ENSEMBLE TECHNIQUE BASED

MODEL FOR INTRUSION DETECTION IN INTERNET OF THINGS ENVIRONMENT

BY

EDWARD, Elizabeth Ozioma
MTech/SICT/2017/7161

DEPARTMENT OF CYBER SECURITY SCIENCE

FEDERAL UNIVERSITY OF TECHNOLOGY, MINNA

SEPTEMBER, 2021

ENSEMBLE TECHNIQUE BASED

MODEL FOR INTRUSION DETECTION IN INTERNET OF THINGS ENVIRONMENT

BY

EDWARD, Elizabeth Ozioma
MTech/SICT/2017/7161

A THESIS SUBMITTED TO THE POSTGRADUATE SCHOOL FEDERAL UNIVERSITY OF TECHNOLOGY, MINNA, NIGERIA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF MASTER OF TECHNOLOGY IN CYBER SECURITY SCIENCE

SEPTEMBER, 2021

ABSTRACT

The rapid growth of interconnected devices gives rise to possible exploits and compromise of devices in Internet of Things (IoTs) environment, which will require an established fact for such compromise to be checkmate. This research focuses on detection of threat attack detection in IoT environment based on ensemble technique. In the proposed ensemble model, an outlier analysis was performed in order to optimize the features for enhanced model performance, while Support Vector Machine and Feed Forward Neural Network serves as the base learners which were combined to form the proposed ensemble model, this was employed in order to enhance performance strength of training model in detection of intrusion threat in IoT environment. The advantage of the proposed model is its ability to generate an enhanced performance evaluation output; as the result proves an excellent performance for intrusion detection in IoTs environment. The obtained accuracy, precision, F-score, and recall of the proposed ensemble model are 99.96 for each respectively outperforming the existing technique with a record of 92.42 and 95.43 for accuracy and detection rate respectively, this is a clear distinction of the superiority of ensemble (Feed Forward Neural Network and Support Vector Machine) model over traditional model, of which the proposed ensemble model can serve as technique in forestall intrusion threat experienced in IoTs environment.

TABLE OF CONTENTS

CHAPTER TWO

CHAPTER THREE

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ANN	Artificial Neural Network
CART	Classification and Regression Trees
DCNN	Distributed Convolutional Neural Network
DDoS	Distributed Denial of Service
DoS	Denial of Service
FAR	False Alarm Rate
FFNN	Feed Forward Neural Network
GA	Genetic Algorithm
HMMs	Hidden Marcov Models
IDS	Intrusion Detection Systems
IoTs	Internet of things
KNN	K-Nearest Neighbor
LDA	Linear Discriminant Analysis
LSTM	Long Short Term Memory
LVQ	Learning Vector Quantization
MCC	Matthew Correlation Coefficient
MLP	Multi-Layer Perception

NSL KDD	Network Security Laboratory Knowledge Discovery in Databases
R2L	Root-to-Local
RBF	Radial Basis Function
SMO	Spider and Network Optimization
SVM	Support Vector Machine
U2R	User-to-Root
WSN	Wireless Sensor Network

CHAPTER ONE

1.0 INTRODUCTION

1.1 Background to the study

The emergence of Internet of Things (IoTs) serves as a developmental change in computing environment experienced in recent years, comes with the potential about the transformational experience every ramification of human existence (Vega-barbas *et al*., 2021)**.**

Also noted is exponential growth in the network of interconnected devices of internet enabled technologies, supply change environment, smart retail, health care services, smart dwelling homes, interconnected vehicles as well as a varied range of applications are some of the environment that IoTs devices are deployed for its services (Rezvy *et al.,* 2019).

The expansion rate of development of IoTs enabled devices comes with its challenges in terms of security as well as privacy compromise, the trajectory growth of connected technologies introduces new security flaws and provides the avenues for threat exploits that can lead to compromise of devices associated in the connectivity flows and of more challenge in the field of IoTs devices is that security is given little regard with adequate checkmating security technology in place (Bambang and Riri., 2020).

Studies revealed that great percentage of current IoTs technologies in the market place have high rate security setbacks Santoyo-gonz (2019), a compromised IoTs technology can be used as a platform of attack vectors if security control is given a blame view, cyber-attack such as intrusion exploit are part of the new trend widely experienced channeled towards IoTs devices (Rajendran *et al*., 2019).

IoTs is exponentially growing, posing as danger zone and challenges for investigators of variety attack ranging from cyberattacks as well as physical cyberattacks. A lot of IoTs

devices autonomously at times create data from embedded sensors in replication to human activities in regard to responses like door opening and motion spotting. The functionality of IoTs devices can be altered from almost every location, in as much as smart devices, smart houses, as well as other IoTs environments are connected. The generated data by IoTs serves as excellent digital feature as well as the traces of activities of IoTs devices captured through investigations can serve as digital evidence (Servida and Casey, 2019).

In the past years, the different number of IoTs devices has experienced an exponential growth with about 50 billion of devices to be connected to the internet as at year 2020. A very large number of new smart devices and products such as smart houses, smart watches, smart TVs, smart appliances are spreading fast in the society which is becoming more pervasive in our everyday life (Meneghello *et al.*, 2019).

The use of IoTs concept in different economic sectors has become an important factor for business enhancement, which is believed by 92% of companies that the IoTs concept will be of use in their various businesses as of 2020 (Cvitić *et al.*, 2021). However, the companies have put into consideration that security, privacy, costs, and regulatory issues pose a great threat in the implementation and application of the IoTs concept.

The exponential growth of the internet is the reason for the emergence of the internet of things (IoTs). In recent time, IoTs paradigm was used to build smart environments, such as smart cities and smart homes with another technology domains and similar services (Kumar *et al.*, 2021).

Ensemble based machine learning model have experienced a surge research efforts due to the ability of the model to handle huge data, ensemble model combines multiple models in order to classify new unseen instances with excellent performance record in classification as well as regression problems Ren *et al.* (2016), the ensemble model function through the

combination of different learning models to enhance performance, ensemble learning were carried out in the 1990's to establish that weak learning classifier could be transformed into strong learning classifier (Dietterich, 2000) .

Ensemble model was employed in this research to enhance performance of intrusion detection in Internet of Things environment, which leads to better techniques in addressing the threat challenges that exist in IoTs environment such as distributed denial of service, privilege exploit and stealth data espionage.

1.2 Statement of the research Problem

Threat emanating from network intrusion, serves as an overwhelming challenge to IoTs environment due to the existing traditional technique's inability to manage the dynamic and sophisticated threat against IoTs devices (Diro and Chilamkurti, 2017)**.** Most of the traditional techniques such as Support Vector Machine and Artificial Neural Network widely adopted in the detection of intrusion threat in IoTs environment suffers a setback characterized by low accuracy as well as high false positive rate (Almiani *et al*., 2020) as a result of the challenge of impending strength to manage the traffic volume of data. Hence, the need for a model that can render performance enhancement in terms of accuracy as well as false positive rate in IoTs environment is required.

1.3 Aim and Objectives to the Study

The aim of the study is to develop an ensemble technique based model for intrusion detection in IoTs environment.

The objectives of this research are to:

i. design architecture for ensemble technique based model for intrusion detection in IoTs environment,

ii. develop an intrusion detection model in IoTs environment based on Ensemble technique (Support Vector Machine and Feed Forward Neural Network),

iii. evaluate the performance of the designed model in (ii) through relevant performance metrics.

1.4 Significance to the Study

The design of an architecture provides a crystal and ease integration of learning models based on intrusion detection, as it relates to IoTs environment. The employment of ensemble model in this research provides an intelligent and enhanced detection model for detection of intrusion threats in IoTs environment, thereby revealing the outperforming strength of ensemble model against existing traditional models, overcoming the drawbacks of performance deficiency.

Finally, the designed model performance evaluation with relevant metrics as reviewed, reflects a proper gauging with existing approaches as recorded in relevant studies.

1.5 Scope to the Study

This research is focused on the design of the proposed intrusion detection model in IoTs environment based on ensemble technique (Support Vector Machine and Feed Forward Neural Network), and evaluating same using standard dataset (NSL-KDD) from recognized dataset repository for experimentation.

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 Internet of Things Preamble

The idea of IoTs was framed by a community of Radio Frequency Identification (RFID) in the year 1999, in recent time, it has become very relevant to the practical world majorly because of the growth of mobile devices embedded and ubiquitous communication, cloud computing and better analysis (Yakubu *et al*., 2018).

The common definition of IoTs is a network of physical objects. The network is not just a network of computers, but it has included a network of devices of all types and sizes, home appliance, medical instruments, vehicles, smart phones, camera's, toys, home appliances, animals, people, buildings and industrial systems all linked , all communicating and all sharing information on an agreed protocol in order of achieving a smart re-organization , tracing ,positioning, save and control and in most cases personal real time online monitoring, online upgrade, process control and administration (Salazar, 2019).

2.1.1 Enabling technology in internet of technology

The global infrastructure for information society can be linked to IoTs, this can enable advanced services by interconnecting devices based on existing and evolving inter-operable information and communication technology. The communication of IoTs can be extended through internet to all devices that surrounds us. The internet of things (IoTs) is more than just a machine- machine communication, wireless sensor networks, sensor networks, WI-FI, GPS, RFID, GSM, GPRS, 2G/3G/4G/5G, microcontroller and microprocessor. These all can

be considered as enabling technologies which make IoTs applications possible. (Yakubu *et al.*, 2018).

IoTs enabling technologies can be grouped into three categories: these includes:

i. Technology that enable things to acquire contextual information.

ii. Technologies that enable things to process contextual information.

iii. Technologies to improve security and privacy.

Yakubu *et al.* (2018) the first two can be understood as functional building blocks that are needed in building intelligence into things, these are the features that segregates the IoTs from usual internet. The last category is not a functional building blocks. It is rather a de-facto requirement without which the penetration and acceptance of IoTs would be severely low.

IoTs is not a single technology but rather a mixture of multiple hardware and software technology. This IoTs provides solution for information technology based on their integration, which refers to the hardware and software used to store, retrieve, and process data and communication technology that includes electronic systems used for communication between individuals or groups (Yakubu *et al.*, 2018).

There is heterogeneous mixture of communication technology that is needed to the adapted in order to address the necessity of IoTs applications which includes energy efficiency, speed, security and reliability. It is possible that in this context, the level of diversity which can be scaled to a number of manageable connectivity technologies which can be used to address the IoTs applications. Some common examples in this category, include wired and wireless technologies, Bluetooth, wi-fi, Gsm, ZigBee, and GPRS (Yakubu *et al.*, 2018).

2.1.2 Characteristics of internet of things

Interconnectivity: Anything can be interconnected globally with information communication infrastructure in regard to IoTs.

Things related services: the capability of IoTs can provide things related services within the constraints of things which can include privacy protection and semantic consistency between the physical things and their associated virtual things. To provide things related services within the constraint of things, both the technologies in physical world and information world will have to change.

Heterogeneity: based on different hardware platforms and networks, the devices in IoTs are heterogeneous. They can easily communicate with other devices or service platforms through different networks.

Dynamic changes: Device states change dynamically, such as sleeping and waking up, being connected and, or disconnected, as well as the context of devices, such as location and speed. Furthermore, the number of devices will change on a regular basis.

Enormous scale: the number of devices that must be handled and communicate with one another would be at least an order of magnitude greater than the number of devices currently connected to the internet.

The management of the data produced, as well as its interpretation for application purposes, will be even more crucial. This has to do with data semantics as well as effective data handling.

Safety: we must not forget about protection while we reap the reward of the IoTs. We must design for protection as both designers and recipients of IoTs. This includes the protection of

our personal information as well as our physical well-being. Securing endpoints, networks, and the data that moves through them all necessitates the creation of a scalable security model.

Connectivity: network accessibility and compatibility are made possible by connectivity. Accessibility is the ability to connect to a network, while compatibility is the ability to consume and generate data in a common way.

2.2 Outlier Analysis Algorithm

The outlier algorithms determine the outlier that exist in a spatial feature, through the evaluation of the differences between the attributes and the summary functions of deviation, thereafter normalizes same with their mean and standard threshold, which then denoted as an outlier and its attribute is taken to be the summary function of its neighbors, this procedure is repeated until no outlier are found Alharbi *et al.* (2012) Table 2.1 presents the outlier analysis algorithm.

Table 2.1 **Outlier Analysis Algorithm** (Alharbi *et al.*, 2012)

Outlier Analysis **Algorithm**:
1: θ be a threshold
2: let n be $\|N\|$
3: outlier = true
4: $\sigma = 0; \mu = 0$
5: for i = 1 $\rightarrow n$ **do**
6: let N($\emptyset_i$) be the set of files in $D(\emptyset_i)$ and $\|N(\emptyset_i)\|$ its cardinality
7: set the attribute function $f(\emptyset_i)$ to be $\|(\emptyset_i)\|$
8: compute the summary function $g(\emptyset_i) = \frac{1}{\|N(\emptyset_i)\|}\sum_{\emptyset \in N(\emptyset_i)} f(\emptyset)$
9: compute the summary function $h(\emptyset_i) = f(\emptyset_i) - g(\emptyset_i)$
10: $\mu = \mu + h(\emptyset_i)$
11: $\sigma = \sigma + h(\emptyset_i)^2$
12: end for
13: $\mu = \sqrt{\frac{\sigma}{n} - \frac{\mu^2}{n}}$ % the standard deviation
14: while outlier = = true **do**
15: outlier = false
16: $\emptyset_q = \arg max_{\emptyset} \left\|\frac{h(\emptyset)-\mu}{\sigma}\right\|$
17: if $\left\|\frac{h(\emptyset)-\mu}{\sigma}\right\| > \theta$ **then**
18: mark $\emptyset_q$ as a outlier % $\emptyset_q$ is a outlier
19: $f(\emptyset_q) = g(\emptyset_q)$
20: update $g(\emptyset)$ and $h(\emptyset)$ for every $\emptyset$ in $N(\emptyset_q)$
21: update μ and σ
22: outlier = =true
23: end if
24: end while

2.3 Support Vector Machine (SVM) Algorithm

SVM is known to be widely used state of the art model, mainly for the problem of classification, thus used also for regression, SVM is based on the principle of margin calculation. SVM draws a margin between the class in a fashion that the distance between the margin and the classes is maximize and thereby minimizing the classification error that may be associated (Boswell, 2002). SVM is noted to have a good generalization performance

with regards to other machine learning models and also efficient in binary classification problems, Lin *et al*. (2013), Table 2.2 present SVM algorithm.

Table 2.2: Support Vector Machine Algorithm (Vishwanathan and Murty, 2002)

SVM **Algorithm:**

candidateSV = {closest pair from opposite classes}
While violation point exists do
Find a violator
candidateSV = candidateSV ∪ violator
If any $\propto_p < 0$ due to addition of c to S then
candidateSV = candidateSV/p
repeat till all such points are pruned
end If
end While

2.4 Related Literature

IoTs face severe threat challenges as well as vulnerability that are exposed to attacks (Smys and Wang, 2020) opined an intrusion detection system with the capacity of detecting attacks in the case of network security breach, in IoTs, the technique employed was a hybrid convolutional neural network method which is model toward IoTs applications, the following results was obtained for precision, recall, f-score, and accuracy, 1.00, 1.00, 0.99 and 98.6 respectively; however, the scores obtained in precision and recall is an indication that proposed research model is perfect in this field which is not quit fixable, meanwhile, improvement can be made for enhanced performance for detection of attacks in IoTs environment.

A Generic Algorithm based model for the detection of Intrusion Detection was proposed, an expected detection rate of 0.97 was targeted; however, a drawback of detection of new attacks is one of the challenges faced by the proposed model in the research by Paliwal and Gupta

(2012), meanwhile, improvement can be done in terms of detection rate for enhancement of threat attacks as it relate to intrusion threats.

Revealed that IoTs devices are prone to cyber exploit due to experience challenge in terms of storage, computational as well as communication strength as a result of the numerous challenge of cyber-attack in IoTs autonomies deep-learning based detection and classification system based on convolution neural networks for IoTs intrusion detection was developed, the evaluation of the developed system scored the following 99.3, and 98.2 respectively for accuracy as it relates to two class and five class classification. A precision rate of 99.04 and 98.27 both for two class and five class classification, an error rate of 00.7% and 01.8% respectively for two class and five class classification, while a recall of 99.33 98.23 for two class and five class classification, f-score rate of 99.18 and 98.22 for two class and five class classification respectively, FAR of 01.28 and 1.73% respectively for two class and five class classification, however, an enhanced performance can be recorded if candidate model is well trained for better classification of IoTs intrusion detection (Rezvy *et al.*, 2019).

In order to determine the performance strength of machine learning of some selected machine learning model, a comparative analysis was carried out by Mol and Mary (2021) the research recorded that AdaBoost model outperformed other selected model in IoTs intrusion detection with an accuracy score rate of 99.8% against Random forest, multi label classifier and deep neural network, which score98.7, 99.6, and 99.2 % respectively, however, other true state efficiency as explored by other relevant performance metric can be employed to determine the true state efficiency as explore by other research work in same field of learning and analysis.

A numerical analysis based method was employed to determine the performance capability of IoTs network intrusion detection, the following machine learning model was used for analysis. Decision tree, XGBoost, Bagging Tree, Random Forest, Bayes Net, Support Vector Machine, Naïve Bayes and Ada Boost, while an ensemble based model of XGBoost recorded an optimal performance of 0.970, 0.970, 09.68 and 0.968 for metric accuracy, precision, recall and f-score respectively outperforming other selected models used in this research, Bayes Net recorded the low false positive rate of 0.006, while XGBoost still maintained an enhanced performance of 0.970, 0.905 and 0. 996 of true positive rate, mcc, and accuracy respectively against the selected models in the research. However, the exploration of other machine learning model and training of the model parameter can serve as a promising field to employ in this research, (Liu *et al.*, 2020).

In order to tackle the security threat in IoTs environment, a new model deep learning based for intrusion detection was opined, combining the features of spider and monkey optimization (SMO) and stacked-deep Polynomial Network to record better accuracy detection rate, the proposed model in this research claimed to have better performance in accuracy, precision, recall and f-score with the following score 99.02, 99.38, 98.29 and 98.83 respectively, however, exploring other classification model can serve as a promising aspect to improve on the performance evaluation in IoTs threat environment (Khare *et al*., 2020).

A model that is targeted at threat in IoTs environment was proposed by Thamilarasu and Chawla (2019)**,** the model known as an intelligent intrusion detection system is based on a deep learning model with the capability of detecting malicious threat flows in IoTs networks, it was further claimed that the proposed model performed effectively in detection intrusion in IoTs environment with the following average performance measure score of 95% and 97%

for precision and recall rate respectively, however, the research was restricted to some selected threat in IoTs environment such as black hole, DDOS, sinkhole and wormhole attacks.

Recognizing the challenges in threat detection in IoTs technology Nazarpour *et al.* (2020) particle swarm optimization algorithm was used to enhance neural network model as well as other TLBO to achieve an optimal performance for IoTs intrusion attack detection, the following score rate of 90%, 91%, and 89% accuracy was achieved for Dos and R2L, U2R and PROB respectively, however, improvement can be achieved if other models are developed for analysis and relevant performance measure such as precision, recall, and f-measure are explored.

An intrusion detection system based K-nearest neighbor was proposed in the research performed by Wazirali (2020), through the training of the hyper parameter of the candidate model for efficient performance evaluation, as well as intrusion attack detection in cyber space, it was ascertained that the proposed model recorded an optimal performance of 0.9849, 0.9871, 0.9815 and 0.9843 respectively for accuracy, precision, recall and f1-score, nevertheless, real time attack detection was pointed out as a setback challenge in this research.

Davahli *et al.* (2020) proposed a light weight model Support Vector Machine (SVM) based intrusion detection system though an incorporated Genetic Algorithm (GA) and Green Wolf Optimizer (GWO) for dimensioning and reduction, it was further ascertained that the proposed model recorded 99.10, 99.32, 96.03, 967.64 and 0.69 respectively for accuracy, detection rate, precision, f1-score, and FPR respectively, however, focus was based on wireless IoTs network. An expanded data sample can be more promising for analysis.

Alsamiri and Alsubhi, (2019) performed a comparative evaluation of some selected model performance in order to determine the model with efficient performance for intrusion attack in IoTs environment, the selected model which are: NB, QDA, RF, ID3, Ada Boost, MLP and KNN recorded these scores, 0.78, 0.88, 0.98, 0.99, 1.0, 0.84 and 0.99 respectively for accuracy, while the precision scores are 0.84, 0.89, 0.99, 1.0, 0.88, and 0.99 respectively, the scores for recall are 0.78, 0.88, 0.98, 0.99, 1.0, 0.88, 0.84 and 0.99 respectively and f-measure recorded 0.75, 0.87, 0.98, 0.99, 1.0, 0.83, and 0.99 respectively, while KNN proved to outperform other selected models, NB takes a least time of 2052.1801secs, while KNN has the highest time record of 2052.180secs, however, the data sample used is not quite enough as it affects the performance of model learning, 84 data samples was employed for this research.

Alshammari and Zohdy (2019) a model named tiered system of hidden markov models (HMMS) was proposed to address the challenge of identification of attacks and detection in Internet of Things, the following performance metrics and scores was deployed using precision and accuracy with score rate of 85% and 96% respectively, however, improvement can be achieved if other relevant model are deployed as in the field of intrusion detection in IoTs.

A model based on Genetic Algorithm and deep belief network for IoTs intrusion detection was proposed by Zhang, *et al.* (2019)**,** it was claimed that the proposed model achieved the following optimal performance 99.45%, 97.78%, 99.37%, 98.68% respectfully for Dos R2l, UTR respectively and also for accuracy rate, however, more dataset sample will be needed to ascertain the effectiveness of the proposed model, as it was reported that small sample of some of the attack types face higher performance.

Samy *et al.*, (2020) proposed a framework that is fog-based attack detection based on deep learning targeted at IoTs intrusion detection, 99.34%, 99.18%, 99.59%, 99.39%, 0.1% and 99.59% performance score was recorded for accuracy, precision, recall, f1-measure, FAR and detection rate respectfully, however, the deployment of other model can serve as a promising exploration for enhanced performance measure output.

Ioannou and Vassiliou (2019) proposed a detection model for IoTs intrusion, the detection model is based on SVM, the performance of the proposed model was ascertained to have recorded up to 100% accuracy for black hole and sinkhole attacks present in IoTs attack and 80% accuracy for a different variant of attacks in IoTs, however, an improved performance can be recorded in a large sample of data used as this research used less than a thousand record of sample.

A model deep transfer learning technique for detection of IoTs attack was opined, the proposed technique which is based on auto encoders was claimed to have achieved an enhanced performance in terms of accuracy in detecting IoTs attacks, while 0.764% was the optimal accuracy achieved by the model as recorded, however, it was noted that more time is required to train the model, this serve as a draw back in this research, though more time to train the proposed model was not revealed (Vu *et al.*, 2020).

Alshahrani, (2021)proposed a model known as a collaborative intruder detection system for internet of things, the proposed model encompasses four layer based which are the IoTs layers, network layers fog layer and cloud layer, with the capability of tracking and analyzing the generated IoTs network traffic, it was ascertained that the proposed system achieved an accuracy of 98.35% with an error rate of 3.61%, while f1-score and precision scores are 97.39% and 96.39% respectively, however, it was reported that the sample dataset has less

feature, which will impede proper pattern learning by candidate model for efficient classification.

An optimization-based early classification model was opined for malware detection through the research conducted by Sharma (2020)**,** the model focuses on using early stopping rules (ESR) based Gaussian process classifier and particle swarm optimization for training candidate model as well as series of probabilistic classifiers in other to achieve optimal performance recorded as it relates to malware detection, it was further claimed that the proposed model delivered a decent balanced accuracy as well as early classification of malware, the following optimal performance record was attained by this research, 0.8380%, 0.8230%, 0.8611% and 0.8416% for accuracy, precision, recall, and f1-score respectively with a tradeoff earliness performance of malware detection model.

A deep learning based model multi-layer perception was proposed by Chaabouni (2020) in order to address the challenge of intrusion detection, an accuracy of 91.41% score was recorded for DoS attack class, however, a robust dataset is required for better performance record as the samples of the dataset used in this research have some void and irrelevant sample that affects the candidate model performance (Krishna, 2020)**.**

In the quest for intrusion detection Fatayer and Azara (2019) carried out a research based on Artificial Neural Network (ANN), learning Vectors Quantization (LVQ) versions, Radial Basis Function (RBF) and MLP, the following accuracy was recorded 97.44%, 99.44% and 99.86% respectively with MLP model outperforming other models, however, time complexity was recorded.

Mliki *et al.* (2020) address the issue of performance enhancement of machine learning model for intrusion detection in IoTs environment, the enhancement of selected model was achieved through a multi-level tweak which is likewise known as mixed models, SVM, KNN and K-mean++ was enhanced to SVM^2, $K-mean^2$and SVM+K, thereafter, it was ascertain that the enhanced model outperforms the traditional existing model in terms of IoTs intrusion detection as it relates to accuracy, detection rate false alarm rate with the following scores, 97.9%, 0.963% and 0.006% with training time of 3.288secs, while K_mean^2 achieved 97.21%, 0.008% respectively for accuracy, detection rate and false alarm rate with training time of 70.32secs, also, SVM+K recorded 93.9%, 0.9686% and 0.087% for accuracy, detection rate and false alarm rate respectively with training time of 1.532secs, however, it is a clear indication that the further model enhancement will yield optimal performance in detection of IoTs intrusion threats.

Proposed in the research experiment by Singh *et al.*, (2020) is a method known as a scheme based ensemble of discriminant classifier with the strength of detecting intrusion threat on a network, it was ascertained that the proposed scheme a superior performance record of 98.9% accuracy in detecting attack types as it relates to intrusion threats, however, a more justifiable performance evaluation can be experienced if other relevant metric are considered and other learning models explored.

A deep model approach opined for detection of IoTs intrusion threat, it was claimed that the proposed model recorded an optimal performance for IoTs threat in terms of accuracy, detection rate and false alarm rate of 99.20%, 99.27%, and 0.85% respectively for a 2-class detection and 98.27%, 96.5% and 2.57% respectively for accuracy, detection rate and FAR as it relates to 4-class detection, however, this study was based on social internet of things ,

meanwhile, a more generalized performance evaluation will give a better view of detection rate as this research is streamlined to social IoTs environment, (Diro and Chilamkurti, 2017).

A cloud based distributed deep learning frame work for fishing and botnet threats, the model encompasses a distributed convolutional neural network (DCNN) as well as long-short Term Memory (LSTM) security mechanism for optimal IoTs threat detections, 94.3%, and 93.58% score rate was recorded for accuracy and f1-score respectively against CNN and 94.8% and 0.6666% for accuracy and f1-score respectively for LSTM, nevertheless, an improved performance record can be achieved if the model used in this research can be turned after hybridization (La *et al.*, 2020)**.**

IoTs a multi-CNN fusion technique was proposed, the proposed technique was to serve as robust detection mechanism for network threat detection in industrial IoTs environment, it was claimed an exceptional accuracy rate of 86.95% was recorded for binary based classification and 81.33% accuracy for a multiclass classification, however, a broaden focus can expand the detection capability of IoTs threat attack as focus was based son industrial IoTs environment (Li *et al.*, 2019).

Almiani *et al.* (2020) proposed fog computing intrusion detection model for IoTs threat, the model was based on recurrent neural network enhanced through a back propagation algorithm, it was observed that the proposed model recorded an enhanced IoTs intrusion detection capability in terms of DR, accuracy, precision, FPR, F1-score,mcc, cohen's Kappa coefficient (K) and FNR of 94.27%, 92.18%, 90.23%, 9.8%, 92.29%, 84.44%, 84.36% and 5.7% respectively, however, more exploit on model can be done to enhance performance record for IoTs threat detection.

Almiani *et al.* (2020) proposed a deep recurrent Neural Network for IoTs intrusion detection system, the proposed multi-layered model was said to have achieved the following optimal performance of 94.27%, 92.18%, 9.8%, 92.29%, 84.44%, 84.36%, and 5.7% respectively for detection rate, accuracy, precision, false positive rate, f1-score, MCC, Kappa Statistics and false negative rate, however, improvement in performance can be experienced if an hybridization of model is employed.

In order to arrest the threat factor associated with IoTs technology, based on wireless sensor network (WSN), a novel architectural model for intrusion detection based on the approach of Map Reduce was proposed, it was claimed that the proposed model achieved a superior performance of 80.95% and 96.20% for anomaly and misuse detection module respectively, while 5.92% and 1.44% was recorded for false alarm rate (FAR). Major focus was streamlined to wireless sensor network (WSN) intrusion in IoTs, broader view of attack vector environment can further curtail IoTs threats (Bostani and Sheikhan, 2016).

Hodo *et al.* (2020) presents a threat analysis model of IoTs based Artificial Neural Network to enhance IoTs threat, the model proposed was focused on Distributed Denial of Service (DDoS) attacks, an accuracy of 99.4% was experienced, nevertheless, other performance evaluation metrics can detailed a clearer view of evaluation strength of the proposed model.

Bambang and Riri, (2020) proposed an algorithm that will detect denial of service (DoS) based on deep neural network in IoTs environment, an accuracy of 91.21% was achieved, however, for enhanced performance, a proposed combination of varied algorithms for detection of threat was suggested.

A convolutional neural network model for the detection of IoTs intrusion threat, was further claimed that the proposed model has a high sensitivity to IoTs network attack threats, the precision, recall, f1-score, miscalculation rate, as well as accuracy score achieved are 1.00%, 1.00%, 0.99%, 0.32%, and 98.60% respectively. However, more tuning can be done on models optimal performance scores (Smys and Wang, 2020).

2.4.1 Application of ensemble model

Moustafa *et al,* (2018) employed ensemble method for the detection of intrusion threat in IoTs environment, the proposed model was based on protection of statistical features flow of network traffic, the composition of ensemble model are Decision Tree, Naïve Bayes and Artificial Neural Network, it was claimed that the proposed model recorded an optimal performance in detection rate as well as low false positive rate, 99.54%, 98.93% and 1.38% was recorded for accuracy, detection rate and false positive rate respectively for DNS data source detection, while 98.97%, 97.02, and 2.58 was achieved for accuracy, detection rate and false positive rate respectively for http data source intrusion.

Saranya *et al.* (2020) explored a comparative study based on machine learning model performance for intrusion detection in order to establish the efficacy of model performance as it relates to IoTs environment, Linear Discriminant Analysis (LDA) Classification and Regression Trees (CART) , Random forest, Support Vector Machine (SVM), modified K-means, J48, Naïve Bayes, Decision Tree, Logistic Regression, and Artificial Neural Network (ANN), was explored, random forest was reported to have outperformed all other model aforementioned in terms of accuracy, with an accuracy rate of 99.81%, however, the employed dataset of KDD'99 cup has some recorded challenge of irrelevant features, which can affect performance record.

Smys and Wang (2020) proposed a hybrid binary classification technique known as DNN-KNN for the detection of intrusion threat in IoTs environment, the following performance scores was achieved, 99.77%, 0.23%, 99.74%, 99.76%, 99.77%, 99.75% and 99.53% for accuracy, error rate, precision, recall, true negative rate, f1-score and Matthew Correlation Coefficient (MCC) respectively, however, this gives a clearer identification of ensemble model, nevertheless, performance enhancement in terms of detection rate can be improved through varied application of other models.

Ensemble method was deployed for the detection of anomaly with a record performance of optimal rate, also Khraisat *et al.* (2019) employed the model of ensemble technique for optimal detection of intrusion threat in IoTs environment and achieved an outperforming record, Abdulrahaman and Alhassan (2018) also suggested the strength of ensemble model for enhanced performance record of detection of intrusion in environment related to information, security, an optimal performance was recorded in this research. Ensemble model noted as a model that enhance performance of learning model through combination of more than one weak or traditional classifier, has recorded an excellent performance in the field of threat detection and edge attack detection Basit *et al*. (2020); Torres (2018); Zhu *et al.* (2020), Figure 2.1 present the ensemble classifier (stacking) general pseudocode, which entails the process of a learner training combined individual traditional learners which is referred to as first level learners, while the combiner is known as the second level learner or meta-learner, the first level learner uses the original dataset to generate a new dataset for training the second level learner, where the output of the first level learner are recognized as the input features of the second level learner (Zhou *et al*., 2019).

Algorithm 1 - Stacking

Input : $D = \{(x_i, y_i) | x_i \in \chi,\ y_i \in Y\}$

Output : An ensemble classifier H

1. **Step 1 :** Learn first-level classifiers
2. For $t \leftarrow 1$ to T do
3. Learn a base classifier h_t based on D
4. **Step 2 :** Construct new data set from D
5. For $i \leftarrow 1$ *to* m do
6. Construct a new data set that contains $\{x_i^{new}, y_i\}$, where $x_i^{new} = \{h_j(x_i)\ for\ j = 1\ to\ T\}$
7. **Step 3 :** Learn a second-level classifier
8. Learn a new classifier h^{new} based on the newly constructed data set
9. **Return** $H(x) = h^{new}(h_1(x), h_2(x), \ldots, h_T(x))$

Figure 2.1 Pseudocode for Ensemble Classifier (Smyth and Wolpert., 1999)

CHAPTER THREE

3.0 RESEARCH METHODOLOGY

3.1 Research Design

Figure 3.1 represents the research design of a model for intrusion detection in IoTs environment. Furthermore, the research design encompasses the following phases; Phase1: problem formulation and planning, Phase 2: Design and implementation phase 3: Testing and Evaluation.

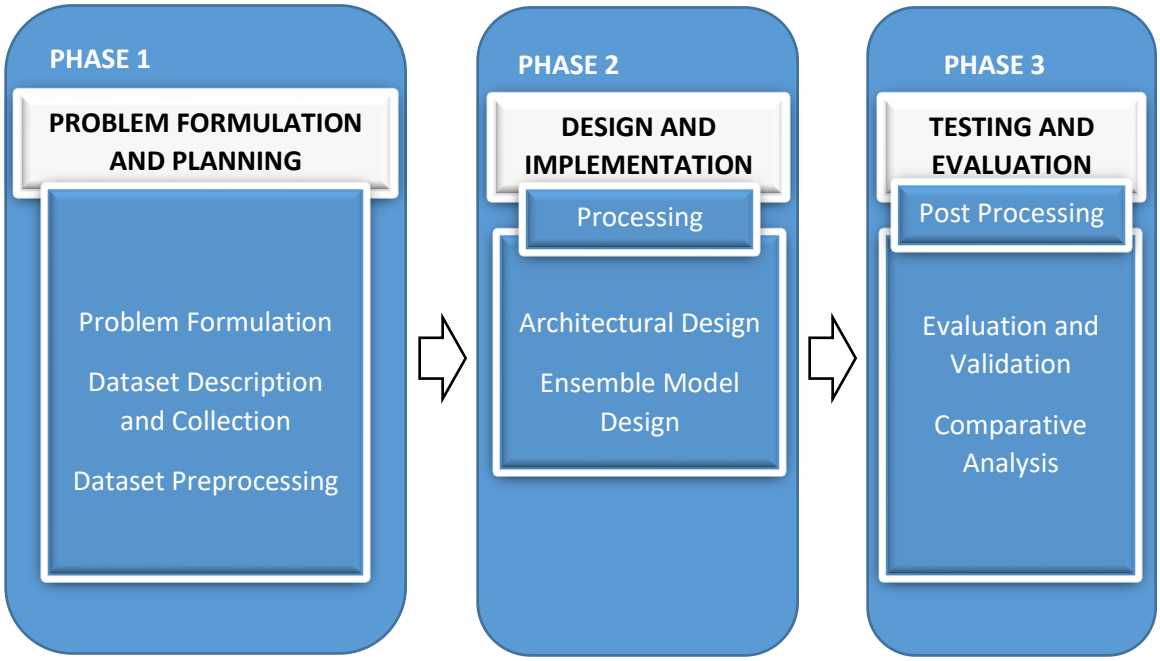

Figure 3.1 Proposed Research Design

3.2 Problem Formulation

Problem formulation is proposed from the following two important domains: namely; Application domain which encompasses the correlational/association, determination problem of outliers, classification and regression problems and the Technique domain which is made up of outlier analysis, feed forward neural network and support vector machine.

3.3 Dataset Collection and Description

The proposed dataset description gives brief definition of the features (input and output) and explanation of the values, this is the refined form of the original KDD Cup 99 IDS benchmark dataset developed in 2009. The KDD Cup 99 was refined to eliminate redundant records and include more reasonable number of instances. Meanwhile, the NSL-KDD has 41 different features with additional one attribute as class, the instances of the dataset is 125973. The dataset is made up of several exploited attacks categorized into four attack types: Remote-to-local (R2L), Denial of Service (DoS), Probe attack, and User-to-root (U2R) attack. In addition, the NSL-KDD data has several attack exploits, but categorized into the following four attack types: basic connection, host, traffic, and content as shown in Table 3.1.

Table 3.1 **NSL-KDD Dataset Description**

S/N	Features Category	Description	Features
1.	Basic Connection	These features provide information about TCP/IP connection	Duration type, flag, urgent, protocol type, wrong fragment, service, src bytes, land, dst bytes (9 features)
2.	Host Based	These features include information about the systems connected to the network	dst host serror rate, dst host count, dst host same src port rate, dst host rerror rate, dst host serror rate, dst host diff srv rate, dst host svr count, dst host svrrerror rate, dst host svr diff host rate, dst host same svr rate, (10 features)
3.	Traffic Based	The features that record system information based on the time window (of 2 seconds)	count, , same srv rate, diff srv rate, srv count, rerror rate, srvrerror rate, srvserror rate, srv diff host rate, serror rate (9 features)
4.	Content Based	Features suggested by domain knowledge	num failed logins, hot, logged in, root shell, num compromised, su attempt, num root, num shell, num file creations, num access files, is host login, num outbound cmds, is guest login (13 features)

3.4 Dataset Preprocessing

The proposed dataset preparation stages will involve the following steps: data discretization and preprocessing, dataset cleaning, dataset normalization and feature selection respectively.

a) Dataset discretization

This is the process of interpolating values with the aim of minimizing the number of possible state a feature can hold. It is a form of data transformation.

b) Dataset cleaning

Dataset cleaning aims at handling outliers and missing values. This stage will result in producing a more reliable dataset and hence more robust model. The following basic method for handling missing values will be deployed; Inputting missing values for samples with missing values that are not supposed to be removed.

c) Dataset normalization

Dataset normalization is therefore the appropriate method that may be utilized to prevent outweighing features that hold larger ranges, and one of the commonly utilized methods is to scale dataset values in a predefined range, which will result in better performance of the proposed ensemble model. Min-Max technique was employed for normalization, which has the strength of converting the dataset into a bound range of [0, 1]; the mathematical representation is presented in equation (3.1).

$$p = \frac{(x - x_{min})(max - min)}{(x_{max} - x_{min}) + min} \tag{3.1}$$

3.4.1 Data preprocessing

This is a method applied on dataset to address the challenge of biasness that may exist in the dataset which can impede on the performance of learning algorithm. Outlier analysis and feature normalization were implemented in this research.

3.4.2 Outlier analysis

The technique known as capping the outlier dataset was implemented. This technique replaces the outlier data, values with the upper and lower bounds respectively. That is replacing upper bound values with outliers that are found to be at more upper bound as well as applying same to lower bound values.

3.4.3 Feature normalization

This method scales each values of the feature into a define range, which will result into a smoother dataset, Min-Max scale was employed in this research

3.4.4 K-fold cross validation

A cross validation value of 5, was implemented in this research whereby the dataset was divided into 5 equal parts, a part is held for testing and the remaining parts are used for training, this continued in an iterative step of 5 iterations in same manner as outlined above.

3.5 Traditional Classifiers

The traditional classifiers which can also be referred to as the base learner algorithms implemented in this research are: SVM and FFNN, these base learner receives the dataset from the cross validation test option, use it for training its model and therefore generates a new training dataset for the next stage learner, which is the metal learner(Ensemble

classifier), the output of the base learner represents the inputs features for the next stage learner and the original class of the dataset still stands as the new class for the new training dataset.

3.6. The Ensemble Classifier

The ensemble classifier which can be called a meta learner, trains the meta- model based on the input received from the base learner as well as the defined FFNN classifier to determine the ensemble model final prediction, Table 3.2 represent the pseudocode of the proposed ensemble model for this research.

3.7 Proposed Architecture Design

The Figure 3.2 represents the block architectural design which is made of data collector module which receives the input data for analysis which will be further fed into dataset preprocessor for data preparation followed by the next stage which is feature selector module where relevant feature selection takes place. Furthermore, the features and instances optimization stage will deploy the following techniques; outlier analysis. The output will be fed into the last stage which is the traditional model optimization (FFNN and SVM) which will generate the improved output. While Figure 3.3 represent the proposed architectural design, encompassing the following components; IoTs collected data packets, structured into a recognized research dataset repository, which in turn is feed into an outlier analysis in order to address the challenge of biasness that exist in the dataset for better and improved performance in terms of training the dataset, the features are feed into the traditional classifiers (FFNN and SVM) for training, the meta learning collects the output of the traditional.

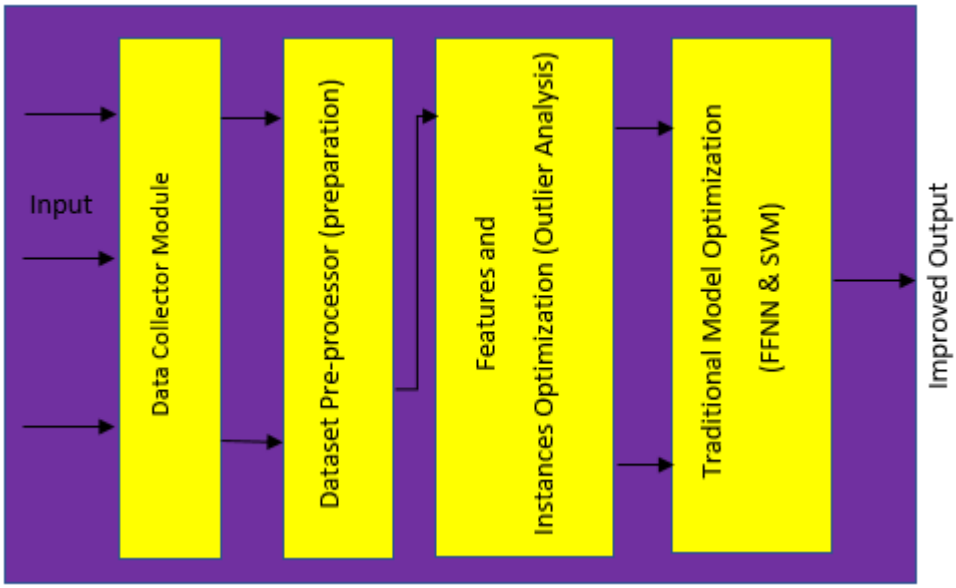

Figure 3.2 Block Architectural Design

Table 3.2 **Proposed Ensemble Learning Model Pseudocode**

Algorithms : Proposed Ensemble Learning Model
1: **Transform Dataset: Outlier Analysis & Normalization**
2: **Input:** Training Dataset DT = $\{P_i, Y_i\}_{i=1}^{n}$ $(P_i \in R^n, Y_i \in \gamma)$
1st level classifier SVM, FFNN
2nd level classifier FFNN
3: **Output**: An Ensemble Algorithm **H**
4: Use a 5-fold cross validation resampling method to generate training set for (2nd) level classifier
5: From DT split **K** equal size subsets DT = {DT1, DT2, DT3, … DT5}
6: **for** k = 0 to 4 do
7: Train the (1st) level classifiers
8: **for** m = 0 to 1 **do**
9: Train a classifier C_{km}
10: **end for**
11: **end for**
13: Create training set for (2nd) level classifier
14: **for** $P_i \in D$ do, obtain instance $\{P_i', Y_i\}$, where $P_i^{`} = \{C_{k1}(P_1), C_{k2}\}$ **end for**
15: Train (2nd) level classifier $C^{`}$ using all dataset of $\{P_i^{`}, Y_i\}$
16: Train (1st) level classifiers
17: **for** m = 0 to 1 **do**
18: Train classifier C_m based on DT
19: **end for**
20: **return** $H(P) = C^{`}(C_1(P), C_2(P))$

classifier, serving as features to be trained, and generation of and ensemble prediction to obtained based on normal or malicious attack.

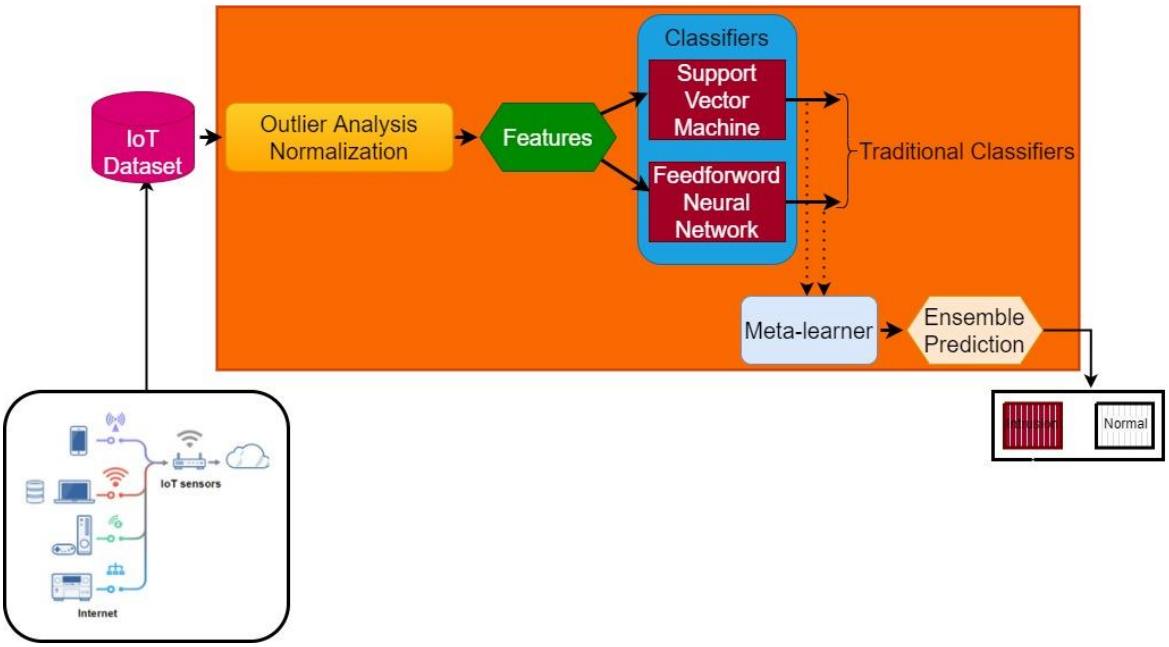

Figure 3.3 Proposed Architectural Design

3.8 Ensemble Model Formulation

The ensemble model is based on the following:

i. Outlier analysis: this was deployed to locate and features that are out of bound and making the features to fall under same bound in respect to IoTs environment dataset, removing outliers from the dataset addresses inconsistent values that make learning difficult.

ii. Feed forward neural network model: it is a machine learning model that was used as part of the ensemble model formulation. Artificial neural network (ANN) known as a computerized model of the human brain and nervous system, it consists of a set of nodes. Each node represents a neuron or a

processing unit, the nodes output depends on a parameter called the connection weights or synapse strengths, Given j is hidden layer of a FFNN, while y_i been the output of j for input $(x_i - x_n)$ is evaluated using $y_i = f\left(\sum_{i=1}^{h} w_{ji}x_i + b_j\right)$, where b_i represents bias on node j, w_{ji} been the matrix of a layers and the function $f(.)$ Is known as the activation function

iii. Support vector machine: It forms a high hyperplane or set of hyperplane in a high or infinite dimension space that can be used for classification, regression or other tasks. SVM can handle multiple, continuous, and categorical data.

3.8.1 Proposed model

The Figure 3.4 depicts the composition of the proposed Ensemble model. The composition encompasses; Dataset preprocessing, deployment of K-fold test option, training of Traditional classifiers, final prediction by the ensemble classifier, thus explained below:

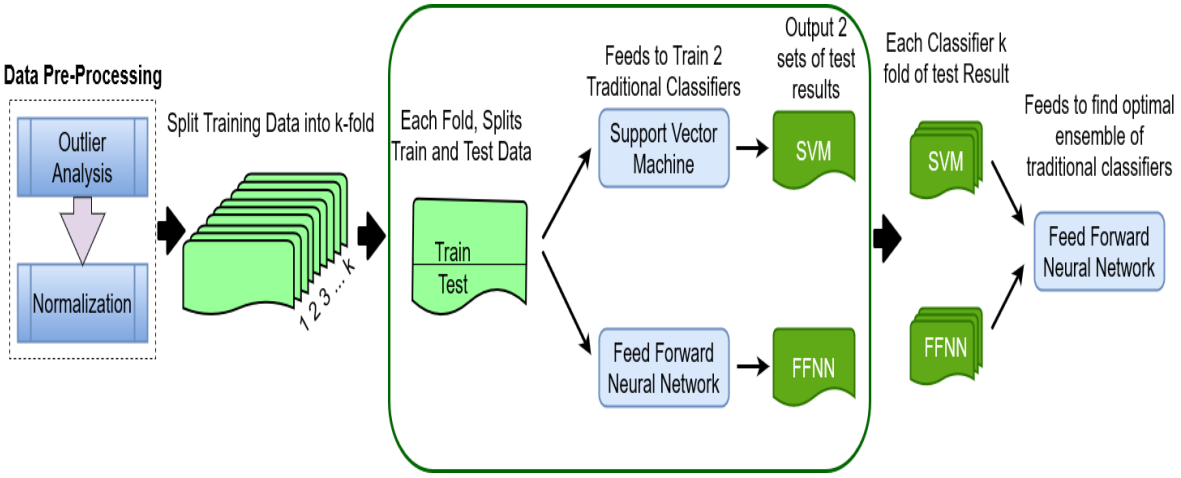

Figure 3.4 Proposed Ensemble Model

3.9 Evaluation and Validation

The formation of the ensemble model will be evaluated based on four metrics: precision, recall, f1-score and accuracy.

i. Precision (p): precision is the fraction of relevant instances among the retrieved instances.

$$P = \frac{TP}{TP+FP} \tag{3.1}$$

ii. Recall (R): This is equivalent to TP-rate (sensitivity).

iii. F1-score (Harmonic mean): It is the weighted average of P and R. this metric weights R and P equally.

$$F1 = \frac{2\,P\,R}{P+R} \tag{3.2}$$

iv. Accuracy (ACC): This is the overall rate of correctly classified examples in the testing dataset.

$$Acc = \frac{TP+TN}{TP+FP+TN+FN} \tag{3.3}$$

In order to assess the error-rate accurately in an unbiased manner,10- cross fold validation was used.

CHAPTER FOUR

4.0 RESULTS AND DISCUSSION

This chapter presents the outcome of proposed model for intrusion detection in IoTs environment based on ensemble technique, and also discussion of the findings.

4.1 Result for Ensemble model (FFNN and SVM) for Intrusion Detection in IoTs

To perform a comparative analysis based on the designed model for intrusion detection in IoTs environment, the architecture and the parameters employed needs to be tuned in order to be established, this entails the parameter values to be used for analysis as well as evaluation, the default configuration of both SVM and FFNN algorithms as present in Google collaboration was adopted for this research with the use of Tensor Processing Unit (TPU) as the processing power platform.

4.2 Result Based on Unprocessed Dataset

The result obtained as represented in Table 4.1 is based on unprocessed dataset and the performance was measured using the following metrics: Precision, Recall, F-score and Accuracy in detection of intrusion threat in IoTs environment based on ensemble technique.

Table 4.1 **Ensemble Model (FFNN and SVM) Result with Unprocessed Data**

Accuracy	Recall	Precision	F-score	FAR
0.9857	0.9777	0.9935	0.9855	0.897

The accuracy of 0.9857 was obtained in the analysis based on the proposed ensemble model, while 0.9777, 0.9935, 0.9855, 0.897 respectively was achieved for recall, precision and f-

score and FAR. However, only precision score was able to attain slightly above 0.9900, improvement can be reached if the dataset is well processed for enhanced performance of the model.

4.3 Result Based on Processed Dataset

The Table 4.2 represents the outcome of the result obtained from the processed intrusion detection dataset that was deployed in the proposed ensemble model.

Table 4.2 **Ensemble Model (FFNN and SVM) Result with Processed Data**

Accuracy	Recall	Precision	F-score	FAR
0.9996	0.9996	0.9996	0.9996	0.342

The Table 4.2, stands for ensemble technique based base model for intrusion detection in IoTs environment. The accuracy, recall, precision and F-score of the proposed ensemble model are 0.9996 respectively each, while FAR scored 0.324 for the evaluation metrics.

4.4 Comparative Analysis of Unprocessed and Preprocessed Intrusion Detection Dataset

The Table 4.3 is a detail comparative result of the output generated from the use of unprocessed and preprocessed intrusion detection dataset based on IoTs environment.

The result reflected in Table 4.3 and Figure 4.1, depicts a distinct out-performance across the accuracy, recall, precision, f-score and FAR model as well as execution period of the proposed ensemble model. The proposed ensemble model looks promising in terms of performance evaluation in regards to the performance metrics employed in this research,

likewise time complexity, while the out-performance experienced notably is as a result of preprocessing methods employed on the dataset in order to address the challenge such as execution period, and performance efficiency.

Table 4.3 **A Comparative Analysis of Unprocessed and Preprocessed Dataset**

	Accuracy	**Recall**	**Precision**	**F-score**	**FAR**	**Execution time (secs)**
Unprocessed Dataset	0.9857	0.9777	0.9935	0.9855	0.897	43200
Preprocessed Dataset	0.9996	0.9996	0.9996	0.9996	0.342	7200

The strength of outlier analysis and data normalization improved the result output of the proposed ensemble model. The accuracy of the proposed ensemble model when the dataset is preprocessed is 0.9996 against 0.9857 with an unprocessed dataset proving the efficiency of outlier analysis and normalization of dataset before training coupled with ensemble model capability across the board, a recall of 0.9996 was also attained likewise precision and f-score based on preprocessed dataset against 0.9777, 0.9935, and 0.9855 respectively for the unprocessed dataset as it regards to recall, precision and f-score respectively, noting the enhanced score of FAR of 0.897 and 0.342 for unprocessed and processed data respectively . The time complexity also proved to have a sharp increase performance comparing the execution time of the preprocessed dataset against the unprocessed dataset. It can be noted that a variation of 36000 seconds which is quite massive is experienced if the dataset is not been addressed to enhance performance as experienced from the output of the result discussions and presented in Figure 4.2.

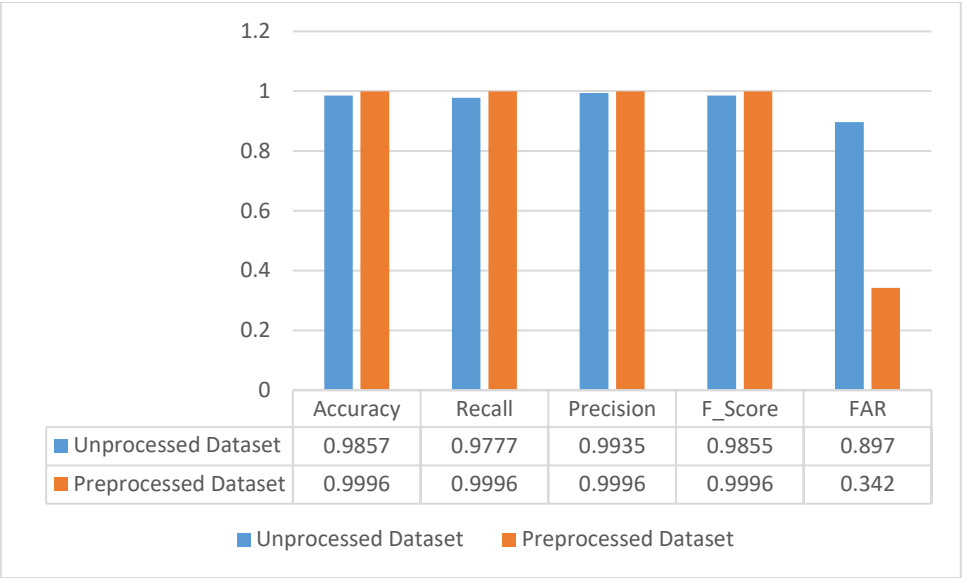

Figure 4.1: A Comparative Analysis of Unprocessed and Preprocessed Dataset

Figure 4.2: A Comparative Analysis of Unprocessed and Preprocessed Dataset Execution Time

4.5 Benchmark Model Analysis with Adopted Intrusion Detection Dataset

The adopted baseline literature technique is compared with the proposed model that was deployed for this research. The model that was used by the baseline literature of this research, was compared with the ensemble model proposed in this research. This is to enable the research to have a balanced view and also to address the challenge of biasness in terms of the performance analysis evaluation, thus proving the enhancement strength of the proposed model.

The Figure 4.3 reflects the outcome of a comparison between the proposed model in this research and the models that was used by baseline literature.

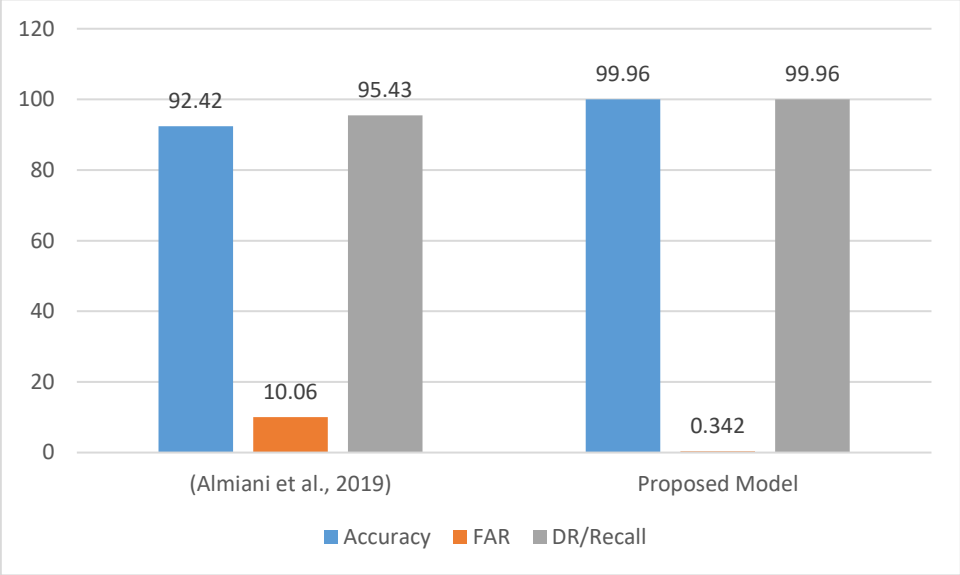

Figure 4.3 Comparison Analyses of Proposed Model and Baseline Literature Models

4.5.1 Comparison analysis

The indication from Figure 4.5 shows that the baseline literature achieves a performance score rate of 92.42, 95.43 and 10.6 respectively for accuracy, recall and FAR. However, the proposed ensemble model in this research attained 99.96 for accuracy and recall respectively,

while FAR of 0.342 was achieved, this is a clear indication of excellent performance as it outperforms the baseline literature model of deep neural network. This further shows the capability of ensemble technique against single model in terms of performance analysis as can be deduced from the result generated in this result which proved that combining multiple traditional classifiers will generate a promising result output.

CHAPTER FIVE

5.0 CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

In the current 21st century of rapid exponential growth of interconnected computing devices experienced across the globe, IoTs threat is inevitable. Therefore, the detection of intrusion threat in IoTs environment is of paramount interest in information security field, with the need for investigative processes and establishment of facts for compromise through exploit by unwanted entities, traditional classification models have been deployed by researchers with good result output, however, performance can be improved on. Ensemble model serves as a promising model that can be employed for optimal performance based on its prove of strength in literatures, which of cause is established in this research.

An ensemble model was proposed in this research with, Support Vector Machine (SVM) and Feed Forward Neural Network (FFNN) adapted as the base learners that generates an input feature for the meta-learner that combine the performance of the base learner for optimal result. Also, feature optimization analysis was deployed based on outlier analysis and model was proposed for intrusion detection in IoTs environment based on ensemble model. 0.9996 was achieved for the performance evaluation accuracy, precision, recall and f-measure respectively, which is a prove of excellent performance over traditional classifier serving as a promising field of machine learning that can be further explored and stands out for optimal performance analysis in intrusion detection exploit.

5.2 Recommendations

The research recommends further exploration of the strength of ensemble model techniques through additional multiple variation of traditional models, likewise other variant of feature optimization algorithm for enhanced performance, not leaving out hybridization of machine learning models and the employment of feature selection techniques for further research which is a projected promising field to explore.

5.3 Contributions to Knowledge

i. Optimal performance for IoTs intrusion threat detection was achieved with an excellent accuracy rate of 99.96% based on the proposed model in this research.

ii. An architecture design for enhanced IoTs intrusion threat detection was achieved.

.

REFERENCES

Abdulrahaman, M. D., & Alhassan, J. K. (2018). Ensemble learning approach for the enhancement of performance of intrusion detection system. *In International Conference on Information and Communication Technology and its Applications,* (pp. 1-8). Minna, Nigeria: Imanager

Alharbi, S., Moa, B., Weber-Jahnke, J., & Traore, I. (2012). High performance proactive digital forensics. *In Journal of Physics: Conference Series*, 385, 1–15. doi:10.1088/1742-6596/385/1/012003

Almiani, M., AbuGhazleh, A., Al-Rahayfeh, A., Atiewi, S., & Razaque, A. (2020). Deep recurrent neural network for internet of things intrusion detection system. *Simulation Modelling Practice and Theory*, 101, 102031. doi:10.1016/j.simpat.2019.102031

Alsamiri, J., & Alsubhi, K. (2019). Internet of Things cyber attacks detection using machine learning. *International Journal Advance Computer Science and Application*, 10(12), 627-634. doi: 10.14569/IJACSA.2019.0101280

Alshahrani, H. M. (2021). A collaborative intruder detection system for internet of things devices. *Electronics*, 10(7), 848. doi: 10.3390/electronics10070848

Alshammari, A., & Zohdy, M. A. (2019). Internet of things attacks detection and classification using tiered hidden Markov model. *In Proceedings of the 8th International Conference on Software and Computer Applications,* (pp. 550-554). New York, United States: ACM

Bambang, S., & Riri, F. S. (2020). Intrusion detection in internet of things networks using deep learning algorithm. *Information,* 11(5), 279. doi: 10.3390/info11050279

Basit, A., Zafar, M., Javed, A. R., & Jalil, Z. (2020). A novel ensemble machine learning method to detect phishing attack. *In 23rd International Multitopic Conference (INMIC),* (pp. 1-5). Bahawalpur, Pakistan: IEEE

Bostani, H., & Sheikhan, M. (2016). A hybrid intrusion detection architecture for internet of things. *In 8th International Symposium on Telecommunications (IST),* (pp. 601-606). Tehran, Iran: IEEE.

Boswell, D. (2002). *Introduction to support vector machines.* California San Diego: IEEE

Chaabouni, N. (2020). *Intrusion detection and prevention for internet of technology systems using Machine Learning* (Doctoral dissertation, Université de Bordeaux).

Cvitić, I., Peraković, D., Periša, M., & Gupta, B. (2021). Ensemble machine learning approach for classification of internet of things devices in smart home. *International Journal of Machine Learning and Cybernetics*, 1-24. doi: 10.1007/s13042-020-01241-0

Davahli, A., Shamsi, M., & Abaei, G. (2020). Hybridizing genetic algorithm and grey wolf optimizer to advance an intelligent and lightweight intrusion detection system for internet of things wireless networks. *Journal of Ambient Intelligence and Humanized Computing*, 11(11), 5581-5609. doi: 10.1007/s12652-020-01919-x

Dietterich, T. G. (2000). Ensemble methods in machine learning. *In International Workshop on Multiple Classifier Systems* (pp. 1-15). Berlin, Heidelberg: Springer

Diro, A. A., & Chilamkurti, N. (2017). Distributed attack detection scheme using deep learning approach for internet of things. *Future Generation Computer Systems*. doi: 10.1016/j.future.2017.08.043

Fatayer, T. S., & Azara, M. N. (2019). Internet of things secure communication using artficial neutral network classification algorithms. *In International Conference on Promising Electronic Technologies (ICPET),* (pp. 142-146). Gaza, Palestine: IEEE.

Hodo, E., Bellekens, X., Hamilton, A., Dubouilh, P. L., Iorkyase, E., Tachtatzis, C., & Atkinson, R. (2020). Threat analysis of internet of things networks using artificial neural network intrusion detection system. *In International Symposium on Networks, Computers and Communications (ISNCC),* (pp. 1-6). Yasmine Hammamet, Tunisia: IEEE.

Ioannou, C., & Vassiliou, V. (2019). Classifying security attacks in internet of things networks using supervised learning. *In 15th International Conference on Distributed Computing in Sensor Systems (DCOSS),* (pp. 652-658). Santorini, Greece: IEEE

Khare, N., Devan, P., Chowdhary, C. L., Bhattacharya, S., Singh, G., Singh, S., & Yoon, B. (2020). Spider monkey optimization and deep neural network hybrid classifier model for intrusion detection. *Electronics*, 9(4), 692. doi: 10.3390/electronics9040692

Khraisat, A., Gondal, I., Vamplew, P., Kamruzzaman, J., & Alazab, A. (2019). A novel ensemble of hybrid intrusion detection system for detecting internet of things attacks. *Electronics*, 8(11), 1210. doi: 10.3390/electronics8111210

Krishna, A., Lal, A., Mathewkutty, A. J., Jacob, D. S., & Hari, M. (2020). Intrusion detection and prevention system using deep learning. *In International Conference on Electronics and Sustainable Communication Systems (ICESC),* (pp. 273-278). Coimbatore, India: IEEE.

Kumar, P., Gupta, G. P., & Tripathi, R. (2021). An ensemble learning and fog-cloud architecture-driven cyber-attack detection framework for internet of medical things networks. *Computer Communications*, *166*, 110–124. doi: 10.1016/j.comcom.2020.12.003

La, G. De, Parra, T., Rad, P., Choo, K. R., & Beebe, N. (2020). Detecting internet of things attacks using distributed deep learning. *Journal of Network and Computer Applications*,

163, 102662. doi:10.1016/j.jnca.2020.102662

Li, Y., Xu, Y., Liu, Z., Hou, H., Zheng, Y., Xin, Y., ... & Cui, L. (2019). Robust detection for network intrusion of industrial internet of things based on multi-convolution neural network fusion. *Measurement*, 154, 107450. doi: 10.1016/j.measurement.2019.107450

Lin, J., & Zhang, J. (2013). A fast parameters selection method of support vector machine based on coarse grid search and pattern search. *In Fourth Global Congress on Intelligent Systems*, 1(4), (pp. 77-81). Hong Kong, China: IEEE

Liu, J., Kantarci, B., & Adams, C. (2020). Machine learning-driven intrusion detection for contiki-next generation-based internet of things networks exposed to network security laboratory-knowledge discovery and data dataset. *In Proceedings of the 2nd Workshop on Wireless Security and Machine Learning*, (pp. 25-30).New York, United States: ACM

Meneghello, F., Calore, M., Zucchetto, D., Polese, M., & Zanella, A. (2019). Internet of threats: A survey of practical security vulnerabilities in real internet of things devices. *Institute of Electrical and Electronics Engineers Internet of Things Journal*, 6(5), 8182-8201. doi: 10.1109/JIOTS.2019.2935189

Mliki, H., Kaceam, A. H., & Fourati, L. C. (2020). Risks and security of internet and systems. *14th International Conference Crisis*, 29–31. Hammamet, Tunisia: Springer Nature

Mol, P. R., & Mary, C. I. (2021). Classification of network intrusion attacks using machine learning and deep learning. *Annals of the Romanian Society for Cell Biology*, 25(2), 1927-1943. Retrieved from https://www.annalsofrscb.ro/index.php/journal/article/view/1137

Moustafa, N., Turnbull, B., & Choo, K. K. R. (2018). An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of internet of things. *Institute of Electrical and Electronics Engineers Internet of Things Journal*, 6(3), 4815-4830. doi: 10.1109/JIOTS.2018.2871719

Nazarpour, M., Nezafati, N., & Shokuhyar, S. (2020). Detection of attacks and anomalies in the internet of things system using neural networks based on training with particle swarm optimization and teaching learning based optimization algorithms. *Signal Processing and Renewable Energy*, 4(4), 81-94. Retrieved from http://www.iaujournals.ir/article_677172.html

Paliwal, S., & Gupta, R. (2012). Denial-of-service, probing and remote to user (r2l) attack detection using genetic algorithm. *International Journal of Computer Applications*, 60(19), 57-62. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.303.3876&rep=rep1&type=pdf

Rajendran, G., Nivash, R. R., Parthy, P. P., & Balamurugan, S. (2019). Modern security

threats in the internet of things: Attacks and countermeasures. *In International Carnahan Conference on Security Technology (ICCST)*, (pp. 1-6). Chennai, India: IEEE

Ren, Y., Zhang, L., & Suganthan, P. N. (2016). Ensemble classification and regression-recent developments, applications and future directions. *Computational Intelligence Magazine*, 11(1), 41-53. doi: 10.1109/MCI.2015.2471235

Rezvy, S., Luo, Y., Petridis, M., Lasebae, A., & Zebin, T. (2019). An efficient deep learning model for intrusion classification and prediction in 5g and internet of things networks. *53rd Annual Conference on Information Sciences and Systems (CISS)*, 1–6. Baltimore, MD, USA: IEEE

Salazar, C. (2019). Internet of things: Definition, characteristics, architecture, enabling technologies, application & future challenges. *International Journal of Engineering Science and Computing*, 6(5). doi: 10.4010/2016.1482

Samy, A., Yu, H., & Zhang, H. (2020). Fog-based attack detection framework for internet of things using deep learning. *Institute of Electrical and Electronic Engineers Access*, 8, 74571-74585. doi: 10.1109/ACCESS.2020.2988854

Santoyo-gonz, A. (2019). High-performance, platform-independent distributrd denial of service detection for internet of things ecosystems. *In 44th Conference on Local Computer Networks (LCN)* (pp. 69-75). Osnabrueck, Germany: IEEE

Saranya, T., Sridevi, S., Deisy, C., Chung, T. D., & Khan, M. A. (2020). Performance analysis of machine learning algorithms in intrusion detection system: A review. *Procedia Computer Science*, 171, 1251-1260. doi: 10.1016/j.procs.2020.04.133

Servida, F., & Casey, E. (2019). Internet of things forensic challenges and opportunities for digital traces. *Digital Investigation*, 28, S22-S29. doi: 10.1016/j.diin.2019.01.012

Sharma, A. (2020). A novel approach for early malware detection. *Transactions on Emerging Telecommunications Technologies*, 32(2), 1–19. doi: 10.1002/ett.3968

Singh, B., Rai, C. S., Balamurugan, B., & Al-turjman, F. (2020). An intrusion detection scheme based on the ensemble of discriminant classifiers. *Computers & Electrical Engineering*, 86, 106742. doi: 10.1016/j.compeleceng.2020.106742

Smys, S., & Wang, H. (2020). Hybrid intrusion detection system for internet of Things. *Journal of Internet of Things in Social, Mobile, Analytics & Cloud*, 2(04), 190-199. doi: 10.36548/ijismac.2020.4.002

Smyth, P., & Wolpert, D. (1999). Linearly combining density estimators via stacking. *Machine Learning*, 36(1), 59-83. doi: 10.1023/A:1007511322260

Thamilarasu, G., & Chawla, S. (2019). Towards deep-learning-driven intrusion detection for the internet of things. *Sensors*, 19(9), 1977. doi: 10.3390/s19091977

Torres, J. F. (2018). Stacking ensemble learning for short-term electricity consumption forecasting. *Energies*, 11(4), 949. 1–31. doi: 10.3390/en11040949

Vega-barbas, M., Characterization, P., Rivera, D., & Rodrigo, M. S. (2021). An internet of things-focused intrusion detection system approach based on preprocessing characterization for cybersecurity datasets. *Sensors*, 21(2), 656. doi: 10.3390/s21020656

Vishwanathan, S. V. N., & Murty, M. N. (2002). Simple support vector machine: a simple SVM algorithm. *In Proceedings of the International Joint Conference on Neural Networks. IJCNN'02,* (Vol. 3, pp. 2393-2398). Honolulu, HI, USA: IEEE

Vu, L., Nguyen, Q. U., Nguyen, D. N., Hoang, D. T., & Dutkiewicz, E. (2020). Deep transfer learning for internet of things attack detection. *Institute of Electrical and Electronic Engineer Access*, 8, 107335-107344. doi: 10.1109/ACCESS.2020.3000476

Wazirali, R. (2020). An improved intrusion detection system based on k-nearest neigbor hyperparameter tuning and cross-validation. *Arabian Journal for Science and Engineering*, 45(12), 10859-10873. doi: 10.1007/s13369-020-04907-7

Yakubu, O., Adjei, O., & Narendra, B. C. (2018). A review of prospects and challenges of Internet of Things. *International Journal of Computer Applications*, 139(10), 33-39. doi: 10.5120/IJCA2016909390

Zhang, Y., Li, P., & Wang, X. (2019). Intrusion detection for Internet of things based on improved genetic algorithm and deep belief network. *Institute of Electrical and Electronics Engineers Access*, 7, 31711-31722. doi: 10.1109/ACCESS.2019.2903723

Zhou, A., Kardani, N., Nazem, M., & Shen, S. L. (2021). Improved prediction of slope stability using a hybrid stacking ensemble method based on finite element analysis and field data. *Journal of Rock Mechanics and Geotechnical Engineering*, *13*(1), 188-201.

Zhu, H., Li, Y., Li, R., Li, J., You, Z. H., & Song, H. (2020). Sedmdroid: An enhanced stacking ensemble of deep learning framework for android malware detection. *IEEE Transactions on Network Science and Engineering*. 8(2), 984–994. doi: 10.1109/TNSE.2020.2996379

APPENDIX A

Published Article

Edward, E. O., & Ojeniyi, J. A. (2019, December). A systematic literature review on digital evidence admissibility: Methodologies, challenges and research directions. *In 15th International Conference on Electronics, Computer and Computation (ICECCO)* (pp. 1-7). Abuja: Nigeria: IEEE