International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 11-No.3 (2025) pp. 5022-5035
http://www.ijcesen.com

ISSN: 2149-9144

Research Article

Intrusion Detection and Prevention Using Machine Learning for IoT-Based

Rajesh[1*], Mridul Chawla[2]

[1]PhD Scholar, ECED, Deenbandhu Chhotu Ram University Science and Technology, Murthal Haryana – 131039, India
***Corresponding Author Email**: 19001903010rajesh@dcrustm.org

[2]Professor, ECED, Deenbandhu Chhotu Ram University Science and Technology, Murthal Haryana – 131039, India
Email: mridulchawla.ece@dcrustm.org

Article Info:

DOI: 10.22399/ijcesen.3323
Received : 22 May 2025
Accepted : 06 July 2025

Keywords

Intrusion detection system (IDS)
Machine Learning (ML)
IOT, Wireless Sensor Network (WSN)
Genetic Algorithm (GA)
Gini Impurity-based Weighted Random Forest (GIWRF)

Abstract:

The intrusion detection system (IDS) is an essential component for enterprises as it safeguards network infrastructure, assets, and confidential data, effectively preventing cybercriminal actions. Various strategies have been devised and put into action in order to prevent malicious activity up to this point. Given the efficacy of machine learning (ML) techniques, the proposed strategy utilized multiple ML techniques for the IDS. The UNSW-NB15 dataset was utilized to conduct an offline analysis of models' performance and to create a tailored integrated classification system for detecting malicious activities in a network. The performance analysis involved training and evaluating the "Decision Tree (DT)", RF, CatBoost, and Hybrid models for a binary classification task. To address the decline in the performance of IDS while using a feature vector with a large number of dimensions, a Gini Impurity-based Weighted Random Forest (GIWRF) model was utilized to choose the most suitable set of features. This approach served as the incorporated choosing features technique. Additionally, feature extraction was performed using the Genetic algorithm (GA). This method utilized Gini impurity in order to enhance the learning algorithm's comprehension of the class distribution. 27 features were chosen from UNSW-NB 15 based on their relevance value. The results of the study showed that the Hybrid model scored better than the other trained models used in the present study. This study offers useful insights on enhancing the security of IoT networks through the utilization of ML. The research also quantified various attacks (DOS, Probe etc.), assessing their detection efficiency using the hybrid model. The findings proved high accuracy in detecting various threats, further confirming the strength of the proposed method. Study highlights the significance of customized strategies and continuous improvements in increasing the resilience of systems to constantly changing cyber-attacks. In addition, the GIWRF-Hybrid method proposed in the paper showed better performance than other methods considered in the paper, that is, accuracy and loss.

1. Introduction

Network security is currently a significant issue in need of solutions due to fast development in communication technologies and Internet services and the growing network application base. Several defence systems leveraging intrusion detection systems (IDSs) and firewalls and authentication methods and cryptography shield networks today [1]. Network traffic examination by IDS systems seeks out abnormal behavior as well as malicious digital attacks [2]. IDS system alerts network administrators in the form of notifications regarding suspicious activity within the network. Devices employed suitable counter-measures to halt ongoing attacks as well as to forestall future cyber-attacks [3]. Machine Learning methods have been found to be extremely efficient IDS building methodologies in the recent past. An assembly of scientific methodologies called ML supports numerical pattern identification and independent analysis to gain important insights from records of data [4]. ML prediction accuracy increases sharply with total relevant data acquisition. Two primary algorithm groups of ML network are supervised and unsupervised algorithms [5]. The supervised machine learning methods KNN algorithm [6] and DT-based models [7] "Deep Learning" techniques [8] and other algorithms employ classified data to

train and construct input variable to output variable mappings [7]. The detection of patterns in unlabeled data utilizes algorithms such as k-means and "Gaussian Mixture Model" (GMM) and isolation forest and other such methods. The evolution of signature-based IDSs most commonly employs SL algorithms in their creation. The algorithms require labelled structure datasets to carry out their training processes. Anomaly-based IDSs can be evolved through UL techniques for their execution. These IDSs possess characteristics that isolate unusual data from original data samples.

The IDS operates as a watch system by scanning your network time and again to detect invasion attempts and prevent them. IDS conducts in-depth analysis of network requests prior to determining them as perilous or harmless entities by utilizing signature-based examination and protocol analysis in addition to statistical packet analysis methods. The system guards itself against numerous threats by detecting "Distributed Denial of Service (DdoS)" attacks because harmless requests get free passage but any dubious signals prompt an immediate warning response.A Network IDS possesses two primary detection mechanisms: signature analysis and anomaly detection. Network threat detection employs two approaches where signature-based searches are familiar hostile patterns but anomaly-based scans differences from normal behaviour patterns [12]. Network signature detection systems authenticate threats through established attack signature identification previously. The detection system is effective against attacks whose characteristics are already known through established patterns. The detection systems cannot stop newly evolved attacks because they are unable to learn from observations of past attacks based on [13]. Detection systems based on anomaly monitoring try to identify any deviation from regular behaviors or patterns in an effort to identify threats. These detection systems have the capability to detect unknown attacks by using approved models that define ordinary patterns of action [14].Although NIDS performance has been consistently improving, there is still potential for future enhancement. This is especially apparent due to the substantial quantity of network traffic data created, the constantly changing surroundings, the extensive collection of characteristics that make up the training, and the need for intrusion detection in real time [15]. For instance, by slowing down the model training process, duplicate or unnecessary characteristics might negatively affect NIDS's ability to identify threats. The best collection of features should be selected, and the machine learning (ML)-based detection models' parameters should be optimized, in order to improve the models' performance [16].

Several data-centric and algorithmic techniques were used in the studies to produce a lightweight, quick, high-performance classifier free of accuracy compromise.

1. The current methodology and existing research for deploying IDS in network traffic and general domains.
2. A comparative analysis is conducted among the XYZ variations, dataset to provide a comprehensive understanding of the attack kinds, dataset size, sample count, and their importance.
3. The methods used for training, verifying, and testing the algorithms on the datasets included oversampling, extraction using genetic algorithm, and selection using GIWRF and Boost LSTM.
4. Different ML algorithms such as Knowledge Base are created for detailed mode information, transfer learning, Behavioural patterns, and recorded patterns, providing a thorough assessment of their effectiveness.
5. A complete explanation of the benefits, drawbacks, and performance of the suggested method.

Investigations on the application of ML for ID in IoT-based WSNs make a substantial contribution to improving system resilience and security. This project seeks to utilize ML methods, namely anomaly identification and categorization models, to identify and address different forms of intrusions in real time. The objective is to protect confidential data and maintain the uninterrupted functioning of IoT devices in the network. By using ML, it becomes possible to proactively detect suspicious actions and deviations from typical conduct patterns. This, in turn, provides strong defense mechanisms against constantly changing cyber threats in WSNs. This contribution enhances the safety of connected devices and promotes the development of adaptive and intelligent intrusion detection systems specifically designed for the unique problems presented by IoT-based WSN scenarios.The study is split into the following parts according to many approaches: The following sections are organized as follows: The "Related works" section examines prior research connected to the current investigation, the "Methodology" part outlines the suggested approach, the "Experimental setup and analysis" section provides a detailed description of the setup of the experiment, the "Result and analysis" section presents the findings and discourse of ML techniques, and finally, the "Conclusion" section summarizes the paper.

2. Literature Review

Alhayali et al., [17] suggested a better ID strategy for binary categorization. Additionally, in a hybrid

approach combining the Rao-SVM algorithm with supervised ML techniques for "feature subset selection (FSS)", the author incorporated several optimizers, including the "Rao optimization" (RO) technique, LR, SVM, and ELM. Abd, Alsajri, and Ibraheem [18] supervised ML techniques for FSS combined with the newly developed RO method, IDS, SVM, ELM, and LR methods. The Rao-SVM FSS system is presented in this paper along with an analysis of its parameter-free and algorithm-specific model. In [19], a Wireless Sensor Network (WSN) intelligent IDS was developed. This system used the KNN algorithm in ML and included the "Arithmetical Optimization Algorithm" (AOA) from evolutionary computation. The purpose of this system was to create an intelligent structure that could effectively detect and respond to Denial of Service (DoS) assaults in the WSN.

Liu, Yang, and Wu [20] 5024unction a feature analysis and SVM-optimized integrated web intrusion detection system. Using their expertise, experts analyze the characteristics of frequent online assaults. The examination of the HTTP protocol selects the relevant data properties. AI-Janabi and Ismail [21] developed a method that integrated the SVM, NTLBO, ELM, and LR algorithms using supervised machine learning techniques for feature subset selection (FSS). In [22], Suggested a method aimed at optimizing the performance of NIDSs. The method used wrapper-based techniques in conjunction with the GA, FFA, PSO, and GWO methods to choose features using the Anaconda Python Open Source platform. Furthermore, the GA, GWO, FFA, and PSO algorithms were utilized to compute the "Mutual information" (MI) through filtering-based methodologies.

Bhattacharya et al., [23] created a mixture of ML approach for IDS dataset classification using PCA and fireflies. Initially, the process transforms IDS datasets using One-Hot encoding. In order to classify the reduced data, the XG-Boost algorithm was used. In, [24] introduces a novel hybrid intelligent system that utilizes an inverted hourglass-based encrusted network classifier to perform feature classification tasks. This technique is skillfully implemented over 3 datasets to distinguish between older and new assault behaviors. It utilizes a hybrid optimization strategy to choose the most important characteristics for categorization, giving them more priority. In addition, the model utilizes an up-sampled layered network architecture to improve the training process, hence increasing its ability to identify and counter-infiltration attempts. Nazir, and Khan [25] present a brand-new FS method for Network IDS called "Tabu Search Random Forest (TS-RF)". This approach uses Rfas the learning algorithm and tabu search as the search mechanism.

A state-of-the-art IDS was introduced in 2022. It merged the X2 statistical model with the "Bi-Directional Long Short-Term Memory" (Bi-LSTM) structure. The system was built and evaluated based on the NSL-KDD dataset. The proposed model achieved a remarkable accuracy of 95.62% [26]. A cutting-edge IDS built on DNNs was released in 2022. Cross-correlation functioned as the feature extraction mechanism to create stable features from the data. Experimental trials revealed that the suggested model performed well thereby suggesting its possible efficacy in network attack detection [27]. Future-generation hybrid DL architecture surfaced in 2021 to offer successful classification of malicious cyber-attacks. The implemented framework took the CRNN architecture by merging CNNs for local features and RNNs to handle sequential aspects of data. Tests conducted through CSE-CIC-DS2018 dataset showcased the outstanding performance capabilities of the model. The proposed method attained precision scores of up to 97.75% during 10 cross-validation tests that exemplified its performance in detecting as well as classifying cyber threats [28]. The deployment of an intrusion detection system with ANNs went operational in 2021 for the identification of normal and abnormal network traffic.

The processing of the DS2oS dataset employed the Flower Pollination approach (FPA). Based on the evaluation of the model the accuracy rate achieved was 99.1% [29]. 2021 introduced a CNN-based technique for intrusion detection systems. The NSL-KDD dataset had records of DoS, Network Probe, U2R and R2L attack profiles. The study utilized Spider Monkey Optimization (SMO) and Conditional Random Field (CRF) methods as per the research results [30]. Studies conducted in 2020 suggested the use of ANN models for the detection of both abnormal and normal security intrusions. The authors utilized the correlation-based Feature Selection (CFS) method for the NSL-KDD datasets to attain this achievement [31]. Year 2019 presented a neural network-based model that differentiated between normal and abnormal intrusions. The researcher utilized information enhancement methods to examine the NSL-KDD database system. [32] Utilization of DNN for IDS development caught the attention of researchers in the year 2019. The study utilized NSL-KDD dataset to identify abnormal and normal security violations [33]. A model of deep learning began its emergence in 2018 to identify abnormal behavior and regular cyber incidents [34].

Disha et al., [35] developed a feature ranking algorithm based on Gini impurities using RF in order to assess the categorization performance of NIDS based on the TON-IoT dataset. Even though

classification effectiveness was prioritized, computational expenses related to feature reduction were not adequately considered. It should be noted that some existing datasets used to evaluate classification methods in NIDS for IoT security are outdated, highlighting the need for more up-to-date benchmark datasets in this area.

Moreover, wrapper-based feature selection is often used in research to identify optimal feature subsets that enhance arrangement performance. Shafiq et al., [36] presented a wrapper-based FS algorithm and a feature selection approach dubbed CorrAUC, both of which use the "Area Under the Curve" (AUC) metric to pick useful features for ML algorithms. Although the method's accuracy was lower for specific assaults, such as keylogging attacks, it successfully picked relevant characteristics when evaluated on the Bot-IoT dataset [37] using 4 algorithms. In addition, numerous studies have focused on creating thin devices to meet the special needs of IoT networks. Liu et al. [38] introduced a method that combines one-class SVM [39] with "Particle Swarm Optimization" (PSO) to identify attacks. The strategy utilizes light GBM to construct efficient models and optimizes PSO for selecting relevant features. Though these feature selection tactics might improve efficiency, they often need a lot of computing power, especially when using approaches like Genetic Algorithms (GA), PSO, or classifiers based on machine learning. Internet of Things (IoT) networks and systems with limited resources can face difficulties due to this processing cost.

Moustafa et al., [40] asserted an ensemble intrusion detection method that trained using, ANN, DT, and NB as its foundational to extract the most useful information from statistical flow characteristics, while Leevy et al., [41] IG, "Chi-squared (Chi2)", and "Information Gain Ratio" FR method were utilized for feature selection, prioritizing enhanced performance metrics. However, the computational cost was not a primary consideration in this pursuit of improved performance. Gavel et al. [42] The AWID dataset for WSN was analyzed using ant lion refinement to choose features, with a correlation-based fitness estimate being used. Zhou et al., [43] The procedure included improving feature selection to raise NIDS accuracy. This was achieved by removing superfluous traits and focusing on the most informative ones, using a correlation threshold as a reference. Although there were significant advances in accuracy using this method, the system became more complicated. Aggarwal [44] explored using a random forest classifier in conjunction with a "Grey-level Co-occurrence Matrix" (GLCM) data extractor to classify MRI images of brain tumors. Results underlined the possibilities of GLCM features, especially when improved, to provide remarkable accuracy by efficiently collecting important texture components within the pictures.

3. Research Methodology

The proposed research methodology for the "Intelligent Framework for Intrusion Detection and Prevention using Optimized Machine Learning" begins with collecting network traffic-based datasets, which serve as the foundation for subsequent analysis.

Data Pre-Processing

Within a specific dataset, it modifies the ranges of the data to improve information processing. Normalization helps alleviate algorithmic challenges, when there's a wide contrast between maximum and minimum values. This normalization is particularly effective in neural networks for classification tasks. Additionally, when employing back-propagation in neural networks, proper input normalization enhances efficiency and accelerates training speed.

Normalization

Information scaling is a crucial component of the procedure for normalization. It involves applying a max and a min. Method to change the data values within the range of [-1, 1] or [0, 1]. The expression below provides the standardizing formula,

$$I = \frac{d - d_{MIN}}{d_{MAX} - d_{MIN}} \quad (1)$$

According to Equation (1), the term I represents the converted input value, which means it is a balanced value. Additionally, the character "d" represents the real value. "dMAX" and "dMIN" refer to the highest and lowest values of the input variable "d", respectively.

Data Reduction

Redundant information, noises, oversights, and undesired data in the dataset are eliminated by the implementation of a data-reducing technique. This procedure allows only the pertinent data to be processed further.

Feature Extraction

IDS success rates are open to various factors utilized in respective environments. Data quality of intrusion detection depends on both its representation method and accuracy of information used in the process.

Utilization of "Genetic Algorithm" (GA) in intrusion detection is aimed at optimizing feature extraction from network traffic data for the sake of IDS efficiency and accuracy. A binary string code represents possible features accessible to candidate solutions at this stage. Evolutionary process in GA framework comprises solution evaluation, followed by selection and then crossover and mutation processes. Fitness function-based evaluation indicates the ability of each solution that is generated in identifying regular traffic from malicious network activity. The termination condition specifies a finish to the iterative optimization process which yields optimal sets of features that improve network threat detection and mitigation performance of the intrusion detection system.

The fitness 5026unctionn f(x) gives a measurement for assessment for potential solution x with regard to specific task achievement. In maximization problems the fitness function gains values for better solutions but it loses values for better solutions in minimization problems.

$$P(x) = \frac{f(x)}{\sum_i (x_i)} \quad (2)$$

Feature Selection Using GIWRF

The "Random Forest" (RF) is a classifier that combines numerous DT and offers different methods to determine the relevance of features. One method involves calculating the significance score by training the classifier. Traditional ML methods disregard possible class disparities by assuming equal significance for every category in the initial training data. In order to tackle this issue, RF utilizes a weight modification mechanism following the calculation of the GI, represented as i(τ), by the classifier. GI measures the degree to which a split successfully separates the total samples of binary classes inside a particular node. Theoretically, it may be expressed as:

$$i(\tau) = 1 - p_p^2 - p_n^2 \quad (3)$$

where $_p$ is the percentage of favorable instances and pn is the fraction of unfavorable tests out of all samples (N) at node τ. The decrease in GI obtained from any most effective split $\Delta(\tau,M)$ is acquired collectively for all the nodes τ in the M quantity of calculated trees in the forest, separately for all of the features.

System Training

System training involves employing a stable ensemble classification approach, incorporating Cat Boost and "Long Short-Term Memory" (LSTM) networks. This guarantees the durability and flexibility of the IDS.

1. **Classification of nodes**

The trained model is utilized to categorize network nodes according to their individual identifiers, e.g., ID and port number, to allow for the identification of potential threats within the network. Through the observation of node information, the system can recognize doubtful or irregular activity involving certain network entities. Proactive threat detection and countermeasures can be applied through this classification, improving the security of the network by identifying potential weaknesses and malicious activity at the node level.

2. The knowledge base is created for detailed mode information.

A rich knowledge base is created, holding extensive information on multiple modes of intrusion, attack types, and suspicious behaviours. This library includes recognized threat signatures, attack vectors, and past history regarding successful and failed intrusions. Through updating its knowledge base with current data, the system improves its capability to detect and prevent advanced attacks, providing proactive cyber-security protection in changing environments.

Figure 1. *Proposed Framework for intrusion detection using knowledge base and transfer learning.*

3. Prevention Mechanism Integration

The system combines preventive measures by utilizing analyzed intrusion data to pre-emptively prevent possible threats from happening. The system evaluates previous intrusion data to identify recurring patterns of attacks which in turn enables it

to take preventive measures that prevent future threats. Identifying known attack vectors in advance, networks receive improved security since those vectors are automatically rejected or prevented from happening before, they can cause harm.

4. Attach prevention based on recorded patterns. Deployment of prevention strategies is based on records of witnessed intrusion patterns from intrusion detection systems. Analysis of past intrusion data by the system allows for the determination of routine patterns of attacks which give rise to preventive security actions that halt predictable future occurrences. Security systems become stronger by predictive actions that prevent or cancel previously recorded means of attacks thus averting successful intrusions and their destructive effects.

5. Behavioral Patterns Generation with LSTM

By employing LSTM networks, operators generate behavioral patterns that help identify abnormal system events as well as possible intrusions. Since they can monitor long time series of network activities, LSTM networks can identify abnormal patterns that indicate possible security risks. Dynamic threat detection is achievable through this method since it employs context-based techniques to identify unusual network patterns that subsequently trigger immediate responses towards new security threats.

6. Hybrid model

The study deploys CatBoost as a gradient boosting algorithm that optimizes operations on categorical data using LSTM as a deep network for the management of sequential data. CatBoost effectively handles structured network traffic features while managing overfitting and delivering great querying ability. Using LSTM, the model is able to recognize long-range attack patterns in network traffic thereby enhancing its capability to recognize complex intrusions. The execution of these two analysis techniques provides enhanced intrusion detection accuracy since they allow structured learning and in-depth feature extraction from the data.

The best of both CatBoost and LSTM integrate perfectly into intrusion detection since they complement each other in this context. The ability of ML models to process sequential data is still limited but DL models require optimal feature choice to form appropriate generalization capabilities. CatBoost provides efficient processing of both categorical and numerical inputs which reduces discrimination-based errors and remains lucid while LSTM identifies temporal attack patterns. Integration of the methods through hybridization improves the entire system to be more accurate and stable than executing individual models separately.

4. Dataset

The experimental procedure for the IDS involved using the UNSW-NB 15 data for offline evaluation [45]. The UNSW- NB15 [46] dataset is a highly utilized dataset in IDS research. Table 1 of the UNSW-NB15 dataset has a total of 27 stated features. The UNSW-NB 15 dataset is relatively more recent than other notable datasets.

Table 1. *Features of the UNSW-NB15 dataset*.

Features	Value	Section feature
dbytes	int	primary
rate	int	content
sttl	int	primary
dmean	int	content
ct_state_ttl	int	general
dload	float	primary
sloss	int	primary
sinpkt	float	time
dinpkt	float	time
dur	nominal	primary
ct_dst_sport_ltm	int	connection
sbytes	int	primary
synack	float	time
dpkts	int	primary
ackdat	float	time
smean	int	connection
swin	int	content
tcprtt	float	time
ct_src_dport_ltm	int	connection
state_INT	nominal	primary
ct_srv_dst	int	connection
proto_tcp	nominal	flow
ct_srv_src	int	connection
dttl	int	primary
ct_dst_ltm	int	connection
ct_dst_src_ltm	int	connection
sload	int	primary

The study utilized 70% of the dataset for training the models, 15% for validation, and the remaining 15% for testing. A verification approach was done to ascertain the attainment of optimal results in the training process. The dataset comprises contemporary internet traffic data, encompassing both typical and abnormal cases, including current low-profile attacks. The data is presented in a clean style without any unnecessary repetition, making it highly ideal for accurate evaluation in IDS.

5. Results

The experiment was carried out using an HP Notebook 14-AL143TX laptop, which was running the most recent version of the Windows Operating System and had been fitted with the following processor: The CPU is an Intel Core(TM) i5-7200U

operating at a base incidence of 2.8 GHz and an elevated turbo frequency of 3.5 GHz. The ML models were constructed, instructed, and assessed utilizing Pandas, Scikit-Learn (sklearn), and additional ML components within the Python environment of Jupyter Notebook, which is a freely available program.

Evaluation Parameters

The efficacy of the proposed methodology was assessed utilizing measures such as accuracy, precision, and loss. The accuracy is calculated by dividing the number of correctly identified remarks and data from the IDS by the total number of assessments in the data set, as indicated by the following equation:

$$Accuracy = \frac{TN+TP}{TN+TP+FN+FP} \quad (4)$$

$$Precision = \frac{TN}{TN+FP} \quad (5)$$

TP refers to the count of accurately identified attacks, while TN shows the count of accurately categorized normal traffic. FP refers to the count of traffic instances that are mistakenly labeled as attacks, even though they are essentially normal data. FN represents the count of attacks that have been incorrectly identified as regular network traffic [47].

6. Experimental Results

This section presents a description of the results obtained from the binary classifications accomplished by the IDS developed using ML techniques. Moreover, this study assessed the precision rates attained in the UNSW-NB15 datasets in comparison to previous research. Ultimately, this study determined the precise rate at which different types of assaults were accurately identified in the utilized data sets.

The experiment was carried out in two states to assess the performance of four machine learning models: “Random Forest” (RF), DT, CatBoost, and Hybrid. In the preliminary stage, we utilized all the attributes of the UNSW-NB 15 dataset and evaluated the efficacy of machine learning models in detecting binary data. During the second stage, the recommended method of selecting features is employed to assess four separate models according to their accuracy and loss. The accuracy readings of RF Model's capacity to identify WSN faults are shown in Figure 2 within 10 epochs of functioning. The measurements of precision illustrate varying trends within the learning process as per the graph.

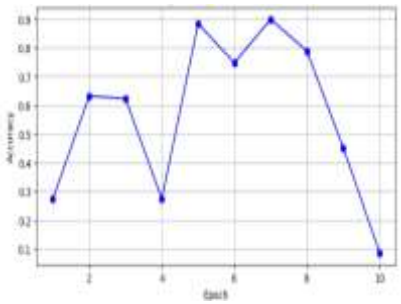

Figure 2. *Relative examination of the accuracy of RF model.*

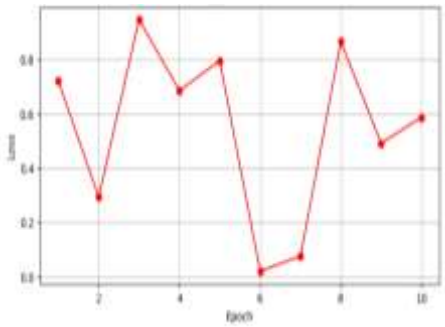

Figure 3. *Loss over the epoch of the RF model.*

The model begins its precision readings at 0.3 within the initial epoch before peaking at 0.9 within epoch five. The evidence suggests volatility with a decrease in the sixth period before increasing to an increase in the eighth period leading to steep decline to roughly 0.1 by period 10. Instability of training accuracy throughout the learning phase probably indicates both overfitting issues in the model and likely adjustments needed to hyperparameters for improving learning process management.

The same model showed its loss values over 10 training epochs in Figure 3. The loss begins at around 0.7 initially. The model decreases its loss value to 0.1 for epoch two and then it suddenly increases to 0.9 in the third epoch. Such huge fluctuations in the loss measures indicate severe uncertainty in the learning operations of the model. The value of loss underwent radical transformations between consecutive epochs in order to reach its minimum in the sixth epoch afterward began to rise. The disorganized trend in loss demonstrates the model has difficulty maintaining stability possibly connected with rate learning and preparation issues concerning data. By enhancing these parameters researchers would be able to obtain more stable outcomes along with lower loss performance in every test cycle. Figure 4. illustrates the accuracy of the second model (DT) over epochs for recognizing and avoiding incidents in an IoT-based WSN. The figure presumably depicts the fluctuation of the model's accuracy over the course of the epochs. Normally, we anticipate observing initial improvements in accuracy as

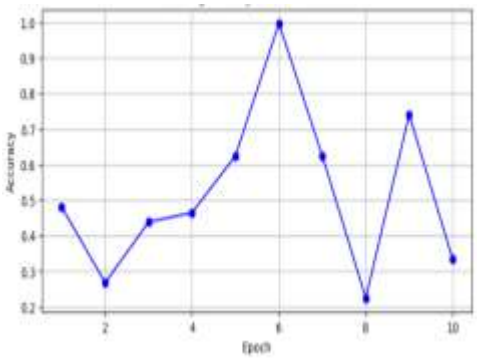

Figure 4. *Accuracy over epoch for model 2 (DT)*

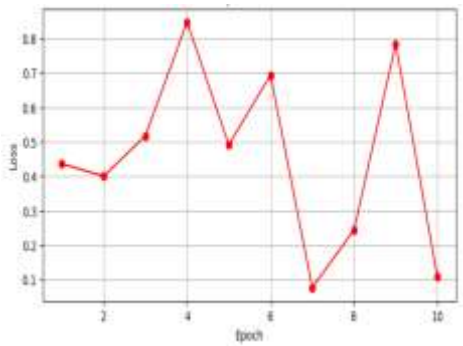

Figure 5. *Loss over Epoch of DT model*

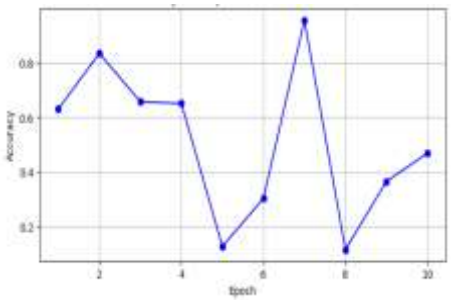

Figure 6. *Accuracy over epoch for CatBoost mod*

the model gains knowledge from the information, followed by possible periods of stability or variations as it refines its results. Differences in accuracy may suggest the model's learning process and identify areas where enhancements in training or data preprocessing may be necessary to enhance and stabilize accuracy.

The performance of DT Model loss is displayed in Figure 5 over its 10 epochs. After Epoch 0.45 the loss starts at 0.45 until it shows around three consecutive epochs of stability. The system shows dramatic growth during epoch four where loss levels reach the peak of 0.8 before decreasing. There appears a ninth epoch loss peak after loss starts decreasing between consecutive epochs. Loss value observations show the DT model experiences learning instabilities that lead to large variations in output loss metrics. This observation indicates evidence of overfitting problems and over-identification of patterns in the data set. The model requires new parameter configurations as well as data pre-treatment operations that must be optimized to obtain stable loss values during training iterations.

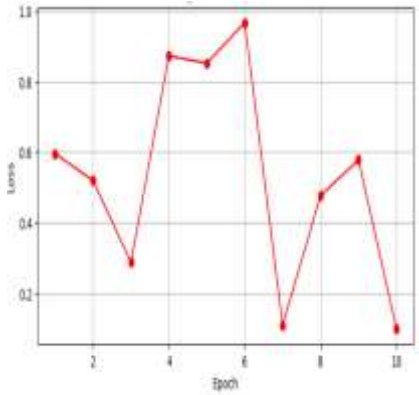

Figure 7. *Loss of Catboost model over epoch*

Figure 6 demonstrates the accuracy of the CB Model. The line graph displays varying levels of accuracy during the learning process. The accuracy starts at roughly 0.6 in the first epoch and reaches its highest point at around 0.9 by the seventh epoch. Subsequently, there is a decrease, indicating a certain level of instability, and then it rises once more around the eighth epoch before sharply declining to approximately 0.1 by the tenth epoch.

Figure 7 indicates the CatBoost model's loss. The loss initiates at approximately 0.6 and gradually diminishes to approximately 0.3 by the 3rd epoch, signifying the initial advancement in learning. Nevertheless, there is an unexpected rise in the magnitude of loss, nearly reaching 1.0 by the fourth epoch. The pattern persists with alternating crests and troughs, indicating another notable surge during the sixth epoch and a rapid decrease by the next epoch. The loss exhibits substantial variability during the training phase, suggesting instability in the model's learning and highlighting the necessity for more refinement to attain consistent and dependable performance.

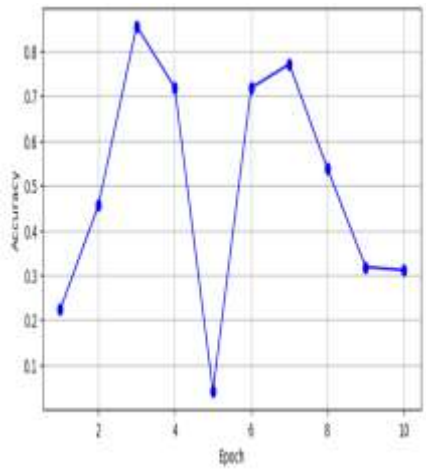

Figure 8. *Accuracy of model 4 (Hybrid model)*

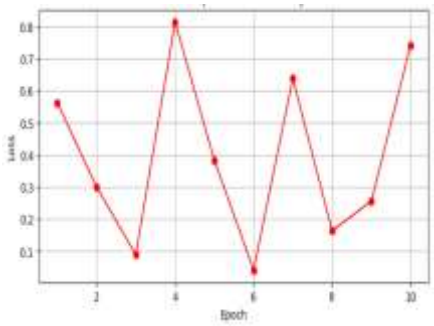

Figure 9. *Loss over epoch for Hybrid model*

Figure 8 displays the accuracy of the hybrid model across epochs. The model initially achieves an accuracy of approximately 0.2, reaches its highest point at around 0.85 during the third epoch, and thereafter plummets to zero during the fifth epoch, suggesting a significant deterioration in performance. The accuracy fluctuates, exhibiting both peaks and troughs over different epochs. The model attains its best accuracy during the third and seventh epochs. However, the overall pattern indicates a lack of stability in the model's capacity to accurately detect intrusions. The lack of consistency in accuracy, together with the variations in loss, emphasizes the necessity for additional improvement of the model's structure and training variables.

Figure 9 exhibits the changes in the loss of a hybrid model that was trained to identify and avoid intrusions in WSNs based on the IoT. The graph displays substantial variations in loss values, suggesting instability in the model's learning process. The initial loss is around 0.5, but it decreases to around 0.1 by the third epoch, indicating considerable progress. Nevertheless, there is a noticeable increase in loss during the fourth epoch, peaking at approximately 0.98, which is then followed by a subsequent fluctuation in a downward and upward trend.

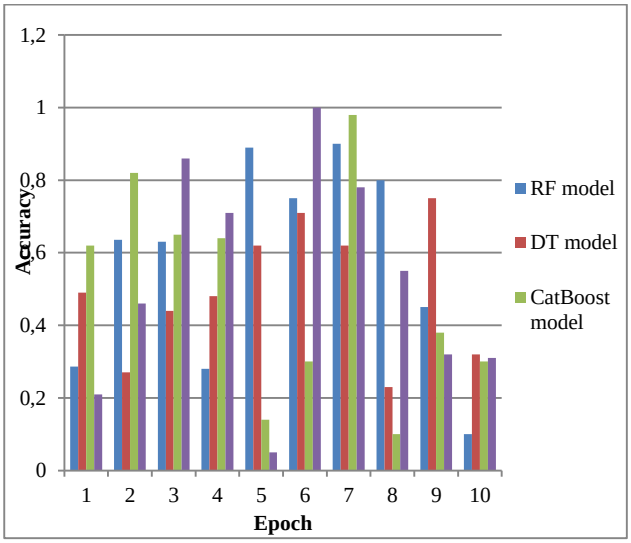

Figure 10. *Comparative research on the accuracy of several models*

Table 2. *Relative examination of the accuracy and Loss of the four models*.

Epoch	Accuracy				Loss			
	RF model	DT model	CatBoost model	Hybrid model	RF model	DT model	CatBoost model	Hybrid model
1	0.286	0.49	0.62	0.21	0.75	0.42	0.59	0.56
2	0.636	0.27	0.82	0.46	0.28	0.4	0.54	0.3
3	0.63	0.44	0.65	0.86	0.9	0.51	0.29	0.09
4	0.28	0.48	0.64	0.71	0.7	0.86	0.86	0.82
5	0.89	0.62	0.14	0.05	0.8	0.49	0.82	0.39
6	0.75	0.71	0.3	0.99	0.02	0.69	0.98	0.05
7	0.9	0.62	0.98	0.78	0.1	0.08	0.09	0.62
8	0.8	0.23	0.1	0.55	0.85	0.25	0.5	0.18
9	0.45	0.75	0.38	0.32	0.5	0.79	0.58	0.25
10	0.1	0.32	0.3	0.31	0.6	0.11	0.08	0.75

Figure 10 shows the accuracy of four ML models. Each model's performance fluctuates across epochs, with the Hybrid model consistently showing higher accuracy than the other models in most epochs. The CatBoost model occasionally reaches similar accuracy levels but is generally outperformed by the Hybrid model. The DT and RF models display more variability, often lagging behind in accuracy. Overall, the Hybrid model demonstrates superior and more stable performance, suggesting its effectiveness in accurately detecting intrusions in IoT-based WSNs.

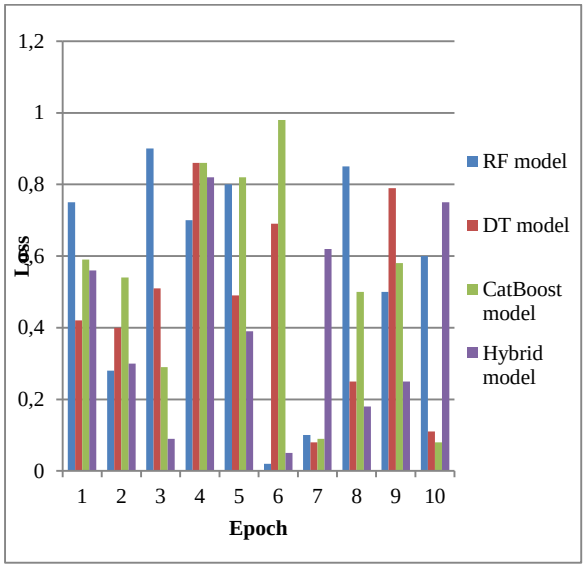

Figure 11. *Analysis of the loss of various ML models*

Figure 11 illustrates the loss values of four ML models that are identical. Lower loss values indicate higher model efficacy, as they reflect a smaller difference between expected and actual results. The Hybrid model consistently exhibits superior performance in terms of decreased loss values across multiple epochs, emphasizing its resilience and efficacy in identity detection. The CatBoost model occasionally experiences infrequent instances of greater loss values, however it generally demonstrates strong performance. The RF and DT models demonstrate higher volatility and frequently experience bigger levels of loss. This inclination enhances the reliability of the Hybrid model in minimizing forecast mistakes, leading to a more effective tool for preventing unwanted access in IoT-based WSNs.

The suggested method attains in-depth evaluation outcomes regarding its ID performance on attacks through Table 3 analysis of DoS, Probe, RPL Rank Attack, Sybil Attack, and Blackhole attack types. The model attains a high level of performance in intrusion detection by sustaining 97% to 99% accuracy, precision, recall and F1-score measurement levels. Sybil attack detection demonstrates the best rates of performance with a 99.1% success rate that is equivalent to the detection rates of other inspected attacks. The threat detection ability of the system achieves a minimum of 98% average Intrusion Detection Rate demonstrated by its capacity to identify threats and block them using few false alarms. The model presents effective and stable security features for securing IoT-based WSN networks against cyber-attacks in its outstanding performance detection.

Table 3. *Overall performance of the Hybrid model for various attacks.*

Attack Type	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	IDR
DoS	98.3	98.1	98.4	98.25	99.1
Probe	97.9	97.8	97.9	97.85	98.7
RPL Rank Attack	98.5	98.3	98.6	98.45	98.9
Sybil Attack	99.1	99.0	99.2	99.1	99.5
Blackhole	97.8	97.6	97.9	97.75	98.3
Average	98.32	98.16	98.4	98.28	98.9

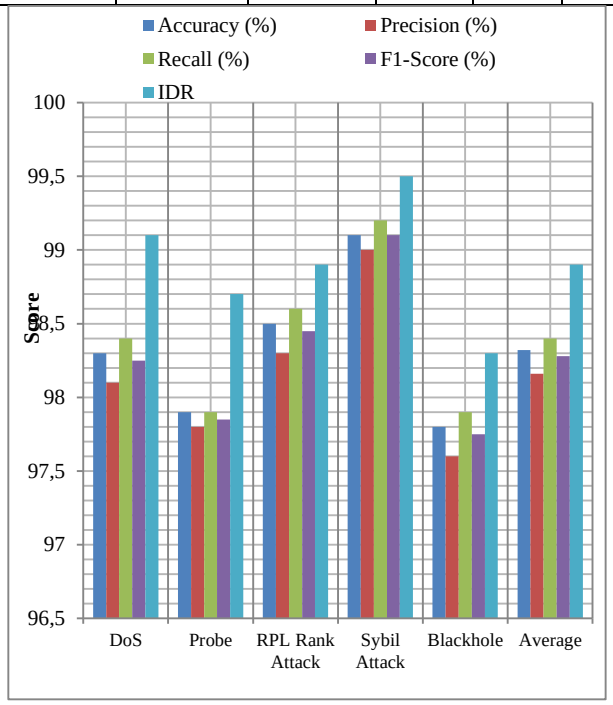

Figure 12. *Performance of Hybrid ML-based Intrusion Detection in IoT-WSN.*

The comparison of various existing model with proposed model are as shown in Figure 12. The system results in outstanding metrics of detection

across various types of attacks as proved by stability on various methods of attack evaluation maintaining accuracy and precision levels, ability to recall and F1-score at rates over 97% and 99% respectively. Sybil attack detection reaches optimal performance but all the attacks show consistent detection effectiveness. Normal and malicious behavior is identified efficiently based on the provided input data thereby showcasing the capacity to protect IoT-based WSN networks from cybersecurity attacks.

Comparative Analysis

Table 4 compares various intrusion detection systems by evaluating DNN, NB, DRNN, DCNN and KNN-PSO with respect to the novel hybrid machine learning paradigm. The suggested model outperforms other models by yielding 98.32% accuracy along with 98.16% precision and 98.4% recall and 98.28% F1-score. The NB model shows very competitive performance but DCNN shows the worst accurate performance with only 89.1% accuracy. The resulted outcomes prove that the suggested approach achieves peak efficiency in detecting intrusions on IoT-based WSN networks.

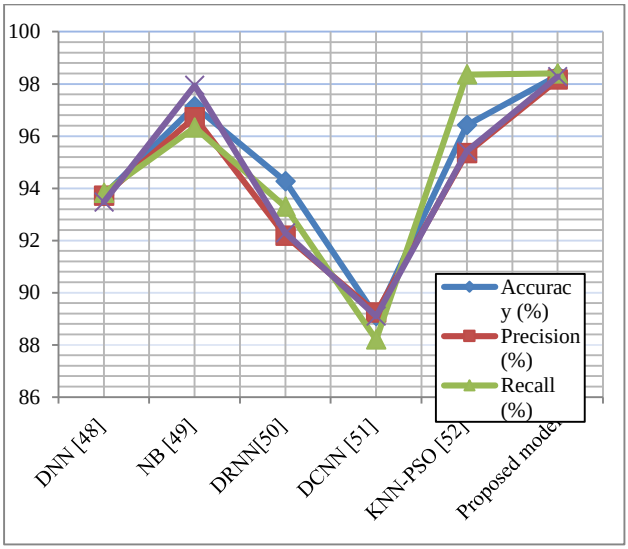

Figure 13. *Performance metrics comparison of different models for intrusion detection.*

Figure 13 represents different performance metrics score utilized in this research for distinct intrusion detection models. It demonstrates performance degradation in the DCNN model, whereas the proposed model maintains the highest values on all metrics throughout. The trends demonstrate that proposed model greatly improves intrusion detection within IoT-based WSN networks, validating its effectiveness over traditional approaches.

Novelty of the Study

The originality of this work is in using GIWRF and GA in optimized feature selection, along with a hybrid model of CatBoost-LSTM to detect intrusion in IoT-based WSNs. As opposed to conventional IDS methods that make use of ML or DL alone, this research uses the complementary powers of both methods—CatBoost for processing categorical data and LSTM for extracting temporal relationships in network traffic. The performance of detection is enhanced by choosing 27 top features from the UNSW-NB15 dataset while maintaining computational expenses at a low level. Performance evaluation by the proposed method is carried out widely over various intrusion environments which resolves model stability issues and generates stronger results compared to conventional methods. The hybridization model offers better IDS flexibility to accommodate in changing IoT networks which results in deployable security solutions for networks. Machine learning is applied in different fields and reported [53-60].

7. Discussion and Conclusion

Table 4. *Performance comparison of different models with the proposed hybrid model*

Models	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
DNN [48]	93.74	93.712	93.824	93.472
NB [49]	97.14	96.72	96.33	97.94
DRNN[50]	94.27	92.18	93.29	92.29
DCNN [51]	89.1	89.23	88.2	89.1
KNN-PSO [52]	96.42	95.35	98.36	95.42
Proposed model	98.32	98.16	98.4	98.28

When comparing the findings of this study on ID&P in IoT-based WSNs using ML and Big Data analytics (BDA) with previous research, some significant discoveries emerge. Data collection methods for IoT operations demonstrated efficiency in managing the complexities of IoT data and building training data instances based on the findings of current research. The integration of BDA methodology for feature gathering with GA and GIRPF data extraction techniques is successful since it reflects earlier methods that enhanced model precision and operating effectiveness. The system achieves its objective through feature maximization and selecting characteristics depending on associated patterns. Various deployment environments of WSN are enhanced by ML models

RF, DT, and Hybrid models when considering node identification. This is consistent with other research that highlights the importance of selecting appropriate models based on specific application requirements. This study contributes to the existing knowledge by providing up-to-date insights on the evolving environment of identification in IoT-based WSNs. The focus is on the ongoing advancement of Hybrid ML approaches in conjunction with BDA, with the goal of improving network dependability and security.

The goal of this study was to train and evaluate the machine learning models, including DT, RF, CatBoost, and Hybrid, for the binary classification function of ML-based IDS. In order to choose an appropriate collection of features from two datasets that have an imbalance in their distribution: The UNSW-NB 15 dataset utilizes a GA tech for feature extraction. Furthermore, the GIRFW approach is presented as the feature evaluation procedure. The decision-making strategy reduced the amount of information in the UNSW-NB 15 data set. The models were assessed based on their accuracy and loss score in order to detect intrusions. Initially, the experiment employed a single ML method to measure both accuracy and loss. Subsequently, the experiment was conducted for each of the four techniques. An evaluation of the model's performance was conducted. For UNSW-NB 15 datasets, the Hybrid model exhibited superior performance when combined with the attribute-selecting technique. However, this study did not include the use of multiclass classification and time complexity analysis. Therefore, future research should focus on developing a multiclass categorization scheme for IDS that takes into account time complexity analysis.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] Khan, Sharfuddin, E. Sivaraman, and Prasad B. Honnavalli, (2020). Performance evaluation of advanced machine learning algorithms for network intrusion detection system. In *Proceedings of International Conference on IoT Inclusive Life (ICIIL 2019), NITTTR Chandigarh, India*, 51-59. Springer Singapore.

[2] Zhao, Ruijie, Guan Gui, Zhi Xue, Jie Yin, Tomoaki Ohtsuki, Bamidele Adebisi, and Haris Gacanin, (2021). A novel intrusion detection method based on lightweight neural network for internet of things. *IEEE Internet of Things Journal* 9(12): 9960-9972.

[3] Yang, Li, Abdallah Moubayed, Abdallah Shami, Parisa Heidari, Amine Boukhtouta, Adel Larabi, Richard Brunner, Stere Preda, and Daniel Migault (2021). Multi-perspective content delivery networks security framework using optimized unsupervised anomaly detection. *IEEE Transactions on Network and Service Management* 19(1). 686-705.

[4] Injadat, MohammadNoor, Abdallah Moubayed, Ali Bou Nassif, and Abdallah Shami. (2021). Machine learning towards intelligent systems: applications, challenges, and opportunities. *Artificial Intelligence Review* 54, no. 5, 3299-3348.

[5] Yang, Li, and Abdallah Shami, (2022). IoT data analytics in dynamic environments: From an automated machine learning perspective. *Engineering Applications of Artificial Intelligence* 116:105366.

[6] Zuo, Wangmeng, David Zhang, and Kuanquan Wang, (2008). On kernel difference-weighted k-nearest neighbor classification. *Pattern Analysis and Applications* 11: 247-257.

[7] Safavian, S. Rasoul, and David Landgrebe, (1991). A survey of decision tree classifier methodology. *IEEE Transactions on Systems, Man, and cybernetics* 21:3. 660-674.

[8] Khalil, Ruhul Amin, Nasir Saeed, Mudassir Masood, Yasaman Moradi Fard, Mohamed-Slim Alouini, and Tareq Y. Al-Naffouri. (2021). Deep learning in the industrial Internet of things: Potentials, challenges, and emerging applications. *IEEE Internet of Things Journal* 8(14): 11016-11040.

[9] Alsabti, Khaled, Sanjay Ranka, and Vineet Singh, (1997). An efficient k-means clustering algorithm.

[10] Li, Lishuai, R. John Hansman, Rafael Palacios, and Roy Welsch, (2016). Anomaly detection via a Gaussian Mixture Model for flight operation and safety monitoring.*Transportation Research Part C: Emerging Technologies* 64, 45-57.

[11] Liu, Fei Tony, Kai Ming Ting, and Zhi-Hua Zhou, (2008). Isolation forest. In *2008 eighth ieee international conference on data mining*, 413-422. IEEE.

[12] Bamakan, Seyed Mojtaba Hosseini, Behnam Amiri, Mahboubeh Mirzabagheri, and Yong Shi, (2015). A new intrusion detection approach using PSO based multiple criteria linear programming. *Procedia Computer Science* 55: 231-237.

[13] Wu, Shelly Xiaonan, and Wolfgang Banzhaf, (2010). The use of computational intelligence in intrusion detection systems: A review. *Applied soft computing* 10(1): 1-35.

[14] H.-J. Liao, C.-H. R. Lin, Y.-C. Lin, and K.-Y. Tung, (2013). Intrusion detection system: A comprehensive review, Journal of Network and Computer Applications, vol. 36(1), 16–24.

[15] Suthaharan, Shan, (2014). Big data classification: Problems and challenges in network intrusion prediction with machine learning. *ACM SIGMETRICS Performance Evaluation Review* 41(4). 70-73.

[16] Zhang, Jiong, and Mohammad Zulkernine, (2006). Anomaly-based network intrusion detection with unsupervised outlier detection. In *2006 IEEE International Conference on Communications*, vol. 5, 2388-2393. IEEE.

[17] Alhayali, Royida A. Ibrahem, Mohammad Aljanabi, Ahmed Hussein Ali, Mostafa Abdulghfoor Mohammed, and Tole Sutikno, (2021). Optimized machine learning algorithm for intrusion detection. *Indonesian Journal of Electrical Engineering and Computer Science* 24(1), 590-599.

[18] Abd, Shamis N., Mohammad Alsajri, and Hind Raad Ibraheem, (2020). Rao-SVM machine learning algorithm for the intrusion detection system. *Iraqi Journal for Computer Science and Mathematics* 1,(1) 23-27.

[19] Liu, Gaoyuan, Huiqi Zhao, Fang Fan, Gang Liu, Qiang Xu, and Shah Nazir, (2022). An enhanced intrusion detection model based on improved kNN in WSNs. *Sensors* 22(4). 1407.

[20] Liu, Chao, Jing Yang, and Jinqiu Wu, (2020). Web intrusion detection system combined with feature analysis and SVM optimization. *EURASIP Journal on Wireless Communications and Networking* 2020, no. 1(33).

[21] Al-Janabi, Mohammed, and Mohd Arfian Ismail, (2021). Improved intrusion detection algorithm based on TLBO and GA algorithms. *Int. Arab J. Inf. Technol.* 18(2) 170-179.

[22] Almomani, Omar, (2020). A feature selection model for network intrusion detection system based on PSO, GWO, FFA, and GA algorithms. *Symmetry* 12(6), 1046.

[23] Bhattacharya, Sweta, Praveen Kumar Reddy Maddikunta, Rajesh Kaluri, Saurabh Singh, Thippa Reddy Gadekallu, Mamoun Alazab, and Usman Tariq, (2020). A novel PCA-firefly based XGBoost classification model for intrusion detection in networks using GPU. *Electronics* 9(2):219.

[24] Kumar, Neeraj, and Sanjeev Sharma, (2023). A Hybrid Modified Deep Learning Architecture for Intrusion Detection System with Optimal Feature Selection. *Electronics* 12(19): 4050.

[25] Nazir, Anjum, and Rizwan Ahmed Khan (2021). A novel combinatorial optimization based feature selection method for network intrusion detection. *Computers & Security* 102, 102164.

[26] Hanif, Sohaib, Tuba Ilyas, and Muhammad Zeeshan, (2019). Intrusion detection in IoT using artificial neural networks on UNSW-15 dataset. In *2019 IEEE 16th international conference on smart cities: improving quality of life using ICT & IoT and AI (HONET-ICT)*, 152-156. IEEE.

[27] Harrison, Onel, (2018). Machine learning basics with the k-nearest neighbors algorithm. *Towards Data Science* 11.

[28] Hassija, Vikas, Vinay Chamola, Vikas Saxena, Divyansh Jain, Pranav Goyal, and Biplab Sikdar, (2019). A survey on IoT security: application areas, security threats, and solution architectures. *IEEE Access* 7, 82721-82743.

[29] He, Wenhao, Hongjiao Li, and Jinguo Li. (2019). Ensemble feature selection for improving intrusion detection classification accuracy. In *Proceedings of the 2019 international conference on artificial intelligence and computer science*, 28-33.

[30] Hodo, Elike and Bellekens, Xavier and Hamilton, Andrew and Dubouilh, Pierre-Louis and Iorkyase, Ephraim and Tachtatzis, Christos and Atkinson, Robert, (2016). Threat analysis of IoT networks Using Artificial Neural Network Intrusion Detection System." 2016 International Symposium on Networks, Computers and Communications (ISNCC). IEEE, 1-6.

[31] Idrissi, Idriss and Boukabous, Mohammed and Azizi, Mostafa and Moussaoui, Omar and El Fadili, Hakim (2021). Toward a deep learning-based intrusion detection system for IoT against botnet attacks. *IAES International Journal of Artificial Intelligence.*

[32] Imrana, Yakubu and Xiang, Yanping and Ali, Liaqat and Abdul-Rauf, Zaharawu and Hu, Yu-Chen and Kadry, Seifedine and Lim, Sangsoon, (2022). χ 2-bidlstm: a feature driven intrusion detection system based on χ 2 statistical model and bidirectional lstm. Sensors 22.5.

[33] Inamdar, Ashwinin (2021). Data Science. Ensemble Learning Techniques in Machine Learning 18-9.

[34] Jabez, Ja and Muthukumar, B. (2015). Intrusion Detection System (IDS): Anomaly Detection using Outlier. Procedia Computer Science. *Elsevier*, 338-346.

[35] Disha, Raisa Abedin, and Sajjad Waheed, (2022). Performance analysis of machine learning models for intrusion detection system using Gini Impurity-based Weighted Random Forest (GIWRF) feature selection technique. *Cybersecurity* 5(1):1.

[36] Shafiq, Muhammad, Zhihong Tian, Ali Kashif Bashir, Xiaojiang Du, and Mohsen Guizani, (2020). CorrAUC: a malicious bot-IoT traffic detection method in IoT network using machine-learning techniques. *IEEE Internet of Things Journal* 8(5): 3242-3254.

[37] Koroniotis, Nickolaos, Nour Moustafa, Elena Sitnikova, and Benjamin Turnbull, (2019). Towards the development of realistic botnet dataset in the internet of things for network forensic analytics:

Bot-iot dataset. *Future Generation Computer Systems* 100: 779-796.

[38] Liu, Jingyu, Dongsheng Yang, Mengjia Lian, and Mingshi Li, (2021). Research on intrusion detection based on particle swarm optimization in IoT. *IEEE Access* 9: 38254-38268.

[39] Chohra, Aniss, Paria Shirani, ElMouatez Billah Karbab, and Mourad Debbabi, (2022). Chameleon: Optimized feature selection using particle swarm optimization and ensemble methods for network anomaly detection. *Computers & Security* 117: 102684.

[40] Moustafa, Nour, Benjamin Turnbull, and Kim-Kwang Raymond Choo, (2018). An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of internet of things. *IEEE Internet of Things Journal* 6(3): 4815-4830.

[41] Leevy, Joffrey L., John Hancock, Taghi M. Khoshgoftaar, and Jared M. Peterson, (2022). IoT information theft prediction using ensemble feature selection. *Journal of Big Data* 9, (1): 6.

[42] Gavel, Shashank, Ajay Singh Raghuvanshi, and Sudarshan Tiwari, (2022). An optimized maximum correlation based feature reduction scheme for intrusion detection in data networks. *Wireless Networks* 28(6): 2609-2624.

[43] Zhou, Lu, Ye Zhu, Tianrui Zong, and Yong Xiang, (2022). A feature selection-based method for DDoS attack flow classification. *Future Generation Computer Systems* 132: 67-79.

[44] Aggarwal, Ashwani Kumar, (2022). Learning texture features from glcm for classification of brain tumor mri images using random forest classifier. *Trans Signal Process* 18: 60-63.

[45] Moustafa N (2021) A new distributed architecture for evaluating AI-based security systems at the edge: *network TON_IoT datasets.* Sustain Cities Soc 72:102994

[46] M.A. Omari, M. Rawashdeh, F. Qutaishat, M. Alshira'H, N. Ababneh, An intelligent tree-based intrusion detection model for cyber security, *Journal of Network and Systems Management 29*. doi: 10.1007/s10922-021-09591-y.

[47] X. Deng, Q. Liu, Y. Deng, S. Mahadevan, (2016). An improved method to construct basic probability assignment based on the confusion matrix for classification problem. *Inf. Sci*. 340–341. 250–261.

[48] Awajan, Albara, (2023). A novel deep learning-based intrusion detection system for IOT networks. Computers 12(2), 34.

[49] Saheed, Yakub Kayode, Aremu Idris Abiodun, Sanjay Misra, Monica Kristiansen Holone, and Ricardo Colomo-Palacios, (2022). A machine learning-based intrusion detection for detecting internet of things network attacks. *Alexandria Engineering Journal* 61(12):9395-9409.

[50] Almiani, Muder, Alia AbuGhazleh, Amer Al-Rahayfeh, Saleh Atiewi, and Abdul Razaque, (2020). Deep recurrent neural network for IoT intrusion detection system. *Simulation Modelling Practice and Theory 101*: 102031.

[51] Mopuru, Bhargavi, and Yellamma Pachipala, (2024). Advancing IoT Security: Integrative Machine Learning Models for Enhanced Intrusion Detection in Wireless Sensor Networks. *Engineering, Technology & Applied Science Research* 14(4). 14840-14847.

[52] Karthikeyan, M., D. Manimegalai, and Karthikeyan RajaGopal, (2024). Firefly algorithm based WSN-IoT security enhancement with machine learning for intrusion detection. *Scientific Reports* 14(1): 231.

[53] Sakshi Taaresh Khanna, Khatri, S. K., & Sharma, N. K. (2025). Advancements in Artificial Intelligence for Oral Cancer Diagnosis. *International Journal of Computational and Experimental Science and Engineering*, 11(2). https://doi.org/10.22399/ijcesen.1666

[54] Ibeh, C. V., & Adegbola, A. (2025). AI and Machine Learning for Sustainable Energy: Predictive Modelling, Optimization and Socioeconomic Impact In The USA. *International Journal of Applied Sciences and Radiation Research* , 2(1). https://doi.org/10.22399/ijasrar.19

[55] G. Prabaharan, S. Vidhya, T. Chithrakumar, K. Sika, & M.Balakrishnan. (2025). AI-Driven Computational Frameworks: Advancing Edge Intelligence and Smart Systems. *International Journal of Computational and Experimental Science and Engineering*, 11(1). https://doi.org/10.22399/ijcesen.1165

[56] Hafez, I. Y., & El-Mageed, A. A. A. (2025). Enhancing Digital Finance Security: AI-Based Approaches for Credit Card and Cryptocurrency Fraud Detection. *International Journal of Applied Sciences and Radiation Research*, 2(1). https://doi.org/10.22399/ijasrar.21

[57] M.K. Sarjas, & G. Velmurugan. (2025). Bibliometric Insight into Artificial Intelligence Application in Investment. *International Journal of Computational and Experimental Science and Engineering*, 11(1). https://doi.org/10.22399/ijcesen.864

[58] Olola, T. M., & Olatunde, T. I. (2025). Artificial Intelligence in Financial and Supply Chain Optimization: Predictive Analytics for Business Growth and Market Stability in The USA. *International Journal of Applied Sciences and Radiation Research*, 2(1). https://doi.org/10.22399/ijasrar.18

[59] ZHANG, J. (2025). Artificial intelligence contributes to the creative transformation and innovative development of traditional Chinese culture. *International Journal of Computational and Experimental Science and Engineering*, 11(1). https://doi.org/10.22399/ijcesen.860

[60] García, R., Carlos Garzon, & Juan Estrella. (2025). Generative Artificial Intelligence to Optimize Lifting Lugs: Weight Reduction and Sustainability in AISI 304 Steel. *International Journal of Applied Sciences and Radiation Research*, 2(1). https://doi.org/10.22399/ijasrar.22