**Journal of Computer Science and Technology Studies**
ISSN: 2709-104X
DOI: 10.32996/jcsts
Journal Homepage: www.al-kindipublisher.com/index.php/jcsts

JCSTS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

**| RESEARCH ARTICLE**

# Next-Generation Cyber Threat Detection and Mitigation Strategies: A Focus on Artificial Intelligence and Machine Learning

**Md Rasheduzzaman Labu[1] ✉ and Md Fahim Ahammed[2]**
*[12]Department of Information Assurance and Cybersecurity, Gannon University, Erie, Pennsylvania, USA.*
**Corresponding Author:** Md Rasheduzzaman Labu, **E-mail**: sourcerweb@gmail.com

**| ABSTRACT**
The principal objective of this research was to examine strategies for detecting and mitigating cyber threats in the next generation, by underscoring Artificial Intelligence (AI) and Machine Learning (ML). This study provides a comprehensive overview of the role of AI, ML, and deep learning (DL) in the domain of cybersecurity. Furthermore, this study highlights the benefits of integrating deep learning into cybersecurity practices. The researcher explored the effectiveness of consolidating AI and ML techniques into the Feedzai security system to reinforce the detection of fraudulent activities. To validate the methodology, the investigator experimented by employing the supervised machine learning random forest algorithm on a dataset comprising historical transaction records in CSV format. The results of the research ascertained that by employing Feedzai's AI-based software combined with the random forest algorithms, future financial institutions can achieve real-time fraud detection and accurate identification of legitimate transactions. The Random Forest framework had the highest accuracy rate, at 83.94%. By contrast, the Naïve Bayes framework had an accuracy rate of 79.23%, and the KNN model had the lowest accuracy rate, of 78.74%. These results ascertained that the Random Forest system was the most effective for pinpointing cyber-attacks.

**| KEYWORDS**

Cyber Threat Detection, Machine Learning, Deep Learning, Supervised Learning, Artificial Intelligence, Random Forest.

**| ARTICLE INFORMATION**

**ACCEPTED:** 02 February 2024 **PUBLISHED:** 13 February 2024 **DOI:** 10.32996/jcsts.2024.6.1.19

## 1. Introduction

According to Alaskar (2021), as cyber threats are escalating exponentially in scope and complexity, traditional signature-based detection and mitigation approaches are increasingly becoming inadequate in every generation. Cyber-criminals are finding more complex ways to maneuver present security controls via methods like obfuscated malware, zero-day exploits, and tailored social engineering. Hasan (2022), indicated that to counter these advanced threats, companies are turning to machine learning and artificial intelligence methods that can detect unknown and novel patterns of attack. This report explores how methods like anomaly detection, deep learning, natural language processing, and other AI/ML methods are being employed for next-generation cyber security mitigation. This report aims to discuss the opportunities and challenges in terms of developing, deploying, and scaling these new AI-powered strategies across organizations.

The principal objective of this study is to explore the transformative potential of the next generation of Artificial Intelligence and Machine Learning in the domain of cybersecurity. By examining the application of Artificial Intelligence and Machine Learning techniques in cyber-attack detection and mitigation, besides, this paper aims to provide insights into their efficiency and the emerging trends in this sector. Furthermore, this study aims to pinpoint the challenges and threats related to these technologies and provide recommendations for companies and researchers to leverage their benefits while mitigating potential drawbacks. The research questions can be presented as follows:

- What are the emerging technologies that are most likely to shape the future of next-generation cyber-attack detection and mitigation strategies, and how can Artificial Intelligence and Machine Learning adapt to these developments?
- How can machine learning and artificial intelligence be efficiently consolidated into the current cybersecurity infrastructures to develop next-generation cyber threat detection and mitigation strategies?
- How do Artificial Intelligence and Machine Learning-based cyber-attack detection and mitigation methods influence the overall preparedness and resilience of companies in the face of rapidly advancing cyber-attacks, and what are the best practices for their successful implementation?
- What are the key limitations and advantages of Artificial Intelligence and Machine Learning techniques in terms of reinforcing cyber attack detection and mitigation, and how do these technologies contrast to traditional approaches?

### *1.1 Background*

The ever-advancing domain of cyberspace presents an accelerating threat to the security of digital infrastructure, therefore, creating the need for efficient cyber-attack detection and prevention strategies more fundamental than ever (Dayyabu, 2023). Traditional methods have proven to be inadequate in terms of combating the dynamic and complex nature of contemporary cyber threats (Iliadis, 2021). In reaction to these obstacles, artificial intelligence (AI) and machine learning (ML) have emanated as game-changing technologies in the domain of cybersecurity. This research paper presents an extensive examination of the role of Artificial Intelligence and Machine Learning in cyber threat detection and mitigation, delving into their advantages, applications, and limitations.

### *1.2 Scope and Limitations*

This study revolves around Artificial Intelligence and Machine Learning methods in cyber-attack detection and mitigation, underscoring their application, advantages, techniques, and real-world implementations in the next generation. While the study aims to provide a comprehensive overview, the research also presents specific Artificial Intelligence and Machine Learning approaches used in cybersecurity. Furthermore, this study acknowledges the dynamic nature of cybersecurity and the advancing landscape of threats, which may result in some methods becoming outdated over time.

## 2. Literature Review

### *2.1 Overview of Traditional Cyber Threat Detection and Mitigation Strategies*

As per Islam et al. (2022), in the ever-advancing landscape of cybersecurity, traditional or mainstream techniques for cyber-attack detection and mitigation have been the cornerstone of many companies' security landscapes. While they have played an instrumental role in terms of protecting digital assets, these techniques have their limitations and are consistently being consolidated by more advanced methods, such as machine learning and artificial intelligence (Dayyabu, 2023). This section discusses the prevalent traditional cyber threat detection and mitigation strategies, outlining their key components, advantages, and challenges.

**a. Firewall Protection:** Mhlanga (2021), argues that Firewalls are one of the prevalent and most instrumental elements of traditional cybersecurity. They act as a barricade between a company's internal network and the external world, controlling and overseeing traffic as per the predefined rules. Firewalls are efficient in terms of blocking unauthorized access and safeguarding against known cyber-attacks by inspecting packet headers and deploying access control policies (Islam et al., 2022) Nevertheless, they struggle to detect and respond to advanced cyber-attacks that may exploit vulnerabilities or use encrypted traffic to bypass inspection.

**b. Antivirus Software:** Antivirus software is tailored to detect and eliminate known viruses and malware. These techniques adopt signature-based detection, contrasting codes and files against a database of known malicious signatures (Montalvo, 2021). While they are effective in terms of detecting and eliminating known dangers, they are less efficient against zero-day threats and complex malware that continuously evolves to evade signature-based detection (Hasan, 2022).

**c. Intrusion Detection Systems (IDS):** IDS are tailored to monitor system logs and network traffic, detecting suspicious activities and possible security attacks. They are designed in two main types, most notably, network-based IDS (NIDS) and host-based IDS (HIDS). NIDS passively regulates network traffic, while HIDS regulates activities on individual devices (Montalvo, 2021). IDS can provide alerts when they pinpoint anomalies or known attack trends. Nevertheless, they frequently generate false positives, making it difficult for security teams to differentiate between real threats and benign events (Montalvo, 2021).

**d. Security Information and Event Management (SIEM) Systems:** SIEM systems gather, assess, and correlate data from a myriad of sources, comprising alerts and logs developed by security solutions. They offer a centralized forum for regulating and resolving security incidents (Raghavendran, 2022). While SIEMs provide valuable insights and reporting capacities, they are not

always real-time resolutions and may struggle with the high volume of data produced by contemporary networks, possibly missing threats that require immediate attention (Islam et al., 2022).

**e. Access Control:** Access control interventions, such as authorization and authentications, restrict access to sensitive resources according to user credentials. While access regulation is pivotal for hindering unauthorized access, it does not directly address arising threats like phishing attacks, malware, or zero-day vulnerabilities. Furthermore, it highly depends on the assumption that user credentials have not been compromised.

**f. Patch Management:** Patch management is the method of maintaining software and operating systems up to date with the current security updates and patches. While this technique is instrumental for discarding known vulnerabilities, it depends on organizations promptly applying patches as soon as they are released. Delayed patching leaves systems vulnerable to exploitation (Ryman-Tubb, 2022).

**g. Security Awareness Training:** Ryman-Tubb (2022), asserts that human error persists as a significant factor in cyber-attacks. Traditional techniques entail educating employees regarding the risks of phishing, social engineering, and best practices for secure behavior. While training is pivotal, it does not guarantee complete protection, as even well-informed users can fall victim to complicated cyber-attacks.

### *2.2 Artificial Intelligence and Machine Learning in Cybersecurity*

Sarwat (2020), contends that machine learning and Artificial intelligence are subfields of computer science that focus on developing systems and algorithms capable of learning from data, detecting trends, and making decisions or predictions. In the setting of cybersecurity, Artificial Intelligence and Machine Learning technologies empower and reinforce systems to adapt, adjust, evolve, and enhance their threat detection and mitigation capabilities over time.

### *2.3 Advantages of AI and ML in Cybersecurity*

Artificial Intelligence and Machine Learning provide a myriad of key advantages in cybersecurity. In particular, they facilitate proactive threat identification, real-time analysis, and automated scenario response (Sarwat, 2020). Their capability to process large volumes of data and detect complex patterns makes them highly efficient against advancing cyber threats. Moreover, these technologies can minimize false positives, enhancing the overall efficiency of cybersecurity operations (Hasan, 2022).

### *2.4 The Role of Artificial Intelligence and Machine Learning in Improving Cyber Threat Detection and Mitigation*

Artificial Intelligence (AI) and Machine Learning (ML) are transforming the domain of cybersecurity by providing innovative techniques for cyber threat identification and mitigation. These technologies can adapt, adjust learn, and enhance over time, making them instrumental assets in the prevailing battle against cyber-attacks (Sarwat, 2020). The following section presents the role of Artificial Intelligence and Machine Learning in cyber threat detection and mitigation.

**1) Proactive Threat Detection:** Islam et al (2022), asserts that Artificial Intelligence and Machine Learning are exceptionally proficient at proactive threat detection. By constantly assessing system logs, network traffic, and user behaviors, these techniques can detect subtle and previously undetectable patterns that may present a cyber-attack. This proactive technique is specifically efficient in terms of detecting zero-day vulnerabilities and advanced persistent threats that evade traditional signature-based methods.

**2) Real-time Analysis:** Artificial Intelligence and Machine Learning facilitate real-time analysis of large volumes of data. This real-time capacity is pivotal in terms of pinpointing and resolving cyber-attacks as they unfold, minimizing the response time to security scenarios (Hasan, 2022). Rapid detection and response are pivotal in terms of minimizing the effects of cyberattacks and preventing data breaches.

**3) Anomaly Detection:** One of the noteworthy applications of Artificial Intelligence and Machine Learning in cybersecurity is anomaly detection. Machine Learning algorithms can instigate baselines of normal network and client behavior. Any deviations from these baselines generate alerts, showcasing possible security breaches (Shukur, 2022). These techniques are particularly effective in terms of identifying insider cyber-attacks and unauthorized access and detecting unusual patterns that may signify an ongoing attack (Shukur, 2022).

**4) Threat Intelligence Analysis:** Dayyabu (2023), holds that Artificial Intelligence and Machine Learning are pivotal in terms of analyzing large volumes of threat intelligence data from many sources. By consolidating threat feeds, open-source intelligence, and historical data, these techniques can pinpoint trends and correlations that assist in terms of comprehending and predicting advancing attacks. This analysis aids in creating more effective defense strategies.

**5) Behavior-based Analysis:** Behavior-based evaluation adopts Artificial Intelligence and Machine Learning to examine user and entity behavior in real time. By tracking how devices and users interact with network resources, these techniques can detect behavioral anomalies that may signal a cyber threat. The behavior-based analysis is valuable in terms of pointing out insider threats, compromised accounts, and lateral movement by cyber attackers within a network (Dayyabu, 2023).

**6) Malware Detection:** Artificial Intelligence and Machine Learning are adopted in malware identification to detect new, unknown malware strains. Machine Learning algorithms can learn to detect malicious behavior patterns and codes, even when traditional signature-based solutions fail (Dayyabu, 2023). These capacities are paramount in terms of identifying polymorphic and file-less malware that consistently evolves to evade detection.

**7) Automated Incident Response:** According to Hasan (2022), Artificial Intelligence and Machine Learning can automate scenario response processes, facilitating companies to react rapidly to identified threats. Automated responses may include isolating infected systems, hindering malicious traffic, and instigating forensics analysis. This capacity minimizes the manual workload on security teams and ensures consistent and swift responses.

## 3. Models and Methodology

### *3.1 Machine Learning Techniques for Mitigating Cyber-Threats*

Islam et al. (2022), indicated that Machine learning techniques are indispensable in combating numerous challenges related to managing large-scale business data. In the arena of cybersecurity, machine learning techniques leverage historical patterns of fraudulent behaviors to pinpoint and detect them in cyber activities. For instance, machine learning algorithms such as Random Forest, which employs a combination of decision trees, are deployed to develop decision trees for data classification. These techniques prove to be more effective in identifying fraudulent characteristics compared to human analysis (Dayyabu, 2023).

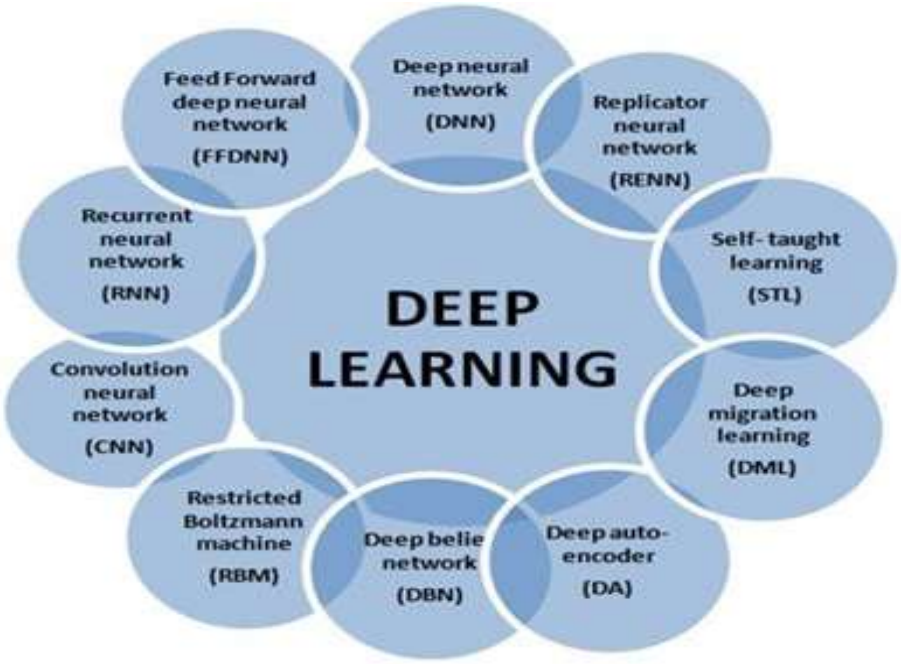


**Figure 1: Exhibits Machine Learning Approaches**

### *3.2 Usage of Deep Learning*

Deep Learning is a sub-section of Machine Learning that adopts Neural Networks, motivated by the structure of human neurons, to emulate human brain-like behavior. Deep Learning techniques emulate the functioning of human neurons and develop intricate neural architectures with sophisticated interconnections. Deep Learning can be classified into distinct types deemed as artificial neural nets or neural networks (Alaskar, 2021). These neural networks are arranged with strategic layers and are commonly termed CNN (Convolutional Neural Networks), frequently adopted for tasks associated with visual and pixel processing, and RNN (Recurrent Neural Networks). Two forms of machine learning algorithms are employed in cyber-attack detection: 1) Supervised, and 2) Unsupervised learning.

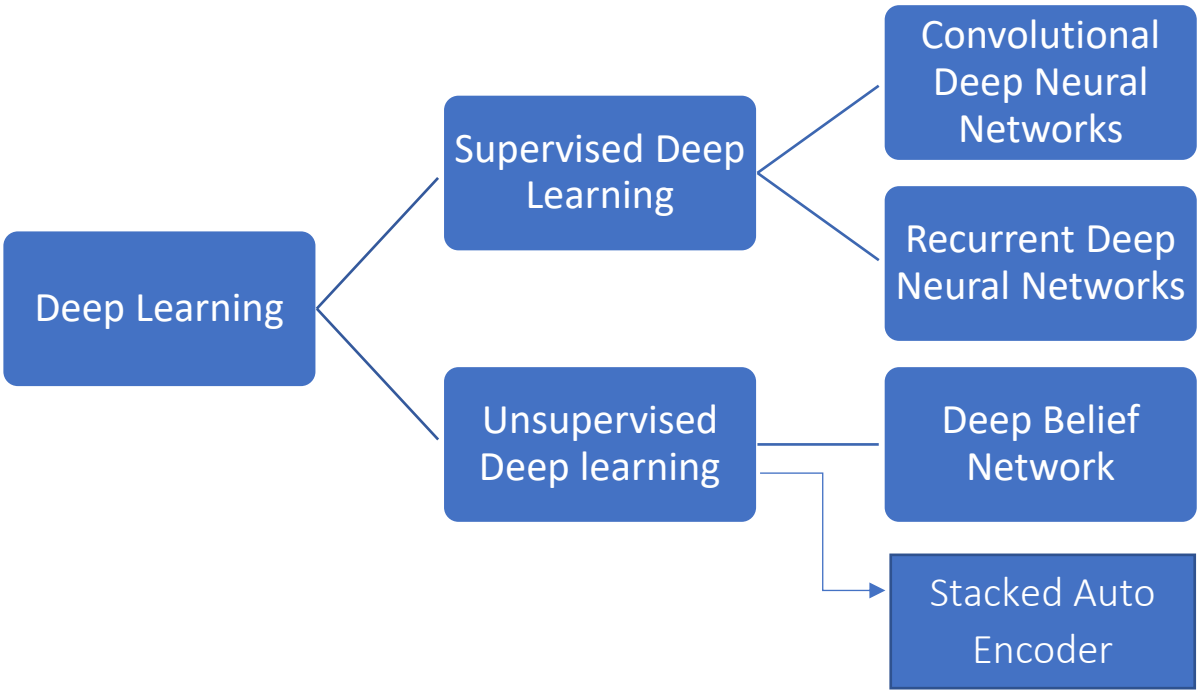


**Figure 2: Displays Deep Learning Categories**

***3.2.1 Supervised Deep Learning Framework***

Supervised Learning is an extensive machine learning methodology that facilitates the categorization and processing of data. It highly depends on the employment of labeled datasets, where every single input data point is designated a corresponding label signifying if it belongs to a "good" or "bad" category (Raghavendran, 2022). This labeled information acts as the premise for equipping a model to make accurate classifications or predictions on new, unseen data. The process of supervised learning entails feeding the framework with the designated dataset, enabling it to learn the underlying trends and associations between the input attributes and their corresponding labels. The framework then generalizes this knowledge to make predictions on new, unlabeled data points (Hasan, 2022).

The power of supervised learning lies in its capability to classify and discern data based on pre-defined classification. By offering candid labels to the training data, the model learns to pinpoint specific trends, attributes, or consolidation thereof that are indicative of the desired category (Raghavendran, 2022). This facilitates the framework to make informed decisions when provided with new, unlabeled data points, accurately classifying them as either "good" or "bad" based on the learned patterns.

***3.2.2 Unsupervised Deep Learning Model***

Ryman-Tubb (2022), contends that this technique revolves around the constant analysis and monitoring of data, combined with the ongoing update of the detection system premised on discoveries. By applying unsupervised learning algorithms, this technique is capable of pinpointing anomalous activities within the dataset. The system constantly inspects incoming data, searching for emerging trends and ascertaining if they align with fraudulent operations (Islam et al., 2022) This dynamic process facilitates the timely identification of new data patterns that may indicate fraudulent activities.

The unsupervised nature of deep learning enables the model to adjust to evolving fraud strategies, as it can detect and pinpoint anomalies without depending on pre-labeled fraudulent activities (Ryman-Tubb, 2022). This flexibility is specifically valuable in fraud detection, since forms of fraudulent activities constantly arise, making it difficult to uphold an up-to-date labeled dataset. By consistently inspecting data and utilizing the power of deep learning, this strategy reinforces the detection capacities in fraud prevention systems. It offers a dynamic and adaptive system to detect and counter fraudulent operations, facilitating companies to stay ahead of ever-evolving fraud schemes and protect their assets and customers.

### *3.2.3 Reinforcement Deep Learning Framework*

Ryman-Tubb (2022), contends that the reinforcement learning (RL) framework strengthens machines to distinguish optimal activities within a provided setting by constantly learning from their surroundings. These frameworks target pinpointing actions that minimize threats and maximize rewards by iteratively examining and exploiting the available alternatives. An instrumental aspect of Reinforcement Learning is the incorporation of a reinforcement feedback signal, which guides the learning process.

Through a trial-and-error process, the Reinforcement Learning framework examines different activities in different states and observes the culminating rewards. By adopting reinforcement signals, such as positive rewards or negative penalties, the framework gradually learns which activities are favorable and which ought to be avoided. Over time, the Reinforced Learning framework adapts its activities to select actions that lead to higher rewards and lower risks within the given context (Hasan, 2022).

### *3.3 Artificial Intelligence in Cyber-Security*

As per Ryman-Tubb (2022), Artificial Intelligence (AI) methods are comprehensively applied within the domain of cybersecurity to resolve various pivotal tasks, entailing vulnerability management, projecting breach risks, scenario response, threat exposure analysis, malware regulation, intrusion identification, and prevention. Islam et al (2022), argues that Artificial Intelligence is defined as a promising resolution to counter the escalating threat of cybercrimes. It holds the capability to proactively prevent and pinpoint anomalies related to fraud, significantly enhancing security measures.


**Figure 3: Exhibits the Application of Artificial Intelligence in Cyber-Security**

### *3.3.1 Artificial Intelligence and Machine Learning in Detecting Cyber Attacks*

#### *3.3.1.1 Machine Learning Process*

- **Preparation Phase:** This training phase comprises feeding the model with labeled data, where every input is related to a corresponding desired output. By offering this labeled data to the model, the system can identify the underlying trends and associations between the input and output. The framework employs the labeled data to train a system, which is a machine learning algorithm or a neural network. The objective is to tailor a model that can generalize its comprehension of the input-output association and accurately predict the output for new, unseen input data.

- **Evaluation Stage:** During the evaluation stage, distinct attributes provided in the dataset are cautiously considered, and their matching attributes are computed. These attributes grant valuable insights and quantitative info that can be utilized to evaluate and analyze particular aspects of the data. For instance, in a banking setting when evaluating a specific account, one significant attribute that can be computed is the average transaction amount related to that account. This attribute presents a measure of the typical monetary value involved in transactions associated with the account, offering insights into the account holder's spending patterns or financial behavior.

**Supervised Technique Using the Rain Forest Algorithms**

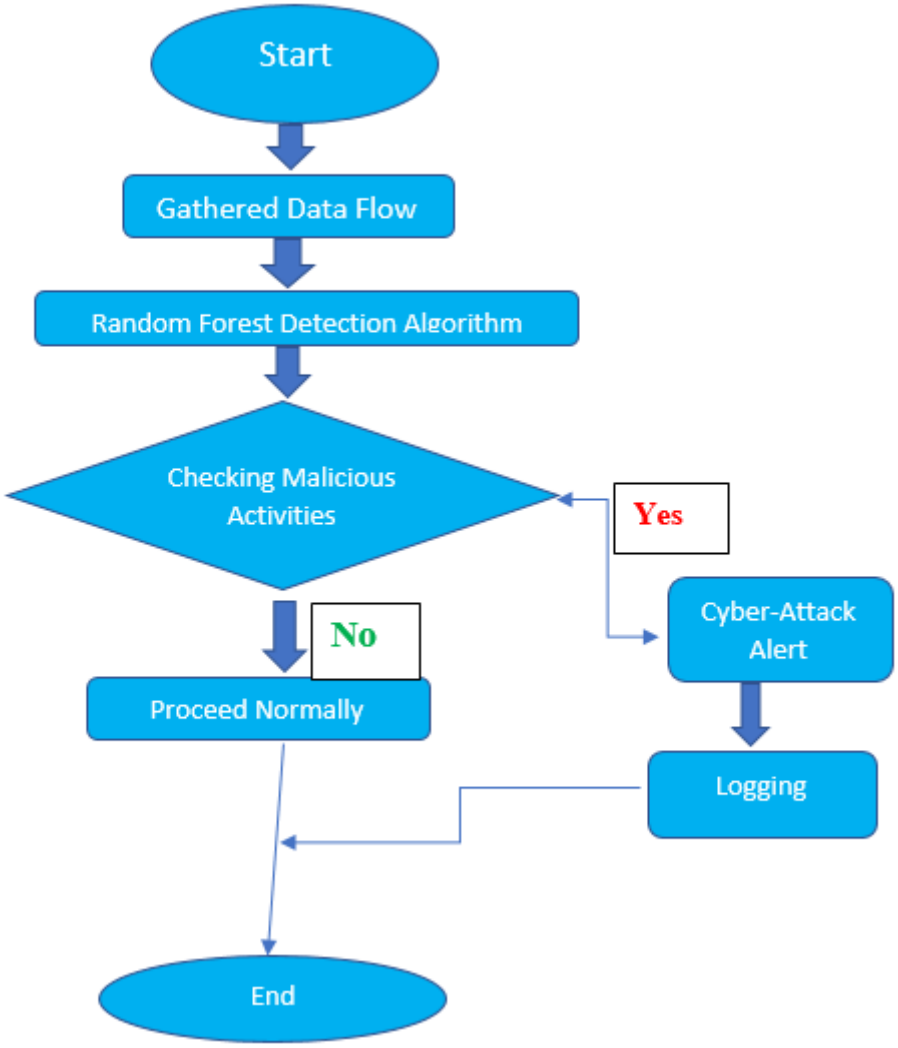


**Fig 4: Displays the Cyber-Attack Detection Using Rain Forest Algorithms**

The flowchart above showcases a cyber-attack detection system that employs a random forest algorithm. The system works as follows:

- **Step 1: Gathered data flow:** The system gathers data from different sources, such as system logs, network traffic, and security events.
- **Step 2: Random forest detection algorithm**: The model employs a random forest algorithm to assess and inspect the gathered data and detect potential cyber-attacks.
- **Step 3: Checking malicious activities:** The model subsequently checks for fraudulent activities as per the output of the random forest algorithm.
  - ***Yes***: If the model detects malicious activities, it proceeds to further analysis.
  - **No:** If the model does not detect malicious behavior, it proceeds to the next stage.

- **Step 4: Cyber-attack alert**- The framework sends a cyber-attack notification to the appropriate personnel.
- **Step 5: Logging:** The system logs the cyber-attack incident.
- **Step 6: End:** The process ends.

### *3.4 Random Forest Algorithms*

Random forests are an ensemble learning technique that employs multiple decision trees to make predictions. Rather than depending on a single decision tree, random forests integrate the outputs of multiple trees to reach a final prediction. Every decision tree in the random forest is trained and equipped on a random subset of the data, and the ultimate prediction is established by majority voting or averaging the predictions from individual trees.

To start the training process, the initial phase comprises importing the Random Forest Classifier category from the ensemble library. This category offers the essential functionalities for executing a random forest classifier. Subsequently, the decision tree classifier is tailored to the training set adopting the Random Forest Classifier category. The classifier is equipped with specific parameters. In this scenario, the parameter estimator is set to 50, showcasing the intended number of trees to be constituted in the random forest. The procedural parameter is set to "entropy," which indicates that the entropy function is adopted to measure the quality of each split in the decision trees. After the classifier is configured, it is equipped with the training data. The p-train and q-train variables denote the attribute matrix and the target variable, accordingly. The fit () function is then called on the classifier object, which starts the training process. The process can be broken down as follows:

1. To start, a random choice of N tuples is undertaken from the dataset D.
2. By adopting the information from the chosen tuples, a decision tree is designed.
3. Subsequently, the algorithm specifies the intended number of decision trees to be developed, and stages 1 and 2 are repeated for every tree.
4. For classification challenges, every decision tree predicts the class or type to which a record belongs.
5. The designated classification is determined for a new record premised on the predictions made by each decision tree.

To predict the result of the test set during the evaluation stage, the classifier is adopted. The classifier is a trained framework retrieved from the random forest algorithm. By employing the predict () function to the test data (x_test), the classifier produces predicted results (y_pred) that signify the predicted class or category for each record in the test set.

### *3.5 Fedzai Anomaly Cyber-Attack Detection*

In this research, the Feedzai Software tool is proposed to detect anomalies in cyber activities. Feedzai is considered one of the leading intelligent forums crafted to counter cyber-attacks. It maximizes the power of Artificial Intelligence to efficiently pinpoint and mitigate fraudulent activities within the banking sector. By adopting Artificial Intelligence frameworks, Feedzai's robust anomaly detection system is capable of identifying behaviors and patterns that may not be easily detectable to human observation. This facilitates the tool to accurately identify and stop fraudulent activities, therefore protecting financial institutions and their customers from potential losses and security breaches.

#### *3.5.1 Procedure:*

- **Phase 1**: The system begins by gathering the exchange data of the users and thoroughly analyzing their profiles and transaction practices.
- **Phase 2**: Subsequently, the system integrates external data sources and maximizes Feedzai's data consolidation capabilities to enhance the collected information.
- **Phase 3:** Adopt machine learning frameworks to monitor and assess the risk factors related to transactions and identify patterns indicative of financial crime.
- **Phase 4**: Employ the random forest model to combat the risk or presence of fraudulent transactions.

### *3.6 Integrating Fedzai Software AI-Based with Machine Learning*

- **Stage 1**: The client's original transaction data is inserted into the Feedzai machine learning (ML) engine.
- **Stage 2**: Data Set Preparation In this specific framework, a historical transaction dataset labeled "transaction.csv" is adopted. The dataset comprises several columns commonly adopted in fraud detection incidents, comprising Timestamp, Entry Mode, Amount, and Card present. Furthermore, there is a specified column labeled "Fraud target" that indicates whether a transaction is fraudulent or not.
- **Stage 3:** Model Training: The Machine Learning framework is trained using the prepared dataset and the provided features. The training process entails optimizing the model's parameters and learning trends from the historical transaction data. By evaluating the association between the input characteristics and the fraud target column, the frameworks learn to make predictions and pinpoint fraudulent transactions accurately.
- **Stage 4: Evaluating the Performance of the Model:** During the evaluation phase, the system is tested on a separate dataset to inspect its generalization capability. Performance metrics such as precision, accuracy recall, and the ROC curve are generated to assess and interpret the framework's efficiency in identifying fraudulent activities.

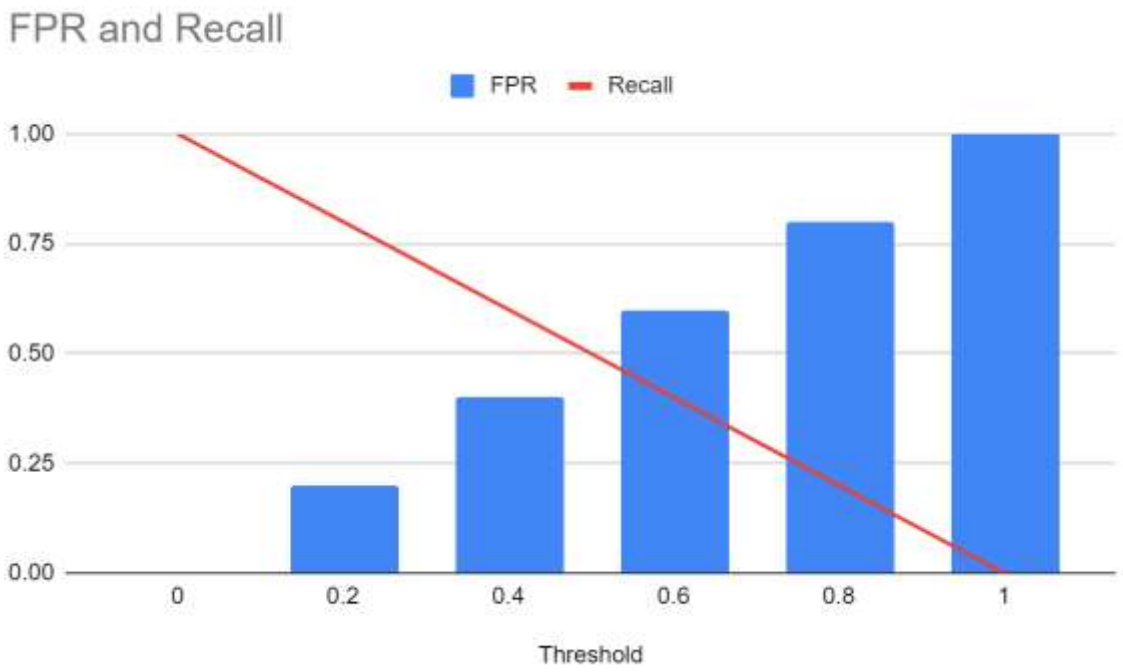


**Fig 5: Exhibits Evaluation of the Performance Model**

- **Stage 5:** Entails configuration of the Fraud Model External Scoring Service
  In this stage, the cyber-attack model's external scoring service is consolidated to seamlessly incorporate with the trained model. This consolidation facilitates real-time or batch scoring of new transactions to ascertain their likelihood of being fraudulent.
- **Step 6: Transaction Listing**-In the workflow, the case manager documents and highlights all the transactions and activities that have gone through scoring by the deployed model. These transactions are now in production and are being evaluated for possible fraud.

The objective of this phase is to organize and centralize the transactions that have been undertaken by the model, facilitating effective regulation and management of the fraud detection process. By highlighting these transactions, the case manager obtains visibility into the ongoing scoring activities and can track the outcomes or predictions generated by the framework for each transaction. This listing serves as a reference point for further analysis, investigation, or action on flagged or suspicious transactions.

**Table 1: Displays Models with their Respective Performance**

| Model | Cyber-Attack Detection Rate | Accuracy |
|---|---|---|
| **Naïve Bayes** | 48.62 | 79.23% |
| **KNN** | 47.46 | 78.74% |
| **Random Forest** | 52.65 | 83.94% |

The table above displays the accuracy rates of distinct cyber-attack detection models, most notably, KNN, Naïve Bayes, and Random Forest. The accuracy rate denotes the percentage of cyber-attacks that were correctly identified by the model. The Random Forest framework had the highest accuracy rate, at 83.94%. By contrast, the Naïve Bayes framework had an accuracy rate of 79.23%, and the KNN model had the lowest accuracy rate, of 78.74%. These results ascertained that the Random Forest system was the most effective for pinpointing cyber-attacks. This is likely because Random Forest frameworks are capable of learning sophisticated relationships between features, which makes them suitable for tasks such as detecting anomalies and fraud.

## 4. Conclusion

In this study, the researcher investigated next-generation cyber threat detection and mitigation strategies, most notably, a concentration on Artificial Intelligence and Machine Learning. From the research, the application of artificial intelligence (AI) and machine learning (ML) approaches, by consolidating the existing Feedzai security system, to reinforce the detection of fraudulent activities is portrayed. The study articulated a summary of the application of Artificial Intelligence and its associated technologies, such as Machine Learning and deep learning (DL), in the domain of cybersecurity. Furthermore, this research has conveyed the

advantages of integrating deep learning into cybersecurity practices. To substantiate our methodology, the investigator executed an experiment by employing the supervised machine learning random forest algorithm, as per the dataset entailing historical transaction records in CSV format. The outcomes of the study demonstrate that future financial institutions can attain real-time detection of fraud and accurate detection of genuine transactions by leveraging Feedzai's software and its open Machine Learning tool.

**Funding:** This research received no external funding.
**Conflicts of Interest:** The authors declare no conflict of interest.
**Publisher's Note**: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

**References**

[1] Alaskar, H. (2021). Deep Learning Approaches for Intrusion Detection in IoT Networks – opportunities and future directions. *www.academia.edu*. https://www.academia.edu/94625867/Deep_Learning_Approaches_for_Intrusion_Detection_in_IIoT_Networks_Opportunities_and_Future_Directions

[2] Dayyabu, Y. Y. (2023). The application of artificial intelligence techniques in credit card fraud detection: a quantitative study. *www.academia.edu*. https://www.academia.edu/103424871/The_application_of_artificial_intelligence_techniques_in_credit_card_fraud_detection_a_quantitative_study

[3] Hasan, M. R., & Ferdous, J. (2024). Dominance of AI and Machine Learning Techniques in Hybrid Movie Recommendation System Applying Text-to-number Conversion and Cosine Similarity Approaches. *Journal of Computer Science and Technology Studies, 6*(1), 94-102. https://doi.org/10.32996/jcsts.2024.6.1.10

[4] Hasan, M. R. (2022). Cybercrime Techniques in Online Banking. Int. J. of Aquatic Science, 13(1), 524-541. Retrieved from: https://www.journal-aquaticscience.com/article_158883.html (January 2022)

[5] Iliadis, L. (2021). AI threat detection and response on smart networks. *Duth*. https://www.academia.edu/84978312/AI_Threat_Detection_and_Response_on_Smart_Networks

[6] Islam, M., Z., Chowdhury, M., & Sarker, M., (2023). The Impact of Big Data Analytics on Stock Price Prediction in the Bangladesh Stock Market: A Machine Learning Approach. *International Journal of Science and Business, 28*(1), DOI: https://doi.org/10.58970/IJSB.pdf

[7] Mhlanga, D. (2021). Financial Inclusion in Emerging Economies: The application of machine learning and artificial intelligence in credit risk assessment. *Johannesburg*. https://www.academia.edu/81548617/Financial_Inclusion_in_Emerging_Economies_The_Application_of_Machine_Learning_and_Artificial_Intelligence_in_Credit_Risk_Assessment

[8] Montalvo, R. M. (2021). Design of a Dynamic and Self-Adapting System, Supported with Artificial Intelligence, Machine Learning, and Real-Time Intelligence for Predictive Cyber Risk Analytics in Extreme Environments – Cyber Risk in the Colonisation of Mars. *www.academia.edu*. https://www.academia.edu/83033349/Design_of_a_Dynamic_and_Self_Adapting_System_Supported_with_Artificial_Intelligence_Machine_Learning_and_Real_Time_Intelligence_for_Predictive_Cyber_Risk_Analytics_in_Extreme_Environments_Cyber_Risk_in_the_Colonisation_of_Mars

[9] Raghavendran, C. (2022). Cyber defense in the age of artificial intelligence and machine learning for financial fraud detection applications. *www.academia.edu*. https://www.academia.edu/82442603/Cyber_Defense_in_the_Age_of_Artificial_Intelligence_and_Machine_Learning_for_Financial_Fraud_Detection_Application

[10] Ryman-Tubb, N. F. (2022). How Artificial Intelligence and machine learning research impacts payment card fraud detection: A survey and industry benchmark. *Surrey*. https://www.academia.edu/40110568/How_Artificial_Intelligence_and_machine_learning_research_impacts_payment_card_fraud_detection_A_survey_and_industry_benchmark

[11] Sarker, I. H. (2022). Machine learning for intelligent data analysis and automation in cybersecurity: Current and Prospects. *Annals of Data Science, 10*(6), 1473–1498. https://doi.org/10.1007/s40745-022-00444-2

[12] Shahriar, S., Allana, S., & Hazratifard, S., (2023). A survey of privacy risks and mitigation strategies in the Artificial intelligence life cycle. *IEEE Journals & Magazine | IEEE Xplore.* https://ieeexplore.ieee.org/document/10155147?denied=

[13] Shukur, H. (2022). A survey on the role of artificial intelligence, machine learning, and deep learning for cybersecurity attack detection. *www.academia.edu*. https://www.academia.edu/75825384/A_Survey_on_the_Role_of_Artificial_Intelligence_Machine_Learning_and_Deep_Learning_for_Cybersecurity_Attack_Detection

[14] Zaman, S., Alhazmi, K., Aseeri, M., & Ahmed, M., (2021). Security Threats and Artificial Intelligence based Countermeasures for Internet of Things Networks: A Comprehensive survey. *IEEE Journals & Magazine | IEEE* Xplore. https://ieeexplore.ieee.org/document/9456954?denied=