Feature Reduction and Anomaly Detection in IoT Using Machine Learning Algorithms

Adel Hamdan[1], Muhannad Tahboush[2], Mohammad Adawy[3], Tariq Alwada'n[4], Sameh Ghwanmeh[5]
Computer Science Dept., The World Islamic Sciences and Education University, Amman, Jordan[1]
Information System and Network Dept., The World Islamic Sciences and Education University, Amman, Jordan[2, 3]
Network and Cybersecurity Dept., Teesside University, Middlesbrough, UK[4]
Computer Science Dept., American University in the Emirates, Dubai, UAE[5]

***Abstract*—Anomaly detection in IoT is a hot topic in cybersecurity. Also, there is no doubt that the increased volume of IoT trading technology increases the challenges it faces. This paper explores several machine-learning algorithms for IoT anomaly detection. The algorithms used are Naïve Bayesian (NB), Support Vector Machine (SVM), Decision Tree (DT), XGBoost, Random Forest (RF), and K-nearest Neighbor (K-NN). Besides that, this research uses three techniques for feature reduction (FR). The dataset used in this study is RT-IoT2022, which is considered a new dataset. Feature reduction methods used in this study are Principal Component Analysis (PCA), Particle Swarm Optimization (PSO), and Gray Wolf Optimizer (GWO). Several assessment metrics are applied, such as Precision (P), Recall(R), F-measures, and accuracy. The results demonstrate that most machine learning algorithms perform well in IoT anomaly detection. The best results are shown in SVM with approximately 99.99% accuracy**.

Keywords*—*Machine learning; Internet of Things (IoT); anomaly detection; feature reduction; Naïve Bayesian (NB); Support Vector Machine (SVM); Decision Tree (DT); XGBoost; Random Forest (RF); K-Nearest Neighbor (K-NN)

I. INTRODUCTION

Detecting anomalies on the Internet of Things (IoT) is a major security issue that has been investigated and studied for centuries. The Internet of Things (IoT) involves several devices capable of processing, collecting, storing data, and communicating. The adoption of the IoT brought many innovations to industries, homes, and businesses, and undoubtedly, it has improved the quality of life.

Recently, the Internet of Things (IoT) has experienced quick growth in many specific applications. Also, IoT has become a driving force for the current technology revolution. IoT captures valuable data daily, allowing individuals or users to make critical decisions. There are many applications for IoT, such as healthcare, transportation, agriculture, and others. Also, there is no doubt that IoT devices have some limitations, such as CPU, memory, and low-energy storage. IoT devices comprise several interconnected sensors, actuators, and other devices [1],[2]. A lot of research expected tremendous growth in IoT. For example, cisco predicted an average of 75.3 billion linked devices by 2025 [3], [4].

IoT devices are extremely vulnerable to cyber-security threats targeting integrity and availability, and it is necessary to prevent cyber-security accidents. Thus, a Network Intrusion Detection System (NIDS) is needed. NIDS can detect any anomaly to protect the IoT network and the device. NIDS has the ability to monitor all traffic across the IoT network and acts as a first defense line. Also, NIDS can identify networks against intruders and suspicious activity. In addition, NIDS can examine and investigate the devices on the network [5], [6], [7].

Anomaly recognition can be divided into three categories based on the function of the training data stated as follows [2], [3], [4].:

Supervised Anomaly Detection: The normal and abnormal training datasets contain labeled cases. Thus, this methodology is about creating a predictive model for the abnormal and normal classes and then comparing both together.

Semi-supervised anomaly detection: The learning here involves only common cases of the class. Thus, anything that cannot be classified as usual is marked as abnormal.

Unsupervised anomaly detection: The training datasets will not be necessary for the methods. Thus, these methods indicate that regular cases are much more common than anomalies in the test data sets. Even if the hypothesis fails, this leads to a high false alarm rate for this practice.

This research proposes a new approach for IoT anomaly detection combined with artificial intelligence (AI) using detection mechanisms. The proposed approach combines three techniques for feature reduction (FR). Principal Component Analysis (PCA), Particle Swarm Optimization (PSO), and Gray Wolf Optimizer (GWO) were implemented for IoT cybersecurity.

Several research papers and surveys related to IoT have been proposed and published. Some of this research discusses security frameworks, privacy issues, security challenges, models, and tools [8], [9], [10], [11]. When Artificial Intelligence (AI) and the IoT combine, anomaly detection becomes more effective and reliable. AI-based anomaly detection can detect a wide range of threats. This paper will focus on machine learning (ML) algorithms and techniques for IoT security; the contribution of this paper can be reviewed in the following points:

- Using several machine learning algorithms for anomaly detection in IoT.

- Using The RT-IoT2022 proprietary dataset taken from a real-time IoT infrastructure.
- Using several up-to-date techniques for feature reduction, such as PSO, GWO, and PCA.

The rest of this paper is organized as follows: Section II will discuss previous studies related to this research. Section III will introduce machine learning algorithms for anomaly detection in an IoT environment. Section IV will discuss feature reduction and the dataset used in this paper. Section V will demonstrate experiments and results. Finally, the paper is concluded in Section VI.

II. Related Work

In this section, the authors will concentrate on some of the most prevailing solutions and demonstrate several research talks about IoT anomaly discovering methods and techniques.

Ayan Chatterjee [12] demonstrates a complete survey of IoT anomaly detection methods and applications. This survey examines 64 articles among publications between 2019 and 2021. The authors explain that they witnessed a shortage of IoT anomaly detection methodologies. Also, the authors present challenges and offer a new perspective where more research is needed. Besides that, the authors show that the publication of IoT detection is still in its early stages. Finally, they present no single best generic algorithm, but several methods are specific to a particular application.

Rafique Saida [13] presents a variety of literature on anomaly detection in IoT using both ML and DL. The authors discuss various challenges in anomaly detection in IoT infrastructure. Also, this research presents an increasing number of attacks. Finally, this work summarizes the most available literature and concludes that further development of the current detection technique is needed.

Maryam Khan [14] presents a machine learning anomaly detection model for cybersecurity using the Canadian Institute for Cybersecurity (CIC) dataset. The dataset presented in this work consists of 33 types of IoT attacks divided into seven categories. Techniques used in this work are Random Forest (RF), Adaptive Boosting (AB), Logistic Regression (LR), and Neural Network (NN). RF performs 99.55% accuracy.

Edwin Omo [15] presents several machine-learning algorithms for anomaly detection. The algorithms used in this work are isolation forest, One-Class SVM, Autoencoders, and Random Forest (RF). The study also examines the performance evaluation, efficiency process, and model selection methods. Besides that, the research sheds light on the main IoT aspects.

Adel Abusitta [16] presents a deep learning-powered anomaly recognition for IoT. The proposed model is designed based on a denoising autoencoder. Also, the denoising autoencoder allows the system to obtain features. Finally, experiments were conducted using the DS2OS traffic dataset.

Bhawana Sharma [17] provides an overview of anomaly detection using both machine learning and deep learning methods. This research addresses the key issues and challenges related to deep anomaly detection techniques in IoT.

Sahu [18] presents a supervised learning model to predict anomalies. This research uses several machine learning algorithms to predict anomalies on the 350K dataset. Two different approaches are used in this research. Also, classification algorithms were applied to the whole dataset in two different ways. The algorithms used were Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), and Artificial Neural Network (ANN). Finally, accuracy achieved an average of 99.4%.

Muhammad Inuwa [19] presents the comprehensive difficulties and challenges of cybersecurity in the context of IoT. This research uses machine learning (ML) methods to detect cyber anomalies within IoT systems. The algorithms used were Support Vector Machine (SVM), Artificial Neural Network (ANN), Decision Tree (DT), Logistic Regression (LR), and K-Nearest Neighbors (k-NN). Results demonstrate that ANN performs better than other models.

Abebe Diro [20] aim to provide a deep review of available works in anomaly discovery based on machine learning methods for IoT protection. This work indicates that blockchain-based anomaly detection can be effective. The future work of this research is to provide the implementation of a blockchain-based anomaly detection system.

A. Pathak [21] addresses the tampering of IoT security sensors in an office environment. Data is collected from real-life settings, and machine learning is applied to detect sensor tampering. The classification accuracy of the proposed model is 91.62%, with the lowest false positive rate.

Grace Hannah [22] explores several ML algorithms for anomaly discovery. This research explores supervised, unsupervised, and semi-supervised techniques. Also, the authors discuss the challenges and difficulties in implementing these algorithms in an IoT environment. Preprocessing techniques are examined. Besides that, this research demonstrates a case study on anomaly discovery in an IoT-based temperature monitoring system using a Gaussian Mixture Model (GMM). Precision, recall, and F1 score are used for evaluation.

III. Machine Learning Algorithms for IoT Anomaly Detection

Machine Learning (ML) algorithms can be used for different objectives and objectives and can impact every part of our lives. ML algorithms can be employed for pattern recognition, speech recognition, fraud detection, spam detection, phishing, and others. Also, machine learning procedures are used for prediction and classification, such as Decision Tree (DT), Random Forest (RF), Support Vector Machines (SVM), K-Nearest Neighbor (k-NN), Naïve Bayes Theorem (NB), K-Mean Clustering, Artificial Neural Network (ANN), and others. [23] [24], [25], [26], [27], [28], [29], [30].

Machine learning can be used for anomaly detection in IoT environments. The noun anomaly comes from the Greek word anomolia, meaning “irregular” which means that something is unusual if compared to similar things around it [31]. This

paper will introduce several machine learning algorithms in IoT anomaly detection. An anomaly in IoT is a pattern or series of samples in the IoT network that is different from a normal pattern. Also, anomaly detection can be defined as suspicious activity that falls outside normal patterns or behavior. Generally, anomalies can be divided into three categories: global outliers, contextual, and collective outliers [32], [33], [34], [35].

IV. Feature Reduction and Selection

Feature reduction or dimensionality reduction is the process of reducing the number of features in a dataset. Minimizing the number of features in a dataset is very important since the number of features could be huge. Also, Reducing the number of features could be useful and retaining the most helpful information. Besides that, reducing features means reducing processing time in CPU, memory usage, and other resources [26], [27], [28]. In other words, feature reductions mean assigning a weight to each feature to decide how important they are. On the other hand, Feature selection means selecting the most powerful features in the training phase. In summary, if feature reduction is done properly, this means that selecting a partial subset of features could be enough to represent all features. In this paper, the authors will use the Principal Component Analysis (PCA), Grey Wolf Optimizer (GWO), and Particle Swarm Optimizer (PSO) [28], [29], [30], [33].

A. Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a feature extraction method and is often used to reduce a higher-dimensional feature space to a lower-dimensional feature space. PCA is a statistical method that is employed to convert a set of possibly correlated variables into linearly unrelated variables known as principal components. The main objective of PCA is to capture the maximum variance available in the dataset with the fewest number of principal components. The transformation is defined mathematically as [25]:

$$\sum = \frac{1}{m-1}\sum_{i=1}^{m} (xi - \mu)(xi - \mu)^T \quad (1)$$

Where:

$\sum$: Covariance

xi: Data point

μ: Mean Vector

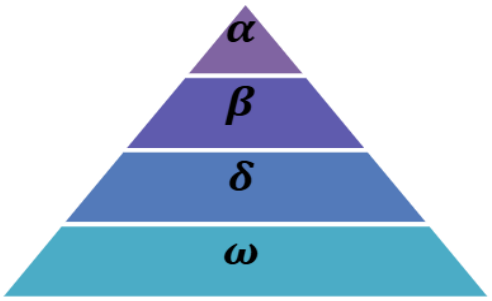
m: Number of data points.

B. Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO) is a powerful meta-heuristics optimization algorithm. This algorithm is inspired by natural swarm activities, such as that of fish and birds. PSO can be used to find the optimal values for specific parameters of a given system. In PSO, particles are moved according to a simple formula. Besides that, swarms move through the search space in order to find the optimal value. Every time a better position is found, movement is done. This process is repeated until finding the optimal solution [36], [37], [38], [39], [40].

C. Gray Wolf Optimizer (GWO)

The Gray Wolf Optimizer (GWO) algorithm is a population-based meta-heuristics algorithm that simulates the leadership hierarchy and hunting mechanism of grey wolves in nature Fig. 1 1[41] [42].



Hierarchy of grey wolf (dominance decreases from top down)

Fig. 1. Wolves' hierarchy.

Alpha wolves (α) wolf is the dominant, and his orders must be followed. Beta wolves (*β*) are subordinate wolves, which support alpha in decision-making. Delta wolves (*δ*) have to submit to alpha and beta. Omega wolves (*ω*) are the least important individuals in the pack [41], [42], [43], [44].

D. Dataset

The RT-IoT-2022 dataset, this dataset is proprietary and derived from a real-time IoT infrastructure. The RT-IoT-2022 provides comprehensive resources and a diverse range of IoT network machines. This dataset contains both normal and adversarial network behaviors. The RT-IoT-2022 contains 123117 instances and 83 features. Table I summarizes the RT-IoT-2022 dataset [45].

TABLE I. RT-IoT-2022 Dataset

No	Service	No of instances	Patterns
1	MQTT	4146	Normal Patterns
2	Thing_speak	8108	Normal Patterns
3	Wipro_bulb	253	Normal Patterns
	Total	**12507**	
4	ARP_poisioning	7750	Attacks patterns
5	DDOS_Slowloris	534	Attacks patterns
6	DOS_SYN_Hping	94659	Attacks patterns
7	Metasploit_Brute_Force_SSH	37	Attacks patterns
8	NMAP_FIN_SCAN	28	Attacks patterns
9	NMAP_OS_DETECTION	2000	Attacks patterns
10	NMAP_TCP_scan	1002	Attacks patterns
11	NMAP_UDP_SCAN	2590	Attacks patterns
12	NMAP_XMAS_T+REE_SCAN	2010	Attacks patterns
	Total	**110610**	

V. Experiments and Results

This section will display the authors' experiments and results. Also, it will display evaluation matrices and important features. This study also uses the Anaconda platform (Python) and MATLAB 2020a. Finally, experiments were done using a Dell Machine, 11th Gen -1165G7 @ 2.80GHz, RAM 32 GB, Windows 11.

A. Experimental Metrics

In machine learning, there are several criteria for evaluation, such as accuracy, precision, recall, F-measure, True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). This is demonstrated in Table II and Eq. (2) to (8).

TABLE II. Matrix of Confusion

		Prediction.	
		Normal.	Phishing
Act.	Normal.	x (TP)	y (FN)
	Phishing	z (FP)	w (TN)

$$\text{Accuracy}= \frac{TP+TN}{TP+FP+TN+FN} \quad (2)$$

$$TPR= x/(x+y) \quad (3)$$

$$FPR=z/(z+w) \quad (4)$$

$$FNR=y/(x+y) \quad (5)$$

$$P = TP/ (TP + FP) \quad (6)$$

$$R= TP/ (TP + FN) \quad (7)$$

$$F\text{-Measure}= 2*P*R / (P+R) \quad (8)$$

B. Experimental Results

In this section. The authors will demonstrate the results of feature reduction and selection using PCA, PSO, and GWO. PCA is evaluated using 10, 20, 30, 40, 50, 60 and 70 features.

Feature Reduction (FR) is done by using PCA, PSO and GWO. The PSO and GWO algorithms are executed independently for (10) iterations; then, the number and the name of the features are written. Then, the most important features of each algorithm are determined and picked for the classification stage. The testing part of the dataset represents only 20% of the dataset, meaning only 24624 instances.

Fig. 2 represents the results using Fine Tree without any feature reduction (PCA Disabled) using MATLAB 2020a. The figure demonstrated a good result, but the high number of features required extensive CPU and RAM resources.

The results of the experiments using feature reduction are demonstrated in Tables III-VIII. Most of the algorithm's performance is highly accepted. Also, FR techniques are very helpful since reducing the number of features from 83 to any number will reduce processing time and memory storage.

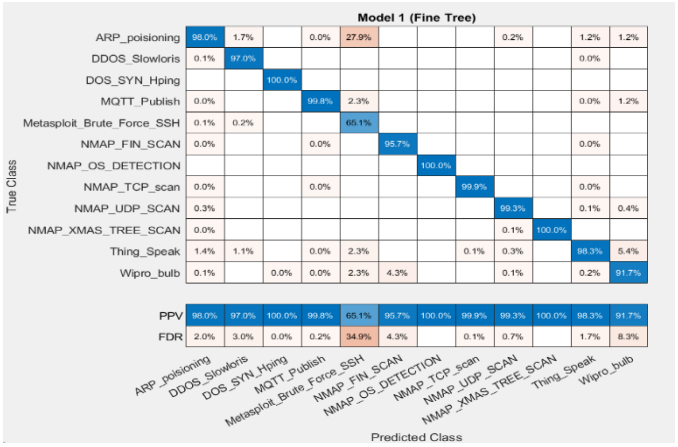


Fig. 2. Fine tree results (PCA disabled).

The above table shows that feature reduction using NB, (PCA-40) provides the best accuracy and optimal values of TP, TN, and FP compared with other types of feature reduction.

The above table shows that feature reduction using SVM, (PCA-50, PCA-60, and PCA-70) provides the best accuracy and optimal values of TP, TN, and FP compared with other types of feature reduction.

TABLE III. Naïve Bayesian Experiments

FR	TP	TN	FP	FN	Pr.	Re.	F-Me.	Acc.
PCA-10	21324	1052	1501	747	93.42%	96.62%	94.99%	90.87%
PCA-20	21281	995	1558	790	93.18%	96.42%	94.77%	90.46%
PCA-30	21211	963	1590	860	93.03%	96.10%	94.54%	90.05%
PCA-40	21352	1055	1498	719	93.44%	96.74%	95.06%	91.00%
PCA-50	21345	886	1667	726	92.76%	96.71%	94.69%	90.28%
PCA-60	21369	834	1719	702	92.55%	96.82%	94.64%	90.17%
PCA-70	21385	793	1760	686	92.40%	96.89%	94.59%	90.07%
GWO-55	21395	800	1750	679	92.44%	96.92%	94.63%	90.14%
PSO-58	21400	815	1740	669	92.48%	96.97%	94.67%	90.22%

TABLE IV. SUPPORT VECTOR MACHINE EXPERIMENTS

FR	TP	TN	FP	FN	Pr.	Re.	F-Me.	Acc.
PCA-10	21874	2485	68	197	99.69%	99.11%	99.40%	98.92%
PCA-20	22039	2524	29	32	99.87%	99.86%	99.86%	99.75%
PCA-30	22067	2552	2	3	99.99%	99.99%	99.99%	99.98%
PCA-40	22067	2552	2	3	99.99%	99.99%	99.99%	99.98%
PCA-50	22070	2551	2	1	99.99%	100.00%	99.99%	99.99%
PCA-60	22070	2552	1	1	100.00%	100.00%	100.00%	99.99%
PCA-70	22070	2552	1	1	100.00%	100.00%	100.00%	99.99%
GWO-55	22040	2520	35	29	99.84%	99.87%	99.86%	99.74%
PSO-58	22041	2519	36	28	99.84%	99.87%	99.86%	99.74%

TABLE V. DECISION TREE EXPERIMENTS

FR	TP	TN	FP	FN	Pr.	Re.	F-Me.	Acc.
PCA-10	22053	2538	15	18	99.93%	99.92%	99.93%	99.87%
PCA-20	22052	2542	11	19	99.95%	99.91%	99.93%	99.88%
PCA-30	22053	2537	16	18	99.93%	99.92%	99.92%	99.86%
PCA-40	22056	2538	15	15	99.93%	99.93%	99.93%	99.88%
PCA-50	22055	2532	21	16	99.90%	99.93%	99.92%	99.85%
PCA-60	22056	2537	16	15	99.93%	99.93%	99.93%	99.87%
PCA-70	22052	2538	15	19	99.93%	99.91%	99.92%	99.86%
GWO-55	22050	2536	17	21	99.92%	99.90%	99.91%	99.85%
PSO-58	22048	2534	19	23	99.91%	99.90%	99.90%	99.83%

The above table shows that feature reduction using DT, (PCA-20, PCA-40) provides the best accuracy and optimal values of TP, TN, and FP compared with other types of feature reduction.

The above table shows that feature reduction using XGBoost, (GWO-55, PSO-58) provides the best accuracy and optimal values of TP, TN, and FP compared with other types of feature reduction.

The above table shows that feature reduction using RF, (PCA-30) provides the best accuracy and optimal values of TP, TN, and FP compared with other types of feature reduction.

TABLE VI. XGBOOST EXPERIMENTS

FR	TP	TN	FP	FN	Pr.	Re.	F-Me.	Acc.
PCA-10	22064	2538	15	7	99.93%	99.97%	99.95%	99.91%
PCA-20	22070	2543	10	1	99.95%	100.00%	99.98%	99.96%
PCA-30	22069	2544	9	2	99.96%	99.99%	99.98%	99.96%
PCA-40	22069	2547	6	2	99.97%	99.99%	99.98%	99.97%
PCA-50	22069	2545	8	2	99.96%	99.99%	99.98%	99.96%
PCA-60	22069	2546	7	2	99.97%	99.99%	99.98%	99.96%
PCA-70	22069	2547	6	2	99.97%	99.99%	99.98%	99.97%
GWO-55	22070	2548	4	2	99.98%	99.99%	99.99%	99.98%
PSO-58	22070	2549	3	2	99.99%	99.99%	99.99%	99.98%

TABLE VII. RANDOM FOREST EXPERIMENTS

FR	TP	TN	FP	FN	Pr.	Re.	F-Me.	Acc.
PCA-10	22064	2533	20	7	99.91%	99.97%	99.94%	99.89%
PCA-20	22064	2538	15	7	99.93%	99.97%	99.95%	99.91%
PCA-30	22066	2542	11	5	99.95%	99.98%	99.96%	99.94%
PCA-40	22064	2537	16	7	99.93%	99.97%	99.95%	99.91%
PCA-50	22063	2538	15	8	99.93%	99.96%	99.95%	99.91%
PCA-60	22063	2534	19	8	99.91%	99.96%	99.94%	99.89%
PCA-70	22061	2537	16	10	99.93%	99.95%	99.94%	99.89%
GWO-55	22060	2541	14	9	99.94%	99.96%	99.95%	99.91%
PSO-58	22061	2540	12	11	99.95%	99.95%	99.95%	99.91%

TABLE VIII. K-NEAREST NEIGHBOR EXPERIMENTS

FR	TP	TN	FP	FN	Pr.	Re.	F-Me.	Acc.
PCA-10	22060	2535	18	11	99.92%	99.95%	99.93%	99.88%
PCA-20	22067	2533	20	4	99.91%	99.98%	99.95%	99.90%
PCA-30	22068	2544	9	3	99.96%	99.99%	99.97%	99.95%
PCA-40	22069	2544	9	2	99.96%	99.99%	99.98%	99.96%
PCA-50	22066	2544	9	5	99.96%	99.98%	99.97%	99.94%
PCA-60	22066	2544	9	5	99.96%	99.98%	99.97%	99.94%
PCA-70	22068	2542	11	3	99.95%	99.99%	99.97%	99.94%
GWO-55	22070	2540	9	5	99.96%	99.98%	99.97%	99.94%
PSO-58	22072	2538	11	3	99.95%	99.99%	99.97%	99.94%

The above table shows that feature reduction using KNN, (PCA-40) provides the best accuracy and optimal values of TP, TN, and FP compared with other types of feature reduction.

As demonstrated in the above tables. The performance of machine learning algorithms with feature reduction techniques is highly acceptable. Having too many processing features makes the ML model complex. There is no doubt that reducing the number of features has a lot of advantages, such as reducing time, improving computational efficiency, and preventing overfitting.

Fig. 3 and Fig. 4 show the accuracy of the machine-learning algorithms used in this paper. The figures demonstrated that the accuracy results are highly acceptable, especially with the number of features selected.

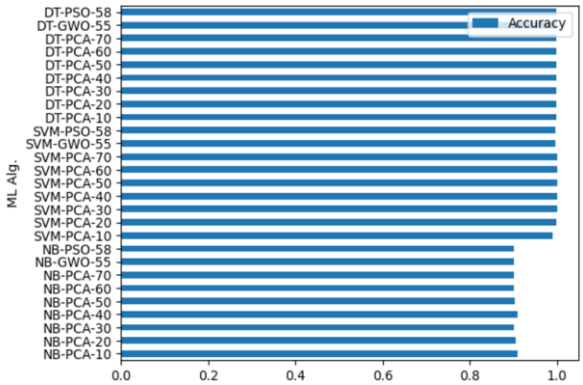


Fig. 3. DT, SVM, and NB algorithms.

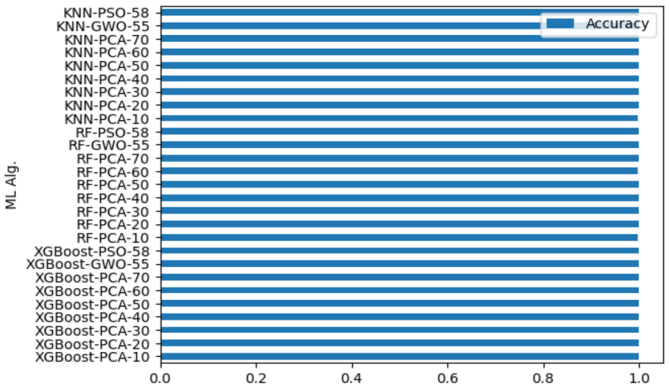


Fig. 4. KNN, RF, and XGBoost algorithms.

VI. CONCLUSION AND FUTURE WORKS

The Internet of Things (IoT) or "Smart Objects" refers to physical devices embedded with sensors, software, and network connectivity. IoT devices can be used in smart homes, smart cities, and complex industries. IoT enables smart devices to communicate with each other and with the Internet. In the last decades, IoT devices have faced several threats and difficulties. This paper demonstrates several machine learning algorithms used in anomaly detection in IoT environments. This paper also uses PCA, GWO, and PSO as feature-reduction techniques. Several criteria are used for evaluation, such as precision, recall, F-measure, and accuracy. Most of the algorithms show excellent performance except the Naïve Bayesian. The support vector machines (SVM) show the best results with 99.99 accuracy with PCA-60 and PCA-70.

ACKNOWLEDGMENT

The researchers would like to provide a special thanks to the editor and reviewers for their time in reviewing the manuscript and overall suggestions to improve the manuscript. In addition, they are very grateful to WISE University.

REFERENCES

[1] E. Gyamfi and A. Jurcut, "Intrusion Detection in Internet of Things Systems: A Review on Design Approaches Leveraging Multi-Access Edge Computing, Machine Learning, and Datasets," *Sensors*, vol. 22, no. 10, 2022, doi: 10.3390/s22103744.

[2] M. Stoyanova, Y. Nikoloudakis, S. Panagiotakis, E. Pallis, and E. K. Markakis, "A Survey on the Internet of Things (IoT) Forensics: Challenges, Approaches, and Open Issues," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 1191–1221, 2020, doi: 10.1109/COMST.2019.2962586.

[3] A. Yastrebova, R. Kirichek, Y. Koucheryavy, A. Borodin, and A. Koucheryavy, "Future Networks 2030: Architecture & Requirements," *2018 10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, pp. 1–8, 2018, [Online]. Available: https://api.semanticscholar.org/CorpusID:59601484

[4] A. Jurcut, T. Niculcea, P. Ranaweera, and N.-A. Le-Khac, "Security Considerations for Internet of Things: A Survey," *CoRR*, vol. abs/2006.10591, 2020, [Online]. Available: https://arxiv.org/abs/2006.10591

[5] I. Butun, S. D. Morgera, and R. Sankar, "A Survey of Intrusion Detection Systems in Wireless Sensor Networks," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 266–282, 2014, doi: 10.1109/SURV.2013.050113.00191.

[6] M. A. Alsoufi *et al.*, "Anomaly-Based Intrusion Detection Systems in IoT Using Deep Learning: A Systematic Literature Review," *Applied Sciences*, vol. 11, no. 18, 2021, doi: 10.3390/app11188383.

[7] L. Njilla, L. Pearlstein, X.-W. Wu, A. Lutz, and S. Ezekiel, "Internet of Things Anomaly Detection using Machine Learning," in *2019 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, 2019, pp. 1–6. doi: 10.1109/AIPR47015.2019.9174569.

[8] M. Ammar, G. Russello, and B. Crispo, "Internet of Things: A survey on the security of IoT frameworks," *Journal of Information Security and Applications*, vol. 38, pp. 8–27, 2018, doi: https://doi.org/10.1016/j.jisa.2017.11.002.

[9] Y. Yang, L. Wu, G. Yin, L. Li, and H. Zhao, "A Survey on Security and Privacy Issues in Internet-of-Things," *IEEE Internet Things J*, vol. 4, no. 5, pp. 1250–1258, 2017, doi: 10.1109/JIOT.2017.2694844.

[10] M. Sain, Y. J. Kang, and H. J. Lee, "Survey on security in Internet of Things: State of the art and challenges," in *2017 19th International Conference on Advanced Communication Technology (ICACT)*, 2017, pp. 699–704. doi: 10.23919/ICACT.2017.7890183.

[11] R. Benabdessalem, M. Hamdi, and T. Kim, "A Survey on Security Models, Techniques, and Tools for the Internet of Things," *2014 7th International Conference on Advanced Software Engineering and Its Applications*, pp. 44–48, 2014, [Online]. Available: https://api.semanticscholar.org/CorpusID:18825070

[12] A. Chatterjee and B. S. Ahmed, "IoT anomaly detection methods and applications: A survey," *Internet of Things*, vol. 19, p. 100568, 2022, doi: https://doi.org/10.1016/j.iot.2022.100568.

[13] S. H. Rafique, A. Abdallah, N. S. Musa, and T. Murugan, "Machine Learning and Deep Learning Techniques for Internet of Things Network Anomaly Detection—Current Research Trends," *Sensors*, vol. 24, no. 6, 2024, doi: 10.3390/s24061968.

[14] M. M. Khan and M. Alkhathami, "Anomaly detection in IoT-based healthcare: machine learning for enhanced security," *Sci Rep*, vol. 14, no. 1, p. 5872, 2024, doi: 10.1038/s41598-024-56126-x.

[15] E. Omol, L. Mburu, and D. Onyango, "Anomaly Detection In IoT Sensor Data Using Machine Learning Techniques For Predictive Maintenance In Smart Grids," *International Journal of Science, Technology & Management*, vol. 5, no. 1, pp. 201–210, Jan. 2024, doi: 10.46729/ijstm.v5i1.1028.

[16] A. Abusitta, G. H. de Carvalho, O. Abdel Wahab, T. Halabi, B. C. M. Fung, and S. Al Mamoori, "Deep learning-enabled anomaly detection for IoT systems," *SSRN Electron. J.*, 2022.

[17] B. Sharma, L. Sharma, and C. Lal, "Anomaly Detection Techniques using Deep Learning in IoT: A Survey," in *2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE)*, 2019, pp. 146–149. doi: 10.1109/ICCIKE47802.2019.9004362.

[18] N. K. Sahu and I. Mukherjee, "Machine Learning based anomaly detection for IoT Network: (Anomaly detection in IoT Network)," in *2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184)*, 2020, pp. 787–794. doi: 10.1109/ICOEI48184.2020.9142921.

[19] M. M. Inuwa and R. Das, "A comparative analysis of various machine learning methods for anomaly detection in cyber attacks on IoT networks," *Internet Things*, vol. 26, p. 101162, 2024, [Online]. Available: https://api.semanticscholar.org/CorpusID:268402373

[20] A. Diro, N. Chilamkurti, V.-D. Nguyen, and W. Heyne, "A Comprehensive Study of Anomaly Detection Schemes in IoT Networks Using Machine Learning Algorithms," *Sensors*, vol. 21, no. 24, 2021, doi: 10.3390/s21248320.

[21] A. K. Pathak, S. Saguna, K. Mitra, and C. Åhlund, "Anomaly Detection using Machine Learning to Discover Sensor Tampering in IoT Systems," in *ICC 2021 - IEEE International Conference on Communications*, 2021, pp. 1–6. doi: 10.1109/ICC42927.2021.9500825.

[22] Dr. J. G. Hannah, Dr. A. S. D. Murthy, Dr. G. Kalnoor, M. Vetriselvan, and Dr. M. S. Nidhya, "Machine Learning Algorithms for Anomaly Detection in IoT Networks," *Migration Letters*, vol. 20, no. S13, pp. 560–565, Dec. 2023, [Online]. Available: https://migrationletters.com/index.php/ml/article/view/6728

[23] R. Mahajan and I. Siddavatam, "Phishing website detection using machine learning algorithms," *Int. J. Comput. Appl.*, vol. 181, no. 23, pp. 45–47, Oct. 2018.

[24] D. T. Mosa, M. Y. Shams, A. A. Abohany, E.-S. M. El-kenawy, and M. Thabet, "Machine Learning Techniques for Detecting Phishing URL Attacks," *Computers, Materials and Continua*, vol. 75, no. 1, pp. 1271–1290, 2023, doi: https://doi.org/10.32604/cmc.2023.036422.

[25] M. Altin and A. Cakir, "Exploring the influence of dimensionality reduction on anomaly detection performance in multivariate time series," 2024.

[26] F. Abbasi, M. Naderan, and S. E. Alavi, "Anomaly detection in Internet of Things using feature selection and classification based on Logistic Regression and Artificial Neural Network on N-BaIoT dataset," in *2021 5th International Conference on Internet of Things and Applications (IoT)*, 2021, pp. 1–7. doi: 10.1109/IoT52625.2021.9469605.

[27] A. G. Ayad, N. A. Sakr, and N. A. Hikal, "A hybrid approach for efficient feature selection in anomaly intrusion detection for IoT networks," *J Supercomput*, vol. 80, no. 19, pp. 26942–26984, 2024, doi: 10.1007/s11227-024-06409-x.

[28] A. H. Mohammad, "Intrusion Detection Using a New Hybrid Feature Selection Model," *Intelligent Automation & Soft Computing*, vol. 30, no. 1, pp. 65–80, 2021, doi: 10.32604/iasc.2021.016140.

[29] A. Mandadi, S. Boppana, V. Ravella, and R. Kavitha, "Phishing Website Detection Using Machine Learning," in *2022 IEEE 7th International conference for Convergence in Technology (I2CT)*, 2022, pp. 1–4. doi: 10.1109/I2CT54291.2022.9824801.

[30] A. H. Mohammad, T. Alwada'n, O. Almomani, S. Smadi, and N. ElOmari, "Bio-inspired Hybrid Feature Selection Model for Intrusion Detection," *Computers, Materials and Continua*, vol. 73, no. 1, pp. 133–150, 2022, doi: https://doi.org/10.32604/cmc.2022.027475.

[31] "https://www.vocabulary.com/dictionary/anomalyWebsite".

[32] A. H. Mohammad, S. Smadi, and T. Alwada'n, "Email Filtering Using Hybrid Feature Selection Model," *CMES - Computer Modeling in Engineering and Sciences*, vol. 132, no. 2, pp. 435–450, 2022, doi: https://doi.org/10.32604/cmes.2022.020088.

[33] S. Alrefaai, G. Özdemir, and A. Mohamed, "Detecting Phishing Websites Using Machine Learning," in *2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, 2022, pp. 1–6. doi: 10.1109/HORA55278.2022.9799917.

[34] M. Tahboush, A. Hamdan, F. Alzobi, M. Husni, and M. Adawy, "NTDA: The mitigation of denial of service (DoS) cyberattack based on network traffic detection approach," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 3, 2024.

[35] A. Hamdan, M. Tahboush, M. Adawy, T. Alwada'n, S. Ghwanmeh, and M. Husni, "Phishing detection using grey wolf and particle swarm optimizer," *Int. J. Electr. Comput. Eng. (IJECE)*, vol. 14, no. 5, p. 5961, Oct. 2024.

[36] J. A. W. A. S. Saeed M. Alshahrani Nayyar Ahmed Khan, "URL Phishing Detection Using Particle Swarm Optimization and Data Mining," *Computers, Materials & Continua*, vol. 73, no. 3, pp. 5625–5640, 2022, doi: 10.32604/cmc.2022.030982.

[37] W. Ali and S. Malebary, "Particle Swarm Optimization-Based Feature Weighting for Improving Intelligent Phishing Website Detection," *IEEE Access*, vol. 8, pp. 116766–116780, 2020, doi: 10.1109/ACCESS.2020.3003569.

[38] F. Marini and B. Walczak, "Particle swarm optimization (PSO). A tutorial," *Chemometrics and Intelligent Laboratory Systems*, vol. 149, pp. 153–165, 2015, doi: https://doi.org/10.1016/j.chemolab.2015.08.020.

[39] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, 1995, pp. 1942–1948 vol.4. doi: 10.1109/ICNN.1995.488968.

[40] K. Ishaque, Z. Salam, M. Amjad, and S. Mekhilef, "An Improved Particle Swarm Optimization (PSO)–Based MPPT for PV With Reduced Steady-State Oscillation," *IEEE Trans Power Electron*, vol. 27, no. 8, pp. 3627–3638, 2012, doi: 10.1109/TPEL.2012.2185713.

[41] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014, doi: https://doi.org/10.1016/j.advengsoft.2013.12.007.

[42] J.-S. Wang and S.-X. Li, "An Improved Grey Wolf Optimizer Based on Differential Evolution and Elimination Mechanism," *Sci Rep*, vol. 9, no. 1, p. 7181, 2019, doi: 10.1038/s41598-019-43546-3.

[43] E. M. R. Devi and R. C. Suganthe, "Feature selection in intrusion detection grey wolf optimizer," *Asian J. Res. Soc. Sci. Humanit.*, vol. 7, no. 3, p. 671, 2017.

[44] Q. M. Alzubi, M. Anbar, Z. N. M. Alqattan, M. A. Al-Betar, and R. Abdullah, "Intrusion detection system based on a modified binary grey wolf optimisation," *Neural Comput Appl*, vol. 32, pp. 6125–6137, 2019, [Online]. Available: https://api.semanticscholar.org/CorpusID:128021795

[45] B. & N. R. (2023). R.-I. [Dataset]. U. M. L. Repository. S., "https://archive.ics.uci.edu/dataset/942/rt-iot2022".