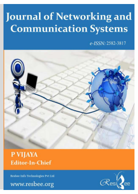
Vol.6 No.2 April 2023

Journal of Networking and Communication Systems

Received 12 April, Revised 26 April, Accepted 10 May



Machine Learning-Based Ids for IoT: A Comparative Analysis

Aouatif Arqane
Department of Computer Science
LAROSERI Laboratory Chouaib Doukkali University, El Jadida, Morocco

Omar Boutkhoum
Department of Computer Science
LAROSERI Laboratory Chouaib Doukkali University, El Jadida, Morocco

Hicham Boukhriss
Department of Computer Science
LAROSERI Laboratory Chouaib Doukkali University, El Jadida, Morocco

Abdelmajid Elmoutaouakkil
Department of Computer Science
LAROSERI Laboratory Chouaib Doukkali University, El Jadida, Morocco

Abstract: Now that we can connect everyday objects to the Internet through integrated terminals, the Internet of Things (IoT) has become a fundamental part of our life and a key component in many fields like Health care, Industry, and Education. IoT refers to the network of physical devices that are embedded with sensors, and software, enabling them to collect and exchange data over the Internet. These devices can communicate with each other, share data, and perform actions based on that data, creating a seamless and automated network of connected objects. However, IoT security has come under intense scrutiny after various high-profile attacks where typical IoT devices were used to penetrate and attack the network of many important organizations. Furthermore, traditional IoT security measures, such as firewalls along with authentication and encryption protocols are no more efficient against modern and sophisticated attacks. Motivated by this fact, we decided to perform a comparative analysis of four well-known Machine Learning algorithms to identify their efficiency in classifying attacks on the newest IoT dataset called EDGE-IIOTSET 2022. In this context, the present paper illustrates and analysis the results of a comparative analysis of the four classifiers named: Support Vector Machine (SVM), Naive Bayes (NB), Decision tree (DT), and Light Gradient Boosted (Light GBM). Moreover, and to reduce dimensionality and enhance the model's performance the Pearson Correlation coefficient is used. Our empirical results of experiments conducted on the aforementioned dataset indicate that the best classification accuracy was achieved by Light GBM followed by SVM.

Keywords: *Machine Learning, Intrusion Detection System, Internet of Things, feature selection, classification.*

Nomenclature

Acronym	Description
ACO	Ant Colony Optimization
GA	Genetic Algorithm
ML	Machine Learning
IoT	Internet of Things
DLHA	Double-Layered Hybrid Approach
PSO	Particle Swarm Optimization
DL-IDS	Deep Learning-based Intrusion Detection System
SMO	Spider Monkey Optimization
SDPN	Stacked-Deep Polynomial Network
ML-IDS	Machine Learning-based Intrusion Detection System
PCA	Principal Component Analysis
DID	Deep Learning-based Intrusion Detection
NIDS	Network Intrusion Detection System
DCNN	Deep Convolutional Neural Network
CFS	Correlation-based Feature Selection
PCC	Pearson Correlation coefficient
SVM	Support Vector Machine
DT	Decision Tree
NB	Naive Bayes
Light GBM	Light Gradient Boosted Machine
EFB	Exclusive Feature Bundling
GOSS	Gradient-based One-Side Sampling

https://doi.org/10.46253/jnacs.v6i2.a2

1. Introduction

Over the past few years, the accessibility of the Internet enables the connection of numerous objects in the same network to white out human interaction. Thus, the IoT has become one of the most important innovations of the 21st century. An IoT ecosystem is comprised of smart interconnected devices that utilize incorporated sensors, terminals, and processors to exchange and collect data easily.

Since the IoT is a nascent market, many product designers and manufacturers are more interested in getting their products to market rapidly than in taking the necessary cost and effort to secure the environment from the beginning [12]. Another common problem with IoT devices is that they are often resource-constrained and lack the computational resources needed to implement strong security. Due to this issue, many devices do not or cannot offer advanced security features, for example, sensors that monitor humidity or temperature cannot handle advanced encryption or other security measures. Thus and because of their nature, these devices are the most vulnerable to intrusions and attacks than gateways [15]. Consequently, the implementation of security measures is essential to guarantee the security of networks and connected IoT devices.

ML is a branch of artificial intelligence encompassing many algorithms to automatically create models from data. Contrary to traditional algorithms, ML models do not follow instructions but learn from experience. Therefore, its performance improves as the algorithm is exposed to more data. It is very effective in situations where insights need to be discovered from large sets of diverse and changing data, i.e. Big Data. Because of the rapid evolution of ML, it becomes a major asset in the intrusion detection and treatment of cyber risks. Indeed, one of the greatest challenges for cyber security experts is to anticipate the attacks of tomorrow, and using ML models facilitates this task [19].

In this context, this paper presents a comparative analysis of four of the most popular ML algorithms to create an efficient IDS. The main contributions of this research are:

- Performing data pre-processing and feature engineering on the EDGE-IIOTSET dataset.
- Using Pearson correlation to select the most relevant features.
- Implementing four ML algorithms.
- Analyzing and choosing the most accurate algorithm.

The rest of the paper is organized into four sections. In Section 2 the literature review of related studies is presented. Section 3 illustrates the research methodology. Section 4 provides an analysis of the empirical results obtained for each algorithm. Finally, Section 5 indicates the conclusions and future work.

2. Literature Review

This study aims to store view some machine learning classifiers in IDS that can help researchers to choose an efficient model for designing a robust ML-based IDS. Thus, we have emphasized vital components of research, like the model used in each paper, datasets, and technical issues.

In 2021, Alsarhan, *et al.* [3] have analyzed the performance of SVM when combined with three optimization algorithms, named: ACO, GA, and PSO. According to experiments conducted on the NSL-KDD dataset, GA has demonstrated better performance in terms of classification accuracy compared to two other optimization algorithms.

In 2021, Wisanwanichthan & Thammawichai, [20] have used a DLHA approach, to detect anomalies and new attacks. In the first layer of DLHA, the NB classifier was used to detect Probe and DoS attacks, while an SVM classifier was deployed in the second layer to differentiate U2R and R2L from normal traffic. The experimental outcomes conducted on the NSL-KDD dataset show that the combination of the two classifiers can achieve better performance than using a single classifier.

In 2019, Ahmim *et al.,* [2] have implemented a novel approach to detect network intrusions by combining a decision tree classifier with three rules-based methods that are: Forest PA, JRip algorithm, and REP Tree. The empirical findings obtained by evaluating the approach using the CICIDS2017 dataset, confirm their superiority in regards to classification accuracy, false alarm, and detection rate compared to other existing studies.

In 2021, Otoum, *et al.,* [22] have executed DL-IDS to identify security threads in an IoT environment. Initially, datasets were cleaned and the SMO algorithm was used to choose the optimal features in the dataset. Then, the SDPN classifies the data as normal or anomalous (DoS, U2R, probe, and R2L). Finally, Performance was evaluated based on accuracy, precision, recall, and F-score.

In 2022, Saheed, *et al.,* [23] have ensembled ML-IDS for detecting IoT network attacks. Initially, the min-max normalization technique was used to attribute values on the same scale. The PCA reduced the features and Six machine learning models were used for analysis, including XGBoost, Random Forest, Decision Tree, K-Nearest Neighbor, Support Vector Machine, and Logistic Regression and finally evaluated in terms of validation data.

In 2023, Awajan, A., [24] have executed DID system for IoT devices. Initially, the emulated network communicates with the feature extractor and network classifier through an interface module. The deep neural network used the extracted features to identify malicious traffic on connected IoT devices.

In 2022, Sahba Baniasadi., *et al.* [25] have implemented NIDS for the IoT using DCNN. Initially, data were passed through convolution and pooling layers and MLP classify the datasets. Furthermore, the NSBPSO algorithm was used to train the deep architecture for better accuracy and performance. Finally evaluated using UNSW-NB15 and Bot-IoT.

In 2021, Islam, *et al.*, [26] have used IoT threats and shallow for intrusion detection in the IoT environment. Initially, Data were preprocessed using the Standard Scaler method. The relevant features were selected through the CFS method. Shallow and deep machine learning models were trained and tested on the benchmark datasets. Performance was evaluated using Accuracy, Precision, Recall, and F1-score. The best-performing model was selected and its hyper parameters were optimized using the grid search method and further deployed for intrusion detection in IoT systems.

2.1 Review

Table 1 portrays the methodology, advantages, and disadvantages of the existing method. We considered eight papers that used a different methodology for this study. Each method has certain benefits and short comings that were explained.

Table 1: *Review Based on Existing Methods.*

Study	Methodology	Advantage	Disadvantage
Alsarhan, *et al.* [3]	SVM	• Special direction at a finite sample and irrelevance between the complexity of the algorithm and the sample dimension. • Three intelligence optimization algorithms were used.	• Effectiveness may depend on the specific characteristics of the VANET being used and the types of attacks that are being targeted.
Wisanwanichthan & Thammawichai, [20]	DLHA	• Robust and accurate in detecting anomalies and unseen attacks. • Helped in identifying common characteristics of different attack categories. • The distinct patterns were easily detected. • Detects rare attacks.	• Effectiveness of the approach may depend on the quality of the data. • May require more computational resources compared to simpler signature-based methods.
Ahmim, *et al.* [2]	Forest PA, JRip algorithm, and REP Tree	• Higher accuracy in detecting and classifying network traffic. • Lower false alarm rate. • Lower time overhead.	• Required further optimization and tuning to achieve optimal performance in different scenarios.
Otoum, *et al.* [22]	DL-IDS	• DL-IDS was more effective than other methods. • Improved accuracy of attack detection. • Helped to detect different types of anomalies. • Handled datasets with uncertain or missing data and redundant values.	• Required high computational resources due to the use of deep learning techniques.
Saheed, *et al.* [23]	ML-IDS	• High accuracy in detecting attacks. • Ability to detect new and unknown attacks. • Handled large amounts of data. • Identify patterns that may not be visible to humans.	• Evaluated on a single dataset, • The performance of the ML-IDS system may be affected by the quality and quantity of the training data used. • May not be suitable for real-time detection, • Required significant computational resources. • May not be able to detect attacks that use sophisticated evasion techniques.
Awajan, A. [24]	DID	• Allows intrusions in real-time.	• Required significant computational resources to operate in real-time.
Sahba Baniasadi., *et al.* [25]	DCNN	• Allowed more accurate and efficient detection of network intrusion in the IoT.	• Cannot be extended to multi-objective feature selection to optimize the classification accuracy.
Islam, *et al.* [26]	SVM and DL	• Can learn complex patterns and relationships in data, • Detect new and unknown attacks. • Can be used in real-time.	• ML models were computationally expensive and required significant computational resources.

2.2 Challenges

According to the reviewed papers, ML classifiers can highly improve the detection capabilities of an IDS. However, the majority of them used old datasets like NSL-KDD which can influence negatively the obtained results because we cannot judge the performance of the proposed approaches against the new attacks scenarios [3][20]. The effectiveness of the approach may vary based on the quality of the input. [22][23][24][26] explained that these methods were difficult to implement in real time because of their computational cost. Thus, the results can be misleading since the accuracy is not enough the show the effectiveness of an approach.

3. Research Methodology

This section includes the description of the used dataset, alongside the various steps performed to prepare it for ML usage.

The experiments were carried out on a machine with 8 GB of RAM, and a 4 GB GPU. To implement the models, we used the Scikit-learn, NumPy, and Pandas libraries.

Figure 1 depicts the flowchart of the methodology used to achieve the experimental results. The dataset was separated into a training set with 70% of records and a testing set with 30% of records.

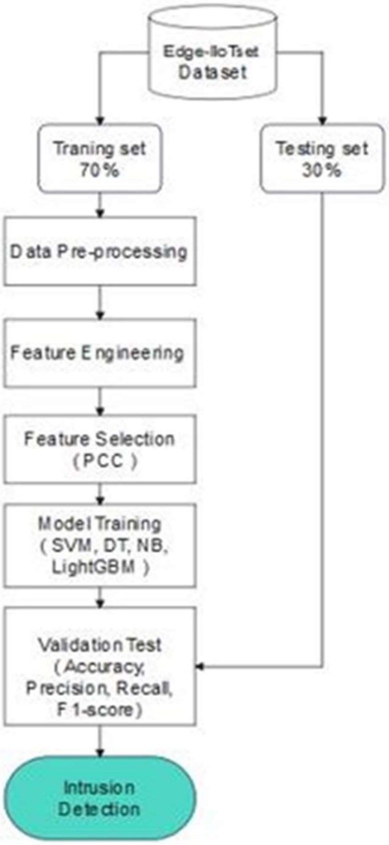


Fig 1. *Flowchart of the methodology*

3. 1 Dataset

The EDGE-IIOTSET dataset [9] is the newest publicly available dataset that is dedicated to IoT and IIoT environments. To obtain the data, the experiments were conducted from November 21, 2021, to January 10, 2022, in a disconnected manner. The used test bed was constructed out of seven interconnected layers: IoT/IIoT perception layer, edge layer, SDN layer, fog layer, Blockchain layer, NFV layer, and cloud computing layer. The dataset contains 11223940 normal records and 9728708 attack records along with 63 features. Also, it consists of fourteen types of attacks characterized by five threats namely: Information gathering, DoS/DDoS attacks, Injection attacks, Malware attacks, and Man in the middle attacks. Besides, there are two reduced versions of this dataset to evaluate ML and DL methods. But for this study, we used the whole original dataset.

3.2 Data Pre-processing

We started the work on the dataset by removing fifteen useless features which are: frame.time,tcp.options, http.request.full_uri, arp.src.proto_ipv4,ip.src_host, arp.dst.proto_ipv4, http.file_data,ip.dst_host, http.request.uri.query, tcp.srcport,mqtt.msg, tcp.dstport,tcp.payload, icmp.transmit_timestamp, and udp.port.

The second step was dropping duplicate records and rows containing missing values.

3.3 Feature engineering

Some ML algorithms cannot handle categorical features [13], We converted the six remaining categorical features namely:

mqtt.topic, mqtt.protoname, mqtt.conack.flags, dns.qry.name.len, http.request.version, and http.referer into dummy features using the *get_dummies* function of pandas.

The "Attack-type" label is also a categorical feature. But we used the *Label Encoder* function that's imply converts each categorical value in a column into a number.

To improve the performance of ML algorithms and reduce the training time, we employed the *Standard Scaler* function that can standardize the records.

3.4 Feature Selection

Feature selection consists of high-dimensional space, in finding a subset of relevant variables. In other words, it seeks to minimize the loss of information resulting from the suppression of all the other variables.

In our study, we utilized the PCC to obtain the optimal subset of features that offer the most relevant information. It is commonly used in ML for feature selection, where the correlation between features and the target variable is analyzed to identify the most relevant features for the model [16][11].

PCC [5] is the association or relationship between two or more features. It indicates if the features are simultaneously changed together or not. Correlation value range from -1 and +1 meaning that the closer the value is to -1 or 1, the stronger the correlation. While 0 means that the correlation is absent.

PCC has several advantages as a measure of the correlation between two variables:

- **Robustness**: It is robust to outliers.
- **Efficient**: It is an efficient measure of correlation that can be computed quickly and easily using standard statistical software.
- **Easy to interpret**: It is easy to interpret and understand.
- **Linear relationship**: It is a useful tool when analyzing data with a linear structure or when testing linear hypotheses.

On the other hand, PCC has also some limitations that must be taken into consideration when using it. Three of the most known issues of this feature selection technique are:

- It only measures the association between two variables and does not take into account other variables that may be important in the relationship.
- It assumes a linear relationship between the two variables being measured. If the relationship is non-linear, the correlation coefficient may not accurately reflect the strength or direction of the relationship.
- It assumes that the variables being measured are normally distributed.

Overall, the PCC is a useful measure of the correlation between two variables that are widely used and easy to interpret. It is robust to outliers and efficient to compute, making it a valuable tool for Feature selection. However, it is important to carefully consider the assumptions and limitations of the Pearson correlation coefficient before using it in any analysis [4].

3.5 Machine Learning Algorithms

3.5.1 Support Vector Machine

SVM [6] is a no-table supervised algorithm that can deal with classification and regression. It is one of the most powerful and accurate methods in the ML field because it offers the likelihood to manage various continuous and categorical variables.

It is commonly utilized on small datasets with an immense number of features or when the number of data points is significantly less than the number of features. However, this classifier is not appropriate for large datasets since it needs a long training time.

3.5.2 Decision Tree

A DT [7] is a graphical illustration to have all the potential answers to a specific problem according to pre-given conditions. In this tree, each node with branches is named an internal node and substitutes a feature or attribute of a dataset. The branch is a decision rule and each leaf node, which is a node without branches, represents an outcome of those decisions.

The benefits of this algorithm are efficiency in solving decision-related issues, high classification accuracy, and fewer requirements for data cleaning. The main disadvantage of the decision tree is the complexity of the hierarchy which makes the rules extraction a complicated task in deeper and wider trees.

3.5.3 Naïve Bayes

NB [17] methods are a set of supervised learning algorithms, known for their simplicity of building and implementation. They can be used for binary classification but perform well in multi-class predictions compared to other complex classifiers. Also, these types of classifiers are robust towards irrelevant features and needles straining data, so they are useful for small and medium datasets.

However, they cannot learn the relationship between features because they assume that the presence of any feature is independent or unrelated to the existence of other features.

3.5.4 Light GBM

Light GBM [14] is a sequential ensemble learning model. It aims to reduce training time and computing power by aggregating the advantages of Gradient Boosting and DT. These are combined in such a way that each new learner adjusts the residuals of the previous tree and the last added tree accumulates the results of each step and a power full earner is achieved.

To find the right tree structure for each learner added, it uses a technique based on key concepts: EFB and GOSS.

3.6 Evaluation metrics

To evaluate the performance of the classifiers, we utilized four well-known evaluation metrics:

- Accuracy measures the proportion of correctly classified instances in the dataset.
- Precision measures the proportion of true positives (correctly predicted positive instances) to the total number of positive predictions.
- Recall indicates the proportion of true positives to the total number of actual positive instances in the dataset.
- F1 Score calculates a weighted average of precision and recall, which provides a balanced measure of the model's performance.

4. Empirical Results and Analysis

The present section illustrates the results of our study that compared the performance of four ML algorithms to detect intrusions in an IoT environment

4.1 Accuracy

As it is shown in Figure 2, the best accuracy is obtained by Light GBM (96.6 %), followed by DT (95%). While SVM and NB are the less accurate algorithms.

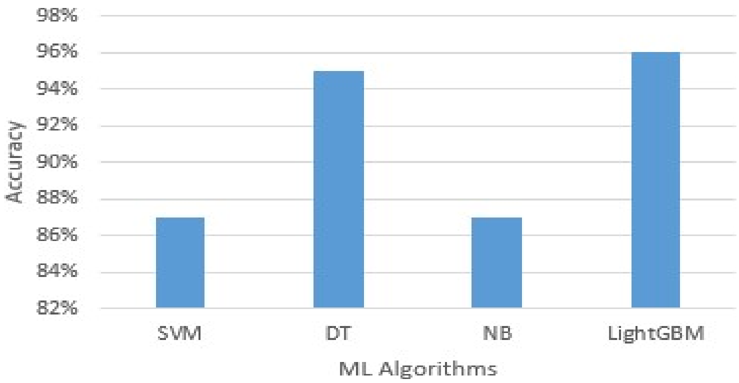


Fig 2. *Accuracy of the four algorithms*

4.2 Model comparison

Table 2 depicts the numerical results of precision, recall, and F1–score of the four algorithms used in this comparative study grouped by type of attack. We adopted the same classification of attacks given by the creators of this dataset.

As we can notice in the following Table 2, all the algorithms were able to identify normal traffic perfectly. Also, the MITM attacks were detected with high precision especially when using SVM and Light GBM algorithms.

Table 2. *Evaluation Metrics of Multi-class Classification*

Algorithm	Metric	Normal	DDOS	Injection	Malware	Scanning	MITM
SVM	Precision	1.00	0.93	0.60	0.79	0.74	1.00
	Recall	1.00	0.81	0.90	0.61	0.67	1.00
	F1-score	1.00	0.76	0.72	0.52	0.87	1.00
DT	Precision	1.00	0.78	0.88	0.87	0.75	1.00
	Recall	1.00	0.83	0.88	0.88	0.82	0.98
	F1-score	1.00	0.80	0.88	0.88	0.85	0.98
NB	Precision	1.00	0.89	0.62	0.90	0.74	0.99
	Recall	1.00	0.87	0.90	0.81	0.89	0.98
	F1-score	1.00	0.86	0.93	0.89	0.84	0.96
Light GBM	Precision	1.00	0.83	0.62	0.81	0.96	1.00
	Recall	1.00	0.90	0.69	0.78	0.64	1.00
	F1-score	1.00	0.74	0.56	0.83	0.62	1.00

4.3 Feature Selection

We started the process of data preprocessing using a dataset with 63 features. Then, after performing the feature engineering, especially encoding the features, we passed to 76 features.

In the first stage of feature selection, we dropped the non-correlated features with the target (Correlation threshold >0.5). Then, we search for the features that a recorrelated with each other to keep one of them and drop the rest. Finally, we got 14 features that contain the most relevant information.

Table 3 shows the precision, recall, and F1-score of the four algorithms after performing the feature selection process. The performance of all algorithms is greatly improved except for Light GBM where the results are almost the same.

Table 3. *Evaluation metrics of Multi-class Classification after feature selection*

Algorithm	Metric	Normal	DDOS	Injection	Malware	Scanning	MITM
SVM	Precision	1.00	0.95	0.61	0.84	0.79	1.00
	Recall	1.00	0.88	0.95	0.70	0.73	1.00
	F1-score	1.00	0.78	0.72	0.58	0.88	1.00
DT	Precision	1.00	0.83	0.93	0.91	0.76	1.00
	Recall	1.00	0.86	0.89	0.89	0.85	0.98
	F1-score	1.00	0.87	0.90	0.93	0.87	0.98
NB	Precision	1.00	0.92	0.65	0.91	0.79	0.99
	Recall	1.00	0.89	0.92	0.86	0.90	0.98
	F1-score	1.00	0.89	0.93	0.92	0.85	0.96
Light GBM	Precision	1.00	0.84	0.62	0.83	0.97	1.00
	Recall	1.00	0.92	0.69	0.78	0.64	1.00
	F1-score	1.00	0.78	0.56	0.83	0.62	1.00

4.4 Comparative Analysis

In this sub-section, we illustrated the empirical results of other works evaluated on the same dataset and their comparison.

Table 4 presents the best accuracy achieved by each ML model in these papers along with the results of our study. From Table 4, Light GBM outperformed other existing studies in terms of accuracy.

Table 4. *Comparative analysis*

Paper	Model	Accuracy (%)
FERRAG, 2022	DNN	94.67
Tareq *et al.*, 2022	Inception Time	94.94
Chatzoglou *et al.*, 2022	Bagging	98.73
Proposed study {Light GBM)	Light GBM	99.98

5. Conclusion and Future Work

Protecting IoT devices and defending networks against intrusions is one of the main concerns of cyber security experts. Thus, because of the success of ML in resolving many complicated problems in various domains, they decide to use it in the cyber security field. The main purpose of this paper is to present different ML algorithms that will help in the implementation of more efficient and robust IDS for the IoT ecosystem. We have illustrated the various steps to prepare the EDGE-IIOTSET dataset and the method used to select the most relevant features. Also, we introduced the results of our experiments with four ML models.

For future work, we intend to utilize more complex models like those of deep learning and other feature selection techniques to build an accurate and light weight IDS that can be implemented on IoT devices.

6. Compliance with Ethical Standards

Conflicts of interest: Authors declared that they have no conflict of interest.
Human participants: The conducted research follows the ethical standards and the authors ensured that they have not conducted any studies with human participants or animals.

References

[1] Ahmad,Z., ShahidKhan,A., WaiShiang,C., Abdullah, J., & Ahmad, F, (2021). "Network intrusiondetection system:Asystematic study of machine learning and deep learning approaches" , Transactions on Emerging Telecommunications Technologies,Vol. 32, no.1, pp. e4150.

[2] Ahmim,A., aglaras,L., Ferrag,M.A., Derdour,M., & Janicke,H. (2019), "A Novel Hierarchical Intrusion Detection System Based on Decision Tree and Rules-Based Models",15 th International Conference on Distributed Computing in Sensor Systems (DCOSS), pp. 228–233.

[3] Alsarhan,A.,Alauthman,M.,Alshdaifat,E.,Al-Ghuwairi,A.-R.,& Al-Dubai, A., (2021) "MachineLearning-driven optimization forSVM-basedintrusiondetectionsysteminvehicularadhocnetworks",JournalofAmbientIntelligenceandHumanizedComputing,pp. 1-10.

[4] Armstrong,R.A.,"ShouldPearson'scorrelationcoefficientbeavoided?",OphthalmicandPhysiologicalOptics, vol. 39, no.5, pp. 316–327, 2019.

[5] Benesty, J., Chen, J., Huang, Y., & Cohen, I.,"Pearson Correlation Coefficient", InI.Cohen, Y. Huang, J. Chen, & J. Benesty(Eds.), Noise Reduction in Speech Processing, pp.1–4, 2009.

[6] Cervantes,J., Garcia-Lamont,F., Rodríguez Mazahua,L., & Lopez,A., "A comprehensive survey on support vector machine classification:Applications,challenges and trends",Neuro computing,Vol. 408,189–215, 2020.

[7] Charbuty,B., & Abdulazeez,A., "Classification based on decision tree algorithm for machine learning", Journal of Applied Science and Technology Trends,Vol. 2, no.01, pp. 20–28, 2021.

[8] Chatzoglou,E., Kambourakis,G., Smiliotopoulos,C., &Kolias,C., "Best of Both Worlds:Detecting Application Layer Attacks through 802.11 and Non-802.11 Features", Sensors, Vol. 22, no.15, Article15, 2022.

[9] Douiba,M., Benkirane,S., Guezzaz,A., & Azrour,M., "An improved anomaly detection model for IoT security using decision tree and gradient boosting",The Journal of Supercomputing, Edge-IIoTset Cyber Security Dataset of IoT& IIoT.(n.d.).Retrieved 21 September, 2022.

[10] FERRAG,M.A.,"Edge-IIoTset:A new Comprehensive Realistic Cyber Security Dataset of IoT and IIoT Applications:Centralized and Federated Learning [Dataset]", IEEE, 2022.

[11] Gottwalt,F., Chang,E., & Dillon,T., "CorrCorr:A feature selection method formultivariate correlation network anomaly detection techniques", Computers & Security, Vol. 83, pp. 234–245, 2019.

[12] IoT Security Still Not a Priority, Survey Reveals.(n.d.). Hot for Security. Retrieved 21 September 2022, from

[13] Johannemann,J., Hadad,V., Athey,S., & Wager,S.(2021), “Sufficient Representations for Categorical Variables”, 2021.

[14] Ke,G., Meng,Q., Finley,T., Wang,T., Chen,W., Ma,W., Ye,Q., & Liu,T.-Y, “Light GBM:A Highly Efficient Gradient Boosting Decision Tree. Advances in Neural Information Processing Systems”, 30, 2017.

[15] Kollolu, R., “A Review on Wide Variety and Heterogeneity of IoT Platforms”, 2020. (SSRNScholarlyPaperNo.3912454).

[16] Liu,Y., Mu,Y., Chen,K., Li,Y., & Guo,J., “Daily Activity Feature Selection in Smart Homes Based on Pearson Correlation Coefficient”, Neural Processing Letters, Vol. 51, no. 2, pp. 1771–1787,2020.

[17] Saritas, M. M., & Yasar,., “Performance Analysis of ANN and Naïve Bayes Classification Algorithm for Data Classification”, International Journal of Intelligent Systems and Applications in Engineering, Vol. 7, no.2, Article, 2,2019.

[18] Tareq,I., Elbagoury,B.M., El-Regaily,S., & El-Horbaty, E.-S.M., “Analysis of ToN-IoT, UNW-NB15, and Edge-IIoT Datasets Using DL in Cybersecurity for IoT”, Applied Sciences, Vol.12, no. 19, Article. 19, 2022.

[19] Thakkar,A. ,&Lohiya,R., “A Review on Machine Learning and Deep Learning Perspectives of IDS for IoT:Recent Updates, Security Issues, and Challenges”, Archives of Computational Methods in Engineering, Vol. 28, no.4,3211–3243, 2021.

[20] Wisanwanichthan, T., & Thammawichai, M., “A Double-Layered Hybrid Approach for Network Intrusion Detection System Using Combined Naïve Bayes and SVM”, IEEE Access, vol. 9, pp. 138432–138450,2021.

[21] Otoum, Y., Liu, D. and Nayak, A., “DL-IDS: a deep learning–based intrusion detection framework for securing IoT”, Transactions on Emerging Telecommunications Technologies, Vol. 33, no. 3, pp.e3803, 2022.

[22] Saheed, Y.K., Abiodun, A.I., Misra, S., Holone, M.K. and Colomo-Palacios, R., “A machine learning-based intrusion detection for detecting internet of things network attacks”, Alexandria Engineering Journal, Vol. 61, no. 12, pp.9395-9409, 2022.

[23] Awajan, A., “A Novel Deep Learning-Based Intrusion Detection System for IoT Networks”, Computers, Vol. 12, no. 2, pp.34, 2023.

[24] Baniasadi, S., Rostami, O., Martín, D. and Kaveh, M., “A novel deep supervised learning-based approach for intrusion detection in IoT systems”, Sensors, Vol. 22, no. 12, p.4459, 2022.

[25] Islam, N., Farhin, F., Sultana, I., Kaiser, M.S., Rahman, M.S., Mahmud, M., Hosen, A.S. and Cho, G.H., “Towards machine learning based intrusion detection in IoT networks”, Comput. Mater. Contin, Vol. 69, no.2, pp.1801-1821, 2021.