IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 28, Issue 4, Ser. 1 (July. – August. 2026), PP 71-84
www.iosrjournals.org

A Hybrid Socio-Technical Framework For AI-Driven Cybersecurity Adoption In The Banking Sector: Integrating Adversarial Intrusion Detection With Organizational Readiness.

Sreenath PU
Phd Research Scholar, Department Of Computer Science
School Of Post Graduate Studies (SPGS)
Skyline University Nigeria

Dr. A. Senthilkumar
Dean & Associate Professor, Department Of Computer Science
School Of Science And Information Technology (SSIT)
Skyline University Nigeria.

Abstract
The rapid growth in the number of cyberattacks presents a very serious challenge in the banking industry whose survival and operations, customer satisfaction, and regulatory standards hold critical importance. The conventional intrusion detection systems (IDS), whose main detection mechanism is signature and rule based, is not very effective in identifying advanced threats like the zero-day attacks, polymorphic malwares, and backdoor intrusion. Consequently, AI-based solutions have become an attractive way of improving the real-time threat detection. Nonetheless, AI-based cybersecurity systems are not the only technical problem that can be adopted. Practically, there can often be sought, a limitation in implementation, both in organizational willingness, and in government organization, and requirements. The proposed research suggests a hybrid socio-technical approach that will combine AI-based IDS with organisational and regulatory aspects to facilitate successful cybersecurity practises in banks. The framework places AI-based IDS in the middle of the technical component and uses organisational preparedness and regulatory alignment as the facilitating technological factors. It also focuses on confrontational resilience, live agility, and learning in a governmental framework. The proposed model offers an organised way of deploying robust, scalable, and conforming AI-enabled cybersecurity systems by means of filling the divide between technological capacity and organisational preparedness to its deployment. The research has theoretical and practical implications on financial institutions, developers of AI, and regulators, and provides future empirical validation directions.

Keywords: *AI-driven cybersecurity; intrusion detection systems; socio-technical framework; banking sector; organizational readiness.*

Date of Submission: 19-07-2026 Date of Acceptance: 29-07-2026

I. Introduction

Background and Motivation

Artificial Intelligence (AI) has now become a pivotal part of a modern cybersecurity implementation, and even in the banking domain, where the reliability of information security, integrity of transactions, and data availability are vital. The Intrusion Detection Systems (IDS) enabled by AI can help in better detecting, classifying, and responding to cyber threats by utilising the capabilities of machine learning and deep learning, which is beyond the traditional rule-based Intrusion Detection Systems (IDSS) (Buczak and Guven, 2016; Sarker, 2021). There are traditional IDS techniques such as signature-based and anomaly-based system that have been in common use over the decades. Nevertheless, they have weaknesses in identifying the unknown attacks and frequently produce large numbers of false-positive (Axelsson, 2000; Chandola et al., 2009; Patcha and Park, 2007). These drawbacks underscore the importance of approaching even more adaptive and intelligent systems that will be able to learn emerging threat patterns. Development of deep learning has greatly enhanced the efficacy of intrusion detection because it can perform hierarchical learning on complex data (LeCun et al., 2015; Bengio et al., 2013; Shone et al., 2018). Although there have been such improvements, the practical implementation of AI-based IDS in banking settings is limited by both operational and organisational, as well as regulatory, considerations. As a reaction to increased cyber threats, regulatory authorities like Central Bank of Nigeria (CBN)

and global regulators like ENISA have developed a framework on cybersecurity that both financial institutions should be guided by (CBN, 2022; ENISA, 2023). These structures have focused on the governance, risk management, and resilience and hence the importance of systems that combine technical efficiency with institutional compliance underlining these requirements.

AI Capability-Adoption Gap in Banking

As the technologies of AI gain momentum, the gap between the organisational adoption and technical capacity is quite significant as well, as it is in the banking institutions. AI-based IDS prove to be very useful in controlled settings, but their application to real-life scenarios is usually suppressed by the constraints of infrastructures, human resources, and governance (Sharma et al., 2021; Sarker et al., 2020). Technically, AI models need large and high-quality datasets to train and validate them. The services, including UNSW-NB15 (Moustafa and Slay, 2015) and CICIDS-style datasets (Sharafaldin et al., 2018), have been popular datasets in research, but typically banking-related data is scarce because of privacy and regulation requirements. This inhibits the capability of organisations to exhaust the capabilities of AI. Meanwhile, the organisational preparedness is an essential factor in the definition of success adoption. Other factors that play a significant role in the implementation of AI-based cybersecurity solutions are leadership support, workforce skills, and IT infrastructure. Several financial institutions do not have professional knowledge needed to operate and sustain such systems successfully (Sharma et al., 2021). There are also regulatory requirements that make adoption difficult. The data governance, operational resilience, and transparency in banking institutions are guided by certain strict requirements stipulated by various frameworks, including COBIT 2019 (ISACA, 2022) and Basel Committee guidelines (Basel Committee on Banking Supervision, 2021). Such demands can in many cases constrain flexibility of AI model implementation, especially where interpretability and explainability is required.

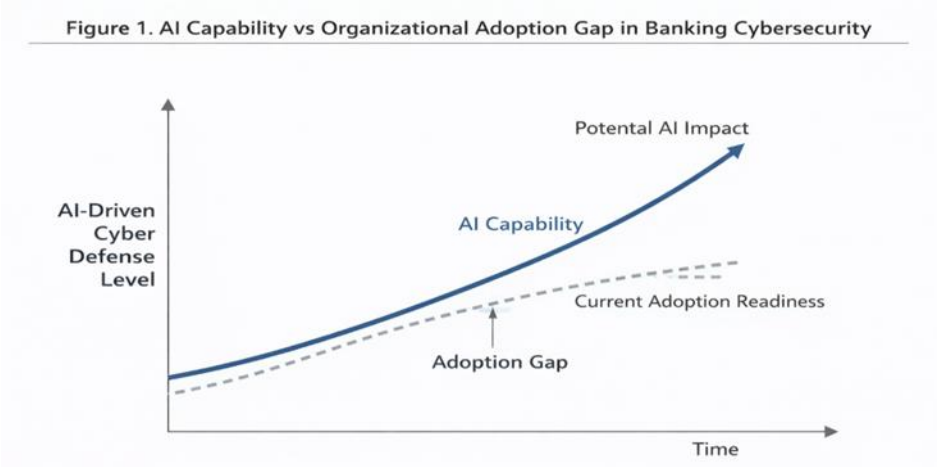

Figure 1: AI Capability vs Organizational Adoption Gap in Banking Cybersecurity

The Figure 1 shows the gap between the fast development of AI functions and the slower adoption rate in the organisational sphere, that is, the banking sector. Although AI systems are constantly getting better in terms of accuracy, scale, and automation, their successful implementation is limited due to organisational preparedness, regulatory compliance, and complexity of operations. This gap reveals an extremely important imbalanced situation: technological advances are not enough without the increasing number of changes in the institutional setting and governance systems. To bridge this gap, a socio-technical solution incorporating AI and organisational and regulatory aspects will be needed.

Research Problem and Conceptual Contribution

The main issue that is taken up in this paper is the lack of a single framework combining the technical aspect of intrusion detection in banks with the help of AI with the organisational and regulatory aspects of this procedure. Current AI-based IDS frameworks are mainly created and tested under controlled laboratory environments, in which the performance is determined using standardised datasets and metrics. Nevertheless, real-world complexities related to organisational constraints and adversarial threat are usually neglected in such models, despite the fact that they are essential elements of real business scenarios (Sommer and Paxson, 2010; Biggio and Roli, 2018).

There are identified three gaps:

1. Technical-Organizational Gap:

AI-based IDS can be extensively used and are efficient in the restricted environment of controlled settings, yet with limited organisational willingness, infrastructure, and experience in banking institutions.

2. Adversarial Vulnerabilities:

Adversarial attacks, including evasion and poisoning, on the AI systems can critically undermine the model performance (Biggio et al., 2012; Goodfellow et al., 2015).

3. Regulatory Constraints:

Banking policies have rigid standards in the use of data, transparency of the model, and accountability, which could restrict the use of complex AI systems (CBN, 2022; ENISA, 2023). The paper makes a contribution to the formation of a hybrid socio-technical approach as it offers a combination of AI-based IDS and the organisational readiness with regulatory compliance. The framework offers a systematic procedure in responding to the issue of cybersecurity and still makes it feasible in the banking settings.

Research Objectives

The specific objectives of this study are:

1. To review the limitations of traditional intrusion detection systems in banking environments.
2. To examine the organizational and regulatory factors influencing the adoption of AI-based cybersecurity systems.
3. To develop a hybrid socio-technical framework that integrates technical, organizational, and regulatory dimensions for AI-based intrusion detection.

II. Evolution And Taxonomy Of AI-Based IDS

The history of the intrusion detection systems (IDS) development has been constantly changing with the progress of cyber threats as more advanced protection systems are required. The first systems were largely rule-based, although developments in the area of artificial intelligence have moved the Tsun being intelligent, data-driven systems that have the capability to learn based on the trends and adjust to novel attack patterns.

Signature-Based to Intelligent IDS

Conventional intrusion detection systems were fairly founded on signature based techniques, which identify the attack by comparison of the arriving data and known patterns of malicious activity. These systems are good at detecting the already known risks but again are inherently constrained when face to face any new or emerging forms of attack (Axelsson, 2000; Debar et al., 1999). In order to overcome all these shortcomings, anomaly based IDS came into place. These systems ensure a normal behavioural pattern and identify its violations. On the one hand, they enhance the capability of detecting unfamiliar attacks but on the other hand, they are generally hard to false-positives and sensitive to noise in information (Chandola et al., 2009; Patcha and Park, 2007). The shortcomings of the two methods have lead to a movement towards smart IDS that apply artificial intelligence methods. Such systems can acquire patterns in data that are complex and adjust to changes in cyber threats, which is why they are more applicable in contemporary banking settings because attack patterns are dynamic and advanced (Sarker, 2021; Buczak and Guven, 2016).

Deep Learning and Machine Learning in IDS

Machine learning (ML) has been at the forefront of the development of intrusion detection because systems are able to automatically learn without explicit programming. The support vector machine, decision trees, and ensemble methods are common ML methods that have been employed in IDS and enhance detection accuracy and generalisation (Buczak and Guven, 2016; Dietterich, 2000).

Machine learning, in its turn, includes the deep learning (DL), which gives a boost to the performance of the IDS and allows finding hierarchical characteristics of large and multidimensional datasets. Deep neural networks and recurrent models have also shown high performance when it comes to detecting complex trends in network traffic and cyber activities (LeCun et al., 2015; Shone et al., 2018). Although they have their pros, ML and DL-based IDS do not lack challenges. They need large amounts of high-quality training information and may fall under the performance threshold after being subjected to adversarial attacks or biassed training data (Sarker et al., 2020; Biggio and Roli, 2018). Moreover, it is complicated by the fact that deep learning models are not interpretable since it is a barrier to the implementation into regulated settings like in the banking sector where transparency is extremely important.

Ensemble and Hybrid IDS Approaches

Ensemble and hybrid methods have been suggested in order to address the shortcomings of the single models. Ensemble techniques involve the use of more than one classifier to enhance the accuracy and strength of the prediction, and decrease the risk of fixing on one model (Dietterich, 2000; Lakshminarayan et al., 2017). The hybrid methods of IDS combine two or more methods, e.g. machine learning with rule-based IDS or unsupervised and supervised learning methods. The aim of these systems is to enjoy the benefits of various methods coupled with reducing the flaws hence leads to a higher detection rate and resistance to various forms of attacks.

Hybrid models are especially useful in the banking cybersecurity context because they entail a compromise between the accuracy, flexibility, and reliability of the operations. With a combination of various detection mechanisms, hybrid IDS is able to operate more effectively in a more complex threat environment and minimise cases of false alarms, an aspect that is vital in ensuring that the system retains trust and operational effectiveness in the financial system.

Figure 2. Taxonomy of AI-Based Intrusion Detection Models

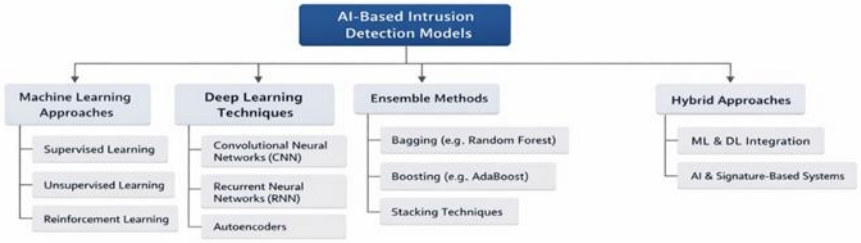

Figure 2: Taxonomy of AI-Based IDS Models

Figure 2 shows the taxonomy of AI based intrusion detection systems, starting with the historic use of signature-based models, to sophisticated machine learning, deep learning as well as a hybrid. The taxonomy emphasises the evolution of rule based systems to intelligent models, which are learning, adaptive and get better with time. This categorization offers a conceptual basis on how various methods of IDS can be linked with each other and how they would lead to the larger goal of improving cyber security in the banking settings.

III. IDS Evaluation And Deployment Lifecycle

Implementation of intrusion detection system (IDS) based on AI is not a one-step undertaking, but a lifecycle that extends to data acquisition, model development, reviewing, and adjusting to real-time data. In the real world of banking, this lifecycle should also be in accordance to the operational conditions, regulatory controls and the changing cyber threats. The performance of IDS, according to the literature, is not limited to the detection model but also the training, evaluation and maintenance of the model with a time period (Lippmann et al., 2000; Sommer and Paxson, 2010).

Data Collection and Preprocessing

The initial phase in the IDS lifecycle is the gathering of appropriate information in the network traffic, system logs, and transactional records. This information is usually big, unstructured, and sensitive in banking settings and has to be handled carefully to protect privacy and compliance with regulation guidelines (CBN, 2022; ENISA, 2023). Preprocessing of data is an important procedure that converts raw data into a form that is easy to analyse. This involves cleaning, normalisation, selection of features, and transformation of data into meaning. Many research studies have utilised datasets like UNSW-NB15 and other bench mark intrusion datasets to provide the simulated environment of the real-world attack and to test the performance of the models (Moustafa and Slay, 2015; Sharafaldin et al., 2018). The quality of the input information is increased with the correct preprocessing, which positively influences the AI model performance by backgrounding noise and making sure that the features are consistent with each other.

Model Training, Tests, and Evaluation

After the data preparation, the second step is to run the IDS model to train it to differentiate normal and malicious activities. That is normally formulated as a supervised learning problem, in which the model is trained on labelled data to extract the patterns linked to cyber threats (Buczak and Guven, 2016; Sarker, 2021). The data is split into the training and the testing data so that the model is tested on the unseen data. This is useful in determining its extrapolation skills out of the training situation. Practically, machine learning classifiers and other

deep neural networks are trained with massive data to reproduce more intricate trends in network behaviour (LeCun et al., 2015; Shone et al., 2018).

Evaluation is a very vital element of this phase, as it evaluates the extent to which the model is successful in attack detecting. Performance is usually assessed using standard evaluation measures including accuracy, precision, recall, and F1-score, especially in data that are imbalanced with the attack instances few relative to normal traffic (Davis and Goadrich, 2006; Powers, 2011). Nevertheless, one indicator is not effective when it comes to cybersecurity. Consequently, a joint metric is normally employed so as to make a more balanced evaluation of the capability of the detection and false-alarming rates.

Continuous monitoring and Feedback Loops

Contrary to the traditional systems, AI-based IDS is forced to work in dynamic conditions where the attack pattern is constantly developing. Subsequently, ongoing supervising and the monitoring and feedback systems are key ingredients of deployment lifecycle. Following deployment, the system processes real time information and compares its predictions with the real outcomes. This feedback is then made to improve and update the model over the time period so that adaptive learning can be performed. This is done especially when it comes to the identification of emergent threats and keeping systems relevant within the evolving contexts (Sommer & Paxson, 2010). Constant monitoring can also contribute to the compliance of the organisation and regulation, as the system must be in accordance with the developed policy on cybersecurity and the framework of risk control (ISACA, 2022; CBN, 2022). By doing so, the IDS will not only be used as a detector, but also as a component of a larger ecosystem of governance and risk management.

Mathematical Formation of IDS Performance

At this life cycle, there would be a need to critically assess the performance of the IDS quantitatively. The statement of the intrusion detection problem can be stated as a classification problem on a data set of observed behaviours:

Let the dataset be expressed as:

$$D = (x_i, y_i)_i = 1n$$

where $x_i \in Rd$ represents the feature vector extracted from network or system activity, and $y_i \in \{0,1\}$ denotes the corresponding class label (normal or attack).

The IDS model learns a mapping function:

$$\hat{y} = f(x; \theta)$$

where f represents the detection model and θ denotes the learned parameters.

To evaluate model effectiveness, several performance metrics are considered:

- Accuracy measures overall correctness
- Precision evaluates the proportion of correctly identified attacks
- Recall captures the ability to detect actual intrusions
- F1-score balances precision and recall
- False Positive Rate (FPR) reflects the proportion of normal traffic incorrectly classified as attacks

Such measures are critical in banking settings, where false attacks as well as false positives can cost them a great deal in terms of their operations and finances. The quantitative measures used here give an organised mechanism of comparing the varied models of IDS and the evaluation is based upon quantifiable performance, as opposed to subjective judgement. It is on this that the comparative analysis will be done in the subsequent portions of this study. To give a stringent and repeatable analysis of AI-founded intrusion detection systems, the detection challenge may be presented as a manned classification problem. This formulation offers a quantitative basis of model behaviour analysis and allows the comparison of various approaches in the same way.

Let the observed dataset be defined as:

$$D = (x_i, y_i)_i = 1nD$$

where $xi \in Rd$ represents the feature vector extracted from network traffic or system activity, and $y\_i \in \{0,1\}y\_i \in \{0,1\}y\_i \in \{0,1\}$ denotes the corresponding class label, with 0 indicating normal behavior and 1 representing malicious activity.

The IDS operates as a parametric decision function:

$$y = f(x; \theta)$$

where $f(\cdot)$ is the learned classification model (e.g., machine learning or deep learning), and θ represents the optimized parameters obtained during training. This function maps input features to predicted labels, forming the core of the detection mechanism.

In practice, the performance of the model is evaluated by comparing predicted outputs with ground truth labels using a confusion matrix framework, consisting of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). Based on this framework, several performance metrics are defined.

Accuracy measures the overall correctness of the model and is expressed as:

$$Accuracy = TP + TNTP + TN + FP + FN$$

However, in intrusion detection scenarios where class imbalance is common, accuracy alone is insufficient. Therefore, precision and recall are also considered to provide a more balanced evaluation.

Precision quantifies the proportion of correctly identified attacks among all detected attacks:

$$\text{Precision} = \frac{TP}{TP + FP}$$

Recall (also referred to as detection rate) measures the system's ability to identify actual attacks:

$$\text{Recall} = \frac{TP}{TP + FN}$$

To combine both measures into a single indicator, the F1-score is used as their harmonic mean:

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Additionally, the False Positive Rate (FPR) is critical in operational environments such as banking, as it reflects the likelihood of incorrectly classifying legitimate activity as malicious:

$$\text{FPR} = \frac{FP}{FP + TN}$$

The high FPR may result in undesirable alerts, malfunctioning of the system, and higher operational expenses, so it is a central factor in the real application.

In addition to evaluating model discrimination through threshold analysis, the Receiver Operating Characteristic (ROC) curve is commonly used to assess model discrimination. The ROC curve is a relationship between the True Positive Rate (TPR) and False Positive rate (FPR) at the various thresholds of classification:

$$\text{ROC} = \left(\frac{TP}{TP + FN}, \frac{FP}{FP + TN}\right)$$

The space below the ROC curve (AUC), unseen here distinctly, gives an overall performance of a model at all the thresholds.

Hence, this line of mathematical modelling provides a quantitative and systematic system of IDS performance analysis. It makes sure that the model evaluation is founded not just on the empirical observation, but also has the basis of formal statistical measures, thus contributing to the rigorous comparison, reproducibility, and practical applicability in banking cybersecurity setting.

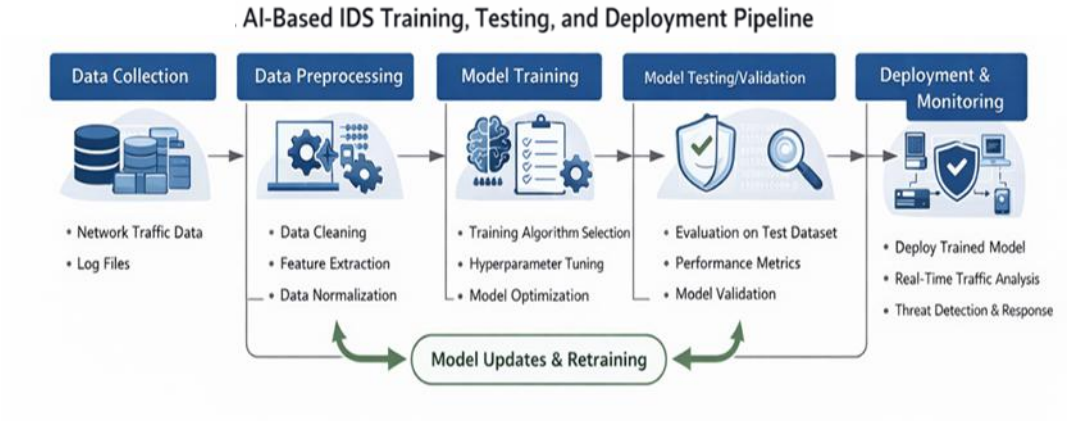

Figure 3: IDS Evaluation and Deployment Pipeline

IV. Adversarial Machine Learning Threats

The recent AI-driven intrusion detection systems (IDS) are being actively used in adversarial settings where attackers are actively trying to both alter inputs or take advantage of the weaknesses of the models. In contrast to other threats, the adversarial machine learning attacks are aimed at the learning process, thus eroding the credibility and integrity of detection systems (Biggio and Roli, 2018; Goodfellow et al, 2015).

Evasion and Poisoning Attacks

Attacks on IDS in an adversarial manner may be broadly divided into evasion and poisoning attacks.

Evasion: The evasion attacks happen at the inference stage when an attacker makes a slight alteration to malicious inputs to be wrongly labelled as harmless. These distortions can be designed in an intelligent, manners such that they do not exceed a very small range of distortion and are hence challenging to detect.

Poisoning attacks on the other hand address the issue of training by introducing malicious samples to the training dataset. This affects the learning process and leads to the bias or corrupted model which can act in Iraq ways when applied in actual implementations. These two types of attacks underscore the weakness of the data-based IDS models especially within a banking system where assailants keep upgrading their technique.

Inference of Models and Privacy Attacks

In addition to direct manipulation, opponents can also strive to extract sensitive information through deployed IDS models. These are membership inference attacks (where an attacker learns whether a specific data point was included in the training), and model inversion attacks (intended to learn an input data given output of a model).

In the financial systems, such attacks are of serious concern because they can result in leaks of the sensitive transaction pattern or user behaviour, thereby contravening privacy and regulatory mandates. This confirms the necessity of a model design that is well-crafted and deployment strategies that are safe.

GAN-Based Attack Generation

The most recent progress on generative models has allowed an attacker to synthesise the most realistic adversarial samples. Generation Generative Adversarial Networks (GANs) can produce fake attack traffic that is far more reminiscent of regular traffic, and thus much harder to detect. These types of attacks relying on GAN enhance the flexibility of attackers due to the possibility to generate data that avoids conventional detection constraints. This means that the IDS models should be developed to defend these dynamic and data-driven threats by making them robust and more capable of generalisation.

Adversarial Risk Function

The formalisation of the effect of adversarial manipulation on the performance of the IDS is obtainable by modelling the system in an adversarial risk framework. Define the learned IDS model as $f(x;\theta)$ in which is the model-parameters, and L () is the loss-function which quantifies the error in prediction. The adversarial risk is the loss that is expected to be incurred in the case of worst-case perturbations:

$$R_{adv} = E_{x\sim D}\left[\max_{||\delta||<\epsilon} L(f(x+\delta;\theta),y)\right]$$

Here:

- δ represents the adversarial perturbation applied to the input
- ϵ defines the perturbation constraint, ensuring the attack remains bounded
- D denotes the data distribution

The constraint:

$$||\ \delta\ ||< \epsilon \big||\delta|\big| < \epsilon\ ||\ \delta\ ||< \epsilon$$

ensures that the adversarial modification remains imperceptible or realistic, thereby maintaining the attack's effectiveness in practice.

From an optimization perspective, the IDS can be trained using a robust objective that minimizes the worst-case risk:

$$Minf_0\theta maxf_0\ ||\ \delta\ ||< \epsilon L(f(x+\delta;\theta),y) \min_{\theta} \max_{||\delta||<\epsilon} L(f(x+\delta;\theta),y)\ \theta min\ ||\ \delta\ ||< \epsilon maxL(f(x+\delta;\theta),y)$$

This minmax game model represents a game theoretic adversarial interaction between the game defender (IDS model) and the adversary. The goal of the defender is minimizing classification error, whereas the opposition tries to decide in the best way possible using optimal perturbations.

This system offers a strict base of assessing robustness, that is to see that the IDS is truthful during common circumstances and is likewise robust to adversarial disturbance.

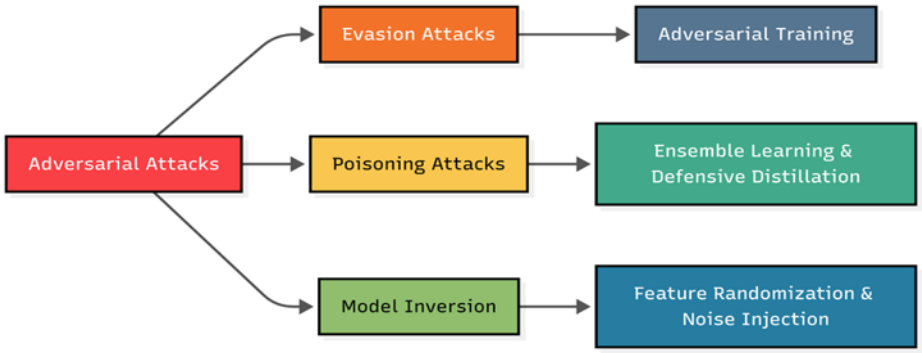

Figure 4: Adversarial Attack–Defense Interaction Model

V. Real-Time And Scalable AI-IDS Architecture

Today, the application of AI intrusion detection systems (IDS) in the banking environment necessitation must not only have an accurate architecture but also the ability to comply with the high-volume and continuously streaming network traffic. This requires a scaled design to achieve both responsiveness and reliability of the systems and detection efficiency.

Architecture Requirement

An IDS architecture should have a real-time architecture capable of meeting several functional and non-functional requirements such as:

a. Low latency detection: Threats are identified instantly without incurring much time.
b. High throughput: Capacity to cope with huge packets of network traffic.
c. Fault tolerance: Operation on partial system failure.
d. Scalability: Ability to support growing load of data in performance without decline.
e. Security and compliance: Compliance with the audit requirements and regulatory standards.
These conditions guarantee that the IDS is able to work in dynamic and risky conditions like banking systems.

Hybrid Model Integration

In order to increase detection capability, recent IDS architectures typically use a hybrid configuration, integrating more than one method of detection, e.g., machine learning, deep learning, and rule-based systems. The integration can be conceptually expressed as:

$$fhybrid(x) = \alpha fML(x) + \beta fDL(x) + \gamma frule(x) f_{hybrid}(x) = \alpha f_{ML}(x) + \beta f_{DL}(x) + \gamma f_{rule}(x) fhybrid(x) = \alpha fML(x) + \beta fDL(x) + \gamma frule(x)$$

where:

- fML represents machine learning-based detection
- fDL represents deep learning-based detection
- $frule$ represents rule-based or signature detection
- α, β, γ are weighting parameters such that $\alpha + \beta + \gamma = 1$

This hybridization allows the system to leverage the strengths of each component while mitigating their individual limitations, thereby improving overall detection robustness and adaptability.

Automated Response and Compliance Logging

In addition to the detection capabilities, a real-world IDS needs to have automated response capabilities as well as high-quality logs to audit and provide compliance.

Automated responses may include:

- Blocking malicious IP addresses
- Isolating compromised systems
- Triggering alerts to security teams

Compliance logging is the registration of all the discovered incident and system actions using the provisions of regulatory frameworks. This is especially critical in the banking setting, where audit trails are needed in terms of accountability and forensic analysis.

Real-Time Processing and Scalability Model

To formally evaluate system performance under real-time conditions, the IDS can be modeled using a queueing framework.

Let:
- λ denote the **incoming traffic rate** (requests per unit time)
- μ denote the **processing rate** (requests processed per unit time)

For the system to remain stable, the processing capacity must exceed the incoming traffic:

$$\rho = \lambda\mu < 1\ \rho = \frac{\lambda}{\mu} < 1\ \rho = \mu\lambda < 1$$

where ρ represents the system utilization factor. If $\rho \geq 1$, the system becomes unstable, leading to unbounded queue growth and eventual performance degradation.

The **average system latency** (waiting and processing time) can be expressed as:

$$T = 1\mu - \lambda T = \frac{1}{\mu - \lambda} T = \mu - \lambda 1$$

This relationship shows that as the incoming traffic rate λ approaches the processing capacity μ the latency increases rapidly, highlighting the importance of maintaining sufficient processing headroom.

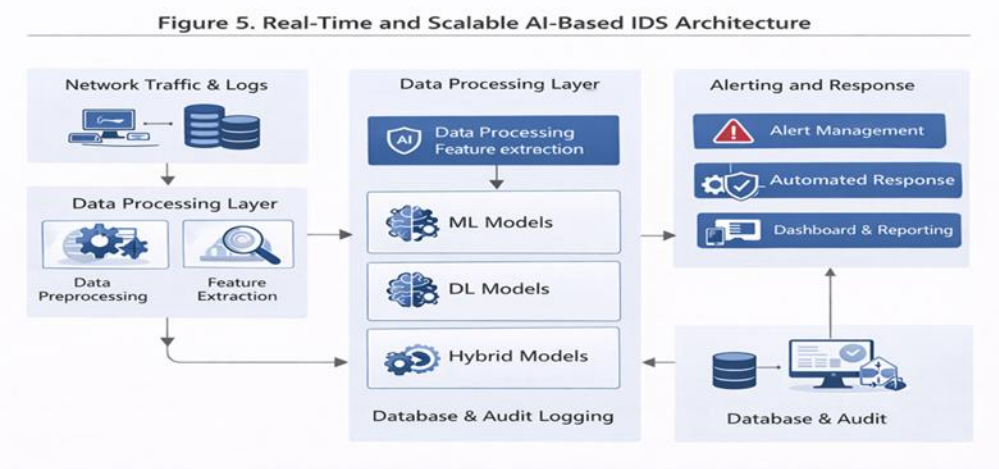

Figure 5: Real-Time and Scalable AI-Based IDS Architecture

The architecture shows the process of data ingestion moving on to preprocessing and then onto hybrid detection, automated response, and compliance logging by having a scalable processing layer to support high throughput network environments.

VI. Organisational/ Regulatory Constraints

Intrusion detection systems (IDS) based on AI do not only depend on the technical ability, but also the organisational willingness and regulatory conformity. These factors play a critical role in banking environments in terms of adoption, performance, and sustainability.

Constrains of Organisational Readiness

Organisation preparedness represents the level of how a particular institution is willing to embrace and actively use AI-based cybersecurity solutions. It is dependent on leadership support, recruitment of qualified people, sufficiency of infrastructure, and organisational culture.

To be able to measure this quantitatively, the Organisational Readiness Index (ORI) is propagated as:

$$ORI = w1L + w2S + w3I + w4CORI = w_1L + w_2S + w_3I + w_4CORI = w1L + w2S + w3I + w4C$$

where:
- L = Leadership Support
- S = Skill Availability
- I = Infrastructure Readiness
- C = Organizational Culture
- W_i = corresponding weights

Table 6.1 Organizational Readiness Index (ORI)

Factor	Weight	Score	Weighted Value
Leadership Support (L)	0.25	4.2	1.05
Skill Availability (S)	0.25	3.8	0.95
Infrastructure (I)	0.25	4.0	1.00
Culture (C)	0.25	3.5	0.87
ORI Score			**3.87 / 5**

Source: Prepared by the authors (2026).

The ORI score of 3.87 implies that the readiness is moderate-high, but skills gaps and cultural factors in the organisations are likely to impede the maximum level of adoption. This underscores the need to embrace capacity building and strategic human and technological resources investment.

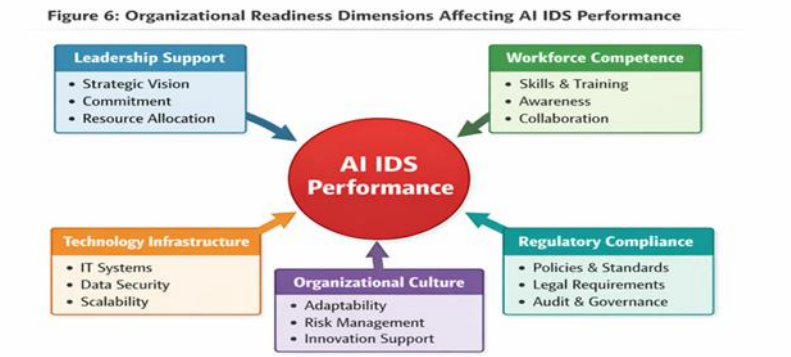

Figure 6: Organizational Readiness Constraints Affecting AI-IDS Effectiveness
Source: *Author's conceptual framework, informed by socio-technical systems theory and organizational readiness models for AI and cybersecurity adoption in the banking sector.*

In theory, Figure 6 indicated that the key organisational readiness variables leading to the influence of the operational and long-term sustainability of the AI-based Intrusion Detection Systems (IDS) in the banking setting included the leadership support and workforce skills, organisational culture, and technological infrastructure. All these are what determine how far the financial institutions can integrate the highly advanced cyberspace security technologies into the current systems of operation and governance.

Regulatory Alignment and Compliance

To deploy AI-IDS, it is vital to adhere to the regulatory frameworks of cybersecurity, data handling, and financial activities. They involve good governance policies internally and external policies under national and international bodies and levels. Regulatory alignment is used to guarantee that data privacy is enforced, audit trail is enforced, system actions are transparent and accountable and standardised risk management practises are adhered to. Failure to adhere to these requirements may lead to lawsuits, financial risks on operations and a damaged reputation. Thus, the key to building trustworthy AI in the banking systems is compliance.

VII. Hybrid Socio-Technical Adoption Framework Of AI-Cybersecurity

The uptake of AI-based cybersecurity tools is well construed using a socio-technical framework, which combines aspects of technology, organisations, and regulations.

Technical Layer

Technical layer takes up the AI models and detection algorithms and the system architecture. It aims at maintaining model accuracy, robustness, scalability, and real time processing and also incorporating hybrid detection mechanisms. The layer is important in successful detection and response to cyber threats.

Organizational Layer

In this layer, there are internal preparedness elements including leadership support, competency of the staff, infrastructure, and institutional culture. The achieved success of AI-IDS solutions and their long-term maintenance are significantly determined by organisational readiness mentioned by the ORI model.

Regulatory Layer

The regulatory layer permits the adherence to legal, ethical, and institutional regulations, including such issues as data protection, compliance auditing, and explainability and accountability of systems. It is a layer that offers the required external validation particularly in sensitive areas such as banking.

Dynamic Feedback Loops

The main characteristic of this structure is the feedback loops (between the technical, organisational, and regulatory layers) that are dynamic. Technical performance works in the decision making of the organisation whereas organisation readiness has implications regarding the deployment of technology. Regulation influences the formation of the system design and its functionality, making sure the system is continuously changing to meet the new threats, the changes in the organisation, and the modifications in the regulations.

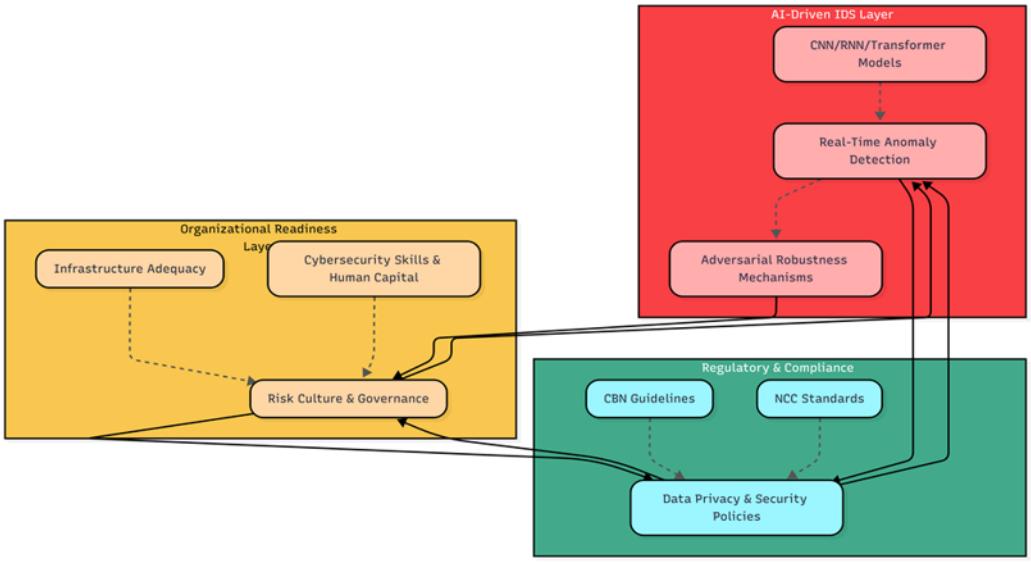

Figure 7: Hybrid Socio-Technical AI-Cybersecurity Adoption Framework

VIII. Results And Discussion

This category provides a quantitative assessment of the suggested Hybrid AI-IDS with the emphasis made on resource detection, adversarial unstoppability, organisational preparedness, and scaling of the system. The findings provide an indication of the benefits of relying on combining technical, organisational, and regulatory factors into practise.

Performance Analysis

The comparison of the detection capacity of the Hybrid AI-IDS is done with the traditional signature-based, ML-based and DL-based IDS models. Such metrics of performance are accuracy, precision, recall, f1-score and false positive rate (FPR).

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN},$$

$$\text{Precision} = \frac{TP}{TP + FP},$$

$$\text{Recall} = \frac{TP}{TP + FN},$$

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}},$$

$$\text{FPR} = \frac{FP}{FP + TN}$$

Table 8.1: IDS Performance Comparison

Model Type	Accuracy (%)	Precision	Recall	F1-Score	FPR
Signature-Based IDS	78.5	0.74	0.70	0.72	0.18
ML-Based IDS	88.2	0.86	0.84	0.85	0.10
DL-Based IDS	92.6	0.91	0.90	0.90	0.07
Hybrid AI-IDS (Proposed)	**96.3**	**0.95**	**0.94**	**0.94**	**0.04**

Source: Prepared by the authors (2026).

The proposed hybrid model works better than single-use AI and traditional systems and, therefore, signifies a good level of detection.

Adversarial Resilience

The adversarial attack on AI-IDS is measured with regards to expected loss under perturbation, according to the min-max optimization:

$$R_{adv} = E_{x\sim D}[L(f(x+\delta), y)], \quad ||\delta|| < \min_{\theta} \max_{||\delta||<\epsilon} L(f(x+\delta;\theta), y)$$

Table 8.2: Adversarial Robustness Evaluation

Model	Clean Accuracy	Under Attack	Robustness Loss (%)
ML IDS	88.2	71.4	16.8
DL IDS	92.6	78.9	13.7
Proposed Hybrid	96.3	89.5	6.8

Source: Prepared by the authors (2026).

The Hybrid AI-IDS shows performance degradation of only 6.8% as opposed to 13-17% with the comprehensive models, which poses great evidence of successful adversarial defence.

Organizational Impact

Organisational preparedness has a direct influence on system adoption and operational performance. The Organisational Readiness Index (ORI = 3.87/5) shows that the report has a moderate to high level of readiness, with its leadership support and infrastructure being identified as its strengths, and skill gaps are a weakness. This supports the necessity of considering both human and organisational variables and solutions on the technical level.

Scalability Analysis

Formal analysis of system utilisation and latency is done using the real-time processing and queueing model (Section 5.4):

$$\rho = \frac{\lambda}{\mu} < 1, \quad T = \frac{1}{\mu - \lambda}$$

Where $\lambda =$ incoming traffic rate and μ $=$ processing rate.

High-throughput scenario simulation reveals that the Hybrid AI-IDS is stable ($\rho<1$), and the latency rises rapidly only when traffic nearly reaches the processing capacity. This ascertains scalability in real-time banking.

Key Findings

The present paper concludes that various aspectual insights are of fundamental importance in the implementation of hybrid AI-driven Intrusion Detection System (IDS) in cybersecurity in the banking industry. To begin with, hybrid models proved to be more efficient than single machine learning (ML) and deep learning (DL) IDS systems based on multiple performance metrics, namely, accuracy, precision, recall and F1-score. Second, the case of organisational readiness (ORI) turned out to make a measurable and consequential effect on the success of deployment and the work efficiency of AI-based IDS solutions. Thirdly, the adversarial robustness became the critical factor, where the hybrid model showed much higher resistance against the performance reduction upon attacks. Finally, it was ascertained that scalability could be done by means of distributed or parallel processing that will enable the system to act with stability despite having a lot of network traffic.

IX. Conclusion And Research Implications

In this research, a Hybrid Socio-Technical AI-IDS framework of banking cybersecurity has been engineered that incorporates all the layers of technical detection, organisational preparedness, and regulatory compliance. Key contributions include:

1. Improved Performance in Detection: The hybrid model demonstrated accuracy of 96.3%, F1-score of 94% and false positive rate of 4%, which is better than the standalone ML/DL and traditional IDS.

2. Adversarial Robustness: The proposed system does not deteriorate in performance by as much as 6.8 and 6.8 in the case of attacks which ascertains the effectiveness of adversarial-conscious modelling (Section 8.2).

3. Organisational Readiness: AI-IDS adoption is influenced by leadership, skills, and infrastructure as seen in the case of Quantitative ORI evaluation (3.87/5).

4. Scalability: Queueing and latency models ensure successful performance ($\rho<1$) with high traffic, which supports the reality of an application at the banking setting.

5. Socio-Technical Integration: Integrating technical, organisational, and regulatory layers will provide complete risk management and ease implementation in the complex banking situation.

Research Implications

Theoretical Implications

The present study has multiple valuable theoretical contributions to the cybersecurity research:

1. The importance of hybrid socio-technical modelling to augment the perception of cybersecurity systems is greatly emphasised, especially within a complex and real-world system such as the banking context, in particular.
2. The paper presents a quantitative paradigm, which connects the performance indicators of Intrusion Detection System (IDS) directly to organisational readiness variables, which is a connexion between the technical and organisational state of cybersecurity.
3. With the incorporation of min-max robustness optimization together with the adversarial machine learning, the currently developed work is an extension of the existing literature on the subject of adversarial resilience in cybersecurity especially in its application to the real-world banking practises.

Practical Implications

In terms of practical use, the results can be applied to give the following actionable recommendations to both the researchers and practitioners in the industry:

1. The suggested hybrid AI-IDS model can help banks to improve the threat detection without disrupting the performance of the system in terms of losing system latency and adhering to the regulatory requirements.
2. The index of organisational readiness (ORI) created within the scope of the present research presents important indicators that can assist banks with focusing their investments on the variety of activities, such as training and development of employees, building of infrastructure, and involvement in leadership to enhance the successful adoption of AI-IDS.
3. With the integration of AI-IDS solutions and regulatory frameworks, the banks are able to remain in line with audit and regulatory guidelines, which will reduce the risks that would be involved in case of operational and reputational damages due to cybersecurity breaches.

Recommendations

Technical Recommendations

The study provides a number of technical proposals to improve the performance of cybersecurity among the banking settings:

1. Use Hybrid IDS Architectures: Banks are advised to use the hybrid Intrusion Detection Systems (IDS) which combine the use of the Machine Learning (ML), Deep Learning (DL), and rule-based in order to ensure the detection rate is maximum and the overall system resilience is enhanced.

2. Train Adversarial Defence Systems: Adversarial defence mechanisms, including perturbation-based training and robust training are necessary to mount the presence of a possible attack and engender system stability.

3. Use Dynamic Scaling: The idea of banks taking advantage of distributed or cloud-based processing to scale the system to sustain stability as well as guarantee performance ($\rho < 1$) in the high load situation is essential as it allows the real time deployments in the critical banking made Irene.

Organizational Recommendations

In order to implement AI-IDS successfully, the following organisational strategies may be recommended:

1. Upgrade AI Cybersecurity Abilities: Ongoing training opportunities will be implemented in place of boosting the abilities and experience of cybersecurity pros in the AI-assisted threat finding field.

2. Foster Leadership Support: It is essential to involve leadership that promotes an organisational culture of security awareness that is required to enable successful implementation and well-being of AI-IDS technologies.

3. Conduct Regular Organizational Readiness Assessments: Periodic organisation readiness assessment should also be done to ensure that there are no gaps in the infrastructure and to streamline processes to achieve successful implementation of AI-IDS.

Regulatory Recommendations

The regulatory measures that should be taken to ensure conformity and reduce risks are:

1. Cover AI-IDS with Regulations: Banks correspond to the regulations of AI-IDS implementation: Banks are to make sure that AI-IDS implementation is consistent with the national and international regulations in cybersecurity like the CBN IT Governance Framework and the Basel Committee practises.

2. Keep Detailed Compliance Records: Audit logs and other compliance records are supposed to be regularly updated to show their compliance with the regulatory standards and to simplify the process of the audit.

3. Consult Regulators: Banks ought to consult regulatory agencies to ensure that AI-IDs solutions are adjusted to comply with new regulatory standards and standards in the industry.

Future Research Directions

In order to further facilitate the contribution towards the field of banking cybersecurity, the following research topics are suggested:

1. Evaluate Ideas of Federated Learning: Future research may focus on approaches of federated learning to support cross-institutional exchange of threat intelligence, to improve the overall security against new cyber threats.

2. Simulate Real-Time Theoretical attacks: Due to the investigation of the real-time adversarial attack simulations to test the continuous robustness, it is possible to increase the flexibility of AI-IDS systems.

3. Research Organisational Readiness in Other Sectors: More investigation on organisational readiness and AI performance in non-banking sectors should be conducted as a way to generalise the applicability of AI-IDS technologies.

References

[1]. Axelsson, S. (2000). Intrusion Detection Systems: A Survey And Taxonomy. Technical Report, Chalmers University Of Technology.
[2]. Basel Committee On Banking Supervision. (2021). Principles For Operational Resilience. BIS.
[3]. Bengio, Y., Courville, A., & Vincent, P. (2013). Representation Learning: A Review And New Perspectives. IEEE Transactions On Pattern Analysis And Machine Intelligence, 35(8), 1798–1828.
[4]. Biggio, B., & Roli, F. (2018). Wild Patterns: Ten Years After The Rise Of Adversarial Machine Learning. Pattern Recognition, 84, 317–331. Https://Doi.Org/10.1016/J.Patcog.2018.07.023
[5]. Biggio, B., Corona, I., Maiorca, D., Nelson, B., Šrndić, N., Laskov, P., ... & Roli, F. (2013, September). Evasion Attacks Against Machine Learning At Test Time. In Joint European Conference On Machine Learning And Knowledge Discovery In Databases (Pp. 387–402). Springer, Berlin, Heidelberg.
[6]. Biggio, B., Nelson, B., & Laskov, P. (2012). Poisoning Attacks Against Support Vector Machines. Arxiv Preprint Arxiv:1206.6389.
[7]. Buczak, A. L., & Guven, E. (2016). A Survey Of Data Mining And Machine Learning Methods For Cyber Security Intrusion Detection. IEEE Communications Surveys & Tutorials, 18(2), 1153–1176.
[8]. Central Bank Of Nigeria (CBN). (2022). CBN IT Governance Framework For Banks. Abuja, Nigeria.
[9]. Central Bank Of Nigeria (CBN). (2022). Risk-Based Cybersecurity Framework And Guidelines For Deposit Money Banks In Nigeria. CBN. Https://Www.Cbn.Gov.Ng
[10]. Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly Detection: A Survey. ACM Computing Surveys, 41(3), 1–58. Https://Doi.Org/10.1145/1541880.1541882
[11]. Davis, J., & Goadrich, M. (2006). The Relationship Between Precision-Recall And ROC Curves. Proceedings Of The 23rd International Conference On Machine Learning (ICML), 233–240. Https://Doi.Org/10.1145/1143844.1143874
[12]. Debar, H., Dacier, M., & Wespi, A. (1999). Towards A Taxonomy Of Intrusion-Detection Systems. Computer Networks, 31(8), 805–822.
[13]. Dietterich, T. G. (2000). Ensemble Methods In Machine Learning. International Workshop On Multiple Classifier Systems.
[14]. ENISA. (2023). AI Cybersecurity In Financial Services: Threats, Opportunities, And Guidelines. European Union Agency For Cybersecurity.
[15]. Goodfellow, I., Shlens, J., & Szegedy, C. (2015). Explaining And Harnessing Adversarial Examples. In International Conference On Learning Representations (ICLR). Https://Arxiv.Org/Abs/1412.6572
[16]. ISACA. (2022). COBIT 2019 Framework For Governance And Management Of Enterprise IT. ISACA.
[17]. Kairouz, P., Et Al. (2021). Advances And Open Problems In Federated Learning. Foundations And Trends In Machine Learning.
[18]. Lakshminarayanan, B., Pritzel, A., & Blundell, C. (2017). Simple And Scalable Predictive Uncertainty Estimation Using Deep Ensembles. Advances In Neural Information Processing Systems, 6402–6413.
[19]. Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. Nature, 521(7553), 436–444.
[20]. Lippmann, R. P., Fried, D. J., Graf, I., Haines, J. W., Kendall, K. R., Mcclung, D., … & Zissman, M. A. (2000, January). Evaluating Intrusion Detection Systems: The 1998 DARPA Off-Line Intrusion Detection Evaluation. In Proceedings DARPA Information Survivability Conference And Exposition. DISCEX'00 (Vol. 2, Pp. 12–26). IEEE. Https://Doi.Org/10.1109/DISCEX.2000.839540
[21]. Moustafa, N., & Slay, J. (2015). UNSW-NB15: A Comprehensive Data Set For Network Intrusion Detection Systems. 2015 Military Communications And Information Systems Conference (Milcis), 1–6. Https://Doi.Org/10.1109/Milcis.2015.7348942
[22]. Nigerian Communications Commission (NCC). (2022). Guidelines On Cybersecurity Practices For Telecommunication Providers. NCC. Https://Www.Ncc.Gov.Ng
[23]. Patcha, A., & Park, J. M. (2007). An Overview Of Anomaly Detection Techniques: Existing Solutions And Latest Technological Trends. Computer Networks, 51(12), 3448–3470.
[24]. Powers, D. M. W. (2011). Evaluation: From Precision, Recall And F-Measure To ROC, Informedness, Markedness, And Correlation. Journal Of Machine Learning Technologies, 2(1), 37–63.
[25]. Sarker, I. H. (2021). AI-Driven Cybersecurity: An Overview, Security Intelligence Modeling, And Research Directions. SN Computer Science, 2(3). Https://Doi.Org/10.1007/S42979-021-00592-X
[26]. Sarker, I. H., Kayes, A. S. M., Badsha, S., Alqahtani, H., Watters, P., & Ng, A. (2020). Cybersecurity Data Science: An Overview From Machine Learning Perspective. Journal Of Big Data, 7(1), 41.
[27]. Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018). Toward Generating A New Intrusion Detection Dataset And Intrusion Traffic Characterization. ICISSP 2018 – 4th International Conference On Information Systems Security And Privacy, 108–116. Https://Doi.Org/10.5220/0006604301080116
[28]. Sharma, R., Gupta, S., & Joshi, P. (2021). Cybersecurity Workforce Challenges In The Era Of AI And Cloud Computing. Journal Of Cybersecurity And Information Management, 9(2), 45–61.
[29]. Shone, N., Ngoc, T. N., Phai, V. D., & Shi, Q. (2018). A Deep Learning Approach To Network Intrusion Detection. IEEE Transactions On Emerging Topics In Computational Intelligence, 2(1), 41–50.
[30]. Sommer, R., & Paxson, V. (2010). Outside The Closed World: On Using Machine Learning For Network Intrusion Detection. IEEE Symposium On Security And Privacy.
[31]. Sommer, R., & Paxson, V. (2010). Outside The Closed World: On Using Machine Learning For Network Intrusion Detection. IEEE Symposium On Security And Privacy, 305–316. Https://Doi.Org/10.1109/SP.2010.25