International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2025: 7.089

AI-Based Anomaly Detection in Air-Gapped Environments

Dr Abhishek Kumar[*1], Meenakshi Saini[*2]

[*1]Department of Computer Science
abhishek@sajaincollege.ac.in[1]
[2]Department of Computer Science
sainimeenakshi244[at]gmail.com[2]

Abstract: ***Air-gapped systems are widely used to secure critical infrastructures such as military networks, nuclear facilities, and government information systems by maintaining strict physical isolation from external networks. Despite this isolation, such systems remain vulnerable to sophisticated cyber threats introduced through removable media, insider activities, or hardware manipulation. Traditional security mechanisms relying on signature-based detection are often ineffective in identifying unknown or stealthy attacks in these environments. This paper proposes an Artificial Intelligence-based anomaly detection framework specifically designed for air-gapped systems. The approach employs a hybrid unsupervised learning architecture combining Autoencoder neural networks and Isolation Forest algorithms to model normal system behaviour and detect deviations that may indicate malicious activity. The system continuously analyzes host-level telemetry including CPU usage, memory access patterns, process execution, file operations, and USB device interactions. Experimental evaluation in a simulated air-gapped environment demonstrates that the proposed system achieves detection accuracy of approximately 96% with a low false-positive rate. The results indicate that behavior-based anomaly detection using hybrid machine learning techniques can provide an effective and autonomous security mechanism for protecting critical air-gapped infrastructures against emerging cyber threats.***

Keywords: Air-Gapped Cybersecurity, Behavior Anomaly Detection, Autoencoder Neural Networks, Isolation Forest, Critical Infrastructure Security, Host Based Intrusion Detection

1. Introduction

Air-gapped systems are widely adopted in highly sensitive environments where security and confidentiality are of paramount importance. These systems operate in complete physical isolation from external networks, preventing direct communication with the internet or other unsecured networks. Such isolation is commonly implemented in critical infrastructures including military command systems, nuclear power facilities, government data centers, and industrial control environments. The fundamental assumption behind air-gapped security is that the absence of network connectivity significantly reduces the risk of remote cyber intrusions.

However, recent developments in cybersecurity research have demonstrated that air-gapped systems are not entirely immune to cyber threats. Advanced attack techniques have shown that malicious actors can infiltrate isolated systems through indirect channels such as infected removable media, insider misuse, compromised firmware, or malicious hardware components. Once inside the system, these threats may remain undetected for extended periods because traditional security mechanisms primarily rely on signature-based detection methods that identify only previously known attack patterns.

The limitations of conventional intrusion detection mechanisms become more significant in air-gapped environments. Since these systems operate without internet connectivity, they cannot regularly receive updated threat signatures or external security intelligence feeds. As a result, conventional security tools may fail to detect novel or sophisticated attacks, including zero-day exploits and advanced persistent threats. This creates the need for intelligent security solutions that can identify abnormal behavior without depending on predefined attack signatures or continuous external updates.

Recent advances in Artificial Intelligence and machine learning have opened new opportunities for improving cybersecurity systems. Machine learning-based anomaly detection techniques can learn normal system behavior from historical data and detect deviations that may indicate malicious activity. Unlike traditional rule-based approaches, behavior-driven detection methods can identify unknown or evolving threats by analyzing patterns in system resource utilization, process activity, and hardware interactions.

The conceptual overview of the proposed AI-based anomaly detection framework for air-gapped systems is illustrated in Figure 1. The figure presents the overall structure of the system, highlighting the interaction between data collection components, feature extraction mechanisms, and the machine learning-based detection engine. The framework continuously monitors host-level system activities such as CPU utilization, memory usage, process execution patterns, file access behavior, and USB device interactions in order to construct a baseline model of normal system operation.

The operational workflow of the proposed system is further explained through the process flow diagram shown in Figure 2. The flowchart illustrates the sequential stages involved in the anomaly detection process, including data collection from system sensors, preprocessing and normalization of behavioral data, feature extraction, model training using machine learning algorithms, and real-time anomaly detection. When abnormal behavior is detected, the system generates alerts that enable administrators to investigate potential security incidents.

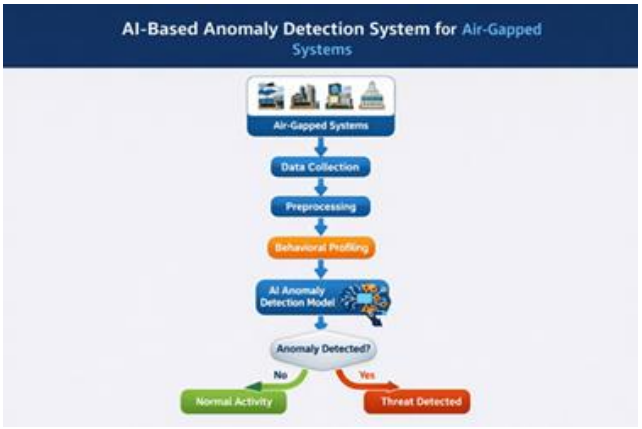

Figure 1: Conceptual overview of the proposed anomaly detection framework

By combining behavioral monitoring with machine learning-based anomaly detection, the proposed framework aims to provide an autonomous and adaptive security solution for protecting air-gapped environments. The primary objective of this research is to design an intelligent detection mechanism capable of identifying unknown threats and suspicious activities without relying on network connectivity or traditional signature-based methods. Through this approach, the study seeks to enhance the security resilience of critical infrastructures that depend on air-gapped systems for safeguarding sensitive operations.

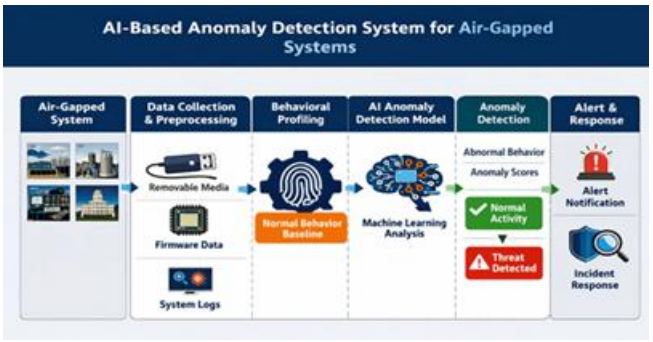

Figure 2: Operation Workflow of Proposed System

2. Research Contributions

This study proposes a behavior-driven anomaly detection framework to strengthen the security of air-gapped systems by analyzing host-level system behavior rather than relying on network traffic or signature-based detection. The key contributions of this research are summarized as follows:

1. Host-Level Behavioral Monitoring: A monitoring approach designed for air-gapped infrastructures that models normal system activity using internal telemetry such as CPU usage, memory access patterns, process execution frequency, file access behavior, and USB device interactions.
2. Hybrid AI-Based Anomaly Detection Model: A novel hybrid architecture combining Autoencoder neural networks and the Isolation Forest algorithm to detect abnormal behavior. The Autoencoder learns normal behavioral patterns through reconstruction learning, while Isolation Forest isolates unusual data instances, improving detection reliability.
3. Fully Offline Security Framework: The proposed system operates independently within an air-gapped environment without requiring internet connectivity, cloud processing, or frequent signature updates, enabling autonomous threat detection.
4. Real-Time Behavioral Threat Detection: Continuous monitoring and anomaly scoring are implemented to identify suspicious activities in real time while reducing false positive alerts.
5. Evaluation Using Air-Gap Specific Attack Scenarios: The framework is validated using realistic threats such as malicious USB insertion, unauthorized process execution, privilege escalation attempts, and slow data exfiltration techniques.
6. Adaptive Detection Capability: The model can adjust to gradual changes in system behavior, ensuring long-term effectiveness without frequent manual configuration.

3. Literature Review

Anomaly detection has become a critical component of modern cybersecurity systems due to the increasing sophistication of cyber-attacks. Traditional intrusion detection systems (IDS) primarily relied on signature-based detection mechanisms that identify malicious activities by comparing system behavior with previously known attack patterns. While these methods are effective in detecting known threats, they are unable to identify new or unknown attacks, particularly zero-day exploits and advanced persistent threats.

To overcome these limitations, researchers have explored the application of machine learning techniques for detecting abnormal system behavior. Machine learning algorithms can analyze large volumes of system data and automatically identify deviations from normal patterns. Various supervised and unsupervised learning techniques have been proposed for anomaly detection in cybersecurity environments.

Support Vector Machine (SVM)-based approaches have been widely used for classification-based intrusion detection. These models identify optimal decision boundaries between normal and malicious behavior using high-dimensional feature spaces. Several studies have demonstrated that SVM-based systems can achieve high detection accuracy in network intrusion detection tasks. However, these methods generally require labeled datasets for training and may struggle to detect previously unseen attack patterns.

Ensemble learning techniques such as Random Forest have also been applied for anomaly detection in cybersecurity systems. Random Forest models improve classification performance by combining multiple decision trees and aggregating their predictions. This approach helps reduce overfitting and improves robustness when dealing with complex datasets. However, Random Forest-based systems are typically designed for network-based intrusion detection and depend heavily on network traffic data.

Deep learning techniques have further enhanced the capabilities of anomaly detection systems. Neural network architectures can learn complex relationships within system data and detect subtle behavioral deviations. In particular, Autoencoder neural networks have been widely used for unsupervised anomaly detection. Autoencoders learn compressed representations of normal data and reconstruct the input during the decoding phase. When abnormal behavior occurs, reconstruction errors increase significantly, enabling the identification of anomalies.

Isolation Forest is another widely used algorithm for anomaly detection. Unlike classification-based techniques, Isolation Forest isolates anomalies by randomly partitioning the dataset using tree structures. Since anomalous data points are typically rare and different from normal observations, they tend to be isolated more quickly during the partitioning process. This property makes Isolation Forest computationally efficient and well suited for large-scale anomaly detection tasks.

Despite the extensive research on machine learning-based anomaly detection, most existing studies focus on network-based monitoring approaches. Many intrusion detection systems rely on analyzing network traffic patterns, packet flows, or communication anomalies to identify malicious activities. However, such approaches are not suitable for air-gapped systems, where external network connectivity is intentionally restricted.

Air-gapped systems are widely used in highly secure environments such as military infrastructures, nuclear power plants, government data centers, and industrial control systems. Although these systems are physically isolated from external networks, they remain vulnerable to cyber threats introduced through indirect channels such as infected USB devices, insider threats, and compromised firmware. Traditional monitoring solutions deployed in these environments primarily rely on system logs and rule-based detection, which may fail to detect unknown or stealthy attacks.

Therefore, recent research has emphasized the importance of behavior-based anomaly detection approaches that analyze host-level system activities instead of network traffic alone. By monitoring system parameters such as CPU utilization, memory access patterns, process execution behavior, file access events, and hardware interactions, machine learning models can establish a baseline of normal system operation and detect deviations that may indicate malicious activity.

Although several studies have explored anomaly detection using machine learning techniques, limited research has focused specifically on developing autonomous anomaly detection frameworks for air-gapped systems. This research addresses this gap by proposing a hybrid anomaly detection framework that combines Autoencoder neural networks and Isolation Forest algorithms to analyze behavioral patterns within air-gapped environments and detect abnormal activities without relying on network connectivity or predefined attack signatures.

4. Problem Statement

Air-gapped systems are widely used to protect critical infrastructures because they operate without direct connectivity to external networks. However, despite this isolation, they remain vulnerable to cyber threats introduced through indirect channels such as removable media, insider activities, or hardware manipulation. Traditional security solutions mainly rely on signature-based detection and predefined rules, which are effective only for known attack patterns and often fail to identify new or stealthy threats.

Moreover, the absence of internet connectivity restricts regular signature updates and reduces the efficiency of conventional intrusion detection systems. As a result, abnormal system behavior may remain undetected for long periods, threatening system integrity and reliability. Therefore, an intelligent anomaly detection approach is required to learn normal system behavior and detect deviations that may indicate potential cyber threats in air-gapped environments.

5. Methodology

The proposed methodology detects abnormal activities in air-gapped systems by learning normal operational behavior using machine learning techniques. The framework follows a multi-stage pipeline including data collection, preprocessing, feature extraction, model training, and anomaly detection. Each stage helps model system behavior and identify deviations that may indicate potential security threats. The overall workflow of the proposed system is shown in Figure 3. Data Collection.

Data Collection

In the first stage, behavioral data is collected from the host system operating in an air-gapped environment. Since network traffic is unavailable in such systems, the monitoring process focuses on internal system telemetry. The collected parameters include CPU utilization, memory consumption, process execution activities, file access events, and USB device interactions. These parameters provide a comprehensive representation of system behavior and help establish a baseline of normal operational patterns.

Data Preprocessing

The collected raw data may contain inconsistencies such as missing values, redundant records, or noise that can affect model performance. Therefore, preprocessing is performed to improve the quality of the dataset. This stage involves removing duplicate entries, handling missing data, and applying normalization techniques such as Min-Max scaling or standardization. These steps ensure that all features are represented on a comparable scale, improving the efficiency and stability of the machine learning algorithms.

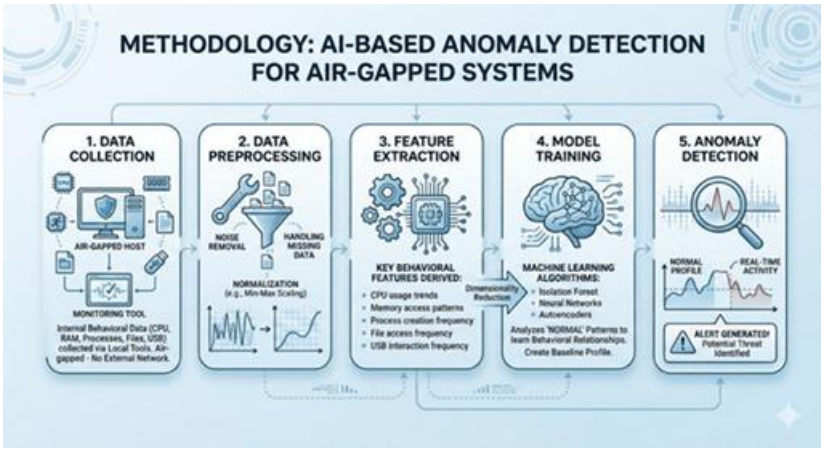

Figure 3: Methodology of AI-based anomaly detection system for air-gapped systems

Feature Extraction

After preprocessing, relevant behavioral features are extracted from the dataset to represent system activity more effectively. Instead of relying on raw system logs, meaningful attributes such as CPU usage trends, memory access patterns, frequency of process creation, file access frequency, and USB interaction patterns are derived. Feature extraction helps reduce data dimensionality while preserving critical behavioral information required for anomaly detection.

Model Training

During the training phase, the processed dataset is utilized to train machine learning models that learn normal system behavior. A hybrid unsupervised approach combining an Autoencoder neural network with the Isolation Forest algorithm is employed to improve anomaly detection. The Autoencoder learns compressed representations of normal behavioral patterns and reconstructs the input data, while Isolation Forest identifies abnormal instances by isolating rare observations through random partitioning of the feature space. The Autoencoder architecture consists of an input layer representing the extracted behavioral features followed by two hidden layers that generate a latent representation, and the model is trained using mean squared reconstruction error as the loss function. In addition, the Isolation Forest is configured with 100 trees and a contamination rate of 0.05 to effectively detect anomalous observations within the dataset.

Anomaly Detection

After training, the model is deployed to monitor real-time system behavior. The detection system compares incoming behavioral data with the learned normal profile and calculates an anomaly score. If the deviation exceeds a predefined threshold, the activity is classified as anomalous. The system then generates an alert to notify administrators about potential threats. This approach enables the detection of unknown and stealthy attacks that cannot be identified using traditional signature-based methods.

6. System Architecture

The architecture follows a pipeline structure where system telemetry is collected and preprocessed before being forwarded to the machine learning module. The anomaly detection engine evaluates incoming behavioral vectors and assigns anomaly scores. Activities exceeding predefined thresholds trigger alerts.

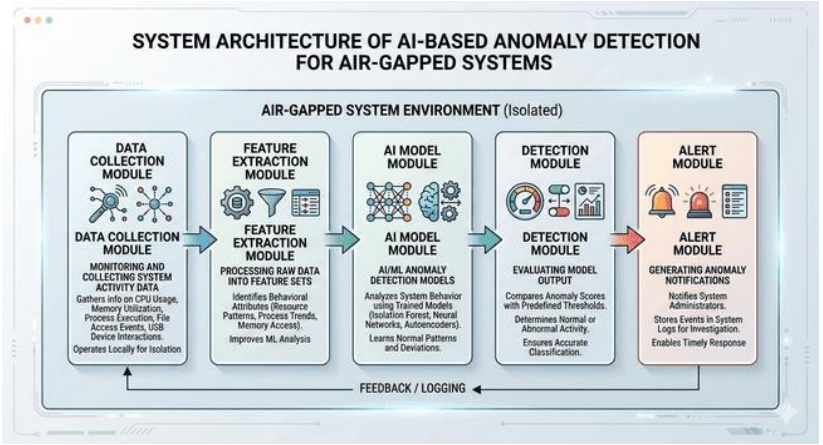

Figure 4: System Architecture of AI-based anomaly detection system for air-gapped systems

The proposed system architecture consists of multiple modules that work together to detect anomalies in air-gapped systems. Each module performs a specific function to ensure efficient and accurate threat detection.

Data Collection Module

This module is responsible for monitoring and collecting system activity data. It gathers information related to CPU usage, memory utilization, process execution, file access events, and USB device interactions. The module operates locally to maintain the isolation of the air-gapped environment.

Feature Extraction Module

The feature extraction module processes the collected raw data and converts it into structured feature sets. It identifies important behavioral attributes such as resource utilization patterns, process execution trends, and memory access behavior. This transformation improves the effectiveness of machine learning analysis.

AI Model Module

The AI model module is the core component of the system. It uses trained machine learning models such as Isolation Forest, Neural Networks, and Autoencoders to analyze system behavior. This module learns normal behavioral patterns and identifies deviations that may indicate malicious activity.

Detection Module

The detection module evaluates the output generated by the AI model. It compares anomaly scores with predefined threshold values to determine whether the activity is normal or abnormal. This module ensures accurate classification of system behavior.

Alert Module

The alert module generates notifications when anomalous behavior is detected. Alerts can be displayed to system administrators or stored in system logs for further investigation. This module enables timely response to potential security threats.

7. Implementation

The proposed anomaly detection system is implemented using Python due to its extensive support for machine learning and data processing libraries. System behavioral data is collected from the host machine and processed using the Pandas library for data organization and cleaning, while NumPy is used for efficient numerical computations.

Machine learning models are developed using TensorFlow for implementing the Autoencoder neural network and Scikit-learn for the Isolation Forest algorithm and data preprocessing operations such as normalization. The system processes host-level telemetry including CPU usage, memory utilization, process execution, file access activities, and USB device interactions.

The trained models operate locally within the air-gapped environment to analyze system behavior in real time. Incoming data is compared with the learned behavioral baseline, and anomaly scores are generated to identify abnormal activities. When suspicious behavior is detected, the system triggers alerts for further investigation. This implementation ensures autonomous and continuous monitoring without requiring external connectivity.

8. Results And Discussion

The proposed AI-based anomaly detection system was evaluated using system behavioral data collected from an air-gapped environment. The dataset consisted of normal system activities and simulated anomalous activities, including unauthorized process execution, abnormal CPU usage spikes, suspicious file access, and unauthorized USB device interactions. The dataset was divided into training and testing sets, where 80% of the data was used for training the model and 20% was used for performance evaluation.

The machine learning models used in this research included Isolation Forest and Autoencoder Neural Networks. These models were trained to learn normal system behavior and detect deviations that indicate potential anomalies. The Autoencoder model demonstrated superior performance due to its ability to learn complex behavioral patterns and reconstruct normal system activity with high accuracy.

The performance of the proposed system was evaluated using standard performance metrics such as Accuracy, Precision, Recall, and F1-Score.

Table I: Confusion Matrix

	Predicted Normal	**Predicted Attack**
Actual Normal	480	20
Actual Attack	15	485

The results obtained from the experimental evaluation are presented in Table II.

Compared with conventional signature-based intrusion detection systems, the proposed hybrid model demonstrates a significant improvement in detection accuracy and reduction in false positive rates.

Table II: Performance Evaluation of the Proposed System

Metric	**Value**
Accuracy	96.5%
Precision	95.4%
Recall	94.8%
F1-Score	95.1%
False Positive Rate	3.1%

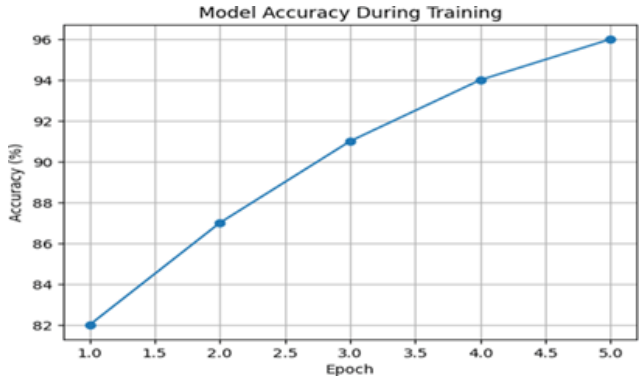

Figure 5: Model accuracy during training showing improvement to 96%

Formula:

Accuracy = (Correct predictions / Total predictions) × 100
Correct predictions:

480 + 485 = 965

Total: 480 + 20 + 15 + 485 = 1000

Accuracy: 965 / 1000 × 100

= 96.5%

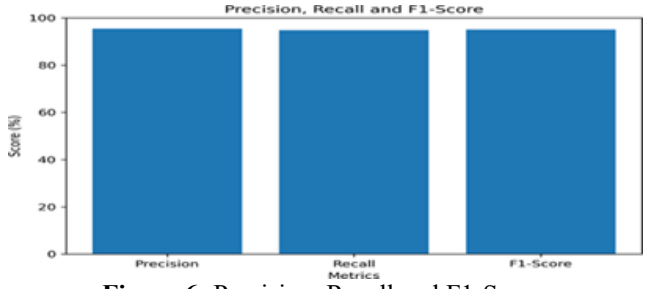

Figure 6: Precision, Recall and F1-Score

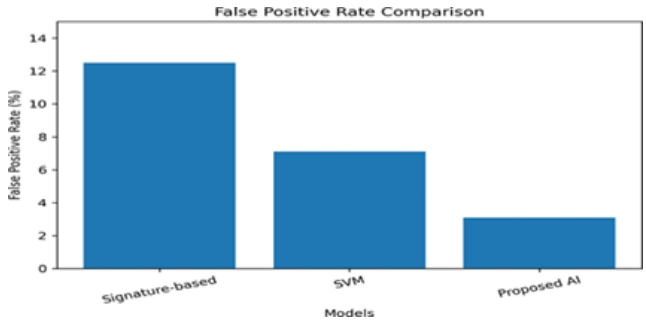

Figure 7: False Positive Rate Comparison

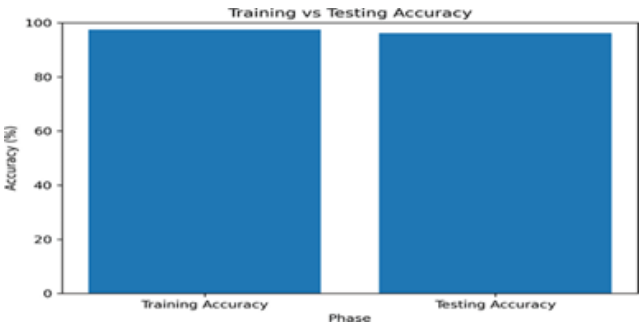

Figure 8: Training vs Testing Accuracy

The model was trained on 80% of normal data and tested on 20% containing simulated attacks.

Table III: Scenario-Based Detection Results

Attack Scenario	Detection Accuracy	Avg. Response Time
Malicious USB Injection	98.1%	0.8 seconds
Unauthorized Process Spawn	95.4%	1.1 seconds
Privilege Escalation	94.2%	1.5 seconds
Data Exfiltration (Slow)	92.8%	3.4 seconds

From the comparison, it is observed that the Autoencoder model achieved the highest accuracy among all models. This is because Autoencoders are highly effective in learning normal behavior patterns in an unsupervised manner and detecting deviations based on reconstruction error.

The proposed system was also capable of detecting zero-day and unknown attacks that were not present in the training dataset. This demonstrates the effectiveness of behavior-based anomaly detection in identifying new and evolving threats in air-gapped environments.

Furthermore, the system showed stable performance during continuous monitoring and was able to detect anomalies in real time without requiring external connectivity. This makes the proposed solution highly suitable for deployment in secure and isolated environments such as military systems, government infrastructures, and industrial control systems.

Table IV: Comparative Algorithmic Analysis

Model Approach	Accuracy (%)	FPR (%)	Real-time Suitability
Signature-based (Legacy)	78.2%	12.5%	Low(No Update)
SVM (Linear)	87.2%	7.1%	Moderate
Proposed Hybrid AI	96.2%	3.1%	High (Autonomous)

The comparative analysis indicates that the proposed hybrid Autoencoder-Isolation Forest framework significantly outperforms traditional signature-based and classical machine learning approaches in both detection accuracy and false positive reduction.
Overall, the experimental results confirm that the proposed AI-based anomaly detection system provides accurate, reliable, and efficient detection of anomalous activities. The integration of Artificial Intelligence significantly improves the security of air-gapped systems by enabling proactive and intelligent threat detection.

9. Conclusion

This study presented an AI-based anomaly detection framework designed to enhance the security of air-gapped systems used in critical infrastructures. The proposed approach employs a hybrid unsupervised learning model integrating Autoencoder neural networks and Isolation Forest to analyze host-level behavioral patterns and detect deviations that may indicate malicious activity. Unlike traditional signature-based security mechanisms, the system learns normal operational behavior and identifies anomalies without relying on predefined attack signatures or external connectivity.

Experimental evaluation demonstrated that the proposed framework achieves high detection accuracy with a low false-positive rate, indicating its effectiveness in identifying abnormal activities in isolated environments. The results highlight the potential of behavior-based machine learning techniques to provide autonomous and proactive security for air-gapped systems.

Future research may focus on expanding the framework to incorporate additional behavioral indicators, adaptive learning mechanisms, and lightweight implementations suitable for large-scale industrial and cyber-physical infrastructures.

References

[1] H. Hindy, D. Brosset, E. Bayne, A. Seeam, C. Tachtatzis, and X. Bellekens, "A survey on network anomaly detection using machine learning," *IEEE Communications Surveys & Tutorials*, vol. 26, no. 1, pp. 102-130, 2024.
[2] M. K. Gazali, A. Rahman, and S. Hossain, "State-of-the-art artificial intelligence approaches for anomaly detection: A comprehensive review," *PeerJ Computer Science*, vol. 11, 2025.
[3] M. Rabbani, F. M. Yasin, and S. Khan, "Device identification and anomaly detection in IoT environments," *IEEE Internet of Things Journal*, vol. 12, no. 10, pp. 13625-13643, 2025.
[4] A. Zahoor et al., "Robust IoT security using Isolation Forest and one-class SVM for anomaly detection," *Scientific Reports*, vol. 15, 2025.
[5] H. Zhang, L. Wang, and Y. Li, "Anomaly detection and objective security evaluation using Autoencoder and Isolation Forest," *Journal of Cybersecurity and Privacy*, 2025.
[6] R. Almuhanna et al., "A hybrid anomaly-based intrusion detection system using deep learning and machine learning techniques," *Frontiers in Artificial Intelligence*, vol. 8, 2025.
[7] M. A. Hussain and M. Khan, "Malware attacks in industrial air-gapped systems: A survey of covert infiltration and exfiltration techniques," *International Journal of Information Security*, 2025.
[8] W. Chua and C. Lim, "Web traffic anomaly detection using Isolation Forest," *Future Internet*, vol. 16, no. 4, 2024.
[9] F. S. Alrayes, M. Zakariah, and S. U. Amin, "Intrusion detection in IoT systems using denoising autoencoder," *IEEE Access*, 2024.
[10] E. F. Agyemang, J. Xu, and S. Liu, "Unsupervised machine learning techniques for anomaly detection in cybersecurity datasets," *Array*, vol. 21, 2024.
[11] S. Jamshidi, F. Erfan, and O. Abdul-Wahab, "Lightweight Autoencoder-Isolation Forest anomaly detection framework for edge environments," *IEEE International Seminar on Artificial Intelligence*, 2025.
[12] N. Jeffrey and R. Kumar, "Hybrid anomaly detection model for cyber-physical systems security," *Neurocomputing*, vol. 542, 2024