Dijlah Journal of Engineering Science (DJES) Vol. 3, No. 1, March, 2026, pp. 127-143
ISSN: Printed: **3078-9656, Online: 3078-9664**, paper ID: 117

Employing Functional Machine Learning Principles for Active Detection of Brute Force Attack Relying on The CICIDS2017 Dataset

Saja Abdulkareem[1], **Ali A-Zerkani**[2]
[1]Information technology, Altınbaş University, Istanbul, Turkey
[2]Information technology, Altınbaş University, Istanbul, Turkey

Article Info

Article history:

Received Oct., 21, 2025
Revised Dec., 5, 2025
Accepted Jan.,10, 2026

Keywords:

Cyber Security threats
CICIDS2017 dataset
Python simulation
Machine learning
Information security

ABSTRACT

Cybercrime and attacks due to brute force have increased in the last decades. Old-fashioned detection methods have become useless and unreliable. Researchers are now working on new ideas and ways to improve information security and privacy. One prominent strategy involves machine learning. The aim of this master's research is to assess how machine learning significantly impacts identifying severe cyber threats due to brute force attack. Thus, analysis was performed on the CICIDS2017 datasets. Furthermore, we used numerical analyses with Python programming to verify that machine learning enables enhanced performance and reliability in network threat detection. The computational simulations and theoretical evaluations indicated that the voting algorithm exhibited superior performance (in terms of accuracy, precision, recall, and F1-score) in identifying benign brute force assaults. Moreover, the voting algorithm demonstrated peak accuracy across three distinct categories of attacks. The Gradient Boosting (GB) model demonstrated the highest F1-score in relation to FTP and SSH brute force vulnerabilities. The peak accuracy levels, precision, recall, and F1-score recorded in the three experimental conditions were 99.1%, 75.4%, 59.6%, and 70.1% for the voting, second voting, decision tree (DT), and gradient boosting (GB) methodologies, respectively.

Corresponding Author:

Saja Abdulkareem
Information technology, Altınbaş University, Istanbul, Turkey
Email: sajaabdulkareem949@gmail.com

1. INTRODUCTION

The past few years saw an increased rate of digitalization and technological innovation that transformed industries, economies, and social life at a global scale. As much as these developments have brought about some clear gains, they have also come with major challenges in terms of information security and privacy of data. As the user base, organizations, and infrastructures attached to the World Wide Web continue to grow, so does the likelihood of a malicious activity engage sensitive information and digital resources [1] - [7]. The current situation is alarming as this situation is highlighted by Figure 1.1, which depicts the increasing rate of cyberattacks throughout the past decades and forecasts the even steep rise in the future [8]. A series of high-profile attacks have occurred in the past decade involving governments, infrastructures of critical importance and even private corporations all over the world. As examples, Stuxnet (2010) interfered with industrial control, ShaMoon (2012) attacked oil companies in Saudi Arabia, and the Mirai botnet (2016) used individual Internet of Things (IoT) devices. Most recently, the SolarWinds attack (2020) adversely affected big organizations, such as Microsoft and FireEye, and ransomware campaigns in the healthcare sector have been disrupting hospitals and jeopardizing patient information [9]. Table 1.1 captures some of the groundbreaking cases, and it shows how cyber threats impact various industries like energy, healthcare, and supply chains.

Table 1: Cyber Threats That Have Received Widespread Media Attention and Affected A Wide Range Of Infrastructures and Organizations Globally [9].

No.	Name of the Cyber Threat	The Targeted System	Year of Attack	Details
1	Stuxnet	ICS/SCADA	2010	Industrial programs were affected
2	ShaMoon	ICS/SCADA	2012	Oil firms in Saudi Arabia were influenced
3	BlackEnergy	Smart Grids	2015	Smart grids related to Ukraine's renewable energy systems
4	Mirai	IoT Equipment	2016	Connected devices that use CCTV cameras
5	WannaCry	CPS/Healthcare	2017	Healthcare systems and hospitals
6	SolarWinds	Large-Scale Systems	2020	Big organizations, like Microsoft and FireEye, and supply chain
7	Hospital Ransomware	CPS/Healthcare	2021	Healthcare organizations in France

Such attacks have more long-term effects than short term inconvenience. Breach of information may result into identity theft, corruption of records, loss of finances and long-term reputation damage [10]-[14]. To organizations, these repercussions further require a significant amount of investment in security controls aimed to mitigating phishing, malware, ransomware, insider threats, and other intrusion vectors. Figure 1.2. demonstrates how different forms of cyberattacks are common across the globe up to July 2020 [9]. As a rule, those attacks are either passive or active. Unlike an active attack that has damaging effects, a passive attack accesses information without any damaging effect at present. An active attack modifies, destroys or steals data. It is often more damaging [15].

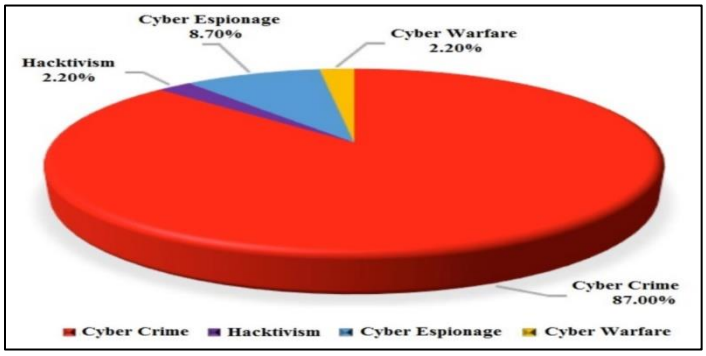

Figure 1: The common types of Cyber Attacks that occurred worldwide in July 2020 [9].

The most prevalent security instruments, such as antivirus software, intrusion detection systems, vulnerability assessment tools, and encryption protocols, despite their utility, have been inadequate in addressing recent cyberattacks. Most malware isn't precise, fast, or adaptable enough to counter the tactics employed by cybercriminals. As a result, machine learning, artificial intelligence, deep learning and neural network have become more popular and everyone's interest. Numerous methods have shown that one can efficiently detect cyberattacks through anomaly detection as it is more accurate, efficient, and reliable than conventional methods [16]. Due to the limitations of the traditional techniques which in a lot of achieving trustworthy and low-cost detection of these advanced attacks with a high-speed time and a high rate of accuracy [25], [26]. Modern technologies utilizing machine learning and deep learning are developing

methods to detect anomalies in network traffic, troubleshoot threats, and predict possible violations. These methods can adjust to any changes in the attack vector as the cyber world is not anything but hostile. Experts in research, innovation and computer security analysis have invested a lot of time in researching on effective and established techniques at the international level of business process' security, finances' security and data confidentiality's security [17]-[24]. Thus, it is clear that the study of ML, AI and DL is important to cybersecurity. Not only does this contribution provide a new theory to academia on the topic of digital defense mechanism but it has practical implication too. Researchers identified desired outcomes and useful information for various stakeholders dealing with cybersecurity issues. Paraphrase this as if a researcher is giving some recommendations in his research paper. As smart systems perform better, as well as more accurate, reliable and trustworthy they are a strong alternative to nickel status security tools and lets organizations start preparing for advanced cyber member.

2. RELATED WORKS

The smart grid is an important area in which machine learning-based intrusion detection systems (IDS) have made significant progress. Sahani et al. [27] The importance of authenticity in anomaly detection for energy systems has been highlighted by a detailed study machine learning-based Intrusion-Detection Systems in smart-grid computing architectures. The research discovered that a data-driven intrusion detection system can recognize attacks on SCADA systems that are the basis of the grid systems. In a similar vein, Almalaq et al. [32] The exploration augmented deep-learning (DL) detection system on cyberattack of smart power infrastructures. Deep Neural Architectures Learn Non-Linear Patterns of Power Consumption Data. As a result, These Attacks Predict & Prevent Cyber Attacks in Fuel Real-Time. These studies reveal that smart grids are one of the important places to employ AI-based cybersecurity technologies because they play a key role in social systems. Healthcare systems are using more medical devices and wireless technologies that are connected to the internet, making them targets of cyberattacks. The EWMCPS is a Wireless Medical Cyber-Physical System by Alzahrani et al. [28] The medical communication network is fortified against malicious traffic using a barrage of machine learning techniques. As said in the document, machine learning based methods can enhance system reliability and reduce the risk of potentially fatal events after a degradation of health information. Concurrently, Rashid et al. [30] The empirical examination clarified that the implementation of machine learning algorithms possesses the capability to identify cyber threats emanating from vulnerabilities associated with the Internet of Things (IoT) within healthcare applications specifically designed for intelligent urban settings. Moreover, an additional area of concern pertains to cloud computing, which serves as the essential infrastructure underpinning modern digital systems and services. El-Kassabi et al. [29] the research examined deep learning methodologies for the implementation of security within the orchestration of cloud workflows and presented a framework that employed deep architectures for the detection of real-time anomalies. The authors concluded that deep learning models outperformed conventional classifiers in their ability to adapt to the dynamic and distributed nature of cloud services. Concurrently, Alkahtani and Aldhyani. [31] concentrated on industrial control systems and food security systems and they suggested ML and DL models to enhance the cyber resilience of critical supply-chain networks. To identify SSH and FTP brute-force intrusion, Hossain et al. [34] introduced a hybrid approach, which is a combination of long short-term memory (LSTM) networks and traditionalML algorithms, and thus underscores the ability of temporal models to record sequential login attempts. A complementary research by Zhang et al. [35] proposed a black-box brute-force approach that can attack the ML-based systems themselves, leaving the issue of the adversarial robustness of intelligent IDS in question. Luxemburk et al. [37] also expanded the analyzing feature of packets at the packet-level to identify intrusion through brute-force attacks on HTTPS and showed that fine-grained traffic analysis increases the ability to detect intrusion at an early stage. Taken together, these works indicate that as the ML and DL methods continue to evolve defensive mechanisms, the opponents also create countermeasures thus highlighting the dual-use aspect of AI in cybersecurity. Data sets are critical towards facilitating successful training and assessment of IDS. Panwar et al. [38] and Boukhamla and Gaviro [39] developed a thorough research on the CICIDS-2017 data that is one of the most used benchmarks in cybersecurity research. Their findings showed that selection of features greatly enhance classification accuracy as well as the cut-off of the computational expenses which is an important consideration in applying IDS to real-time settings. Jairu and Mailewa [43] enhanced these findings by using supervised AI methods on the same data, whereas Imran et al. [44] the implementation of ensemble learning techniques that integrate multiple classifiers has been suggested to improve the accuracy of anomaly detection. Moreover, the concept of transfer learning has been further

developed to enrich the academic research associated with Intrusion Detection Systems (IDS). Otoum et al. [40] used transfer learning methods in intrusion detection of Internet of Vehicles (IoV) as the traditional IDS models fail frequently because of the dynamic and mobile characteristics of car networks. In the same manner, Kebede et al. [41] came up with predictive ML based models to prevent and identify the DDoS attacks by combining IDS and prevention strategies in multimedia intensive settings. In addition to domain-specific uses, complex surveys of the higher role of ML and DL in IDS have also been reviewed. Asharf et al. [33] took a systematic review of the IDS use in the IoT setting, whereby they noted the specific challenges, including data heterogeneity, scalability, and energy limitations of the IoT devices. They also stressed on the need to have light but powerful models with a balance between performance and resource conservation. An elaboration of this discussion is given by Lee et al. [36], who introduced a DL-capable IDS on software-defined networks (SDN), showing how the central control planes can be combined with intelligent anomaly detection to strengthen network agility and resilience. In addition, Dat -Thinh et al. [42] created a multistage idis, MidSiot, designed to operate on Ioot which integrated several steps of detecting data in order to obtain a greater accuracy and lesser false-positive results. These findings were still supported by Khan et al. [45] in their review of DL to IoT security regarding the fact that the current problems are interpretability of the model, adversarial vulnerabilities, and scaling to deployment. Collectively, these publications can solidify the position of the IDS as a key cybersecurity mechanism, where ML and DL solutions have become more and more important. Although this has gone a long way, there are still difficulties in the implementation of smart algorithms to deal with cybersecurity. The adversarial robustness of ML models is one of such issues, and Zhang et al. [35] explain that in such a way, ML-based systems themselves may be targeted by brute-force adversarial attacks. The scarcity of standardized data other than CICIDS 2017 is also a limitation since the real-life traffic data may be not accessible in many cases because of privacy and confidentiality issues. In addition, deep neural networks have computational overhead that poses a problem to deploy ability in resource-constrained systems like IoT. Overall, the literature on smart grids [27, 32], healthcare [28, 30], cloud systems [29], industrial applications [31], IoT [33, 42, 45], and vehicular networks [40] indicates the ubiquitous use of ML and DL in improving cybersecurity. Of particular interest to research has become the IDS that can combat a range of threats that start with brute-force intrusions [3437] and extend to massive DDoS attacks [41]. Even though this is the case, adversarial resilience, scale, and resource limitation continue to shape research agenda. We can say that, all the above pieces of evidence confirm that the intelligent algorithm integration in IDS is indeed a paradigm shift in cyber defense. However, constant innovation and multidisciplinary collaboration are required to keep pace with the innovative developments of new threats.

3. METHOD

The research method adopted in this research work explains the various machine learning and deep learning algorithms that are used for the identification of intrusions particularly brute force intrusions. The selection of datasets, preprocessing of raw data, implementation of learning algorithm and evaluation of their efficiency with the help of suitable metrics is the procedural framework. The following steps explain the process in a clear way.

3.1 Dataset Description

The empirical investigations were conducted utilizing the CICIDS2017 dataset, which is widely regarded as one of the foremost benchmarks for evaluating intrusion detection systems. The data was created within a managed network setting and has a wide range of benign and malicious actions. It consists of flow-based traffic records based on packet captures, and each such record has 80 features like the flow duration, type of protocol, source and destination bytes, flag counts and other features of behavior.In this research, a subsample of data was selected to use only the detection of brute-force attacks. Three categories of traffic were chosen:

- Traffic of good users,
- FTP Brute Force attacks, and
- SSH Brute Force attacks.

The subset therefore presents a binary classification problem the ability to differentiate between legitimate and brute-force traffic plus giving the models the opportunity to learn patterns related to two different brute-force protocols. The statistical data regarding the traffic across these three classifications is encapsulated in Table 3.1, while the frequency distribution pertaining to the various categories of attacks within the collected dataset is illustrated in Figure 2.

Table 2: Attack Classes in CICIDS2017 Dataset – Brute force Attack

Attack	Label	Frequency
Benign	Normal	667,626
FTP-Brute Force	Malicious	193,360
SSH-Brute force	Malicious	187,589

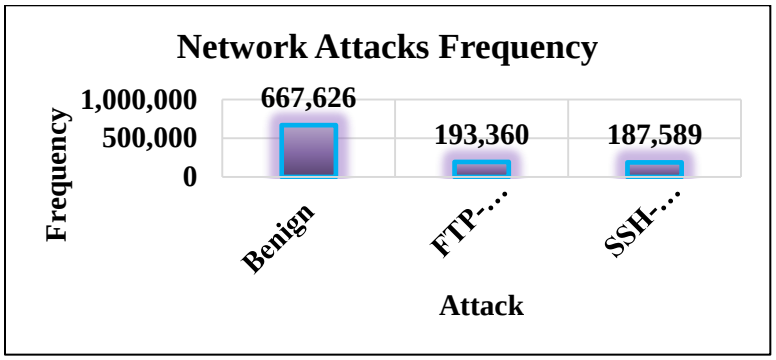

Figure 2: The Frequency of The Cyber Threats in The CICIDS2017 Dataset, Considering The Brute-Force Cyber Threat.

3.2 Data Preprocessing

Raw data used in a network intrusion detection system are usually heterogeneous, noisy and mixed. To have good quality input to the learning algorithms, several preprocessing steps were followed:

3.2.1 Label Encoding of Nominal Features:

Several attributes within the dataset, such as protocol type, service, or flag indicators, exhibit a categorical nature. Given the imperative for numerical data within machine learning frameworks, these categorical features were transformed into a numerical representation through the application of label encoding techniques. For instance, protocol designations, including TCP, UDP, and ICMP, were systematically represented as integers.

3.2.2 Timestamp Transformation:

The data contains flow start and end time, as normal date-time values. Their conversion into numerical numbers was done through the extraction of the applicable time-based elements like hours, minutes, and seconds. This modification enabled the incorporation of time behavioral patterns and yet compatible with the models.

3.2.3 Cleaning and Feature Expansion:

The process of encoding and timestamp processing resulted in the increase of the number of usable features by 85 features instead of the initial 80 features. These were some other features that improved the representation of the entire dataset on traffic characteristics. All the gaps or inconsistent values were imputed using imputation methods where the values were required and there were no null values in the record that could skew the model.

3.2.4 Balancing Considerations:

Even though non-dangerous flows prevail in the real world, this experiment favored a rather balanced dataset to promote efficient training. In this way, it can be ensured that the models will be able to learn the attack signatures without being bombarded with the regular traffic patterns.

3.3 Models Used

To solve the issue of the detection of brute-force intrusion attempts, this research utilised six supervised learning algorithms that included decision-tree-based learners, boosting techniques and a meta-classifier. The models that have been selected were selected because of their effectiveness in structured data analytics and previous studies on intrusion detection.

Decision Tree (DT)

A Decision Tree constitutes an approach of supervised learning which recursively divides the feature space as based on threshold criteria. The internal nodes represent the decision rule and the branches represent the outcome results and the terminal nodes represent the categorical predictions. In the framework of intrusion detection, decision trees are especially beneficial since they explicitly show which attributes, e.g. connection count or flow duration, are used to make the classification decision. They can be interpreted and trained quickly, thus making them suitable to large

datasets, such as CICIDS2017. However, decision trees are also vulnerable to overfitting particularly in the case of unlimited depth of the trees. Irrespective of this disadvantage, their simplicity provides a rich point of reference to which a more advanced strategy of an ensemble can be observed.

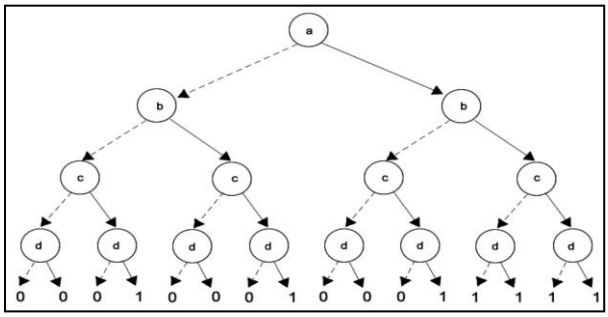

Figure 3: An Architecture Related to The Principles of Binary DT Algorithm [57].

Random Forest (RF)

The Random Forest algorithm represents a significant advancement over the elemental decision tree paradigm, which produces an ensemble of decision trees. To augment the variability among the individual models, each tree is constructed utilizing a bootstrap-sampled subset of the dataset alongside a randomly selected subset of features. The predictions are synthesized through a majority voting process. The bagging method helps to make models stronger, decreases variance, and lessens overfitting or fitting in decision trees. Random Forests are really good at capturing complex interactions between features. For example, if we wanted to do brute-force detection, they would link failed login attempts to packet-size distributions. Network traffic data often contains noisy and missing values; however, they are resilient to these issues. The primary trade-off is that this model has a larger computation cost during training, while its prediction cost is still low.

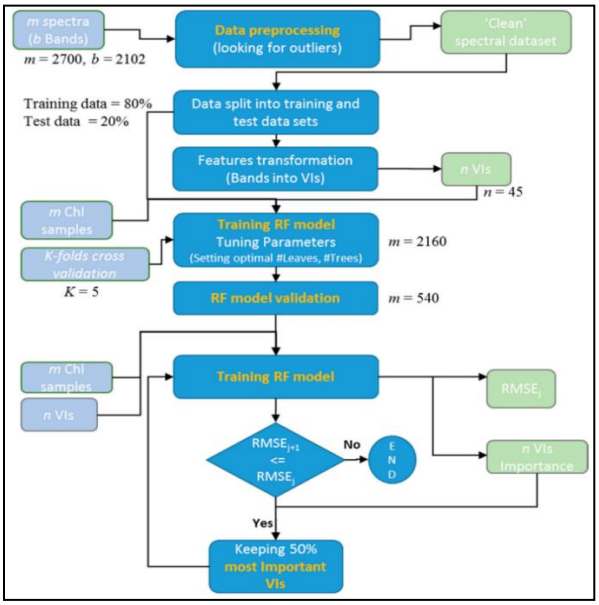

Figure 4: The Major Workflow and Critical Principles Related To The RF Algorithm [59].

Gradient Boosting (GB)

Gradient Boosting helps reduce the errors of the trees involved in the earlier iterations. The successive decision trees result from their combinations. The model can identify the smaller, more intricate patterns in the data because to this. Gradient boosting is very important in intrusion detection because the strong class here can be set as small but

systematic differences to normal that a weak learner fails sees at the beginning but Gradient boosting does see eventually. To improve classification, the model corrects itself by reducing errors. Nevertheless, the gradient boosting algorithm is prone to hyper-parameter configuration and improper tuning may result in over-fitting or excessive generalisation. However, its ability to outline intricate decisions makes it an effective network security application.

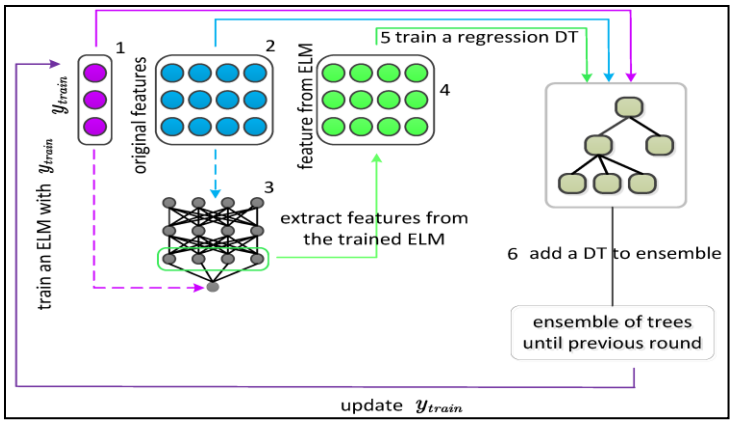

Figure 5: The Operation Principles and Techniques Considered In The GB Algorithm [62].

Extreme Gradient Boosting (XGBoost)

XGBoost is basically an enhanced version of gradient boosting that offers regularization and also deals with missing values and parallelized training. Many enhancements were added to the algorithm that greatly improves performance and scalability. These enhancements make XGBoost one of the top algorithms in competitive machine learning. XGBoost is great at identifying normal and malicious traffic in intrusion detection. It can handle high dimensionalities of features and class imbalance which is often present in cyber-attacks. The model has built-in ways to prevent overfitting. Furthermore, because it understands sparsity, it shows robustness to incomplete traffic flows as well. Also, the fast speed of the algorithm makes it useful for real-time applications that need fast detection.

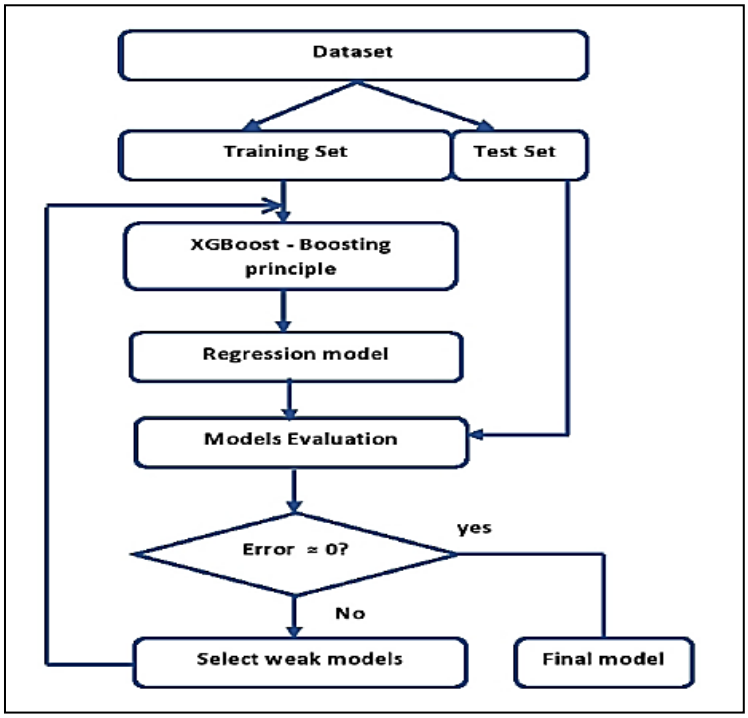

Figure 6: The Fundamental Operation Of The Xgboost Algorithm [58].

Adaptive Boosting (AdaBoost)

Ada Boost is a boosting technique, which at each boosting pass re-weights all misclassified examples, causing the weak learners to focus on more difficult examples. The ultimate forecast is established by the superficial decision trees that constitute the weak learners via a weighted plurality vote. The repeated focus on problem cases is useful in intrusion detection as some of those brute-force patterns resemble normal traffic closely and are difficult to classify. Although AdaBoost has a lot of advantages, it is still vulnerable to noisy data and outliers that can make it perform worse. But, the adaptive weighting system ensures that even the weak learners have their share in enhancing the overall ensemble accuracy.

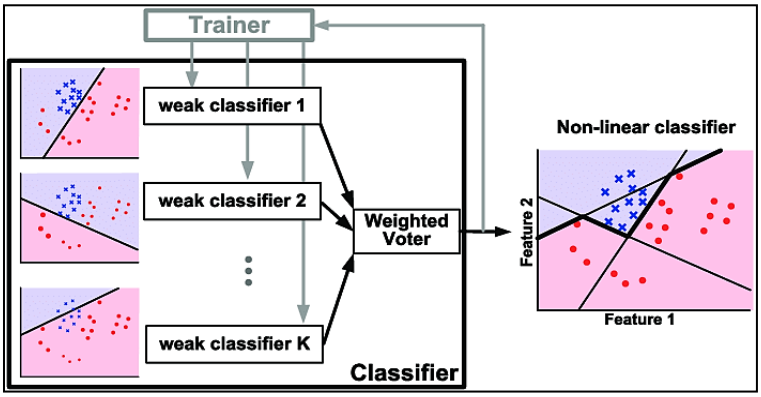

Figure 7: The Fundamental Working Principles And Operation Steps Of The Adaboost Algorithm [69].

Voting Classifier

The Voting Classifier constitutes a meta-algorithmic framework. It combines together the predictions made by many base classifiers. So, the Classifier Voting makes a final prediction. Majority Voting was performed using the outputs obtained from Decision Tree, Random Forest, Gradient Boosting, XGBoost, and AdaBoost. By amalgamating strategies it is possible to exploit the complementary merits of different algorithms. *For example, tree-based algorithms have been interpretable, boosting models have shown great predictive accuracy, and bagging ones are usually robust to variability.* A voting classifier takes the majority vote with several classifiers to create a more robust solution that is less sensitive to individual classifiers. In the case of intrusion detection, these hybrid techniques reduce false classification in different networks. An algorithm that makes predictions more reliable is worth the significant computation expenditure when the security of something is at stake.

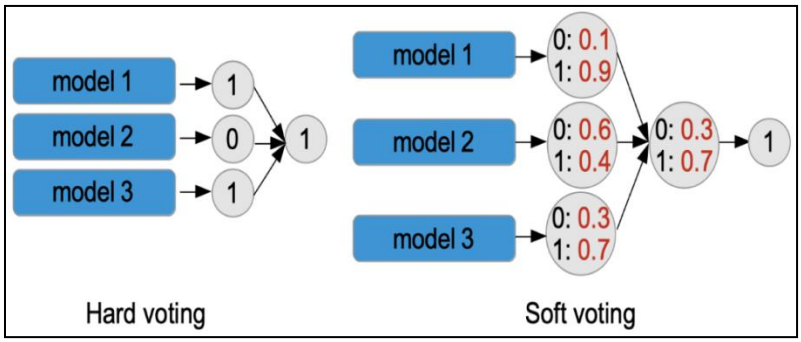

Figure 8: The Working Principles Related to The Voting Classifier Used in Detection Problems [75].

1.1 Performance Evaluation Metrics

The study used different assessment methods and evaluation techniques to validate the effectiveness and reliability of the six machine learning algorithms used in this research. To make this metrics easier to implement, several indicators were tackled within these dimensions:

Accuracy

The term accuracy refers to the ratio of correctly predicted data to the total number of data or the percentage of correctly predicted data which includes all of the data is the simplest performance measure. Secondly, accuracy could be explained as the ratio of correctly predicted data of the all models.
The accuracy can be calculated and assessed using the following formula:

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \times 100\% \quad (3.1)$$

Where:

TP : True Positives
TN : True Negatives
FP : False Positives
FN : False Negatives

Precision

A second term will be introduced to define and evaluate the effectiveness of the machine learning algorithms analyzed and discussed in this master's thesis. The metric refers to the Precision. This can be defined as the "correctly and accurately identified positive outcome in relation to the target positive outcome". The precision metric (in percentage ratio) can be computed using the below-mentioned formula:

$$Precision = \frac{(TP)}{(TP + FP)} \times 100\% \quad (3.2)$$

Recall

In addition to the two above-mentioned tools, there can be another tool which can be used for ascertaining the effectiveness of the machine learning algorithms proposed in this study. The recall is known as the fraction of positive correctly identified divided by the total number of positive cases. We can illustrate this third evaluative standard through the following equation:

$$Recall = \frac{(TP)}{(TP + FN)} \times 100\% \quad (3.2)$$

F1-Score

The F1-score is the fourth and last measurement used to assess the performance and reliability of the six machine learning algorithms in this study. The F1-score is the harmonic mean of the recall and precision score based on the number of true instances. This fourth evaluative criterion can be expressed by:

$$F1Score = \frac{2 \times (Precision \times Recall)}{(Precision + Recall)} \times 100\% \quad (3.4)$$

4. RESULTS AND DISCUSSION

This study carried out large numerical simulations to evaluate how well different six models work in non-malicious brute-force attacks. The six different machine learning algorithms performance with respect to accuracy, precision, recall and f1 score regarding benign brute-force cyber event are shown in table 3.

Table 3: Performance Findings of The Six Algorithms in Detecting Benign Attacks in The CICIDS2017 Dataset.

Type of Model	Accuracy	Precision	Recall	F1-Score
DT	92.8%	59.8%	50.3%	55.3%
RF	91.6%	65.2%	50.8%	58.1%
GB	94.7%	69.1%	52.4%	60.4%
XGB	93.1%	67.4%	52.5%	59.3%
AdaBoost	93.8%	67.3%	52.7%	59.7%
Voting	95.4%	74.6%	53.2%	61.2%

The most accurate voting algorithm had a 95.4% accuracy. Gradient boosting model was second with 94.7% accuracy. As a result, the outcomes obtained in the performance evaluation of six different machine learning techniques for classifying benign cyber brute force assaults indicate the effectiveness of the algorithms in recognizing harmful actions. Moreover, the results show that the precision, recall, and F1 score metrics remain optimized in the voting model. The results are illustrated in Figure 9 along with additional information pertaining to the results.

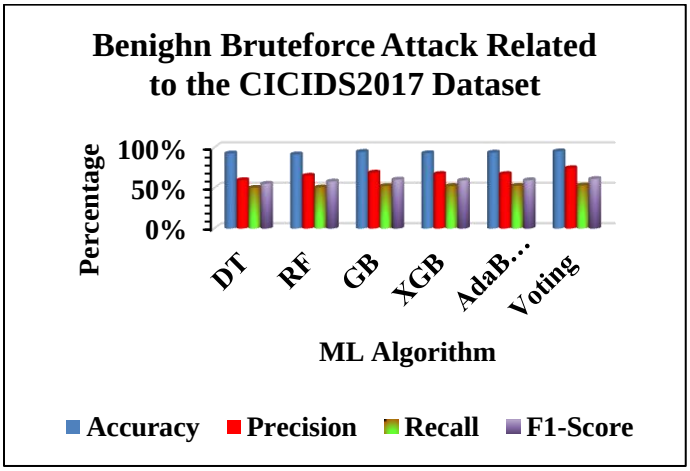

Figure 9: A Graphical Illustration of Performance Results Connected with The Six Algorithms Detecting the Benign Brute Force Attacks in The CICIDS2017 Dataset.

The data in Figure 9 indicates that accuracy had the largest values of each of the metrics, followed by precision, F1-score and finally the recall metric. The accuracy of random forest technique is least effective among all techniques. On the contrary, Decision Rule scheme produced the least proportions of precision, recall and F1 score. In Table 4.2, we present the numerical results of accuracy, precision, recall and F1-score of the six machine learning algorithms for this type of attack.

Table 4: Results of The Six Algorithms in Detecting FTP-Brute force Attacks In The CICIDS2017 Dataset.

Type of Model	Accuracy	Precision	Recall	F1-Score
DT	96.9%	73.6%	57.4%	57.5%
RF	95.8%	64.5%	56.7%	57.8%
GB	97.1%	68.2%	53.8%	68.4%
XGB	96.5%	66.7%	58.7%	60.7%

AdaBoost	97.9%	68.6%	55.9%	63.4%
Voting	98.8%	72.2%	54.6%	62.6%

Based on the results of the performance assessment on the effectiveness six machine learning models in detecting FTP-brute force attacks related to the CICIDS2017 dataset proved that the voting model gave the highest accuracy . The AdaBoost model followed with accuracy rates at a rate of 98.8% and 97.9% respectively. Also, the metrics of precision, recall, and F1-score peaked in the decision tree (DT), the DT in the second execution, and GB models, respectively, the findings supported further. In Figure 10, you can see a visual representation coupled with additional information regarding these findings.

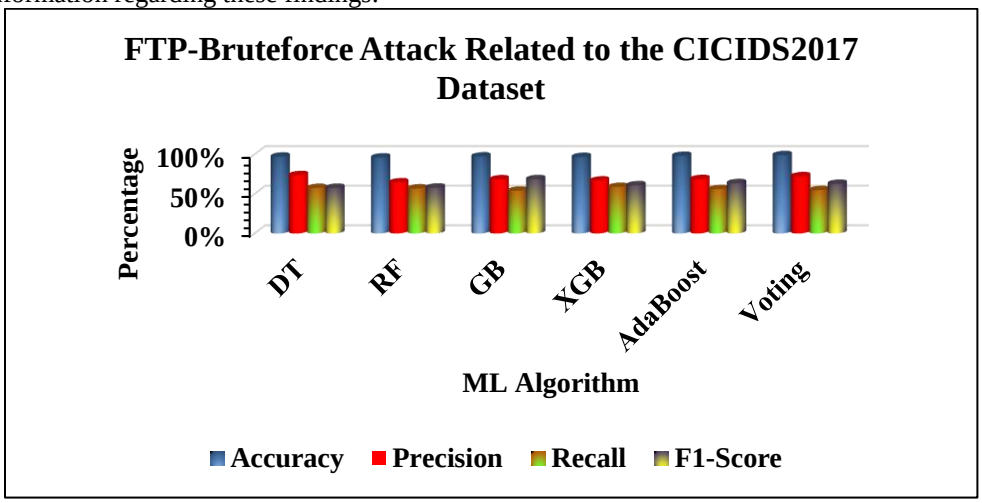

Figure 10: A Graphical Illustration of Performance Results Connected with The Six Algorithms Detecting The FTP-Brute force Attacks In The CICIDS2017 Dataset.

It can be inferred from the quantitative results depicted in Figure 4.2 that the most significant proportions are associated with accuracy, followed by precision, F1-score, and subsequently the recall metric. The Random Forest (RF) algorithm exhibited the lowest effectiveness regarding accuracy (95.8%). Concurrently, the Decision Rule (DR) model recorded the minimal values for both precision and recall. Moreover, the least values of the F1-score metric were achieved in the Decision Tree (DT) algorithm. Table 5 clarifies the quantitative results concerning accuracy, precision, recall, and F1-score for the six machine learning algorithms pertinent to this category of attack.

Table 5: Results of The Six Algorithms In Detecting SSH-Brute force Attacks In The CICIDS2017 Dataset.

Type of Model	Accuracy	Precision	Recall	F1-Score
DT	97.4%	72.2%	59.6%	58.6%
RF	96.2%	65.6%	57.8%	59.7%
GB	98.0%	67.5%	54.3%	70.1%
XGB	97.2%	63.6%	57.9%	62.6%
AdaBoost	96.8%	67.9%	56.0%	64.5%
Voting	99.1%	75.4%	55.8%	63.4%

It can be inferred from the results of the performance evaluation concerning the efficacy of the six machine learning algorithms in detecting SSH-brute force attacks associated with the CICIDS2017 dataset that the highest accuracy rate was achieved by the voting algorithm (with a rate of 99.1%), followed closely by the Gradient Boosting (GB) model (98.0%), and subsequently the Decision Tree (DT) model (with an accuracy of 97.4%). The findings further corroborated that the metrics of precision, recall, and F1-score reached their peak values within the voting, DT, and GB models, respectively. A graphical representation and additional information pertinent to these findings are depicted in Figure 4.3.

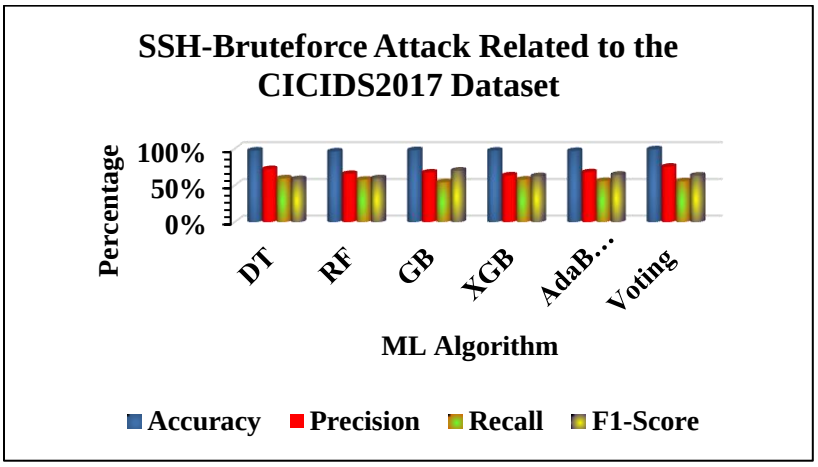

Figure 12: A Graphical Illustration of Performance Results Connected With The Six Algorithms Detecting The SSH-Brute force Attacks In The CICIDS2017 Dataset.

It can be noted from the numerical outputs represented in Figure 4.3 that the more significant percentages were related to the accuracy, followed by precision, F1-score, and then recall metric (like the two previous scenarios). The RF algorithm attained the minimum performance in terms of accuracy (96.2%). Meantime, the XGB model recorded the minimum proportions of precision (63.6%). The minimum recall was recorded at 54.3%. Further, the minimum ratios of the F1-score metric were attained in the DT algorithm (58.6%).To provide further clarification regarding the effectiveness of the machine learning (ML) and deep learning (DL) algorithms scrutinized in this research, a comparative evaluation was performed to determine the optimal accuracy, precision, recall, and F1-Score of the six algorithms employed in the identification of diverse brute force cyber threats. The findings of this evaluation are depicted in Table 4.4.

Table 7: Recorded Maximum Ratios of Four Assessment Metrics Recorded In The Three Brute force Attacks.

Type of Attack	Maximum Rates of				Which ML Algorithm Provided this Highest Rate in:			
	Accuracy	Precision	Recall	F1-Score	Accuracy?	Precision?	Recall?	F1-Score?
Benign Brute force	95.4%	74.6%	53.2%	61.2%	Voting	Voting	Voting	Voting
FTP-Brute force	98.8%	73.6%	58.7%	68.4%	Voting	DT	XGB	GB
SSH-Brute force Attack	99.1%	75.4%	59.6%	70.1%	Voting	Voting	DT	GB

It is indicated from the results represented in Table 4.4 that the ultimate ratios recorded in case of accuracy corresponded with the voting algorithm (in the three cases of brute force attack). At the same time, the voting

algorithm gave the ultimate values of the four metric ratios only in the benign brute force cyber-Attack. A graphical representation of those data is shown in Figure 4.4.

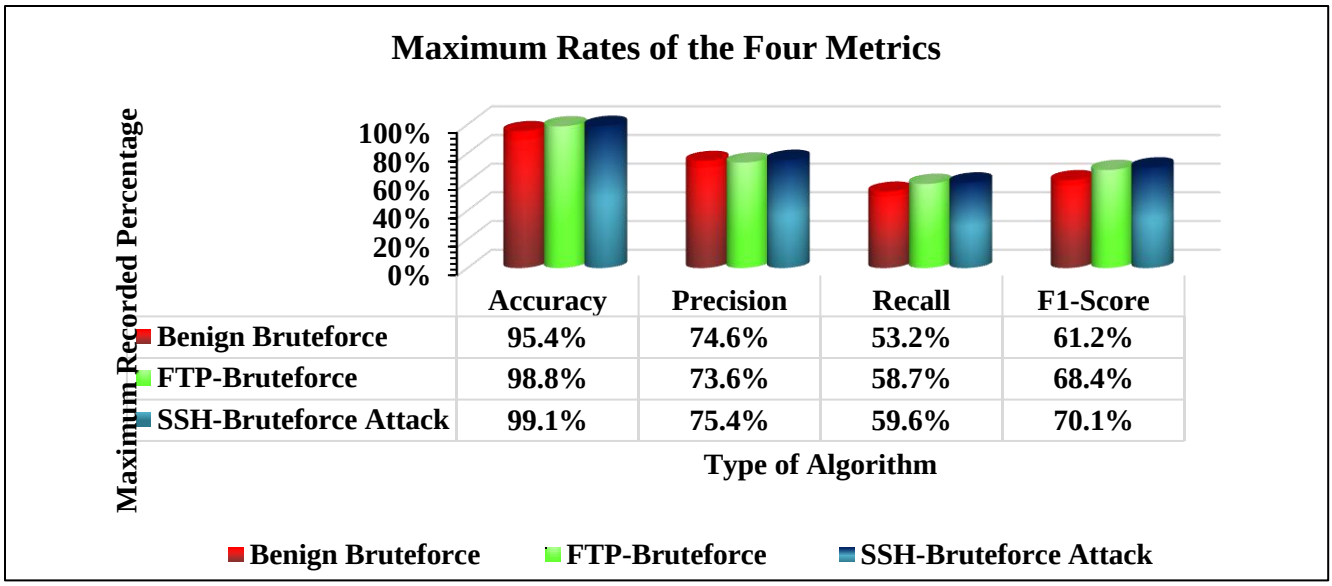

Figure 13: Illustration of Ratios of Four Assessment Metrics Recorded In The Three Brute force Attacks.

It is derived from the quantitative analyses elucidated in Figure 4.4 that the Decision Tree (DT) model exhibited the highest levels of precision in the context of the FTP-brute force attack. At the same time, the voting mechanism showed the best precision ratio for identifying SSH-brute force attacks. Also, based on the research outcomes of this study, it can be seen that the XGBoost (XGB) model produced a higher recall ratio in FTP-brute force detection. On the other hand, the Decision Tree (DT) model achieved the highest recall rates for Secure Shell (SSH) brute force intrusion detection. Also, the F1-score metrics in both detection contexts, specifically FTP and SSH brute force attacks, are better with GB model. The simulation results from this study show that the Voting Classifier is far more effective and reliable in detecting brute-force attacks compared to the other five models tested. According to the set criteria assessment metric, the Voting mechanism achieved a score accuracy of 99.1%, thus proving superior results. The Decision Tree and Gradient Boosting algorithms produced similar results with a precision of 75.4, a recall of 59.6, and an F1 of 70.1. A more detailed analysis shows that the FTP brute-force attack recall is highest with the XGBoost model while the SSH brute-force attack recall is achieved with the Decision Tree model. The F1-scores of the Gradient Boosting algorithms were quite high regarding the FTP and SSH brute-force attacks. Machines that work with algorithms can tell one brute-force method from the other. This may have uses in network security. Such findings may be placed within the scope of existing literature [34]– [47], The research looks into how effective various machine-learning and deep-learning models are at detecting complex cyberattacks. Some of the models we are going to use are Decision Tree, Logistic Regression, XGBoost, K-Nearest Neighbours, and more. According to our research results, advanced ML designs can be incorporated into the databases, without causing businesses any financial losses and also protecting other forms of data from a brute-force attack as well as other advanced attacks.

5. CONCLUSION

The study explains the use of machine learning and deep learning techniques for identifying brute-force attacks in the financial institution in Turkey that needs a significant amount of data security. The analysed attacks are Benign, FTP brute-force, and SSH brute-force. The research employed six different predictive modelling techniques such as Decision Tree (DT), Random Forest (RF), Gradient Boosting (GB), XGBoost (XGB), AdaBoost, and Classifier Ensemble (Voting). Results found using real analytical processes, the Voting ensemble performed best overall with a score of 99.1 for accuracy, 75.4 for precision, 59.6 for recall, and F1-score of 70.1. On the other hand, GB achieved the highest F1 scores for both FTP and SSH attacks, while DT had the highest recall for SSS attacks. Given the results of investigation, it is suggested to implement above mentioned algorithms. Moreover,

organizational structures should take on these methods. Further experiments should reconsider existing datasets, while existing data should be labelled for efficient processing.

ACKNOWLEDGEMENT
I dedicate this work to my family.

REFERENCES

[1] A. Djenna, S. Harous, and D. E. Saidouni, "Internet of Things Meet Internet of Threats: New Concern Cyber Security Issues of Critical Cyber Infrastructure," *Applied Sciences*, vol. 11, no. 10, p. 4580, May 2021, doi: 10.3390/app11104580.
[2] H. Habibzadeh, B. H. Nussbaum, F. Anjomshoa, B. Kantarci, and T. Soyata, "A survey on cybersecurity, data privacy, and policy issues in cyber-physical system deployments in smart cities," *Sustain Cities Soc*, vol. 50, p. 101660, Oct. 2019, doi: 10.1016/j.scs.2019.101660.
[3] F. Badrouchi *et al.*, "Cybersecurity Vulnerabilities in Biomedical Devices: A Hierarchical Layered Framework," in *Internet of Things Use Cases for the Healthcare Industry*, Cham: Springer International Publishing, 2020, pp. 157–184. doi: 10.1007/978-3-030-37526-3_7.
[4] B. Venkatesh and J. Anuradha, "A Review of Feature Selection and Its Methods," *Cybernetics and Information Technologies*, vol. 19, no. 1, pp. 3–26, Mar. 2019, doi: 10.2478/cait-2019-0001.
[5] N. Chowdhury and V. Gkioulos, "Cyber security training for critical infrastructure protection: A literature review," *Comput Sci Rev*, vol. 40, p. 100361, May 2021, doi: 10.1016/j.cosrev.2021.100361.
[6] C. Ma, "Smart city and cyber-security; technologies used, leading challenges and future recommendations," *Energy Reports*, vol. 7, pp. 7999–8012, Nov. 2021, doi: 10.1016/j.egyr.2021.08.124.
[7] H. Mehraj *et al.*, "Protection motivation theory using multi-factor authentication for providing security over social networking sites," *Pattern Recognit Lett*, vol. 152, pp. 218–224, Dec. 2021, doi: 10.1016/j.patrec.2021.10.002.
[8] Fleck A., "Cybercrime Expected To Skyrocket in Coming Years," *Statista*, 2022.
[9] A. Djenna, S. Harous, and D. E. Saidouni, "Internet of Things Meet Internet of Threats: New Concern Cyber Security Issues of Critical Cyber Infrastructure," *Applied Sciences*, vol. 11, no. 10, p. 4580, May 2021, doi: 10.3390/app11104580.
[10] O. K. Tosun, "Cyber-attacks and stock market activity," *International Review of Financial Analysis*, vol. 76, p. 101795, Jul. 2021, doi: 10.1016/j.irfa.2021.101795.
[11] M. G. Porcedda, "Patching the patchwork: appraising the EU regulatory framework on cyber security breaches," *Computer Law & Security Review*, vol. 34, no. 5, pp. 1077–1098, Oct. 2018, doi: 10.1016/j.clsr.2018.04.009.
[12] V. Wang, H. Nnaji, and J. Jung, "Internet banking in Nigeria: Cyber security breaches, practices and capability," *Int J Law Crime Justice*, vol. 62, p. 100415, Sep. 2020, doi: 10.1016/j.ijlcj.2020.100415.
[13] S. Hasan, M. Ali, S. Kurnia, and R. Thurasamy, "Evaluating the cyber security readiness of organizations and its influence on performance," *Journal of Information Security and Applications*, vol. 58, p. 102726, May 2021, doi: 10.1016/j.jisa.2020.102726.
[14] M. Parasol, "The impact of China's 2016 Cyber Security Law on foreign technology firms, and on China's big data and Smart City dreams," *Computer Law & Security Review*, vol. 34, no. 1, pp. 67–98, Feb. 2018, doi: 10.1016/j.clsr.2017.05.022.
[15] S. Tufail, I. Parvez, S. Batool, and A. Sarwat, "A Survey on Cybersecurity Challenges, Detection, and Mitigation Techniques for the Smart Grid," *Energies (Basel)*, vol. 14, no. 18, p. 5894, Sep. 2021, doi: 10.3390/en14185894.
[16] M. Wazid, A. K. Das, V. Chamola, and Y. Park, "Uniting cyber security and machine learning: Advantages, challenges and future research," *ICT Express*, vol. 8, no. 3, pp. 313–321, Sep. 2022, doi: 10.1016/j.icte.2022.04.007.
[17] J. Hua, Y. Chen, and X. (Robert) Luo, "Are we ready for cyberterrorist attacks—Examining the role of individual resilience," *Information & Management*, vol. 55, no. 7, pp. 928–938, Nov. 2018, doi: 10.1016/j.im.2018.04.008.
[18] J. I. Alcaide and R. G. Llave, "Critical infrastructures cybersecurity and the maritime sector," *Transportation Research Procedia*, vol. 45, pp. 547–554, 2020, doi: 10.1016/j.trpro.2020.03.058.
[19] Y. Li and Q. Liu, "A comprehensive review study of cyber-attacks and cyber security; Emerging trends and recent developments," *Energy Reports*, vol. 7, pp. 8176–8186, Nov. 2021, doi: 10.1016/j.egyr.2021.08.126.
[20] R. Knight and J. R. C. Nurse, "A framework for effective corporate communication after cyber security incidents," *Comput Secur*, vol. 99, p. 102036, Dec. 2020, doi: 10.1016/j.cose.2020.102036.
[21] H. S. Lallie *et al.*, "Cyber security in the age of COVID-19: A timeline and analysis of cyber-crime and cyber-attacks during the pandemic," *Comput Secur*, vol. 105, p. 102248, Jun. 2021, doi: 10.1016/j.cose.2021.102248.
[22] T. Tam, A. Rao, and J. Hall, "The good, the bad and the missing: A Narrative review of cyber-security implications for australian small businesses," *Comput Secur*, vol. 109, p. 102385, Oct. 2021, doi: 10.1016/j.cose.2021.102385.
[23] N. F. Syed, S. W. Shah, R. Trujillo-Rasua, and R. Doss, "Traceability in supply chains: A Cyber security analysis," *Comput Secur*, vol. 112, p. 102536, Jan. 2022, doi: 10.1016/j.cose.2021.102536.
[24] K. Kimani, V. Oduol, and K. Langat, "Cyber security challenges for IoT-based smart grid networks," *International Journal of Critical Infrastructure Protection*, vol. 25, pp. 36–49, Jun. 2019, doi: 10.1016/j.ijcip.2019.01.001.
[25] A. Corallo, M. Lazoi, and M. Lezzi, "Cybersecurity in the context of industry 4.0: A structured classification of critical assets and business impacts," *Comput Ind*, vol. 114, p. 103165, Jan. 2020, doi: 10.1016/j.compind.2019.103165.

[26] I. Agrafiotis, J. R. C. Nurse, M. Goldsmith, S. Creese, and D. Upton, "A taxonomy of cyber-harms: Defining the impacts of cyber-attacks and understanding how they propagate," *J Cybersecur*, vol. 4, no. 1, Jan. 2018, doi: 10.1093/cybsec/tyy006.
[27] N. Sahani, R. Zhu, J.-H. Cho, and C.-C. Liu, "Machine Learning-based Intrusion Detection for Smart Grid Computing: A Survey," *ACM Transactions on Cyber-Physical Systems*, Jan. 2023, doi: 10.1145/3578366.
[28] A. Alzahrani, M. Alshehri, R. AlGhamdi, and S. K. Sharma, "Improved Wireless Medical Cyber-Physical System (IWMCPS) Based on Machine Learning," *Healthcare*, vol. 11, no. 3, p. 384, Jan. 2023, doi: 10.3390/healthcare11030384.
[29] H. T. El-Kassabi, M. A. Serhani, M. M. Masud, K. Shuaib, and K. Khalil, "Deep learning approach to security enforcement in cloud workflow orchestration," *Journal of Cloud Computing*, vol. 12, no. 1, p. 10, Jan. 2023, doi: 10.1186/s13677-022-00387-2.
[30] M. M. Rashid, J. Kamruzzaman, M. M. Hassan, T. Imam, and S. Gordon, "Cyberattacks Detection in IoT-Based Smart City Applications Using Machine Learning Techniques," *Int J Environ Res Public Health*, vol. 17, no. 24, p. 9347, Dec. 2020, doi: 10.3390/ijerph17249347.
[31] H. Alkahtani and T. H. H. Aldhyani, "Developing Cybersecurity Systems Based on Machine Learning and Deep Learning Algorithms for Protecting Food Security Systems: Industrial Control Systems," *Electronics (Basel)*, vol.11,no.11,p.1717, May 2022, doi: 10.3390/electronics11111717.
[32] A. Almalaq, S. Albadran, and M. A. Mohamed, "Deep Machine Learning Model-Based Cyber-Attacks Detection in Smart Power Systems," *Mathematics*, vol. 10, no. 15, p. 2574, Jul. 2022, doi: 10.3390/math10152574.
[33] J. Asharf, N. Moustafa, H. Khurshid, E. Debie, W. Haider, and A. Wahab, "A Review of Intrusion Detection Systems Using Machine and Deep Learning in Internet of Things: Challenges, Solutions and Future Directions," *Electronics (Basel)*, vol. 9, no. 7, p. 1177, Jul. 2020, doi: 10.3390/electronics9071177.
[34] M. D. Hossain, H. Ochiai, F. Doudou, and Y. Kadobayashi, "SSH and FTP brute-force Attacks Detection in Computer Networks: LSTM and Machine Learning Approaches," in *2020 5th International Conference on Computer and Communication Systems (ICCCS)*, May 2020, pp. 491–497. doi: 10.1109/ICCCS49078.2020.9118459.
[35] S. Zhang, X. Xie, and Y. Xu, "A Brute-Force Black-Box Method to Attack Machine Learning-Based Systems in Cybersecurity," *IEEE Access*, vol. 8, pp. 128250–128263, 2020, doi: 10.1109/ACCESS.2020.3008433.
[36] T.-H. Lee, L.-H. Chang, and C.-W. Syu, "Deep Learning Enabled Intrusion Detection and Prevention System over SDN Networks," in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, Jun. 2020, pp. 1–6. doi: 10.1109/ICCWorkshops49005.2020.9145085.
[37] J. Luxemburk, K. Hynek, and T. Cejka, "Detection of HTTPS Brute-Force Attacks with Packet-Level Feature Set," in *2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC)*, Jan. 2021, pp. 0114–0122. doi: 10.1109/CCWC51732.2021.9375998.
[38] S. Singh Panwar, Y. P. Raiwani, and L. S. Panwar, "Evaluation of Network Intrusion Detection with Features Selection and Machine Learning Algorithms on CICIDS-2017 Dataset," *SSRN Electronic Journal*, 2019, doi: 10.2139/ssrn.3394103.
[39] A. Boukhamla and J. C. Gaviro, "CICIDS2017 dataset: performance improvements and validation as a robust intrusion detection system testbed," *International Journal of Information and Computer Security*, vol. 16, no. 1/2, p. 20, 2021, doi: 10.1504/IJICS.2021.117392.
[40] Y. Otoum, Y. Wan, and A. Nayak, "Transfer Learning-Driven Intrusion Detection for Internet of Vehicles (IoV)," in *2022 International Wireless Communications and Mobile Computing (IWCMC)*, May 2022, pp. 342–347. doi: 10.1109/IWCMC55113.2022.9825115.
[41] S. D. Kebede, B. Tiwari, V. Tiwari, and K. Chandravanshi, "Predictive machine learning-based integrated approach for DDoS detection and prevention," *Multimed Tools Appl*, vol. 81, no. 3, pp. 4185–4211, Jan. 2022, doi: 10.1007/s11042-021-11740-z.
[42] N. Dat-Thinh, H. Xuan-Ninh, and L. Kim-Hung, "MidSiot: A Multistage Intrusion Detection System for Internet of Things," *Wirel Commun Mob Comput*, vol. 2022, pp. 1–15, Feb. 2022, doi: 10.1155/2022/9173291.
[43] P. Jairu and A. B. Mailewa, "Network Anomaly Uncovering on CICIDS-2017 Dataset: A Supervised Artificial Intelligence Approach," in *2022 IEEE International Conference on Electro Information Technology (eIT)*, May 2022, pp. 606–615. doi: 10.1109/eIT53891.2022.9814045.
[44] Imran, F. Jamil, and D. Kim, "An Ensemble of Prediction and Learning Mechanism for Improving Accuracy of Anomaly Detection in Network Intrusion Environments," *Sustainability*, vol. 13, no. 18, p. 10057, Sep. 2021, doi: 10.3390/su131810057.
[45] A. R. Khan, M. Kashif, R. H. Jhaveri, R. Raut, T. Saba, and S. A. Bahaj, "Deep Learning for Intrusion Detection and Security of Internet of Things (IoT): Current Analysis, Challenges, and Possible Solutions," *Security and Communication Networks*, vol. 2022, pp. 1–13, Jul. 2022, doi: 10.1155/2022/4016073.
[46] V. Ravi, R. Chaganti, and M. Alazab, "Recurrent deep learning-based feature fusion ensemble meta-classifier approach for intelligent network intrusion detection system," *Computers and Electrical Engineering*, vol. 102, p. 108156, Sep. 2022, doi: 10.1016/j.compeleceng.2022.108156.
[47] P. Kumar, R. Kumar, G. P. Gupta, R. Tripathi, A. Jolfaei, and A. K. M. Najmul Islam, "A blockchain-orchestrated deep learning approach for secure data transmission in IoT-enabled healthcare system," *J Parallel Distrib Comput*, vol. 172, pp. 69–83, Feb. 2023, doi: 10.1016/j.jpdc.2022.10.002.
[48] Gartner, "What Is a Brute Force Attack?," *Fortinet, Inc.*, 2023.

[49] S. S. Shetty, R. R. Shetty, T. G. Shetty, and D. J. D'Souza, "Survey of hacking techniques and it's prevention," in *2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)*, Sep. 2017, pp. 1940–1945. doi: 10.1109/ICPCSI.2017.8392053.
[50] A. Das and P. Pathak, "Risk assessment and mitigation techniques of cyber attacks in emerging technologies," 2022, p. 030042. doi: 10.1063/5.0109616.
[51] M. Shafiq, Z. Gu, O. Cheikhrouhou, W. Alhakami, and H. Hamam, "The Rise of 'Internet of Things': Review and Open Research Issues Related to Detection and Prevention of IoT-Based Security Attacks," *Wirel Commun Mob Comput*, vol. 2022, pp. 1–12, Aug. 2022, doi: 10.1155/2022/8669348.
[52] J. P. Barrowclough and R. Asif, "Securing Cloud Hypervisors: A Survey of the Threats, Vulnerabilities, and Countermeasures," *Security and Communication Networks*, vol. 2018, pp. 1–20, Jun. 2018, doi: 10.1155/2018/1681908.
[53] R. Banda, J. Phiri, M. Nyirenda, and M. M. Kabemba, "Technological Paradox of Hackers Begetting Hackers: A Case of Ethical and Unethical Hackers and their Subtle Tools," *Zambia ICT Journal*, vol. 3, no. 1, pp. 40–51, Mar. 2019, doi: 10.33260/zictjournal.v3i1.74.
[54] M. Jena, R. K. Behera, and S. Dehuri, "Hybrid Decision Tree for Machine Learning: A Big Data Perspective," 2022, pp. 223–239. doi: 10.1007/978-981-16-8930-7_9.
[55] J. Wang, P. Li, R. Ran, Y. Che, and Y. Zhou, "A Short-Term Photovoltaic Power Prediction Model Based on the Gradient Boost Decision Tree," *Applied Sciences*, vol. 8, no. 5, p. 689, Apr. 2018, doi: 10.3390/app8050689.
[56] M. A. Hafeez, M. Rashid, H. Tariq, Z. U. Abideen, S. S. Alotaibi, and M. H. Sinky, "Performance Improvement of Decision Tree: A Robust Classifier Using Tabu Search Algorithm," *Applied Sciences*, vol. 11, no. 15, p. 6728, Jul. 2021, doi: 10.3390/app11156728.
[57] P. Santos *et al.*, "ImplicPBDD: A New Approach to Extract Proper Implications Set from High-Dimension Formal Contexts Using a Binary Decision Diagram," *Information*, vol. 9, no. 11, p. 266, Oct. 2018, doi: 10.3390/info9110266.
[58] S. K. Kiangala and Z. Wang, "An effective adaptive customization framework for small manufacturing plants using extreme gradient boosting-XGBoost and random forest ensemble learning algorithms in an Industry 4.0 environment," *Machine Learning with Applications*, vol. 4, p. 100024, Jun. 2021, doi: 10.1016/j.mlwa.2021.100024.
[59] S. H. Shah, Y. Angel, R. Houborg, S. Ali, and M. F. McCabe, "A Random Forest Machine Learning Approach for the Retrieval of Leaf Chlorophyll Content in Wheat," *Remote Sens (Basel)*, vol. 11, no. 8, p. 920, Apr. 2019, doi: 10.3390/rs11080920.
[60] S. Georganos, T. Grippa, S. Vanhuysse, M. Lennert, M. Shimoni, and E. Wolff, "Very High Resolution Object-Based Land Use–Land Cover Urban Classification Using Extreme Gradient Boosting," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 4, pp. 607–611, Apr. 2018, doi: 10.1109/LGRS.2018.2803259.
[61] M. Al-kasassbeh, M. A. Abbadi, and A. M. Al-Bustanji, "LightGBM Algorithm for Malware Detection," 2020, pp. 391–403. doi: 10.1007/978-3-030-52243-8_28.
[62] Y. Zou and C. Gao, "Extreme Learning Machine Enhanced Gradient Boosting for Credit Scoring," *Algorithms*, vol. 15, no. 5, p. 149, Apr. 2022, doi: 10.3390/a15050149.
[63] S. Raschka, J. Patterson, and C. Nolet, "Machine Learning in Python: Main Developments and Technology Trends in Data Science, Machine Learning, and Artificial Intelligence," *Information*, vol. 11, no. 4, p. 193, Apr. 2020, doi: 10.3390/info11040193.
[64] Y. Abakarim, M. Lahby, and A. Attioui, "A Bagged Ensemble Convolutional Neural Networks Approach to Recognize Insurance Claim Frauds," *Applied System Innovation*, vol. 6, no. 1, p. 20, Jan. 2023, doi: 10.3390/asi6010020.
[65] W. J. Al-Mudhafar, M. A. Abbas, and D. A. Wood, "Performance evaluation of boosting machine learning algorithms for lithofacies classification in heterogeneous carbonate reservoirs," *Mar Pet Geol*, vol. 145, p. 105886, Nov. 2022, doi: 10.1016/j.marpetgeo.2022.105886.
[66] S. K. Kalagotla, S. v. Gangashetty, and K. Giridhar, "A novel stacking technique for prediction of diabetes," *Comput Biol Med*, vol. 135, p. 104554, Aug. 2021, doi: 10.1016/j.compbiomed.2021.104554.
[67] B. Jia and Y. Liang, "Anti-D chain: A lightweight DDoS attack detection scheme based on heterogeneous ensemble learning in blockchain," *China Communications*, vol. 17, no. 9, pp. 11–24, Sep. 2020, doi: 10.23919/JCC.2020.09.002.
[68] F. Zhang, Y. Wang, J. Ni, Y. Zhou, and W. Hu, "SAR Target Small Sample Recognition Based on CNN Cascaded Features and AdaBoost Rotation Forest," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 6, pp. 1008–1012, Jun. 2020, doi: 10.1109/LGRS.2019.2939156.
[69] Z. Wang, J. Zhang, and N. Verma, "Realizing Low-Energy Classification Systems by Implementing Matrix Multiplication Directly Within an ADC," *IEEE Trans Biomed Circuits Syst*, pp. 1–1, 2015, doi: 10.1109/TBCAS.2015.2500101.
[70] J. Sun, H. Li, H. Fujita, B. Fu, and W. Ai, "Class-imbalanced dynamic financial distress prediction based on Adaboost-SVM ensemble combined with SMOTE and time weighting," *Information Fusion*, vol. 54, pp. 128–144, Feb. 2020, doi: 10.1016/j.inffus.2019.07.006.
[71] Z. Long *et al.*, "Motor fault diagnosis using attention mechanism and improved adaboost driven by multi-sensor information," *Measurement*, vol. 170, p. 108718, Jan. 2021, doi: 10.1016/j.measurement.2020.108718.
[72] T. K. K. Ho, J. Gwak, C. M. Park, and J.-I. Song, "Discrimination of Mental Workload Levels From Multi-Channel fNIRS Using Deep Leaning-Based Approaches," *IEEE Access*, vol. 7, pp. 24392–24403, 2019, doi: 10.1109/ACCESS.2019.2900127.
[73] Kunal and M. Dua, "Attribute Selection and Ensemble Classifier based Novel Approach to Intrusion Detection System," *Procedia Comput Sci*, vol. 167, pp. 2191–2199, 2020, doi: 10.1016/j.procs.2020.03.271.

[74] I. Ahmad, M. Yousaf, S. Yousaf, and M. O. Ahmad, "Fake News Detection Using Machine Learning Ensemble Methods," *Complexity*, vol. 2020, pp. 1–11, Oct. 2020, doi: 10.1155/2020/8885861.
[75] A. Manconi, G. Armano, M. Gnocchi, and L. Milanesi, "A Soft-Voting Ensemble Classifier for Detecting Patients Affected by COVID-19," *Applied Sciences*, vol. 12, no. 15, p. 7554, Jul. 2022, doi: 10.3390/app12157554.
[76] M. Alsaedi, F. Ghaleb, F. Saeed, J. Ahmad, and M. Alasli, "Cyber Threat Intelligence-Based Malicious URL Detection Model Using Ensemble Learning," *Sensors*, vol. 22, no. 9, p. 3373, Apr. 2022, doi: 10.3390/s22093373.

BIOGRAPHIES OF AUTHORS

Author 1 picture	**Saja Abdulkareem** earned a master's degree from *INFORMATION TECHNOLOGY, ALTINBAS UNIVERSITY, ISTANBUL, TURKEY.*