932 (7.23%)<br>7,134,943 (92,77%) |
| Distribution of data (3 classes) | # of "Benign" records<br># of "Bashlite" records<br># of "Mirai" records | 555,932 (7.23%)<br>2,838,272 (36,90%)<br>4,296,671 (55,87%) |

The dataset has also been constructed to server binary classification as well as multi-class classification, where the target class labels take values of "benign" or "TCP attack" for binary classification and "Bashlite" or "Mirai" attack types for multi-class classification. Prior experimental studies on the detection of IoT botnets or IoT traffic anomalies typically relied on emulated or simulated data. In contrary, this dataset enables empirical evaluation with *real* traffic data, gathered from nine commercial IoT devices infected by authentic botnets from two families in an isolated network. It facilitates the examination of Mirai and Bashlite, two of the most common IoT-based botnets, which have already demonstrated their harmful capabilities.

*Performance Parameters:*

The confusion matrix is used to determine the true positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN) respectively. The parameters used for the evaluation are defined as follows-

$$precision = \frac{TP}{TP + FP} \quad (3)$$

$$recall = \frac{TP}{FN + TP} \quad (4)$$

$$F_{1_{Score}} = 2\,x\,\frac{precision * recall}{precision + recall} \quad (5)$$

$$Accuracy = \frac{TP + FN}{TP + FP + FN + TN} \quad (6)$$

### *7.2 Result of the Single layer Binary Classifications*

In this alternate method have used DNN for binary classification, one of the class is representing the 'attack state (0) as malicious and the other one is representing safe state (1) as Benign. For the experiment the attacks class data are combined (Mirai and Gafgyt attack data) are combined and represented as malicious (1). And then as binary classifier case attack shows 1 and normal data shows 0. The 3135 samples are considered as Benign under 0 class and 2683 data of combined attacks are taken as the class1. The data count under the consideration for binary classification test is shown in Figure 8.

To calculate the loss proposed method used binary cross entropy loss as our problem is of binary classification. The activation used in the hidden layer is ReLU to solve the problem of vanishing gradient problem and use sigmoid activation function at the output layer. The optimizer which we used is RMSPROP that takes care of momentum and adaptive learning rate. To train the model we used 80% of the original dataset and 20% for the testing. Our model is converged in 100 epochs. In testing we have got the accuracy of 99.9% and loss of 0.001.

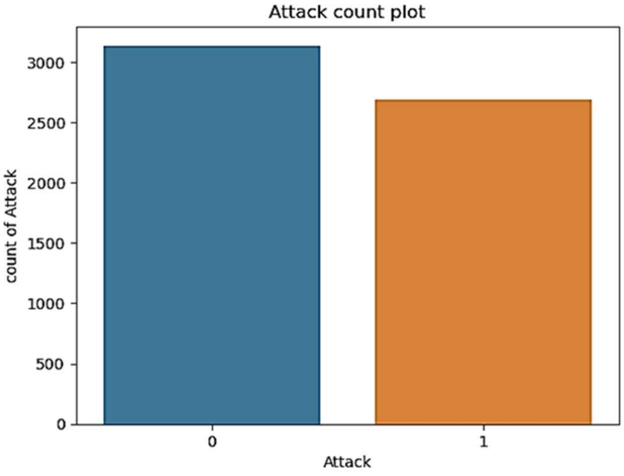

Figure 8. Instances count used for the experimentation of binary classification

Split data into train and test i.e.70%, 30% respectively. This will be required for avoid over fitting for our model. The dataset which we used is having one hundred fifteen features which are very high so we reduced the feature count to eight by the help of PCA. The PCA not only reduced the number of features but also extracted the new features in such a way that which the new features become more informative in comparison to the previous features. For the DL we used vanilla deep neural feed word network with one hidden layer. The configuration of the neural network [NN] is shown in the Table 2 given below. Number of neurons used in the input layer is 8 and in the hidden layer 38 neurons are used.

This proposed experiment of the binary classification with RNN uses the only single hidden layer and large number of hidden layer neurons and the large number of epoch (as shown in Table 2) are required for achieving the optimum accuracy. The comparison of the accuracy and the loss during training and testing phase are shown for different epochs in Figure 9 and 10 respectively. It is clear from the Figure 9 and 10 that minimum loss calculated for the classification is 0.0014596, and the maximum achievable accuracy is 99.948436%. The major drawback of this single layer classifier is its requirement of large epoch to converge. There is lot of possibility of further improvement in accuracy and time performance of this approach.

TABLE 2. EVALUATION OF THE BINARY CLASSIFICATIONS NETWORK

| Number of hidden layer | Number of neurons used in the hidden layer | Optimizer | Epochs | Activation function at hidden layer | Activation function at output layer |
|---|---|---|---|---|---|
| 1 | 38 | RMSPROP | 100 | ReLU | Sigmoid |

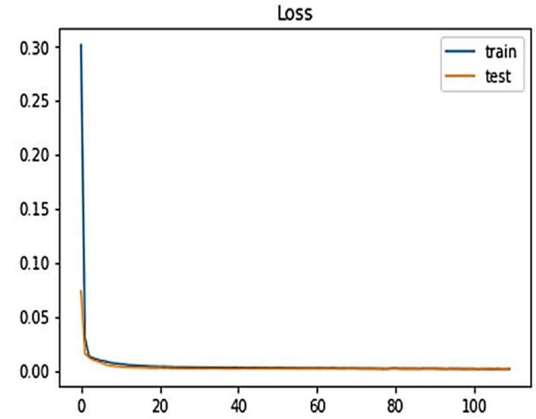

Figure 9. Loss during training and testing

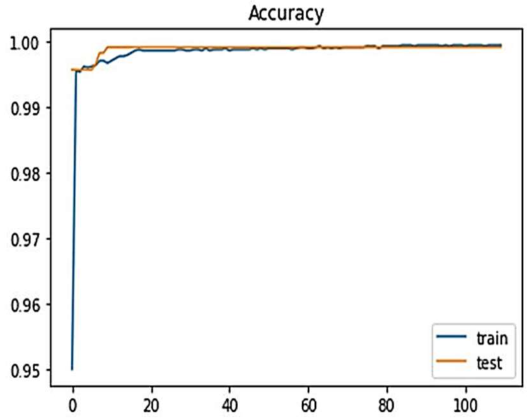

Figure 10. Accuracy for training and testing

### *7.3 Binary Classification DNN Results*

Another experiment the multi-layer DNN with single hidden layers are compared with proposed approach of data oversampling using SMOTE approach. The binary classification is implemented for predicting the Benign or Malicious classes over the Doorbell dataset. In this section the three binary classification models as DNN, RF, and decision tree (DT) are used for the classification using data oversampling to make balanced dataset. Results of the single layer binary classification are compared in terms of accuracy and recall with proposed methods.

The results of confusion matrixes for DNN, RF, and DT classification models are presented in the Figure 11 and 12 with SMOTE and ROU sampling approaches respectively. It is clear from the Figure 11 and 12 that SMOTE oversampling method outperforms over the ROU method. There is significant reduction in false positive and improvement in true positives using the DNN approach. Thus it is observed that DNN performs best over DT and RF approach.

In order to quantitatively justify the results accuracy and the F1 score for three classifiers are compared for two sampling methods as shown in the Table 3. It is clear that with 99.98% accuracy and 0.99981221 as F1 score the proposed DNN based method outperforms.

It can be observed from Table 3 that using the SMOTE oversampling approach might increase the accuracy of the classification .the maximum improvement is observed for the DT approach with 2.1174 % increment. The average accuracy for three classifiers with ROU oversampling is 99.24 % which is increased to average of 99.96 % with SMOTE approach. As an experiment the accuracy and F1 score are plotted for the 5 epoch with DNN classifiers and compared for ROU and SMOTE approaches as shown in the Figure 13 and Figure 14 respectively. It is also clear from that within 5 epochs accuracy of SMOTE method increases from 98.88 to 99.982 %.

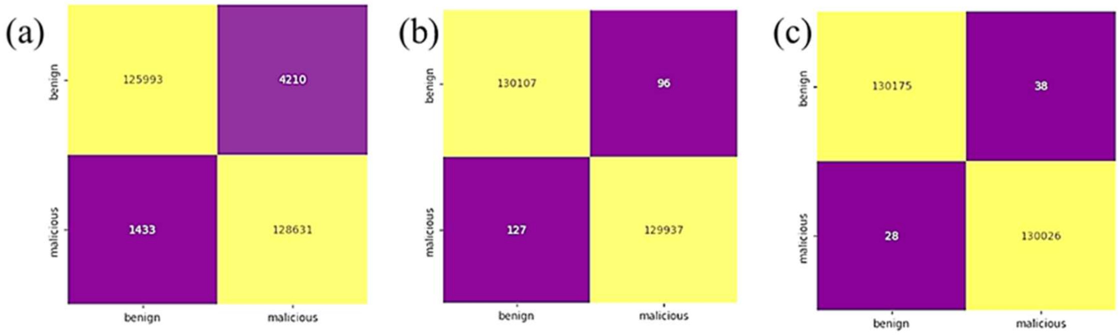

Figure 11. Results of binary class confusion matrixes.with ROU oversampling method
a) Column 1 for DT, b) Column 2 for RF c) Column 3 for proposed DNN classifier

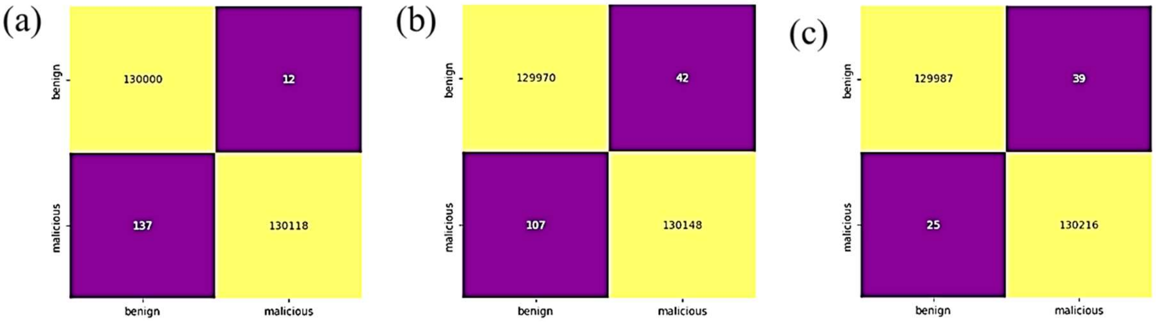

Figure 12. Results of binary class confusion matrixes with SMOTE oversampling method
a) Column 1 for DT, b) Column 2 for RF c) Column 3 for proposed DNN classifier

TABLE 3 QUANTITATIVE COMPARISON OF THE IOT VULNERABILITIES FOR BINARY CLASSIFICATIONS

| **Parameters** | **RNN Binary Classification** | **With ROU Oversampling** | | | **With proposed SMOTE Oversampling** | | |
|---|---|---|---|---|---|---|---|
| | | ***For DT*** | ***For RF*** | ***For DNN*** | ***For DT*** | ***For RF*** | ***For DNN*** |
| F1 Score | 0.9994732 | 0.97831619 | 0.99914263 | 0.99974626 | 0.9994928 | 0.9993675 | 0.99981221 |
| Accuracy | 99.94843 % | 97.8318 % | 99.9144 % | 99.9746 % | 99.94928% | 99.9366 % | 99.98117 % |

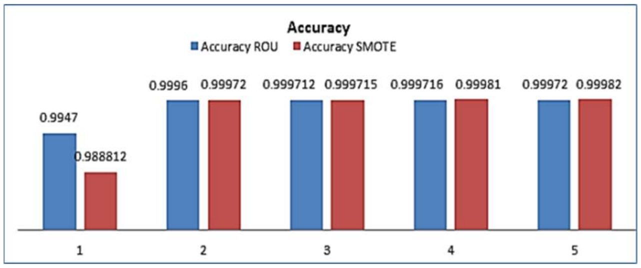

Figure 13. Accuracy for ROU and SMOTE based DNN for epochs

### *7.4 Multi Class Classification Results*

This dataset addresses the lack of public bot net datasets, especially for the IoT. The attack count plot is shown in the Figure 15 where benign means normal data and Gafgyt and Mirai are anomaly data. Here Gafgyt means Bashlite attack. Data was originally gathered from 9 commercial IoT devices authentically infected by Mirai and Bashlite. There are total 7062606 numbers of instances. For an experiment of multiclass testing total 141254 records were randomly chosen for the study extracted as 2% random sample from around 89 files. The comparison of the precision and recall values for DT and DNN classifiers are shown in the Table 4 and Table 5 respectively. The proposed DNN classifier using SMOTE with two hidden layer via 10 and 8 neuron activation function ReLU. The loss function used was sparse categorical cross entropy. The DNN proposed model offers nearly maximum accuracy.

The confusion matrix of the Multi class classification of attacks is shown in the Figure 16 for three different classifiers. There is significant reduction in false positive and improvement in true positives using the DNN approach.

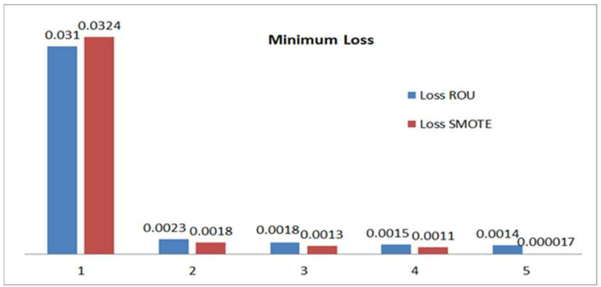

Figure 14. F1 Score for ROU and SMOTE based DNN for epochs

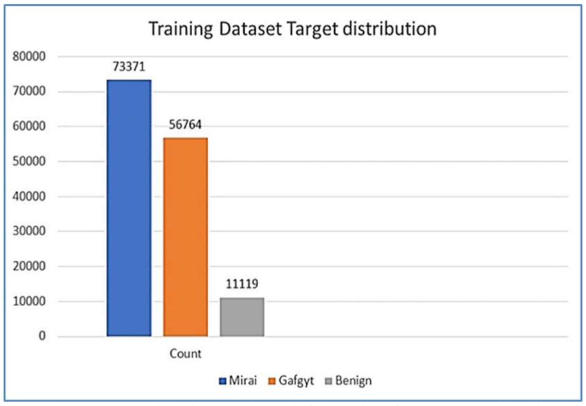

Figure 15. Multi class vulnerabilities data base target distribution considered for evaluation

The loss calculation for DNN epochs are shown in Figure 17. To make the result much clear accuracy and F1 score for multi class classification with proposed SMOTE oversampling approach is shown in the Table 6 it can be observed the even for multi class classification significant higher accuracy is achieved and RF classifier has slight edge over DNN. Thus it is still an open field of research to improve accuracy of DNN of this case in future.

TABLE 4. CLASSIFICATION RESULTS ANALYSIS FOR DT APPROACH

| | Precision | Recall | f1-score | Support |
|---|---|---|---|---|
| benign | 0.98 | 1 | 0.99 | 2234 |
| gafgyt | 0.98 | 0.98 | 0.98 | 11361 |
| Mirai | 0.98 | 0.98 | 0.98 | 14656 |
| | | | | |
| Accuracy | | | 0.98 | 28251 |
| macro avg | 0.98 | 0.98 | 0.98 | 28251 |
| Weighted avg | 0.98 | 0.98 | 0.98 | 28251 |

TABLE 5. CLASSIFICATION RESULTS ANALYSIS FOR PROPOSED DNN APPROACH

| | Precision | Recall | f1-score | Support |
|---|---|---|---|---|
| Benign | 1 | 1 | 1 | 2171 |
| Gafgyt | 1 | 1 | 1 | 11358 |
| Mirai | 1 | 1 | 1 | 14722 |
| | | | | |
| Accuracy | | | 1 | 28251 |
| macro avg | 1 | 1 | 1 | 28251 |
| Weighted avg | 1 | 1 | 1 | 28251 |

## 8. CONCLUSIONS AND FUTURE SCOPE

The studies helped to design and build a DL IoT security scheme for categorising and predicting vulnerabilities. The paper evaluated the binary classification effectiveness to multi-class classification models. The primary purpose of detection models is to determine whether a sample is benign or malicious. The classification algorithm' aim is to assign every sample to one of the following categories: Benign, Mirai, or Gafgyt and. Attack classification is implemented utilising a DNN-based learning model. The N-BaIoT multiple vulnerability dataset is utilised, which is unbalanced. In first part dataset dimensions is reduced using PCA analysis. And RNN is tested on reduced features using ReLU activation to avoid VGP. The 98.9 % accuracy is achieved for binary classification.

In second experiment paper proposed to use oversampling approaches for balancing the data. Performance of SMOTE and ROU methods are compared for three classifiers DT, RF and DNN for binary classification problem,

The concluded DNN-based technique clearly outperforms, as seen by its 99.98% accuracy and F1 score of 0.99981221. It is observed that employing the SMOTE oversampling strategy may boost the classification accuracy. The DT strategy shows the greatest improvement, with a 2.1174% increase. The average accuracy for three classifiers using ROU oversampling is 99.24%, which increases to an average of 99.96% with the SMOTE technique.

TABLE 6 QUANTITATIVE COMPARISON OF THE IOT VULNERABILITIES FOR MULTI CLASS CLASSIFICATIONS FOR DOORBELL DATA

| Parameters | Kim J, et al [14] KNN | Abbasi et al. [23] ANN | Hodo, E et al. [33] | With proposed SMOTE Oversampling | | |
|---|---|---|---|---|---|---|
| | | | | ***For DT*** | ***For RF*** | ***For DNN*** |
| F1 Score | 0.99 | 0.9335 | - | 0.9996514 | 0.9998121 | 0.9988256 |
| Accuracy | 99.01% | 93.58% | 99.4% | 99.9786 % | 99.9893 % | 99.8825 % |

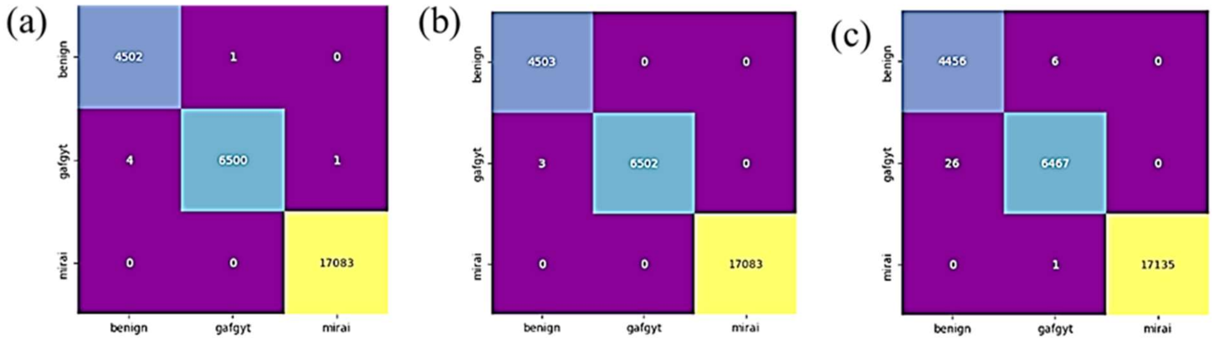

Figure 16. Results of Multi class confusion matrixes with SMOTE oversampling method
a) Column 1 for DT, b) Column 2 for RF c) Column 3 for proposed DNN classifier

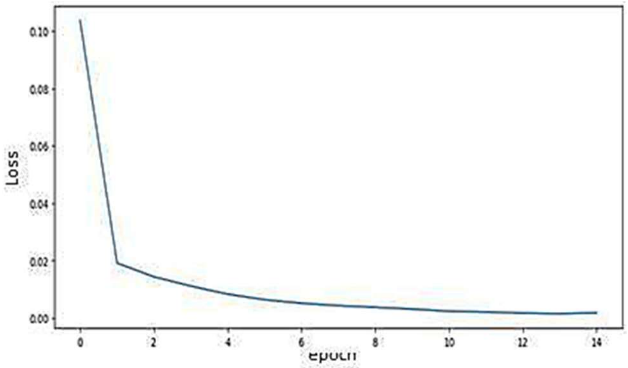

Figure 17. Loss calculation for the multi class IoT attack classification using DNN

There is significant reduction in false positive and improvement in true positives using the DNN approach. It is concluded that even for multi class classification significant higher accuracy of 98 to 99% is achieved and RF classifier has slight edge over DNN.

## Future scope

According to the data used, the risk assessment will enable future research to focus on the building's most dangerous areas and the deployment of appropriate security measures. It is still an open field of research to improve accuracy of DNN of this case in future. The multiple devices as class can be used for the classification problem in future studies. In addition multiple attacks may be incorporated for classification. The performance may be tested over other datasets too in future. iple devices as class can be used for the classification problem in future studies. In addition multiple attacks may be incorporated for classification. The performance may be tested over other datasets too in future.

## Declaration

It is to declare that this research is part of own research and is not copied from any other sources. There is no finding involved in this research and also no conflict of interest is there. The dataset is globally available and referred.

## REFERENCES

[1]. S.B. Hulayyil, S. Li and L. Xu, "Machine-learning-based vulnerability detection and classification in internet of things device security," *Electronics*, Vol. 12(18),2023 ,3927..

[2]. V. Adat and B. Gupta, "A DDoS attack mitigation framework for internet of things," in *Proceedings of the 2017 International Conference on Communication and Signal Processing (ICCSP), Chennai, India*, 2017, pp. 2036–2041.

[3]. R. Singhai and R. Sushil, "An investigation of various security and privacy issues in Internet of Things," *Materials Today: Proceedings*, Vol. 80, 2023, 3393-3397.

[4]. M. A. Ferrag and L. Maglaras, ''DeepCoin: A novel deep learning and blockchain-based energy exchange framework for smart grids,'' *IEEE Trans. Eng. Manage.*, vol. 67, 2020, pp. 1285–1297.

[5]. M. Ge, X. Fu, N. Syed, Z. Baig, G. Teo, and A. Robles-Kelly, ''Deep learning-based intrusion detection for IOT networks,'' in *Proc. IEEE 24th Pacific Rim Int. Symp. Dependable Comput. (PRDC)*, 2019, pp. 256–25609.

[6]. M. Ge, N. F. Syed, X. Fu, Z. Baig, and A. Robles-Kellyd, ''Towards a deep learning-driven intrusion detection approach for Internet of Things,'' *Computer Network*, Vol. 186, 2021, 107784.

[7]. S. Aldhaheri, D. Alghazzawi, L. Cheng, B. Alzahrani, and A. Al-Barakati, ''DeepDCA: Novel network-based detection of IOT attacks using artificial immune system,'' *Appl. Sci.*, vol. 10, 2020, pp. 1909.

[8]. Q. Wang, W. Zhao, and J. Ren, ''Intrusion detection algorithm based on image enhanced convolutional neural network,'' *J. Intell. Fuzzy Syst.*, vol. 41, 2021, pp. 2183–2194.

[9]. A. M. Lonzetta, P. Cope, J. Campbell, B.J. Mohd and T. Hayajneh, "Security vulnerabilities in bluetooth technology as used in IoT." *Journal of Sensor and Actuator Networks*. Vol. 7, 2018, p. 28.

[10]. S. Rajagopal, P. P. Kundapur and K.S. Hareesha, "A stacking ensemble for network intrusion detection using heterogeneous datasets," *Security and Communication Networks*, 2020, pp. 1-9.

[11]. I. Cvitić and M. Vujić, "Classification of security risks in the IoT environment", *26th DAAAM Int. Symposium on Intelligent Manufacturing and Automation*, 2016, pp.0731-0740.

[12]. V. A. Knights and Z. Gacovski, "Methods for detection and prevention of vulnerabilities in the IoT (Internet of Things) Systems' Web of science IntecOpen Book chapter 2020.

[13]. U. Shafiq, M. K. Shahzad, M. Anwar, Q. Shaheen, M. Shiraz,and A. Gani, "Transfer learning autoencoder neural networks for anomaly detection of DDoS generating IoT devices," *Security and Communication Networks,* 2022, 8221351.

[14]. R. Singhai and R. Sushil, "Towards identification of security threats and vulnerabilities in internet of things," in *2022 2nd*

*International Conference on Innovative Practices in Technology and Management (ICIPTM)*, Vol. 2, 2022, pp. 58-62.

[15]. A. Marzano, D. Alexander, O.Fonseca, E. Fazzion, C. Hoepers, K. Steding-Jessen, M.H Chaves, I. Cunha, D. Guedes and W. Meira, "The evolution of bashlite and mirai iot botnets," in *2018 IEEE Symposium on Computers and Communications (ISCC),* 2018,pp. 00813-00818).

[16]. J. Kim, M. Shim, S. Hong, Y. Shin and E. Choi, "Intelligent detection of IOT botnets using machine learning and deep learning," *Applied Sciences*, Vol. 10, 2020, 7009.

[17]. S. Choudhary and N. Kesswani, "A Hybrid Classification Approach for Intrusion Detection in IOT Network," *Journal of Scientific & Industrial Research*, Vol. 80, 2021, pp. 809-816.

[18]. R. Singhai, and R. Sushil, "Real Time Security Enhancement for IOT Enabled Intelligent Network," *International Journal of Intelligent Systems and Applications in Engineering*, Vol. 11(6s), 2023, pp.775-781.

[19]. H. Alasmary, A.Khormali, A. Anwar, J. Park, J. Choi, A. Abusnaina, A. Awad, D. Nyang and A. Mohaisen, "Analyzing and detecting emerging Internet of Things malware: A graph-based approach," *IEEE Internet of Things Journal*, Vol. *6*(5), 2019 pp.8977-8988.

[20]. C.S. Htwe, Y.M. Thantm and M.S. Thwin, "Botnets attack detection using machine learning approach for iot environment," in *Journal of Physics: Conference Series,* Vol. 1646, 2020, p. 012101.

[21]. N.V. Chawla, K.W. Bowyer, L.O. Hall and W.P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *J. Artif. Intell. Res,* 16, 2002, pp.321–357.

[22]. J.H. Joloudari, A. Marefat, M.A. Nematollahi, S.S. Oyelere and S. Hussain, "Effective class-imbalance learning based on SMOTE and convolutional neural networks," *Appl. Sci.,* Vol. 13, 2002, 4006.

[23]. Q. Chen, Z.-L. Zhang, W.-P. Huang, J. Wu and X.-G. Luo, "PF-SMOTE: A novel parameter-free SMOTE for imbalanced datasets," *Neurocomputing,* Vol. 498, 2022, pp. 75–88.

[24]. X. Yi, Y. Xu, Q. Hu, S. Krishnamoorthy, W. Li and Z. Tang, "ASN-SMOTE: a synthetic minority oversampling method with adaptive qualified synthesizer selection," *Complex Intell. Syst.*, Vol. 8, 2022, pp. 2247–2272.

[25]. Y.G. Damtew and H. Chen, "SMMO-CoFS: Synthetic multi-minority oversampling with collaborative feature selection for network intrusion detection system," *Int J Comput Intell Syst,* Vol. 16, 2023, p.12.

[26]. F. Abbasi, M. Naderan and S.E. Alavi, "Intrusion detection in IoT with logistic regression and artificial neural network: further investigations on N-BaIoT dataset devices," *Journal of Computing and Security*, Vol. 8, 2021, p.27-42.

[27]. C.D. McDermott, F. Majdani and A.V. Petrovski, "Botnet detection in the internet of things using deep learning approaches," *2018 International Joint Conference on Neural Networks (IJCNN)*, 2018, 1-8.

[28]. Koggle Database link https://www.kaggle.com/datasets/mkashifn/nbaIOT-dataset/code .

[29]. Doorbell Dataset link https://research.google.com/audioset/dataset/doorbell.html

[30]. S. Bagui, and K. Li, "Resampling imbalanced data for network intrusion detection datasets," *J Big Data,* Vol. 8, 2021, pp.1-41.

[31]. Y. Meidan, M. Bohadana, Y. Mathov, Y. Mirsky, D. Breitenbacher, A. Shabtai and Y. Elovici, "N-BaIOT: network-based detection of IOT botnet attacks using deep autoencoders," *IEEE Pervas Comput*., Vol. 13,2018,pp. 1–8.

[32]. A. Luque, A. Carrasco, A. Martin and A. Heras de las, "The impact of class imbalance in classification performance metrics based on the binary confusion matrics," *Pattern Recogn.* Vol. 19, 2019, 216–31.

[33]. C. Obimbo, D. Mand and S. Singh, "Oversampling techniques in machine learning detection of credit card fraud," *Journal of Internet Technology and Secured Transactions (JITST)*, Vol. 9, 2021, pp. 741-746.

[34]. J.L. Leevy, J. Hancock, J., T.M. Khoshgoftaar, and J.M. Peterson, "IoT information theft prediction using ensemble feature selection," *Journal of Big Data*, Vol. *9*(1), 2022, pp.1-48.

[35]. S. I. Popoola, B. Adebisi, R. Ande, M. Hammoudeh, K. Anoh and A.A. Atayero, "SMOTE-DRNN: A deep learning algorithm for botnet detection in the internet-of-things networks," *Sensors*, Vol. *21*,2021, 2985.

[36]. R. Biswas and S. Roy, ''Botnet traffic identification using neural networks,'' *Multimedia Tools Appl.*, Vol. 2021, pp. 24147–24171.

[37]. H. Wasswa, T. Lynar and H. Abbass, "Enhancing IoT-Botnet detection using variational auto-encoder and cost-sensitive learning: a deep learning approach for imbalanced datasets," *2023 IEEE Region 10 Symposium (TENSYMP)*, Canberra, Australia, 2023, pp. 1-6.

[38]. C. -C. Chen, Y. -R. Chen, W. -C. Lu, S. -C. Tsai and M. -C. Yang, "Detecting amplification attacks with Software Defined Networking," *2017 IEEE Conference on Dependable and Secure Computing*, Taipei, Taiwan, 2017, pp. 195-201.

[39]. Joshi, C., Ranjan, R. K., & Bharti, V, "ANN based Multi-Class classification of P2P Botnet" 2021 *International Journal of Computing and Digital System*, 2021, pp. 1319-1325.

**Richa Singhai** Richa singhai is a PhD Research scholar at DIT University, Dehradun, India. She has completed his M.tech (CSE) from RGPV, Bhopal (M.P). Her research area is Machine Learning, IoT, Block chain, Cyber Security and Deep learning.

**Dr.Rama Sushil** Dr.Rama Sushil currently working as an Professor & Academic Head at DIT University Dehradun, india . She has completed her Ph.D. degree from IIT Roorkee, India. Her research interests include Data Warehousing, Machine Learning, Deep Learning and Cyber Security.