2023 4th International Conference for Emerging Technology (INCET)
Belgaum, India. May 26-28, 2023

Intrusion Detection in Wireless Sensor Networks Using Deep Learning

Aman Goyal
Dept. of Information Technology
Indian Institute of Information Technology Allahabad
Prayagraj, India
amang13298@gmail.com

Sourav Mishra
Dept. of Information Technology
Indian Institute of Information Technology Allahabad
Prayagraj, India
rsi2019005@iiita.ac.in

Dr. Vijay K Chaurasiya
Dept. of Information Technology
Indian Institute of Information Technology Allahabad
Prayagraj, India
vijayk@iiita.ac.in

2023 4th International Conference for Emerging Technology (INCET) | 979-8-3503-3575-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/INCET57972.2023.10170243

***Abstract*—WSNs are currently seeing widespread adoption across a variety of settings, including the medical industry, manufacturing environments, and a number of other domains. WSNs are distinguished by a number of distinguishing qualities, including their limited power supply, use of bandwidth, and overall energy consumption. Because of this, traditional networks can be secured using a wide variety of methods, whereas wireless sensor networks (WSNs) cannot be protected using the same methods. In order to make wireless sensor networks (WSNs) safer as a whole, it was necessary to develop new ideas and approaches. The prevention of intrusion is the most important concern in wireless sensor networks (WSNs). This study develops a Deep Learning (DL) method for Intrusion Detection Systems (IDS) in Wireless Sensor Networks (WSNs) by making use of Dense Artificial Neural Networks (Deep-ANN). When compared to earlier models, the one that was developed using Ann achieved the highest level of accuracy, which was 96.45 percent. The ANN model outperforms the other ML models that are currently available thanks to its superior precise recall as well as F1 scores of 96.38, 98.94, and 97.64.**

***Index Terms*—Deep Learning, Wireless Sensor Networks, Intrusion Detection Systems, Artificial Neural Network.**

I. INTRODUCTION

The schemes of IoT have attained a lot of attention in current years to gather sensing data and develop intelligent services and applications. The IoT defines the association of devices or any devices with the internet through embedded software and sensors which are utilized for communicating, gathering and exchanging the data with one another. The world becomes wide open with the deployment of Internet of Things that provides a virtually endless array of opportunities and connections at home, at work or at play. The connectivity is integrated with sensors, devices and people via Internet of things and a form of free-flowing conversation is established amid man and machine, software and hardware. These conversations allow enable devices for anticipating, reacting, responding and improving the physical world in almost the similar way that the networks and computer screens are carried out in the internet for the improvement of information world due to the development of AI and ML [26].

It is supposed that the Internet of Things is expected to expand at rapid rate over the coming years and a new dimension of services will be set free in this convergence which assisted in enhancing the quality of life of consumers and efficiency of enterprises, to unlock an opportunity that the GSMA terms as the Connected Life. The IoT is capable for delivering solutions using which energy efficiency and various other aspects of daily life are enhanced for the customers [27]. This technology has provided solutions which are useful for improving the decision-making and production in manufacturing, retail, agriculture and other sectors for the enterprises.

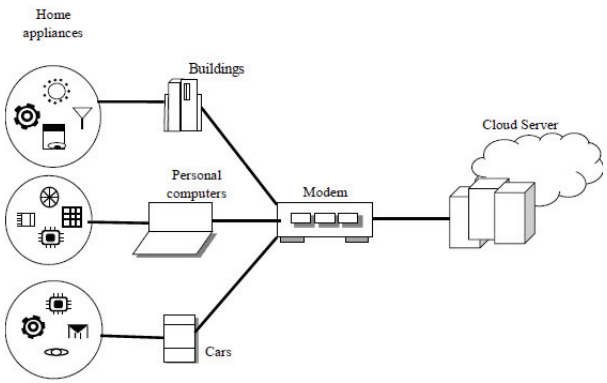

Fig. 1.

The above figure shows the whole scenario of IoT system in diagrammatic manner which represents a number of physical devices in which inter-connectivity amongst home appliances, buildings, personal computers and cars has been established. The environment is set up in this diagram to carry out sensors, actuators, essential electronics, software, and network connectivity in order to make communication between all connected entities possible. Each object in the IoT network

is given a distinct identity and IP address. It is known as M2M [28]. The things or devices connected in this network are used to share information so that many tasks can be completed according to the network owner's requirements.

A. IoT applications

Almost all of the applications of Internet of Things have taken the term smart such as Smart Home or connected namely Connected Health. The main applications of the recent has been consisted of:

- **Smart home**: The connectivity inside the houses is defined using the term Smart Home. In this, smoke detectors, light bulbs, devices, entertainment systems, windows, door locks etc. are comprised. Nest, Apple and Philips are some well-known companies which are included in it.
- **Wearables**: It assisted in making a large part of the customer who utilized IoT applications whether it is Jawbone Up, the Fitbit Flex or the Apple Smartwatch.
- **Smart City**: A huge range of use cases, from managing the traffic to the distribution of water, to managing the waste, urban security and to monitor the environment, all are included in the smart city. It assisted in mitigating the real pains of people who are lived in cities now-a-days. It is helpful to solve the issues of traffic congestion, to mitigate noise and pollution and to make the cities safer.
- **Smart grids**: It enables the deployment of information related to the behaviours of electricity suppliers and consumers in an automated fashion so that the efficiency and economics of electricity will be enhanced.
- **Industrial internet**: There are a number of market research in which the industrial internet is considered as the Internet of Things concept that has the highest overall potential. The smart factories or connected industrial equipment are comprised in the applications among others.
- **Connected Health**: The connected health care system and smart medical devices have shown to be useful for both companies and people. Its expected benefits include new types to track a patient's health in real time and better medical decisions based on a lot of patient data.
- **Smart farming**: The framing becomes a remarkable case for IoT due to the detachment of farming operations and the great number of livestock which is easily monitored.

B. Security in IoT

The security considerations have been present since long time in the context of information technology. The new and unique security challenges generate regularly as consequences of attributes of IoT implementations with multiple features. Therefore, it is necessary to address these challenges and ensuring security in IoT products and services at utmost priority. IoT users must be ensured about the security of IoT devices and related data services from vulnerabilities. This is critical since technology has become increasingly universal and ingrained in everyone's daily lives. Unsecured IoT devices and services could be exploited to launch a cyberattack. Attackers can take advantage of this circumstance by making it simple to steal user data and failing to safeguard data streams adequately. Because Internet of Things devices are interconnected, each one with wide security that is connected online has the potential to disrupt the overall security and reliability of the Internet. This issue is exacerbated by factors such as the adoption of uniform IoT devices, the chance that some devices would automatically link to other devices, and the possibility that these devices will be utilized in non-unsecure locations. [5]. There is a collective responsibility of developers and users of IoT devices and systems for ensuring that users and the Internet is not explored by them to potential harm because of the matter of principle. Therefore, a collaborative approach will be required for the security so that the effectual and suitable solutions are designed to Internet of Things security challenges which are well-fitted to the scale and complexity of the issues.

C. Attacks in IoT

Some commonly launched security attacks in IoT have been discussed below:

- **Denial of Service (DoS) attacks**: DoS attacks on computing nodes can be classified into three types: sleep deprivation, outages, and battery draining attacks. Sleep deprivation is a type of Dos attack. During this assault, a large number of requests may be sent to battery-powered nodes. These enquiries appear to be legitimate and sent by an intruding party. IoT nodes are powered by tiny batteries due of their compact size. This is the reason that battery-draining intrusions are prevailing, causing damaging effects, such as a power failure. The third sorts of attacks occur when an IoT object does not perform its necessary operations. This may happen because of unnecessary error in the manufacturing stage, lack of sleep, and code inoculation.
- **Physical attack**: Some objects are susceptible to the physical access that may cause the hardware or firmware attacks because the employment of some objects is necessitated in hostile environments within the Internet of Things [6].An assailant is capable to obtain precious cryptographic information, for altering the operating system and for vandalize circuiting, all of which lead to long-term destruction through the physical access to an object.
- **Side channel attack**: The Internet of Things objects execute its normal operations, thus there is a likelihood of disclosure of significant information. It may be executed even in the case, when no wireless protocol is implemented by them for transferring the data.
- **Eavesdropping**: This attack is usually connected with the communication protocols, thus there is a probability of its occurrence at this level, specifically for RFID tags. The fundamental purpose of eavesdropping attack is that the

messages are intercepted, read and modified to perform the further exploration.

- **Tag cloning**: It is a kind of attack which is extremely helpful for the hackers and it could be hazardous also for the reputation of company. An attacker is able to obtain the access of sensitive data and closed areas with the help of duplicating the tags [7].
- **Collision attacks**: These attacks can attack on the link layer. In one scenario, an adversary utilizes an intentionally-generated noise against communication links for generating a collision. There-transmission of packets which are affected due to the collision is required in this collision. An attacker employs this kind of technique to easily drain the battery of an object for which various collisions are generated that leads to various retransmissions.

D. Version Number Attack in IoT

The IoT (Internet of Things) is a new platform that is growing rapidly under the umbrella of universal networks and amenities. This paradigm expands Internet to the objects that occur in the real world, communicate with each other to reach common objectives. The massive expansion of this technology has given rise to the comprehensive deployment of LLN (Low power and Lossy Networks) including home automation systems and WSNs (Wireless Sensor Networks). RPL is a well-known routing protocol for Low power and Lossy Networks (LLN) based on IPv6. RPL-based networks may be exposed to a huge range of security intrusions because of their restricted nature [8]. The version number attack makes use of an RPL feature that is typically used to ensure that there are no loops or errors in the topology. A hostile node modifies the version number of a topology in this attack. As a result, the complete routing tree must be rebuilt. Because parents include the version number in control packets, the established protocol has no mechanism of ensuring that the number is correct. A forced rebuild can result in more effort, less energy, concerns about channel availability, and even routing architectural loops. Several studies suggest that these types of attacks have a significant impact on RPL networks, emphasizing the importance of dealing with them. The RPL protocol refers to a distance-vector routing protocol based on IPv6. The interconnectivity among RPL devices is established on the basis of a definite topology. This topology combines mesh and tree topologies together to form Destination Oriented Directed Acyclic Graphs (DODAG). A root node, also known as data sink of the graph constructs a DODAG graph. A network can function on single or many RPL instances containing multiple DODAG graphs represented by Figure 2. Every RPL instance belongs to an objective function. This objective function is liable to measure the optimal route according to a group of parameters or restraints. For example, this function is capable of minimizing the energy consumption or simply computing the directroute. An RPL node can be linked with many several instances simultaneously, however it can be linked with just single DODAG graph per instance such as nodes 13 and 17 as depicted in the figure. These manifold instances allow the

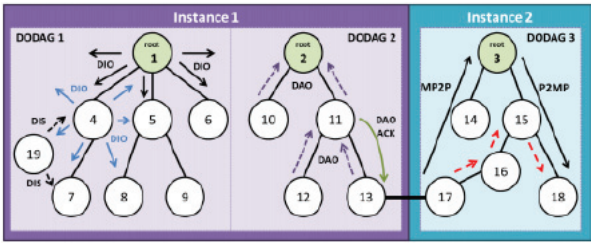

Fig. 2. Example of a RPL network with two instances and three DODAGs

RPL protocol to carry out diverse optimizations just like in QoS (quality-of-service). As shown in figure, in the third DODAG, it is possible to forward RPL packets depending on the three traffic patterns. The first one is multipoint-to-point traffic (MP2P) in which RPL packets are transferred from leaves to the root through upward routes. The next traffic pattern called point-to-multipoint traffic (P2MP) makes use of downward routes to transfer RPL packets from the root to leaves [9]. The red doted arrows denote third point-to-point traffic (P2P) pattern. This traffic pattern uses both up and downward routes for packet forwarding. The RPL network uses version number metric as a sign of a global repair operation. Also, only the root of the DODAG can make changes in this metric for building the topology again. The version number is a vital arena of every DIO ((DODAG Information Object) message. The version is broadcasted unbothered down the DODAG graph. Only, root can increment it, whenever it is required to reconstruct the DODAG, also termed as global repair. An former value specifies that the node has not moved to the novel DODAG graph and it is not possible to use it as a parent node. An intruder can illegally increase this arena of DIO messages by forwarding them to its neighbors to change the version number. A redundant reformation of the entire DODAG graph occurs during this intrusion [10]. This means that the intrusion can trigger many loops, causing data packets to be lost. Furthermore, the graph is modified repeatedly, which adds a lot of unnecessary work to the control messages. This produces network congestion and exhausts the nodes' resources. VeRa (Version Number and Rank Authentication), a security approach, can prevent attacked nodes from impersonating the root and delivering an illegally larger Version Number. Authentication mechanisms based on hash operations are employed in this approach. In that circumstance, a node can easily verify whether the Version Number has been updated by the root node or another rogue node. As a result, someone else cannot take over the DODAG root.

E. Networking Attacks

This section provides an overview of the four main forms of networking assaults. Each network assault can easily be classified into one of these kinds.

- **Remote to User Attacks (R2L)** – "A user sends packets to a system across the internet to expose the machine's

vulnerabilities and exploit privileges that a local user would have on the computer, such as xnsnoop, dictionary, guest, xlock phf, sendmail, and so on".

- **User to Root Attacks (U2R)** – "Attacks on the system begin with a standard user account and attempt to exploit weaknesses in the system to achieve super user privileges, such as xterm and perl".
- **Probing** – "These attacks examine a machine or a networking device for flaws or vulnerabilities that can be exploited later to compromise the system. Portsweep, mscan, nmap, saint, and other data mining tools use this technique".

F. Intrusion Detection Systems

Penetration testing is a method of examining what is happening on within a computer or a network of computers to identify user actions that do not correspond to how the system is supposed to work. Figure 3 depicts how an Intrusion Detection System (IDS) like this one searches network traffic for suspicious patterns. In every secure network, IDSs are employed as the second and ultimate line of protection to halt assaults that have gone beyond the first line of defense.

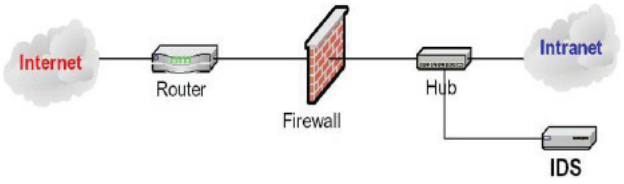

Fig. 3. Intrusion detection System

Firewalls, cryptography, and other network security solutions are not designed to handle network plus application layer threats including DoS and DDoS attacks, worms, viruses, and Trojans. The increased incidence of dangers over the Internet, together with the rapid growth of the Internet, has prompted security officials to consider IDSs. These are the systems that detect network assaults and take countermeasures to prevent them. They are a collection of approaches for detecting suspicious activities on both the network and the host level. Intrusions are identified in a misuse-based intrusion detection system by checking for activities that match known intrusion or vulnerability signatures. An anomaly-based intrusion detection system, on the other hand, looks for unusual network activity to detect intrusions. A patient's breach of the legitimate profile set for regular behaviour or a violation to recognised norms for the events occurring in a connection can be defined as an unexpected traffic pattern. The first phase in an anomalous detection approach is training, which entails generating a normal traffic picture; the second step is abnormality detection, which entails applying the learned profile to actual traffic in order to look for anomalies. To detect such abnormalities, a number of anomalous detection mechanisms have lately been presented, which can be divided into statistical techniques, data-mining methods, and machine learning-based methods. The integration of machine learning and data mining technologies is discussed in this work.

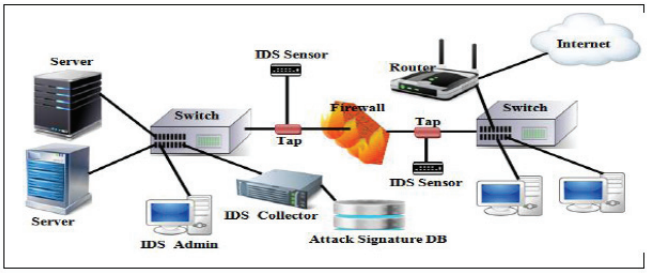

Fig. 4. Signatures – Based IDS

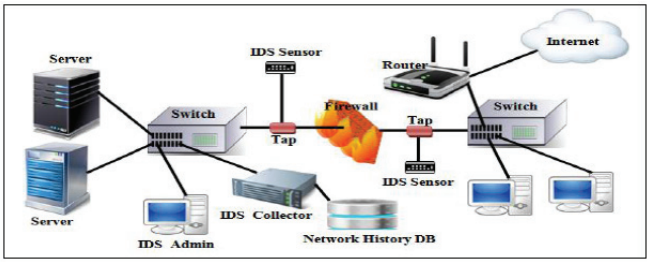

Fig. 5. Anomaly – Based IDS

G. Real-Time Intrusion Detection System

It is currently in development, and includes neural networks, data gathering, deep learning, and other techniques. It has disseminated several intelligent techniques for intrusion detection system application and study. The study's main goal is to lower the likelihood of false alarms and detecting system false alarms, as well as to increase the system's self-learning capabilities. Cyber-attacks on large data have recently increased as existing security solutions have failed to identify them. For various forms of network attacks, there are numerous intrusion detection systems available. The majority of them are unable to detect new unknown threats, and others do not provide a real-time ways to resolve obstacles. The focus of intrusion prevention research is shifting to distributed, high accuracy, fast detection, and intelligent detection. And will contain the ways listed below.

- **Distributed Intrusion Detection**- Distributed intrusion detection systems are typically used in heterogeneous systems and large networks, and they employ collaborative processing, dispersed structure, as well as analysis of a variety of data, as well as a single intrusion detection system architecture compared to greater detection ability.
- **Intelligent Intrusion Detection**- It is currently in development, and includes neural networks, data gathering, deep learning, and other techniques. It has disseminated several intelligent techniques for intrusion detection system application and study. The study's main goal is to lower the likelihood of false alarms and detecting system

false alarms, as well as to increase the system's self-learning capabilities.

- **High-Speed Packet Capture Technology**- High-speed packet capture can help network intrusion detection system, to improve the detection speed and reduce the resource consumption.
- **Efficient Pattern Matching Algorithm**- As intrusions become more complex and varied, complex models need to be stored in the rule base. For that reason, it is important to improve and enhance the pattern matching algorithm.

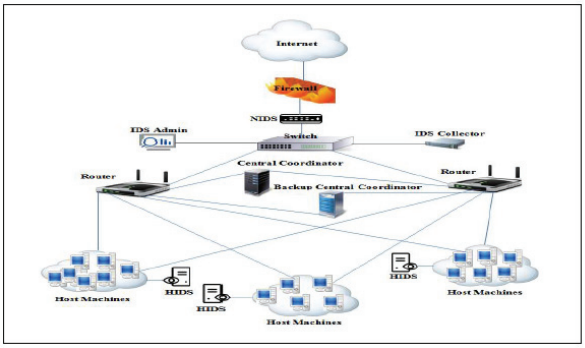

Fig. 6. Framework of a collaborative (IDS) in Big Data environment

- **Apache Spark**- High-speed and large-scale data applications necessitate quick processing and extensive storage space. Hadoop is an excellent tool for delivering scalable distributed storage, despite the fact that it is a distributed system. However, because of the way it processes multi-disk I/O operations, it lacks efficient data processing capabilities. To address this issue, we incorporated Apache Spark. Apache Spark is an open-source clustering framework for fast data processing. Spark provides the functionality to store the data generated in memory because of the different steps in the data transformations. For cache/memory and disc data, Spark works up to 100 times faster and processes data 10 times faster than Hadoop. Spark acts as an interface on a collection of objects termed Resilient Digital Data (RDD) that is partitioned between both the working nodes and may be exploited in parallel, which is one of the most essential notions. The other significant attribute of (RDD) is that it is fault-tolerant and immutable in read-only mode. For calculations, Spark can cache them. Map, reduce, flatmap, filtering, extract, collect, and count are just a few of the ways Spark may help you change the state of (RDDs). Libraries like Spark's own are also beneficial, such as Apache Kafka (Spark SQL, MLlib, and Dataframe). An abstraction of the datasets is provided by these top-level APIs. It is possible to perform relational functions on datasets similar to those in a database using Spark SQL. An (RDD) is converted into a databases, which is similar to a database in the (Spark) framework (RDBMS). Users can query data with Spark SQL in the Spark database.

- **Host Based Intrusion Detection System(HIDS)**- HIDSs look at the operating system, system, and application file contents from one or more host computers [7]. HIDS collects information from sources inside the computer, usually at the level of the operating system (different logs, etc.). It also keeps an eye on what users do and how system programs run. Although chat show IDS is

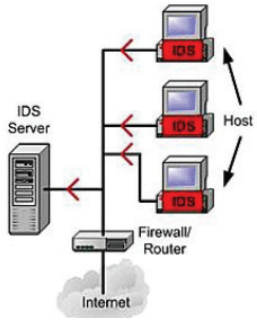

Fig. 7. Host Based IDS

not as powerful as NIDS in general, it does have a few advantages versus network-based IDS". HIDS can gather significantly more precise information about exactly what happens during an attack. Because the granularity of tracing occurrences in the monitored system is increased, recovery after a successful event is usually faster. Due to network heterogeneity and a plethora of computer systems, no single chat show IDS can translate all applications, software platforms, and file systems. Furthermore, with in absence of a corporate secret, no IDS can decrypt encrypted data. Instead of actively monitoring activity, host-based IDS typically rely primarily or entirely on an audit record of activity generated by a system or application. The quality and quantity of this audit record vary greatly between systems and applications, which has a significant impact on IDS efficacy. On the monitored system, host-based intrusion detection systems are implemented. There are thousands of workstations on very large networks. Providing IDS on such a large scale is both costly and difficult to administer.

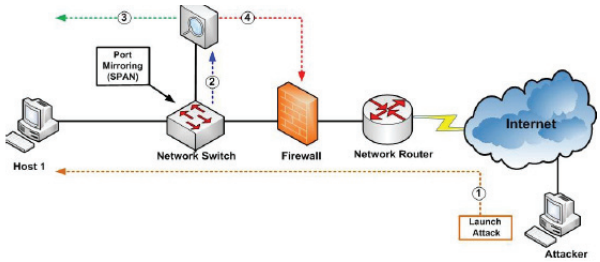

Fig. 8. Network Based IDS

- **Network Based IDS (NIDS)**- It monitors network traffic via connected devices to detect intrusions (e.g. Network Interface Cards, switches and Routers). Its data is mostly

gathered through network streams that move over the network, such as internet packets. Only NIDS can detect all attacks in a LAN, including those that host-based IDS couldn't detect, like DOS. NIDS provide a greater degree of integrity to infrastructure of an organization. They are able to trace user activity from entry point to exit point of attack. They record alteration of data and give report. They also perform analysis of abnormal activity pattern. However there are a few issues that plague these systems: The intrusion prevention system obtains its information from audit logs or packets on the a network. Data must travel a longer path from its source to the IDS, and an attacker might potentially destroy or modify it along the way. In addition, the intrusion prevention system must deduce the system's behaviour from the data acquired, which might lead to misinterpretations or missing occurrences. This is called the fidelity issue. Even when no intrusions are detected, the intrusion prevention system consumes additional resources inside the systems it monitors because the intrusion prevention system's components should function at all times. This is a problem with resource allocation. The intrusion detection system's components are vulnerable to tampering because they are implemented as distinct applications. An intruder has the ability to disable or modify programmes running on a computer, rendering the intrusion detection system worthless or ineffective. This is the issue of dependability.

H. Machine Learning Approaches

Machine learning (ML) is an artificial intelligence (AI) field that researches learning strategies that help computers improve its behaviour based on previous observations. The purpose of these strategies is for them to adapt to their surroundings and learn from their previous experiences, which has drawn researchers from a variety of domains like as computer science, engineering, mathematics, or cognitive science. A learning approach, in general, turns knowledge gained from the world into a new form that can be used later. Classification and prediction are the two most prevalent types of machine learning algorithms. The details of analysis and regression are briefly explained in this section.

1) Classification:

- Learning and classification are two-step processes, according to data mining literature.
- Learning: A classification algorithm analyses training data, and the learnt model or classifier is specified using classification rules.
- Classification: Data are used to assess the classification rules' accuracy. The rules are applied to the categorization of fresh data tuples if the accuracy is deemed acceptable. This step is also called supervised learning because each training tuple has a class label. Supervised learning approaches include support vector machines, neural networks, decision trees, KNN, and genetic algorithms (GAs). This thesis focuses on supervised categorization as one of the most common jobs performed by intelligent systems.
- Like in clustering, a classifier's trust on a given set of tests is determined by the proportion of test set tuples properly classified by the classifier, unlike in clustering, where the class labels of each training tuple are unknown and how many or how many classes to be taught are unknown.

2) Supervised Learning:

- The current decision aids in the judgement process when using a decision tree to guide a subject thru a series of decisions. The decision-making process is represented by a series of trees. Each end leaf node represents a classification group as a specimen has been classified from of the root node to the suitable end leaf node.
- **K-Nearest Neighbors** The most fundamental and extensively used non-parametric strategy for identifying specimens is (k-NN). It assigns the unmarked point to a K-nearest neighbor's house class after estimating the distances among various points using only the input vectors. K is a crucial parameter in the creation of a k-NN classification model, and different values of k will vary in various conducts. The accuracy of the prediction will be impacted if k is large because classifying the neighbours in use for prediction may be time-consuming. Precedent-based learning is what it is known as, and it is distinct from preparatory learning. In order to find and classify input vector instances, it skips the prototype training phase. K-NN "on-line" train services and finds examples as a result.
- **Artificial Neural Networks (ANNs)** - An artificial neural network (ANN) is a computational model based on biological central nervous, comparable to how the real brain processes information. This paradigm relies heavily on the particular structure of a information processing system. An ANN is customised for a given application, including such pattern classification or data classification, through the learning process.
- ANNs are a sort of artificial intelligence that tries to mimic the functioning of the human brain. Instead of employing a digital model in which all computations are done using zeros and ones, a neural network operates by connecting processing elements, which are the computer counterpart of neurons. The output is determined by the arrangement and weights of the connections. Network architecture, setting weights, and activation function are the three essential building pieces of ANN.
- **Support Vector Machines (SVM)** - Before selecting the optimum separating high energy inside the multidimensional space, SVM turns the input vector into the a higher-dimensional feature space. Support vectors also generate a decision boundary, i.e. the separation hyperplane, rather than using the full training set, making it highly resilient against outliers. It was designed with binary categorization in mind. The SVM additionally has a penalty factor criterion that may be set by the

user. It enables users to link the number of misclassified specimens with the breadth of a decision boundary.

- **Genetic Algorithms** - This is an evolutionary technique in which natural selection and evolution are implemented using computers. Adaptive survival in natural organisms is the source of this idea. The algorithm begins by producing a huge number of candidate programmes at random. To analyse the behaviour of each person in a population, some form of fitness metric is used. The fittest chromosomes are then chosen after a huge number of iterations. Recombination for new populations is achieved by crossover and mutation activities.
- The Adaptive GA form of GA is used in this thesis. It calculates the percentage of the population that will be replaced by new people (generation gap). It selects a crossover solution and adjusts the amount of mutations based on the current population condition. The status of the population is assessed by looking at some of its features, such as individual fitness levels, the population average, and so on.
- **Rough Set Approach** - This method can be used to investigate structural relationships in noisy or approximated data. As a result, before using continuous-valued attributes, they must be normalised. The construction of equivalence classes inside the given training data is the basis of rough set theory. All the data tuples in a class label are opaque, meaning that the samples are identical in terms of the data attributes.
- **Fuzzy Logic (FL)** - It really is based on the idea of fuzzy aspects that are common in reality. For inference, FL analyzes the set membership values, which range from 0 to 1. That is, a statement's degree of truth can range from 0 to 1, and it is not limited to those two truth values.

3) Unsupervised Learning:

- It organises items into related physical and abstract groupings. A cluster is a group of data that are comparable to one another but not to data from other clusters. A cluster may be thought of as data compression since it can be handled as a single unit. Clustering methods density-based include partitioning methods, hierarchical methodologies, grid-based methodologies, and techniques, model-based methodologies.
- A Partitioning approach that splits the data into j divisions, with k indicating the number of partitions to be formed. After that, an incremental re motion technique is used to enhance partition by shifting objects from one group to the next. K-means, k-Medoids, CLARANS, and their improvements are a few such examples
- A Density-Based approach that groups objects together depending on their density. It either creates groups based on the number of surrounding objects (as in DBSCAN) and on some posterior distribution (e.g. DENCLUE). Beam is a density-based approach for improving data clustering structure ordering.
- A Model-Based method that generates a model for each cluster and identifies the data that fits that model the best (e.g. EM algorithm).

I. Prediction

Prediction models can make predictions about continuous-valued functions. For example, we can create a model that categorizes bank loan applications as safe or risky, or we can create a model that predicts how much money a potential consumer will spend on computer equipment based on their income or work.

J. Problem Formulation

The Internet of Things is a technology that is decentralized. This network's main concerns are security and routing due to its decentralised structure. Because the network is decentralised in design, hostile devices can infiltrate the network, resulting in a variety of possible attacks. The position protection is indeed an active type of attack that causes the network delay to increase during a location security attack, the rogue node takes the longest route to send data, lowering network performance. The strategy proposed in previous research work is based on threshold delay. Because of the explanations of threshold delay, the detection performance of malicious nodes is quite low. This research suggests a new approach for detecting a rogue node in the shortest period of time.

K. Research Gaps

- The techniques which are proposed to improve security of the data routing can have high latency. A technique needs to be proposed which should be light weighted so as to improve security of the network.
- The routing technique which is already proposed can establish a path from source to destination but version number attack is still possible which affects network performance.
- A novel method needs to be proposed which helps in not only the detection but also isolation of malicious nodes from the network in the least amount of time and increased accuracy.

L. Objectives

- To learn, analyze and implement various secure routing schemes used for detection of malicious devices on the basis of triggering version number attacks in IoT networks and evaluate the effect of the attack on network operations.
- To design energy level calculation techniques for trust-based mechanisms.
- To design a malicious node detection algorithm for multi-path routing techniques.
- To implement the proposed algorithm for the detection of the version number attack in an IoT network and compare it with existing methods in terms of various parameters.

II. RELATED WORK

Samir Ifzarne et al. Present an intelligent identification approach based on machine learning. Using a clustered WSN network design, the model detects in real time if there is an invasion and what type of presence it is. The suggested model ID-GOPA identifies intrusions quickly and efficiently while avoiding resource waste. To reduce overall characteristics and processing burden, it used the gain ratio as just a feature selection. Feature selection is a crucial component that enhances the algorithm's effectiveness when using passive-aggressive methods as just an incremental learning machine. When contrasted to offline models, the simulation results demonstrate a 96 percent accuracy rate, demonstrating that the model is extremely accurate. The model is superior to earlier systems since it may be utilized for any purpose[7]. In this work, Francesco Cauteruccio examines how a WSN's anomalies can be accurately detected using a combination of short and long-term methods. The planned short-term strategy is as follow was successful in identifying 785 potential irregularities locally and highlighting temporal periods of potential importance, which were subsequently For longer-term analysis, the data was uploaded to a cloud service. The long-term approach is beneficial for discovering anomalous temporal periods. Overall, they showed how integrating short and long-term techniques might reduce the disadvantages of both, including such false positives or processing requirements, while increasing the benefits, such as timeliness and accuracy[8]. Xianhao Shen et alresearch .'s focuses on the data anomaly detection challenge in WSN. The CNN model is developed using the characteristics of marked mode and deep neural network structure to identify anomalous data. In the studies, they presented three fresh network models and compared them to a previous cart model, using DA, TPR, and PRE to assess performance. Experiments show that the three models described in this paper outperform the cart model, with the M2 model doing the best[9]. Anton Kanev proposes an ANN-based[10] A method for locating unusual objects in the network of a "smart home" system. This is the first time that a hybrid ANN technique has been applied in research to detect anomalies in "smart home" or building automation systems. The experiment's findings show that the ANN beats typical ML techniques, with a 0.9689 area under the ROC curve[11].

III. ALGORITHMS USED

- **ALGORITHM 1** - This study combined the improved low energy cluster based hierarchy (Improved LEACH) routing protocol with an attacker node detection model based on reputation (MNDREL). MNDREL is a new method targeted at more effectively identifying rogue nodes in wireless sensor networks (WSNs). The upgraded LEACH routing protocol is used to choose cluster-head nodes first. Other WSN nodes then construct separate clusters by selecting cluster-head nodes and determining packet delivery channels. Before transmitting the packet to the sink node, each node adds its base station number or reputation assessment value to it. The node numbers obtained from the sink node's packet parsing are then compared to the source host numbers to create a list of suspect nodes. The ratio of each node's suspect value to its trusted value is then calculated and compared to a specified threshold to identify malicious nodes in the network. In the same scenario, the algorithm suggested in this study is tested alongside HRTM and FMATM. In results, MNDREL model is more effective at detecting faulty nodes while also having a reduced false alarm rate.
- **ALGORITHM 2** - The suggested methodology takes into account harmful behaviour first for testing. The insider attack pseudocode is described in Algorithm 1. The program is separated into two sections; the first section loads libraries along with the sensor dataset. The data is then categorised based on distinct actions when a single sensor is chosen. The data is then prepared and organised per week.
  END Procedure: The data was smoothed in the second section to eliminate false positives. The latency, threshold value, and influence value are the three parameters that make up the threshold algorithm. The sliding window's "lag," the algorithmic signals' "threshold value," and the "effect value" of fresh signals just on standard deviation all range from 0 to 1. The average value (mean) and standard deviation of the data are compared to determine if a signal was positive or negative. The affected value is used, the filter is adjusted, and the dial tone performance improvement is saved in the final step.
  Malicious Distance Calculation: This phase takes the information from the previous stage and processes this to calculate measurements for any potentially dangerous behaviour. When the loops is initialised with length of the dataset, the distance metric takes into account using the "stringdist" package. The function used is sequential length, in which the LV method is selected and the resulting week value is given a +5 increment, with first Five weeks being considered safe. The week value is then compared to a distance value to verify the prevalence of malicious activities. The Threshold algorithm settings are gradually modified and the process repeats if no malicious behaviour is discovered. The distance-measuring procedure is repeated, with threshold method's parameter "lag".
  Benign/Malicious: The final method is for innocuous behaviour and is only required if no harmful activity has already been found. The Threshold algorithm is repeated after the parameters for benign activity have been changed. Following that, the distance measuring algorithms are run. This might be thought of as an improved version of the previous method with different parameters. The Algorithm 3 presents the pseudo code for classifying activities as benign or harmful.
- **ALGORITHM 3** - Static filtering is the second level packet filtering in the 6LoWPAN PFDL (Border Router). All the incoming data traffic from the external world to

IoMT network will be passed through the PFDL Static filtering. In this filtering process, all the data packets (Pi) are filtered based on the source IP index which has already listed in the firewall rule set. While verifying the SrcIP, if the incoming packet is presented in the blacklisted IP then it will be dropped into honeypot. Otherwise, it will be verified for the packet size constraint. If the data size is greater than the data threshold that is 120 bytes per second then the packet will be dropped into honeypot. Finally, the packet will be verified frog GreyList IP verification and forwarded to threshold-based filtering.

- **ALGORITHM 4 Threshold based filtering** - In the threshold-based filtering, all the incoming packets are verifying for the data size in a particular interval of time. If the same data size packets are coming from the same IP for a particular period of time, then it is considered that the suspected Src IP sending packets with intention of DoS attack also the IP will be added into the Blacklist. Also, the same behavior of data is sent from different Src IP with particular interval of time, then it is suspected that the multiple sources are targeting a single targeted medical device with the intention of DDoS attack, finally the SrcIP from various sources are blocked and send to the honeypot.

- **ALGORITHM 5** - Implementing encryption algorithms including such Rivest-Shamir-Adleman (RSA) or Advanced Encryption Standard (AES) algorithms mitigates the assault. Cryptool is used to implement these algorithms, and the performance of loT devices is assessed.
  a.)RSA (Rivest–Shamir–Adleman) - For secure data transport, is a widely used public-key cryptosystem. The ”factoring issue,” which is really the stress of factoring a product of two enormous prime numbers, is the foundation of RSA’s security. The method of breaking Encryption algorithm is known as the RSA problem. It’s not apparent if it’s as difficult as the factorization problem[3]. There have been no known ways to get around the mechanism if a large enough password is used. The Rsa is a lengthy process. As a result, it’s rarely utilised to directly encrypt client data. RSA is commonly used to exchange shared symmetric key cryptography keys, which are then utilised for mass encryption and decryption.
  b.)Advanced Encryption Standard(AES) algorithms - The AES cipher is not a Feistel cipher. It is, instead, an iterative cipher. It uses a technique known as a ”substitution-permutation network.” It is composed of a sequence of linked processes, some of which require changing out outputs for inputs (known as replacements) and others which entail moving bits around (permutations). AES considers 128 bits in a plaintext block to be 16 bytes. Here’s how these eight bits are connected: AES uses all of its calculations in bytes rather than bits. So there are four sections and four rows for matrix processing. The number of bullets in AES can be modified and is determined by the length of the key. This is not true for DES. AES employs ten rounds for 128-bit keys, twelve rounds for 192-bit keys, and fourteen rounds for 256-bit keys. Each of these rounds uses a 128-bit key distinct from the initial AES key.

IV. METHODOLOGY

The proposed model used the CIC KDD NSl[12] dataset from https://www.unb.ca/cic/datasets/nsl.html, which contains two files for training and testing machine learning algorithms: KDDTrain.txt and KDDTest.txt. Finally, we turned the.txt data into a pandas data frame, which contains 42 characteristics containing numerical and textual data. After that, we used the describe() method to compute and display summary statistics for a Python data frame in terms of count, mean, standard deviation, minima, and maxima, which has been giving statistical information. It also works with Pandas series objects and data frame columns. Count the number of different attacks in the data frame. For visualization, data were analyzed using univariate, bivariate, and multivariate features for analytical attacks (normal, dos, probe, r2l, and u2r) on various networking payload protocols such as TCP, UDP, and ICMP. For density, we used a bar graph analysis of the label, protocol, flag, and duration graphs.

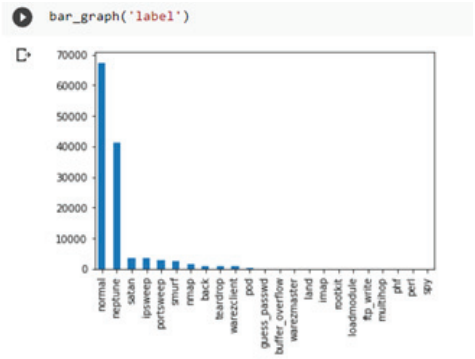

Fig. 9. Bar Plot of Sub Attacks

The multivariate bar graph evaluation for the label is shown in Figure 9. It displays the number of attack labels as well as the value counts for each attack, such as normal, nNeptune devil, ipsweep port sweep, and so on. The bar graph analysis for labels in figure 10 indicates the number of counts for regular, Dos, Probe, R2L, and U2R types of attacks.

Figure 11 shows a study of several numbers for various network payload protocols such as TCP, UDP, and ICMP. Figure 4 shows the frequency duration histogram after binning durations data for various frequencies. The graphic depicted a bar graph for a count of several flags. Counts for SF, S0, REJ, RSTR, RSTO and other flags were displayed. The samples in a dataset, as well as their attack type and methods, were evaluated using bivariate analysis. In addition, the number of

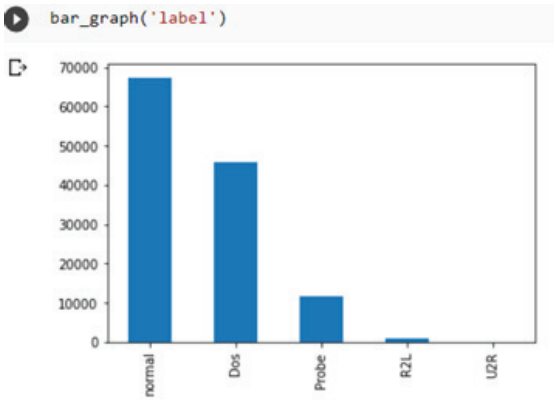

Fig. 10. Bar Plot of 5 Main Attacks

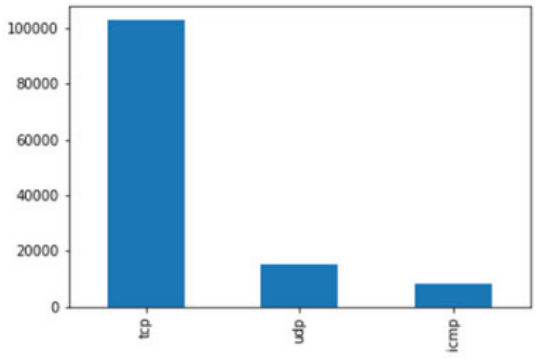

Fig. 11. Plot of Protocols Used

samples in a dataset, as well as the attack class and services. This is the result of the 1323 matrix. We used Label and

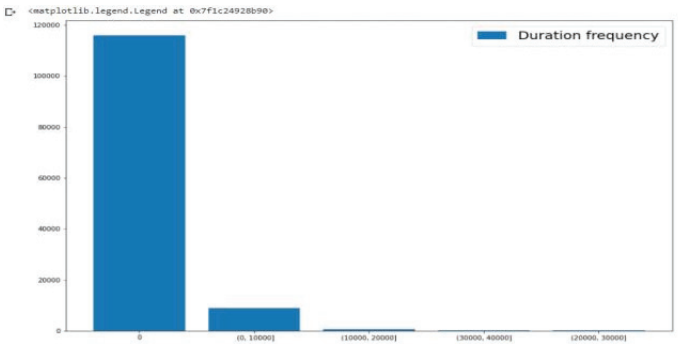

Fig. 12. Frequency of Duration

One-Hot encoding to prepare data for training and testing for binary and multi-class classification.[13] Sk-learn is a powerful technique for transforming category feature levels into numerical values, and it has been used in a variety of assaults. Label-Encoder is a program that Labels with values scale from zero to n classes-1 are encoded, with n indicating the number of unique labels. The encoding of categorical data as binary vectors is one example of such encoding. The categorical variables must first be converted to integers. Then, for each numerical value, a binary vector with in all zero values (save for the integer's position, which is labelled with a 1) is generated.

Feature selection[14], We choose the "intrusion" function after doing one hot and label encoding. Selection is either an automatic or manual method of selecting the attributes that have the greatest influence on the output variables or output that you are interested in. Irrelevant qualities in your data may lead the model to be less accurate and force it to train on unlabeled data, causing the model to fail. The final shapes of X and Y were (125973, 42) and (125973, 42), respectively (125973, 1). The importance of a character performs a correlation[15]. The number of samples in the dataset that used the TCP protocol type and resulted in intrusion, and the number of samples in the dataset that used the TCP protocol type and resulted in the intrusion.[17] type and led to the normal situation.

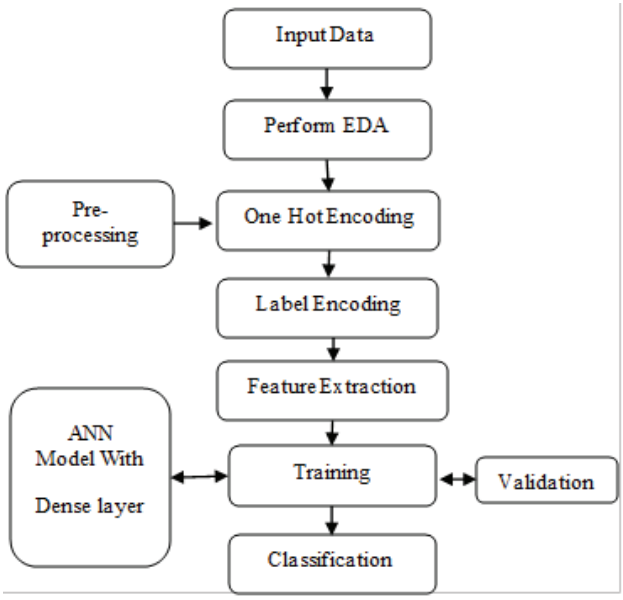

Fig. 13. Flowchart of Proposed Model

To train the model for binary classification, the Deep ANN Classifier is employed. The Dense layer is used to train the sequential model, and the output layer is finished with Softmax activation from the Dense layer. The ADAM optimizer is used to train the model. A CNN-based RS comprises mostly of the process follows, which suggests things for the target user. The data is first entered into the input layer, which then preprocesses the properties of the data. The embedded layer is then utilized to extract features and generate each typical feature vector from the pre-processed data. After the embedding procedure, the full-connection layer executes the program to interconnect attribute features and generate user and item attributes. The independent and dependent attributes are then used to produce the predictions attack.The Deep ANN Classifier is used to train the model for binary classification. The sequential model is trained using 29 neurons in the Dense layer, and the output layer is completed with the Softmax activation function in the Dense layer.

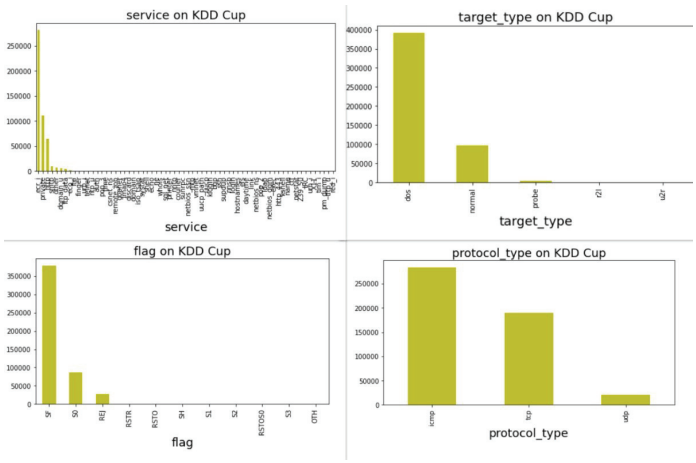

Fig. 14. Services used in KDD Cup dataset, target type (5) or attacks, Flags used in attacks, Protocol used in dataset

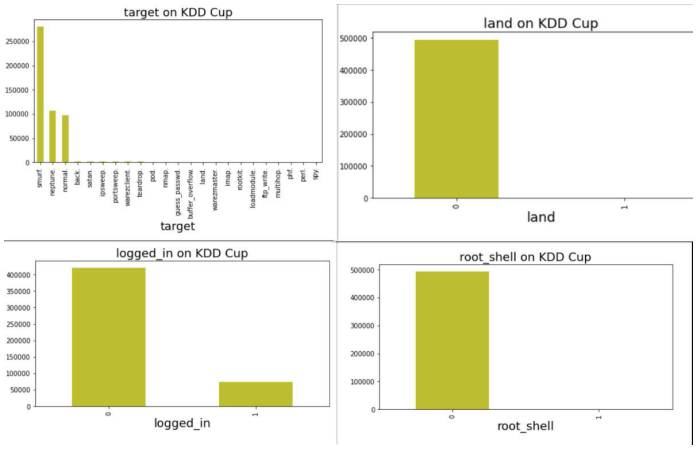

Fig. 15. 13 target type or 13 attacks, Land on KDD cup, Logged in on KDD cup, Root Shell on KDD cup

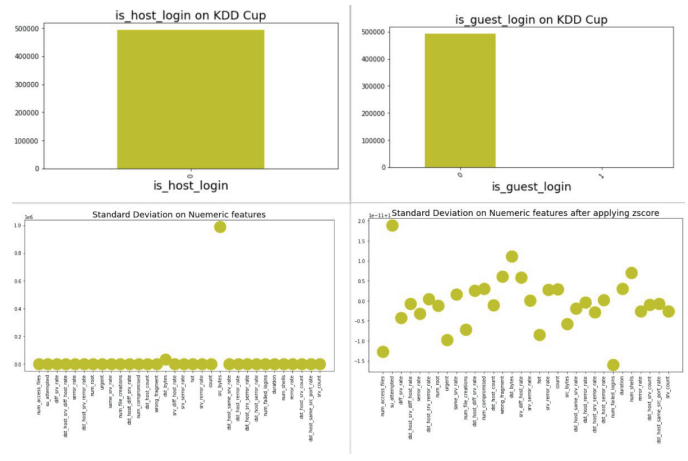

Fig. 16. Host Login on KDD cup, Guest Login on KDD cup, Standard Deviation on Numeric Features, Standard Deviation on Numeric Features after Zscore

Table 1 lists the many parameters that were used in model training.

TABLE I
PARAMETERS USED FOR TRAINING

Model	Sequential
RNN	ANN
Training data	80 percent
Testing data	20 percent
Data shape	(81581,42)
Layers	DENSE with 29 neurons
Output layer	Dense with 5 neuron
Output function	SOFTMAX
Loss function	Sparse categorical cross-entropy
Optimizer	ADAM
Metrics	ACCURACY

Figure 17 displays the accuracy versus epoch graph for the train and test datasets after training the model, as well as the graph plot of loss vs epoch for the train and test datasets.

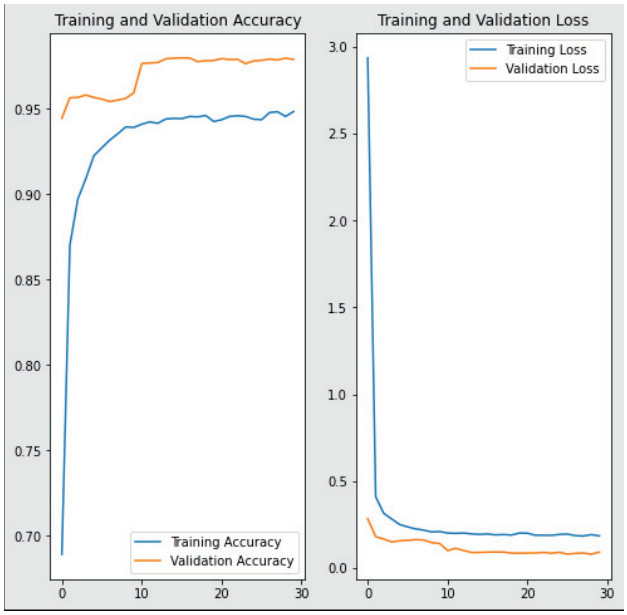

Fig. 17. Accuracy and Loss Graph Of Train and Test Data

As demonstrated in table 3, the suggested ANN model outperforms the other ML techniques. The accuracy of the ANN model is 94.45 percent, with the precision of 96.38, recall of 98.94, and F1 score of 97.64. Our ANN model's results were compared to those of other machine learning models such as[24], Decision Tree, Support Vector Machine[25], etc.

The accuracy, precision, recall, and F-measure are among the performance parameters used to assess the outcome.

A. True Negatives (TN) - Negative values are entirely correct (meaning the real The expected class value is 'not,' whereas the current class value is 'no.' When a class value

reveals that a passenger won't survive and the predicted class verifies it, for example. The genuine class develops when the expected category, a false affirmative, and a negative result all occur in opposition.

B. True Positives (TP) - Both the true class value and the forecasted class value are accurate. Consider the difference between the actual class value, which states that "this passenger has survived," and the predictive class value, which states that "this passenger is likely to be the same one next time."

C. False Negatives (FN) – The answer is yes for real class, but no for predicted class. In other words, the value of each passenger class may indicate whether passengers have survived or are likely to perish.

D. False Positives (FP) – Whenever the actual class is "No," but the expected class is "Yes," this passenger has died. To put it another way, if indeed the class real shows that this individual did not live, but the class prediction says that he did, this passenger has died.

E. Accuracy - It's a basic accuracy metric that's proportional to the total number of measurements. Because the number of false negatives and false positives is nearly equal in symmetrical data-sets, statistical accuracy is improved.

$$Accuracy = TP + TN/TP + FP + FN + TN \quad (1)$$

F. Precision – "It's the proportion of correctly predicted positive observations to total positive observations forecast".

$$Precision = TP/TP + FP \quad (2)$$

G. Recall (Sensitivity)- Yes, it's a ratio of favorable comments that was precisely anticipated based on all of the actual class observations.

$$Recall = TP/TP + FN \quad (3)$$

H. F1 score – "It is a weighted average of recall and precision. As a result, this score accounts for both false negatives and false positives".

TABLE II
MODEL MATRICES AND PERFORMANCE EVALUATION

Model Name	Accuracy	Precision	Recall	F1 Score
Propose ANN	**94.45**	**96.38**	**98.94**	**97.64**
DT	88	95	94	89
SVM	85	85	87	86

V. CONCLUSION

Identification is one of the most difficult aspects of practical applications of wireless sensor networks. Given the ongoing increase in the service area and the significant rise in data volume, the threat and consequences of network assaults in WSN cannot be overlooked. ID systems, for the most part, are only capable of coping with specific types of attacks and are impotent in the face of unanticipated threats. It is challenging to provide security services in the WSN depending on IDS that can reliably identify attacks. In this study, we have presented how incremental machine learning can be used to make a smart ID strategy. The model detects attacks using a clustered approach and then classifies it using the Wireless Sensor Network topology.

REFERENCES

[1] G. Divyashree, A. Durgabhavani, M. Kavya, A. Gudoor, and M. B. Shetty, "Intrusion detection system in wireless sensor network," Int. J. Recent Technol. Eng., vol. 8, no. 1, pp. 2047–2051, 2019.

[2] P. R. Chandre, P. N. Mahalle, and G. R. Shinde, "Machine Learning Based Novel Approach for Intrusion Detection and Prevention System: A Tool Based Verification," Proc. - 2018 IEEE Glob. Conf. Wirel. Comput. Networking, GCWCN 2018, pp. 135–140, 2019, doi: 10.1109/GCWCN.2018.8668618.

[3] L. Sheeba and V. . Meenakshi, "A Brief survey on Intrusion Detection System for WSN," Int. J. Comput. Trends Technol., vol. 40, no. 3, pp. 109–113, 2016, doi: 10.14445/22312803/ijctt-v40p121.

[4] A. H. Farooqi and F. A. Khan, "A survey of intrusion detection systems for wireless sensor networks," Int. J. Ad Hoc Ubiquitous Comput., vol. 9, no. 2, pp. 69–83, 2012, doi: 10.1504/IJAHUC.2012.045549.

[5] S. H. A. H. Baddar, A. Merlo, and M. Migliardi, "Anomaly detection in computer networks: A state-of-the-art review," J. Wirel. Mob. Networks, Ubiquitous Comput. Dependable Appl., vol. 5, no. 4, pp. 29–64, 2014, doi: 10.22667/JOWUA.2014.12.31.029.

[6] I. Butun, S. D. Morgera, and R. Sankar, "A survey of intrusion detection systems in wireless sensor networks," IEEE Commun. Surv. Tutorials, vol. 16, no. 1, pp. 266–282, 2014, doi: 10.1109/SURV.2013.050113.00191.

[7] S. Ifzarne, H. Tabbaa, I. Hafidi, and N. Lamghari, "Anomaly Detection using Machine Learning Techniques in Wireless Sensor Networks," J. Phys. Conf. Ser., vol. 1743, no. 1, 2021, doi: 10.1088/1742-6596/1743/1/012021.

[8] F. Cauteruccio et al., "Short-long term anomaly detection in wireless sensor networks based on machine learning and multi-parameterized edit distance," Inf. Fusion, vol. 52, pp. 13–30, 2019, doi: 10.1016/j.inffus.2018.11.010.

[9] "A Method for Detecting Abnormal Dataof Network Nodes Based on Convolutional Neural Network," pp. 1–12

[10] N. C. Steven Walczak, "Artificial Neural Network," Science Direct, 2003. https://www.sciencedirect.com/topics/engineering/artificial-neural-network/.

[11] A. Kanev et al., "Anomaly detection in wireless sensor network of the 'smart home' system," Conf. Open Innov. Assoc. Fruct, vol. 2017-April, pp. 118–124, 2017, doi: 10.23919/FRUCT.2017.8071301

[12] CIC, "NSL-KDD DATASET," UNB. https://www.unb.ca/cic/datasets/nsl.html

[13] P. Vadapalli, "Label Encoder vs One Hot Encoder in Machine Learning," upGrad, 2021. https://www.upgrad.com/blog/label-encoder-vs-one-hot-encoder/: :text= One hot encoding takes a section which has,The input should always be a 2-D array

[14] Rahul Bajaj, "Feature Selection Techniques in Machine Learning," Geeks for Geeks, 2021. https://www.geeksforgeeks.org/feature-selection-techniques-in-machine-learning/

[15] A. Upadhyay, "What Is Correlation in Machine Learning?," Analytics Vidhya, 2020. https://medium.com/analytics-vidhya/what-is-correlation-4fe0c6fbed47: :text=1 If two variables are closely correlated2C then,understanding of causal relationship 28ifany29 More

[16] V. Lendave, "Gini Impurity vs Information Gain vs Chi-Square – Methods for Decision Tree Split," Analyticsindiamag, 2021. https://analyticsindiamag.com/about/

[17] JayGala, "What is Transmission Control Protocol (TCP)?," Geeks for Geeks, 2021, [Online]. Available:https://www.geeksforgeeks.org/what-is-transmission-control-protocol-tcp/

[18] Zixuan Zhang, "Understand Data Normalization in Machine Learning," towards data science, 2019. https://towardsdatascience.com/understand-data-normalization-in-machine-learning-8ff3062101f0

[19] Manikanth, "What is the use of data standardization and where do we use it in machine learning," Analytics Vidhya, 2021.https://medium.com/analytics-vidhya/what-is-the-use-of-data-standardization-and-where-do-we-use-it-in-machine-learning-97b71a294e24
[20] A. Chugh, "Deep Learning — Introduction to Long Short Term Memory," Geeks for Geeks, 2019. https://www.geeksforgeeks.org/deep-learning-introduction-to-long-short-term-memory/
[21] Aishwarya, "Introduction to Recurrent Neural Network," Geeks for Geeks, 2018. https://www.geeksforgeeks.org/introduction-to-recurrent-neural-network/
[22] Q. Liao, M. Zhu, L. Wu, X. Pan, X. Tang, and Z. Wang, "Deep Learning for Air Quality Forecasts: a Review," Curr. Pollut. Reports, vol. 6, no. 4, pp. 399–409, 2020, doi: 10.1007/s40726- 020-00159-z
[23] S. Jadhav, "What Is ROC Curve?," Analytics Vidhya, 2020.https://medium.com/analytics-vidhya/what-is-roc-curve-1f776103c998.
[24] M. Gupta, "ML — Linear Regression," Geeks for Geeks, 2018. https://www.geeksforgeeks.org/ml-linear-regression/
[25] S. Morris, "Image classification using SVM," Rpubs.Com, 2018, [Online]. Available:https://rpubs.com/Sharon$_1$684/454441
[26] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A Survey" 2010, Computer Networks, vol.54, no.15, pp.2787–2805
[27] J. Gubbi, R. Buyya, S. Marusic, M. Palaniswami, "Internet of things (IoT): A vision, architectural elements, and future directions," Elsevier Future Generation Computer System, Vol. 29, No. 7, pp. 1645–1660, 2013
[28] Mashal, O. Alsaryrah, T.-Y. Chung, C.-Z. Yang, W.-H. Kuo, and D. P. Agrawal, "Choices for interaction with things on Internet and underlying issues," 2015, Ad Hoc Networks, vol. 28, pp. 68–90O. Said and M. Masud, "Towards internet of things: survey and future vision," 2013, International Journal of Computer Networks, vol. 5, no. 1, pp. 1–17