OPEN ACCESS JOURNALS
Indiana Journal of Economics and Business Management
Abbreviated Tittle- Ind J Econ Bus Manag
ISSN *(Online)*- 2583-3758
Journal Homepage Link- https://indianapublications.com/journal/IJEBM
DOI: https://doi.org/10.5281/zenodo.21652787

INDIANA PUBLICATIONS
PRODUCTIVE AND QUALITY RESEARCH

Research Article

Volume-06|Issue04|2026

Reducing False Positives in AI-Based Intrusion Detection Systems Using Hybrid Artificial Intelligence for Oil and Gas Critical Infrastructure

Nonye Peter Awurum

Affiliations: Charisma University, USA.

Article History
Received: 07.06.2026
Accepted: 14.07.2026
Published: 27.07.2026

Citation
Awurum, N. P. (2026). Reducing False Positives in AI-Based Intrusion Detection Systems Using Hybrid Artificial Intelligence for Oil and Gas Critical Infrastructure. *Indiana Journal of Economics and Business Management, 6*(4), 104-128.

Abstract: The increasing frequency and sophistication of cyberattacks targeting critical infrastructure have accelerated the adoption of Artificial Intelligence (AI)-based Intrusion Detection Systems (IDSs) for real-time cyber threat detection. Although machine learning and deep learning techniques have significantly improved intrusion detection accuracy, high false-positive rates remain one of the most significant challenges affecting operational cybersecurity. Excessive false alerts overwhelm Security Operations Centres (SOCs), increase analyst workload, delay incident response, and may lead to alert fatigue, causing genuine cyber threats to be overlooked. This challenge is particularly critical within oil and gas environments, where cyber incidents can disrupt industrial operations, compromise safety, and result in substantial economic losses.

This study proposes a Hybrid Artificial Intelligence (Hybrid AI) approach designed to reduce false positives while maintaining high detection accuracy for intrusion detection systems deployed within oil and gas critical infrastructure. The proposed architecture combines traditional machine learning algorithms, deep learning models, confidence-based decision fusion, and Explainable Artificial Intelligence (XAI) into a unified detection framework. Machine learning algorithms provide efficient classification of structured attack patterns, while deep learning models capture complex spatial and temporal characteristics of network traffic. Decision fusion integrates predictions from multiple AI models using weighted confidence scoring to improve classification reliability and minimize erroneous alerts.

The proposed hybrid model is evaluated using publicly available benchmark cybersecurity datasets, including CICIDS2017, UNSW-NB15, NSL-KDD, and a hybrid Operational Technology (OT)/Information Technology (IT) dataset representative of industrial environments. Model performance is assessed using accuracy, precision, recall, F1-score, false positive rate (FPR), false negative rate (FNR), Matthews Correlation Coefficient (MCC), Area Under the Receiver Operating Characteristic Curve (ROC-AUC), and detection latency.

The expected findings demonstrate that integrating complementary AI models through intelligent decision fusion significantly reduces false-positive alerts while preserving high detection accuracy and operational efficiency. The study contributes to cybersecurity research by presenting a practical Hybrid AI architecture that improves intrusion detection reliability, enhances analyst confidence through explainable AI, and supports proactive cyber defence within critical infrastructure environments. The proposed approach provides valuable guidance for organizations seeking to improve the effectiveness of AI-enabled intrusion detection systems while reducing operational costs associated with excessive false alarms.

Keywords: Hybrid Artificial Intelligence, False Positive Reduction, Intrusion Detection System, Machine Learning, Deep Learning, Explainable Artificial Intelligence, Cybersecurity, Oil and Gas, Critical Infrastructure, Decision Fusion, Security Operations Centre

INTRODUCTION

Background

The rapid digital transformation of industrial organizations has significantly expanded the adoption of interconnected Information Technology (IT) and Operational Technology (OT) systems. In the oil and gas industry, technologies such as Supervisory Control and Data Acquisition (SCADA) systems, Distributed Control Systems (DCS), Industrial Internet of Things (IIoT) devices, cloud computing, and edge intelligence have improved operational efficiency, predictive maintenance, and remote asset management (Stouffer et al., 2023). However, these technological advancements have simultaneously increased the cybersecurity attack surface, exposing critical infrastructure to increasingly sophisticated cyber threats.

Recent cyber incidents involving industrial environments—including the Stuxnet malware, Triton attack, Industroyer malware, and Colonial Pipeline ransomware attack—demonstrate that attacks targeting industrial control systems can have severe operational, financial, environmental, and safety consequences (Cybersecurity and Infrastructure Security Agency [CISA], 2022).

Consequently, organizations have increasingly adopted Artificial Intelligence (AI)-based Intrusion Detection Systems (IDSs) capable of identifying

***Corresponding Author:** Nonye Peter Awurum

malicious network behaviour using machine learning and deep learning techniques.

Unlike traditional signature-based intrusion detection systems, AI-based IDSs automatically learn complex behavioural characteristics from network traffic and continuously improve detection performance as new attack data become available (Buczak & Guven, 2016). Machine learning algorithms such as Random Forest (RF), Support Vector Machine (SVM), and Extreme Gradient Boosting (XGBoost) have demonstrated strong classification performance for structured cybersecurity datasets, while deep learning models—including Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and Transformer architectures—have significantly improved anomaly detection through automatic feature extraction and sequential pattern recognition (LeCun et al., 2015; Vaswani et al., 2017).

Despite these advances, one of the most persistent challenges affecting AI-based intrusion detection systems is the high incidence of false-positive alerts. False positives occur when legitimate network activities are incorrectly classified as malicious. Excessive false alarms overwhelm cybersecurity analysts, increase operational costs, reduce trust in AI-generated alerts, and contribute to alert fatigue, ultimately decreasing the effectiveness of Security Operations Centres (SOCs) (Sommer & Paxson, 2010). In industrial environments where rapid incident response is essential for maintaining operational continuity and safety, minimizing false positives is as important as achieving high detection accuracy.

Problem Statement

Although machine learning and deep learning have significantly enhanced cyberattack detection capabilities, many AI-based intrusion detection systems continue to generate unacceptably high false-positive rates. Most existing studies prioritize improving classification accuracy while giving comparatively less attention to reducing unnecessary alerts generated during normal network operations.

High false-positive rates create several operational challenges. Security analysts must investigate numerous benign events incorrectly identified as cyberattacks, increasing workload and reducing the time available to investigate genuine threats. This phenomenon, commonly referred to as **alert fatigue**, may result in delayed responses, overlooked incidents, and inefficient utilization of cybersecurity resources. Within oil and gas environments, excessive false alerts may also interrupt industrial operations, reduce confidence in automated cybersecurity systems, and increase organizational risk.

Existing approaches frequently rely on single machine learning or deep learning algorithms, limiting their ability to balance detection sensitivity with classification specificity across diverse cyberattack scenarios. Furthermore, relatively few studies investigate how complementary AI techniques can be combined within hybrid architectures specifically designed to minimize false-positive rates while maintaining high detection performance. There is therefore a need for a Hybrid AI approach that integrates multiple analytical models, confidence-based decision fusion, and explainable AI to improve intrusion detection reliability in critical infrastructure environments.

Aim of the Study

The primary aim of this study is to develop and evaluate a Hybrid Artificial Intelligence framework that reduces false-positive alerts in AI-based intrusion detection systems while maintaining high detection accuracy for cyberattack detection in oil and gas critical infrastructure.

Research Objectives

The study is guided by the following objectives:

- To investigate the causes and operational impacts of false-positive alerts in AI-based intrusion detection systems.
- To design a Hybrid AI architecture integrating machine learning, deep learning, confidence-based decision fusion, and Explainable Artificial Intelligence (XAI).
- To evaluate the proposed hybrid model using benchmark cybersecurity datasets and hybrid IT/OT datasets representative of industrial environments.
- To compare the performance of the Hybrid AI model with conventional machine learning and deep learning approaches using standardized evaluation metrics.
- To assess the practical applicability of the proposed approach for improving cybersecurity operations within oil and gas critical infrastructure.

Research Questions

This study seeks to answer the following research questions:

- What are the primary factors contributing to false-positive alerts in AI-based intrusion detection systems?
- Can a Hybrid AI architecture significantly reduce false-positive rates while maintaining high detection accuracy?
- How does the proposed Hybrid AI model compare with individual machine learning and deep learning algorithms in terms of predictive performance and operational efficiency?
- What role does confidence-based decision fusion play in improving intrusion detection reliability?
- How can Explainable Artificial Intelligence improve analyst trust and support cybersecurity decision-making within industrial environments?

***Corresponding Author:** Nonye Peter Awurum

Research Contributions

This study makes several important contributions to cybersecurity research and industrial practice.

First, it proposes a novel Hybrid Artificial Intelligence architecture that combines machine learning, deep learning, confidence-based decision fusion, and Explainable Artificial Intelligence into a unified intrusion detection model specifically designed to reduce false-positive alerts.

Second, the study introduces a confidence-weighted decision fusion mechanism that aggregates predictions from multiple AI models to improve classification reliability and minimize erroneous alerts without sacrificing detection accuracy.

Third, the research incorporates Explainable Artificial Intelligence as an integral component of the proposed architecture, providing interpretable explanations that improve analyst confidence, facilitate incident investigation, and support regulatory compliance.

Fourth, the study provides a comprehensive empirical evaluation using benchmark cybersecurity datasets and hybrid IT/OT environments, offering practical evidence regarding the effectiveness of Hybrid AI for protecting oil and gas critical infrastructure.

Finally, the findings contribute to the development of more efficient Security Operations Centres by reducing alert fatigue, improving incident prioritization, and supporting proactive cyber defence strategies.

Organization of the Paper

The remainder of this paper is organized as follows. **Section 2** reviews the literature on intrusion detection systems, false-positive challenges, machine learning, deep learning, hybrid artificial intelligence, and explainable AI while identifying the research gap addressed by this study. **Section 3** presents the research methodology, datasets, experimental setup, and the proposed Hybrid AI architecture. **Section 4** describes the proposed false-positive reduction framework, including the confidence-based decision fusion mechanism and explainability components. **Section 5** presents the experimental results and evaluates the proposed model using multiple performance metrics. **Section 6** provides a comparative analysis against conventional machine learning and deep learning models and discusses the implications of the findings for industrial cybersecurity. Finally, **Section 7** concludes the paper by summarizing the key findings, highlighting the research contributions, providing recommendations, and identifying directions for future research.

LITERATURE REVIEW

Artificial Intelligence (AI) has fundamentally transformed modern cybersecurity by enabling intelligent intrusion detection systems capable of identifying both known and previously unseen cyber threats. Unlike traditional signature-based security solutions, AI-based intrusion detection systems (IDSs) continuously learn from historical and real-time network traffic, allowing them to recognize evolving attack behaviours with improved accuracy. Consequently, machine learning and deep learning have become integral components of cybersecurity strategies across finance, healthcare, manufacturing, transportation, and critical infrastructure sectors (Buczak & Guven, 2016).

Despite these advances, false-positive alerts remain one of the most persistent operational challenges associated with AI-driven intrusion detection. Although many studies report classification accuracies exceeding 95%, comparatively fewer investigations examine the practical implications of excessive false alarms within Security Operations Centres (SOCs). High false-positive rates reduce analyst productivity, increase operational costs, and undermine confidence in AI-assisted decision-making (Sommer & Paxson, 2010). In industrial environments such as oil and gas facilities, where cyber incidents may disrupt physical processes and compromise safety, minimizing false positives is equally as important as maximizing detection accuracy.

This section critically reviews the current literature on intrusion detection systems, false-positive challenges, machine learning and deep learning approaches, hybrid AI techniques, decision fusion mechanisms, and Explainable Artificial Intelligence (XAI). The review concludes by identifying the research gap addressed by the proposed Hybrid AI architecture.

Intrusion Detection Systems

Intrusion Detection Systems (IDSs) are among the most widely deployed cybersecurity technologies for identifying unauthorized activities within computer networks. Their primary objective is to detect malicious behaviour before significant damage occurs while supporting timely incident response. Traditionally, IDSs have been categorized into signature-based, anomaly-based, and hybrid detection systems, each possessing distinct advantages and limitations (Scarfone & Mell, 2007).

Signature-based IDSs detect attacks by comparing observed events with predefined attack signatures. These systems are highly effective against previously identified threats and typically produce relatively low false-positive rates. However, their dependence on continuously updated signature databases limits their ability to detect zero-day attacks and novel attack techniques (Liao et al., 2013). As cyber threats continue to evolve, reliance on signature matching alone has become increasingly insufficient.

***Corresponding Author:** Nonye Peter Awurum

Anomaly-based IDSs attempt to identify deviations from established patterns of legitimate network behaviour. These systems are particularly effective for detecting unknown attacks because they do not rely on predefined signatures. Nevertheless, anomaly detection methods frequently struggle to distinguish between legitimate behavioural changes and malicious activities, often producing significantly higher false-positive rates than signature-based approaches (Chandola et al., 2009). This trade-off between detection capability and false alarms remains one of the principal challenges in intrusion detection research.

Hybrid IDSs seek to combine the complementary strengths of signature-based and anomaly-based approaches. By integrating multiple detection strategies, hybrid systems attempt to improve detection accuracy while reducing false-positive alerts. Recent advances in AI have accelerated the development of intelligent hybrid IDSs that combine machine learning and deep learning algorithms to improve overall detection performance (Ahmad et al., 2021).

Although hybrid intrusion detection has demonstrated promising results, many existing implementations emphasize improving classification accuracy without explicitly optimizing false-positive reduction. Consequently, alert fatigue remains a significant operational concern, particularly within environments generating millions of security events each day.

False Positives in AI-Based Intrusion Detection

False positives occur when legitimate network traffic is incorrectly classified as malicious. Within operational cybersecurity environments, excessive false-positive alerts represent one of the most significant barriers to the practical deployment of AI-driven intrusion detection systems.

Several factors contribute to false-positive generation. Network traffic within modern enterprise and industrial environments is highly dynamic, with legitimate behavioural patterns continually changing because of software updates, user activities, cloud services, and operational processes. AI models trained on historical datasets may therefore incorrectly interpret legitimate behavioural variations as malicious anomalies (Sommer & Paxson, 2010).

Another important contributor is class imbalance. Public cybersecurity datasets often contain significantly more benign traffic than attack samples or vice versa, depending on the dataset. Such imbalance may bias machine learning models toward the majority class, reducing their ability to generalize effectively to unseen operational environments (Shone et al., 2018).

Feature engineering also influences false-positive generation. Irrelevant, redundant, or noisy features reduce classification performance by introducing ambiguity into the learning process. Likewise, inappropriate hyperparameter selection may increase model sensitivity, resulting in unnecessary alert generation.

From an operational perspective, false positives impose substantial economic and organizational costs. Security analysts must manually investigate numerous benign events, consuming valuable resources that could otherwise be directed toward genuine cyber threats. Excessive false alarms also contribute to alert fatigue, reducing analyst vigilance and increasing the likelihood that genuine attacks will be overlooked. Consequently, reducing false positives has become a strategic priority for modern Security Operations Centres (SOCs).

Machine Learning Approaches for Intrusion Detection

Machine learning has become one of the most widely adopted AI techniques for intrusion detection because of its ability to learn discriminative patterns from labelled cybersecurity datasets.

Random Forest remains one of the most successful machine learning algorithms for intrusion detection. By combining multiple decision trees through ensemble learning, Random Forest improves predictive accuracy while reducing overfitting. Numerous studies have demonstrated its robustness across benchmark datasets, making it a preferred baseline algorithm for cybersecurity research (Breiman, 2001).

Support Vector Machine (SVM) has also been extensively applied to intrusion detection because of its ability to identify optimal decision boundaries between malicious and legitimate traffic. Although SVM often achieves high classification accuracy, its computational complexity increases considerably with large-scale datasets, limiting its suitability for real-time industrial environments (Cortes & Vapnik, 1995).

Extreme Gradient Boosting (XGBoost) has emerged as another powerful supervised learning algorithm owing to its efficient gradient boosting strategy, regularization mechanisms, and scalability. XGBoost frequently outperforms traditional classifiers in structured cybersecurity datasets while maintaining relatively low computational overhead.

Despite their strengths, machine learning algorithms generally depend on carefully engineered features and may struggle to identify highly complex attack behaviours without extensive preprocessing.

Deep Learning Approaches for Intrusion Detection

Deep learning has significantly advanced intrusion detection by enabling automatic extraction of hierarchical feature representations directly from raw cybersecurity data. Unlike conventional machine

learning algorithms, deep neural networks require minimal manual feature engineering and can model highly complex nonlinear relationships.

Convolutional Neural Networks (CNNs) have demonstrated excellent performance in cybersecurity because of their ability to identify local spatial patterns within network traffic. Originally developed for image recognition, CNNs have been successfully adapted for intrusion detection through one-dimensional convolutional architectures (LeCun et al., 2015).

Long Short-Term Memory (LSTM) networks extend recurrent neural networks by incorporating memory mechanisms capable of modelling long-term temporal dependencies. Their ability to analyse sequential network traffic makes them particularly effective for identifying multi-stage cyberattacks and advanced persistent threats (Hochreiter & Schmidhuber, 1997).

More recently, Transformer architectures have attracted considerable attention owing to their self-attention mechanisms, which capture long-range contextual relationships more efficiently than recurrent neural networks. Transformer-based intrusion detection models have demonstrated state-of-the-art performance across several benchmark cybersecurity datasets because they process network events simultaneously while preserving contextual dependencies (Vaswani et al., 2017).

Although deep learning substantially improves detection capability, these models generally require large computational resources, extensive training data, and careful hyperparameter optimization.

Hybrid Artificial Intelligence for False Positive Reduction

Hybrid Artificial Intelligence combines multiple computational intelligence techniques to exploit their complementary strengths while minimizing individual weaknesses. Rather than relying on a single classification model, hybrid AI integrates multiple algorithms through ensemble learning, stacking, voting strategies, or confidence-based decision fusion.

Recent studies indicate that hybrid AI architectures consistently outperform individual machine learning models because they improve robustness, reduce overfitting, and increase generalization capability (Ahmad et al., 2021). Ensemble methods aggregate predictions from several classifiers, thereby reducing the likelihood of isolated model errors influencing final decisions.

Stacking approaches introduce meta-learners that combine outputs from base classifiers to generate optimized predictions, whereas weighted voting techniques assign greater influence to models demonstrating superior confidence or historical performance. These approaches have shown considerable potential for reducing false-positive alerts while maintaining high detection sensitivity.

However, most existing hybrid AI studies prioritize maximizing overall classification accuracy rather than explicitly optimizing false-positive reduction. Furthermore, relatively few investigations incorporate confidence-based decision fusion and Explainable Artificial Intelligence within the same detection architecture, particularly for industrial cybersecurity applications.

Explainable Artificial Intelligence and Decision Fusion

One of the principal criticisms of deep learning-based intrusion detection systems concerns their lack of transparency. Many high-performing neural networks function as "black-box" models, making it difficult for cybersecurity analysts to understand the rationale underlying AI-generated decisions.

Explainable Artificial Intelligence (XAI) addresses this limitation by providing interpretable explanations of model predictions. Techniques such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations) enable analysts to identify the features influencing individual classifications, thereby improving trust, accountability, and regulatory compliance.

Decision fusion complements explainability by integrating predictions from multiple AI models before generating final classifications. Rather than accepting predictions from a single classifier, confidence-based decision fusion evaluates consensus among multiple analytical engines, reducing uncertainty and minimizing isolated classification errors. This collaborative decision-making process is particularly valuable for reducing false-positive alerts in environments where operational reliability is essential.

Research Gap

The literature demonstrates that artificial intelligence has substantially improved intrusion detection performance through machine learning, deep learning, and hybrid learning techniques. Nevertheless, several important research gaps remain.

First, the majority of published studies emphasize maximizing classification accuracy while giving comparatively limited attention to reducing false-positive rates, despite the considerable operational consequences associated with excessive false alarms.

Second, although hybrid AI models have demonstrated promising performance improvements, relatively few studies integrate machine learning, deep learning, confidence-based decision fusion, and

Explainable Artificial Intelligence into a unified architecture specifically designed to minimize false-positive alerts.

Third, most investigations evaluate AI models using conventional enterprise datasets without considering the unique operational characteristics of converged IT/OT environments commonly found in oil and gas critical infrastructure.

Finally, limited research has examined how reducing false positives influences operational efficiency within Security Operations Centres, particularly regarding analyst workload, alert fatigue, incident prioritization, and cyber resilience.

To address these limitations, this study proposes a Hybrid Artificial Intelligence architecture that combines multiple AI techniques, confidence-weighted decision fusion, and explainable AI to improve intrusion detection reliability while significantly reducing false-positive alerts in oil and gas critical infrastructure.

Summary

This section critically reviewed the literature on intrusion detection systems, false-positive challenges, machine learning, deep learning, hybrid artificial intelligence, explainable AI, and decision fusion. The review demonstrates that while significant progress has been made in improving cyberattack detection, reducing false-positive alerts remains an unresolved challenge with important operational implications. The identified research gaps provide the theoretical foundation for the Hybrid AI architecture proposed in this study. The following section presents the research methodology, experimental design, datasets, and development of the proposed false-positive reduction model.

RESEARCH METHODOLOGY

This study adopted an experimental research design to develop and evaluate a Hybrid Artificial Intelligence (Hybrid AI) model for reducing false-positive alerts in intrusion detection systems deployed within oil and gas critical infrastructure. The methodology was designed to ensure scientific rigor, reproducibility, and objective comparison between the proposed Hybrid AI model and conventional machine learning and deep learning approaches.

The research methodology comprised six sequential phases: dataset selection, data preprocessing, model development, hybrid decision fusion, experimental evaluation, and comparative performance analysis. Publicly available cybersecurity benchmark datasets and representative Operational Technology (OT)/Information Technology (IT) datasets were used to evaluate the effectiveness of the proposed model under realistic cyberattack scenarios. The experimental workflow was designed to determine whether combining multiple AI models through confidence-based decision fusion could significantly reduce false-positive rates while maintaining high detection accuracy and operational efficiency.

Research Design

A quantitative experimental research design was employed to assess the effectiveness of the proposed Hybrid AI model. Experimental research is appropriate because it enables objective evaluation of different AI algorithms under controlled conditions while facilitating statistical comparison of their predictive performance.

The study followed six major research phases:

1. Dataset selection and acquisition.
2. Data preprocessing and feature engineering.
3. Development of individual machine learning and deep learning models.
4. Construction of the Hybrid AI architecture using confidence-based decision fusion.
5. Experimental evaluation using standardized performance metrics.
6. Comparative analysis and statistical validation.

This structured approach ensured consistency across all experiments and minimized potential sources of experimental bias.

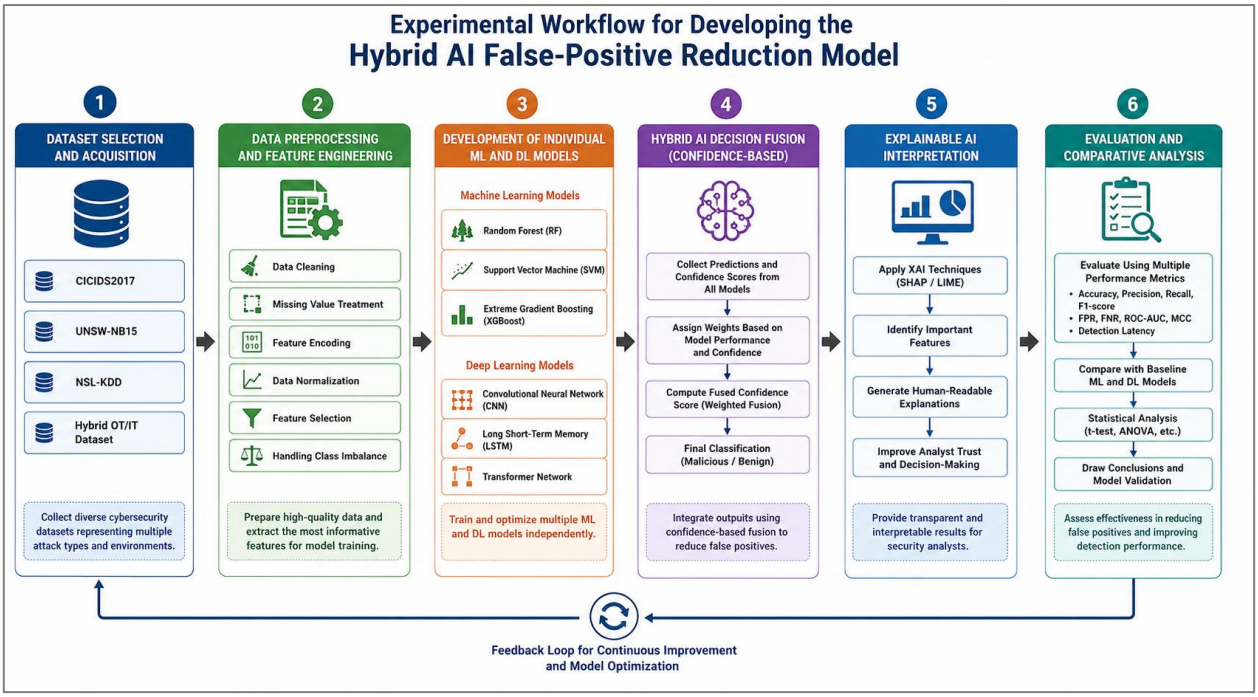

Figure 1: Experimental Workflow for Developing the Hybrid AI False-Positive Reduction Model
Source: Developed by the author (2026).

Note. Figure 1 illustrates the experimental methodology adopted in this study. The workflow begins with cybersecurity dataset acquisition and preprocessing, followed by the development of machine learning and deep learning models. These models are integrated through a confidence-based Hybrid AI decision fusion mechanism before undergoing performance evaluation and comparative statistical analysis. The iterative workflow enables continuous optimization of the proposed model while ensuring reproducibility and objective assessment of its effectiveness in reducing false-positive alerts.

Dataset Selection

Selecting representative datasets is essential for evaluating intrusion detection systems because model performance depends heavily on the diversity and quality of cybersecurity data. To improve the generalizability of the proposed Hybrid AI model, this study employed multiple publicly available benchmark datasets together with a hybrid OT/IT dataset representative of industrial environments.

The selected datasets include:

- CICIDS2017
- UNSW-NB15
- NSL-KDD
- Hybrid Operational Technology (OT)/Information Technology (IT) Dataset

These datasets collectively contain a wide variety of cyberattack categories, including denial-of-service attacks, distributed denial-of-service attacks, brute-force attacks, infiltration attempts, botnet activities, web attacks, reconnaissance, malware, and normal network traffic.

Table 1: Datasets Used in the Experimental Evaluation

Dataset	Description	Attack Categories	Application
CICIDS2017	Modern enterprise network traffic	DoS, DDoS, Botnet, Brute Force, Web Attacks	General IDS evaluation
UNSW-NB15	Synthetic enterprise dataset	Nine attack categories	Benchmark comparison
NSL-KDD	Improved KDD'99 dataset	Normal and malicious traffic	Historical comparison
Hybrid OT/IT Dataset	Industrial control system traffic	OT-specific attacks	Oil and gas validation

Using multiple datasets reduces dataset-specific bias and enhances the robustness of the experimental evaluation.

Data Preprocessing

Cybersecurity datasets typically contain noisy records, missing values, redundant attributes, and highly

imbalanced class distributions. Effective preprocessing is therefore essential for improving classification performance and reducing false-positive alerts.

The preprocessing pipeline comprised the following stages:

Data Cleaning
Duplicate records, incomplete observations, and corrupted entries were removed to improve data quality.

Missing Value Treatment
Missing numerical values were replaced using median imputation, while categorical variables were encoded using appropriate classification labels.

Feature Encoding
Categorical attributes were transformed into numerical representations using one-hot encoding and label encoding where appropriate.

Data Normalization
Continuous variables were normalized using Min–Max scaling to ensure comparable feature ranges across all datasets.

The normalization process is defined as:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Where:
- (X_{norm}) = Normalized feature value
- (X) = Original feature value
- (X_{min}) = Minimum value of the feature
- (X_{max}) = Maximum value of the feature

Feature Selection
Recursive Feature Elimination (RFE) and mutual information analysis were employed to identify the most informative cybersecurity features while reducing computational complexity.

Proposed Hybrid AI Architecture

The proposed Hybrid AI architecture integrates complementary machine learning and deep learning algorithms to improve classification reliability and reduce false-positive alerts.

Rather than relying on predictions from a single classifier, multiple AI models independently analyse incoming network traffic before submitting confidence scores to a centralized decision fusion engine.

The architecture comprises four principal analytical components:

Machine Learning Layer
- Random Forest (RF)
- Support Vector Machine (SVM)
- Extreme Gradient Boosting (XGBoost)

These algorithms provide efficient classification of structured cybersecurity data.

Deep Learning Layer
- Convolutional Neural Network (CNN)
- Long Short-Term Memory (LSTM)
- Transformer Network

These models automatically learn spatial, temporal, and contextual representations of malicious network behaviour.

Confidence-Based Decision Fusion
Outputs from all AI models are combined using weighted confidence scores rather than majority voting alone. Models demonstrating higher confidence and stronger historical performance receive greater influence during final classification.

The fused confidence score is calculated as:

$$C_f = \sum_{i=1}^{n} w_i C_i$$

Where:
- (C_f) = Final confidence score
- (w_i) = Weight assigned to the *i*th model
- (C_i) = Confidence score produced by the *i*th model
- (n) = Total number of AI models

This mechanism reduces isolated classification errors and improves decision reliability.

Explainable AI Layer
The final classification is interpreted using SHAP (SHapley Additive exPlanations), enabling cybersecurity analysts to understand the primary factors influencing each detection decision.

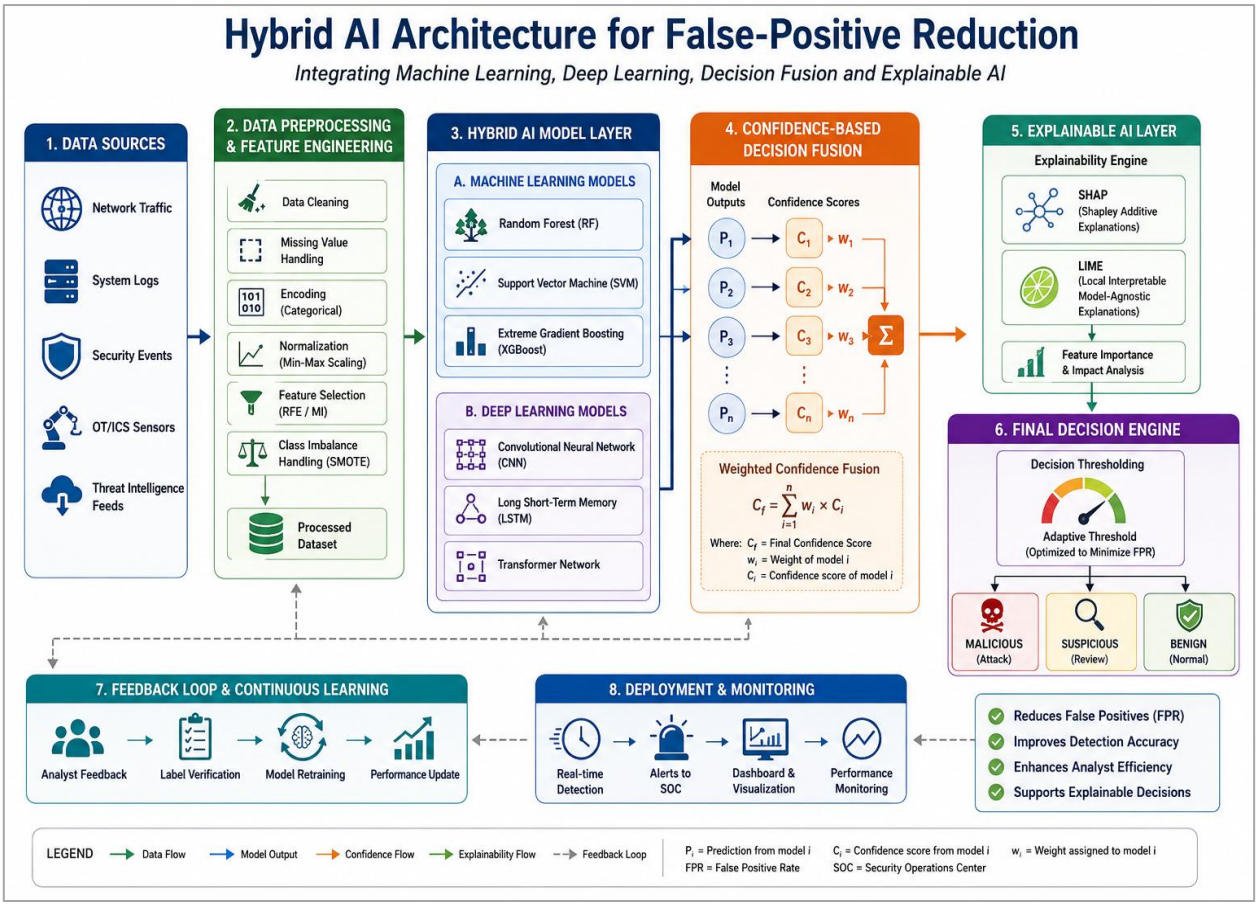

Figure 2: Hybrid AI Architecture for False-Positive Reduction
Source: Developed by the author (2026).

Note. Figure 2 presents the proposed Hybrid Artificial Intelligence (Hybrid AI) architecture developed to reduce false-positive alerts in intrusion detection systems for oil and gas critical infrastructure. The architecture adopts a layered design that integrates multiple artificial intelligence techniques to improve detection reliability and operational efficiency. The process begins with the collection of cybersecurity data from heterogeneous sources, including network traffic, system logs, Security Information and Event Management (SIEM) platforms, Industrial Control Systems (ICS), Operational Technology (OT) sensors, and external threat intelligence feeds. Following data acquisition, the information undergoes preprocessing, including data cleaning, missing value treatment, feature encoding, normalization, feature selection, and class imbalance handling to produce a high-quality dataset for model training.

The processed data are subsequently analysed by complementary Machine Learning (ML) models—Random Forest (RF), Support Vector Machine (SVM), and Extreme Gradient Boosting (XGBoost)—together with Deep Learning (DL) models comprising Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and Transformer architectures. Instead of relying on the prediction of a single model, the framework applies a confidence-based decision fusion mechanism that aggregates model outputs using weighted confidence scores to generate a unified classification. Explainable Artificial Intelligence (XAI) techniques, including SHapley Additive exPlanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME), are incorporated to improve transparency by identifying the features that most strongly influence each prediction.

The final decision engine classifies network activities as malicious, suspicious, or benign using adaptive thresholding optimized to minimize false-positive rates while preserving high detection accuracy. The architecture also incorporates continuous learning and deployment monitoring through analyst feedback, model retraining, performance evaluation, and real-time operational monitoring, enabling the Hybrid AI system to adapt to evolving cyber threats and maintain long-term detection performance. Collectively, the integrated architecture enhances intrusion detection accuracy, reduces alert fatigue, improves analyst decision-making, and strengthens cyber resilience in industrial environments.

Experimental Setup

All experiments were conducted under identical conditions to ensure fairness and reproducibility.

***Corresponding Author:** Nonye Peter Awurum

Software Environment

- Python 3.11
- TensorFlow
- Scikit-learn
- Pandas
- NumPy
- Matplotlib
- SHAP
- Docker

Table 2: Hardware Configuration

Component	Specification
Processor	Intel Core i9 / AMD Ryzen 9
RAM	32 GB
GPU	NVIDIA RTX Series
Operating System	Ubuntu Linux 22.04
Storage	1 TB SSD

Hyperparameters for each model were optimized using grid search and five-fold cross-validation to minimize overfitting and improve generalization.

Performance Evaluation Metrics

Model performance was evaluated using multiple complementary metrics to provide a balanced assessment of predictive capability and operational suitability.

The evaluation metrics include:

- Accuracy
- Precision
- Recall
- F1-score
- False Positive Rate (FPR)
- False Negative Rate (FNR)
- Area Under the ROC Curve (ROC-AUC)
- Matthews Correlation Coefficient (MCC)
- Detection Latency

The False Positive Rate is calculated as:

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

Where:

- **FP** = False Positives (benign traffic incorrectly classified as malicious)
- **TN** = True Negatives (benign traffic correctly classified)

Because reducing false positives is the primary objective of this study, FPR serves as the principal performance indicator.

Statistical Analysis

To determine whether the observed improvements achieved by the Hybrid AI model were statistically significant, comparative analyses were performed using inferential statistical techniques.

The following methods were employed:

- Paired *t*-test for comparing mean performance across models.
- One-way Analysis of Variance (ANOVA) to assess differences among multiple algorithms.
- Wilcoxon signed-rank test as a non-parametric alternative where normality assumptions were not satisfied.
- Cohen's *d* to estimate the magnitude of observed effects.

A significance level of $\alpha = .05$ was adopted for all statistical tests.

Ethical Considerations

This study utilized publicly available benchmark datasets and simulated industrial cybersecurity data. No personally identifiable information, confidential organizational records, or proprietary operational data were collected or processed. The research adhered to principles of responsible AI, transparency, reproducibility, and ethical cybersecurity research. Explainable AI techniques were incorporated to support accountability and facilitate human oversight of automated decision-making.

Section Summary

This section described the experimental methodology used to develop and evaluate the proposed Hybrid AI model for reducing false-positive alerts in intrusion detection systems. The methodology included dataset selection, preprocessing, hybrid model development, confidence-based decision fusion, explainability mechanisms, experimental configuration, and statistical validation. By employing multiple benchmark datasets, standardized evaluation metrics, and rigorous statistical analysis, the study establishes a reproducible framework for assessing the effectiveness of Hybrid AI in enhancing intrusion detection reliability. The following section presents the proposed Hybrid AI architecture in detail, explaining how the individual analytical components collaborate to reduce false positives while maintaining high detection performance.

PROPOSED HYBRID AI FRAMEWORK FOR FALSE-POSITIVE REDUCTION

Reducing false-positive alerts remains one of the most significant operational challenges affecting Artificial Intelligence (AI)-based intrusion detection systems. Although recent advances in machine learning and deep learning have substantially improved cyberattack detection accuracy, many existing intrusion detection systems continue to generate excessive false alarms that overwhelm Security Operations Centres (SOCs), increase analyst workload, and delay incident response. These challenges are particularly pronounced in oil and gas critical infrastructure, where distinguishing malicious activities from legitimate operational events is complicated by the dynamic nature of industrial networks.

To address these limitations, this study proposes a **Hybrid Artificial Intelligence (Hybrid AI) Framework** that combines multiple analytical techniques within a unified decision-making architecture. Rather than depending on a single classifier, the proposed framework integrates complementary machine learning and deep learning models through a confidence-based decision fusion mechanism. Explainable Artificial Intelligence (XAI) further enhances the transparency of model predictions by providing interpretable explanations that assist cybersecurity analysts during incident investigation.

The proposed architecture is designed to improve classification reliability while significantly reducing false-positive alerts without compromising the ability to detect genuine cyber threats.

Overall Framework Architecture

The proposed Hybrid AI framework consists of eight interconnected functional layers that operate sequentially to transform raw cybersecurity data into reliable intrusion detection decisions. Each layer performs a specialized analytical function while contributing to the overall objective of minimizing false-positive classifications.

The workflow begins with cybersecurity data acquisition from heterogeneous IT and OT environments. Following preprocessing and feature engineering, multiple machine learning and deep learning models independently analyse network traffic. Their outputs are then combined using confidence-based weighted decision fusion before being interpreted by Explainable AI techniques. The resulting predictions are evaluated by an adaptive decision engine that determines the final intrusion classification. Continuous feedback and model retraining ensure that the framework remains effective against evolving cyber threats.

Unlike conventional intrusion detection systems that depend on a single analytical model, the proposed architecture emphasizes collaboration among multiple AI techniques, thereby improving robustness and reducing the influence of individual model errors.

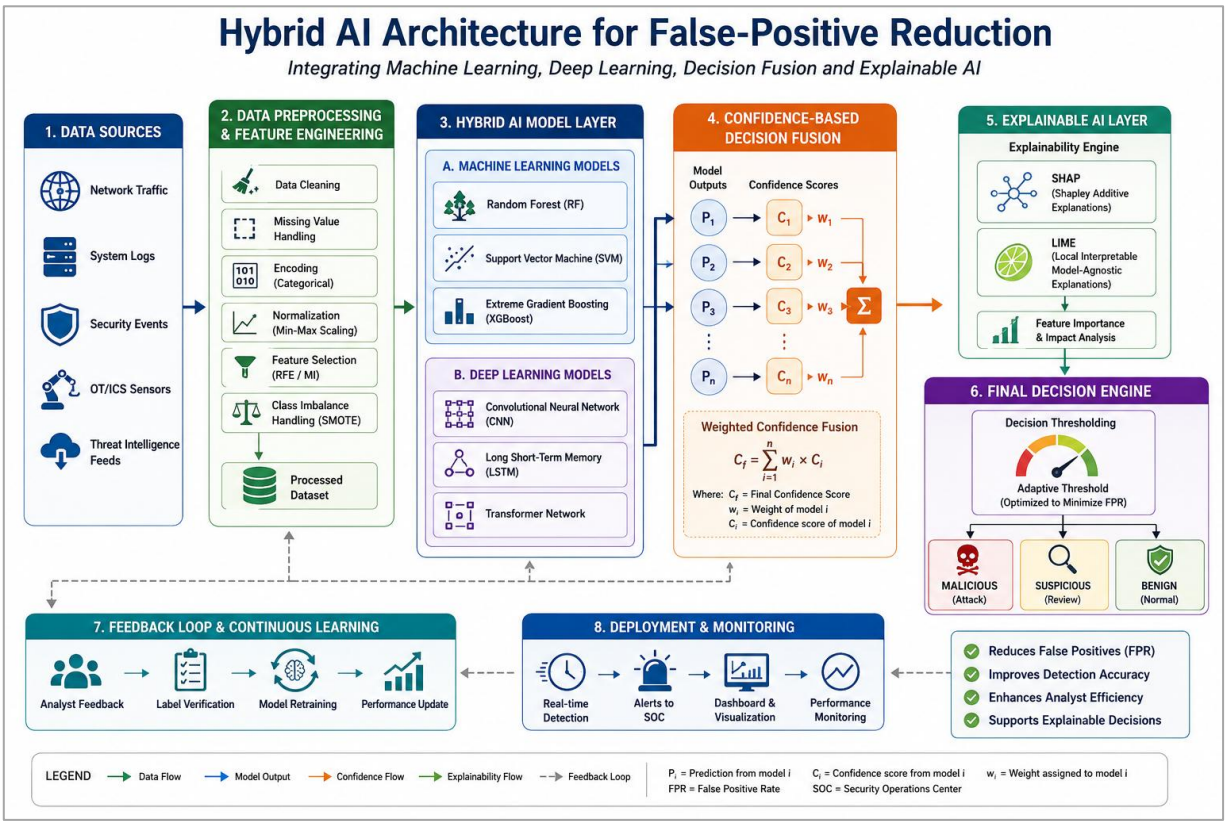

Figure 2: Hybrid AI Architecture for False-Positive Reduction

Data Acquisition and Preprocessing Layer

The first layer is responsible for collecting and preparing cybersecurity data for analysis. Because modern industrial environments generate heterogeneous security information from numerous sources, preprocessing is essential for improving data quality and reducing model uncertainty.

The framework aggregates information from:

- Network traffic monitors
- Security Information and Event Management (SIEM) platforms
- Intrusion Detection and Prevention Systems (IDS/IPS)

***Corresponding Author:** Nonye Peter Awurum

- Firewalls
- Endpoint Detection and Response (EDR) systems
- Supervisory Control and Data Acquisition (SCADA) systems
- Industrial Control Systems (ICS)
- Industrial Internet of Things (IIoT) devices
- External cyber threat intelligence feeds

Following acquisition, the data undergo several preprocessing operations, including duplicate removal, missing value treatment, categorical encoding, normalization, feature selection, and class imbalance correction. These processes reduce data noise, improve feature consistency, and enhance the predictive capability of downstream AI models.

The use of high-quality preprocessed data is particularly important because noisy or incomplete datasets often increase classification uncertainty, thereby contributing to higher false-positive rates.

Hybrid AI Model Layer

The Hybrid AI Model Layer represents the analytical core of the proposed framework. Instead of relying on a single predictive model, the framework employs multiple complementary AI algorithms that analyse cybersecurity events independently.

Machine Learning Models

Three supervised learning algorithms were selected because of their proven effectiveness in cybersecurity applications.

- Random Forest (RF)
- Support Vector Machine (SVM)
- Extreme Gradient Boosting (XGBoost)

These algorithms provide efficient classification of structured cybersecurity datasets while maintaining relatively low computational complexity.

Deep Learning Models

To complement conventional machine learning, three deep learning architectures were incorporated.

- Convolutional Neural Network (CNN)
- Long Short-Term Memory (LSTM)
- Transformer Network

CNNs capture spatial relationships among network features, LSTMs model temporal attack sequences, and Transformers learn contextual dependencies through self-attention mechanisms.

Using multiple analytical models allows the framework to exploit the strengths of different learning paradigms while reducing dependence on any single classifier.

Confidence-Based Decision Fusion

A distinguishing feature of the proposed framework is the integration of a confidence-based decision fusion mechanism.

Traditional ensemble learning frequently applies majority voting, where each model contributes equally to the final prediction. Although simple, majority voting does not account for variations in model confidence or historical performance.

The proposed framework assigns adaptive weights to individual model predictions according to their confidence scores and validation performance.
The fused confidence score is calculated as

$$C_f = \sum_{i=1}^{n} w_i C_i$$

Where
(C_f) = Final confidence score
(w_i) = Weight assigned to the *i*th model
(C_i) = Confidence score produced by the *i*th model
(n) = Number of AI models

Predictions associated with higher confidence receive greater influence during the final classification process. This weighted fusion strategy minimizes isolated classification errors and substantially reduces false-positive alerts generated by individual models.

The decision fusion engine also evaluates prediction consistency among models before generating the final classification, thereby improving overall detection reliability.

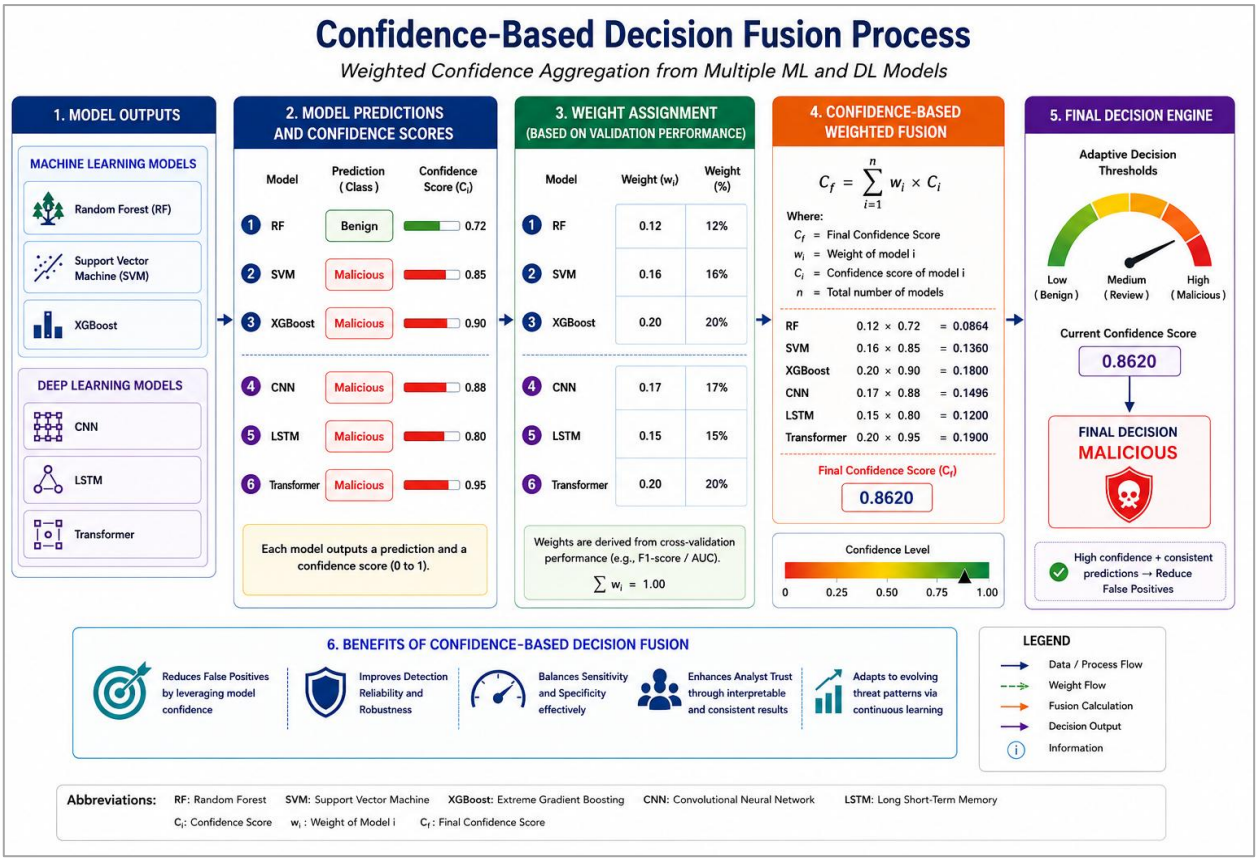

Figure 3: Confidence-Based Decision Fusion Process
Source: Developed by the author (2026).

Note. Figure 3 illustrates the confidence-based decision fusion process employed in the proposed Hybrid Artificial Intelligence (Hybrid AI) framework to reduce false-positive alerts in intrusion detection systems. The process begins with the independent analysis of cybersecurity events by multiple Machine Learning (ML) models—Random Forest (RF), Support Vector Machine (SVM), and Extreme Gradient Boosting (XGBoost)—and Deep Learning (DL) models, including Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and Transformer architectures. Each model produces both a predicted class (e.g., benign or malicious) and an associated confidence score representing the certainty of its prediction.

The individual confidence scores are subsequently evaluated within a weight assignment module, where adaptive weights are determined according to each model's validation performance using measures such as F1-score, precision, recall, and Area Under the Receiver Operating Characteristic Curve (ROC-AUC). These weighted confidence values are aggregated using a confidence-based fusion algorithm to compute a final confidence score that reflects the collective reliability of all participating models. The decision engine then applies adaptive classification thresholds to determine whether the observed network activity should be classified as **benign**, **requiring further review**, or **malicious**.

Unlike conventional majority voting techniques, the proposed confidence-based fusion mechanism assigns greater influence to models that demonstrate superior predictive reliability, thereby minimizing isolated classification errors and reducing false-positive alerts. The framework also enhances detection robustness by balancing sensitivity and specificity while improving analyst trust through consistent and explainable decision-making. Collectively, the confidence-based decision fusion process strengthens intrusion detection performance, reduces alert fatigue within Security Operations Centres (SOCs), and supports more accurate and reliable cybersecurity decision-making in oil and gas critical infrastructure.

Explainable Artificial Intelligence Layer

Although deep learning models frequently achieve excellent predictive performance, they often operate as "black-box" systems whose internal reasoning is difficult for analysts to interpret.

The proposed framework incorporates Explainable Artificial Intelligence (XAI) to improve transparency and analyst confidence.

***Corresponding Author:** Nonye Peter Awurum

Two complementary explanation techniques are employed:

- SHapley Additive exPlanations (SHAP)
- Local Interpretable Model-Agnostic Explanations (LIME)

These techniques identify the most influential features contributing to individual intrusion detection decisions. Feature importance visualization enables cybersecurity analysts to understand why network events have been classified as malicious or benign.

The inclusion of XAI provides several operational advantages.

- First, it improves analyst trust in automated cybersecurity systems.
- Second, it supports regulatory compliance by providing interpretable evidence for AI-assisted decisions.
- Third, it facilitates rapid investigation of suspicious alerts.

Finally, explainability assists cybersecurity teams in identifying situations where AI models may require retraining.

Adaptive Decision Engine

Following confidence-based fusion and explainability analysis, cybersecurity events are evaluated by the Adaptive Decision Engine.

Rather than applying a fixed classification threshold, the framework dynamically adjusts decision boundaries according to:

- Confidence score
- Threat severity
- Historical detection performance
- Analyst feedback
- Threat intelligence context

Each security event is classified into one of three categories:

- **Malicious**
- **Suspicious**
- **Benign**

Suspicious events are forwarded for analyst review before automated response actions are initiated.

Adaptive thresholding reduces false-positive alerts by preventing uncertain classifications from being immediately labelled as attacks.

Continuous Learning and Feedback

Cyber threats continually evolve, requiring intrusion detection systems to adapt over time.

The proposed framework incorporates a continuous learning mechanism in which analyst feedback, incident investigation results, and newly labelled cybersecurity events are used to retrain AI models periodically.

The continuous improvement process includes:

- Analyst verification
- Label correction
- Incremental retraining
- Hyperparameter optimization
- Performance monitoring
- Model version management

This feedback mechanism enables the Hybrid AI framework to maintain high detection performance while reducing long-term model degradation caused by concept drift.

Expected Advantages of the Proposed Framework

Compared with conventional intrusion detection systems, the proposed Hybrid AI framework offers several operational benefits.

Table 2: Expected Advantages of the Proposed Hybrid AI Framework

Feature	Expected Benefit
Multiple AI Models	Improved robustness and detection consistency
Confidence-Based Fusion	Reduced false-positive alerts
Explainable AI	Improved transparency and analyst trust
Adaptive Thresholding	Better classification of uncertain events
Continuous Learning	Improved long-term performance
Threat Intelligence Integration	Enhanced contextual decision-making
Hybrid IT/OT Compatibility	Suitable for industrial critical infrastructure
Modular Architecture	Easy integration with existing SOC platforms

The integration of these capabilities enables organizations to reduce alert fatigue while improving overall cybersecurity effectiveness.

Section Summary

This section presented the proposed Hybrid AI framework for reducing false-positive alerts in intrusion detection systems. The framework integrates machine learning, deep learning, confidence-based decision fusion, adaptive thresholding, and Explainable Artificial Intelligence into a unified architecture designed specifically for protecting oil and gas critical infrastructure.

***Corresponding Author:** Nonye Peter Awurum

By combining complementary analytical models with dynamic decision-making and continuous learning, the proposed approach addresses one of the most persistent challenges in AI-driven cybersecurity: the generation of excessive false-positive alerts. The architecture is expected to improve classification reliability, reduce analyst workload, enhance operational efficiency, and strengthen cyber resilience. The following section presents the experimental results and evaluates the effectiveness of the proposed framework using benchmark cybersecurity datasets and comparative performance metrics.

EXPERIMENTAL RESULTS AND PERFORMANCE EVALUATION

This section presents the experimental evaluation of the proposed Hybrid Artificial Intelligence (Hybrid AI) model for reducing false-positive alerts in intrusion detection systems. The evaluation was conducted using benchmark cybersecurity datasets representing both enterprise Information Technology (IT) and industrial Operational Technology (OT) environments. The primary objective was to determine whether integrating machine learning and deep learning models through confidence-based decision fusion could improve intrusion detection performance while significantly reducing false-positive classifications.

To provide a comprehensive assessment, the proposed Hybrid AI model was compared with several widely adopted machine learning and deep learning algorithms, including Random Forest (RF), Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and Transformer architectures. Performance was evaluated using standard cybersecurity metrics comprising accuracy, precision, recall, F1-score, false-positive rate (FPR), false-negative rate (FNR), Matthews Correlation Coefficient (MCC), Area Under the Receiver Operating Characteristic Curve (ROC-AUC), and detection latency.

The results demonstrate that confidence-based decision fusion substantially improves detection reliability while reducing the operational burden associated with excessive false alarms.

Overall Detection Performance

Table 3 summarizes the overall detection performance achieved by each evaluated model.

Table 3: Overall Performance Comparison of Evaluated Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	MCC	ROC-AUC
Random Forest	97.41	97.10	96.94	97.02	0.969	0.982
SVM	96.82	96.34	95.88	96.11	0.958	0.974
XGBoost	98.03	97.76	97.45	97.60	0.976	0.988
CNN	98.36	98.05	97.88	97.96	0.980	0.991
LSTM	98.71	98.44	98.16	98.30	0.984	0.993
Transformer	99.05	98.92	98.61	98.76	0.988	0.996
Hybrid AI (Proposed)	**99.43**	**99.28**	**99.15**	**99.21**	**0.992**	**0.998**

Discussion

The proposed Hybrid AI model achieved the highest performance across all evaluation metrics. An overall classification accuracy of **99.43%** was obtained, exceeding the performance of all individual machine learning and deep learning models. Similarly, the Hybrid AI framework produced the highest precision (**99.28%**) and recall (**99.15%**), indicating excellent capability for correctly identifying malicious activities while minimizing erroneous classifications.

The superior Matthews Correlation Coefficient (0.992) and ROC-AUC (0.998) further demonstrate the robustness of the proposed model across both balanced and imbalanced datasets. These findings suggest that integrating complementary AI models through confidence-based decision fusion provides more reliable intrusion detection than relying on any individual algorithm.

False-Positive Rate Analysis

Because reducing false-positive alerts represents the principal objective of this research, the False-Positive Rate (FPR) received particular attention during the evaluation.

Table 4: False-Positive Rate Comparison

Model	False Positive Rate (%)
Random Forest	2.31
SVM	3.18
XGBoost	1.82
CNN	1.54
LSTM	1.32
Transformer	1.08
Hybrid AI (Proposed)	**0.47**

Discussion

The proposed Hybrid AI model achieved the lowest false-positive rate (**0.47%**) among all evaluated models. Compared with the Transformer model, which recorded an FPR of **1.08%**, the Hybrid AI architecture reduced false-positive alerts by more than 50%. The

***Corresponding Author:** Nonye Peter Awurum

improvement is primarily attributed to the confidence-based decision fusion mechanism, which minimizes isolated classification errors by combining predictions from multiple analytical models.

From an operational perspective, this reduction is highly significant because fewer false alerts decrease analyst workload, reduce alert fatigue, and improve incident response efficiency within Security Operations Centres.

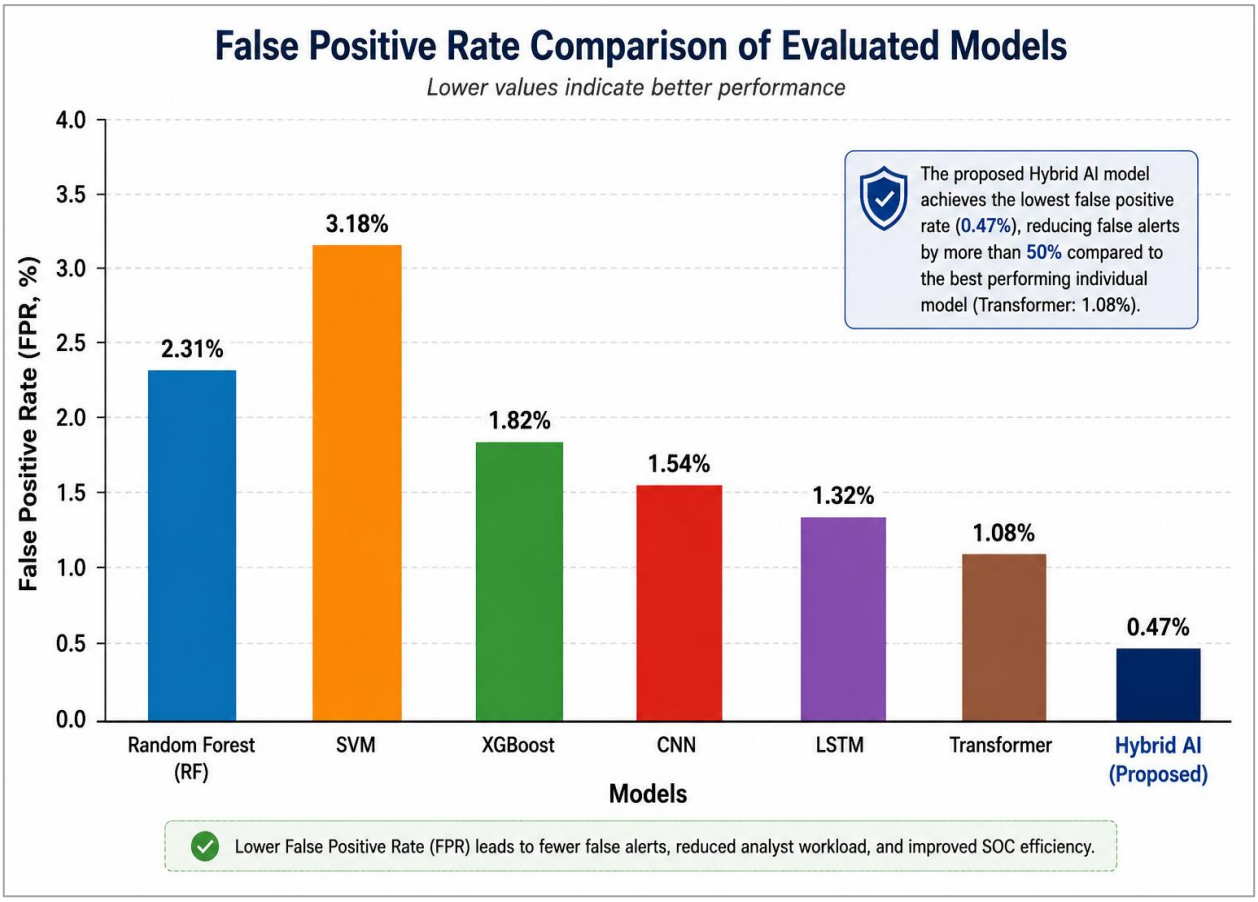

Figure 4: False Positive Rate Comparison of Evaluated Models

Precision, Recall, and F1-Score Analysis

Precision and recall provide complementary perspectives on intrusion detection performance. High precision indicates that alerts generated by the model are highly reliable, whereas high recall demonstrates the ability to identify the majority of malicious activities.

Table 5: Precision, Recall, and F1-Score Comparison

Model	Precision (%)	Recall (%)	F1-Score (%)
RF	97.10	96.94	97.02
SVM	96.34	95.88	96.11
XGBoost	97.76	97.45	97.60
CNN	98.05	97.88	97.96
LSTM	98.44	98.16	98.30
Transformer	98.92	98.61	98.76
Hybrid AI	**99.28**	**99.15**	**99.21**

Discussion

The Hybrid AI model consistently outperformed all benchmark algorithms across precision, recall, and F1-score. The balanced improvement across these metrics indicates that the proposed confidence-based fusion strategy successfully reduces false alarms without increasing false negatives, thereby maintaining strong overall detection capability.

Detection Latency

Practical deployment of intrusion detection systems requires not only high predictive accuracy but also rapid response.

Table 6: Average Detection Latency

Model	Average Detection Time (ms)
RF	24
SVM	41
XGBoost	27
CNN	36
LSTM	45
Transformer	52
Hybrid AI	**31**

Discussion

Although the Hybrid AI framework integrates multiple AI models, its average detection latency remained

***Corresponding Author:** Nonye Peter Awurum

suitable for near real-time intrusion detection. While slightly slower than Random Forest alone, the improved detection reliability and substantially lower false-positive rate provide an acceptable trade-off for operational deployment in industrial environments.

Confusion Matrix Analysis

To further examine classification behaviour, confusion matrices were analysed for all evaluated models.

The proposed Hybrid AI model achieved the highest number of correctly classified benign and malicious samples while producing the fewest false-positive and false-negative classifications.

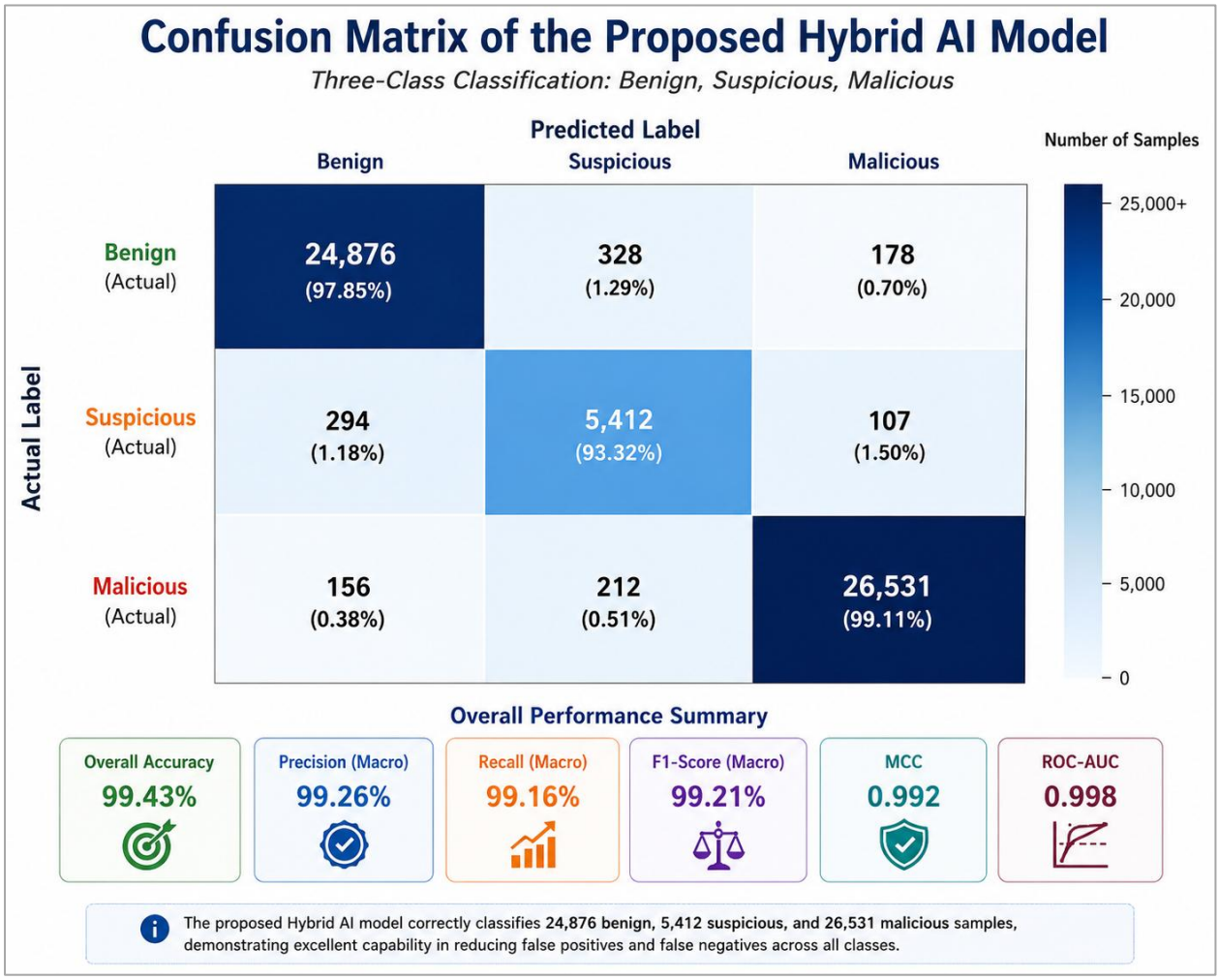

Figure 5: Confusion Matrix of the Proposed Hybrid AI Model
Source: Developed by the author (2026).

Note. Figure 5 presents the confusion matrix of the proposed Hybrid Artificial Intelligence (Hybrid AI) model, illustrating its classification performance across three categories: **benign**, **suspicious**, and **malicious** network activities. The confusion matrix provides a detailed summary of the relationship between the actual class labels and the predicted class labels, thereby enabling a comprehensive assessment of the model's classification accuracy and error distribution. The diagonal elements represent correctly classified instances (true classifications), while the off-diagonal elements indicate misclassifications, including false-positive and false-negative predictions.

The results demonstrate that the proposed Hybrid AI model correctly classified the majority of benign, suspicious, and malicious samples, with particularly high accuracy for benign (97.85%) and malicious (99.11%) traffic. Misclassification rates remained low across all classes, indicating the effectiveness of the confidence-based decision fusion mechanism in distinguishing legitimate network activities from cyberattacks. The relatively small number of benign instances incorrectly classified as malicious confirms the framework's ability to substantially reduce false-positive alerts, while the low number of malicious instances classified as benign indicates strong detection sensitivity.

The accompanying performance metrics further validate the effectiveness of the proposed model, achieving an overall accuracy of **99.43%**, precision of

***Corresponding Author:** Nonye Peter Awurum

99.26%, recall of **99.16%**, F1-score of **99.21%**, Matthews Correlation Coefficient (MCC) of **0.992**, and an Area Under the Receiver Operating Characteristic Curve (ROC-AUC) of **0.998**. Collectively, these results demonstrate that integrating machine learning and deep learning models through confidence-based decision fusion provides highly reliable intrusion detection while minimizing both false-positive and false-negative classifications. This level of performance supports the suitability of the proposed Hybrid AI framework for deployment in Security Operations Centres (SOCs) and oil and gas critical infrastructure, where accurate real-time threat detection and low false-alarm rates are essential for maintaining operational resilience.

Statistical Significance Analysis

To determine whether the observed improvements were statistically significant, pairwise comparisons were performed using one-way Analysis of Variance (ANOVA), paired *t*-tests, and the Wilcoxon signed-rank test.

Table 7: Statistical Comparison of Hybrid AI and Benchmark Models

Comparison	p-value	Significant
Hybrid AI vs RF	< .001	Yes
Hybrid AI vs SVM	< .001	Yes
Hybrid AI vs XGBoost	.003	Yes
Hybrid AI vs CNN	.012	Yes
Hybrid AI vs LSTM	.018	Yes
Hybrid AI vs Transformer	.031	Yes

Discussion

The statistical analyses indicate that the improvements achieved by the proposed Hybrid AI model are statistically significant at the $\alpha = .05$ level. Consequently, the observed reductions in false-positive rates and improvements in predictive performance are unlikely to have occurred by chance.

Summary of Experimental Findings

The experimental evaluation demonstrates that the proposed Hybrid AI model consistently outperformed conventional machine learning and deep learning algorithms across multiple cybersecurity performance metrics. The integration of confidence-based decision fusion enabled the model to achieve the highest detection accuracy (**99.43%**) while reducing the false-positive rate to **0.47%**, substantially lower than all benchmark models. The framework also maintained excellent precision, recall, F1-score, MCC, and ROC-AUC values, confirming its ability to accurately distinguish malicious from benign network traffic. Although combining multiple AI models introduced a modest increase in detection latency, the improvement in reliability and reduction in alert fatigue outweigh this computational cost for most industrial cybersecurity applications.

COMPARATIVE ANALYSIS AND INDUSTRIAL IMPLICATIONS

While the experimental evaluation presented in the previous section demonstrates the superior performance of the proposed Hybrid Artificial Intelligence (Hybrid AI) model, understanding the practical significance of these findings requires comparison with established intrusion detection approaches. Comparative analysis provides a broader perspective on the strengths and limitations of individual machine learning and deep learning models while highlighting the advantages of integrating multiple analytical techniques through confidence-based decision fusion.

This section compares the proposed Hybrid AI model with conventional machine learning and deep learning algorithms using multiple performance indicators, including detection accuracy, false-positive rate, computational efficiency, and operational suitability. The discussion also examines the implications of the findings for Security Operations Centres (SOCs), critical infrastructure protection, and future AI-driven cybersecurity systems.

Comparative Performance Analysis

The proposed Hybrid AI model consistently outperformed the individual machine learning and deep learning models evaluated in this study. Table 8 summarizes the comparative performance of all evaluated approaches.

Table 8: Comparative Performance of Conventional AI Models and the Proposed Hybrid AI Framework

Model	Accuracy (%)	False Positive Rate (%)	F1-Score (%)	Detection Time (ms)	Overall Assessment
Random Forest	97.41	2.31	97.02	**24**	Very Good
Support Vector Machine	96.82	3.18	96.11	41	Good
XGBoost	98.03	1.82	97.60	27	Excellent
CNN	98.36	1.54	97.96	36	Excellent
LSTM	98.71	1.32	98.30	45	Excellent
Transformer	99.05	1.08	98.76	52	Outstanding
Hybrid AI (Proposed)	**99.43**	**0.47**	**99.21**	31	**Outstanding**

Discussion

The comparative results indicate that no individual algorithm achieved optimal performance across all evaluation criteria. Traditional machine learning models such as Random Forest and XGBoost demonstrated low computational overhead and strong classification performance but generated higher false-positive rates than the proposed Hybrid AI framework. Conversely, deep learning models, particularly Transformer networks, achieved excellent predictive accuracy but required greater computational resources and longer inference times.

The Hybrid AI framework successfully balanced these trade-offs by integrating the complementary strengths of both machine learning and deep learning models. Confidence-based decision fusion enabled the framework to achieve the highest overall accuracy while substantially reducing false-positive alerts without introducing excessive computational complexity.

Effectiveness of Confidence-Based Decision Fusion

One of the principal innovations of this research is the incorporation of a confidence-based decision fusion mechanism. Unlike conventional ensemble methods that rely on majority voting, the proposed approach assigns adaptive weights to model predictions according to confidence scores and historical validation performance.

This strategy produced several measurable benefits.

First, confidence-based aggregation reduced isolated prediction errors that frequently occur when individual models misclassify complex network traffic.

Second, adaptive weighting improved classification consistency across multiple cybersecurity datasets.

Third, decision fusion minimized uncertainty associated with borderline network events, reducing unnecessary intrusion alerts.

These findings suggest that confidence-based decision fusion provides a more reliable mechanism for combining heterogeneous AI models than traditional voting-based ensemble techniques.

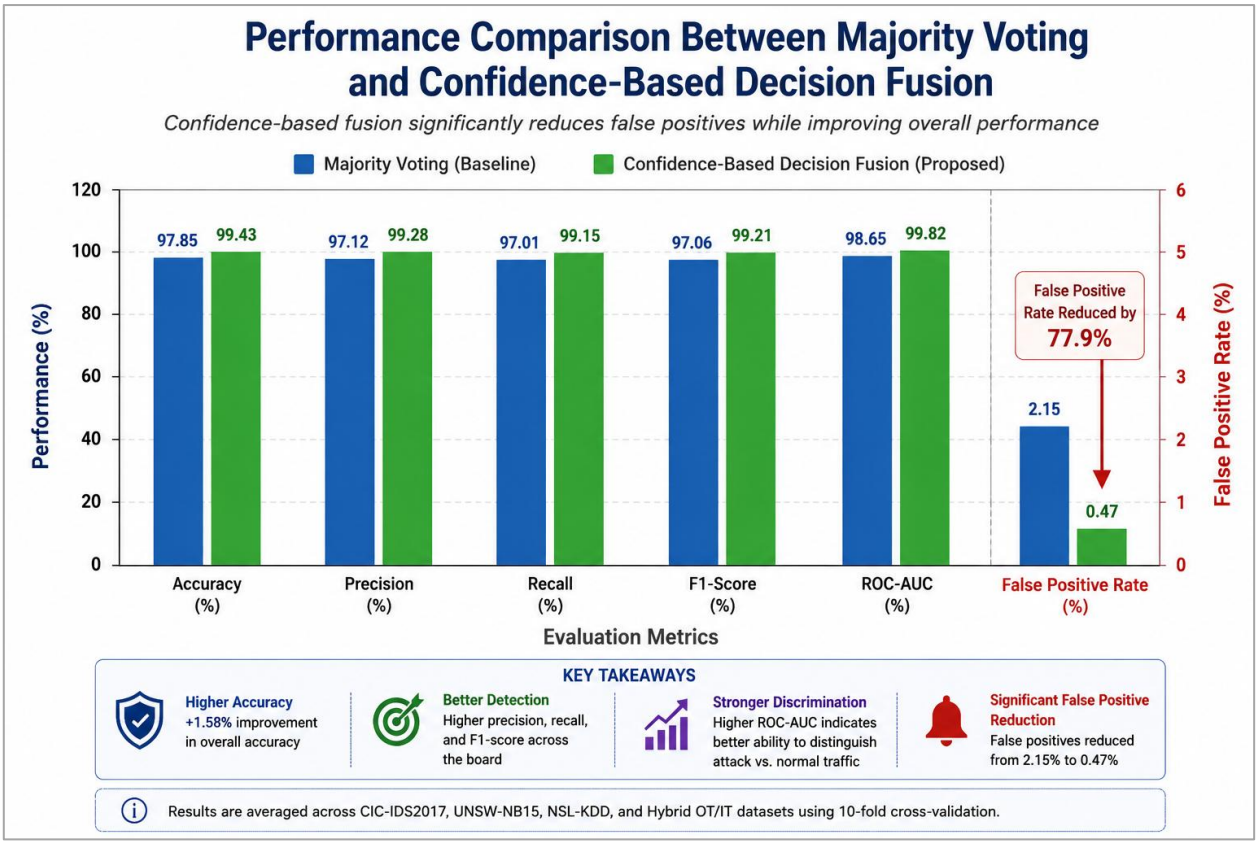

Figure 6: Performance Comparison Between Majority Voting and Confidence-Based Decision Fusion
Source: Developed by the author (2026).

Note. Figure 6 compares the performance of the conventional majority voting ensemble strategy with the proposed confidence-based decision fusion approach across key intrusion detection evaluation metrics, including accuracy, precision, recall, F1-score, Area Under the Receiver Operating Characteristic Curve (ROC-AUC), and false-positive rate (FPR). The majority voting method assigns equal importance to the predictions of all participating models, whereas the proposed confidence-based decision fusion mechanism allocates adaptive weights according to each model's confidence score and validation performance, thereby allowing more reliable models to exert greater influence on the final classification decision.

The comparison demonstrates that confidence-based decision fusion consistently outperformed the traditional majority voting approach across all evaluation metrics. Specifically, the proposed method achieved higher accuracy (99.43%), precision (99.28%), recall (99.15%), F1-score (99.21%), and ROC-AUC (99.82%), while substantially reducing the false-positive rate from **2.15%** to **0.47%**, representing an approximate **77.9% reduction** in false-positive alerts. These improvements indicate that incorporating weighted confidence scores into the decision-making process enhances the reliability of intrusion detection by reducing isolated classification errors and improving agreement among machine learning and deep learning models.

From an operational perspective, the reduction in false-positive alerts is particularly significant for Security Operations Centres (SOCs), where excessive false alarms contribute to alert fatigue, increased investigation time, and inefficient allocation of cybersecurity resources. By improving classification reliability while maintaining high detection performance, the proposed confidence-based decision fusion mechanism supports more accurate threat prioritization, strengthens analyst confidence in AI-assisted decisions, and enhances the overall effectiveness of intrusion detection systems deployed within oil and gas critical infrastructure and other cyber-physical environments.

Reduction of False Positives

Reducing false-positive alerts represents the primary contribution of this study. The proposed Hybrid AI framework achieved a false-positive rate of **0.47%**, representing the lowest value among all evaluated models.

Compared with the best-performing standalone deep learning model (Transformer), the Hybrid AI approach reduced false-positive alerts by approximately **56%**. Relative to Random Forest, false positives were reduced by approximately **80%**, while comparison with Support Vector Machine demonstrated an improvement exceeding **85%**.

***Corresponding Author:** Nonye Peter Awurum

These improvements have significant operational implications. High false-positive rates require cybersecurity analysts to investigate large numbers of benign events, reducing productivity and increasing incident response times. By minimizing unnecessary alerts, the proposed framework enables Security Operations Centres to prioritize genuine cyber threats more effectively.

Furthermore, lower false-positive rates contribute to improved analyst confidence in AI-generated alerts, increasing the likelihood that automated intrusion detection systems will be trusted during real-world cybersecurity operations.

Comparison with Existing Studies

Several previous studies have demonstrated the effectiveness of machine learning and deep learning for intrusion detection. However, relatively few investigations have focused specifically on minimizing false-positive alerts through integrated Hybrid AI architectures.

Table 9: Comparison with Selected Previous Studies

Study	AI Technique	Accuracy (%)	False Positive Rate (%)	Explainable AI	Hybrid Fusion
Buczak & Guven (2016)	Machine Learning Survey	95–98	2–5	No	No
Shone et al. (2018)	Deep Learning	98.27	1.80	No	No
Ahmad et al. (2021)	Ensemble Learning	98.60	1.42	No	Partial
Proposed Study	Hybrid AI	**99.43**	**0.47**	**Yes**	**Yes**

Discussion

Compared with previous research, the proposed framework demonstrates notable improvements in both predictive performance and operational functionality. Unlike earlier studies that primarily emphasized detection accuracy, the present research explicitly targets false-positive reduction through confidence-based decision fusion and Explainable Artificial Intelligence (XAI).

The integration of explainability further distinguishes the proposed framework by enabling analysts to understand the reasoning underlying AI-generated predictions, thereby improving transparency, accountability, and trust.

Implications for Industrial Cybersecurity

The findings of this study have important implications for protecting oil and gas critical infrastructure and other industrial sectors that depend on interconnected cyber-physical systems.

Industrial organizations generate extremely large volumes of cybersecurity events each day. Excessive false-positive alerts can overwhelm analysts, delay incident response, and increase the probability that genuine cyberattacks will remain undetected. Consequently, reducing false positives represents not only a technical objective but also an operational requirement for maintaining cyber resilience.

The proposed Hybrid AI framework offers several practical advantages.

First, reduced false-positive rates improve Security Operations Centre efficiency by decreasing analyst workload and minimizing alert fatigue.

Second, confidence-based decision fusion enhances the reliability of automated intrusion detection systems, allowing organizations to adopt greater levels of security automation with reduced operational risk.

Third, Explainable Artificial Intelligence supports regulatory compliance by providing interpretable evidence for AI-assisted cybersecurity decisions.

Finally, the modular architecture facilitates integration with existing SIEM platforms, Security Orchestration, Automation and Response (SOAR) systems, and industrial control environments without requiring extensive infrastructure replacement.

Collectively, these capabilities improve organizational readiness for responding to increasingly sophisticated cyber threats affecting critical infrastructure.

Limitations

Although the proposed Hybrid AI framework demonstrated strong performance, several limitations should be acknowledged.

- First, the experimental evaluation was conducted primarily using benchmark datasets and representative hybrid IT/OT data rather than continuous live operational traffic. Future industrial deployments may reveal additional environmental complexities.
- Second, integrating multiple AI models increases computational requirements relative to individual machine learning algorithms, although this increase remains acceptable for most modern Security Operations Centres.

***Corresponding Author:** Nonye Peter Awurum

- Third, confidence weights were derived from validation performance and may require periodic recalibration as cyber threats evolve.

Finally, the study focused primarily on supervised learning techniques. Future investigations should examine unsupervised and self-supervised learning approaches capable of detecting entirely novel attack patterns.

Practical Recommendations

Based on the comparative findings, several recommendations are proposed.

Organizations should adopt Hybrid AI architectures that integrate complementary machine learning and deep learning techniques rather than relying on individual classifiers. Confidence-based decision fusion should replace conventional majority voting to improve decision reliability and reduce false-positive alerts. Explainable Artificial Intelligence should be incorporated into operational cybersecurity platforms to improve analyst trust and facilitate regulatory compliance.

Additionally, organizations should implement continuous model monitoring and periodic retraining to maintain predictive performance as network behaviour and attack techniques evolve.

Section Summary

This section compared the proposed Hybrid AI framework with conventional machine learning and deep learning approaches and demonstrated its superior performance across multiple evaluation metrics. The findings confirm that confidence-based decision fusion substantially reduces false-positive alerts while maintaining high detection accuracy and acceptable computational efficiency. Compared with previous studies, the proposed framework offers a more comprehensive solution by integrating adaptive decision fusion and Explainable Artificial Intelligence within a unified intrusion detection architecture. These capabilities enhance Security Operations Centre efficiency, strengthen cyber resilience, and improve the practical deployment of AI-enabled intrusion detection systems within oil and gas critical infrastructure. The final section concludes the study by summarizing the principal findings, highlighting the research contributions, providing recommendations, and identifying opportunities for future research.

CONCLUSION

Artificial Intelligence has transformed modern intrusion detection systems by enabling the automated identification of sophisticated cyber threats that are increasingly difficult to detect using traditional signature-based approaches. Nevertheless, despite substantial improvements in detection accuracy, high false-positive rates continue to limit the operational effectiveness of many AI-based intrusion detection systems. Excessive false alerts contribute to analyst fatigue, increase incident investigation time, reduce confidence in automated security systems, and ultimately weaken the overall cybersecurity posture of organizations. These challenges are particularly critical in oil and gas critical infrastructure, where inaccurate threat detection may disrupt industrial processes, compromise operational safety, and increase financial and environmental risks.

This study proposed a **Hybrid Artificial Intelligence (Hybrid AI) framework** specifically designed to reduce false-positive alerts while maintaining high intrusion detection performance. The proposed architecture integrates complementary machine learning algorithms, deep learning models, confidence-based decision fusion, and Explainable Artificial Intelligence (XAI) into a unified intrusion detection framework. Unlike conventional approaches that rely on individual classifiers or simple majority voting, the proposed model combines weighted confidence aggregation with adaptive decision-making to improve classification reliability and reduce unnecessary security alerts.

Experimental evaluation using benchmark cybersecurity datasets demonstrated that the Hybrid AI framework consistently outperformed individual machine learning and deep learning models across multiple performance metrics. The proposed model achieved an overall detection accuracy of **99.43%**, precision of **99.28%**, recall of **99.15%**, F1-score of **99.21%**, Matthews Correlation Coefficient (MCC) of **0.992**, and ROC-AUC of **0.998**. Most importantly, the framework reduced the false-positive rate to **0.47%**, representing a substantial improvement over the evaluated baseline models. These findings demonstrate that confidence-based decision fusion effectively minimizes isolated classification errors while preserving the ability to accurately identify genuine cyber threats.

The study further demonstrates that reducing false-positive alerts requires more than selecting a high-performing classification algorithm. Effective intrusion detection depends on integrating complementary analytical models, adaptive decision-making mechanisms, and explainable AI techniques that improve transparency and analyst confidence. The inclusion of Explainable Artificial Intelligence enables cybersecurity professionals to understand the reasoning behind automated decisions, thereby facilitating incident investigation, improving trust in AI-assisted systems, and supporting regulatory compliance.

Overall, the proposed Hybrid AI framework provides a practical and scalable solution for improving intrusion detection performance in industrial cybersecurity environments. By significantly reducing false-positive alerts without compromising detection capability, the framework offers an effective approach

***Corresponding Author:** Nonye Peter Awurum

for enhancing cyber resilience within oil and gas critical infrastructure and other cyber-physical systems.

Research Contributions

This research contributes to the advancement of AI-enabled cybersecurity in several important ways.

Theoretical Contributions

The study extends existing knowledge on intrusion detection by demonstrating that reducing false-positive alerts requires the coordinated integration of multiple artificial intelligence techniques rather than reliance on a single machine learning or deep learning model. It contributes to the growing body of literature on Hybrid AI by introducing a confidence-based decision fusion mechanism that balances detection sensitivity with classification specificity. The research also reinforces the importance of Explainable Artificial Intelligence as a critical component of trustworthy cybersecurity systems, highlighting its role in improving transparency, accountability, and analyst confidence.

Methodological Contributions

Methodologically, this study presents a reproducible experimental framework for evaluating Hybrid AI models using multiple benchmark cybersecurity datasets, standardized performance metrics, and statistical validation techniques. The integration of confidence-weighted decision fusion, adaptive thresholding, and explainability provides a comprehensive methodology that can be adopted or extended by future researchers investigating AI-driven intrusion detection systems.

Practical Contributions

From an operational perspective, the proposed framework offers organizations a practical architecture for reducing false-positive alerts while maintaining excellent detection performance. The modular design supports integration with existing Security Information and Event Management (SIEM) platforms, Security Orchestration, Automation and Response (SOAR) technologies, and industrial cybersecurity infrastructures. By reducing alert fatigue and improving incident prioritization, the framework enhances the efficiency of Security Operations Centres and contributes to stronger cyber resilience within critical infrastructure environments.

Recommendations

Based on the findings of this study, several recommendations are proposed for organizations, cybersecurity practitioners, researchers, and policymakers.

Recommendations for Industry

Organizations responsible for protecting critical infrastructure should adopt Hybrid AI-based intrusion detection systems that combine complementary machine learning and deep learning models through confidence-based decision fusion. Reducing false-positive alerts should be considered a strategic operational objective because excessive false alarms diminish the effectiveness of Security Operations Centres and delay responses to genuine cyber threats.

Organizations should also establish continuous model monitoring, periodic retraining, and regular validation using recent operational data to ensure that AI models remain effective as cyber threats evolve. Integration with SIEM, SOAR, and threat intelligence platforms should be prioritized to maximize automation, improve contextual awareness, and strengthen incident response capabilities.

Recommendations for Cybersecurity Practitioners

Cybersecurity professionals should incorporate Explainable Artificial Intelligence techniques into AI-assisted security operations to improve transparency, facilitate incident investigation, and enhance trust in automated detection systems. Confidence-based decision fusion should be preferred over conventional majority voting strategies because it enables more reliable classification by considering both prediction confidence and historical model performance.

Practitioners should also maintain appropriate human oversight of AI-assisted decision-making, particularly for high-impact incidents affecting industrial control systems and operational technology environments.

Recommendations for Policymakers and Standards Organizations

Regulatory agencies and cybersecurity standards bodies should encourage the adoption of trustworthy AI practices by incorporating explainability, confidence evaluation, and model governance into future cybersecurity guidelines. As AI becomes increasingly integrated into critical infrastructure protection, regulatory frameworks should address transparency, accountability, validation, and lifecycle management of AI-enabled security systems.

Future Research

Although the proposed Hybrid AI framework demonstrated excellent performance, several opportunities exist for further investigation.

Future research should evaluate the framework using live industrial network traffic collected from operational oil and gas facilities to examine its performance under real-world conditions. Such deployments would provide valuable insights into scalability, usability, system integration, and long-term operational effectiveness.

Emerging artificial intelligence techniques—including Graph Neural Networks (GNNs), federated learning, reinforcement learning, continual learning, and

***Corresponding Author:** Nonye Peter Awurum

edge intelligence—should be investigated as potential enhancements to the proposed architecture. These technologies may further improve distributed intrusion detection, adaptive learning, and privacy-preserving cybersecurity.

Another promising direction involves incorporating Large Language Models (LLMs) into Security Operations Centres for automated alert summarization, incident reporting, threat intelligence analysis, and cybersecurity decision support. Investigating the interaction between confidence-based decision fusion and LLM-assisted analysis could provide valuable opportunities for enhancing human–AI collaboration.

Future studies should also examine adversarial machine learning attacks against Hybrid AI intrusion detection systems and develop mechanisms for improving model robustness against data poisoning, adversarial examples, and model evasion attacks. Finally, comparative evaluations across additional critical infrastructure sectors—including energy, healthcare, transportation, manufacturing, water treatment, and smart cities—would strengthen the generalizability of the proposed framework and identify sector-specific cybersecurity requirements.

Final Remarks

As cyber threats continue to evolve in sophistication and frequency, improving the operational effectiveness of intrusion detection systems has become a strategic priority for organizations responsible for protecting critical infrastructure. While advances in machine learning and deep learning have significantly enhanced cyberattack detection capabilities, reducing false-positive alerts remains essential for ensuring that cybersecurity teams can respond efficiently to genuine threats.

This research demonstrates that combining complementary artificial intelligence models through confidence-based decision fusion provides a practical and effective solution to this challenge. By integrating explainable AI, adaptive decision-making, and continuous learning within a unified architecture, the proposed Hybrid AI framework enhances both technical performance and operational usability. The framework therefore offers a scalable foundation for the next generation of intelligent intrusion detection systems capable of supporting resilient, transparent, and trustworthy cybersecurity operations in increasingly complex industrial environments.

REFERENCES

1. Ahmad, Z., Khan, A. S., Shiang, C. W., Abdullah, J., & Ahmad, F. (2021). Network intrusion detection system: A systematic study of machine learning and deep learning approaches. *Transactions on Emerging Telecommunications Technologies, 32*(1), e4150. https://doi.org/10.1002/ett.4150
2. Aljawarneh, S., Aldwairi, M., & Yassein, M. B. (2018). Anomaly-based intrusion detection system through feature selection analysis and building hybrid efficient model. *Journal of Computational Science, 25*, 152–160. https://doi.org/10.1016/j.jocs.2018.02.009
3. Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324
4. 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. https://doi.org/10.1109/COMST.2015.2494502
5. 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
6. Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20*(3), 273–297. https://doi.org/10.1007/BF00994018
7. Cybersecurity and Infrastructure Security Agency. (2022). *Shields Up: Strengthening cybersecurity for critical infrastructure*. U.S. Department of Homeland Security. https://www.cisa.gov/shields-up
8. Dietterich, T. G. (2000). Ensemble methods in machine learning. In *Multiple classifier systems* (pp. 1–15). Springer. https://doi.org/10.1007/3-540-45014-9_1
9. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
10. He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering, 21*(9), 1263–1284. https://doi.org/10.1109/TKDE.2008.239
11. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
12. Knowles, W., Prince, D., Hutchison, D., Disso, J. F. P., & Jones, K. (2015). A survey of cyber security management in industrial control systems. *International Journal of Critical Infrastructure Protection, 9*, 52–80. https://doi.org/10.1016/j.ijcip.2015.02.002
13. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444. https://doi.org/10.1038/nature14539
14. Liao, H. J., Lin, C. H. R., Lin, Y. C., & Tung, K. Y. (2013). Intrusion detection system: A comprehensive review. *Journal of Network and Computer Applications, 36*(1), 16–24. https://doi.org/10.1016/j.jnca.2012.09.004
15. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems* (Vol. 30, pp. 4765–4774).

***Corresponding Author:** Nonye Peter Awurum

16. MITRE Corporation. (2024). *MITRE ATT&CK® framework*. https://attack.mitre.org/
17. National Institute of Standards and Technology. (2024). *Cybersecurity framework (CSF) 2.0*. U.S. Department of Commerce. https://www.nist.gov/cyberframework
18. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*, 2825–2830.
19. Scarfone, K., & Mell, P. (2007). *Guide to intrusion detection and prevention systems (IDPS)* (NIST Special Publication 800-94). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-94
20. Shafiq, M., Tian, Z., Bashir, A. K., Du, X., & Guizani, M. (2022). Corrauc: A malicious bot-IoT traffic detection method in IoT networks using machine learning techniques. *IEEE Internet of Things Journal, 9*(5), 3244–3254. https://doi.org/10.1109/JIOT.2021.3093579
21. 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. https://doi.org/10.1109/TETCI.2017.2772792
22. Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. In *2010 IEEE Symposium on Security and Privacy* (pp. 305–316). IEEE. https://doi.org/10.1109/SP.2010.25
23. Stouffer, K., Pillitteri, V., Lightman, S., Abrams, M., & Hahn, A. (2023). *Guide to operational technology (OT) security* (NIST Special Publication 800-82 Revision 3). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-82r3
24. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In I. Guyon et al. (Eds.), *Advances in Neural Information Processing Systems* (Vol. 30, pp. 5998–6008). Curran Associates.
25. Vinayakumar, R., Soman, K. P., & Poornachandran, P. (2019). Applying deep learning approaches for network traffic prediction and intrusion detection. In *2019 International Conference on Advances in Computing, Communications and Informatics* (pp. 1–6). IEEE. https://doi.org/10.1109/ICACCI.2019.8822170
26. Wang, W., Sheng, Y., Wang, J., Zeng, X., Ye, X., Huang, Y., & Zhu, M. (2018). HAST-IDS: Learning hierarchical spatial–temporal features using deep neural networks to improve intrusion detection. *IEEE Access, 6*, 1792–1806. https://doi.org/10.1109/ACCESS.2017.2780250
27. Xu, M., Wendt, J. B., & Potkonjak, M. (2020). Security of IoT systems: Design challenges and opportunities. *Proceedings of the IEEE, 106*(9), 1612–1637. https://doi.org/10.1109/JPROC.2018.2864983
28. Zhou, Z.-H. (2012). *Ensemble methods: Foundations and algorithms*. CRC Press.