International Journal Of Engineering Research And Development
e- ISSN: 2278-067X, p-ISSN: 2278-800X, www.ijerd.com
Volume 20, Issue 3 (March. 2024), PP. 91-107

Methods, Potentials and Challenges of Machine Learning Based Artificial Intelligence Systems in Cyber Security

Chidiebere C. Nweke[1], Paulinus C. Eze[2*], Isaac A. Ezenugu[2], Victor N. Okorogu[3]

[1]*Department of Computer Science, University of Lagos, AkokaYaba, Lagos, Nigeria*
[2]*Department of Electrical and Electronic Engineering, Imo State University, Owerri, Nigeria*
[3]*Department of Electronic and Computer Engineering, NnamdiAzikiweUniversity,Awka, Nigeria*
**Corresponding Author*

ABSTRACT
The need for a solution to address the growing and changing pattern of cyber-attacks has increased in recent times. Conventional systems such as signature-base detection systems, manual analysis, and rule-base systems have proven to be very weak against new methods of attacks or threats. Hence, research in cyber security recent times has largely focused on the use of artificial intelligence (AI), which is considered to provide robust and adaptive security for networked public and privative infrastructures connected via internet. In this paper, some recent methods in literature using AI models that are mostly machine learning (ML) have been reviewed considering the approaches used in developing them and possible applications. In addition, the potentials of AI systems such as adaptation (or ability to update), monitoring capability, defence against attack, response, and prediction were highlighted. Furthermore, challenges of the AI technologies, which include lack of security guarantee, moral concerns, financial constraint, the need for large volume of data, as tool for cybercriminals, and lastly as a tool against opponent were mentioned. In the next study, the focus will be to examine solutions designed to address the issues in the use of AI systems in cyber security.

Keywords: Artificial intelligence, Cyber security, Machine learning, Potentials and challenges

---

Date of Submission: 10-03-2024 *Date of Acceptance: 23-03-2024*

I. INTRODUCTION

Cyber security is regarded as any technique, measure or an act taken to detect, prevent, or reduce the impact of cyber threats. It is therefore a strategy designed to ensure that systems, applications, computing devices, sensitive data and financial assets of individuals and organisations (public or private enterprises) are protected against malicious attacks involving the use of computer viruses, sophisticated and costly ransom ware [1]. Cyber security has become an issue of concern because the contemporary society, which basically carries out its activities using computer devices (information computer and technology (ICT) based electronic systems) and software applications, is increasingly faced with a devastating threat or action emanating from the activities of cyber criminals. Hence, there is no individual or organisation that is directly or indirectly involved in computer related activities that is immune or shielded against cyber threats. This is because even personal computer (P.C.) of individuals or the computer systems of corporate organisations can be rendered useless or out of services by the action of simple or sophisticated virus attack on personal files or database.

In this modern-day, which is usually regarded as digital age, every individual or organisation, regardless of the location or nature of business (small and medium enterprises, SMEs), or large corporations such as banks, or government agencies, is prone to the menace of system attack. This means that cyber security is everybody's business because from government to internet banking the global society is under networked technology [2]. Immediate attention is required for ensuring the integrity of any data because of the global impact of cyber security. Hence, cyber security is described in [3] as an attempt by users to ensure that personal and professional information is kept undamaged from the attacks on the internet. The fact that cyber threat is not limited to any sector or area of human activities, means that it is now a worldwide problem and as such the foremost task of cyber security application is to protect networks, computers, software, and database from unauthorised access and loss.

Prior to the emergence artificial intelligence (AI), cyber security has relied on traditional approach such as signature-base detection systems (incoming traffic is compared to a database of known malicious code or threat), manual analysis (security alerts and logs are manually investigated by security analysts by looking for patterns or indicators of a security breach), and rule-base systems (setting up rules or polices that define the acceptable behaviour on a network) [4]. The traditional approaches are characterized by time consumption, drain on resources, and inadequate protection (for example in the easy bypassing of the signature-based detection system by cybercriminals) [4]. Therefore, in recent times, to ensure reduction in process timing, minimised resources, and provide adequate protection of public and corporate facilities, strong and sophisticated techniques have been advocated and applied for efficient security especially in cyberspace. These modern techniques, usually AI based method, are used for mining data, identifying pattern, and predicting future events [5]. The application of AI in cyber security is on the increase because of the evolving patterns of threats associated with malicious acts of cybercrime that require techniques that provide better and robust detection, protection, and recovering strategies. The AI technology for cyber security comes in different forms such as artificial immune systems, intelligent agent, expert systems, and neural networks and this provides unique advantages with respect to its distinctive applicability to different cases of cyber threats [6].

The transformation from traditional approach to AI based method in cyber security can be attributed to advancement and unprecedented change in digital space, which has resulted in the quest to keep pace with recent advances and maintain a safe and secure computer networked system. This can be associated with the way ICT devices have evolved, where in early computer systems, if any security strategies were available; it was minimal and only involved basic authentication and simple logging [7]. Also, before the introduction of AI, the traditional approach was largely ineffective in tackling new and unknown threats (that is, it was not adaptive), and there is the possibility of it generating a high number of false positive that could result in draining of resources. These traditional security systems use such tools as IDS, IPS, and Firewalls for protection against simple attacks that repeatedly use the same tactics and tools, and are independently implemented such that no contact exists between them that will cause firewall for instance to block intrusion detected in an IDS [8, 9].

The approach to cyber security has completely changed since the introduction of AI, which offers advance and robust technique to cyber threats detection and prevention. Generally, because of the effectiveness of the AI technique in cyber security, it is increasingly being sort after and adopted by many public and private entities as a vital cyber security management tool. The AI based cyber security strategies are designed to be adaptive such that they continuously learn and adapt with proactive and predictive potential to detect anomalies, which distinguish them from traditional approach. With the AI revolutionising the way cyber security is managed and the sophistication and the robustness it offers, this study is designed to review AI methods that basically use machine learning (ML) models that have been implemented and reported in literature and the applications, possibilities, and issues facing the system.

II. ARTIFICIAL INTELLIGENCE METHOD

Some of the AI based techniques that have been proposed in literature for cyber security purposes are presented in this section. The focus is on AI solutions that have been basically developed using ML in recent past years (2020 and beyond) in order to limit the scope of the review and concisely present the section. The discussed methods are not limited to any field or area of human endeavour, but have been considered on account of different approaches applied for achieving specific purpose in cyber security applications.

Considering the multiple vulnerabilities threats faced by online social network users, an AI technique that used ML algorithm to predict potential attack or threat on social network users, which was primarily designed to examine the usual perception and continuous reactions taken towards socially engineered threats daily was presented in [10]. Unsupervised learning approach was used to identify unknown patterns in data and features. The ML based solution was used to demonstrate how an AI based model of user’s behavioural characteristics, perceptions, and socio-emotions that would aid in identifying features, which in addition become significant in determining the vulnerability of a user to social engineer threats and attacks on online social network. A comparative analysis of the theories and the findings from primary data used in the study revealed that certain behaviours of Facebook users are threat to other users. The results in the study were used to validate the ability of ML method as an AI-algorithm, to identify scamming messages and scammers. Figure 1 shows the structure of the model using k-means clustering where data is collected, processed, features identified and extracted, semantics applied on the data for more precise extraction, appropriate data classification performed based on the algorithm predetermined models, and eventually, pertinent action taken by system network management personnel or an advanced system protection integration module automatically prevent any threat or attack.

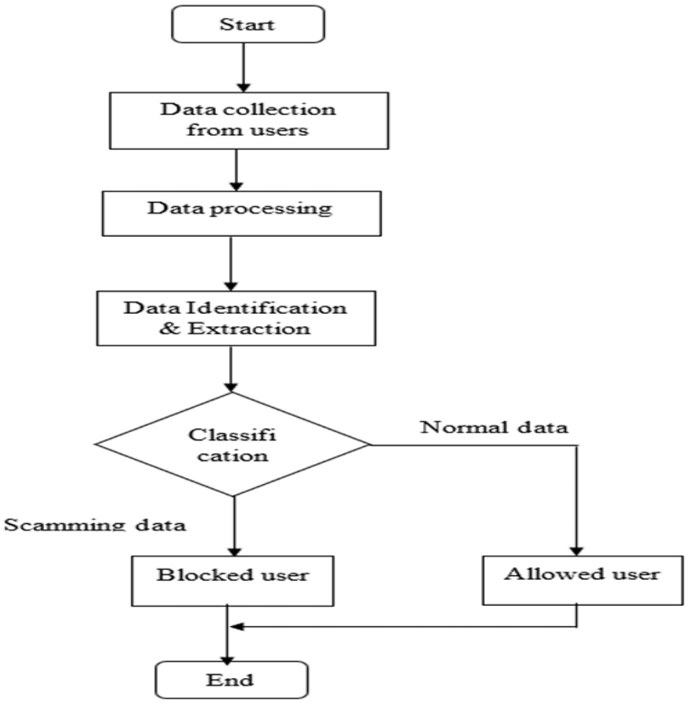


Figure1: Flowchart of ML algorithm for threat prediction on social media based on user's behaviour [10]

In the software development industry, vulnerabilities of the software security is a significant concern and as a result, data sources that could provide software development team the insight to proactively act upon software security vulnerability were explored by conducting exploration on software development industry landscape, software development eco-system landscape, and software system customer landscape in [11]. That is the sources of information are integrated data generated in industry, within software development processes and the conservation between customer and software development team. The possibility of integrating software development practices and data science practices to improve the control of software system vulnerabilities was explored. A smart knowledge management system that combines the information processed across module into a central system was built. The AI methods were used to identify security insights leveraging data sources across the modules of software landscape, customer landscape, and industry landscape. A data classifier that could refine the data to identify security-related information and in addition offer granular data classification in the customer conservation module was built. Two-level processing of information was achieved using deep learning algorithms. An integrated approach combining the risk assessment and threat modelling methods using ML algorithms was proposed for industry landscape module. Also, it was reported in the study that proactive management of software development activities can leverage the proposed central intelligence system. The ML and deep learning algorithms were developed for modelling data and learn from across modules. The model of the proposed AI knowledge processing system for optimizing security of software system is shown in Figure 2.

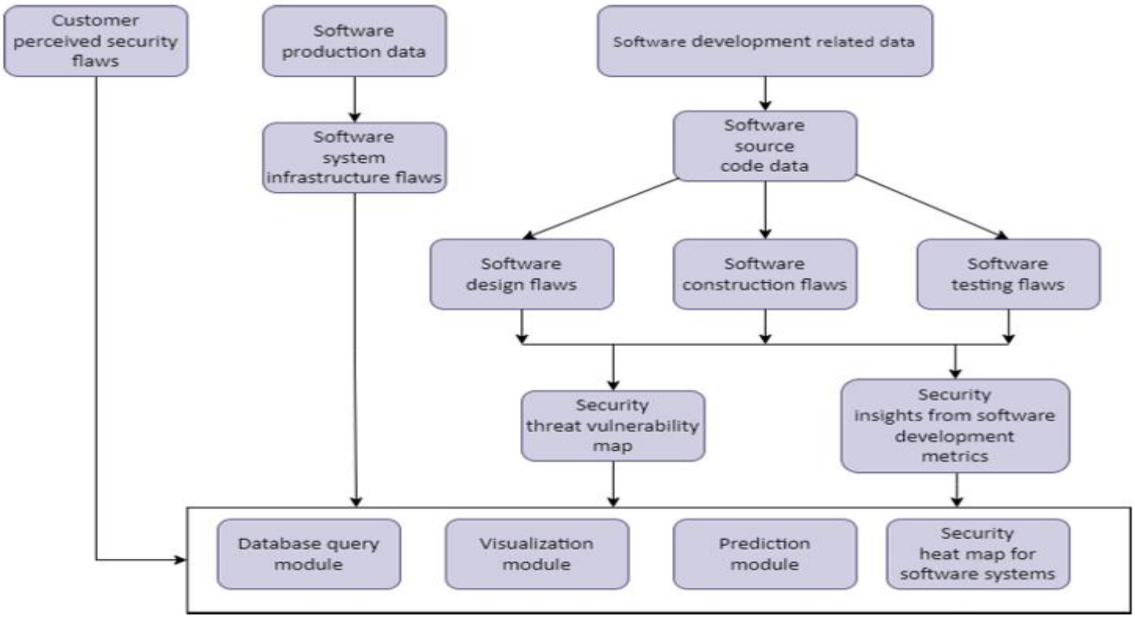

Customer perceived security flaws
Software production data
Software development related data
Software system infrastructure flaws
Software source code data
Software design flaws
Software construction flaws
Software testing flaws
Security threat vulnerability map
Security insights from software development metrics
Database query module
Visualization module
Prediction module
Security heat map for software systems

Figure 2: Structure of AI knowledge processing system for optimizing security software system [11]

As a result of the inadequacy of published generic standards and guidelines regarding lack of description of technologies or methods that could be used to have substantial knowledge about cyber security and how to implement it, a vision model based on Quantum Artificial Intelligence (QAI) that generates secure software development (SSD) rules to educate and train developers and tester during the different phases of Software Development Life Cycle (SDLC) was proposed in [12]. Data from industry standards, vulnerability information, and proprietary and historical data were used to train the QAI algorithms in order to create rules that could be quickly adapted by developers and tester. The proposed system was called Industrial Cyber Security Education with Quantum Artificial Intelligence (ICSEQAI), which is a high-level abstraction of learning framework to provide assistance during SDLC for stakeholders. The impact of the proposed learning system was tested in cyber security education in the automotive industry. The learning framework can be efficiently used to continuously train and educate company resources such as developers, testers, and architects on cyber security proficiencies, and was considered as an interesting method for solving the problem of cyber security education in academics and industry. The block diagram of the ICSEQAI vision model is shown in Figure 3.

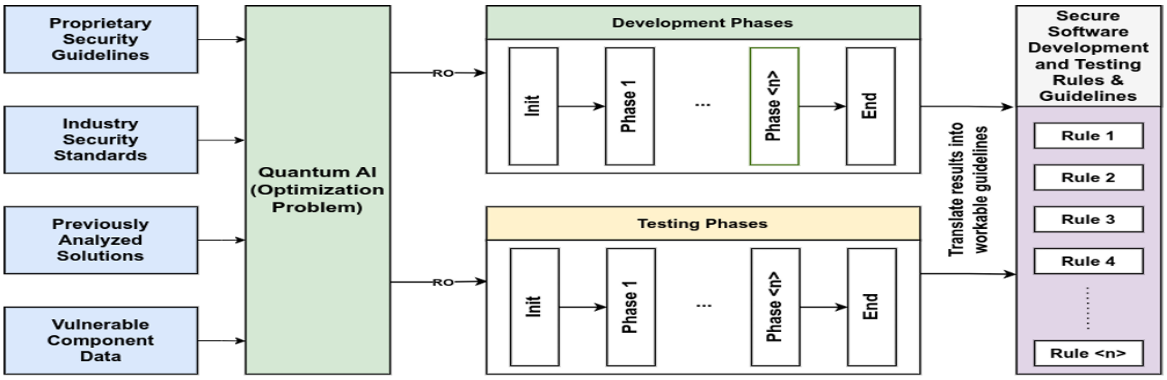

Proprietary Security Guidelines
Industry Security Standards
Previously Analyzed Solutions
Vulnerable Component Data
Quantum AI (Optimization Problem)
RO
RO
Development Phases
Init
Phase 1
...
Phase <n>
End
Testing Phases
Init
Phase 1
...
Phase <n>
End
Translate results into workable guidelines
Secure Software Development and Testing Rules & Guidelines
Rule 1
Rule 2
Rule 3
Rule 4
Rule <n>

Figure 3: Block diagram of ICSEQAI vision model to assist developers and testers during SDLC [12]

An AI based network intrusion detection system for Supervisory Control and Data Acquisition (SCADA) systems was presented in [13] to address number of vulnerabilities in these systems that have gradually increased and could be exploited because of the incorporation of advanced technology such as Internet of things (IoT). The system was designed to protect smart Industrial Control Systems (ICSs) against conventional and SCADA-specific network-based attacks. This is because ICSs are crucial components of modern critical infrastructures such as nuclear and thermal power plants, power grids, oil refineries, and water treatment facilities. A number of machine learning such as k-nearest neighbour (KNN) and random forest (RF), and deep learn such as gated recurrent unit (GRU) and convolution neural network plus gated recurrent unit

(GRU) (CNN-GRU), which are classification techniques, were empirically evaluated for the network intrusion detection system. The proposed algorithms were examined using genuine SCADA traffic datasets. The tests indicated the network intrusion detection in SCADA systems yielded high detection accuracy and offered the ability to tackle new emerging threats. With KNN and RF algorithms the system achieved a near perfect model score accuracy of 99.99% against 99.98% with CNN-GRU algorithm using WUSTL-IIoT-2018 dataset. A model score accuracy of > 99.75% was provided by RF and GRU based network intrusion detection system using WUSTL-IIoT-20121 dataset. Statistical analysis of KNN, RF and CNN-GRU techniques successfully achieved an R^2> 99%, which showed that the proposed system was able to tackle new or unknown threats in ICSs cyber space. The block diagram of the SCADA based cyber security system framework is shown in Figure 4.

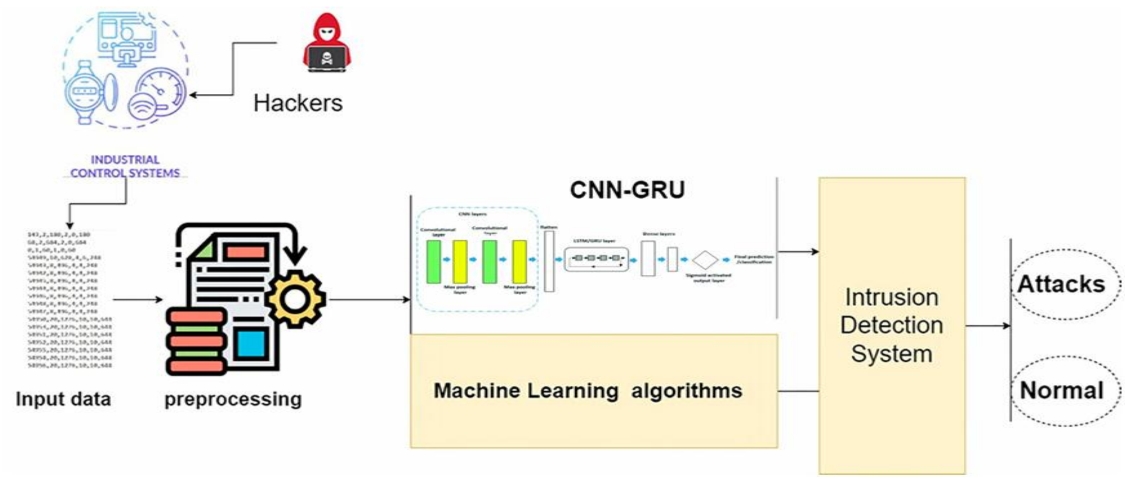
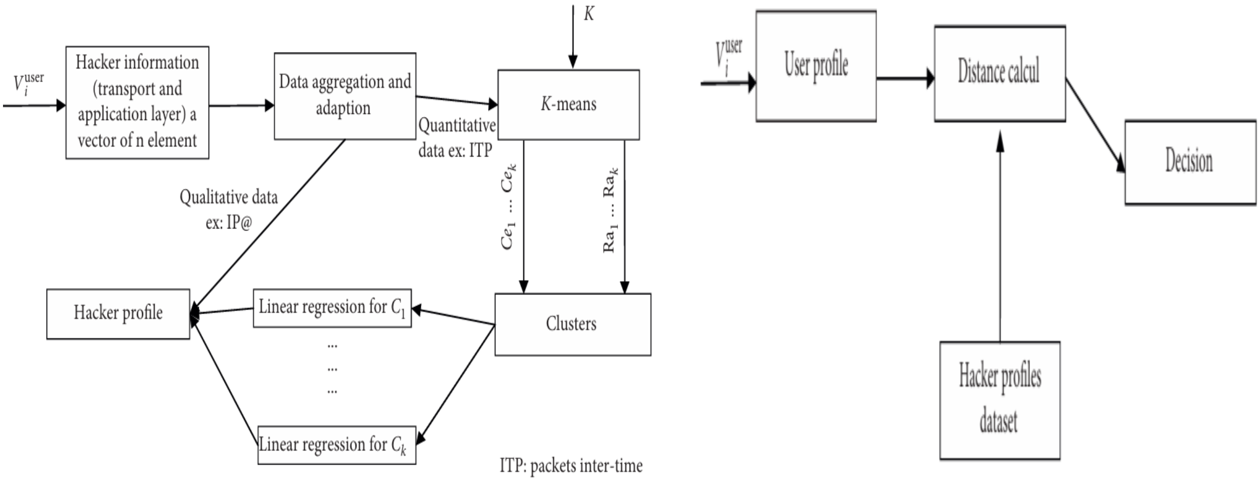

Hackers
INDUSTRIAL CONTROL SYSTEMS
Input data
preprocessing
CNN-GRU
Machine Learning algorithms
Intrusion Detection System
Attacks
Normal

Figure 4: Block diagram of SCADA cyber security system based on AI [13]

Since the honeypot technique can be used passively as an information system and also to reinforce the traditional defence systems against future attacks, it has been used together with ML to design a smart agent for preventing and predicting of cyber-attack in [8]. The honeypot systems and ML techniques are combined as tools for gathering information, analysis, and prediction of threats so as to guarantee the security of companies' networks. The system is such that for every interaction with a suspicious flow, collected information vector (regarding user activities) such as IP, logging, packet length are stored by the honeypot in a database so as to create a user profile. Submission of the collected data is made to a combination of algorithms. The k-means algorithm is used to carry out clustering of data into homogenous classes for creating a profile. The profile is classified into an attacker or non-attacker based on another algorithm. For better significant and homogeneous presentation to the data, each class was modelled using linear regression. The system has a learning stage tagged algorithm 1 and decision stage tagged algorithm 2. Three parameters (namely: number of clusters, initialization of centroids, and parameters of linear function) are required in algorithm. The fitting model achieves higher precision as the precision of initialization and calculation of these parameters becomes higher. In algorithm 2, the following information are returned: creation and classification of profile based on the attacker model. Decision is made depending on the new profile projections on the hackers' profiles. The learning stage and the decision stage of the system are shown in Figure 5. The application of the proposed smart agent was illustrated as shown in Figure 6 considering a company that deploys three servers: a mail server, an FTP server for transfer of files, and a web server for its site web hosting. For the company to monitor suspicious profiles via the services it provides in the same honeypot server, three virtual machines can be implemented. This enables the detection of suspicious profile patterns on services and the prediction of attacker profiles based on ML analysis.



(a) (b)

Figure 5: Honeypot and machine learning cyber security system: (a) Learning stage, (b) decision stage [8]

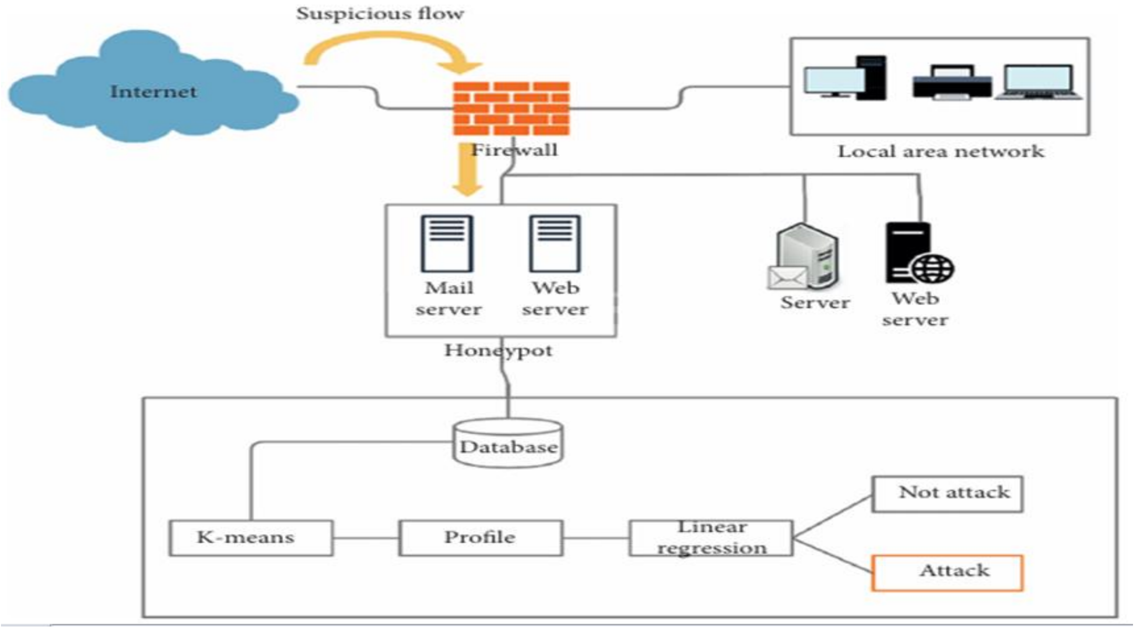


Figure 6: Illustration of the application of honeypot and ML based smart agent for cyber security [8]

With the new and continuous emerging cyber threats in the e-health system, it is obvious that security of telemonitoring systems cannot be neglected any longer and as such, Ardito et al. [14] proposed an AI based cyber-attack detection system (CADS) that detects threats (or malicious activities) without requiring a dedicated security analyst, but autonomously explains the threat and output (or display) suspected attack data to healthcare workers for feedback. The system was designed specifically for detection of cyber-attack considering the case of a hacked remote patient telemonitoring system. A specific test application was considered for heartbeat telemonitoring. Automatic detection of anomaly was achieved using deep learning algorithms. Particularly, the system detects cases of anomaly in heartbeats using a robust deep autoencoder. Subsequently, the detected abnormality in heartbeat is interpreted with a combination of cutting-edge explainable security models and by means of two new explainable scores introduced that show the user the causes of a malicious action interrupting the heartbeat telemonitoring. The structure of the system is shown in Figure 7.

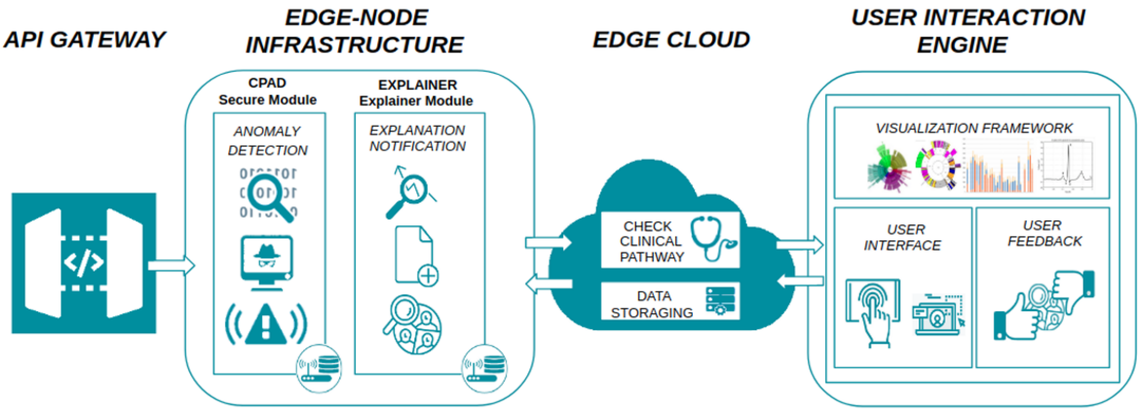


Figure 7: Cyber-attack detection system architecture [14]

The dire consequences of cyber-attacks such as losses of large materials and leakage (or falsification) of critical data, is of immense concern to network users (especially companies and institutions) and the fact that existing methods utilized for detection of threats and attacks are slow or operate after the events have occurred and analyse them and issue response, has resulted in the development of a real-time AI based network cyber security model using ML algorithms in [15]. The system is a conceptual framework that contains new rules with previous rules so as to create a network monitoring system that detect and prevent cyber-attacks in real-time using AI models that provide classification and prediction of user behaviour. The system consisted of dynamic programming techniques with clustering and making use of multiple algorithms. The system was designed with the goal of discovering the optimum solutions that can identify and detect malware. The Gaussian mixture model (GMM) was selected amongst ML based clustering algorithms because it does not depend on distance like other popular models such as k-means, but relies on similar behaviour of samples. The GMM algorithm was used by the proposed tool for monitoring and securing the system by creating the models that contain groups representing previous attacks that were logged and compared with the new activities for the protection of the network. The performance of the system was evaluated using data from VirusShare, which were filtered to obtain suitable data for the purpose of study and thereafter 107 samples were run in the sandbox environment. The system calls were collected by depending on DrStrace programme and were converted into an iteration matrix. The conceptual framework of the system is such that samples from previous cyber-attacks and malware are collected and stored in the form of groups utilizing AI techniques so as to generate models that are later trained and employed for comparison process with user action as shown in Figure 8.

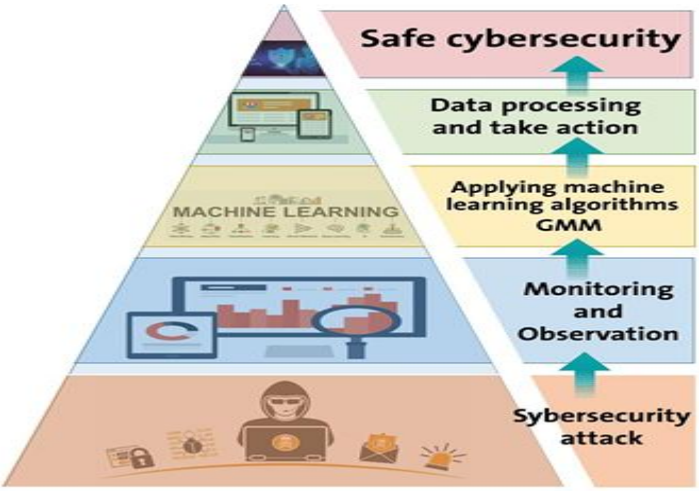

Safe cybersecurity
Data processing and take action
MACHINE LEARNING
Applying machine learning algorithms GMM
Monitoring and Observation
Sybersecurity attack

Figure 8: Framework of ML based networks security model [15]

Network integration of cyber-physical systems (CPS) is common because it allows remote access, surveillance, and analysis, which have make them vulnerable to cyber-attack due to integration with insecure network. Hence, a new framework for the detection of cyber-attacks in CPS using AI and ML was presented in [16]. The implemented system utilized a self-tuned fuzzy logic hidden Markov model (SFL-HMM) together with heuristic multiple swarm optimization (HMSO) for detecting anomaly was designed for the purpose of detecting cyber-attack in CPS. A range of alternative techniques for launching cyber-attack was examined because in recent times they have undergone a significant transformation as cybercriminals are continuously devising means of manoeuvring security methods. The KDD99 attack datasets, with 41 features that collectively show 22 different kinds of attacks were used for the study. Clean-up of data in CPS database was performed by the means of normalisation to get rid of errors and duplication. Linear discriminant analysis (LDA) was used to acquire features. The proposed framework was evaluated via MATLAB simulation in terms of detection rate, the rate of the true positive, and the rate of the false positive. It was observed to outperform traditional detection method. The structure of the system is depicted in Figure 9.

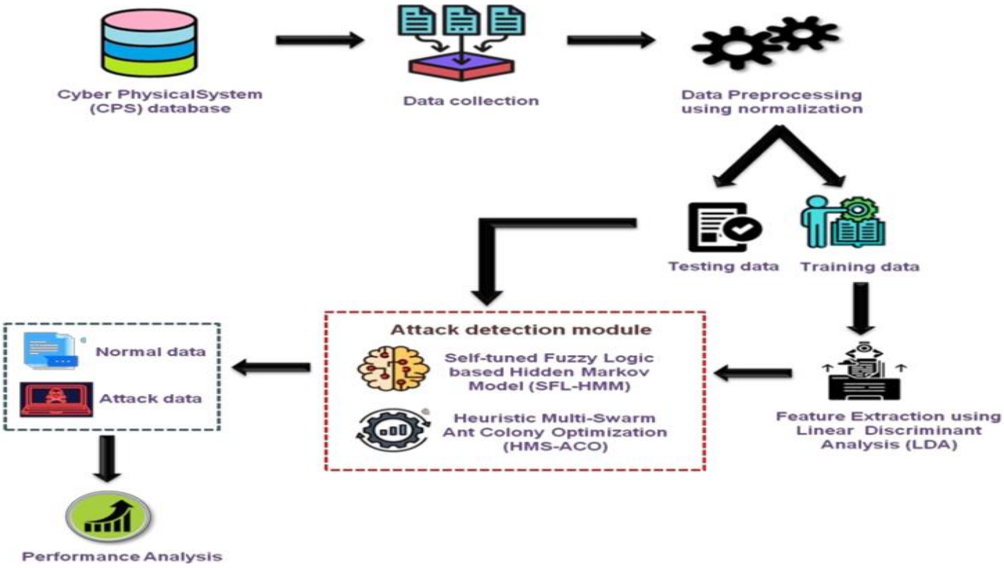

Cyber PhysicalSystem (CPS) database
Data collection
Data Preprocessing using normalization
Testing data
Training data
Feature Extraction using Linear Discriminant Analysis (LDA)
Attack detection module
Self-tuned Fuzzy Logic based Hidden Markov Model (SFL-HMM)
Heuristic Multi-Swarm Ant Colony Optimization (HMS-ACO)
Normal data
Attack data
Performance Analysis

Figure 9: Structure of AI based system for detecting cyber-attacks in CPS [16]

Though the concept of smart city has been considered from the perspective of its ease and comfort, several security issues are being pointed out that limit its pace of growth and as such ML based intrusion detection system (IDS) that intelligently detect threats in network are used to monitor the entire network traffic and take decision about the legit ability of data and alerts the user when any anomaly occurs. Therefore, Chohan et al. [17] carried comparative analysis of different ML techniques that have been used to improve the IDS detection accuracy. The ML algorithms considered were ADA Boost, linear support vector machine (LSVM), auto encoder classifier, quadratic support vector machine (QSVM), and multi-layer perceptron algorithms. The ML algorithms were trained over UNSW-NB15 dataset. Simulation platform was developed using Python. The analysis indicated the IDS with ADA Boost yielded best performance in terms of accuracy. The results from the analysis are presented in Table 1.

Table 1: Cyber-attack detection accuracy of ML techniques [17]

Algorithms	**Accuracy (%)**
ADA Boost	98.3431
Auto encoder classifier	96.1133
Linear support vector machine	97.8503
Quadratic support vector machine	84.7305
Multi-layer perceptron	97.9735

An intrusion detection system (IDS) for Naval Facilities Engineering Command (NAVFAC) smart grid was proposed in [18]. The system consisted of feature extractor, classifier, anomaly detector, and response manager. For accurate response management, k-nearest neighbours (KNN) algorithm was used to group various attacks such as web attacks, File Transfer protocol/secure shell (FTP/SSH) attacks, denial of service (DoS), distributed DoS (DDoS) and port scanning into broader attack classes of Active, Denial and Probe. The accuracy of the classifier was maximized by optimizing the k value so as to reduce the load on security operations centre (SOC). Increase in k and disable principal component analysis (PCA) was taken as an approach to reduce the false positive rate. It is expected of the system to prompt specific responses to protect the smart grid integrity and separate the areas under malicious action. The CICIDS2017 dataset was introduced and different attacks were grouped in terms of the identified objectives. The ML toolbox of the MATLAB was used to show that KNN algorithm was able to isolate various classes of attack traffic from benign traffic. The structure of the proposed IDS cyber security system is shown in Figure 10.

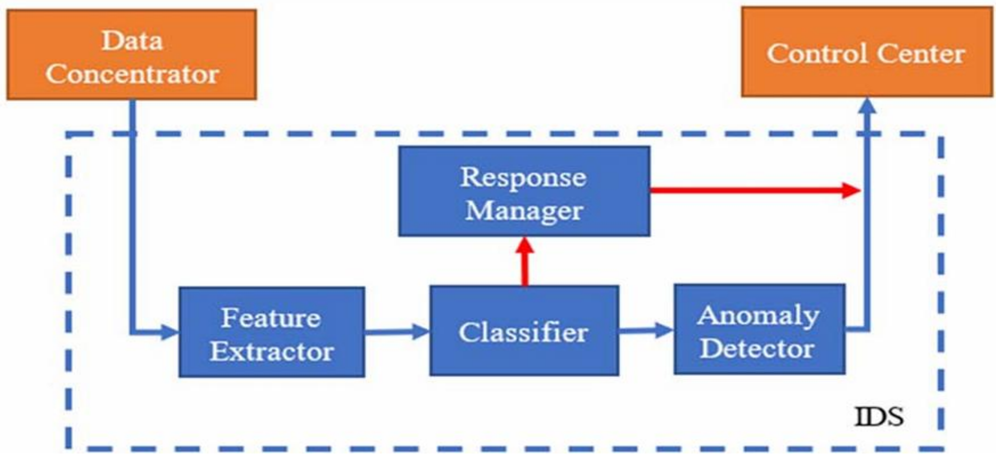


Figure 10: Technique for IDS for navy smart grid [18]

It is essential to provide security against cyber-attacks for industrial control systems (ICSs) because of the fact that their function is extensively sort for and these systems are frequently used in critical infrastructure and though several methods for detecting various types of cyber-attacks have been provided with each method having its uniqueness, [19] proposed ML and deep learning based system to detect anomaly in ICSs. The system was primarily designed to detect cyber-attacks in ICSs The system employed logistic regression, k-nearest neighbour (KNN), linear discriminant analysis, and decision tree (DT) algorithms for the ML and the long short-term memory (LSTM) and the convolution neural network and long short-term memory (CNN-LSTM) as the deep learning algorithms for the detection of malicious activities. Real ICS datasets from Necon Automation and the International Islamic university, Malaysia (IIUM) were used to examine the algorithms. These algorithms were trained on three types of attacks namely; man-in-the middle (MITM) attacks, web-server access attacks, Telnet attacks, and then with normal packets. Two stages were taken in developing the system, binary classification and multiclass classification in order to achieve high performance approach for detecting attacks in CSs. Malware was detected as normal or attacks by the binary classification whereas the multiclass was employed for all individual attack detection. Amongst the ML algorithms, KNN and DT yielded the highest level of accuracy. In terms of percentage error between targeted and predicted output values during the validation phase using sensitivity (statistical) analysis such as mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE), and R^2, the KNN and DT algorithms achieved the best result with fewer prediction errors in binary classification and multiclass classification respectively. The KNN and DT algorithms outperformed other algorithms by yielding 100% accuracy in both binary classification and multiclass classification, and in the same vein provided R^2 = 100% for sensitivity analysis. Performance comparison with other ML and deep learning algorithms in existing systems revealed that the proposed system outperformed them. The performance of KNN and DT indicated that the algorithms can outperform other cutting-edge classifier models [19]. It was recommended that future study intends to apply the proposed system with ICS system for protecting food security. The conceptual framework of the ML and deep learning based AI system for security of the ICS system is shown in Figure 11.

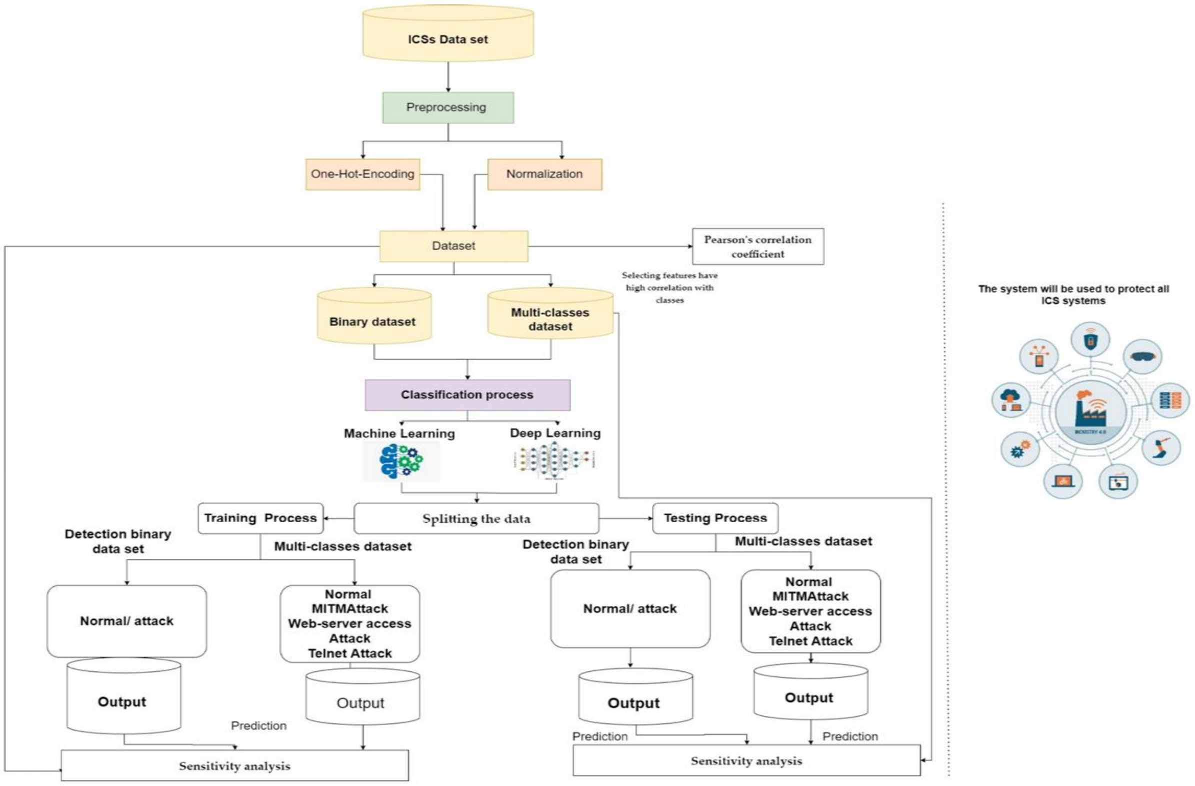


Figure 11: Framework of cyber security system for ICS [19]

The application of Internet of Things (IoT) (in cars, unmanned aerial vehicles, and medical equipment etc.) has unimaginably increased and impacted nearly all aspect of human endeavour due to rapid revolution and development of low power, smart, and autonomous (or self-reliant) devices, which has also prompted increased threat and attack capabilities and thereby making it important for new technologies to be devised so as to secure these critical infrastructures. Hence, [20] presented efficient and generic top-down system for detection of intrusion, and using non-conventional ML technique for classification in IoT networks. The proposed system was designed in such a way that it can be adapted and employed for detection and classification integrating any IoT cyber-attack datasets. Example of such datasets provided were CICIDS dataset, MQTT dataset, KDD99 dataset, NSL-KDD dataset, AWID dataset, DDoS dataset, and UNSW-NB15 dataset. There are three subsystems that make up the system and are feature engineering (FE), feature learning (FL), and detection and classification (DC) as shown in Figure 12. The FE subsystem comprised of IoT attacks dataset, and data pre-processing and cyber data encoding (discrete input features). The FL subsystem uses three machine learning algorithms which include shallow neural networks (SNN), convolution neural networks (CNN),and deep neural networks (DNN) (model parameters updates), while the DC subsystem consisted of fully connected neural network (majority voting method, MVM, and support vector machine (SVM) with softmax function or multiclass classifier (traffic categories/classes). Deep learning algorithms have been used for the system to facilitate the detection of slightly modified attacks of IoT network while achieving high detection/classification accuracy for IoT traffic obtained from either real-time system or an off-line dataset.

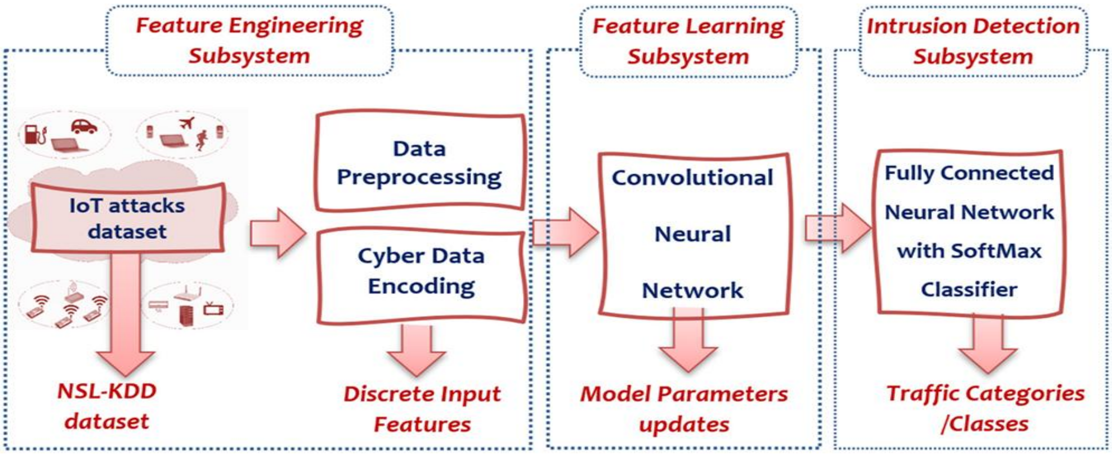

Feature Engineering Subsystem
Feature Learning Subsystem
Intrusion Detection Subsystem
IoT attacks dataset
Data Preprocessing
Cyber Data Encoding
Convolutional Neural Network
Fully Connected Neural Network with SoftMax Classifier
NSL-KDD dataset
Discrete Input Features
Model Parameters updates
Traffic Categories /Classes

Figure 12: Architecture of generic top-down for intrusion detection/classification in IoT network [20]

On the account of the rapid development in ICT in nations across the world, there is corresponding increase in cybercrimes and as such there is need for users of digital systems to be secured. In accordance to this, ML based predictive model for cyber security threats prediction was developed in [21]. The system was developed for predicting the users' risk based on their online behaviour. Real-life data was obtained based on the behaviour of 207 undergraduates from a private university. The system was modelled using five supervised machine learning algorithms, which included naïve Bayesian classifier, support vector machine (SVM), decision tree (DT), k-nearest neighbour (KNN), and regression logistics with the assistance of RapidMiner. The construction, testing, and validation of three types of cybercrime threats (malware, social engineering, and password attack) predictive models were performed using the algorithms. The outcome of the study indicated that the KNN algorithm produced the highest accuracy and least classification error with respect to the three types of cybercrime threats. The system was basically considered as vital security measure to alert users with features of high or low risk behaviour and the kind of actions that could be taken to increase awareness. Structure of supervised learning model is shown in Figure 13.

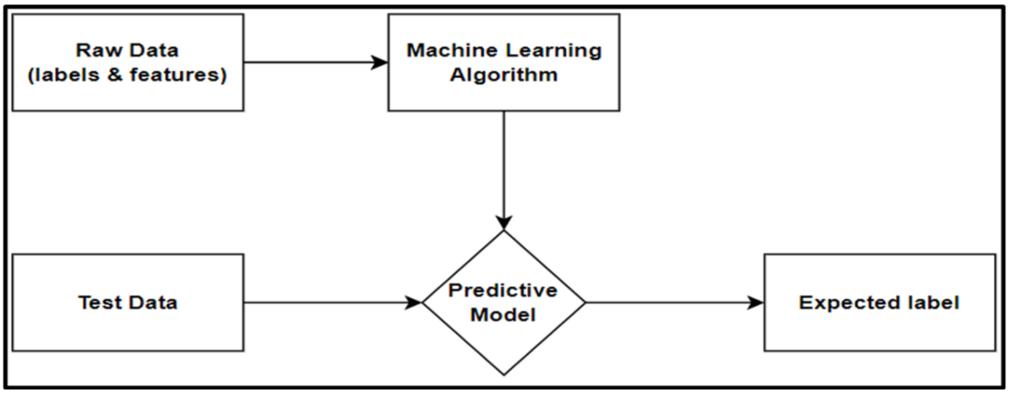


Figure 13: Diagram of supervised machine learning system [21]

The development of network security is considered a critical research interest because of the rapid expansion of data in internet and communication network, which has also resulted in corresponding cyber-attacks on the systems and thereby complicating the effectiveness of detecting breaches by network security. In order to address the excess new attacks being created that makes it difficult for network security to adequately detect breaches and also overcome several errors associated with current intrusion detection systems (IDS) including their not being able to thwart attacks, [22] used combined data analysis technique with four robust ML collaborative algorithms with proposed robust genetic ensemble model. The collaborative modes included voting classifier, bagging classifier, gradient boosting classifier, and random forest-based bagging algorithm in addition to the proposed robust genetic ensemble classifier. Each of the algorithms was used to create and test a model using a network dataset. The performance of the models was evaluated based on the ability to detect the occurrence of anomaly. The algorithms were experimentally evaluated using network traffic dataset available on Kaggle, which is on the Python platform and has limited samples. The robust genetic optimize ensemble model yielded the finest results amongst the others by showing the lowest error for MSE and MAE, and with correlation coefficient of 0.9985 between actual and predicted variables. The structure of the proposed method is shown in Figure 14.

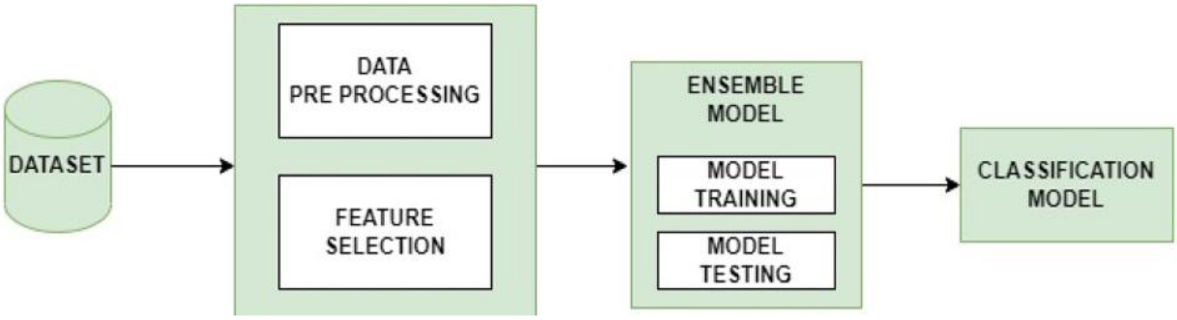


Figure 14: Block diagram description of system approach [22]

The ML as a power tool for intelligent cyber analysis to facilitate proactive security, has been reported to face two significant challenges to its application, and these are the computational costs incur by state-of-the -art ML systems and the fact that security must develop unique frameworks for ML computation for different applications. The first drawbacks hinder the widespread adaptation of ML based cyber security system in the security of businesses and the second cased requires that cyber analyst with must be dedicated for a given application such that ML system designed for detecting vulnerabilities in 5G core network (5 GCN) cannot not be used by security analyst to analyse the security of a connect vehicle [23]. A generalized ML based frameworks called machine learning framework for efficient exploit detection (ML-FEED) was proposed in Saha (2022) to address the drawbacks. The ML-FEED used ML and rule-base algorithms to offer effective detection of new threats. A new method for identifying vulnerable fingerprints from public cyber-attack databases was introduced by the ML-FEED. A generic framework that could be used for security (risk) analysis of diverse platforms called smart hacking approaches for risk scanning (SHARKS) was introduced. SHARKS was designed for risk analysis of IoT and CPS. The SHARKS was used to initially extracted intelligence from cyber-attacks datasets on IoT and CPS environments and then ML was used to learn the fundamental patterns of these attacks. With the ML algorithm, the SHARKS was able to protect unknown attacks on IoT and CPS. SHARKS was used to perform threat analysis on 5GCN, which resulted in discovering 119 new potential threats in a generic 5GCN architecture. These attacks are largely cause by communication among different vulnerabilities of evolving 5GCN technologies such as network function virtualisation and software-defined networking [23]. The inference pipeline of the ML-FEED is shown in Figure 15. Figure 16 shows flow chart of the methodology used in the ML-FEED based cyber security system.

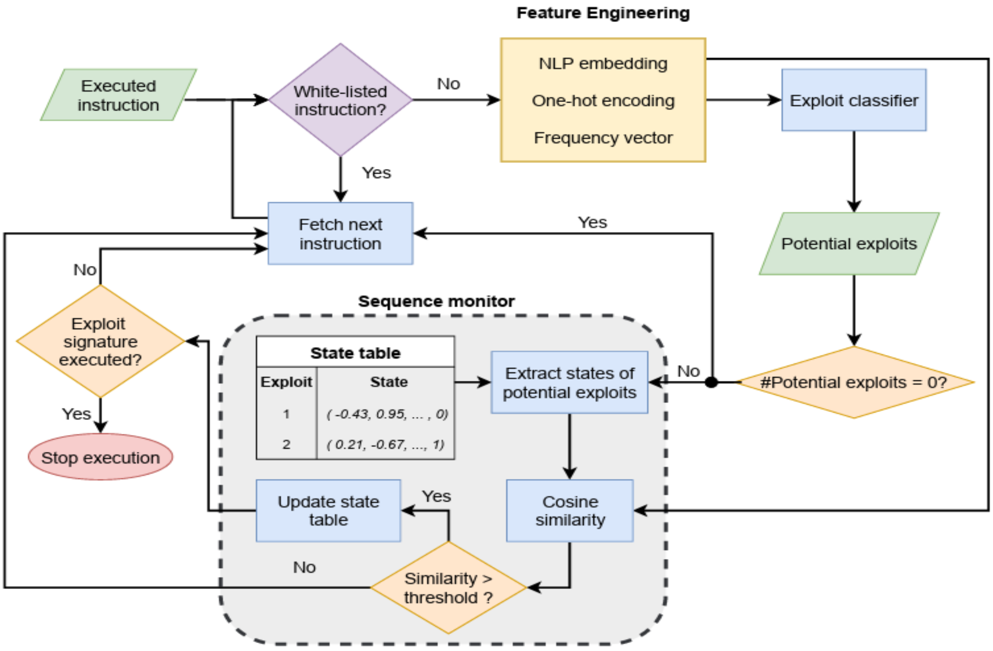


Figure 15: Inference pipeline of the ML-FEED [23]

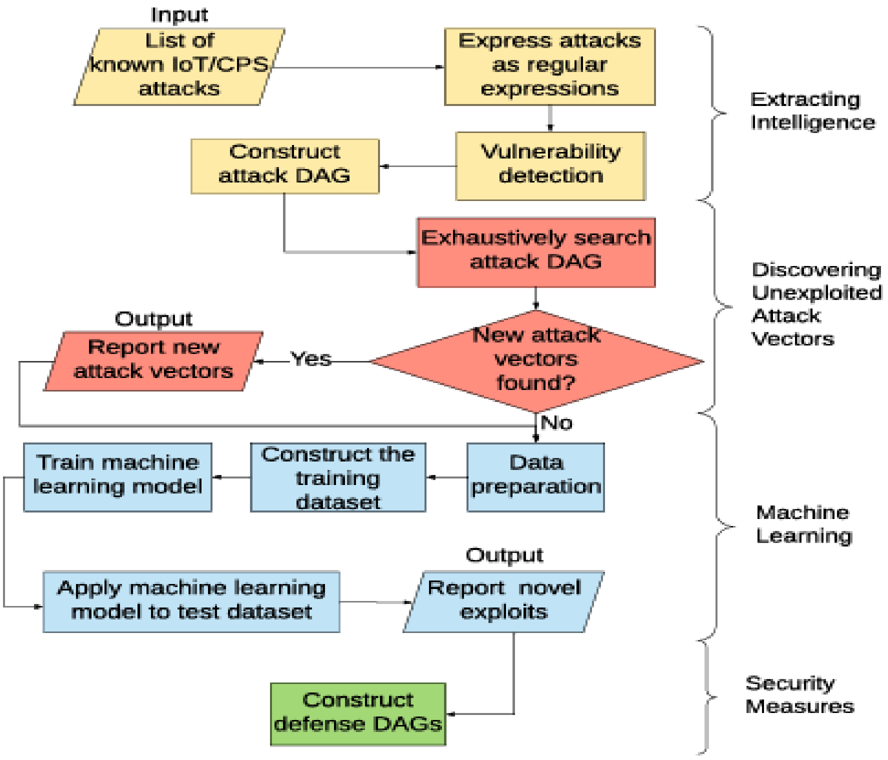
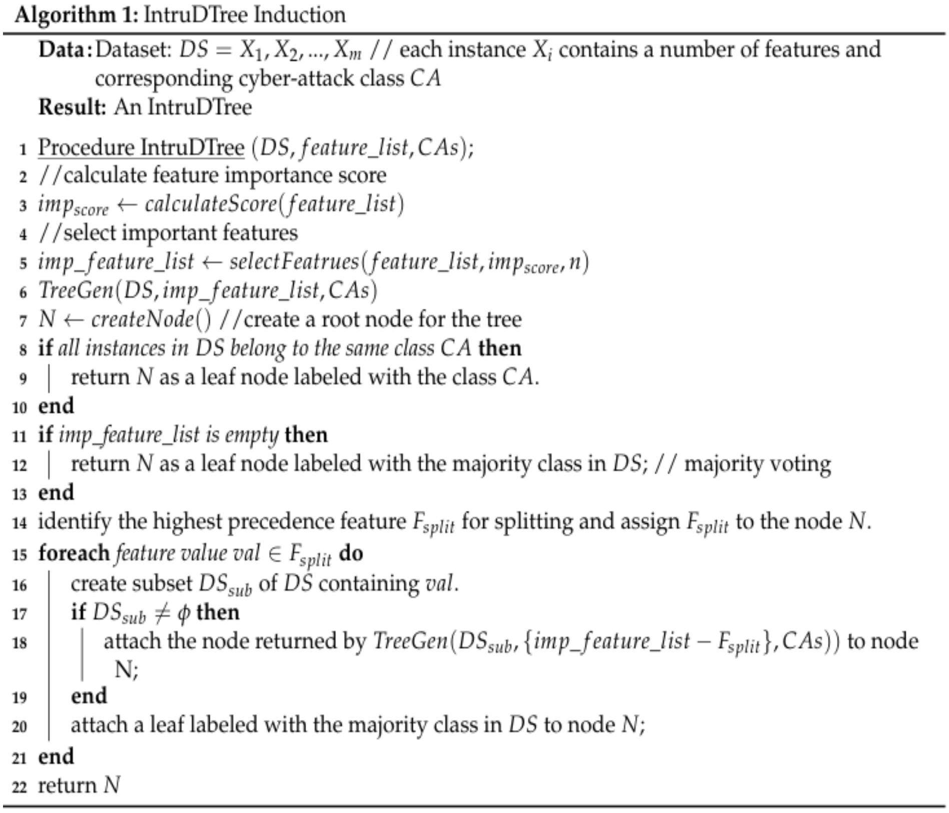

Input
List of known IoT/CPS attacks
Express attacks as regular expressions
Vulnerability detection
Construct attack DAG
Exhaustively search attack DAG
New attack vectors found?
Yes
Output
Report new attack vectors
No
Data preparation
Construct the training dataset
Train machine learning model
Apply machine learning model to test dataset
Output
Report novel exploits
Construct defense DAGs
Extracting Intelligence
Discovering Unexploited Attack Vectors
Machine Learning
Security Measures

Figure 16: Flowchart of the system design [23]

It is becoming more important today to develop an efficient intrusion detection system (IDS) that can perform the role of detecting various cyber-attacks or anomalies a network of computer and IoT devices. Machine line based security system that initially considered the ranking of security features based on their importance called intrusion detection tree ("IntruDTree") in [24]. A tree-based generalisation intrusion detection system was developed based on the selected important features. Prior to the building of the tree-based generalisation intrusion model, the ML based security system first allows for the ranking of security features consistent with their importance. In exploring security dataset, an intrusion dataset that was classified into two categories: normal and anomaly obtained from Kaggle was used in the study. The total features in the dataset were 41 of which 3 features were qualitative and 38 features were quantitative. In the dataset were more than 25,000 cases collected from a various kinds of intrusions simulated in a military network environment. An environment was created by simulating a typical US Air Force local area network (LAN) in order to collect the raw data. The raw security data was prepared based on feature encoding and scaling. Label encoding was used to carry out direct conversion of the feature values into specific numerical values. Feature scaling also known as data normalisation was achieved using a Standard Scaler. Feature importance and ranking was determined based on important score using Gini Index formula. The efficiency of the IntruDTree model was evaluated by carrying out experiments on dataset of cyber security in terms of precision, recall, fscore, accuracy, and true positive rate against false positive rate (i.e. ROC curve). Experimental showed that the IntruDTree produced 98% efficiency for precision, recall, fscore, and accuracy for each individual class of data (normal and anomaly) respectively. The IntruDTree was compared with other ML models such as naïve Bayes (NB), logistic regression (LR), KNN, and SVM and it was shown to best performance in terms of precision, recall, fscore, and accuracy. Figure 17 shows the algorithm describing the overall process for building the IntruDTree.



```
Algorithm 1: IntruDTree Induction
Data: Dataset: $DS = X_1, X_2, ..., X_m$ // each instance $X_i$ contains a number of features and
      corresponding cyber-attack class $CA$
Result: An IntruDTree
Procedure IntruDTree ($DS, feature\_list, CAs$);
//calculate feature importance score
$imp_{score} \leftarrow calculateScore(feature\_list)$
//select important features
$imp\_feature\_list \leftarrow selectFeatrues(feature\_list, imp_{score}, n)$
$TreeGen(DS, imp\_feature\_list, CAs)$
$N \leftarrow createNode()$ //create a root node for the tree
if all instances in DS belong to the same class $CA$ then
    return $N$ as a leaf node labeled with the class $CA$.
end
if imp_feature_list is empty then
    return $N$ as a leaf node labeled with the majority class in $DS$; // majority voting
end
identify the highest precedence feature $F_{split}$ for splitting and assign $F_{split}$ to the node $N$.
foreach feature value $val \in F_{split}$ do
    create subset $DS_{sub}$ of $DS$ containing $val$.
    if $DS_{sub} \neq \phi$ then
        attach the node returned by $TreeGen(DS_{sub}, \{imp\_feature\_list - F_{split}\}, CAs))$ to node
         N;
    end
    attach a leaf labeled with the majority class in $DS$ to node $N$;
end
return $N$
```

Figure 17: Algorithm for IntruDTree [24]

III. POTENTIALS AND CHALLENGES OF AI SYSEMS IN CYBER SECURITY

The reviewed articles have shown that AI based models can be used for defence purposes in a number of ways. It is believed that in the future, AI systems with better performance than existing models will be available. This is because even attackers can as well develop Algorithms for attacks. Hence, the AI models for cyber security need to be flexible with regard to creating possibility for modifications in order to give room for improvement.

Generally, when considering the potentials of AI in cyber security, diversity of cyber-attacks and threats that exists should be well understood. This as mentioned earlier, is largely due to the fact that as experts in the field of cyber security continuously and tirelessly work to develop systems to analyse and detect new and emerging threats and attacks, corresponding synergy is being put towards actions that can deceive or compromised the AI algorithms in cyber space [25]. Therefore the potentials of the AI systems in cyber security and the challenges are briefly presented in subsections 3.1 and 3.2.

3.1. Potentials of Artificial Intelligence Systems in Cyber Security

The AI cyber security system serves as an autonomous agent that is capable of detecting actions of cyber criminals or foreign agents designed to interrupt or damage an established working process both now and in future using advanced computing algorithms or functions. The primary objective of an AI in cyber security is to track the cyber space environment and execute actions that are directed towards achieving its operational set targets such as identification and prevention of cyber threats and attacks, and the protection of the cyber-physical system (CPS), the Industrial Control System (ICS), or the general digital networked infrastructure of public and private enterprise. The following are possibilities the AI system provides in cyber security.

a) *Adaptation*:Though the pattern of attacks often changes, the application of AI algorithms in cyber security will enable cyber- physical systems (CPSs), the Industrial Control Systems (ICSs), and the general digital ecosystem to continuously adapt to attempted threats and attacks. This is because up to date knowledge of specific threats and attacks can be provided by AI based cyber security technologies and as such ensures protection against exposure to malicious cyber actions. Thus, systems used in digital space take advantage of the predictive analytics of the AI technique in dealing with the difficulty of cyber malicious actions [26].

b) *Monitoring*: The dynamic nature of cyber-attacks is itself a complex process with varying pattern of execution. Therefore, robust and sophisticated approaches are required to deal with it. AI helps in monitoring to detect and identify patterns of known and unknown actions with features of threat. It is possible to mitigate and contain attacks by means of monitoring [27] using AI based cyber security systems.

c) *Defence*: Generally, the primary responsibility of any cyber security system is to provide protection and AI based systems are no exceptions. With the level of attacks increasing faster than the cyber defences provided to protect networked digital facilities, conventional machine learning approach is now being modified to contain two or more combined AI algorithms such as a combination of ML and deep learning techniques, which is utilized for cyber security purposes to increase the ability to detect and protect critical public and private networked digital infrastructures against malicious actions.

d) *Response*: The use of AI enabled technology in cyber security can aid in providing better framework for prioritising and responding to security concerns. This way, faster response to cyber-attacks or security alerts and including prompt exposure of the nature and cause of the attack is achieved. Therefore, by providing prompt response to security threat, vulnerabilities of digital infrastructures are mitigated and future occurrence is prevented early and faster enough.

e) *Automation*: AI systems used in cyber security are designed to improve defence capability taking into consideration the possibility of automating the threat identification, detection, protection, and prevention to avoid future issues while exploiting (big) data analytic capabilities with accuracy, speed and efficiency.

f) *Prediction*: The application of AI based technologies can aid in predicting the possibility of compromised areas in digital infrastructure in order to make arrangement and put resources towards securing that parts of the digital ecosystem with highest vulnerabilities [28] or breach risk.

3.2. Challenges of Artificial Intelligence Systems in Cyber Security

Though in subsection 3.1, the potentials of AI based technologies in security management in cyber space have shown to be promising and serve as the advantages they have over the traditional approaches, there are some challenges using AI and are highlighted in this subsection.

a) *Security guarantee*: There is actually no assurance of absolute security for industrial environment by AI systems [26].

b) *Moral concerns*: In the implementation of AI technologies, ethical issues such as the lack of moral code for machines arise such that when it comes to decision making that could have moral impacts, these systems may not be capable to identify these impacts and therefore resulting in a challenge caused by the lack of ability to sense and make decisions taking into consideration moral issues [26, 29].

c) *Financial constraint*: Investment in AI requires sufficient financing and this can impact on its viability and implementation because organisations would need to commit substantial resources into having it. This is because in training an AI model for cyber security purpose, large volumes of datasets with many distinctive sets of malware codes, non-malicious codes, and inconsistencies must be acquired and the process of getting these datasets is time-consuming and needs financial commitment that most establishments may not be able to pay for [28].

d) *Large volumes of data*: AI technologies require enormous volumes of data and trials otherwise these technologies will give inaccurate results and/or false positives, but collection of more data can results in privacy and protection issues [26, 28].

e) *Tool for cybercriminals*: In the hands of attackers, AI can be a powerful tool to analyse, effectively find and exploit vulnerabilities in digital ecosystem and thereby carry out more advance and far-reaching attacks.

f) *Tool against opponent*: Though AI systems can be effectively used for monitoring and subsequently detects intruder activities on digital (information technology) infrastructure, despotic states and governments can use these technologies to monitor and track their opponents or adversaries. Today, it has become a weapon to wage war by nations against public and private infrastructures or facilities of their opponents. In the same vein, despite the fact that AI techniques allows cyber security experts to develop models or search for emerging threats, it can be used for personal gains such privacy monitoring, tracking and other abuses [26].

IV. CONCLUSION

In this paper, approaches to cyber security using AI technologies that mostly involved the use of ML algorithms including the potentials and challenges of these methods have been presented. The study has shown that many AI based methods are emerging continuously in order to counter the activities of cyber criminals, whose actions are constantly increasing and changing pattern of dynamisms geared towards carrying out cyber-attacks on public or private infrastructures. The AI technologies offer promising security in cyberspace and gives provision for advancement such as hybrid combination of two more methods, but there are some shortcomings associated with these technologies. Further study will focus on approaches that are being proposed to solve the challenges facing AI based cyber security system.

REFERENCES

[1]. IBM, n.d. What is Cybersecurity? https://www.ibm.com/topics/cybersecurity

[2]. Ramasubramanian, K., Venkateswarlu, L. and Yerram, S. 2021. Applications and techniques of artificial intelligence in cyber security. Turkish Journal of Computer and Mathematics Education. 12(4):332-339.

[3]. Das, R. and Patel, M. 2017. Cyber security for social networking sites: issues, challenges and solutions. International Journal for Research in Applied Science & Engineering Technology. 5(4): 833-838.

[4]. Moisset, S. 2023. How security analysts can use AI in cybersecurity. FreeCodeCamp. https://www.freecodecamp,org/news/how-to-use-artificial-intelligence-in-cybersecurity/

[5]. Chakraborty, A., Biswas, A. and Khan, A. K. 2023. Artificial intelligence for cybersecurity: threats, attacks and mitigation. In:Biswas, A., Semwal, V. B., Singh, D. (eds) Artificial Intelligence for Societal Issues. Intelligent Systems Reference Library, 231. Springer, Cham. https://doi.org/10.1007/978-3-031-12419-8_1

[6]. Kharbanda, V., Seetharaman, A. and Maddulety, K. 2023. Application of artificial intelligence in cyber security. International Journal of Security and Privacy in Pervasive Computing. 15(1): 1-13. https://doi.org/10.4018/ijsppc.318676

[7]. Whitlock, P. 2023. Artificial intelligence: the next evolution in cyber threat detection. ISC2. https://www.isc2.org/insights/2023/06Artficial-itelligence-Next-Evolution-Cyber-Threat_Detection

[8]. El Kamel, N., Eddabbah, M., Lmoumen, Y. and Touahni, R. 2020. A smart agent design for cyber security based on honeypot and machine learning. Hindawi, Security and Communication Networks Volume 2020, Article ID 8865474: 1-9. https://doi.org/10.1155/2020/8865474

[9]. Feng, G., Zhang, C. andZhang, Q. 2014. A design of linkage security defense system based on honeypot: trustworthy computing and services, Springer, Berlin, Heidelberg, Germany.

[10]. Saeed, R. A. and Shareef, S. M. 2020. Implementation of artificial intelligence to predict threats in social media based on user's behavior. International Journal of Advanced Trends in Computer Science and Engineering. 9(5): 6931 – 6938. https://doi.org/10.30534/ijatcse/2020/10952020

[11]. Althar, R. R., Samanta, D., Purushotham, S. and Sengar, S. S. 2023. Design and development of artificial intelligence knowledge processing system for optimizing security of software system. SN Computer Science. 4(331): 1-12. https://doi.org/10.1007/s42979-023-01785-2

[12]. Baldassarre, M. T., De Vincentiis, M., Pal, A. and Scalera, M. 2023. Quantum artificial intelligence for cyber security education in software engineering. IS-EUD 2023: 9th International Symposium on End-User Development, 6-8 June 2023, Cagliari, Italy. https://owasp.org/www-project-application-security-verification-standard/

[13]. Alzahrani, A. and Aldhyani, T. H. H. 2023. Design of efficient based artificial intelligence approaches for sustainable of cyber security in smart industrial control system. Sustainability 2023, 15, 8076. https://doi.org/10.3390/su15108076

[14]. Ardito, C., Di Noia, T., Di Sciascio, E., Lofù, D., Pazienza, A. and Vitulano, F. 2021. An artificial intelligence cyberattack detection system to improve threat reaction in e-health. ITASEC21: Italian Conference on Cybersecurity, April 07–09.

[15]. Sari, F. A., Alrammahi, A. A. H., Hameed, A. S., Alrikabi, H. M. B., Abdul–Razaq, A. A., Nasser, H. K. and AL-Rifaie, M. F. 2022. Networks cyber security model by using machine learning techniques. International Journal of Intelligent Systems and Applications in Engineering. 10(3s):257–263.

[16]. Almajed, R., Ibrahim, A., Abualkishik, A. Z., Mourad, N. and Almansour, F. A. 2022. Using machine learning algorithm for detection of cyber-attacks in cyber physical systems. Periodicals of Engineering and Natural Sciences. 10(3):261-275.

[17]. Chohan, M. N., Haider, U., Ayub, M. Y., Shoukat, H., Bhatia, T. K. and UI Hassan, M. F. 2023. Detection of cyber attacks using machine learning based intrusion detection system for IoT based smart cities. EAI Endorsed Transactions on Smart Cities. 7(2):1-7.

[18]. Chan, V. C. 2020. Using a K-nearest neighbors machine learning approach to detect cyberattacks on the navy smart grid. Master thesis, Naval Postgraduate School, Monterey, California.

[19]. Alkahtani, H. and Aldhyani, T. H. H. 2022. Developing Cybersecurity Systems Based on Machine Learning and DeepLearning Algorithms for Protecting Food Security Systems: Industrial Control Systems. Electronics 2022, 11, 1717. https://doi.org/10.3390/electronics11111717

[20]. Abu Al-Haija, Q. 2022. Top-down machine learning-based architecture for cyberattacks identification and classification in iot communication networks. Front. Big Data. 4:782902. https://doi.org/10.3389/fdata.2021.782902

[21]. Tin, T. T., Xin, K. J., Aitizaz, A., Tiung, L. L., Keat, T. C., and Sarwar, H. 2023. Machine learning based predictive modelling of cybersecurity threats utilising behavioural data. International Journal of Advanced Computer Science and Applications. 14(9):832-840.

[22]. Akhtar, M. A., Qadri, S. M. O., Siddiqui, M. A., Mustafa, S. M. N., Javaid, S. and Ali, S. A. 2023. Robust genetic machine learning ensemble model for intrusion detection in network traffic. Nature Scientific Reports. 13:17227. https://doi.org/10.1038/s41598-023-43816-1

[23]. Saha, T. 2022. Machine learning-based efficient and generalizable cybersecurity frameworks. PhD Dissertation, Princeton University.

[24]. Sanker, I, H. Abushak, Y, B., Alsolami, F. and Khan, A. I. (2020). IntruDTree: A machine learning based cyber security intrusion detection model. Symmetry.12, 754. https://doi.org/10.3390/sym12050754

[25]. Szymanski, T. H. 2022. The “Cyber security via determinism” paradigm for a quantum safe zero trust deterministic internet of things (IoT). IEEE Access 2022, 10, 45893–45930.

[26]. deAzambuja, A. J. G., Plesker, C., Schützer, K., Anderl, R., Schleich, B., Almeida, V. R. 2023. Artificial intelligence-based cyber security in the context of industry 4.0—a survey. Electronics 2023, 12, 1920. https://doi.org/10.3390/electronics12081920

[27]. Trung, N. D. Huy, D. T. N., Huong, L. T. T., Van Thanh, T., Thanh, N. T. P., Dung, N.T. 2021. Digital transformation, AI applications and iots in blockchain managing commerce secrets: and cybersecurity risk solutions in the era of industry 4.0 and further. Webology 2021, 18, 453–465.

[28]. Belani, G. 2021. The use of artificial intelligence in cybersecurity: a review. IEEE Computer Society. https://www.computer.org/publications/tech-news/trends/the-use-of-artificial-intelligence-in-cybersecurity

[29]. Laghari, S. U. A., Manickam, S., Al-Ani, A. K., Rehman, S. U., Karuppayah, S. 2021. SECS/GEMsec: A Mechanism for Detection and Prevention of Cyber-Attacks on SECS/GEM Communications in Industry 4.0 Landscape. IEEE Access 2021, 9, 154380–154394.