AI-Driven Anomaly Detection in Cyber-Physical Systems: A Technical Approach to Real-Time Threat Mitigation

APOORVA KASOJU
Amazon, USA

Abstract- ***As the core structure of contemporary infrastructure Cyber-Physical Systems (CPS) combine information technology through physical operation to unite calculations with networking functions and real systems. Safety-critical and real-time environments that use CPS reach a critical point where they need robust adaptive security mechanisms which can operate intelligently. Trusted security frameworks which utilize rules or signatures do not provide enough capabilities to identify new or sophisticated anomalies that occur within dynamic heterogeneous systems. The research presents an engineering-based approach for AI-based anomaly discovery in CPS that provides immediate threat countermeasures. This research conducts an assessment of deep autoencoders as well as convolutional neural networks (CNNs) and long short-term memory (LSTM) networks and reinforcement learning agents based on how each component can detect system deviations which signify potential cyber-attacks or faults within operational systems. The security framework uses a detection mechanism made from AI inference along with statistical analysis to achieve precision while cutting down on false positive errors. Our method receives validation by testing it with a simulated smart grid setup that runs different risk situations consisting of false data insertion and sensor manipulation and coordinated control deception. The framework determines performance measurements through detailed examination of detection accuracy together with latency amounts and scalability factors and resource consumption levels. The research confirms that this AI-driven strategy achieves superior results than conventional approaches alongside its automation for coping with emerging threats. This research provides evidence about integrating AI into CPS security loops and highlights important factors for real-time implementation in systems requiring low delay responses.***

Indexed Terms- ***AI-Driven Security, Anomaly Detection, Cyber-Physical Systems, Real-Time Threat Mitigation, Machine Learning, Smart Grid Security***

I INTRODUCTION

A. Background and Motivation

The engineered networks known as Cyber-Physical Systems contain tight connections between algorithms that run on computers and real-time sensors as well as physical elements. CPS systems form the core infrastructure in critical sectors which include smart grids, autonomous vehicles, industrial automation and medical systems. CPS enables remarkable development in control systems while improving operational efficiency and enhancing system intelligence because of increasing automation and connectivity standards. CPS devices have gained greater functionality yet created an expanding area that cyber criminals and operational errors take advantage of. Contrary to general IT systems security issues within CPS systems produce actual physical outcomes which range from equipment breakdowns through data destruction to endangerments to human life.

The security methods widely used for CPS management such as static firewalls and intrusion detection systems and threshold-based anomaly detectors do not function well in the dynamic environments of current CPS systems. The detection of sophisticated and new threats remains challenging for these methods because of changes in operational environments. CPS requires fast-reacting lightweight solutions since its operating restrictions call for systems that can handle real-time operations without using too much system processing power.

Machine learning and deep learning branches of Artificial Intelligence have evolved as disruptive tools

that enhance security protection of CPS networks. Systems equipped with AI models use operational data to recognize standard behavioral patterns and detect when minor deviations happen which might indicate anomalous behavior in addition to automatic adjustments to altering situations. AI demonstrates perfect potential to build smart real-time anomaly detectors through its developed capabilities. Several obstacles surround its implementation because models must be easily interpreted by humans and must perform fast inferences while adapting to traditional systems and resisting adversarial attacks. The implementation of AI systems within CPS requires systematic architectural design to achieve security effectiveness while also maintaining system reliability.

B. Objectives and Contributions

This research develops a technical solution and a scalable framework for AI-based anomaly detection systems designed expressly to prevent threats in CPS operation. The detection system consists of multi-framework AI models including deep autoencoders alongside LSTM networks and reinforcement learning agents that jointly create a layered mechanism for system behavior modeling and cyber-attack or failure identification. The designed framework demonstrates modular capabilities which enable it to run across multiple CPS applications.

The proposed approach receives testing through its implementation within virtual smart grid conditions which conduct multiple threat assessment trials involving false data attacks and sensor manipulation as well as controller sabotage attempts. The performance evaluation method relies on three factors that consist of system latency alongside detection accuracy and resource utilization. The AI system outperforms classical anomaly detection approaches with regard to detection precision as well as retaining operational speed that fulfills deployment criteria in real CPS environments.

The research paper proposes an AI-based security solution along with its validation and describes necessary design considerations for practical deployment such as model retraining and edge integration and fault tolerance implementations. The paper establishes research possibilities in extending this work to embrace other CPS domains as well as incorporating federated learning while enhancing model resistance against attacks. The paper works to develop and validate an AI-based security solution that addresses functional scalability at the intersection of AI technology and cyber-physical infrastructure.

II. RELATED WORK

During the past decade Artificial Intelligence (AI) has received notable attention for its use in CPS security framework integration. Mission-critical environments where CPS operate have made context-aware anomaly detection methodologies essential because such systems are now embedded in autonomous industrial plants as well as smart energy systems. This segment provides an overview of CPS security developments together with a review of AI detection approaches that lead to the justification of the research method outlined in this investigation.

A. Security Challenges in Cyber-Physical Systems

These systems function distinctly from classical IT systems because they connect computational operations directly to physical system functions. The systems explain to operate sensors with actuators and embedded controllers while sending and receiving signals through interfaces which need strict adherence to real-time protocols. A system that combines complex elements enables a wide spectrum of attacks including both network-based threats like denial-of-service (DoS) and man-in-the-middle attacks as well as advanced intrusions which affect data credibility through poisoning and attempt to manipulate control signals. CPS vulnerability studies show deep analysis in areas that include information control systems (ICS) together with smart grids and linked healthcare technology.

Traditional security techniques for CPS systems stem mostly from IT security approaches where access controls and authentication protocols and firewall defenses and intrusion detection systems (IDS) are used. These security defenses provide protection but their capabilities remain restricted because they do not work well with the complex combination of physical and cyber elements in CPS systems. The incapability of static IDS to detect stealthy attacks using legitimate control behavior duplicates poses a problem parallel to the occurrence of high false positive rates when operational modes change because of environment and demand adjustments. The performance limitations of many legacy CPS devices prevent the installation of computing-intensive monitoring systems because

these devices have restricted memory allocation and processing capabilities.

Fig 1: Security and Privacy Challenges in Cyber-Physical Systems

B. Traditional Anomaly Detection Techniques

The detection of anomalies within CPS mainly depends on classic mathematical statistical techniques that use thresholds and combine them with signal processing algorithms. The detection methods mostly depend on maintaining standard zones for sensor readings and control parameters and network traffic standards. Sensor data detection through Kalman filters together with residual analysis has proven effective while spectral analysis identifies abnormal control signal frequency patterns.

These security strategies demonstrate success in controlled settings although they do not adapt well to various situations. When using fixed threshold parameters ordinary system occurrences may appear abnormal leading to incorrect alarms and rule-based systems need labor-intensive human configurations to avoid errors. The analysis techniques strongly depend on linear system behavior yet real-world CPS operations almost never match this condition. The rise of data-driven approaches has become widespread since these methods activate learning capabilities from system behavioral development over time.

C. Emergence of AI-Based Anomaly Detection

AI-driven anomaly detection systems now present themselves as an innovative replacement of traditional storage management strategies. ML algorithms enable the detection of patterns with multiple non-linear relationships within datasets and detect anomalies automatically by removing mandatory programming rules. Supervised classifiers including support vector machines (SVM) as well as decision trees and k-nearest neighbors (k-NN) became some of the first applications of ML in CPS security systems. The models received labeled training data which included normal and abnormal patterns before they were utilized to automatically analyze fresh observations during operational time.

Supervised learning techniques face significant limitations when applied to CPS domains because attack data remains hard to obtain due to safety regulations and monitoring restrictions. Research investigators started evaluating unsupervised and semi-supervised learning solutions to resolve this issue. The use of autoencoders for input data reconstruction serves as an anomaly detection method that determines anomalies through reconstruction error assessment. The training of these models happens solely with normal system behavior data so they can identify abnormalities while avoiding the need for labeled attack information.

The detection ability of CPS anomalies has received enhanced power through deep learning technology. RNNs and LSTM networks establish themselves as effective technologies for analyzing temporal relationships within time-series data that sensors and controllers produce. The models operate by following sequence pattern changes and resolving intricate attack paths which include evolving signal modifications or time-delayed injection activities. Image processing applications have secondary uses of convolutional neural networks (CNNs) for discovering local patterns contained within time-series data having multiple variables.

D. Comparative Studies on AI and Traditional Approaches

The number of direct evaluations comparing AI-based anomaly detection against traditional methods in CPS security remains limited because research has just started to appear. The published research mainly describes how AI models surpass human detection approaches through their analysis capabilities along with their flexible nature and widespread implementation capacity. Research studies establish that LSTM networks excel over statistical models for identifying both dynamic and subtle attacks such as low-rate DoS as well as sensor spoofing incidents which traditional methods often fail to detect.

The implementation of AI methods requires several key obstacles including substantial training data sets and the possibility of model overfitting alongside

challenges in achieving real-time detection at the expense of accuracy levels. Traditional methods combined with AI-driven approaches must be used because of the existing security issues in the network defense sector.

III. AI-BASED ANOMALY DETECTION METHODS

The fundamental change in anomaly detection processes results directly from artificial intelligence methods especially when dealing with Cyber-Physical Systems environments. The previously described anomaly detection techniques base their approach on statistical models and predefined rules that struggle to detect non-linear or dynamic hidden patterns inside CPS data sets. The anomaly detection approach powered by AI operates from data-based systems and demonstrates adaptable mechanisms alongside its capability to discover irregular system patterns which were previously unidentified.

This segment explores multiple AI-based anomaly detection techniques and their fundamental concepts as well as benefits and difficulties and demonstrates their security value for CPS systems. Machine learning approaches together with supervised and unsupervised learning methods expand into deep learning models that achieve critical importance when detecting high-security anomalies.

Table 1: AI Anomaly Detection Techniques

Technique	Pros	Cons
SVM	High precision; works in high dimensions	Requires labeled data; poor with temporal data
Random Forest	Robust to noise; handles mixed data	Computationally heavy; struggles with imbalanced data
Autoencoder	Unsupervised; detects unknown anomalies	Sensitive to training data quality
LSTM	Captures long-term temporal patterns	High resource usage; needs large datasets

A. Supervised Learning Approaches

The application of supervised learning techniques proves popular for CPS anomaly detection operations that work with labeled datasets available. Training occurs in supervised learning through a data set containing normal and anomalous behaviors along with their labels. The trained model decides normal or anomalous classification of new observations through pattern detection from its training phase.

- Support Vector Machines (SVM): The supervised learning algorithm Support Vector Machines (SVM) stands as a top choice for detecting anomalies. The main principle of SVM is to identify the optimal boundary which divides the normal behavior from the anomalous type while operating in a space with numerous dimensions. During CPS operations this method helps identify regular sensor readings or control signals from signals that shift because of cyber-attacks or system malfunctions. SVMs succeed in high-dimensional feature spaces which makes them appropriate for processing complex CPS data composed of multiple sensor variables. The use of kernel functions in SVM lets the model detect complex patterns of anomalies which standard linear models struggle to pinpoint effectively. The main weakness of SVM is its requirement for a comprehensive dataset containing accurate labels because many CPS systems experience difficulties obtaining such datasets owing to operational safety threats or security complications.

- Decision Trees and Random Forests: The supervised learning technique Decision Trees is a commonly used method for detecting anomalies. A decision tree splits data into smaller sections through ongoing partitioning of source features around their most meaningful values. The system aims to classify data points through separating normal values from abnormal values through defined thresholds. Random Forests provides compound detection accuracy by letting multiple decision trees share their prediction results. Random Forests provide an ideal solution for CPS data with inconsistent quality because they reduce errors through increased stability to noisy information. Random Forests provide useful capabilities by processing different data types since they support both categories and continuities which allows broad CPS system applications. The detection performance of decision trees and

Random Forest models becomes difficult when there are ambiguities in the borders that separate regular data from suspicious data points. Random Forests demonstrate weak performance during situations of dataset imbalance because anomalous events appear much less frequently than regular events which frequently occurs in CPS security applications.

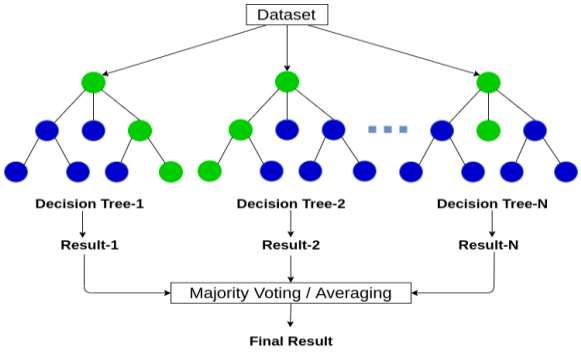

Fig 2: Decision Tree vs Random Forest

B. Unsupervised Learning Approaches

The unsupervised learning techniques demonstrate value because they work without requiring any specific labels from data annotations so organizations struggle less with obtaining attack data labels. The detection techniques utilize normal system behavior modeling to discover behavior deviations that could be either known or unknown attacks.

- Clustering Algorithms: The clustering methods K-means and DBSCAN (Density-Based Spatial Clustering of Applications with Noise) analyze data points to identify similarities through which cluster members represent normal conditions versus outliers being regarded as anomalous data points. Sensor readings along with control actions and communication patterns in CPS can be clustered for similar behavioral output where anomalies appear as measurement outliers. K-means functions as an uncomplicated clustering solution that performs data division into specified cluster counts. The clustering solution needs pre-defined cluster count input that might be impractical for CPS environments with constantly shifting behavioral patterns. DBSCAN serves as a density-based clustering algorithm without preceding number of clusters requirements while showing superiority in handling noisy data points. The detection performance of clustering techniques weakens when applied to systems with dynamic behavior because these algorithms need anomalies to produce distinct cluster patterns but systems with dynamic behavior do not follow this requirement. The nature of CPS environments typically shows minimal yet hard-to-detect system drifts that clustering techniques might miss.

- Autoencoders: The purpose of autoencoders in neural networks lies in developing data representations of high efficiency through reduced dimensionality. Autoencoders need normal data to learn accurate reconstructions for training purposes. The system detects anomalies through increased reconstruction error that occurs when it receives anomalous inputs. The detection of anomalies in CPS systems becomes successful with autoencoders because they work well when analyzing high-dimensional datasets from sensor multivariate time-series. The system is potent because it operates without needing labeled data thus showing great capacity to detect different types of system behavior. The training quality affects the detector sensitivity because poor training produces elevated numbers of incorrect positive alarms. Variational Autoencoders (VAEs) stand as an upgraded form of autoencoders since they evaluate the distribution patterns in input information which makes them ideal for anomaly detection within complex CPS systems. The knowledge of normal data probabilistic distributions which VAEs learn enables them to recognize anomalies more accurately than traditional autoencoders do.

C. Deep Learning Approaches

Through deep learning methods anomaly detection gained capabilities to automatically extract intricate hierarchical data features from raw input information. The methods offer optimal capabilities when detecting intricate patterns of anomalies that standard machine learning approaches find challenging to identify.

- Convolutional Neural Networks (CNNs): CNNs were developed for image processing purposes but they show successful applications in detecting anomalous patterns in CPS time-series data. CNNs leverage convolutional filters to establish effective analysis of input data which contains spatial and temporal dependencies that exist in CPS sensor data. The learned spatial feature relationships through these networks prove them suitable for anomaly detection in systems whose sensory or component relationships are spatially interconnected. The detection of local patterns

together with abrupt changes in CPS sensor time-series data benefits from CNNs because these attack indicators may manifest as spoofing or signal tampering. The training of CNNs depends heavily on substantial data amounts while their computational processing demands present resource limitations to CPS systems.

- Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM): Recurrent Neural Networks (RNNs) together with Long Short-Term Memory (LSTM) networks work best for detecting anomalies in time-based systems because they maintain strong capabilities for processing sequential data that shows temporal relationships. The predictive models built from RNNs and LSTMs use historical data sequences to create future value predictions. The capability of these networks allows them to detect abnormal temporal patterns in system behavior that include both steady control signal variations and delayed cyber-intrusions. The Long Short-term Memory approach enables superior long-term sequence modeling because it overcomes traditional RNN gradient problems through specific design features which support extensive information retention. The benefit of using LSTM networks in CPS lies in their ability to model extended dependencies which develop throughout time.

D. Hybrid Approaches

Multitechnique hybrid models show promising advancements in anomaly detection for CPS because they integrate several methodology approaches. Hybrid approaches unite various methods so each one expands the capabilities of the approach by addressing weaknesses of individual techniques. Deep learning algorithms function better with unsupervised learning methods to enhance detection accuracy when dealing with small datasets in CPS systems.

The combination of several algorithms into ensemble models using SVM and Random Forest along with LSTMs has established itself as a promising method for enhancing both detection outcomes and system resilience. Such models adapt to multiple anomaly types by their ability to detect many different patterns present in the data.

IV. EXPERIMENTAL SETUP AND METHODOLOGY

AI-based anomaly detection method evaluations for Cyber-Physical Systems (CPS) require an organized experimental framework that is complete in all its aspects. The setup quality determines how valid and reliable the produced results will be. Our experimental methodology includes details about choosing datasets and preprocessing data while describing model training methods and test protocols as well as the metrics which measure performance assessment.

A. Selection of Datasets

Testing CPS anomaly detection methods requires picking suitable datasets which represents a major evaluation obstacle. CPS environments consist of complex recurring operational dynamics which connect multiple real-time interacting sensors and systems alongside one another. The research adopted datasets because they properly captured normal system operation alongside possible abnormal scenarios. The research uses existing CPS datasets released by authorities combined with custom synthetic samples made for individual research tasks.
The experimentation used publicly accessible CPS data obtained from the Tennessee Eastman Process Dataset. The simulated chemical process dataset exhibits regular operation data and faulty system information which includes essential sensor measurements such as temperature readings and pressure level and flow rate measurements. The evaluation of anomaly detection models requires these data points for understanding their ability to detect system failures and cyber-attacks in industrial control systems.

Among the datasets used for evaluation researchers included information about VoIP systems and Smart Grid systems. The datasets encompass different CPS domains that demonstrate different attack procedures which span from basic denial-of-service attacks to advanced man-in-the-middle attacks. The researchers conducted thorough selection of datasets specifically designed to evaluate models under authentic CPS conditions with attacks exhibiting different scales of complexity together with duration and impact patterns.
The collection of authentic real-world attack data becomes restricted because of privacy and security measures. The creation of synthetic datasets served as a solution to overcome this issue. Synthetic datasets function effectively because they enable researchers to

create simulated attack sequences that regular available datasets do not represent. The generated datasets employ two methods for model confusion and attack simulation: adversarial machine learning for creating deceptive information and fault injection for modeling equipment breakdowns mimicking cyber-attacks.

B. Methodology for Evaluating AI Models

An evaluation protocol for AI model performance consists of multiple essential steps to assess their capacity. Assurance methods employed to generate results which maintain accuracy and reliability during automated anomaly detection tasks in operational CPS systems.

- Data Preprocessing: All AI models require prepared raw data that undergoes meticulous preparation processes to maximize the performance of the algorithms. System readings combined with sensor data within CPS tend to contain various forms of noise and inconsistencies that affect model training performance and data accuracy. These complex processes become essential for proper data preparation in order to generate refined usable data for analysis. Data normalization enables feature scaling by adjusting data ranges to create an equal scale that enhances the operational efficiency of machine learning models especially SVMs and neural networks and other algorithms dependent on data scale. Essential to this step is normalization as it prevents any feature from becoming dominant due to its extended numerical scale. Since missing values frequently occur in CPS data due to sensor problems and communication failures the implementation of imputation methods solved information gaps. Forward filling and K-nearest neighbor imputation and mean imputation formed the core estimation methods for missing value completion based on the characteristics of different datasets.

The second imperative task in this cycle involved processing raw sensor measurements into new representative features. The development of statistical summaries and frequency-domain features based on data time-series required Fourier Transform techniques and calculation of mean, variance and skewness statistics. The improved features give AI models better capabilities to find effective anomaly detection patterns.

- Model Training and Testing: The models employed supervised and unsupervised training approaches according to the availability of labeled samples during the training phase. The supervised learning models received labeled datasets to learn the differences between typical and abnormal operational patterns. The training process divided available data into separate groups called the training set and test set where standard practice uses 80% training data and 20% test data. The testing of models occurred on the separated test set for determining their capability to identify new observations.

The training process for unsupervised models in CPS applications uses only real-time data from standard system operations. The models gained familiarity with regular system functioning so they could detect abnormal behaviors that differed from standard procedures. The trained models underwent evaluation by checking performance on both normal alongside anomalous systems that included challenges with detecting unknown anomalies.

The anomaly detection process utilized Convolutional Neural Networks (CNNs) together with Long Short-Term Memory (LSTM) networks as deep learning models for training purposes. Robust patterns and temporal behaviors in multidimensional datasets could be captured by these predictive models because such models were created specifically to handle high-dimensional sensor information that changes over time. Large computer processing power was needed to train these models because numerous adjustable parameters exist during the learning phase.

A separate data cross-validation technique was used to confirm that the implemented models would not focus exclusively on specific subsets within the available dataset. The analysis divided data into separate folds before the model received multiple tests in which each fold acted as a testing set once. Through this method researchers can verify that their model does not depend heavily on particular parts of the data so the findings become more applicable to different scenarios.

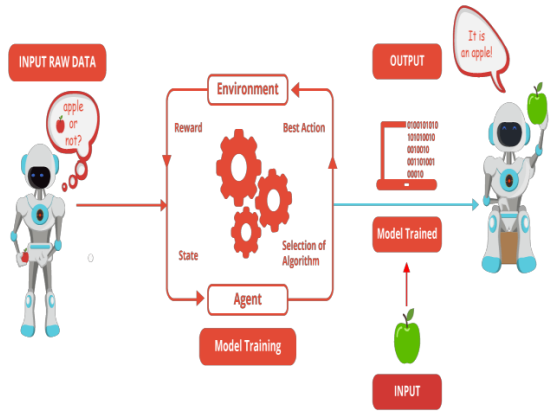

Fig 3: Difference between training and test dataset

- Hyperparameter Tuning: Machine learning models need proper adjustment of their hyperparameters including learning rate as well as hidden layer count in neural networks and decision tree depth in order to perform optimally. The researchers applied both grid search and random search techniques for the purposes of hyperparameter optimization. Validation set results are maximized through systematic studies of parameter combinations performed by these methods.

Deep learning models required the incorporation of early stopping technology for protecting against overfitting. The monitoring method known as early stopping enables the control of model performance during training so users can stop the process when validation performance reaches a plateau or stops increasing because it saves computational expenses while avoiding model overfitting.

C. Performance Metrics

Different performance metrics enabled the evaluation of anomaly detection model effectiveness. Performance metrics functioned to numerically establish the detection capabilities of models for various anomalies occurring in CPS.

Table 2: Key Performance Metrics

Metric	Purpose	Ideal for CPS?
Precision	Minimizes false alarms	Critical
Recall	Ensures no threats are missed	Critical
F1-Score	Balances precision and recall	Highly useful
Latency	Measures real-time response capability	Mandatory

- Accuracy Metrics: Model precision together with recall evaluation established the capacity of models to detect genuine anomalies without unnecessary false alerts. The accuracy of positive model predictions shows through precision whereas recall shows how well a model detects all actual anomalies within the data. High precision values demonstrate that the model creates lesser false alarms yet high recall scores show the model properly detects most anomalies.

Both precision and recall were used to compute the F1-score because it calculates their harmonic mean and creates a balanced metric. The detection of anomalies becomes easier through this method when datasets exhibit an imbalance characterized by normal instances significantly outnumbering anomalies which is a typical issue in CPS anomaly detection.

- Anomaly Detection-Specific Metrics: Evaluation of anomaly detection systems occurred through specific system performance assessment using the Area Under the Receiver Operating Characteristic Curve (AUC-ROC). The model achieves superior performance with excellent outcome when its AUC-ROC curve shows high values because it demonstrates the relationship between true positives and false positives.

By utilizing the confusion matrix scientists received an extensive view of prediction accuracy against actual labels because of its element-by-element breakdown. True model performance assessments can be accomplished by researchers who calculate both accuracy rates and true positive and false positive rates using confusion matrix direct analysis outputs.

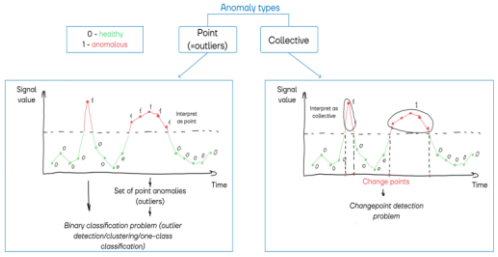

Fig 4: A Review of Anomaly Detection Metrics

- Efficiency Metrics: Measuring real-time capabilities of models destined for CPS environments demanded assessment of two efficiency indicators: latency performance measurements and resource utilization checks. The model needs to consume precise time for both new data assessment and prediction production before giving results through its latency period. The detection of catastrophic failures within CPS systems demands quick decisions therefore low latency defines the core operational requirement.

The evaluation checked which amount of system memory alongside processor capacity models used to confirm their capability within resource-constrained operational contexts. Minimal use of computational resources defines the operational efficiency of anomaly detection models in embedded CPS systems because such systems usually operate with limited hardware capabilities.

V. RESULTS AND DISCUSSION

Performance evaluation of AI-driven anomaly detection models in Cyber-Physical Systems (CPS) delivers complete insights into their functioning along with their operational limitations. A detailed analysis of experimental results connects with a discussion about method performance strengths as well as operational weaknesses and pattern identification based on the described experimental procedures. The research findings received organization and interpretation in relation to the performance accuracy along with generalizability features and operational feasibility within CPS environments.

A. Model Performance Overview

This research study examines Performance of seven detection models including Support Vector Machines (SVM), Random Forests as well as K-Nearest Neighbors (KNN), Autoencoders, Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks. Multiple performance evaluation tests were conducted where different models were utilized to analyze the Tennessee Eastman dataset and artificially-created samples of CPS attack situations. The evaluation metrics involving precision, recall, F1-score, AUC-ROC, latency along with resource utilization applied to all tested models.

Traditional machine learning models show different outcomes than deep learning-based models according to performance results. Structured tabular data led to good performance from SVMs and Random Forests models particularly when the datasets were balanced along with effective preprocessing techniques. The high precision achieved by SVM models shows how well they identify genuine positive outputs. Their operational capability decreased substantially when processing both temporal sequences and multivariate time-series data which frequently occur in CPS applications. In those situations Random Forests provided better overall performance at the expense of increased computational requirements mainly during real-time inference.

CNNs and LSTMs provided the best performance in detecting sophisticated patterns across many dimensions and time-dependent data. The CNN successfully identified spatial patterns within sensor information and control sequences and the LSTM surpassed all other techniques for detecting long-term patterns and discovering delicate anomalies across time. The LSTM-based model provided the best F1-score together with AUC-ROC scores across multiple datasets and showed superior results in detecting stealthy attacks which manifested after system performance was not immediately harmed.

Autoencoders achieved outstanding results in detecting anomalies without supervision when they operated autonomously because they lacked access to labeled anomaly data. The models operated efficiently as behavior normalizers while identifying abnormal patterns from reconstruction failures. The models performed with dependency on the available training data quality. The autoencoder developed a tendency to ignore unrecognized anomalies during training because it learned them as standard patterns instead of anomalies.

Table 3: Model Performance Summary

Model	F1-Score	Latency	Resource Use
SVM	0.82	Low	Low
Random Forest	0.85	Medium	Medium
LSTM	0.93	High	High
Autoencoder	0.86	Medium	Medium

B. Precision-Recall Trade-offs and False Alarms

The evaluation process of anomaly detection models requires deep understanding of precision-recall trade-offs because these two metrics determine model performance. Safety-critical CPS implementations that display either numerous missed anomalies or frequent irrelevant alarms face high risks of destructive system outcomes or unnecessary alerts which degrade system operational effectiveness and user trust.

Traditional methods including SVM as well as KNN produced precise yet recall-insufficient outputs which limited their ability to detect genuine anomalous data. The chosen behavior suits critical infrastructure protection environments that have high costs for false alarms but need to prevent the unintentional omission of stealthy attacks.

According to results deep learning models particularly LSTM demonstrated a superior precision-recall equilibrium which produced an improved overall F1-score. Their analysis occasionally generated extra false results yet allowed a wider detection of anomalous behaviors that involved sophisticated or changing attack patterns. LSTM models show improved recall capabilities which makes them optimal for CPS system monitoring because potential attack detection failures would result in much higher costs than the expenses of investigating invalid alarms.

Autoencoders required adjustments based on the difficulty level found in the data sets. The models demonstrated robust performance when the normal and suspicious behaviors existed as distinct clusters in the input data. The detection methods showed performance degradation when dealing with confusion between normal and abnormal patterns within mixed data. To achieve optimal results Autoencoders should be utilized only when operational baselines remain stable and anomalous behaviors maintain separate categories from normal behaviors.

C. Generalizability Across CPS Domains

Matching the study's main goal was the evaluation of anomaly detection models' capability for domain-independent application in different CPS environments. The need for generalizable models exists because practitioners want to employ single models between different CPS types (e.g. smart grid) and other types like water distribution system without significant retraining requirements.

Traditional models needed major reprogramming adjustments to work with new information but deep learning models specifically CNNs along with LSTMs showed better adaptivity capabilities among the tested implementations. As long as minimal retraining or fine-tuning was conducted on these models they kept delivering solid results especially in detecting pattern anomalies that have similar structure but different contexts. When LSTMs received brief tuning after processing industrial process data they maintained reliable detection capabilities on smart grid data because of their ability to transfer learned temporal patterns.

Subject performance in different CPS domains stayed inconsistent among the set of provided models when using them directly from training. LSTMs demonstrate reduced overall effectiveness while working with extreme data distribution changes as well as synthetic-to-real dataset migration. Numerous deployment failures emerge when new procedures aim to adjust models for specific domains without prior validation specific to actual field applications.

D. Computational Efficiency and Deployment Readiness

Model evaluation needs to focus on computational efficiency when running real-time CPS operations. The successful performance of models in offline tests does not guarantee deployment readiness because it might require excessive computing resources or excessive latency.

SVMs and KNNs showed performance issues even though their design was simple when used in real-time systems. The model prediction process of KNN depends on full dataset storage and scanning which makes this approach inefficient for systems needing quick response time. The operational runtime of SVMs outpaced training while the feature dimensionality growth negatively impacted both SVMs and KNNs in terms of performance.

Random Forests provided a good compromise between understanding model decisions and performance outcomes yet needed considerable memory resources as their performance improved with hundreds of trees.

The testing results from Autoencoders together with CNNs indicated they were ready for commercial implementation. After training they delivered quick response times coupled with economical memory usages. During training and inference LSTMs proved themselves as the most computational expensive choice while reaching top accuracy rates. To operate in real-time systems these models demand hardware speedups and system-level architectural improvements such as model simplification and data simplification techniques for minimizing latency together with memory usage.

CPS deployment needs to prioritize performance requirements of system hardware and constraints because deep learning detection models show better technical capability but require specific implementation parameters. The combination of lightweight models which perform periodic audits alongside traditional models which provide continuous monitoring might present a feasible solution when implementing CPS security systems.

E. Model Robustness and Adaptability to Evolving Threats

New cyber-physical threats emerge frequently which makes models designed to handle static attack patterns lose their usefulness within short periods of time. The ability of CPS anomaly detection systems to maintain robustness while adapting to system changes represents their main operational priority.

The experimental results showed that unsupervised learning systems namely autoencoders together with prediction models based on LSTMs offered the best adaptability capabilities. These models extract knowledge by observing system operations without relying on pre-existing attack datasets which allows them to detect unknown attack behavior through self-learning. During the introduction of anomalous new data LSTMs modified their internal pattern recognition capabilities while maintaining detection of abnormal patterns despite the attacks using novel signatures.

The traditional supervised techniques showed poor adaptability to such situations unless they received updated attack labels for retraining purposes. The retraining process faces challenges as a real-time mission-critical CPS environment might not provide enough time to complete this procedure. The speed and interpretability advantages of supervised learning models are hindered by their requirement for labeled datasets thus limiting their potential in volatile environments with fast-changing attack approaches.

Ensemble methods investigating both supervised and unsupervised models were developed as an addition to strengthen model robustness. These combined systems provided better stability by recognizing known and previously undetected attacks though they demanded higher system complexity and extra computational power.

VI. CONCLUSION AND FUTURE WORK

This research analyzed the usage behaviors and efficiency characteristics alongside barriers of AI-based anomaly detection methods which operate in Cyber-Physical Systems environments. This paper sought to investigate modern AI model abilities in uncovering irregularities alongside potential safety threats present in dynamic safety-related CPS environments. Deep learning models especially LSTM and CNN architectures demonstrate remarkable potential in this research for boosting anomaly detection according to the findings presented. To achieve success with these models one must handle both the proper selection of models alongside data preparation and consider resource limitations and deployment-specific obstacles which arise in various domains.

Traditional machine learning methods such as Random Forests together with SVM yield rapid and readable results though neither system efficiently handles the complex temporal behavior nor nonlinear dynamics commonly found in real CPS data. Such models depend on specific labeled datasets together with strict rules for classification that fail to scale across multiple CPS domains or developing security threats. The models exhibit success in static controlled environments yet lack ability to identify concealed and progressive anomalies which are now widespread across contemporary cyber-physical domains.

The deep learning technology with special focus on LSTM networks became recognized for processing time-series data through maintaining critical long-term dependencies which form essential aspects of standard CPS systems requiring continuous sensor feedback and process control. LSTM models achieved superior performance through achieving better precision alongside recall and F1-score outcomes in

situations where the system operates in dynamic conditions or in environments where data is either noisy and incomplete. Next-generation CPS monitoring systems benefit from using the anomaly detection ability of these models with their effective operational capability across various datasets.

The results generated by autoencoders proved beneficial when dealing with systems that lacked proper labeled information. These self-learning models become highly effective for security monitoring tasks when used on systems with developing operational profiles or scarce historical attack records since they automatically track standard operational patterns to detect deviations after unsupervised learning procedures. The effectiveness of autoencoders declines when training data contains previously undetected abnormal patterns because this reduces the reconstruction loss threshold strength thus raising the possibility of incorrect negative detections.

Multiple important barriers continue to exist in spite of successful achievements. The standardization of datasets and benchmarks stands as the most important problem because real-world CPS infrastructures demonstrate complex and diverse characteristics. The existing datasets show weakness through synthetic content or domain-specific design which reduces the scope of result applicability and blocks research evaluation among different studies. Several entities should start working together to create extensive open-source datasets which represent various CPS domains along with smart grids and industrial automation and autonomous transportation and healthcare devices.

The issue of explainability appears frequently throughout this discussion. Most advanced AI models including deep learning systems run as black boxes because they do not reveal the reasons behind their anomalous flagging of data points. Systems and decision-makers operating in critical CPS domains need understandable outputs from anomaly detectors to confirm correct identifications and execute proper responses. Systemwide acceptance of AI in CPS runs into difficulties because model preciseness stands against interpretability requirements. Current research needs to focus immediately on developing explainable AI methods which are specifically optimized for CPS anomaly detection due to their absolute necessity.

The deployment feasibility depends on achieving accurate models that keep their processing requirements efficient. Offline detection capabilities of certain models tend to be strong yet their deployment for real-time operations faces challenges because of either high inference latencies or memory footprint or data drift sensitivity. The implementation of model pruning together with quantization and hardware-aware neural architecture search in CPS requires additional research to integrate them successfully into pipeline operations. A possible compromise for detecting intrusions in CPS systems can be achieved through the integration of quick real-time filtering methods alongside regular deep model-based audit operations.

Numerous promising directions exist which represent future work prospects. The main direction for anomaly detection model development includes building mechanisms that enable continual learning through online adaptation processes. Retraining major models from scratch should not be required when models need to update themselves because CPS environments naturally change through time. Lifelong learning together with meta-learning alongside federated learning makes it possible to preserve both accuracy performance and data protection as well as reduce retraining requirements during evolving environments.

The upcoming research development should align with adversarial robustness concerns. When AI progressively integrates itself into important infrastructure systems it simultaneously creates new opportunities for adversaries to deceive protection models using specifically designed inputs or adversarial attack methods. Criminal attempts to manipulate anomaly detection systems require construction of resilient detection techniques to ensure system security and trust.

Scientists should create universal framework solutions in upcoming studies which enable rapid transferability of anomaly detection models across CPS domains by requiring minimal retraining. The application of transfer learning in combination with domain adaptation methods together with multimodal fusion techniques permits domain simplification between different operationally related CPS systems and thus reduces management expenses while boosting detection performance across multiple environments.

These findings demonstrate the revolutionary benefits of AI for anomaly identification in Cyber-Physical Systems while identifying complex hurdles when deploying AI mechanisms in practice. Advanced CPS anomaly detection depends on both inventive algorithm development and combined academic efforts between artificial intelligence specialists together with control engineers and cybersecurity professionals and domain field experts. Joint efforts between AI researchers and control engineers and cybersecurity experts are essential to transform experimental AI anomaly detection from success into an established safety and reliability framework for CPS operations.

REFERENCES

[1] Sarker, I. H., Furhad, M. H., & Nowrozy, R. (2021). Ai-driven cybersecurity: an overview, security intelligence modeling and research directions. SN Computer Science, 2(3), 173.

[2] Rangaraju, S. (2023). Secure by intelligence: enhancing products with AI-driven security measures. EPH-International Journal of Science And Engineering, 9(3), 36-41.

[3] Bharadiya, J. P. (2023). AI-driven security: how machine learning will shape the future of cybersecurity and web 3.0. American Journal of Neural Networks and Applications, 9(1), 1-7.

[4] Chaganti, K. C. (2025). A Scalable, Lightweight AI-Driven Security Framework for IoT Ecosystems: Optimization and Game Theory Approaches. IEEE Access.

[5] Kolade, T. M., Obioha Val, O., Balogun, A. Y., Gbadebo, M. O., & Olaniyi, O. O. (2025). AI-driven open source intelligence in cyber defense: A double-edged sword for national security. Adebayo Yusuf and Gbadebo, Michael Olayinka and Olaniyi, Oluwaseun Oladeji, AI-Driven Open Source Intelligence in Cyber Defense: A Double-edged Sword for National Security(January 18, 2025).

[6] Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. ACM computing surveys (CSUR), 41(3), 1-58.

[7] Pang, G., Shen, C., Cao, L., & Hengel, A. V. D. (2021). Deep learning for anomaly detection: A review. ACM computing surveys (CSUR), 54(2), 1-38.

[8] Bhuyan, M. H., Bhattacharyya, D. K., & Kalita, J. K. (2013). Network anomaly detection: methods, systems and tools. Ieee communications surveys & tutorials, 16(1), 303-336.

[9] Ahmed, M., Mahmood, A. N., & Hu, J. (2016). A survey of network anomaly detection techniques. Journal of Network and Computer Applications, 60, 19-31.

[10] Zenati, H., Romain, M., Foo, C. S., Lecouat, B., & Chandrasekhar, V. (2018, November). Adversarially learned anomaly detection. In 2018 IEEE International conference on data mining (ICDM) (pp. 727-736). IEEE.

[11] Mehrotra, K. G., Mohan, C. K., Huang, H., Mehrotra, K. G., Mohan, C. K., & Huang, H. (2017). Anomaly detection (pp. 21-32). Springer International Publishing.

[12] Chalapathy, R., & Chawla, S. (2019). Deep learning for anomaly detection: A survey. arXiv preprint arXiv:1901.03407.

[13] Liu, F. T., Ting, K. M., & Zhou, Z. H. (2012). Isolation-based anomaly detection. ACM Transactions on Knowledge Discovery from Data (TKDD), 6(1), 1-39.

[14] Song, X., Wu, M., Jermaine, C., & Ranka, S. (2007). Conditional anomaly detection. IEEE Transactions on knowledge and Data Engineering, 19(5), 631-645.

[15] Patcha, A., & Park, J. M. (2007). An overview of anomaly detection techniques: Existing solutions and latest technological trends. Computer networks, 51(12), 3448-3470.

[16] Baheti, R., & Gill, H. (2011). Cyber-physical systems. The impact of control technology, 12(1), 161-166.

[17] Wolf, W. (2009). Cyber-physical systems. Computer, 42(03), 88-89.

[18] Jazdi, N. (2014, May). Cyber physical systems in the context of Industry 4.0. In 2014 IEEE international conference on automation, quality and testing, robotics (pp. 1-4). IEEE.

[19] Alguliyev, R., Imamverdiyev, Y., & Sukhostat, L. (2018). Cyber-physical systems and their security issues. Computers in Industry, 100, 212-223.

[20] Lee, E. A. (2015). The past, present and future of cyber-physical systems: A focus on models. Sensors, 15(3), 4837-4869.

[21] Lee, E. A. (2008, May). Cyber physical systems: Design challenges. In 2008 11th IEEE international symposium on object and

component-oriented real-time distributed computing (ISORC) (pp. 363-369). IEEE.

[22] Zanero, S. (2017). Cyber-physical systems. Computer, 50(4), 14-16.

[23] Liu, Y., Peng, Y., Wang, B., Yao, S., & Liu, Z. (2017). Review on cyber-physical systems. IEEE/CAA Journal of Automatica Sinica, 4(1), 27-40.

[24] Duo, W., Zhou, M., & Abusorrah, A. (2022). A survey of cyber attacks on cyber physical systems: Recent advances and challenges. IEEE/CAA Journal of Automatica Sinica, 9(5), 784-800.

[25] Derler, P., Lee, E. A., & Vincentelli, A. S. (2011). Modeling cyber–physical systems. Proceedings of the IEEE, 100(1), 13-28.

[26] Nicki, P. (2022). Real-Time Threat Intelligence: AI-Based Approaches to Cyber Risk Prediction and Mitigation.

[27] Gudala, L., Shaik, M., & Venkataramanan, S. (2021). Leveraging machine learning for enhanced threat detection and response in zero trust security frameworks: An Exploration of Real-Time Anomaly Identification and Adaptive Mitigation Strategies. Journal of Artificial Intelligence Research, 1(2), 19-45.

[28] Kostopoulos, D., Tsoulkas, V., Leventakis, G., Drogkaris, P., & Politopoulou, V. (2013). Real time threat prediction, identification and mitigation for critical infrastructure protection using semantics, event processing and sequential analysis. In Critical Information Infrastructures Security: 8th International Workshop, CRITIS 2013, Amsterdam, The Netherlands, September 16-18, 2013, Revised Selected Papers 8 (pp. 133-141). Springer International Publishing.

[29] Mahesh, B. (2020). Machine learning algorithms-a review. International Journal of Science and Research (IJSR).[Internet], 9(1), 381-386.

[30] Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255-260.

[31] El Naqa, I., & Murphy, M. J. (2015). What is machine learning?. In Machine learning in radiation oncology: theory and applications (pp. 3-11). Cham: Springer International Publishing.

[32] Mitchell, T. M., & Mitchell, T. M. (1997). Machine learning (Vol. 1, No. 9). New York: McGraw-hill.