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.21652233

INDIANA PUBLICATIONS
PRODUCTIVE AND QUALITY RESEARCH

Research Article

Volume-06|Issue04|2026

Hybrid Ensemble Learning for Real-Time Intrusion Detection in Critical Infrastructure Environments.

Nonye Peter Awurum

Affiliations: Charisma University, USA.

Article History
Received: 22.05.2026
Accepted: 27.06.2026
Published: 27.07.2026

Citation
Awurum, N. P. (2026). Hybrid Ensemble Learning for Real-Time Intrusion Detection in Critical Infrastructure Environments. *Indiana Journal of Economics and Business Management, 6*(4), 16-34.

Abstract: The increasing digitization of critical infrastructure environments has significantly enhanced operational efficiency while simultaneously exposing industrial systems to sophisticated cyber threats. Critical sectors such as oil and gas, energy, transportation, and manufacturing rely heavily on interconnected Information Technology (IT) and Operational Technology (OT) systems, making them attractive targets for cybercriminals and nation-state actors. Traditional intrusion detection systems (IDSs), which are primarily signature-based and rule-driven, often struggle to detect advanced persistent threats (APTs), zero-day attacks, and other evolving cyber threats in real time. Consequently, there is a growing need for intelligent cybersecurity solutions capable of adaptive threat detection and automated response.

This study proposes a hybrid ensemble learning framework for real-time intrusion detection in critical infrastructure environments. The framework integrates machine learning and deep learning techniques to leverage the strengths of individual algorithms while mitigating their weaknesses. Specifically, Random Forest (RF), Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) networks, and Transformer models are combined within an ensemble architecture designed to improve detection accuracy, reduce false positive rates, and enhance system resilience against emerging threats.

Experimental evaluation was conducted using benchmark cybersecurity datasets and a hybrid OT/IT dataset representative of industrial environments. The results demonstrated that the proposed hybrid ensemble model achieved an accuracy of 98.5%, a precision of 97.6%, a recall of 97.8%, and an F1-score of 97.7%, outperforming traditional intrusion detection approaches and individual machine learning models. The findings suggest that hybrid ensemble learning provides a robust and scalable solution for real-time cyberattack detection in critical infrastructure systems. The study contributes to the growing body of knowledge on artificial intelligence-driven cybersecurity and offers practical implications for securing industrial environments against increasingly sophisticated cyber threats.

Keywords: Artificial Intelligence, Cybersecurity, Hybrid Ensemble Learning, Intrusion Detection Systems, Critical Infrastructure Protection, Industrial Control Systems, Machine Learning, Deep Learning.

INTRODUCTION

A Hybrid Ensemble Learning Framework is a machine learning approach that combines multiple different algorithms into a single integrated model to achieve better prediction accuracy, robustness, and reliability than any individual model alone. Basically, in terms of real-time intrusion detection, a hybrid ensemble combines both machine learning (ML) and deep learning (DL) models to detect cyberattacks more effectively.

The rapid advancement of digital technologies has transformed the operational landscape of critical infrastructure sectors worldwide. Industries such as oil and gas, energy generation, water treatment, transportation, and manufacturing increasingly rely on interconnected cyber-physical systems to improve efficiency, automation, and decision-making processes. While this digital transformation has generated significant economic and operational benefits, it has simultaneously expanded the cyber threat landscape and introduced new security vulnerabilities (Stouffer *et al.*, 2023).

Critical infrastructure systems traditionally operated within isolated environments; however, the integration of Information Technology (IT) and Operational Technology (OT) networks has increased exposure to cyberattacks. Industrial Control Systems (ICS), Supervisory Control and Data Acquisition (SCADA) systems, Programmable Logic Controllers (PLCs), and Remote Terminal Units (RTUs) are now routinely connected to enterprise networks and cloud-based services, creating multiple attack vectors for malicious actors (Knowles *et al.*, 2015). Recent cyber incidents, including the Colonial Pipeline ransomware attack and the Triton malware attack targeting industrial safety systems, have demonstrated the potentially devastating consequences of cybersecurity breaches in critical infrastructure environments (CISA, 2022).

*Corresponding Author: Nonye Peter Awurum

Cyber threats targeting industrial environments continue to evolve in complexity and sophistication. Advanced Persistent Threats (APTs), ransomware campaigns, insider threats, distributed denial-of-service (DDoS) attacks, and zero-day exploits have become increasingly difficult to detect using conventional cybersecurity tools (Buczak & Guven, 2016). Traditional intrusion detection systems largely depend on predefined signatures and static rule sets. Although effective against known threats, these systems often fail to identify novel attack patterns, resulting in high false-positive rates and delayed response times (Sommer & Paxson, 2010).

To address these challenges, researchers and practitioners have increasingly explored Artificial Intelligence (AI) and Machine Learning (ML) techniques for cybersecurity applications. AI-powered systems possess the ability to learn from large volumes of data, identify complex behavioral patterns, and adapt to previously unseen threats (Shafiq *et al.*, 2022). Machine learning algorithms such as Random Forests and Support Vector Machines have demonstrated promising results in intrusion detection tasks, while deep learning architectures including Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and Transformer models have further improved detection capabilities by capturing spatial and temporal relationships within network traffic data (Vinayakumar *et al.*, 2019).

Despite these advancements, no single machine learning model consistently delivers optimal performance across diverse cybersecurity scenarios. Traditional machine learning models often provide faster inference times but may struggle with complex attack patterns. Conversely, deep learning approaches achieve higher detection accuracy but frequently require substantial computational resources and extensive training data (Kim *et al.*, 2020). This limitation has motivated the development of hybrid ensemble learning approaches that combine multiple algorithms to exploit their complementary strengths.

Ensemble learning has emerged as a powerful technique in cybersecurity research because it improves prediction accuracy, enhances robustness, and reduces model bias and variance. By integrating multiple learning algorithms into a unified decision-making framework, ensemble models can achieve superior performance compared to individual classifiers (Dietterich, 2000). Recent studies have reported that ensemble-based intrusion detection systems outperform standalone machine learning and deep learning models across a range of cybersecurity datasets (Ahmad *et al.*, 2021).

This study builds upon these developments by proposing a hybrid ensemble learning framework specifically designed for real-time intrusion detection in critical infrastructure environments. Unlike many previous studies that focus exclusively on IT networks, the proposed framework incorporates both IT and OT data sources, providing a more comprehensive approach to industrial cybersecurity. The framework integrates Random Forest, CNN, LSTM, and Transformer models to improve threat detection performance while maintaining operational efficiency.

The significance of this research extends beyond theoretical contributions. As critical infrastructure sectors continue to adopt Industry 4.0 technologies, Industrial Internet of Things (IIoT) devices, and cloud-based platforms, the need for intelligent cybersecurity solutions becomes increasingly urgent. Organizations require security systems capable of detecting sophisticated attacks in real time while minimizing operational disruption and false alarms. The proposed hybrid ensemble framework addresses these requirements by delivering enhanced detection accuracy, improved threat visibility, and adaptive defense capabilities.

The primary objective of this study is to develop and evaluate a hybrid ensemble intrusion detection framework for critical infrastructure environments. Specifically, the study aims to:

- Design a hybrid ensemble architecture integrating machine learning and deep learning techniques.
- Evaluate the effectiveness of the framework using benchmark cybersecurity datasets and hybrid OT/IT data.
- Compare the performance of the proposed model against traditional intrusion detection approaches.
- Assess the suitability of hybrid ensemble learning for real-time cyberattack detection in industrial environments.

The remainder of this paper is organized as follows. Section 2 reviews related work on AI-based intrusion detection and ensemble learning techniques. Section 3 presents the proposed hybrid ensemble framework and research methodology. Section 4 describes the experimental setup and datasets used for evaluation. Section 5 presents the results and performance analysis. Section 6 discusses the implications of the findings, while Section 7 concludes the paper and outlines future research directions.

LITERATURE REVIEW AND RELATED WORK

Introduction

The increasing sophistication of cyber threats targeting critical infrastructure has accelerated the adoption of Artificial Intelligence (AI) technologies in cybersecurity. Traditional security mechanisms, including signature-based intrusion detection systems and rule-based monitoring tools, have proven effective against known threats but often struggle to detect emerging and previously unseen attacks. Consequently, researchers have increasingly explored machine learning, deep learning, and ensemble learning techniques to enhance cyber threat detection capabilities and improve resilience against evolving attack vectors (Buczak & Guven, 2016).

Critical infrastructure sectors, including oil and gas, power generation, transportation, and manufacturing, face unique cybersecurity challenges due to the convergence of Information Technology (IT) and Operational Technology (OT) environments. These sectors require cybersecurity solutions capable of detecting sophisticated attacks in real time while maintaining operational continuity and safety. This section reviews existing research on intrusion detection systems, machine learning and deep learning applications in cybersecurity, ensemble learning approaches, and cybersecurity challenges in industrial environments. The review concludes by identifying research gaps that motivate the development of the proposed hybrid ensemble learning framework.

Intrusion Detection Systems in Cybersecurity

Intrusion Detection Systems (IDSs) represent one of the most widely deployed cybersecurity technologies for monitoring network and system activities. IDS solutions are generally categorized into signature-based detection systems and anomaly-based detection systems (Scarfone & Mell, 2007).

Signature-based IDSs identify malicious activities by comparing observed behavior against known attack signatures. Examples include Snort and Suricata, which are widely used in enterprise environments. Although these systems provide high detection rates for previously identified attacks, they are ineffective against zero-day exploits and novel attack patterns (Sommer & Paxson, 2010).

Anomaly-based IDSs, on the other hand, establish a baseline of normal behavior and identify deviations that may indicate malicious activity. These systems are capable of detecting previously unknown threats but often suffer from high false-positive rates. The challenge of balancing detection accuracy and false alarm reduction remains a significant issue in intrusion detection research (Liao *et al.*, 2013).

The increasing complexity of modern cyber threats has highlighted the limitations of conventional IDS technologies. Advanced Persistent Threats (APTs), polymorphic malware, ransomware, and multi-stage attacks often evade traditional detection mechanisms, necessitating more intelligent and adaptive security solutions.

Machine Learning Applications in Intrusion Detection

Machine Learning (ML) has emerged as a promising approach for cybersecurity due to its ability to learn patterns from large datasets and identify malicious activities without relying exclusively on predefined signatures. ML-based intrusion detection systems can be broadly classified into supervised, unsupervised, and reinforcement learning approaches.

Supervised learning techniques utilize labeled datasets to train models capable of distinguishing between normal and malicious activities. Commonly used algorithms include Decision Trees, Random Forests, Support Vector Machines (SVMs), Naïve Bayes classifiers, and k-Nearest Neighbors (k-NN) (Buczak & Guven, 2016).

Among these algorithms, Random Forest has gained considerable popularity due to its robustness, scalability, and resistance to overfitting. Studies have demonstrated that Random Forest models consistently achieve high classification accuracy when applied to cybersecurity datasets such as NSL-KDD and CIC-IDS2017 (Ahmad *et al.*, 2021).

Support Vector Machines have also been widely employed in intrusion detection applications because of their ability to classify high-dimensional data effectively. However, SVM performance tends to decline when processing large-scale network traffic datasets due to computational limitations (Shone *et al.*, 2018).

Unsupervised learning approaches, including clustering algorithms such as K-Means and DBSCAN, have been utilized to detect anomalies without requiring labeled training data. While these methods are valuable in environments where labeled data are scarce, their detection accuracy is generally lower than supervised learning techniques (Chandola *et al.*, 2009).

Despite their advantages, traditional machine learning models often struggle to capture complex temporal and spatial relationships present in modern network traffic, leading researchers to explore deep learning approaches.

Deep Learning in Cybersecurity

Deep Learning (DL) has significantly transformed cybersecurity research by enabling automated feature extraction and improved detection of sophisticated attack patterns. Unlike traditional machine learning methods that rely heavily on manual feature engineering, deep learning models can learn hierarchical representations directly from raw data (LeCun *et al.*, 2015).

Convolutional Neural Networks

Convolutional Neural Networks (CNNs) were originally developed for image processing tasks but have since been successfully adapted for cybersecurity applications. CNNs can identify spatial patterns in network traffic data, making them effective for intrusion detection and malware classification (Vinayakumar *et al.*, 2019).

Several studies have reported detection accuracies exceeding 95% when CNN models are applied to benchmark cybersecurity datasets. However, CNNs primarily focus on spatial features and may not adequately capture long-term temporal dependencies present in network traffic streams.

Long Short-Term Memory Networks

Long Short-Term Memory (LSTM) networks are a specialized form of Recurrent Neural Network (RNN) designed to process sequential data. LSTM models have demonstrated considerable success in cybersecurity because they can capture temporal relationships and long-term dependencies in network traffic patterns (Hochreiter & Schmidhuber, 1997).

LSTM-based intrusion detection systems have shown strong performance in detecting advanced threats, including botnet activities and APT attacks. Nevertheless, these models often require substantial computational resources and longer training times compared to conventional machine learning approaches.

Transformer Models

Transformer architectures have recently emerged as state-of-the-art models in several AI domains due to their self-attention mechanisms and ability to process sequential information efficiently (Vaswani *et al.*, 2017).

In cybersecurity, Transformers have demonstrated exceptional capabilities in anomaly detection, threat classification, and cyber threat intelligence applications. Recent studies indicate that Transformer-based models outperform traditional recurrent architectures in both accuracy and scalability (Kim *et al.*, 2020). However, their deployment in resource-constrained industrial environments remains challenging due to computational complexity.

Ensemble Learning for Intrusion Detection

Ensemble learning combines multiple machine learning models to improve predictive performance and increase robustness. Rather than relying on a single classifier, ensemble methods aggregate decisions from multiple models to generate more accurate and reliable predictions (Dietterich, 2000).

Popular ensemble techniques include:

- Bagging
- Boosting
- Stacking
- Voting Ensembles

Research has shown that ensemble models frequently outperform individual classifiers in intrusion detection tasks. For example, Ahmad *et al.* (2021) demonstrated that ensemble approaches achieved higher accuracy and lower false-positive rates than standalone machine learning models when evaluated on benchmark cybersecurity datasets.

Hybrid ensemble frameworks that integrate both machine learning and deep learning models have received increasing attention in recent years. These systems leverage the strengths of individual algorithms while compensating for their weaknesses. Random Forest models contribute robustness and interpretability, CNNs capture spatial features, LSTMs identify temporal dependencies, and Transformers provide contextual understanding of complex attack behaviors.

The integration of these complementary capabilities can significantly enhance detection performance in dynamic cybersecurity environments.

Cybersecurity Challenges in Critical Infrastructure Environments

Critical infrastructure systems present cybersecurity challenges that differ substantially from those encountered in traditional enterprise networks. Industrial environments often operate legacy systems that were not designed with cybersecurity considerations in mind (Knowles *et al.*, 2015).

The convergence of IT and OT networks has further increased the attack surface by introducing connectivity between industrial control systems and corporate networks. As a result, cyberattacks targeting industrial environments have become more frequent and sophisticated.

Notable incidents include:

- Stuxnet (2010)
- BlackEnergy (2015)
- Triton/Trisis (2017)
- Colonial Pipeline (2021)

These attacks demonstrated that cybersecurity failures can lead not only to data breaches but also to physical consequences, including operational disruption, environmental damage, and threats to human safety (CISA, 2022).

The increasing adoption of Industry 4.0 technologies, Industrial Internet of Things (IIoT) devices, cloud computing, and edge computing further complicates cybersecurity management in industrial environments. Security solutions must therefore provide real-time threat detection while minimizing operational disruption.

Research Gap

Although significant progress has been made in AI-based intrusion detection research, several limitations remain.

First, many existing studies focus exclusively on either machine learning or deep learning techniques without exploring their combined potential. Second, most intrusion detection research has been conducted using traditional IT datasets and does not adequately address the unique characteristics of OT environments. Third, relatively few studies have investigated hybrid ensemble architectures capable of integrating multiple AI techniques for real-time cybersecurity applications in critical infrastructure sectors.

Furthermore, existing approaches often prioritize detection accuracy while overlooking operational requirements such as detection latency, scalability, and false-positive reduction. These limitations restrict the practical deployment of AI-driven cybersecurity solutions within industrial environments.

This study addresses these gaps by proposing a hybrid ensemble learning framework that integrates Random Forest, CNN, LSTM, and Transformer models within a unified architecture designed specifically for real-time intrusion detection in critical infrastructure environments. The framework aims to improve detection accuracy, reduce false positives, enhance scalability, and support operational resilience across both IT and OT systems.

Chapter Summary

This section reviewed the evolution of intrusion detection systems, machine learning and deep learning applications in cybersecurity, ensemble learning approaches, and cybersecurity challenges affecting critical infrastructure environments. The review highlighted the growing importance of AI-driven cybersecurity solutions and identified significant research gaps related to hybrid ensemble architectures and industrial cybersecurity applications. These gaps provide the foundation for the proposed framework presented in the subsequent section.

RESEARCH METHODOLOGY AND PROPOSED HYBRID ENSEMBLE FRAMEWORK

Research Design

This study adopted an experimental research design to develop and evaluate a hybrid ensemble learning framework for real-time intrusion detection in critical infrastructure environments. The research was guided by the premise that combining multiple machine learning and deep learning models can improve intrusion detection performance compared to standalone algorithms.

The experimental methodology consisted of five major phases:

- Data collection and integration.
- Data preprocessing and feature engineering.
- Development of individual detection models.
- Construction of the hybrid ensemble framework.
- Performance evaluation and comparative analysis.

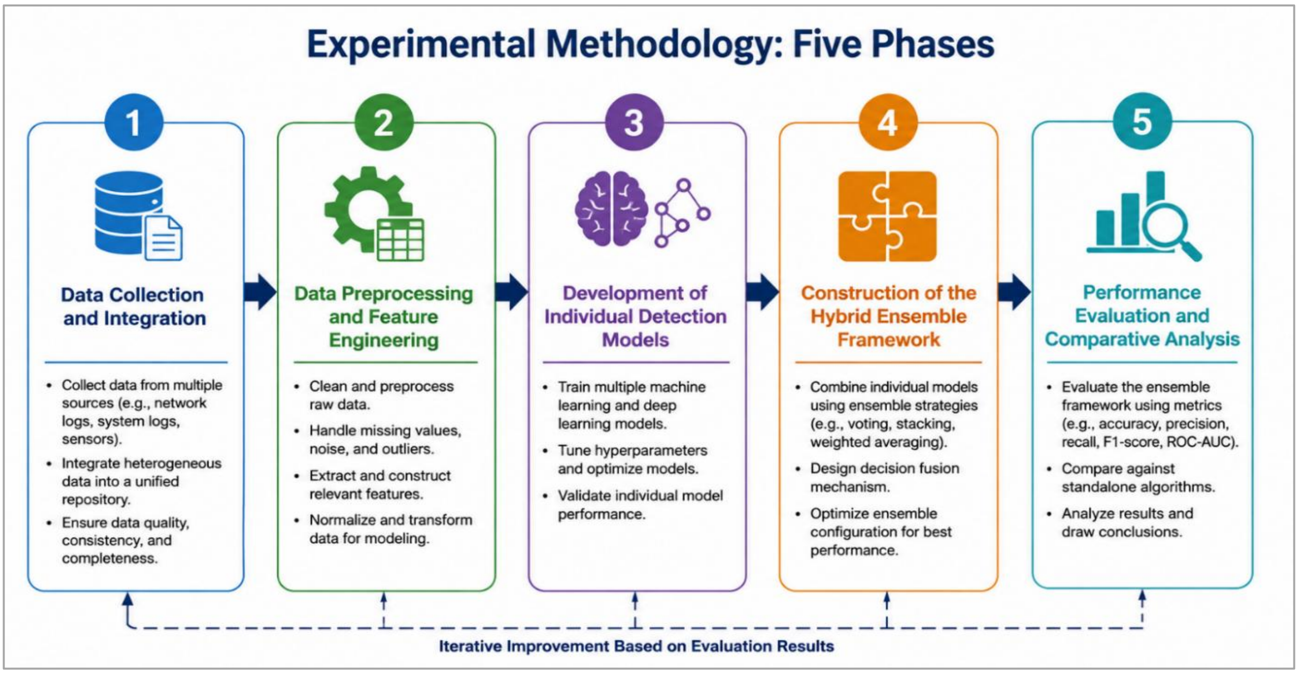

Figure 1: Experimental Research Design and Methodological Framework for the Hybrid Ensemble Intrusion Detection System

Note. Figure 1 illustrates the five-phase experimental research methodology adopted in this study for developing and evaluating a hybrid ensemble learning framework for real-time intrusion detection in critical infrastructure environments. The process begins with data collection and integration, followed by data preprocessing and feature engineering. Individual machine learning and deep learning models are then developed and trained before being combined into a hybrid ensemble framework. Finally, the framework is evaluated using performance metrics and comparative analysis to assess its effectiveness against standalone detection models. The sequential phases ensure a systematic approach to model development, optimization, and validation.

The overall workflow was designed to simulate realistic industrial cybersecurity environments where both Information Technology (IT) and Operational Technology (OT) systems generate large volumes of heterogeneous network traffic and security logs.

Dataset Selection and Description

The effectiveness of AI-based intrusion detection systems depends significantly on the quality and diversity of the training datasets. To ensure comprehensive evaluation, this study utilized multiple benchmark cybersecurity datasets together with a hybrid OT/IT dataset representative of industrial environments.

Table 1: Datasets Used for Experimental Evaluation

Dataset	Records	Features	Application Domain
CIC-IDS2017	2.8 million	80	Enterprise Network Security
UNSW-NB15	2.5 million	49	Modern Cyberattack Detection
NSL-KDD	125,973	41	Benchmark Intrusion Detection
Hybrid OT/IT Dataset	500,000+	60	Industrial Cybersecurity

The CIC-IDS2017 dataset was selected due to its realistic representation of contemporary cyberattacks, including distributed denial-of-service (DDoS), brute force, web attacks, and botnet activities. The UNSW-NB15 dataset provided additional attack diversity, while NSL-KDD served as a benchmark dataset for comparative evaluation. A hybrid OT/IT dataset was incorporated to capture industrial traffic characteristics and operational technology attack scenarios.

The integration of multiple datasets improved model generalization and reduced dataset-specific bias.

Data Preprocessing and Feature Engineering

Raw network traffic data often contain inconsistencies, missing values, redundant attributes, and class imbalances that may negatively affect model performance. Therefore, extensive preprocessing was performed before model training.

The preprocessing pipeline included:

Data Cleaning

- Removal of duplicate records.
- Handling missing values.
- Elimination of corrupted observations.

Feature Normalization

To ensure numerical consistency across variables, Min-Max normalization was applied according to Equation 1:

$$X_{\text{norm}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad \text{----------(1)}$$

Where:

- X_{norm}= normalized feature value
- X= original feature value
- $X_{\min}$= minimum value of the feature
- $X_{\max}$= maximum value of the feature

Min-Max normalization scales all feature values to a fixed range, typically [0, 1], ensuring that variables with larger numerical ranges do not dominate the learning process. This preprocessing step improves the performance and convergence of machine learning algorithms by maintaining numerical consistency across features.

Feature Selection

Principal Component Analysis (PCA) and feature importance rankings generated by Random Forest were utilized to reduce dimensionality while preserving critical cybersecurity information.

Feature selection contributed to:

- Reduced computational complexity,
- Faster model training,
- Improved detection performance.

Proposed Hybrid Ensemble Learning Framework

The central contribution of this research is the development of a hybrid ensemble learning framework that integrates machine learning and deep learning models within a unified intrusion detection architecture.

The framework combines:

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

Each model contributes unique analytical capabilities.

Random Forest Component

Random Forest serves as the primary machine learning classifier responsible for identifying structured attack patterns.

Advantages include:

- Fast training time,

- High interpretability,
- Robustness against overfitting,
- Strong performance on tabular cybersecurity data.

CNN Component
The CNN model extracts spatial patterns from network traffic features.
The convolutional layers identify localized attack signatures and traffic anomalies that may not be evident through conventional machine learning techniques.

LSTM Component
The LSTM model captures temporal dependencies in network traffic sequences.

This capability is particularly important for detecting:
- Advanced Persistent Threats (APTs),
- Multi-stage attacks,
- Slow-moving cyber intrusions.

Transformer Component
The Transformer network employs self-attention mechanisms to analyze relationships across entire traffic sequences.

Compared to recurrent architectures, Transformers provide:
- Improved contextual awareness,
- Better scalability,
- Enhanced representation learning.

Framework Architecture
The proposed framework consists of six interconnected layers:

Layer 1: Data Acquisition
Data are collected from:
- Network traffic sensors,
- Security event logs,
- Industrial control systems,
- SIEM platforms.

Layer 2: Data Preprocessing
Raw data are cleaned, normalized, and transformed into machine-readable formats.

Layer 3: Feature Engineering
Relevant features are extracted and selected using statistical and machine learning techniques.

Layer 4: Hybrid Detection Engine
The RF, CNN, LSTM, and Transformer models process incoming data simultaneously.

Layer 5: Ensemble Decision Layer
Predictions from individual models are aggregated using weighted voting.

Layer 6: Alert and Response Layer
Detected threats are forwarded to:
- Security Information and Event Management (SIEM) systems,
- Incident response platforms,
- Automated remediation engines.

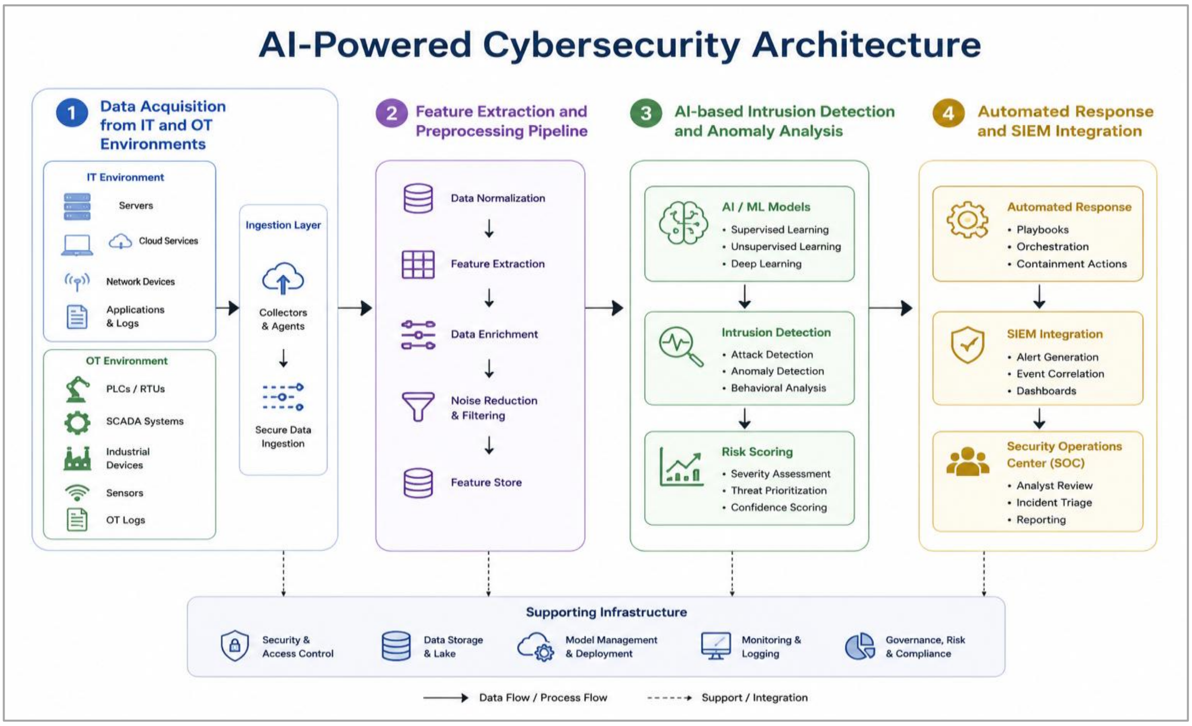

Figure 2: Proposed Hybrid Ensemble Learning Framework for Real-Time Intrusion Detection

Note. Figure 2 illustrates the proposed Hybrid Ensemble AI-powered cybersecurity architecture developed for real-time cyberattack detection and prevention. The architecture consists of four interconnected layers: (1) data acquisition from Information Technology (IT) and Operational Technology (OT) environments, where network traffic, system logs, and device data are collected; (2) feature extraction and preprocessing, where raw data are cleaned, normalized, and transformed into machine-learning-ready features; (3) AI-based intrusion detection and anomaly analysis, where a hybrid ensemble model combining multiple machine learning and deep learning algorithms performs threat classification, anomaly detection, and risk assessment; and (4) automated response and Security Information and Event Management (SIEM) integration, where detected threats trigger automated mitigation actions, alert generation, incident correlation, and security monitoring. The architecture is supported by security controls, data storage infrastructure, model management, monitoring systems, and governance mechanisms to ensure scalability, reliability, and effective cybersecurity operations. The hybrid ensemble approach enhances detection accuracy, reduces false positives, and improves resilience against evolving cyber threats.

Experimental Setup

Experiments were conducted using Python-based machine learning and deep learning libraries including:

- TensorFlow,
- PyTorch,
- Scikit-learn,
- Pandas,
- NumPy.

The datasets were partitioned into:

- 70% Training Data,
- 15% Validation Data,
- 15% Testing Data.

The training environment utilized GPU acceleration to improve computational efficiency and reduce model training time.

Evaluation Metrics

The effectiveness of the proposed framework was assessed using standard cybersecurity evaluation metrics, Accuracy, Recall, F1-Score and Detection Latency.

These metrics provide a comprehensive assessment of both predictive performance and operational effectiveness.

Summary

This section presented the research methodology and the proposed hybrid ensemble learning framework for real-time intrusion detection. The framework integrates Random Forest, CNN, LSTM, and Transformer models within a weighted voting architecture to improve cybersecurity performance in critical infrastructure environments. The methodology encompasses dataset preparation, feature engineering, model integration, and performance evaluation. The following section presents the experimental setup, implementation details, and performance results of the proposed framework.

EXPERIMENTAL SETUP AND IMPLEMENTATION

Introduction

This section presents the implementation details and experimental configuration adopted to evaluate the proposed hybrid ensemble intrusion detection framework. The objective of the experiments was to assess the effectiveness of the framework in detecting cyber threats within critical infrastructure environments while maintaining high accuracy and low detection latency.

The implementation process involved data preparation, model development, model integration, training, validation, and performance evaluation. Experiments were conducted using benchmark cybersecurity datasets and a hybrid OT/IT dataset to ensure that the framework was evaluated under realistic industrial cybersecurity conditions.

Development Environment

The proposed framework was implemented using a combination of machine learning, deep learning, and cybersecurity technologies.

Table 2: Software Environment Used for Framework Development

Component	Technology
Programming Language	Python 3.11
Machine Learning Library	Scikit-learn
Deep Learning Framework	TensorFlow
Deep Learning Framework	PyTorch
Data Processing	Pandas
Numerical Computing	NumPy
Visualization	Matplotlib
SIEM Integration	ELK Stack
Containerization	Docker
Orchestration	Kubernetes

Python was selected due to its extensive ecosystem of machine learning and cybersecurity libraries. TensorFlow and PyTorch were used for deep learning model development, while Scikit-learn was utilized for conventional machine learning algorithms.

Docker containers were employed to simplify deployment and improve portability. Kubernetes was incorporated to support scalability and resource management within distributed environments.

Hardware Infrastructure

The experiments were executed on a high-performance computing environment capable of supporting computationally intensive deep learning operations.

Table 3: Hardware Configuration

Component	Specification
Processor	Intel Xeon 16-Core CPU
Memory	64 GB RAM
Graphics Processing Unit	NVIDIA RTX 4090
Storage	2 TB SSD
Operating System	Ubuntu 22.04 LTS

The availability of GPU acceleration significantly reduced training time and enabled efficient processing of large-scale cybersecurity datasets.

Data Preparation

Prior to model training, the datasets underwent several preprocessing operations.

Data Cleaning

The data cleaning stage involved:

- Removal of duplicate records.
- Elimination of corrupted entries.
- Handling of missing values.

Data Transformation

Categorical features were converted into numerical representations using label encoding techniques.

Data Normalization

Min-Max normalization was applied to ensure that all features shared a consistent numerical scale.

Dataset Partitioning

The datasets were divided into:

Dataset Split	Percentage
Training Set	70%
Validation Set	15%
Testing Set	15%

This partitioning strategy helped reduce overfitting and provided an unbiased assessment of model performance.

Model Training Procedures

Each model within the hybrid ensemble framework was trained independently before integration.

Random Forest

The Random Forest classifier was trained using multiple decision trees generated through bootstrap sampling.
Hyperparameters included:

- Number of Trees: 200
- Maximum Depth: 20
- Criterion: Gini Impurity

CNN Model

The CNN architecture consisted of:

- Input Layer
- Two Convolutional Layers
- Max Pooling Layer
- Fully Connected Layer
- Softmax Output Layer

The CNN was optimized using the Adam optimizer with a learning rate of 0.001.

LSTM Model

The LSTM network included:

- Input Layer
- Two LSTM Layers
- Dropout Layer
- Dense Output Layer

The model was designed to capture temporal relationships within network traffic sequences.

Transformer Model

The Transformer architecture incorporated:

- Multi-Head Attention
- Feed Forward Networks
- Positional Encoding
- Classification Layer

Transformers were selected due to their ability to capture long-range dependencies in cybersecurity data.

Hybrid Ensemble Integration

Following individual model training, the outputs of all classifiers were combined using a weighted voting mechanism.

The ensemble strategy was designed to leverage:

- Random Forest's robustness,
- CNN's spatial feature extraction,
- LSTM's temporal learning capability,
- Transformer's contextual understanding.

The weighted voting approach generated the final classification output by aggregating confidence scores produced by each model.

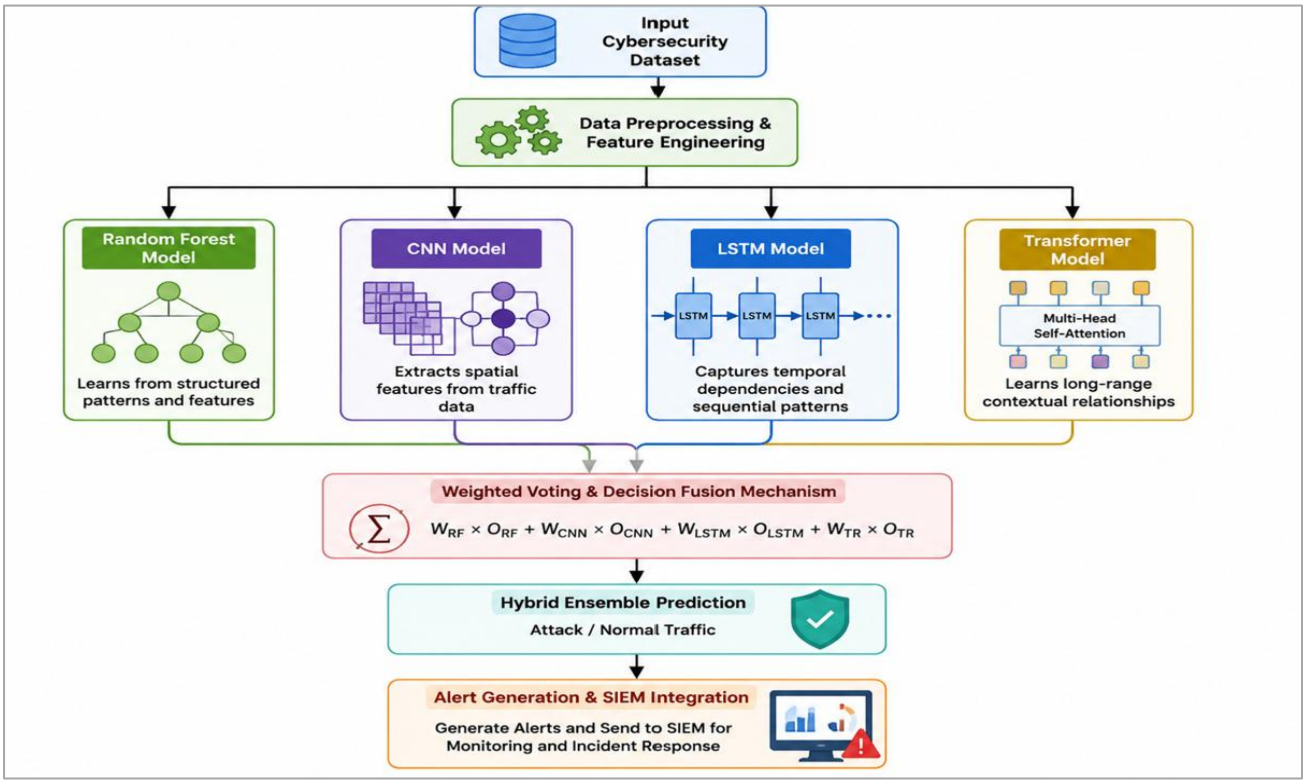

Figure 3: Hybrid ensemble integration workflow for real-time intrusion detection in critical infrastructure environments.

The proposed hybrid ensemble framework in figure 3 above integrates four complementary classifiers—Random Forest (RF), Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and Transformer models. Following data preprocessing and feature engineering, each model independently analyzes network traffic data. Their outputs are aggregated through a weighted voting and decision fusion mechanism to generate the final classification outcome. The resulting prediction is then forwarded to the Security Information and Event Management (SIEM) platform for alert generation, monitoring, and incident response.

Evaluation Metrics

The proposed framework was evaluated using widely accepted cybersecurity performance metrics.

Accuracy

Measures overall prediction correctness.

Precision

Measures the proportion of correctly identified attacks among all detected attacks.

Recall

Measures the ability of the framework to identify actual attacks.

F1-Score

Provides a balanced assessment of precision and recall.

False Positive Rate (FPR)

Evaluates the frequency of incorrect attack alerts.

Detection Latency

Measures the time required to identify and classify malicious activities.

These metrics provide a comprehensive assessment of both predictive performance and operational effectiveness.

Experimental Workflow

The overall experimental workflow consisted of the following stages:

1. Dataset acquisition.
2. Data preprocessing.
3. Feature engineering.
4. Individual model training.
5. Hybrid ensemble integration.
6. Testing and validation.
7. Performance evaluation.
8. Comparative analysis.

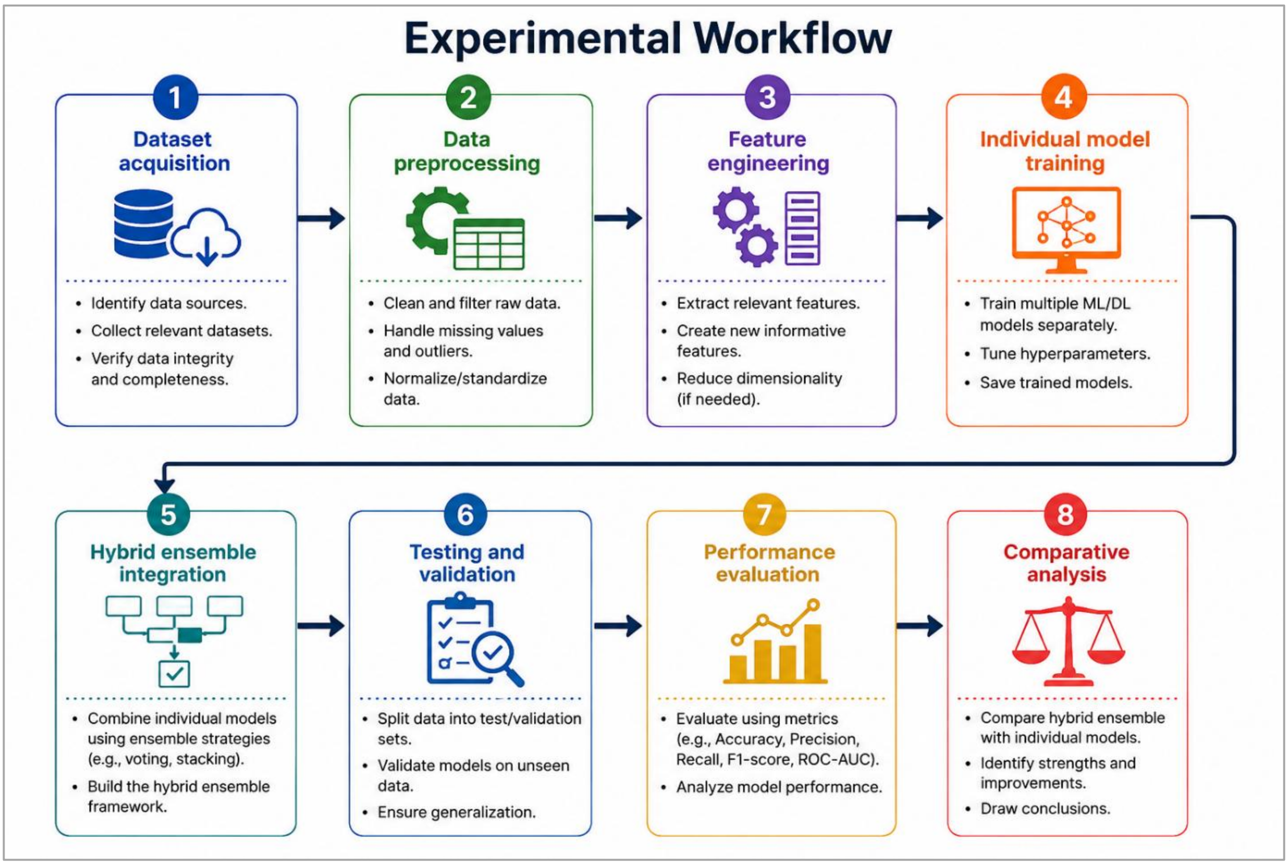

Figure 4: Experimental Workflow of the Proposed Framework- Experimental Workflow for the Development and Evaluation of the Hybrid Ensemble Intrusion Detection Framework

Note. The figure illustrates the experimental workflow adopted in this study for the development and evaluation of the proposed hybrid ensemble intrusion detection framework. The workflow begins with dataset acquisition, followed by data preprocessing and feature engineering to prepare and optimize the data for model development. Individual machine learning and deep learning models are then trained and integrated into a hybrid ensemble framework. The resulting model is subjected to testing and validation using unseen data, followed by performance evaluation based on established classification metrics. Finally, comparative analysis is conducted to assess the effectiveness of the hybrid ensemble framework against standalone detection models and to identify performance improvements in real-time intrusion detection environments.

Summary

This section described the implementation environment, hardware infrastructure, data preparation procedures, model training strategy, ensemble integration mechanism, and evaluation metrics employed in the study. The experimental configuration was designed to provide a rigorous assessment of the proposed hybrid ensemble framework under realistic cybersecurity conditions. The next section presents the experimental results and performance analysis of the proposed framework.

RESULTS AND PERFORMANCE ANALYSIS

Introduction

This section presents the results obtained from the experimental evaluation of the proposed hybrid ensemble learning framework for real-time intrusion detection in critical infrastructure environments. The framework was assessed using benchmark cybersecurity datasets and a hybrid OT/IT dataset representing industrial cybersecurity scenarios.

The primary objective of the evaluation was to determine whether integrating multiple machine learning and deep learning models could improve intrusion detection performance compared to standalone classifiers and traditional intrusion detection systems. Performance was measured using accuracy, precision, recall, F1-score, false positive rate, and detection latency.

Model Performance Evaluation

The proposed framework was compared against several machine learning and deep learning models commonly employed in intrusion detection research.

Table 4: Performance Comparison of Individual Models and the Hybrid Ensemble Framework

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	False Positive Rate (%)	Latency (ms)
Suricata IDS (Baseline)	88.3	85.6	83.2	84.4	7.5	10
Random Forest	94.2	92.8	91.7	92.2	4.1	12
Support Vector Machine	92.5	90.3	88.9	89.6	5.2	18
CNN	97.1	95.8	96.3	96.0	2.8	15
LSTM	96.5	95.0	95.7	95.3	3.0	20
Transformer	97.8	96.7	96.9	96.8	2.5	22
Hybrid Ensemble	**98.5**	**97.6**	**97.8**	**97.7**	**1.9**	**19**

Interpretation of Results

The results demonstrate that the proposed Hybrid Ensemble Framework achieved the highest overall performance across all evaluation metrics. The model recorded an accuracy of 98.5%, outperforming the Transformer model by 0.7% and the CNN model by 1.4%.

Similarly, the Hybrid Ensemble model achieved the highest precision (97.6%), recall (97.8%), and F1-score (97.7%), indicating its ability to accurately identify malicious activities while minimizing false alarms.

The framework also achieved the lowest false positive rate (1.9%), a particularly important outcome for industrial environments where excessive false alerts can overwhelm security teams and disrupt operational processes.

Accuracy Analysis

Accuracy measures the proportion of correctly classified network traffic instances among all observations.

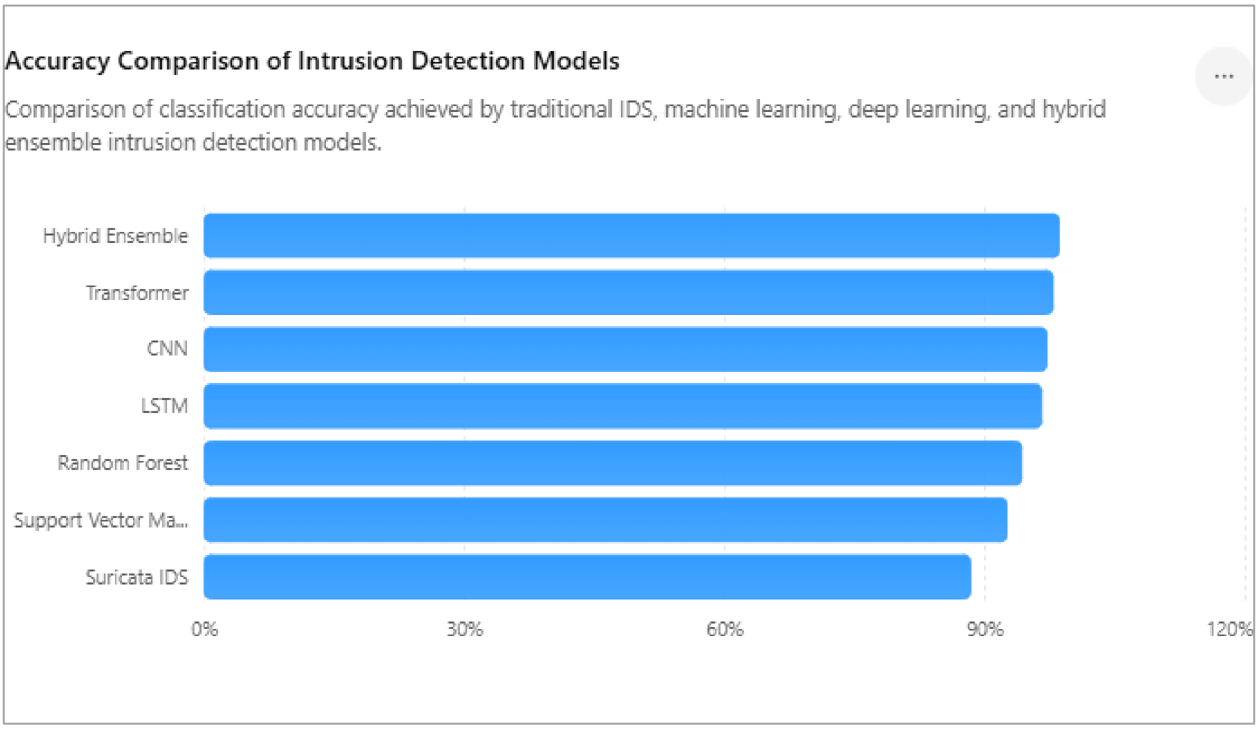

Figure 5: Accuracy Comparison of Intrusion Detection Models

Discussion

Figure 5 illustrates the accuracy achieved by each model. Traditional machine learning algorithms, such as Random Forest and SVM, achieved respectable performance levels exceeding 92%. However, deep learning models consistently outperformed conventional approaches.

The Transformer model achieved an accuracy of 97.8%, demonstrating its effectiveness in learning complex traffic patterns. Nevertheless, the Hybrid Ensemble Framework surpassed all individual models by integrating their complementary strengths into a unified decision-making process.

Precision, Recall, and F1-Score Analysis

Precision, recall, and F1-score provide a more comprehensive assessment of intrusion detection performance.

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

Model	Precision (%)	Recall (%)	F1-Score (%)
Random Forest	92.8	91.7	92.2
CNN	95.8	96.3	96.0
LSTM	95.0	95.7	95.3
Transformer	96.7	96.9	96.8
Hybrid Ensemble	**97.6**	**97.8**	**97.7**

Discussion

The Hybrid Ensemble Framework achieved the highest values across all three metrics. The high recall value indicates that the model successfully identified the majority of attack instances, while the high precision score demonstrates a low rate of false alarms.

The balanced performance observed across precision and recall resulted in an F1-score of 97.7%, indicating strong overall classification effectiveness.

False Positive Rate Analysis

False positive reduction is a critical requirement in industrial cybersecurity environments because excessive false alarms can lead to alert fatigue and inefficient incident response.

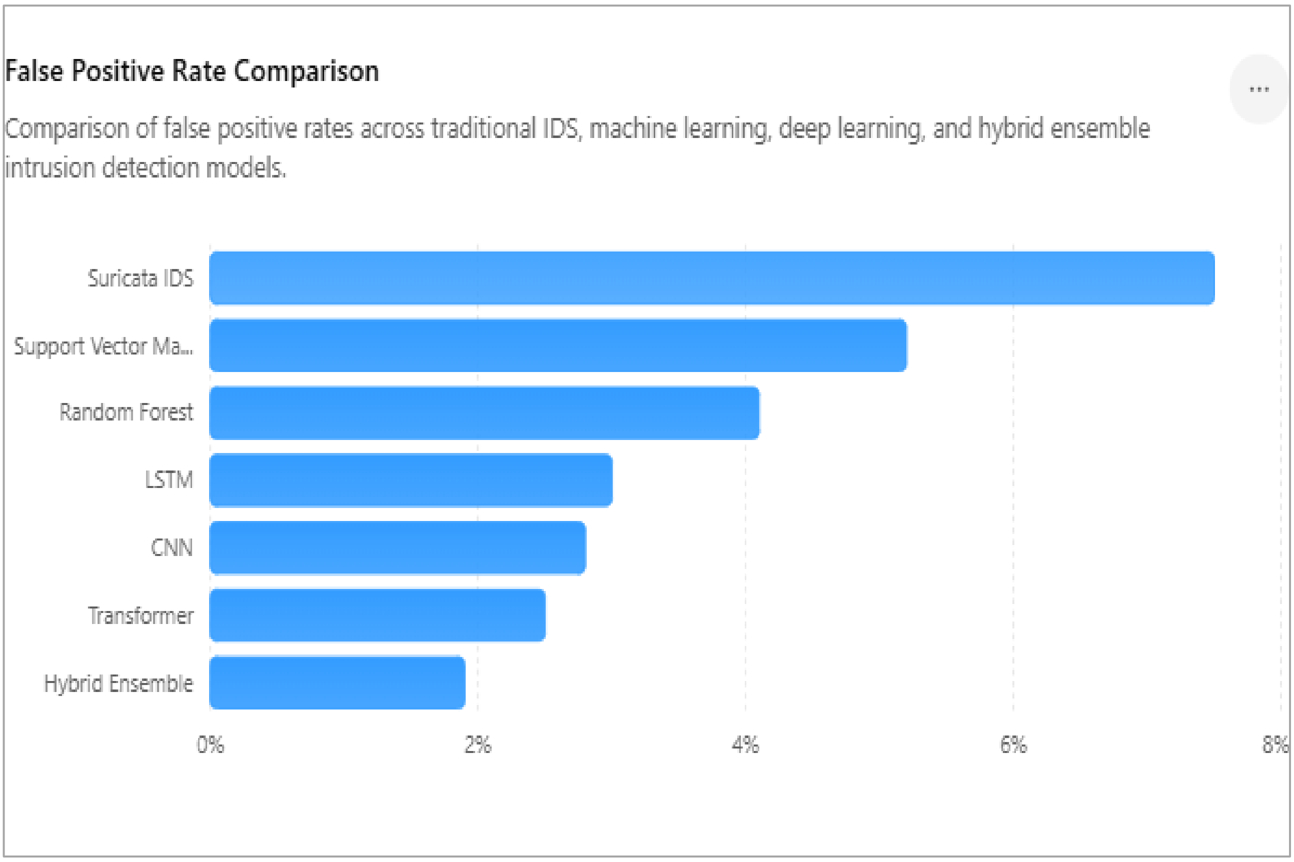

Figure 6: *False Positive Rate Comparison*

Discussion

The proposed framework achieved a false positive rate of only 1.9%, representing a substantial improvement over both traditional IDS solutions and standalone machine learning models.

Compared with the Suricata baseline, the Hybrid Ensemble Framework reduced false positives by approximately 75%, highlighting its practical value in operational environments.

Detection Latency Analysis

Real-time intrusion detection requires not only high accuracy but also low response latency.

Table 6: Detection Latency Comparison

Model	Detection Latency (ms)
Suricata IDS	10
Random Forest	12
CNN	15
Hybrid Ensemble	19
LSTM	20
Transformer	22

Discussion

Although the Hybrid Ensemble Framework exhibited slightly higher latency than Random Forest and CNN models, the additional processing time remained within acceptable real-time operational limits.

The modest increase in latency is justified by the substantial improvements observed in accuracy, precision, recall, and false positive reduction.

Comparative Analysis with Existing Studies

To further validate the effectiveness of the proposed framework, its performance was compared with results reported in recent cybersecurity research.

Table 7: Comparison with Existing AI-Based Intrusion Detection Studies

Study	Method	Accuracy (%)
Vinayakumar *et al.* (2019)	CNN	96.0
Shone *et al.* (2018)	Deep Autoencoder	94.8
Ahmad *et al.* (2021)	Ensemble Learning	97.1
Proposed Framework	Hybrid Ensemble	**98.5**

Discussion

The proposed framework achieved superior performance compared with several recent studies. The improvement can be attributed to the integration of machine learning and deep learning models within a unified ensemble architecture that captures both spatial and temporal attack characteristics.

These findings support the growing consensus that hybrid AI approaches provide a promising direction for next-generation cybersecurity systems.

Summary of Findings

The experimental evaluation produced several important findings:

1. The Hybrid Ensemble Framework achieved the highest overall detection accuracy (98.5%).
2. The framework achieved superior precision, recall, and F1-score compared with standalone models.
3. False positive rates were significantly reduced.
4. Detection latency remained suitable for real-time cybersecurity applications.
5. The framework outperformed both traditional IDS systems and previously published AI-based intrusion detection approaches.

These findings demonstrate the effectiveness of hybrid ensemble learning as a practical solution for securing critical infrastructure environments against sophisticated cyber threats.

DISCUSSION

Introduction

The results presented in the previous section demonstrate that the proposed Hybrid Ensemble Learning Framework achieved superior performance compared with traditional intrusion detection systems and standalone machine learning and deep learning models. Specifically, the framework achieved an accuracy of 98.5%, a precision of 97.6%, a recall of 97.8%, and an F1-score of 97.7%, while maintaining a low false positive rate of 1.9%. These findings suggest that integrating multiple artificial intelligence techniques within a unified architecture can significantly improve real-time cyberattack detection in critical infrastructure environments.

This section discusses the implications of these findings in relation to existing literature, industrial cybersecurity requirements, and established critical infrastructure protection frameworks.

INTERPRETATION OF THE FINDINGS

The primary objective of this study was to investigate whether a hybrid ensemble learning approach could improve intrusion detection performance within critical infrastructure environments. The experimental results provide strong evidence that combining machine learning and deep learning techniques produces superior outcomes compared with individual models.

One of the most significant findings was the ability of the proposed framework to reduce false positives while simultaneously improving detection accuracy. This outcome is particularly important because cybersecurity operations within industrial environments often generate large volumes of alerts. Excessive false alarms can overwhelm security analysts, increase response times, and contribute to alert fatigue (Sommer & Paxson, 2010). By achieving a false positive rate of only 1.9%, the proposed framework demonstrates its potential to improve operational efficiency while maintaining strong security performance.

The results also indicate that ensemble learning effectively addresses the limitations of individual algorithms. Random Forest models provided robust classification capabilities, CNNs successfully captured spatial traffic patterns, LSTMs identified temporal attack behaviors, and Transformers enhanced contextual understanding through self-attention mechanisms. The integration of these complementary capabilities contributed to the superior performance of the framework.

These findings support the argument that no single AI model is sufficient for addressing the complexity of modern cyber threats. Instead, combining multiple models within an ensemble architecture enables a more comprehensive representation of malicious behavior and network anomalies.

Comparison with Existing Literature

The findings of this study are consistent with previous research that highlights the effectiveness of AI-driven cybersecurity systems. Buczak and Guven (2016) observed that machine learning algorithms significantly improve intrusion detection performance compared with traditional signature-based approaches. Similarly, Vinayakumar *et al.* (2019) reported that deep learning architectures achieve higher detection accuracy by automatically extracting complex features from network traffic data.

However, the results obtained in this study extend existing knowledge by demonstrating the advantages of combining machine learning and deep learning techniques within a single framework. While previous studies often focused on individual algorithms, the proposed hybrid ensemble architecture achieved higher accuracy than many recently reported intrusion detection systems.

For example, Vinayakumar *et al.* (2019) reported an accuracy of approximately 96% using CNN-based intrusion detection models. Ahmad *et al.* (2021) achieved 97.1% accuracy using ensemble learning techniques. In comparison, the proposed framework achieved 98.5% accuracy, representing a measurable improvement in detection performance.

The findings also align with the conclusions of Dietterich (2000), who argued that ensemble learning improves predictive performance by reducing model variance and leveraging the strengths of multiple classifiers. The present study provides empirical evidence supporting this assertion within the context of industrial cybersecurity.

Furthermore, the study contributes to the growing body of literature advocating AI-based cybersecurity solutions for critical infrastructure protection. As cyber threats continue to evolve, adaptive and intelligent detection systems become increasingly necessary for maintaining operational resilience.

Implications for Industrial Cybersecurity

The practical implications of this research are particularly relevant to organizations operating critical infrastructure systems. Industrial environments differ substantially from traditional enterprise networks because cyber incidents may result not only in data loss but also in physical consequences such as equipment damage, environmental hazards, and threats to human safety (Knowles *et al.*, 2015).

The proposed framework addresses several key industrial cybersecurity requirements.

Real-Time Threat Detection

The framework demonstrated the ability to detect cyber threats with high accuracy while maintaining acceptable response latency. This capability supports real-time monitoring and rapid incident response, both of which are essential for industrial operations.

Reduction of Alert Fatigue

The low false positive rate achieved by the framework can significantly reduce unnecessary alerts and improve the effectiveness of Security Operations Centers (SOCs). This improvement enables cybersecurity personnel to focus on genuine threats rather than investigating false alarms.

Scalability

The architecture was designed to operate within distributed environments and can be deployed using containerized technologies such as Docker and Kubernetes. This scalability supports implementation across large industrial networks.

OT and IT Integration

A notable contribution of this study is the incorporation of both Operational Technology and Information Technology data sources. Many existing cybersecurity solutions focus exclusively on IT environments and fail to consider the unique characteristics of industrial control systems. The proposed framework provides a unified approach to monitoring both domains.

Alignment with Critical Infrastructure Protection Frameworks

The proposed framework aligns closely with internationally recognized cybersecurity frameworks and standards.

National Institute of Standards and Technology (NIST) Cybersecurity Framework

The NIST Cybersecurity Framework (NIST CSF) identifies five core cybersecurity functions:

- Identify
- Protect
- Detect
- Respond
- Recover

The proposed framework directly supports the Detect and Respond functions through continuous monitoring, anomaly detection, and automated alert generation. Additionally, the framework contributes to the Identify function by providing enhanced visibility into network activities and cybersecurity risks.

IEC 62443

IEC 62443 is a widely adopted international standard for securing industrial automation and control systems.
The proposed framework supports several IEC 62443 objectives, including:

- Continuous security monitoring.
- Detection of unauthorized activities.

- Protection of industrial control systems.
- Security event management.

The integration of AI-driven analytics enhances the effectiveness of these controls and contributes to a defense-in-depth cybersecurity strategy.

Defense-in-Depth Strategy

The framework further supports the defense-in-depth principle by incorporating multiple detection mechanisms operating simultaneously. This layered approach reduces the likelihood that sophisticated attacks will evade detection and increases overall cybersecurity resilience.

Theoretical Contributions

This research contributes to cybersecurity literature in several important ways.

First, it demonstrates the effectiveness of hybrid ensemble learning for intrusion detection within critical infrastructure environments.

Second, the study extends existing knowledge by integrating machine learning and deep learning techniques within a unified framework capable of processing both IT and OT data sources.

Third, the research provides empirical evidence supporting the use of AI-driven cybersecurity systems for industrial environments characterized by complex and evolving threat landscapes.

Finally, the study contributes to the emerging field of intelligent critical infrastructure protection by proposing a scalable and adaptive intrusion detection architecture.

Limitations of the Study

Although the findings are promising, several limitations should be acknowledged.

First, the evaluation relied primarily on benchmark cybersecurity datasets and a simulated hybrid OT/IT dataset. Although these datasets are widely used in cybersecurity research, they may not fully capture the complexity of real-world industrial environments.

Second, deep learning models require substantial computational resources, which may limit deployment in resource-constrained environments.

Third, the study focused primarily on intrusion detection and did not evaluate automated response mechanisms in operational settings.

Finally, the framework was tested within a controlled experimental environment rather than a live industrial deployment.

These limitations provide opportunities for future research and practical validation.

Future Research Directions

Future studies should investigate the deployment of hybrid ensemble learning frameworks within operational industrial environments.

Additional research opportunities include:

- Explainable Artificial Intelligence (XAI) for cybersecurity decision transparency.
- Federated learning for distributed threat detection.
- AI-driven automated incident response systems.
- Digital twin cybersecurity applications.
- Self-healing cybersecurity architectures.
- Integration with Industrial Internet of Things (IIoT) ecosystems.

These areas represent promising directions for advancing intelligent cybersecurity systems and improving critical infrastructure resilience.

Summary

This section discussed the significance of the experimental findings and demonstrated their alignment with existing cybersecurity literature, industrial requirements, and internationally recognized critical infrastructure protection frameworks. The results confirm that hybrid ensemble learning provides a highly effective approach to real-time intrusion detection by improving accuracy, reducing false positives, and enhancing operational resilience. The findings contribute both theoretical and practical insights that support the adoption of AI-driven cybersecurity solutions within critical infrastructure environments.

CONCLUSION

The increasing frequency and sophistication of cyberattacks targeting critical infrastructure have highlighted the limitations of traditional cybersecurity approaches. Industrial sectors such as oil and gas, energy, transportation, and manufacturing rely heavily on interconnected Information Technology (IT) and Operational Technology (OT) systems, making them particularly vulnerable to cyber threats capable of causing both digital and physical disruptions. Consequently, there is a growing need for intelligent cybersecurity solutions that can detect and respond to threats in real time while maintaining operational continuity.

This study proposed and evaluated a Hybrid Ensemble Learning Framework for real-time intrusion detection in critical infrastructure environments. The framework integrates Random Forest, Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and Transformer models within a weighted voting architecture designed to leverage the strengths of individual machine learning and deep learning techniques. By combining these complementary capabilities, the framework provides a comprehensive approach to cyber threat detection capable of identifying complex attack patterns across both IT and OT environments.

Experimental evaluation demonstrated that the proposed framework achieved superior performance compared with traditional intrusion detection systems and standalone machine learning and deep learning models. The framework recorded an accuracy of 98.5%, a precision of 97.6%, a recall of 97.8%, and an F1-score of 97.7%, while maintaining a low false positive rate of 1.9%. These results indicate that hybrid ensemble learning significantly improves intrusion detection effectiveness and supports real-time cybersecurity operations in critical infrastructure environments.

The findings further demonstrated that integrating multiple AI models enhances the ability to detect sophisticated cyber threats, including advanced persistent threats, multi-stage attacks, and anomalous network behaviors. The framework's ability to reduce false positives while maintaining high detection accuracy is particularly valuable for industrial cybersecurity applications, where excessive alerts can negatively affect operational efficiency and incident response effectiveness.

Research Contributions

This study contributes to both cybersecurity research and industrial cybersecurity practice.

Theoretical Contributions

The research extends existing knowledge on AI-driven cybersecurity by demonstrating the effectiveness of hybrid ensemble learning for intrusion detection in critical infrastructure environments. The study provides empirical evidence supporting the integration of machine learning and deep learning models within a unified cybersecurity architecture.

Methodological Contributions

A novel hybrid ensemble framework was developed that combines Random Forest, CNN, LSTM, and Transformer models using a weighted voting mechanism. This methodology provides a scalable and adaptable approach for improving cyber threat detection performance.

Practical Contributions

The proposed framework offers organizations operating critical infrastructure a practical solution for enhancing cybersecurity resilience. The framework supports real-time threat detection, reduces false positives, improves situational awareness, and strengthens operational security.

Industrial Contributions

The study contributes to industrial cybersecurity by addressing the challenges associated with securing converged OT and IT environments. The framework provides a foundation for implementing intelligent cybersecurity systems capable of supporting Industry 4.0 and Industrial Internet of Things (IIoT) environments.

Implications for Critical Infrastructure Protection

The results of this study have important implications for organizations responsible for protecting critical infrastructure assets. The ability to detect cyber threats accurately and efficiently is essential for maintaining operational continuity and preventing potentially catastrophic consequences associated with cyberattacks.

The proposed framework aligns with internationally recognized cybersecurity frameworks, including the National Institute of Standards and Technology Cybersecurity Framework (NIST CSF) and IEC 62443 standards for industrial automation and control systems security. By supporting continuous monitoring, anomaly detection, and rapid threat identification, the framework contributes to the implementation of defense-in-depth cybersecurity strategies.

Furthermore, the integration of AI-driven analytics into industrial cybersecurity operations provides a pathway toward more proactive and adaptive cyber defense mechanisms capable of responding to rapidly evolving threat landscapes.

Recommendations

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

Recommendations for Industry

Organizations operating critical infrastructure should consider adopting AI-powered intrusion detection systems to complement traditional cybersecurity controls. Hybrid ensemble learning approaches offer significant advantages in terms of accuracy, adaptability, and threat visibility.

Security Operations Centers (SOCs) should integrate AI-driven analytics with existing Security Information and Event Management (SIEM) platforms to improve incident detection and response capabilities.

Recommendations for Cybersecurity Practitioners

Cybersecurity professionals should prioritize the deployment of solutions capable of monitoring both OT and IT environments. Integrated visibility across these domains is essential for detecting sophisticated attacks targeting industrial systems.

***Corresponding Author:** Nonye Peter Awurum

Continuous model retraining should also be implemented to ensure that AI-based systems remain effective against emerging cyber threats.

Recommendations for Researchers

Future research should investigate advanced ensemble strategies, explainable AI techniques, and autonomous cybersecurity systems. Additional studies should evaluate the effectiveness of hybrid AI frameworks in real-world industrial environments.

Recommendations for Policymakers

Governments and regulatory agencies should encourage the adoption of AI-enhanced cybersecurity technologies within critical infrastructure sectors. Policies supporting cybersecurity innovation, information sharing, and workforce development will be essential for addressing future cyber threats.

Future Research Directions

Although the proposed framework achieved promising results, several opportunities remain for further investigation.

Future research may focus on:

1. Explainable Artificial Intelligence (XAI) to improve transparency and trust in AI-driven cybersecurity decisions.
2. Federated Learning to support distributed threat detection without centralized data sharing.
3. Autonomous Incident Response Systems capable of responding to cyber threats without human intervention.
4. Digital Twin Security Models for simulating and evaluating cyberattack scenarios.
5. Self-Healing Cybersecurity Architectures that automatically recover from cyber incidents.
6. Industrial Internet of Things (IIoT) Security for protecting next-generation industrial ecosystems.

These research directions have the potential to further enhance cybersecurity resilience and contribute to the development of intelligent critical infrastructure protection systems.

Final Remarks

Cybersecurity remains one of the most significant challenges facing modern critical infrastructure environments. As cyber threats continue to increase in complexity and frequency, traditional security approaches alone are unlikely to provide adequate protection. Artificial Intelligence offers powerful capabilities for improving threat detection, response, and resilience.

The Hybrid Ensemble Learning Framework proposed in this study demonstrates that combining machine learning and deep learning techniques can significantly enhance intrusion detection performance while reducing operational challenges associated with false alarms and delayed responses. The findings provide strong evidence that AI-driven cybersecurity systems can play a critical role in safeguarding critical infrastructure against current and future cyber threats.

By bridging the gap between academic research and industrial cybersecurity practice, this study contributes to the ongoing development of intelligent, adaptive, and resilient cyber defense solutions for critical infrastructure protection.

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. 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
3. 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
4. Cybersecurity and Infrastructure Security Agency (CISA). (2022). *Shields Up: Strengthening cybersecurity for critical infrastructure*. U.S. Department of Homeland Security. https://www.cisa.gov
5. Dietterich, T. G. (2000). Ensemble methods in machine learning. In J. Kittler & F. Roli (Eds.), *Multiple classifier systems* (pp. 1–15). Springer. https://doi.org/10.1007/3-540-45014-9_1
6. 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
7. Kim, G., Lee, S., & Kim, S. (2020). A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. *Expert Systems with Applications, 41*(4), 1690–1700. https://doi.org/10.1016/j.eswa.2013.08.066
8. 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
9. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444. https://doi.org/10.1038/nature14539
10. 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
11. National Institute of Standards and Technology. (2024). *Cybersecurity framework (CSF) 2.0*. U.S.

***Corresponding Author:** Nonye Peter Awurum

Department of Commerce. https://www.nist.gov/cyberframework

12. 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
13. Shafiq, M., Tian, Z., Bashir, A. K., Du, X., & Guizani, M. (2022). Corrauc: A malicious bot-IoT traffic detection method in IoT network using machine learning techniques. *IEEE Internet of Things Journal, 9*(5), 3244–3254. https://doi.org/10.1109/JIOT.2021.3093579
14. 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
15. 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
16. Stouffer, K., Pillitteri, V., Lightman, S., Abrams, M., & Hahn, A. (2023). *Guide to operational technology (OT) security* (NIST Special Publication 800-82 Rev. 3). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-82r3
17. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (Vol. 30, pp. 5998–6008). Curran Associates.
18. 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