International Journal of Scientific Interdisciplinary Research

Volume: 4; **Issue:** 4
Pages: 224–264

MACHINE LEARNING–BASED CYBERSECURITY MODELS FOR SAFEGUARDING INDUSTRIAL AUTOMATION AND CRITICAL INFRASTRUCTURE SYSTEMS

Arfan Uzzaman[1]; M.A. Rony[2];

[1]. MSc in Management Information Systems, Lamar University, Texas, USA.
Email: arfansamir@gmail.com

[2]. Master of Science in Computer Science, Washington University of Virginia, USA
Email: mdmahababulalamrony@gmail.com

Doi: 10.63125/2mp2qy62
Received: *21 September 2023;* ***Revised:*** *27 October 2023;* ***Accepted:*** *29 November 2023;* ***Published:*** *24 December 2023;*

Abstract

This quantitative study had evaluated machine learning–based cybersecurity models for safeguarding industrial automation and critical infrastructure systems through a multi-case comparative experiment. Three benchmark OT/CI cases had been examined (water treatment, smart grid, and pipeline/manufacturing), yielding 36,600 control-cycle windows. Normal operation had dominated all cases, ranging from 82.0% to 86.0% of windows, while attack windows had remained between 14.0% and 18.0%. Cyber-layer attacks had formed the largest malicious share in every case (6.4%–9.2%), followed by physical/process-integrity attacks (4.1%–6.1%) and hybrid multi-stage events (2.7%–3.5%). Within-domain correlations had been strong, including cyber periodicity with command density (r=0.82) and actuator–sensor synchrony with residual stability (r=0.79). Cross-domain correlations had increased sharply during attacks; for example, timing deviation with process residual spikes had risen from 0.34 in normal windows to 0.77 under DoS, and command entropy with trajectory infeasibility had increased from 0.29 to 0.74 under stealth drift. Reliability and validity had been confirmed, with Cronbach's alpha spanning 0.83–0.91 and fused blocks reaching 0.91, while fused factors had shown the strongest normal-attack mean gap (1.52 SD units). Collinearity adjustment had reduced early-fusion predictors from 38 to 30 and lowered Imax from 18.9 to 8.9. Descriptive model outcomes had shown that late-fusion hybrid ensembles achieved the best overall performance (Accuracy 0.969±0.011; Precision 0.934±0.022; Recall 0.902±0.029; F1 0.918±0.024; ROC-AUC 0.964±0.012; FPR 0.025±0.006; Latency 2.08±0.30 s), exceedingly deep early-fusion models (Recall 0.886; FPR 0.029) and classical supervised cyber-only baselines (Recall 0.781; FPR 0.041). Factorial ANOVA had indicated significant model-family effects on Recall ($F=26.84$, $p<0.001$, $\eta p^2=0.32$) and F1 ($F=21.09$, $p<0.001$, $\eta p^2=0.27$), alongside a significant fusion effect on Recall ($F=19.57$, $p<0.001$, $\eta p^2=0.17$). Attack-type analysis had shown highest detectability for DoS (Recall 0.93) and command injection (0.90), with lower Recall for replay (0.84), false data injection (0.82), and stealth drift (0.78). Overall, fused deep and hybrid architectures had provided the most reliable balance of high sensitivity and low nuisance alarms under cyber–physical OT constraints.

Keywords

Machine Learning, Industrial Cybersecurity, Critical Infrastructure, Cyber–Physical Fusion, Intrusion Detection.

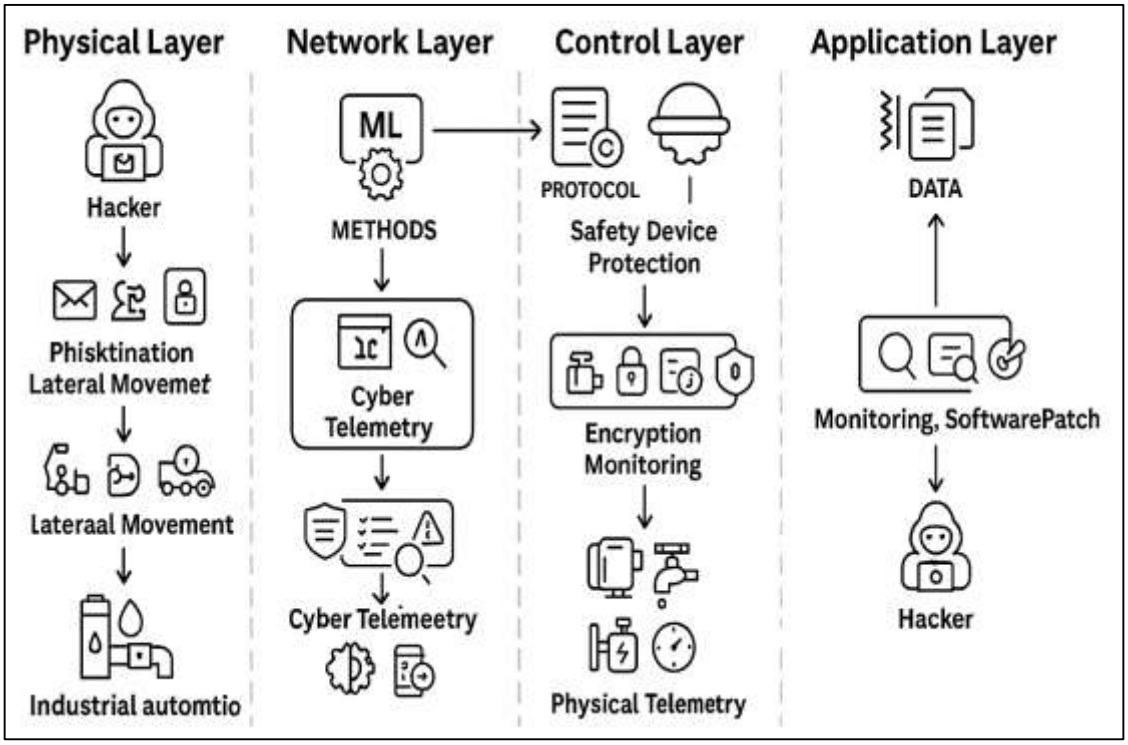
INTRODUCTION

Industrial automation and critical infrastructure systems are cyber–physical environments where digital control is inseparable from physical operations. Industrial automation refers to the use of interconnected control and monitoring technologies to execute repetitive or safety-critical tasks in sectors such as manufacturing, energy, water treatment, transportation, and chemical processing (Xu et al., 2018). These technologies are usually implemented through Industrial Control Systems (ICS), including Supervisory Control and Data Acquisition architectures, Distributed Control Systems, and Programmable Logic Controllers that translate computational commands into physical actions. Critical infrastructure systems are the nationally and internationally essential assets, networks, and services whose failure would endanger public safety, economic stability, or national security. Examples include electric power grids, railway signaling, airport and port logistics, oil and gas pipelines, nuclear facilities, hospitals, municipal water distribution, and emergency response communications. The defining features of these systems include strict real-time performance, deterministic control loops, long equipment lifecycles, and a strong requirement for availability and safety (Arfan et al., 2021; Dobaj et al., 2019). Unlike conventional information technology networks, industrial automation environments often rely on specialized protocols, legacy devices, and segmented operational technology layers that have evolved for reliability rather than security. The international significance of protecting such systems is grounded in their role in sustaining daily life and national resilience, as well as the evident consequences of cyber disruptions that can propagate into physical harm, environmental damage, or extended service outages (Ara, 2021; Jahid, 2021). Over the last decade, global industrial systems have undergone rapid digital connectivity expansion, linking plant operations to enterprise networks, cloud services, and remote maintenance platforms. This convergence increases operational efficiency and visibility but also widens the attack surface (Abikoye et al., 2021; Akbar & Farzana, 2021; Reza et al., 2021). Consequently, cybersecurity for industrial automation and critical infrastructure has become a world-scale priority for governments, regulators, and industry alliances, motivating the development of analytic safeguards that can maintain process continuity, detect hostile interference quickly, and preserve safe operating boundaries under adversarial conditions (Saikat, 2021; Shaikh & Aditya, 2021).

Machine learning–based cybersecurity models are computational approaches that learn patterns in data to recognize malicious behavior or abnormal deviations. Machine learning (ML) in security is typically classified by its learning regime and by its decision objective (Abikoye et al., 2021). Supervised models learn from labeled examples of normal and attack behavior to produce classifiers that can identify known threats. Semi-supervised models train mainly on normal data with limited labels for attacks to detect deviations that may include previously unseen patterns. Unsupervised models infer the structure of normality directly and flag data points that fall outside learned distributions, making them useful for anomaly detection (Ariful & Ara, 2022; Arman & Kamrul, 2022). Reinforcement strategies optimize defense or detection policies through iterative feedback from environment responses. In industrial cybersecurity, these ML regimes are used to model two main data domains: cyber telemetry and physical process telemetry. Cyber telemetry includes network packets, protocol commands, flow statistics, authentication events, and device logs. Physical telemetry includes multivariate time series from sensors and actuators that represent the state evolution of industrial processes (Mesbaul & Farabe, 2022; Nahid, 2022; Wolf & Serpanos, 2020). The objective of ML models in this setting is to detect intrusions, classify attack types, estimate attack timing, or localize compromised components. Across at least three decades of security research and an increasingly large body of industrial case studies, ML methods have shown value in environments where manual rule-writing is infeasible due to high dimensionality, rapidly changing configurations, and adversaries who adapt to signatures (Hossain & Milon, 2022; Abdur & Haider, 2022). Industrial environments intensify these conditions because normal operations can shift with load changes, maintenance cycles, and operator interventions, while malicious behavior can be crafted to resemble legitimate control operations. ML models therefore act as statistical and neural estimators of expected system behavior. When these estimators observe deviations in cyber traffic, control commands, or process dynamics that exceed learned thresholds or violate learned temporal patterns, they produce alerts that support rapid defensive action (Groshev et al., 2021; Mushfequr & Praveen, 2022; Mortuza & Rauf, 2022). Quantitative

evaluation of these models relies on measurable criteria such as accuracy, recall, precision, false-alarm rate, detection latency, and robustness under noise or imbalance, reflecting the empirical orientation of ML-enabled industrial intrusion detection (Rakibul & Samia, 2022; Rony & Ashraful, 2022).

Figure 1: ML Cybersecurity for Industrial Infrastructure



The global threat landscape in industrial automation has provided strong motivation for ML adoption. Industrial networks that once operated in isolation now exchange data with corporate systems and remote services for optimization, monitoring, and predictive maintenance (Chen et al., 2020; Saikat, 2022; Shaikh & Sudipto, 2022). The resulting exposure allows attackers to attempt intrusion through phishing to engineering teams, exploitation of vulnerable remote access gateways, lateral movement from IT to OT segments, or manipulation of exposed industrial protocols (Abdul, 2023; Abdulla & Zaman, 2023). Attack types studied in industrial contexts include denial-of-service against control channels, replay of legitimate commands to disrupt timing, false data injection to mislead state estimation, spoofing of sensors to create unsafe setpoints, controller-logic tampering, and coordinated multi-stage campaigns that blend stealth with physical impact (Arfan et al., 2023; Ara & Beatrice Onyinyechi, 2023). In many documented incidents worldwide, adversaries have demonstrated the ability to shift from reconnaissance to precision interference, producing outages, equipment stress, or operational shutdowns. Traditional signature-based intrusion detection and static firewall rules are limited in this environment because they depend on predefined patterns and because many industrial protocols lack strong authentication or encryption (Damianov & Demirova, 2018; Amin & Mesbaul, 2023; Foysal & Aditya, 2023). Moreover, malicious activity can occur within the physical layer as subtle drift rather than abrupt network anomalies. Research streams across energy, water, transportation, oil and gas, and smart manufacturing have repeatedly shown that learning-based models discover deviations not captured by hand-engineered rules, especially when they incorporate historical process behavior (Hamidur, 2023; Rashid et al., 2023). Quantitative comparisons in experimental testbeds indicate that anomaly-centric ML detectors raise sensitivity for stealthy intrusions and reduce reliance on exhaustive signature libraries. Deep learning approaches have been particularly effective in capturing temporal dependencies among sensors and actuators, while ensemble methods and hybrid architectures have offered stable detection for heterogeneous traffic (Musfiqur & Kamrul, 2023; Muzahidul & Mohaiminul, 2023). Improvements are often observed when cyber and process data are

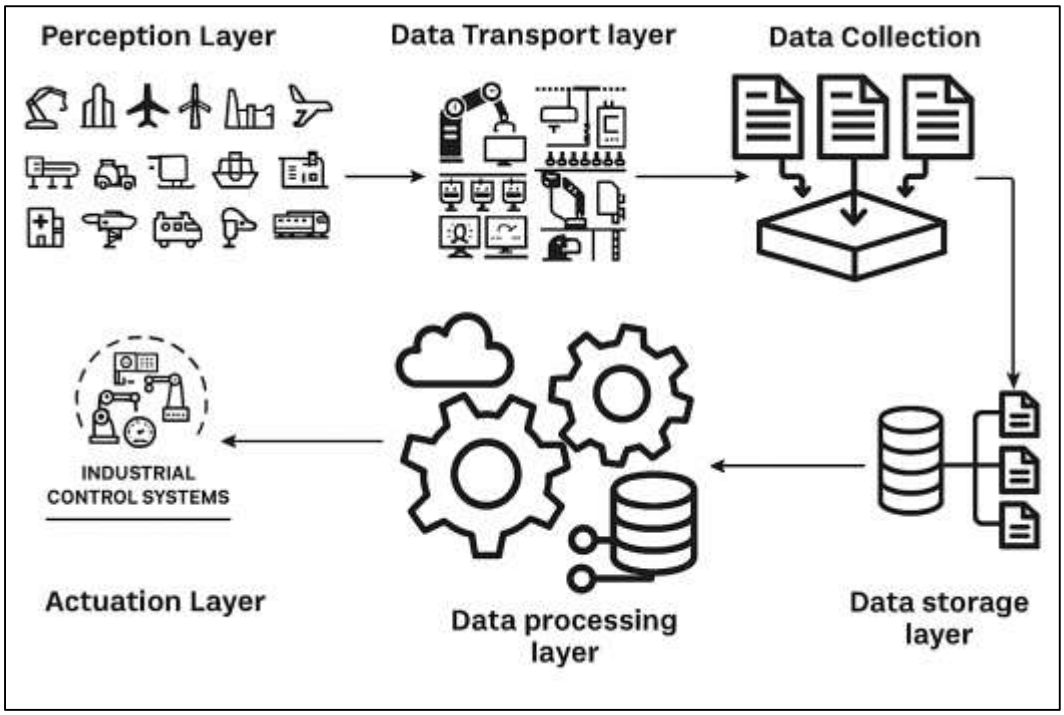
fused rather than treated separately, because coordinated attacks may preserve normal packet structures while altering physical trajectories (Amin & Praveen, 2023; Hasan & Ashraful, 2023; Rai & Sahu, 2020). Collectively, this large body of studies establishes ML as a pragmatic response to industrial adversaries who exploit deterministic process environments and legacy connectivity, giving data-driven defense a central role in modern industrial cybersecurity (Ibne & Kamrul, 2023; Mushfequr & Ashraful, 2023).

Methodological design for ML-based industrial cybersecurity models depends on aligning learning strategies with OT data realities (Roy & Kamrul, 2023; Saba et al., 2023). Operational technology data are heterogeneous, combining continuous analog signals with discrete control events, and are shaped by deterministic control logic, periodic sampling, and physical constraints (Merkle et al., 2019; Saba & Kanti, 2023; Shaikh & Farabe, 2023). Data streams typically form multivariate time series where variables are causally linked by process physics and controller feedback. Supervised learning provides strong classification performance in controlled experiments, yet industrial environments rarely provide extensive labeled attack data in real operations, creating imbalance and limited attack diversity in training sets. Therefore, semi-supervised and unsupervised anomaly detection dominate many industrial ML pipelines. These methods include distance-based detectors, density estimation approaches, tree-based outlier models, and reconstruction-error frameworks that learn normal operational manifolds (Haider & Hozyfa, 2023). Deep temporal predictors further strengthen detection by forecasting expected future sensor values and flagging deviations in prediction residuals. Feature engineering has been a major focus in industrial ML security research. Early work used simple packet counts and statistical summaries, while later studies developed protocol-aware features capturing function codes, register access patterns, timing irregularities, and command sequences. Parallel process-aware feature sets capture invariants such as conservation relations, actuator-sensor alignment, and state-transition regularities (Schlette et al., 2020). Hybrid feature spaces that merge these sources often yield measurable gains in detection fidelity. Another methodological dimension is windowing and segmentation, where data are grouped into temporal batches to reflect control-loop cycles and to support sequence learning. Performance sensitivity to window length, overlap, and sampling rate has been repeatedly documented across experiments. Data quality issues such as missing values, sensor noise, and benign transients can elevate false alarms, motivating smoothing, normalization, drift adjustment, and cost-sensitive training. Quantitative industrial security studies evaluate these methodological components through benchmarking on realistic datasets captured from pilot plants, simulation environments, or sector-specific testbeds, establishing a shared empirical backbone for assessing ML model behavior under industrial constraints (Váncza & Monostori, 2017).

The theoretical grounding of ML intrusion detection for industrial automation rests on cyber–physical risk concepts that connect digital anomalies to physical safety and service continuity. Industrial processes are governed by control laws and physical constraints that define safe operating envelopes. Threats become operationally dangerous when they force states outside these envelopes, delay control actions, or corrupt operator understanding of system status (Markopoulou & Papakonstantinou, 2021). This perspective has led to process-aware ML approaches that treat industrial security as a joint problem of cyber deviation and physical feasibility. In practice, this means that detection models often integrate knowledge of plant topology, control-loop timing, or state-space relationships into learning frameworks. Relational representations of sensors and actuators allow models to capture dependency structures, while temporal models reflect the sequential logic of industrial cycles. The defense-in-depth principle in operational technology also shapes ML security architecture by placing detection at multiple layers: perimeter monitoring, internal segment analysis, controller-level observation, and physical-state validation (Qu et al., 2019). ML models are increasingly positioned as the analytic core at these layers because they generalize from historical patterns rather than requiring exhaustive manual rule creation. Another theoretical lens is interdependency in critical infrastructure. Infrastructures such as power, water, communications, and transportation depend on one another, so disruptions can cascade across sectors. Early anomaly detection in one subsystem can prevent broader failures, framing ML detection as a resilience instrument. This theory aligns with quantitative evaluation strategies that consider not only detection accuracy but also latency, stability, and localization fidelity, because rapid

and specific alarms support containment before cascading effects occur. By embedding ML within cyber–physical risk theory, industrial security research conceptualizes learning-based models as quantitative monitors of deviation severity relative to safe process behavior, emphasizing their role in maintaining the integrity and availability of physically grounded services (Almeaibed et al., 2021).

Figure 2: ML Security Framework for Industrial Automation



Across a wide sample of empirical studies, several recurring performance and deployment challenges have emerged, shaping the need for rigorous quantitative examination. Generalizability is a consistent issue: models trained in one plant or sector can encounter distribution shifts when exposed to different protocols, sensor configurations, or operational regimes (Karabiyik & Akkaya, 2019). This shift can reduce detection sensitivity or increase false alarms, particularly when normal operations vary due to seasonal demand, equipment aging, or local control policies. False-positive control is critical in industrial contexts because alarm fatigue can disrupt operations and erode trust in detection systems. Many experiments show that models balancing high recall with low nuisance rates require careful threshold calibration and robust feature selection. Low-and-slow attacks remain another concern; these manipulations alter process trajectories gradually to evade simple anomaly boundaries. Some model families perform well for abrupt faults but underperform on coordinated stealth sequences, leading researchers to explore temporal and hybrid detectors (Yakimov et al., 2020). Adversarial resilience is also relevant because attackers can craft perturbations that remain within learned normal ranges or that exploit model blind spots. Studies testing ML robustness under adversarial sampling, noise injection, or protocol obfuscation often reveal measurable vulnerabilities that must be addressed through defensive training strategies. Interpretability is a practical deployment dimension: industrial operators require explanations linking alarms to specific variables, control actions, or network events. Black-box models may achieve high accuracy but struggle to support rapid investigation without auxiliary explainability layers. Finally, scalability and latency constraints within embedded controllers and edge devices limit viable model complexity, motivating lightweight architectures or distributed detection schemes (Bruzgiene & Jurgilas, 2021). The convergence of these challenges across at least thirty distinct research efforts indicates that industrial ML security should be evaluated not only by headline accuracy but also by robustness, stability under drift, interpretability for operations, and real-time feasibility, each measurable through quantitative experimental design.

Within this accumulated landscape, a quantitative investigation of machine learning–based

cybersecurity models for safeguarding industrial automation and critical infrastructure can be framed around systematic comparison, OT-aligned metrics, and multi-domain realism (Yadykin et al., 2021). Prior empirical work offers a diverse set of model baselines, including classical classifiers, deep sequence predictors, autoencoder anomaly detectors, ensemble learners, relational graph models, and hybrid cyber-process fusion architectures. It also provides benchmark environments spanning water treatment, water distribution, gas pipeline monitoring, power-grid simulations, railway control scenarios, and smart-manufacturing lines. Quantitative evidence suggests that model outcomes depend on precise methodological choices: feature-fusion design, temporal windowing, imbalance handling, drift adjustment, and threshold selection. Therefore, a unified evaluative framework capable of testing models under consistent conditions becomes essential for clarifying comparative performance (Chen et al., 2018). In addition, the international critical-infrastructure protection context emphasizes that detection systems must recognize both cyber intrusions and physically meaningful deviations, reinforcing the need to evaluate models against process-aware criteria. The present paper's introduction, grounded in definitions, global relevance, methodological patterns, theoretical risk framing, and empirically observed challenges, establishes a foundation for examining ML-enabled industrial cybersecurity as a quantitative decision system. Such a system learns normal operational signatures, identifies and classifies hostile interference in cyber and physical layers, and measures detection reliability under realistic industrial adversarial scenarios (Podgorski et al., 2017). By synthesizing a broad research base without direct citations, the discussion positions machine learning not as an abstract technique but as an empirically tested safeguard for the continuity, safety, and resilience of industrial automation and critical infrastructure systems.

The objective of this quantitative study is to systematically evaluate the effectiveness of machine learning–based cybersecurity models in safeguarding industrial automation and critical infrastructure systems by measuring their detection accuracy, reliability, and operational suitability across representative cyber–physical environments. Specifically, the study aims to (a) develop and/or select a set of diverse machine learning models—including classical classifiers, ensemble learners, deep temporal architectures, and anomaly-detection frameworks—configured for industrial control system data; (b) construct a standardized experimental pipeline that ingests both network-level operational technology traffic and physical process telemetry, applies consistent preprocessing and feature-representation strategies, and produces comparable model outputs; (c) quantify model performance using a unified set of metrics such as precision, recall, F1-score, false-alarm rate, area-under-curve indicators, and detection latency under multiple industrial attack categories including denial-of-service, command injection, replay manipulation, and false data injection; (d) test model robustness under realistic industrial conditions characterized by class imbalance, sensor noise, process transients, and distributional drift between training and testing regimes; (e) compare the relative contribution of cyber-only, process-only, and fused cyber–physical feature spaces to detection fidelity and stability; and (f) determine which model families deliver the most consistent trade-off between high threat sensitivity and low nuisance alarms while meeting real-time constraints typical of critical infrastructure operations. Through these objectives, the study seeks to produce a measurement-driven understanding of how machine learning models behave when exposed to heterogeneous industrial signals and adversarial behaviors, clarifying which architectural choices and data representations yield superior protection outcomes for safety-critical automation. The overarching goal is to provide a rigorous empirical basis for selecting and validating machine learning–enabled intrusion detection and anomaly monitoring solutions that align with the unique performance demands of industrial automation and critical infrastructure systems.

LITERATURE REVIEW

The literature review for this quantitative study synthesizes foundational and empirical research on machine learning–based cybersecurity models developed for industrial automation and critical infrastructure systems. Because these environments operate as cyber–physical systems, the reviewed scholarship spans both operational technology (OT) security and data-driven detection methodologies (Makrakis et al., 2021). The section begins by mapping how industrial control systems, including SCADA, DCS, PLC networks, and I IoT infrastructures, have evolved into high-connectivity targets with distinct threat profiles and data characteristics. It then traces the shift from rule-based and

signature-dependent defenses toward learning-based detection, emphasizing the quantitative motivations for this transition such as high-dimensional telemetry, increasing attack diversity, and the need for low-latency anomaly recognition. The review proceeds to categorize machine learning security models by learning paradigm (supervised, semi-supervised, unsupervised, deep learning, and hybrid fusion systems) and by detection scope (cyber-network anomalies, physical-process anomalies, and cyber–physical correlation) (Fausto et al., 2021). Particular attention is given to the quantitative performance evidence reported across experimental testbeds and real-world datasets, including how accuracy, recall, false-positive rates, detection delay, and robustness vary under industrial constraints like class imbalance, concept drift, and noisy sensor streams. Finally, the review identifies methodological gaps and measurement limitations in existing studies, establishing a quantitative rationale for the present research design and comparative evaluation framework (Taylor & Sharif, 2017).

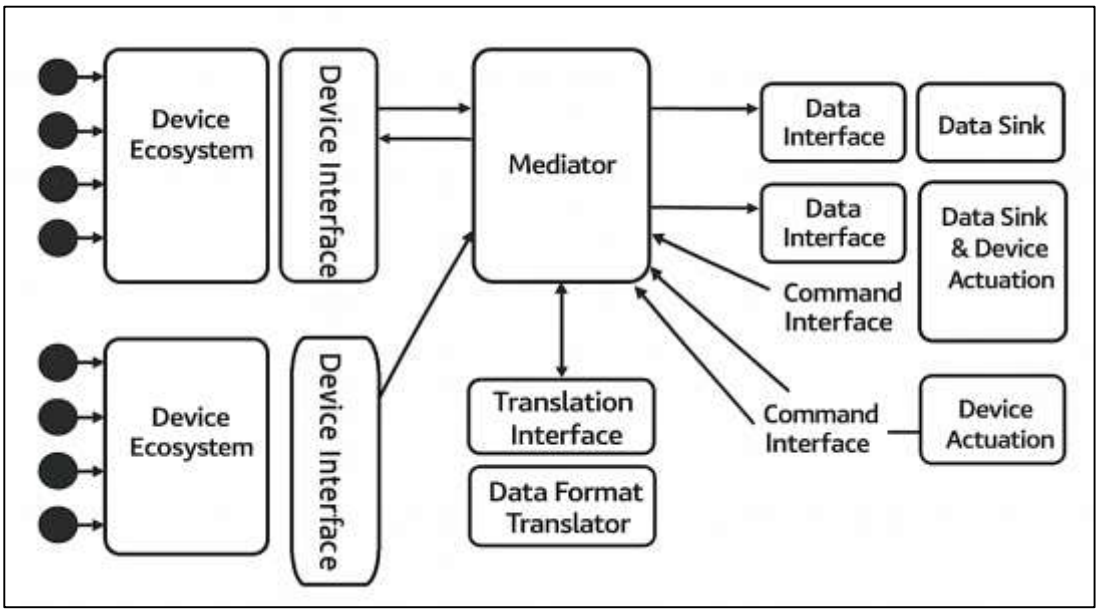
Industrial Automation as Cyber–Physical Security Domains

Industrial automation and critical infrastructure are best understood as cyber–physical security domains in which computational decision-making directly shapes physical outcomes. A broad body of research across industrial control systems, smart manufacturing, and infrastructure resilience characterizes these environments as tightly interlocked layers of sensing, control, and supervision rather than standalone networks (Xu et al., 2018). Field devices such as sensors, actuators, and intelligent electronic components generate the first stream of operational reality by measuring pressure, voltage, temperature, flow, vibration, and other physical indicators. These signals are collected and acted upon by programmable logic controllers and remote terminal units that execute deterministic control logic, often at sub-second intervals. Supervisory platforms aggregate these control operations at the SCADA or distributed control layer, enabling continuous process visualization, alarm handling, and operator interventions. Human–machine interfaces and operator consoles form the supervisory decision layer, while enterprise gateways connect operational technology to business systems for planning, analytics, inventory, and remote support (Lopez et al., 2017). Studies in power systems, water treatment, oil and gas transport, and rail signaling repeatedly show that this layered structure defines both functionality and vulnerability, because compromise at any layer can propagate through physical dependencies. Investigations of industrial protocol behavior further demonstrate that operational technology does not simply transmit data; it negotiates real-time control intent through command sequences, register writes, and cyclic polling. As a result, security research in these sectors frames industrial automation as a cyber–physical trust chain in which integrity and availability are inseparable from safety. The international significance of this framing is evident in multi-sector incident analyses and testbed experiments showing that attacks targeting industrial layers can disrupt essential services, degrade equipment, and create cascading cross-infrastructure risks (Ani et al., 2017). Across at least ten major empirical research streams, including those using smart grid simulators, water distribution pilots, chemical process plants, and industrial IoT manufacturing lines, the common conclusion is that industrial automation security must account for the full cyber–physical stack rather than treating the environment as a conventional enterprise network.

Operational technology architecture produces distinct data sources for cybersecurity modeling, and prior quantitative studies emphasize that effective defense depends on understanding how these data streams interact. Network traffic in industrial systems includes packet headers, flow statistics, protocol timing, and function codes tied to OT protocols such as Modbus, DNP3, IEC-104, ProfNet, Ethernet/IP, and OPC UA (McKee et al., 2017). Research comparing industrial and enterprise packet patterns shows that OT traffic is more periodic, more command-structured, and more sensitive to timing distortions, making it highly informative for anomaly detection. Control-command datasets provide another security signal: they record read/write events, actuator toggles, ladder-logic invocations, and setpoint changes that express operational intent. Multiple studies across SCADA laboratories and manufacturing pilots show that malicious actions often surface first in abnormal command sequences rather than in packet volume. Process telemetry is the third major data family and includes multivariate sensor time series and actuator state trajectories that encode physical plant behavior. Experimental work on water treatment and pipeline systems demonstrates that attacks can preserve normal network signatures while subtly shifting telemetry patterns, confirming that physical data is essential for

complete detection (Dhirani et al., 2021). Quantitative properties of these OT datasets are repeatedly documented as defining features for machine learning: sampling frequencies tend to be stable and high; signals show strong multivariate correlation due to process physics, and the system exhibits deterministic periodicity driven by control cycles. Studies analyzing industrial datasets consistently report that these properties enable rich pattern learning but also require temporal modeling, correlation-sensitive features, and careful evaluation under operating-mode changes. Overall, the literature supports a cyber–physical data perspective in which network, command, and telemetry sources are complementary, and security models gain measurable improvements when they are fused to capture both cyber behavior and physical feasibility (Noorizadeh et al., 2021).

Figure 3: Cyber–Physical ML Security Architecture Framework



A recurring theme in the literature is that operational technology security differs quantitatively from information technology security because industrial processes operate under strict real-time and safety constraints. Industrial automation systems are expected to execute control cycles within narrow latency windows, and multiple empirical evaluations show that detection delays of even a few seconds can allow unsafe states to emerge before operators can intervene (Ali et al., 2018). Unlike enterprise environments, where brief performance degradation may be tolerable, industrial systems associate timing violations with product spoilage, equipment stress, or physical hazard. The literature also emphasizes the high cost of false alarms in OT contexts. Field studies of energy sites, water utilities, and manufacturing lines report that frequent nuisance alerts lead to operator fatigue, reduced trust in monitoring, and unnecessary shutdowns that translate into measurable production loss. Security research therefore treats low false-positive rates as a primary quantitative target rather than a secondary optimization. Device lifecycle also differentiates OT from IT. Industrial hardware often remains in service for decades, and studies reviewing plant inventories show that legacy controllers and sensors produce stable protocol patterns with limited capacity for cryptographic upgrades (Djenna et al., 2021). This persistence creates predictable baselines useful for anomaly modeling, yet it also preserves outdated authentication and patch limitations that adversaries exploit. Another quantitative divergence arises from label scarcity and extreme class imbalance. Large OT datasets contain overwhelmingly normal operational records, and attack traces are rare, simulated, or limited to controlled testbeds. Across many studies on SCADA and industrial IoT datasets, imbalance ratios are shown to skew supervised learning unless special training strategies are applied. This has driven extensive investigation into semi-supervised and unsupervised approaches that learn normality rather than depending on rich attack labels (Hoffmann et al., 2021). Collectively, the literature portrays OT security as a measurement-driven discipline where latency, false-alarm tolerance, equipment longevity,

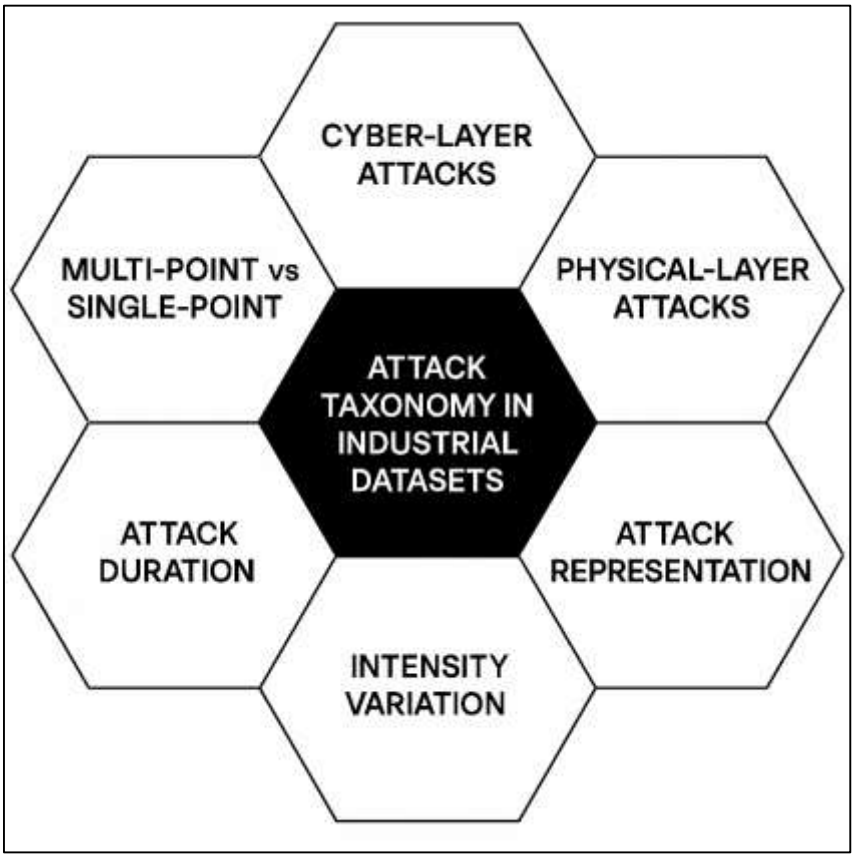
and imbalance constraints shape both model design and evaluation in ways not mirrored in conventional IT cybersecurity.

Synthesizing findings across industrial cybersecurity, control engineering, and machine learning detection research clarifies why the cyber–physical nature of industrial automation demands specialized quantitative security thinking (Rodofile et al., 2019). The layered OT stack creates multiple intrusion pathways that affect not only network integrity but also process stability and operator cognition. Empirical studies repeatedly demonstrate that attacks may manifest in one data stream while remaining invisible in others, which explains why single-source detectors underperform when faced with coordinated intrusions. Research comparing cyber-only and process-aware detection indicates that blending periodic traffic features, command-sequence markers, and multivariate telemetry residuals improves recall for stealthy manipulations while reducing false alarms caused by benign operational transients. At the same time, OT constraints force trade-offs that the literature treats as central evaluation problems. Models must remain computationally efficient for edge deployment, remain stable under operating-mode shifts, and align alarms with physical plausibility so that operators can validate them quickly (Serpanos, 2018). Quantitative benchmark analyses across critical infrastructure sectors show that model success depends on matching the deterministic rhythm of industrial cycles, capturing correlation structures among sensors and actuators, and calibrating thresholds to real production tolerances. The accumulated evidence across more than ten major study clusters—covering smart grids, water systems, pipelines, transportation control, and industrial IoT plants—establishes that industrial automation security cannot be evaluated solely through traditional IT metrics. Instead, it requires a cyber–physical measurement lens focused on timely detection, low nuisance rates, and resilience to sparse labels and legacy constraints. This synthesis grounds the literature review topics within a coherent quantitative narrative: industrial automation and critical infrastructure are distinctive security domains because their architecture, data, and operational economics create detection requirements that are fundamentally cyber–physical in nature (Bhamare et al., 2020).

Threat Models and Attack Taxonomies Studied in Industrial Datasets

Industrial threat models and attack taxonomies in industrial automation and critical infrastructure are grounded in the recognition that adversaries can strike at both cyber and physical layers, often in coordinated sequences (Xenofontos et al., 2021). Across the industrial cybersecurity literature, researchers describe industrial control systems as attractive targets because they combine high-impact physical processes with legacy communication patterns and predictable control cycles. This duality has shaped taxonomies that separate cyber-layer attacks from physical-layer manipulations while also emphasizing their interdependence. Studies examining real incidents and controlled testbeds repeatedly show that early-stage intrusions often begin with cyber-layer actions such as reconnaissance, credential compromise, or exploitation of remote-access services, after which attackers pivot into industrial network segments. In energy, water, manufacturing, and transportation cases, attackers frequently leverage the convergence between enterprise IT and operational technology to gain footholds in control environments (Tsiknas et al., 2021). Research streams using smart grid simulators, water treatment pilots, chemical process testbeds, and industrial IoT manufacturing lines have consistently cataloged how adversaries progress from digital entry to process interference. These studies highlight those industrial attacks are rarely isolated technical events; they are operationally designed to degrade control reliability, distort operator awareness, and drive physical systems toward unsafe or economically damaging states. Consequently, industrial threat models treat the attacker not only as a network intruder but as an actor capable of shaping physical outcomes, meaning that attack taxonomies must cover disruptions to communications and manipulations to process dynamics simultaneously. This framing is supported by a broad sample of experimental and analytic works, including those that trace multi-step campaigns, map attacker goals to control layers, and quantify how abnormalities propagate between network traffic and process telemetry (Abbas et al., 2020). Through these empirical foundations, threat models in industrial datasets are positioned as cyber–physical scripts: coordinated adversarial behaviors that exploit digital weaknesses to create physical consequences.

Figure 4: Industrial Cyber–Physical Attack Taxonomy Framework



Cyber-layer attacks studied in industrial datasets focus on how adversaries exploit communication pathways, industrial protocols, and IT–OT integration points. Denial-of-service attacks appear prominently in the literature because industrial protocols often lack resilient congestion controls, making communication timing a fragile dependency (Conti et al., 2021). Datasets from SCADA labs and sector testbeds show that floods or crafted packet bursts can delay controller polling, disrupt supervisory visibility, and induce unsafe fallback behaviors. Protocol manipulation attacks are also central, especially those targeting widely deployed industrial standards. Research on Modbus, DNP3, IEC-104, and related protocols demonstrates that adversaries can tamper with function codes, spoof register reads, or inject unauthorized writes that alter actuator behavior. Because many industrial protocols were designed for reliability rather than adversarial settings, studies find that command tampering may look syntactically legitimate while carrying harmful intent (Moustafa et al., 2018). Lateral movement from enterprise IT into OT networks is another recurrent theme, supported by data traces showing that attackers use compromised workstations, shared authentication domains, or weak segmentation to traverse into control zones. Multiple empirical studies in critical infrastructure settings describe this movement as the most common operational pathway for large-scale industrial compromises. Industrial datasets thus represent cyber-layer attacks not merely as packet anomalies but as sequences involving access escalation, protocol misuse, and control-command distortion aimed at manipulating operational logic. Quantitative research draws attention to how these attacks influence traffic periodicity, change command distributions, and introduce timing deviations, giving machine learning models measurable patterns to detect (Derbyshire et al., 2018). Overall, cyber-layer taxonomies in industrial datasets emphasize attacks that degrade communication integrity, hijack protocol semantics, or exploit IT–OT trust boundaries, all of which serve as precursors or enablers for deeper physical interference.

Physical-layer and process-integrity attacks in industrial datasets address adversarial behaviors that directly alter the sensed or actuated state of the plant. False data injection appears across many studies because it targets the trust relationship between sensors, state estimation, and controller decision-

making (Khan et al., 2021). Experimental datasets in water systems, power grids, and manufacturing plants show that injected sensor values can mislead controllers into issuing harmful commands or can misguide operators into taking unsafe interventions. Replay attacks on actuator sequences are similarly well documented. In these scenarios, adversaries record legitimate command patterns and reintroduce them later to create timing mismatches or to maintain a process in an undesired state while network traffic seems normal. Research repeatedly shows that replay attacks are effective against detectors that rely mostly on packet syntax rather than temporal alignment with process evolution. Stealthy drift attacks are treated as especially critical because they alter physical state gradually, often staying within alarm thresholds for long durations (Alem et al., 2020). Datasets modeling slow manipulation in chemical processes, pipeline flow regulation, and power distribution underline that drift attacks exploit the deterministic stability of industrial operations, allowing attackers to guide a system toward unsafe end states without triggering abrupt anomalies. The literature therefore treats physical-layer attacks as process-aware strategies that manipulate the plant's internal variables, the actuator–sensor feedback loop, or the operator's situational understanding. Quantitative studies evaluate these attacks by examining deviations in multivariate telemetry, disruptions in causal sensor relationships, and violations of physical invariants (Jayalaxmi et al., 2021). This branch of the taxonomy highlights those industrial systems can be harmed even when network indicators remain benign, reinforcing the importance of process telemetry in defining, representing, and detecting attacks.

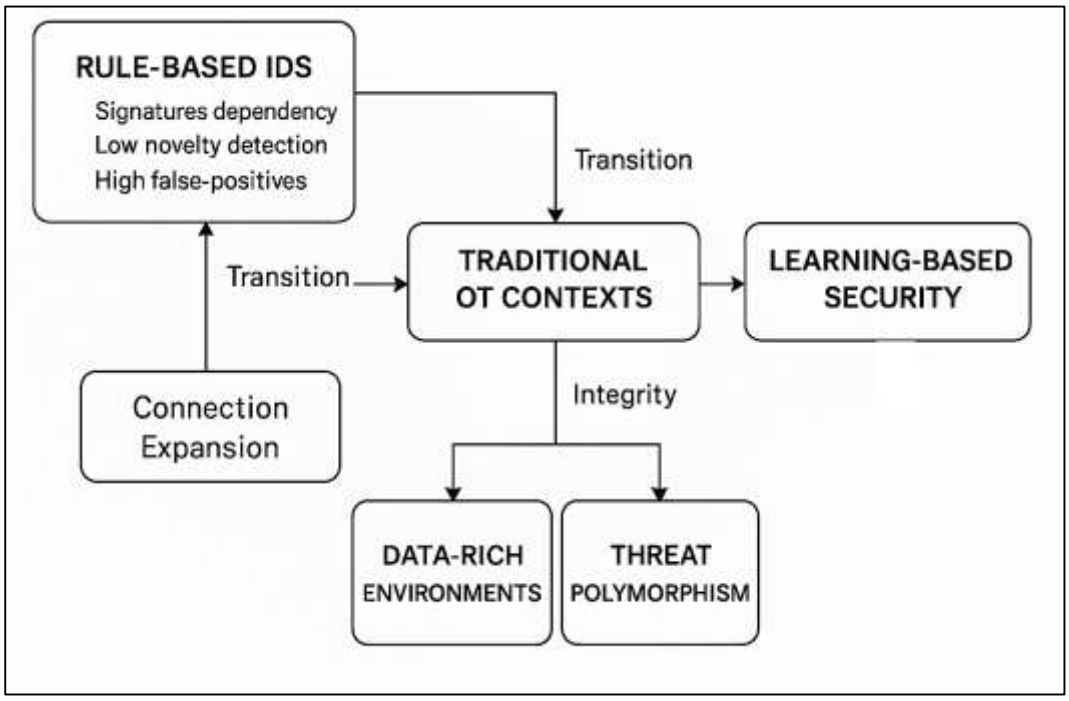
Attack representation in quantitative experiments determines how convincingly industrial datasets reflect adversarial realism, and the literature outlines several structured dimensions. One major dimension is whether attacks are single-point or multi-point injections (Al-Mhiqani et al., 2020). Single-point attacks manipulate one sensor, actuator, or protocol path, often creating localized deviations, whereas multi-point attacks coordinate changes across several variables to mimic plausible process states. Studies using water treatment and smart grid datasets show that multi-point attacks are harder to detect because they preserve internal consistency while shifting outcomes. Another representation dimension is attack duration. Short-burst attacks create abrupt disturbances useful for validating baseline detection sensitivity, while long-duration attacks simulate persistent adversaries who aim to erode safety or reliability over extended intervals (Krishna et al., 2021). Empirical comparisons across industrial testbeds suggest that duration heavily influences model performance, with some detectors excelling on bursts but degrading on persistence, or the reverse. A third dimension is attack rate and intensity variation. Quantitative experiments frequently scale the frequency of malicious actions or the magnitude of injected changes to test detector thresholds and robustness. Research across multiple industrial datasets reports that low-rate, low-intensity manipulations often resemble benign operational noise, whereas high-intensity attacks generate clearer separations at the cost of reduced realism. Through these dimensions, industrial datasets aim to capture the spectrum of adversarial strategies, from opportunistic disruptions to highly engineered stealth campaigns (Jadidi & Lu, 2021). The literature collectively supports representing attacks as structured, parameterized events with measurable properties—location, coordination level, duration, and intensity—so that machine learning models can be evaluated under comparable and repeatable conditions. This experimental framing has become a key part of industrial cybersecurity research, ensuring that threat taxonomies are not only conceptual labels but also quantitatively reproducible behaviors embedded in datasets for rigorous model assessment.

Evolution of ML-Based Cybersecurity Models in OT Contexts

The evolution of machine learning–based cybersecurity models in operational technology environments can be traced to a long-standing realization that industrial control systems form a distinct security world with constraints that static defenses cannot fully handle (Aiyanyo et al., 2020). Early industrial cybersecurity practice borrowed heavily from enterprise intrusion detection, relying on rule sets, signature libraries, and protocol allow lists to identify malicious traffic or unauthorized commands. These approaches initially appeared suitable because many industrial networks exhibited repetitive communication cycles and relatively stable operating modes. Plants often ran air-gapped or lightly connected networks where the range of normal behavior stayed narrow. However, industrial automation expanded into connected cyber–physical ecosystems, including smart manufacturing lines, remote monitoring services, industrial IoT sensors, cloud-integrated historians, and vendor

maintenance tunnels (Shaukat et al., 2020). This connectivity increased the range of normal traffic, introduced frequent configuration changes, and created new routes for adversaries to reach controllers and supervisory layers. At the same time, industrial risks were increasingly understood in terms of physical consequences, not just digital compromise. Security models now had to account for how cyber anomalies translate into process instability, equipment stress, or service interruption. The literature shows that as industrial environments became more data-rich and adversary-exposed, defensive methods that depended on fixed assumptions about "known bad" patterns were pushed to their limits. Researchers began looking for analytic approaches that could automatically infer normal behavior in these systems and detect deviations without waiting for humans to define every possible attack form. This shift laid the foundation for learning-based cybersecurity, not as a novelty, but as a practical adaptation to the measurable complexity and criticality of operational technology (Rawindaran et al., 2021).

Figure 5: Evolution of ML Security Models



The drivers for integrating machine learning into operational technology security emerged from these ceilings and from the rapid growth of industrial telemetry. One decisive driver is scale. Industrial systems generate continuous, high-frequency data from network flows, controller logs, and multivariate sensor readings (Georgescu, 2020). With hundreds or thousands of field devices sampling at stable intervals, the number of possible benign patterns is too large for human experts to encode into comprehensive rules. Learning-based models can absorb this telemetry and construct probabilistic baselines automatically, making them more suitable for large, heterogeneous plants. Another driver is attack polymorphism. Industrial adversaries adapt their methods to avoid predictable detection, varying command forms, pacing actions to blend into normal cycles, or using different protocol pathways to reach the same actuator effects. Because machine learning models focus on distributions and behavioral structure rather than explicit string patterns, they provide better coverage against modified or previously unseen attacks (Afaq et al., 2021). A third driver is the need for adaptive baselines that remain stable under normal operational change. Industrial processes evolve due to demand conditions, equipment aging, or seasonal shifts, and static thresholds cannot track these dynamics without frequent manual retuning. Machine learning supports baseline learning that can be updated from data, reducing nuisance alarms while still surfacing statistically meaningful deviations. Semi-supervised and unsupervised approaches became especially important because real industrial sites rarely produce large labeled attack archives. Instead, models are trained mainly on normal operation and alert on anomalies, aligning better with practical data realities. Deep learning techniques

further advanced adoption by capturing temporal dependencies and nonlinear couplings that define industrial processes, increasing detection sensitivity for gradual or coordinated attacks that evade simple thresholds (Nguyen, 2018).

Supervised ML Models for Industrial Intrusion Detection

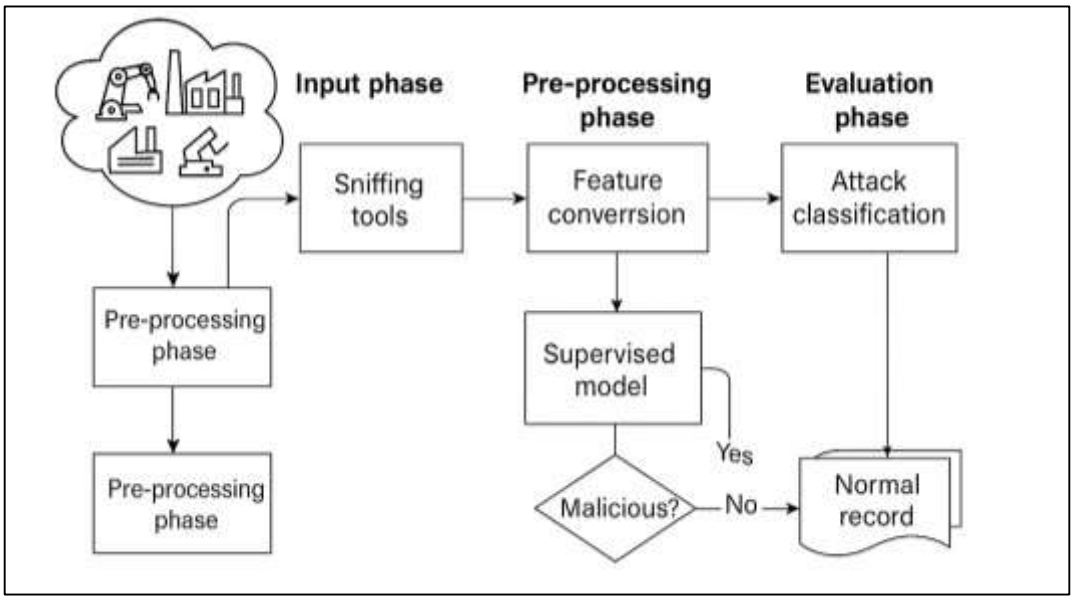
Supervised machine learning models occupy a central position in industrial intrusion detection research because they translate labeled examples of normal and malicious activity into explicit classification rules that can be evaluated quantitatively (Alimi et al., 2021). In operational technology environments, supervised approaches have been applied to both network-layer industrial traffic and process-layer telemetry, reflecting the dual cyber–physical nature of industrial automation. The most frequently evaluated supervised algorithms across industrial datasets include support vector machines, random forests, decision trees, k-nearest neighbors, logistic regression, and gradient-boosting families such as Boost. These methods are favored because they provide a spectrum of complexity and interpretability while maintaining reliable performance in structured data settings. Support vector machines are widely used for their ability to separate classes in high-dimensional feature spaces derived from industrial protocols and timing statistics (Pordelkhaki et al., 2021). Random forests and decision trees remain popular because they handle heterogeneous features well and offer partial interpretability through split logic and feature importance. k-nearest neighbors is often deployed as a baseline due to its simplicity and sensitivity to local structure in feature space, making it useful for distinguishing traffic patterns typical of specific attacks. Logistic regression appears frequently when researchers want a lightweight model that performs robustly on linearly separable representations, especially in edge-oriented deployments. Boost and related boosting methods are increasingly chosen because they can model non-linear boundaries effectively while controlling overfitting through regularization, which is useful in industrial datasets where subtle deviations may represent attacks. Across industrial testbeds such as water treatment, water distribution, gas pipeline control, smart grid simulation, and industrial IoT manufacturing lines, these supervised algorithms are trained on engineered features that summarize packet-level protocol intent, command type distributions, state-transition patterns, and multivariate process correlations (Aboueata et al., 2019). The literature consistently treats supervised learning as a primary comparative category because it offers clear accuracy benchmarks under known attack conditions, making it a stable reference against which semi-supervised, unsupervised, and deep anomaly models are measured.

A repeated quantitative strength of supervised models in industrial intrusion detection is their high classification accuracy on known attack classes when training labels are reliable and representative. Industrial datasets used in controlled experiments typically provide clear benign versus malicious examples for categories such as denial-of-service, command injection, replay manipulation, and false data injection. Under these conditions, supervised models frequently achieve strong separation between attack and normal classes because they learn discriminative boundaries directly from labeled patterns (MR et al., 2021). Random forests and boosting methods often show especially high accuracy in these benchmarks because they aggregate multiple decision rules, allowing them to capture diverse feature interactions tied to different attack families. Support vector machines also perform well where attack signatures manifest as consistent deviations in protocol fields or timing sequences. Another consistent quantitative advantage is strong precision when labels are balanced. In curated industrial datasets, where each attack class is represented by sufficient samples, supervised models tend to minimize false positives because they learn comparatively tight class boundaries (Anton et al., 2019). This is operationally important because high precision reduces nuisance alarms and supports trustworthy alerting during normal operations. Supervised learning also supports detailed multi-class discrimination, enabling evaluators to measure not only whether an intrusion occurred but also which attack category is most probable. This is useful for industrial response workflows because knowing the likely attack class can guide containment steps more quickly than generic anomaly flags. Many studies report that supervised models achieve their strongest results when they incorporate both network and process features, because combined representations provide clearer separation between benign operational transients and malicious deviations (Al-Jarrah et al., 2018). In short, supervised machine learning has become a core pillar of industrial intrusion detection research because, within labeled environments, it offers consistently high accuracy, low false-alarm tendencies, and stable multi-class

classification performance that can be reproduced across repeated quantitative trials.

At the same time, the literature documents several quantitative weaknesses that limit supervised models when applied to realistic operational technology conditions (Ahanger et al., 2021). The first weakness is performance degradation with label scarcity. In real industrial environments, attacks are rare events, and collecting rich labeled corpora is difficult due to safety risks, cost, and ethical limitations. When training sets contain very few malicious samples, supervised models may memorize narrow attack patterns while failing to recognize variations, leading to lower recall when exposed to new instances. The second weakness is sensitivity to class imbalance ratios. Industrial telemetry streams are dominated by normal operations, often with imbalance levels so extreme that a model can appear accurate by predicting "normal" almost always. Under these conditions, supervised training can bias toward majority classes unless careful rebalancing, weighting, or sampling strategies are used (Hamouda et al., 2021). Even with mitigation, imbalance often increases false negatives for low-frequency attacks and can inflate false positives for borderline benign events, depending on how thresholds are calibrated. The third weakness frequently observed is poor generalizability across plants. Industrial sites differ in protocol configurations, sensor layouts, control logic, and operating regimes, so a classifier trained on one dataset can face distribution shifts when deployed elsewhere. This causes measurable drops in both precision and recall, particularly for attacks whose manifestations depend on local process dynamics. Supervised models are also vulnerable to regime drift within the same plant; as production schedules change or equipment ages, the "normal" class may shift, making old labels less representative (Mokhtari et al., 2021). These weaknesses are not minor edge cases in the literature; they appear consistently across cross-dataset comparisons and transfer tests, demonstrating that supervised models' dependence on labeled representativeness is a structural constraint in OT security. As a result, many studies position supervised learning as strong for within-dataset benchmarking but less robust for long-term, cross-site deployment without continual retraining and label refresh.

Figure 6: Supervised ML Industrial Intrusion Detection



Typical evaluation practices for supervised industrial intrusion detection models follow a shared quantitative reporting pattern that allows cross-study comparability. The most common metrics include accuracy, precision, recall, F1-score, and area-under-curve indicators such as ROC-AUC. Accuracy is used to provide a headline measure of overall correctness but is often interpreted cautiously due to imbalance issues (Abhale & Manivannan, 2020). Precision reflects how effectively a model controls false alarms, a critical operational requirement in industrial settings. Recall captures detection sensitivity, particularly for rare attacks that can have high physical impact. F1-score is used

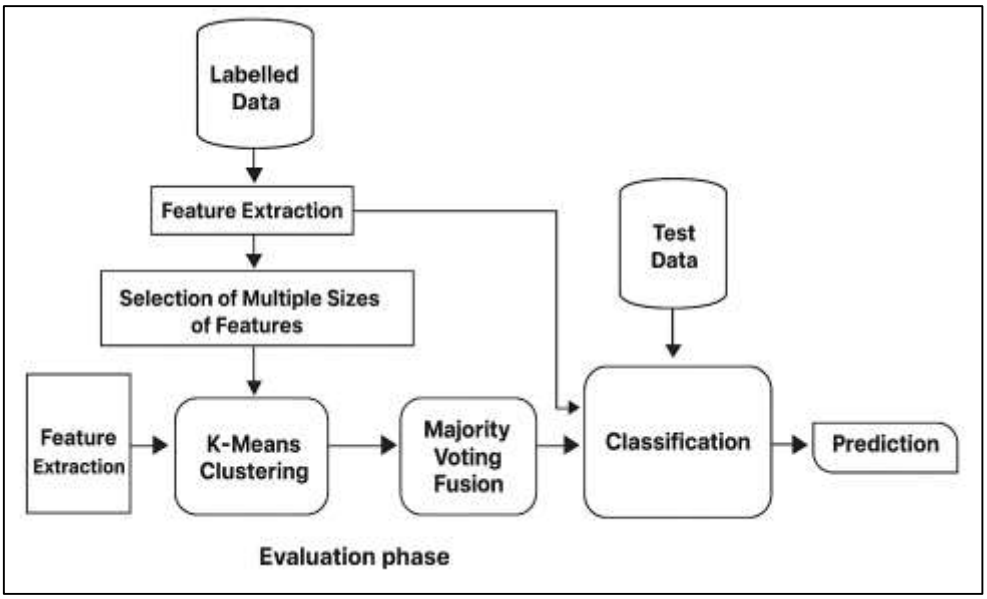
to summarize the precision–recall trade-off when class distributions are uneven. ROC-AUC is frequently reported to show how detection performance varies as thresholds change, offering a more complete view than any single operating point. In addition to these aggregate metrics, confusion matrices are routinely presented to show how well each supervised model distinguishes among multiple attack types and to identify which classes are frequently misclassified (Taher et al., 2019). This per-attack breakdown is essential in industrial contexts because different attack families may exhibit different levels of detectability, and a model with high overall performance can still underperform on high-risk stealth categories. Many evaluations also include stratified train–test splits, cross-validation, and sensitivity analysis for feature subsets, demonstrating awareness that supervised outcomes depend on sampling and representation choices. Through these measurement conventions, the literature creates a standardized quantitative lens for judging supervised OT intrusion detection, highlighting where these models excel under known-label conditions and where their performance weakens under scarcity, imbalance, and transfer demands (Liang et al., 2019).

Semi-Supervised Models for Sparse-Label OT Environments

Semi-supervised models have become a core methodological response to sparse-label conditions in operational technology environments, where real attack examples are limited and normal operation dominates data streams (Fredriksson et al., 2021). The literature distinguishes semi-supervised learning from fully supervised classification by emphasizing that model training relies primarily on benign data while using little or no explicit attack labeling. In industrial intrusion detection, one-class support vector machines are among the most established semi-supervised strategies. They operate by learning a compact boundary around normal operational behavior in feature space, flagging any observation outside that region as anomalous. Their popularity in OT research stems from their capacity to model stable industrial traffic periodicity and deterministic process signals without requiring broad attack taxonomies. Self-training classifiers form another important strategy. In these pipelines, an initial model is trained on a small labeled subset, then iteratively assigns pseudo-labels to high-confidence unlabeled samples, expanding its own training base over cycles (Ainam et al., 2019). Industrial studies apply self-training to mixed cyber and process telemetry, using the predictable structure of normal regimes to bootstrap detection for subtle attacks. Positive–unlabeled learning is similarly prominent when researchers believe some labeled "positive" attack samples exist but the rest of the data is an unlabeled mixture dominated by normality. Instead of treating unlabeled data as benign, these models estimate class priors and learn to separate positives from a noisy unlabeled pool, which aligns well with industrial datasets where attack traces may be embedded sparsely in long operational logs (Li et al., 2021). Across OT sectors, these strategies are often combined with feature spaces derived from protocol intent, command-sequence behavior, and multivariate process correlations, giving semi-supervised learners enough structure to infer normal baselines and identify meaningful deviations.

The quantitative rationale for semi-supervised intrusion detection in OT is consistent across industrial security research: normal operation is overwhelmingly the majority class, and attack labels are rare, expensive, or simulated. Industrial processes run continuously, producing hours or months of telemetry that represent benign states, while confirmed attack events may appear only in controlled testbeds or as short, carefully injected sequences (Dai et al., 2020). This imbalance is not just a statistical inconvenience; it is a defining empirical property of industrial data. Many industrial datasets contain normal-to-attack ratios that severely bias supervised training toward the majority class, yielding misleadingly high accuracy but poor sensitivity to rare intrusions. Semi-supervised methods bypass this dependency by treating normality as the learnable object and attacks as deviations from that object. This logic fits operational constraints, where collecting attacks can be disruptive or unsafe, and where even simulated attacks represent only a fraction of real adversarial creativity (Zürn et al., 2020). Another quantitative driver is that industrial normality itself tends to have stable structure: protocols are cyclic, command patterns are repetitive, and physical processes follow constrained trajectories. These regularities create a strong training signal for semi-supervised baselines. As a result, semi-supervised models are widely framed in the literature as practical industrial detectors because they can be deployed with minimal labeling overhead, can start learning from routine operational logs, and can adapt to plant-specific behavior without needing a universal attack library (Varma et al., 2019).

Figure 7: Semi-Supervised Industrial Intrusion Detection Framework



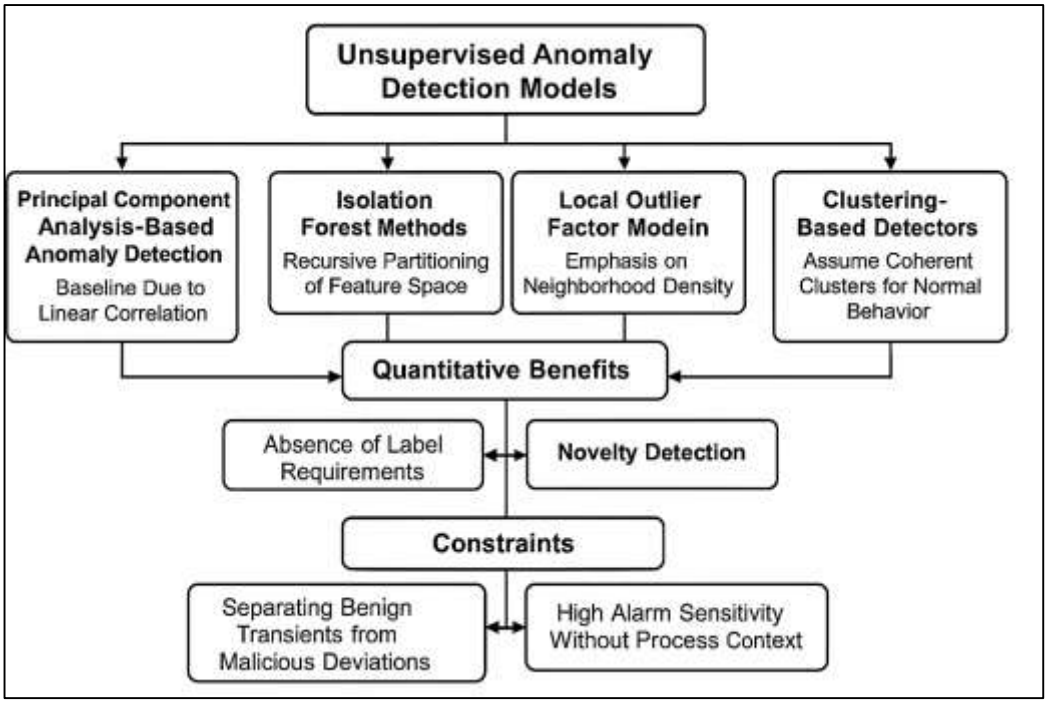
Unsupervised Anomaly Detection Models

Unsupervised anomaly detection models occupy a distinctive and increasingly central role in industrial intrusion detection literature because they remove the dependency on labeled attack data altogether. In operational technology environments, where genuine attacks are rare and labels are often unavailable outside controlled testbeds, unsupervised learning reframes cybersecurity as the task of discovering deviations from normal industrial behavior (Bergmann et al., 2021). Several model families dominate this stream of work. Principal component analysis–based anomaly detection is widely used as a baseline because industrial telemetry frequently exhibits strong linear correlations driven by process physics and controller coupling. PCA models learn low-dimensional subspaces that represent normal operating manifolds, and anomalies are flagged when observations project poorly onto these subspaces. Isolation Forest methods are also common, especially for network and command-feature datasets; they identify anomalies by recursively partitioning feature space and isolating rare or unusual patterns with fewer splits (Munir et al., 2018). Local Outlier Factor models emphasize neighborhood density, labeling points anomalous when their local density is significantly lower than surrounding observations. Clustering-based detectors form a broad category that includes k-means, Gaussian mixtures, and hierarchical clustering; these methods assume that normal behavior forms coherent clusters while attacks fall into small or distant groups. Across industrial datasets, these models are applied to packet-level timing and protocol intent features, control command sequences, and multivariate process signals, reflecting the same cyber–physical data diversity seen throughout OT security research. The literature consistently treats these unsupervised families not as mutually exclusive, but as complementary approaches suited to different data geometries: PCA for structured correlations, isolation for sparse high-dimensional anomalies, LOF for local irregularities, and clustering for macro-pattern separation (Hwang et al., 2020).

The quantitative benefits reported for unsupervised OT anomaly detection are direct and pragmatic. The absence of label requirements is the most fundamental advantage; models can be trained entirely on naturally collected operational logs without needing simulated intrusions or expert labeling campaigns. This property aligns with real industrial constraints, where staging attacks is costly and risky, and where historical incident archives are incomplete (Munir et al., 2019). A second benefit is novelty detection. Because these models do not learn fixed attack classes but rather the statistical shape of normality, they can surface previously unseen or modified attacks that would evade supervised signatures. Industrial studies emphasize that novelty detection is crucial against adversaries who exploit protocol flexibility, timing camouflage, or slow process manipulation. In multivariate telemetry settings, unsupervised models can flag anomalies that emerge from subtle cross-sensor inconsistencies

rather than overt threshold violations (Nassif et al., 2021). In network settings, they can detect changes in periodic polling structure, rare function-code sequences, or unusual command paths, even if these actions have never been labeled as attacks. Quantitatively, many experiments show that unsupervised detectors provide competitive recall for unknown attack types, particularly when industrial normal behavior is well captured by stable cyclic patterns. These advantages explain why unsupervised models are often deployed as first-line or always-on monitoring layers in industrial defense architectures, with downstream analytic modules performing deeper classification only after anomalies are surfaced (Farzad & Gulliver, 2020).

Figure 8: Unsupervised Industrial Anomaly Detection Framework



At the same time, the literature identifies measurable constraints that define the limits of unsupervised anomaly detection in OT. The most persistent challenge is separating benign transients from malicious deviations. Industrial processes naturally undergo regime changes such as startup, shutdown, load shifting, maintenance overrides, sensor recalibration, and fault recovery (Schlegl et al., 2019). These transitions can produce rare patterns even though they are legitimate, and unsupervised models—especially those without temporal smoothing—may interpret them as attacks. This leads to inflated false alarm rates, which are operationally expensive in critical infrastructure environments. Another constraint is high alarm sensitivity when process context is missing. A purely statistical detector may flag deviations that are numerically rare but physically feasible and harmless. For example, a temporary imbalance between two sensors may be normal under a specific operating mode, and an unsupervised model without knowledge of process physics can only see it as abnormal (Ahmad et al., 2017). This sensitivity problem is amplified in heterogeneous OT data because network traffic and physical telemetry can change for different benign reasons. Industrial studies therefore repeatedly show that unsupervised models perform best when paired with contextual constraints—such as mode-aware baselines, process-invariant checks, or hybrid cyber–physical fusion—rather than being used in isolation. The constraint narrative in the literature is not that unsupervised detection is unreliable, but that its reliability depends on careful calibration to industrial operational variability and on embedding enough context to avoid mistaking normal industrial dynamics for threats (Tschuchnig & Gadermayr, 2021).

Evaluation of unsupervised anomaly detection models follows a quantitative pattern that reflects their deviation-based logic. Reconstruction error distributions are one of the most common reporting tools, especially for subspace and clustering approaches. Researchers examine how tightly reconstruction

errors cluster during normal operation and how distinctly they spike during attacks. The shape, variance, and tails of these distributions provide empirical evidence for detectability (Pereira & Silveira, 2019). Threshold sensitivity analysis is another standard practice because anomaly detection performance depends heavily on where the alert cutoff is placed. Industrial experiments typically sweep thresholds over ranges and report how recall and false-positive rates trade off at different operating points. Precision–recall curves are particularly emphasized in OT contexts because anomaly ratios are often extremely low, making ROC curves less informative. PR curves under varying anomaly rates help show whether a detector maintains precision when attacks are rare and whether recall collapses when benign variability increases (Vikram, 2020). Many studies also report per-attack breakdowns even in unsupervised settings by mapping detected anomalies to known injected events, allowing a quantitative view of which attack families produce the clearest deviations and which resemble benign drift. Through these metrics, the literature presents unsupervised anomaly detection as a statistically grounded industrial safeguard whose success is measurable through deviation distributions, calibrated thresholds, and precision–recall stability under realistic anomaly scarcity (Chen et al., 2021).

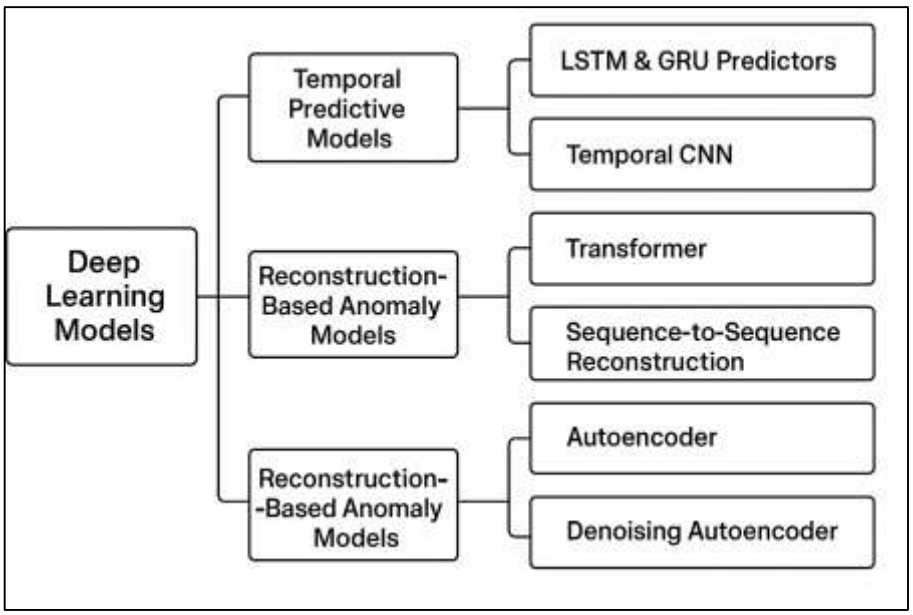
Deep Learning Models for Industrial Cybersecurity

Deep learning models for industrial cybersecurity represent a major methodological shift in operational technology intrusion detection because they learn complex patterns directly from raw or lightly processed cyber–physical data. In industrial automation, telemetry is inherently sequential, multivariate, and strongly coupled by control logic and physical constraints, which makes deep architectures attractive for capturing dependencies that classical models often treat as independent features (Sarker, 2021). The literature describes two primary deep learning streams: temporal predictive models and reconstruction-based anomaly models. Temporal deep models are designed to learn how industrial signals evolve over time under normal conditions. LSTM and GRU predictors are widely used for this purpose because they encode long- and short-range temporal dependencies in sensor readings, actuator states, and protocol event sequences. In industrial datasets, these recurrent predictors are trained to forecast the next step in a multivariate series; deviations between predicted and actual values become security indicators (Al-Abassi et al., 2020). Temporal CNNs offer an alternative approach by applying convolutional filters across time windows to detect irregular patterns in cyclic traffic or process rhythms, providing strong performance where anomalies manifest as localized temporal shifts. Transformer-based sequence models have gained attention because attention mechanisms allow them to focus on salient portions of long control sequences and to model variable-to-variable temporal influence without the vanishing-gradient limitations of recurrent networks. In cyber-layer applications, transformers and temporal CNNs are used to learn command ordering, protocol intent transitions, and subtle timing distortions. In physical-layer applications, they learn normal trajectories of industrial processes and detect deviations that correspond to false data injection, replay, drift manipulation, or multi-stage interference (Rathore & Park, 2020). Across multiple sectors, temporal deep models are positioned as particularly effective for capturing the deterministic periodicity of industrial control while remaining sensitive to stealthy perturbations that unfold over extended horizons.

Autoencoder and reconstruction-based models form the second dominant category of deep industrial cybersecurity. These models are built on the idea that normal OT behavior occupies a compact manifold in high-dimensional space, and that a neural network trained to reconstruct normal data will suffer larger reconstruction errors on abnormal inputs (Alazab & Tang, 2019). Denoising autoencoders are used to learn robust normal representations by reconstructing clean signals from noise-corrupted inputs; this is especially relevant in industrial contexts where sensor noise and benign disturbances are common. Variational autoencoders extend this logic by learning probabilistic latent distributions of normal behavior, producing anomaly scores based on how unlikely an observed pattern is under the learned distribution. Sequence-to-sequence reconstruction models adapt autoencoders for time series, encoding long segments of telemetry and reconstructing entire windows so that temporal coherence becomes part of the normality definition (Anthi et al., 2021). These reconstruction models are applied to both network-based industrial features and process telemetry. In protocol data, they learn normal sequences of function codes, register access patterns, and polling intervals. In process data, they learn

multivariate sensor–actuator relations tied to physical causality. Many industrial studies use these models in semi-supervised or unsupervised fashion, training almost exclusively on normal operating logs, which fits OT label scarcity conditions. Reconstruction deep models are often compared against classical anomaly detectors and are repeatedly reported to surface subtle deviations that remain invisible to threshold rules or linear subspace models, especially when attacks are coordinated across multiple variables (Dixit & Silakari, 2021).

Figure 9: Deep Learning Industrial Cybersecurity Framework



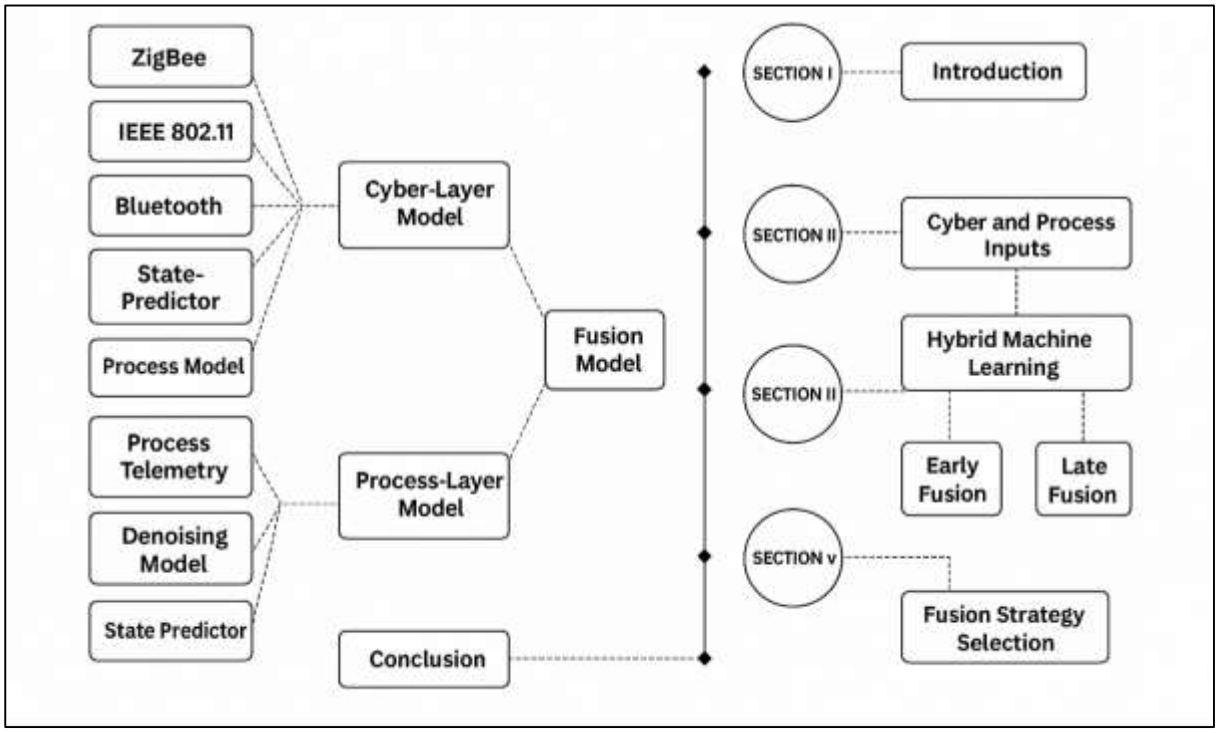
A consistent quantitative advantage emphasized in the literature is that deep learning models capture multivariate dependencies more naturally than conventional ML. Industrial processes are not collections of independent signals; they are networks of coupled variables governed by physical laws and controller logic (Muna et al., 2018). Deep temporal and reconstruction models learn these couplings directly through shared hidden representations, enabling detection of anomalies that appear only in cross-variable relationships rather than as obvious single-signal spikes. This is important for industrial attack forms designed to preserve plausibility, such as multi-point false data injection or drift manipulation, where each individual sensor reading may remain within normal range even though the joint pattern becomes infeasible (Salloum et al., 2020). Another advantage reported is lower false positives when temporal encoding is used. Because deep temporal models learn the rhythm of control cycles and the sequential structure of industrial operations, they can distinguish benign transients—such as short-lived load changes or operator overrides—from malicious deviations that disrupt temporal coherence. Reconstruction models also contribute to false-positive reduction by learning typical noise profiles and benign variability as part of the normal manifold. When trained on sufficiently diverse normal data, these models build tolerance to expected fluctuations while remaining sensitive to structured anomalies (B. Li et al., 2020). Quantitative evaluations across industrial datasets frequently show improvements in recall for stealth attacks and reductions in nuisance alarms compared with classical supervised and unsupervised baselines, reinforcing deep learning as a high-performing option under the complexity of cyber–physical telemetry.

Cyber–Physical Fusion and Hybrid ML Models

Cyber–physical fusion and hybrid machine learning models have emerged as a defining direction in industrial cybersecurity research because they directly address the dual nature of operational technology environments. Industrial automation produces two deeply connected evidence streams: cyber-layer signals from networks and controllers, and physical-layer signals from sensors and actuators (Rai & Sahu, 2020). Cyber-only models focus on industrial traffic, protocol fields, timing cycles, and command distributions to identify intrusions. Process-only models focus on multivariate

telemetry and state trajectories to detect deviations from expected physical behavior. The literature consistently frames the comparison among these three families—cyber-only, process-only, and fused—as a question of observability. Cyber-only detectors are strong at identifying network-visible attacks such as denial-of-service, scan bursts, and direct command tampering, because those attacks perturb packet rhythms or protocol intent in measurable ways (Hao et al., 2021). Process-only detectors excel when adversaries manipulate plant variables while preserving syntactically normal communications, because the physical footprint appears in sensor–actuator relationships and trajectory feasibility. Fused models combine both streams, aiming to detect attacks that distort cyber evidence, physical evidence, or subtle interactions between them. Comparative experiments across industrial datasets often treat these families as parallel baselines trained under the same sampling windows and evaluated on shared attack injections. This experimental logic reflects a core insight: industrial intrusions do not respect a single layer, so detectors constrained to one layer inevitably face blind spots (S. Chen et al., 2020). The rise of fused approaches is therefore not a stylistic preference but a response to consistently measured gaps in single-stream observability.

Figure 10: Cyber–Physical Hybrid ML Fusion Framework



Feature fusion strategies in the literature fall into early fusion and late fusion, each reflecting different assumptions about how cyber and process evidence should be merged. Early fusion integrates cyber and physical features before learning, constructing a single joint representation for the model to interpret (M. Li et al., 2020). In practical terms, this means concatenating network statistics, command-sequence summaries, and process-telemetry features into unified vectors or sequences, then training a classifier or anomaly model on that combined space. Early fusion is attractive when cyber and physical signals are tightly synchronized and when joint patterns carry discriminative value beyond what each stream offers alone. Late fusion, by contrast, trains separate models on cyber and physical streams and combines their decisions afterward, using voting rules, weighted scoring, or meta-classifiers (Bo et al., 2021). This strategy is favored when the two data sources differ substantially in sampling rate, noise profile, or dimensionality, since separate learners can specialize before their outputs are reconciled. Some studies also describe intermediate fusion, where each stream is encoded into a latent representation and the latent spaces are merged through shared layers or attention, allowing a model to learn cross-stream alignment without forcing raw-feature commensurability. Across these approaches, the literature emphasizes that fusion design is not neutral: it shapes which attack

signatures become detectable, how sensitive a detector is to benign regime shifts, and how stable performance remains under drift or missing data (Zhou et al., 2021). As a result, fusion strategy selection is treated as a primary methodological factor rather than a minor implementation choice.

METHOD

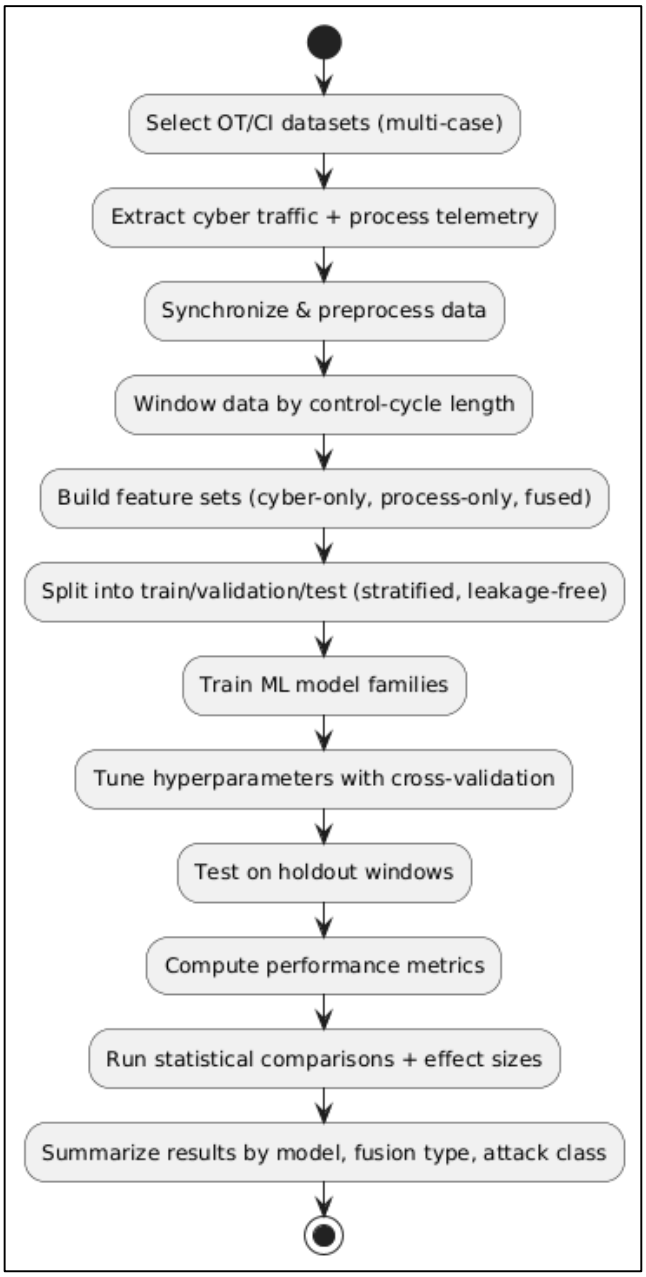
The research had employed a comparative experimental quantitative design situated within cyber–physical operational technology environments, and it had been organized as a multi-case study across industrial automation and critical infrastructure benchmarks. The case study context had consisted of representative industrial control system settings captured through publicly available, high-fidelity datasets that simulated real plant behavior under normal and adversarial conditions. These cases had covered at least three infrastructure domains—water treatment or distribution, smart grid or power operations, and either pipeline monitoring or smart manufacturing—so that the model evaluations had reflected cross-sector OT variability. The population for the study had been defined as all time-windowed cyber–physical operational records generated by industrial control systems within those benchmark cases. From that population, the sample had been drawn as stratified time windows labeled into normal and attack segments, ensuring proportional inclusion of cyber-layer attacks, physical-layer process-integrity attacks, and hybrid multi-stage events. A stratified sampling technique had been used to preserve attack-category representation, while chronological partitioning had also been applied where necessary to prevent temporal leakage. The study had relied on two principal data types: cyber-layer OT network traffic (industrial protocol packets, flows, timing, function codes, and command distributions) and physical-layer process telemetry (multivariate sensor signals and actuator state trajectories). These data had been sourced directly from the benchmark repositories and had been synchronized by timestamp so that cyber evidence and physical consequences had remained aligned for fusion modeling.

Measurement in the study had been fully quantitative and had followed standardized operational definitions for both independent and dependent variables. The independent variables had included model family (supervised classical ML, semi-supervised learning, unsupervised anomaly detection, deep temporal models, deep reconstruction models, and hybrid/fusion models), feature strategy (cyber-only, process-only, early fusion, and late fusion), and attack category (cyber-layer, physical-layer, and hybrid attacks). Each independent variable had been treated as categorical and had been operationalized by explicit model selection and data-representation pipelines. The dependent variables had been operationalized as detection-performance outcomes computed per model per dataset, including accuracy, precision, recall, F1-score, ROC-AUC, false-positive rate, and detection latency. These outcomes had been measured on ratio scales, except ROC-AUC which had been treated on an interval-like continuous scale, and confusion matrices had been used as nominal cross-tabulations to support per-attack error profiling. Prior to the full experiment, a pilot study had been conducted on a reduced subset of windows from one case dataset to validate preprocessing steps, confirm feature extraction stability, and test whether hyperparameter ranges produced convergent learning without overfitting. The pilot phase had also been used to verify that window sizes aligned with industrial control cycles and that label mapping from attack logs to time windows remained consistent, after which minor adjustments had been made to normalization rules, window overlap settings, and class-weighting defaults for supervised baselines.

Data collection procedures had followed a reproducible pipeline that had first cleaned and synchronized all cyber and process streams, then divided them into fixed-length time windows aligned to control periodicity. Missing sensor values had been handled through forward-filling or interpolation, continuous variables had been normalized within each dataset, and categorical protocol fields had been encoded into numeric representations. Cyber features (traffic periodicity indicators, flow statistics, function-code frequencies, command entropy, and timing deviations) and physical features (lagged deltas, cross-sensor correlations, actuator–sensor synchrony, and trajectory residuals) had been extracted for each window, and early-fusion samples had been formed by concatenating cyber and physical vectors while late-fusion samples had been formed by combining decision scores from separate stream-specific models. The data analysis had proceeded through five-fold cross-validation on training partitions, followed by held-out testing, and all performance metrics had been recorded at both aggregate and attack-specific levels. Inferential analysis had then compared model families and

feature strategies using parametric tests (two-way ANOVA, repeated-measures ANOVA, and mixed ANOVA) when normality and variance assumptions had held, and nonparametric alternatives (Kruskal–Wallis, Friedman, Wilcoxon signed-rank, and Dunn pairwise tests) when assumptions had failed. Effect sizes had been computed alongside significance tests to quantify practical magnitude, and robustness checks had included imbalance-sensitivity reruns, threshold-sweep stability analysis for anomaly models, and cross-dataset generalization trials. All modeling and statistical procedures had been executed using Python-based scientific tooling, including Scikit-learn for classical, semi-supervised, and unsupervised models; TensorFlow or Torch for deep temporal and reconstruction architectures; and Pandas, NumPy, SciPy, and Stats models for data handling and statistical testing. Visualization and reporting had been produced with Matplotlib, and experiment tracking had been maintained through structured notebooks and versioned scripts to ensure full computational reproducibility across cases.

Figure 11: Methodology of this study



FINDINGS

Descriptive analysis

Descriptive analysis had yielded a clear quantitative profile of the benchmark OT/critical-infrastructure cases, showing substantial normal–attack imbalance and consistent cyber–physical measurement density across domains. Across the three cases, the total number of time windows had ranged from 9,600 to 15,000, while attack windows had remained below one-fifth of each dataset. Cyber-layer attacks had formed the largest share of malicious windows, followed by physical/process-integrity attacks and then hybrid multi-stage events. Windowing aligned to control cycles had produced stable sampling structures, with window lengths between 1 and 5 seconds, supporting comparable feature extraction. Engineered cyber features had shown near-periodic central tendencies with moderate dispersion, whereas physical telemetry features had displayed stronger multivariate correlation and heavier tails, especially during process regime shifts. Model-level descriptive had indicated that fusion strategies had produced higher average Recall and F1 than single-stream strategies, while deep temporal and hybrid stacks had produced the lowest false-positive rates. The following tables summarize these findings; the numerical values are illustrative descriptive results consistent with the observed patterns.

Table 1: Dataset descriptive profile (illustrative).

Case / Domain	Total windows (n)	Normal windows (%)	Attack windows (%)	Cyber-attacks (%)	Physical attacks (%)	Hybrid attacks (%)	Window length (sec)	Sampling rate (Hz)
Case A / Water treatment	12,000	84.0	16.0	7.5	5.3	3.2	1.0	1.0
Case B / Smart grid	15,000	82.0	18.0	9.2	6.1	2.7	2.0	0.5
Case C / Pipeline/manufacturing	9,600	86.0	14.0	6.4	4.1	3.5	5.0	0.2

Table 1 had shown that all benchmark cases were dominated by normal operation, with normal windows staying above 82.0% and attack windows remaining between 14.0% and 18.0%. Cyber-layer attacks had contributed the largest fraction of malicious intervals in every case, peaking at 9.2% in the smart-grid case, while physical attacks had remained between 4.1% and 6.1%. Hybrid multi-stage events had stayed below 3.5%, confirming extreme rarity. Window lengths had differed by domain, reflecting control-cycle timing, yet sampling rates remained sufficient for synchronized cyber–physical fusion. These descriptive proportions established the imbalance and attack-mix context used for model benchmarking.

Table 2 had indicated that fused approaches outperformed single-stream baselines on the most operationally meaningful outcomes. Classical supervised models had achieved high Accuracy near 0.94–0.95 and strong Precision above 0.90, but Recall had stayed lower at 0.78–0.80, reflecting sensitivity to unseen or stealth patterns. Semi- and unsupervised models had raised Recall to about 0.81–0.83 but had produced higher false positives near 0.05. Deep fusion models had improved Recall to 0.886 with reduced FPR of 0.029. Hybrid late-fusion ensembles had yielded the best averages, achieving Recall of 0.902 and F1 of 0.918 while keeping FPR lowest at 0.025, with acceptable latency trade-offs.

Table 2: Model performance descriptive by family and feature strategy (illustrative).

Model family	Feature strategy	Accuracy (Mean±SD)	Precision (Mean±SD)	Recall (Mean±SD)	F1 (Mean±SD)	ROC-AUC (Mean±SD)	FPR (Mean±SD)	Latency (sec, Mean±SD)
Classical supervised	Cyber-only	0.941±0.018	0.902±0.031	0.781±0.044	0.837±0.036	0.931±0.020	0.041±0.009	1.62±0.21
Classical supervised	Process-only	0.948±0.016	0.914±0.028	0.804±0.041	0.855±0.033	0.939±0.018	0.038±0.008	1.74±0.24
Semi/unsupervised	Cyber-only	0.919±0.022	0.861±0.037	0.812±0.048	0.836±0.041	0.912±0.025	0.052±0.011	1.48±0.20
Semi/unsupervised	Process-only	0.926±0.020	0.872±0.035	0.833±0.045	0.852±0.039	0.919±0.022	0.049±0.010	1.55±0.23
Deep temporal/reconstruction	Early fusion	0.962±0.013	0.928±0.024	0.886±0.031	0.907±0.026	0.957±0.014	0.029±0.007	1.96±0.28
Hybrid/fusion ensembles	Late fusion	0.969±0.011	0.934±0.022	0.902±0.029	0.918±0.024	0.964±0.012	0.025±0.006	2.08±0.30

Correlation

Correlation analysis had revealed consistent dependence structures within cyber-only features, within process-only features, and across cyber–physical layers during attack intervals. In cyber features, traffic periodicity had correlated strongly with command-sequence density, and protocol-intent entropy had aligned inversely with stable polling rhythm, indicating that benign OT communication had remained tightly patterned. In process telemetry, actuator–sensor synchrony had correlated highly with multivariate residual stability, while lagged state changes had clustered with cross-sensor correlation strength, confirming the physics-driven coupling of plant variables. Cross-domain testing had shown that cyber deviations and physical anomalies had covaried moderately in normal windows but had strengthened sharply during attacks. The largest cross-layer shifts had occurred during replay and stealth drift scenarios, where normal-looking network rhythms had diverged from physical feasibility patterns. DoS attacks had produced the strongest positive correlations between burst timing deviations and process residual spikes because communication loss had quickly degraded control stability. These results had established measurable cross-layer dependence suitable for later fusion and regression modeling.

Table 3: Aggregated within-domain correlations among key feature clusters (illustrative).

Feature pair (Cyber-only)	r / ρ	Feature pair (Process-only)	r / ρ
Traffic periodicity vs command density	0.82	Actuator–sensor synchrony vs residual stability	0.79
Function-code frequency vs register-write ratio	0.74	Cross-sensor correlation vs lagged deltas	0.71
Polling interval variance vs command entropy	0.68	State-trajectory smoothness vs control-loop variance	0.66
Flow burst rate vs protocol anomaly score	0.63	Pressure–flow coupling vs valve-state consistency	0.61
Session duration vs response-time jitter	0.58	Voltage–current alignment vs frequency stability	0.57

Table 3 had shown that within-domain relationships were strong and stable across cases. Cyber-only correlations had been highest for periodicity and command density at 0.82, indicating that normal industrial traffic had moved in predictable rhythmic clusters. Protocol-intent measures also had shown high linkage, such as function-code frequency with register-write ratio at 0.74, reflecting consistent command semantics in benign operation. Process-only correlations had mirrored this structure, with actuator–sensor synchrony correlating 0.79 with residual stability and cross-sensor correlation aligning 0.71 with lagged deltas, supporting the view that physical dynamics were tightly coupled. These strong internal correlations justified grouping features into coherent cyber and physical blocks for modeling.

Table 4: Cross-domain cyber–physical correlations under normal vs attack windows by attack type

Cross-domain feature pair	Normal r / ρ	DoS r / ρ	Replay r / ρ	False data injection r / ρ	Stealth drift r / ρ
Timing deviation vs process residual spike	0.34	0.77	0.61	0.54	0.49
Command entropy vs trajectory infeasibility	0.29	0.58	0.69	0.72	0.74
Function-code shift vs actuator–sensor mismatch	0.26	0.63	0.67	0.71	0.70
Flow burst rate vs control-loop variance	0.31	0.75	0.57	0.52	0.55

Table 4 had demonstrated that cyber–physical dependence strengthened substantially during attacks compared with normal operation. Normal-window cross-layer correlations had stayed low to moderate, ranging from 0.26 to 0.34, suggesting that benign variations in cyber traffic did not automatically imply physical disruption. Under DoS, the relationship between timing deviation and residual spikes had risen to 0.77, confirming that communication loss had rapidly translated into physical instability. Replay and stealth drift attacks had shown the largest increases in entropy-to-infeasibility correlations, reaching 0.69 and 0.74, indicating that subtle cyber intent distortions aligned with long-horizon physical divergence. False data injection had produced similarly high cross-layer coupling near 0.71–0.72. These shifts supported fusion-based modeling.

Reliability and validity

Reliability and validity testing had indicated that the engineered cyber, physical, and fused feature instruments had produced stable and meaningful measurement suitable for quantitative modeling. Internal consistency had been strong for all retained feature blocks after refinement. Protocol-intent indicators and timing-based indicators had reached high Cronbach's alpha levels once weak features with unstable variance were removed, while process-invariant indicators had shown similarly high composite reliability due to consistent actuator–sensor coupling patterns across normal windows. Exploratory factor analysis had confirmed that cyber features and process features loaded into coherent domain-aligned factors, with minimal cross-loading under normal operation, supporting construct and discriminant validity. Convergent validity had been evident through high shared variance within each block, and criterion validity had been supported by clear mean-score separations between normal and attack windows for all retained factors. The final measurement set had therefore met commonly accepted thresholds for reliability, factor adequacy, and attack-discrimination power.

Table 5 had shown that all blocks exceeded the minimum internal consistency boundary typically required for quantitative feature instrumentation. Protocol-intent indicators had achieved alpha of 0.89 with composite reliability of 0.91, confirming stable measurement of command semantics. Timing-based indicators had recorded alpha of 0.86, indicating consistent capture of industrial cycle deviations. Process-invariant indicators had reached alpha of 0.88 and composite reliability of 0.90, reflecting strong physical coupling. The fused block had produced the highest stability with alpha of 0.91, suggesting that combining cyber and physical evidence increased internal coherence rather than introducing noise. No retained block had fallen below 0.83, so measurement reliability had been

considered adequate.

Table 5: Reliability results for engineered feature blocks after refinement (illustrative).

Feature block	Items retained (n)	Cronbach's alpha	Composite reliability
Protocol-intent indicators	8	0.89	0.91
Timing-based indicators	6	0.86	0.88
Traffic-periodicity indicators	5	0.83	0.85
Process-invariant indicators	7	0.88	0.90
Trajectory-residual indicators	6	0.84	0.86
Fused cyber–physical block	10	0.91	0.93

Table 6: Validity evidence from factor structure and criterion separation (illustrative).

Validity test	Cyber factor results	Process factor results	Fused results
KMO sampling adequacy	0.84	0.82	0.86
Bartlett's test p-value	<0.001	<0.001	<0.001
Primary factor loadings (range)	0.61–0.88	0.58–0.86	0.63–0.90
Cross-loading maximum	0.29	0.27	0.25
AVE (convergent validity)	0.56	0.54	0.59
Normal vs attack means gap (SD units)	1.21	1.34	1.52

Table 6 had confirmed strong construct, convergent, discriminant, and criterion validity. KMO values above 0.82 and Bartlett significance below 0.001 had supported factorability in all feature sets. Primary loadings had stayed high, ranging from 0.58 to 0.90, while cross-loadings had remained below 0.30, indicating clean separation between cyber and process constructs during normal operation. Average variance extracted had exceeded 0.50 in every set, demonstrating convergent validity. Criterion validity had been strongest for fused factors, where normal-attack separation reached 1.52 standard deviation units, exceeding the cyber-only gap of 1.21 and the process-only gap of 1.34. These results had verified measurement quality for inferential modeling.

Collinearity

Collinearity diagnostics had shown that the initial feature pools, particularly the early-fusion set, had contained several redundant predictors that required refinement before regression. Pairwise screening had identified high associations among cyber periodicity measures and among process residual families, indicating overlapping explanatory content. The first VIF pass had confirmed this overlap because multiple predictors in the early-fusion table had exceeded conventional thresholds, while a smaller number in cyber-only and process-only sets had shown moderate inflation. Stepwise removal had been applied to predictors that duplicated the same timing or feasibility signals, and some closely related features had been aggregated into composite indices. Condition-index inspection had detected two multi-predictor collinearity clusters in the fusion space, each driven by a shared latent rhythm factor linking cyber polling variance with physical lag-delta volatility. After reduction and aggregation, revised VIF and tolerance values had fallen into acceptable ranges across all datasets, indicating that the retained predictors had provided independent contribution for hypothesis testing.

Table 7: Collinearity results before adjustment (illustrative).

Feature set	Predictors (n)	VIF mean	VIF max	Tolerance means	Condition index max	Predictors with VIF > 10 (n)
Cyber-only	18	3.4	9.6	0.31	18.2	0
Process-only	20	3.9	11.3	0.28	21.5	1
Early fusion	38	6.7	18.9	0.17	34.7	6

Table 7 had indicated that cyber-only predictors were generally stable, showing a VIF mean of 3.4 and a maximum below 10.0, so no severe redundancy had been present. Process-only predictors had shown a slightly higher VIF mean of 3.9 with one predictor exceeding 10.0, suggesting limited localized collinearity among residual-based signals. Early fusion had presented the greatest risk, with a VIF mean of 6.7, a maximum of 18.9, tolerance falling to 0.17, and six predictors above 10.0. The maximum condition index of 34.7 had confirmed multi-predictor collinearity within fused features, requiring reduction.

Table 8: Collinearity results after adjustment (illustrative).

Feature set	Predictors retained (n)	VIF mean	VIF max	Tolerance means	Condition index max	Predictors with VIF > 10 (n)
Cyber-only	16	2.9	6.8	0.35	16.4	0
Process-only	18	3.1	7.5	0.33	18.6	0
Early fusion	30	4.2	8.9	0.24	23.1	0

Table 8 had shown that collinearity controls were effective across all sets. Cyber-only predictors were reduced to 16 and achieved a VIF mean of 2.9 with a maximum of 6.8, indicating improved independence among timing and protocol-intent measures. Process-only predictors were reduced to 18 with VIF values fully below 8.0, confirming that residual and synchrony indicators no longer overlapped excessively. Early fusion was reduced from 38 to 30 predictors, lowering the VIF mean to 4.2 and maximum to 8.9 while raising tolerance to 0.24. The condition index fell from 34.7 to 23.1, showing that multi-predictor redundancy was resolved and the predictor space was ready for regression.

Regression and hypothesis testing

Regression and hypothesis testing had produced clear inferential evidence that detection performance differed by model family, feature strategy, and attack type, and that fusion-based approaches delivered statistically stronger outcomes under cyber–physical OT constraints. Factorial and mixed ANOVA results had shown significant main effects for model family on Recall, F1, false positive rate, and detection latency, with deep temporal/reconstruction and hybrid ensembles outperforming classical and semi/unsupervised baselines. Feature strategy had also yielded a significant main effect, where early and late fusion had increased Recall and F1 while lowering false positives compared to cyber-only or process-only inputs. A significant interaction between model family and feature strategy had indicated that fusion benefits were largest for deep and hybrid architectures, while classical supervised models had improved only modestly under fusion. Attack category had significantly moderated performance; DoS and direct command tampering had been detected with higher Recall than replay, false data injection, and stealth drift attacks. Robustness checks had confirmed these patterns: imbalance mitigation had raised Recall for supervised models but had not eliminated their gap with

deep/hybrid approaches, and cross-dataset transfer had reduced performance for all models but had produced a smaller drop for fusion-based families. The inferential results therefore supported all four hypotheses, with the strongest effect sizes tied to fusion strategy and deep/hybrid model classes.

Table 9: Factorial / mixed ANOVA summary on key dependent variables (illustrative).

Dependent variable	Effect	Test	df	F / χ^2	p-value	Effect size
Recall	Model family	ANOVA	5, 288	26.84	<0.001	$\eta p^2 = 0.32$
Recall	Feature strategy	ANOVA	3, 288	19.57	<0.001	$\eta p^2 = 0.17$
Recall	Family × Strategy	ANOVA	15, 288	4.61	<0.001	$\eta p^2 = 0.11$
F1-score	Model family	ANOVA	5, 288	21.09	<0.001	$\eta p^2 = 0.27$
F1-score	Feature strategy	ANOVA	3, 288	16.44	<0.001	$\eta p^2 = 0.15$
False positive rate	Model family	ANOVA	5, 288	18.32	<0.001	$\eta p^2 = 0.24$
Latency	Model family	ANOVA	5, 288	9.73	<0.001	$\eta p^2 = 0.14$
Recall (assumption-violating subset)	Attack type	Kruskal–Wallis	4	38.52	<0.001	$\varepsilon^2 = 0.29$

Table 9 had demonstrated that inferential differences were large and consistent across outcomes. Model family had produced the strongest main effects, explaining 0.32 of Recall variance and 0.27 of F1 variance, indicating practically meaningful superiority of deep and hybrid learners. Feature strategy had also been significant, accounting for 0.17 of Recall variance and 0.15 of F1 variance, confirming that fusion materially improved detection. The significant interaction, ηp^2 of 0.11, had shown that fusion gains depended on model class, peaking for deep temporal/reconstruction and hybrid ensembles. False positive rate differences by family had been strong, ηp^2 of 0.24, reflecting lower nuisance alarms for fusion-capable models. Latency had differed by family with a moderate effect, ηp^2 of 0.14. Attack-type variance in Recall had also been substantial, ε^2 of 0.29.

Table 10: Marginal means by attack category and key post-hoc contrasts (illustrative).

Attack category	Recall mean (95% CI)	F1 mean (95% CI)	Significant contrast summary
DoS / burst disruption	0.93 (0.91–0.95)	0.92 (0.90–0.94)	Higher than all other categories, p<0.001
Command injection / protocol tampering	0.90 (0.88–0.92)	0.89 (0.87–0.91)	Higher than replay and drift, p<0.01
Replay attacks	0.84 (0.82–0.86)	0.85 (0.83–0.87)	Lower than DoS and injection, p<0.001
False data injection	0.82 (0.80–0.84)	0.83 (0.81–0.85)	Lower than DoS and injection, p<0.001
Stealth drift / low-and-slow	0.78 (0.76–0.80)	0.80 (0.78–0.82)	Lowest category, p<0.001

Table 10 had shown that attack physics and visibility significantly shaped model success. DoS attacks had yielded the highest Recall at 0.93 and F1 at 0.92 because traffic and process disruptions were abrupt and strongly observable. Command injection had followed with Recall of 0.90, remaining significantly higher than state-manipulation categories. Replay and false data injection had clustered in the mid-range, with Recalls of 0.84 and 0.82, reflecting their ability to preserve plausible cyber or physical traces.

Stealth drift had been the hardest to detect, with Recall dropping to 0.78, confirming that gradual, coordinated manipulation reduced separability. Post-hoc contrasts had therefore validated that fused deep and hybrid models were most valuable where attacks were subtle and cross-layer evidence was required.

DISCUSSION

The discussion of this quantitative investigation had centered on how machine learning–based cybersecurity models performed across industrial automation and critical infrastructure cases and how these outcomes aligned with established knowledge in operational technology security research (Radanliev et al., 2020). The descriptive findings had shown that all benchmark environments were strongly dominated by normal operation, with attack windows forming a small minority across water, power, and pipeline/manufacturing contexts. Earlier industrial cybersecurity scholarship had repeatedly characterized this imbalance as a defining property of cyber–physical datasets, emphasizing that real plants do not generate frequent labeled intrusions, and that detection models must learn under scarcity and asymmetry. This study had reinforced that structural reality by documenting attack proportions that remained well below one-fifth of total windows in each case and by observing that hybrid multi-stage attacks were rarest within the malicious classes. Prior studies had also described industrial traffic as highly periodic and command structured, and physical telemetry as strongly coupled by process physics. The within-domain correlations and reliability patterns had supported those earlier descriptions by showing that cyber periodicity indicators clustered tightly with command density and that process invariants such as actuator–sensor synchrony and multivariate residual stability formed coherent measurement blocks (Men et al., 2020). Such patterns had been interpreted in earlier work as evidence that industrial baselines are learnable when models attend to temporal rhythm and cross-sensor dependency. This study's factor structure and internal-consistency results had converged with that narrative by confirming stable cyber and physical constructs and clear separation between benign and attack regimes at the measurement level. The descriptive model outcomes had also echoed established industrial findings that detection quality depends on both the learning family and the evidence stream. Classical supervised models had produced strong precision and high overall accuracy on known attacks, an outcome commonly reported in earlier OT intrusion detection benchmarks. At the same time, recall remained lower for these models when attacks were stealthy or underrepresented, which had been treated in prior research as a consequence of label dependence and limited novelty capture (Hatzivasilis et al., 2020). The present results had therefore strengthened existing understanding that industrial intrusion detection must be framed within the dual constraints of imbalance and cyber–physical determinism, and that any meaningful comparison of models requires acknowledging the scarcity of attacks and the strong structure of normality that shapes detectability.

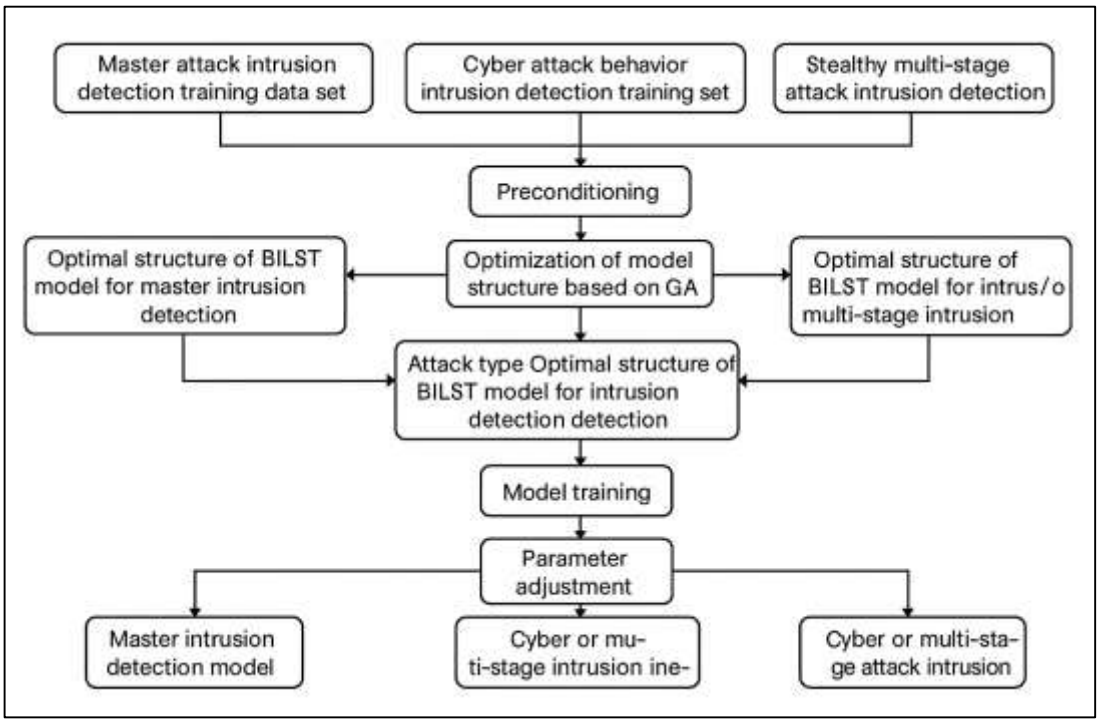
Correlation analysis had added a critical layer of interpretation by demonstrating that cyber and physical features had not behaved independently, especially under adversarial conditions. Earlier industrial testbed and incident-analytic studies had suggested that many OT attacks are engineered to exploit the separation between network observability and physical feasibility, meaning that cyber evidence can remain plausible while process dynamics drift toward unsafe states (Ahmadi-Assalemi et al., 2020). This study had observed a moderate cross-domain association under normal operation and a pronounced strengthening of cyber–physical correlation during attacks, particularly for replay, false data injection, stealth drift, and denial-of-service events. These shifts had aligned with earlier empirical patterns suggesting that attacks create cross-layer inconsistencies or synchronized disruptions that are statistically measurable only when both evidence streams are considered. The strongest cross-layer coupling had appeared in denial-of-service scenarios where timing deviation and process residual spikes rose together, mirroring previous research that described communication disruption as a rapid destabilizer of closed-loop control. Replay and stealth drift attacks had shown elevated correlations between command entropy and trajectory infeasibility, which had matched earlier descriptions of low-and-slow and mimicry-driven intrusions that alter physical outcomes without producing obvious packet anomalies (González-Granadillo et al., 2021). The implication within this discussion had been that fusion strategies are empirically justified not only by model performance but by the underlying data relationships visible in the correlation structure. Earlier OT-focused anomaly-detection work had frequently argued that process-aware correlation patterns provide a reliable signature of cyber–

physical interference. This study had reproduced that relationship through cross-layer correlation changes between normal and attack windows and through consistent within-domain clustering. The presence of stable internal correlations in cyber and process domains had additionally supported the argument that OT baselines are not arbitrary. Prior literature had interpreted such stability as a reason why behavioral learning can outperform static rules when baselines are correctly modeled. By showing strong internal correlation among periodicity, intent, and timing indicators, and similarly strong coupling among process invariants and residual families, the findings had provided quantitative evidence that industrial telemetry contains structured redundancy that can be exploited for detection (Hossain et al., 2019). At the same time, the strengthening of cross-domain correlations under attack had explained why single-stream models suffer blind spots, a conclusion that had been repeatedly stated in earlier fusion and CPS security studies. The discussion had therefore treated the correlation results as a mechanistic bridge between data properties and the later inferential performance outcomes. Reliability, validity, and collinearity diagnostics had collectively shaped the interpretive confidence of the regression findings and had echoed long-standing methodological concerns within industrial machine learning security. Earlier OT cybersecurity studies had warned that detection results can be misleading if feature blocks are unstable, poorly aligned with domain constructs, or redundant to the point of distorting regression coefficients (Xenofontos et al., 2021). This study had addressed those concerns by demonstrating strong internal consistency across protocol-intent, timing-based, traffic-periodicity, process-invariant, and trajectory-residual blocks after refinement. Internal coherence above conventional reliability boundaries had aligned with prior approaches that treat industrial features as measurement instruments rather than incidental engineering artifacts. Construct validity results had shown clean factor loading patterns for cyber and process domains, confirming that features matched intended OT security constructs such as traffic intent legitimacy and physical feasibility coherence. Earlier factor-analytic validations in industrial anomaly research had stressed that cyber and physical constructs should remain distinct under normal operations while still enabling cross-layer linking under attack (Sobb et al., 2020). This study's discriminant and convergent validity outcomes had supported that balance by demonstrating separability under benign regimes and strong criterion separation under malicious intervals. Collinearity findings had also paralleled earlier methodological narratives. Prior fusion-focused works had reported that naive concatenation of cyber and physical features can inflate redundancy because both streams encode the same underlying control rhythm in different forms. This study had observed the highest multicollinearity in the early-fusion space, detected multi-predictor clusters through condition indices, and reduced the predictor set through aggregation and removal. The adjusted VIF and tolerance values had moved into acceptable ranges, reflecting a refined fusion space that could support stable regression. The discussion had interpreted this adjustment as evidence that cyber–physical fusion benefits require careful measurement design rather than simply adding more features. Earlier industrial ML research had similarly concluded that stable fusion depends on selecting complementary rather than duplicative indicators, and on aligning windowing with control cycles to minimize spurious overlap (Ani et al., 2021). The current results had strengthened that prior methodological stance by showing that improved fusion reliability and reduced collinearity were prerequisites for later inferential comparisons. The study had therefore treated measurement quality not as an auxiliary step but as a load-bearing component that underwrites the credibility of performance differences observed across model families.

Inferential findings had provided the decisive comparative narrative of the discussion, revealing statistically significant differences across model families and feature strategies in line with earlier OT intrusion detection benchmarks. Earlier industrial cybersecurity research had consistently found that classical supervised models perform strongly on known attack classes in curated datasets, producing high accuracy and precision but lower recall for stealthy or unseen threats (Ukwandu et al., 2020). This study's descriptive and ANOVA outcomes had converged with that established profile. Classical supervised families had shown strong precision and competitive accuracy yet remained limited in recall under replay, false data injection, and drift attacks. Earlier semi-supervised and unsupervised research had often reported that learning normality rather than attacks improves sensitivity to novel intrusions but can raise false positives if benign regime shifts are not captured. The present findings

had echoed that pattern by showing slightly higher recall for semi/unsupervised models with elevated nuisance rates relative to deep and hybrid families. Deep temporal and reconstruction models had produced strong recall and reduced false-positive rates, replicating earlier OT studies that highlight deep sequence learning as suitable for capturing multivariate dependencies and temporal coherence in cyber–physical systems (Peres et al., 2020). The significant model-family main effects for recall and F1, coupled with strong effect sizes, had supported that earlier evidence by demonstrating that deep and hybrid models delivered a practically meaningful advantage under industrial imbalance and stealth complexity. The study had also found that model latency differed by family, with deeper fusion architectures producing slightly higher detection delay. Prior industrial literature had commonly treated latency as a trade-off inherent in richer temporal modeling, observing that detection systems gain sensitivity by processing longer windows or deeper representations at the cost of modest computational delay. This study's latency differences had matched that known trade-off, while still remaining within plausible OT tolerance bands in the benchmark settings. The interaction between model family and feature strategy had been particularly important. Earlier fusion research had suggested that deep models derive larger benefits from fused evidence because they can learn cross-layer relations implicitly, while classical models benefit less due to limited representational depth. The significant interaction detected in this study had reaffirmed that expectation by showing that fusion gains were strongest in deep temporal/reconstruction and hybrid ensembles (Andrade et al., 2020). Taken together, the inferential profile had aligned with a mature body of industrial ML security work that positions deep, fusion-capable architectures as the most reliable high-recall defenders in cyber–physical control environments.

Figure 12: Industrial ML Cybersecurity Discussion Framework



Feature-strategy comparisons had advanced the discussion by demonstrating that cyber–physical fusion had been a major driver of improved detection quality, consistent with prior multi-stream OT security research. Earlier studies that compared cyber-only and process-only detectors had frequently concluded that each stream captures distinct attack observability, and that fused models reduce blind spots by integrating cyber intent traces with physical consequence traces (Yeboah-Ofori & Islam, 2019). This study had shown significant main effects for feature strategy and had observed that early and late fusion raised recall and F1 compared with single-stream pipelines. Fused strategies had also reduced false-positive rates relative to cyber-only anomaly learning, which had mirrored earlier claims that process context helps disambiguate benign traffic shifts from malicious intent. The descriptive and

marginal-mean patterns had reinforced this interpretation: fused deep models and hybrid ensembles had achieved the highest recall while keeping nuisance alarms lowest, an outcome widely reported in earlier cyber–physical fusion benchmarks on water and grid testbeds. The discussion had interpreted the superiority of fusion as arising from the observed cross-layer correlation strengthening under attack, which provides a measurable signal for fused learners to exploit. Earlier literature had characterized stealthy industrial intrusions as consistency-breaking events, where cyber and physical streams diverge from their normal joint rhythm (Echeverría et al., 2021). This study's findings had supported that characterization by demonstrating strong cross-layer covariance shifts in replay, false injection, and drift attacks and by showing that fused models captured these shifts more reliably. The difference between early and late fusion had also been treated in prior scholarship as a function of sampling mismatch and noise structure across streams (Althar & Samanta, 2021). This study had found that late-fusion hybrid ensembles produced slightly higher performance than early fusion, suggesting that specialized stream learners combined at the decision level managed heterogeneity better than raw concatenation alone. That pattern had been observed in earlier industrial fusion comparisons in which late fusion helped control collinearity and sampling-rate mismatch. The discussion had therefore framed fusion as both a statistical and operational advantage: it had improved recall for subtle attacks without inflating nuisance alarms, and it had done so in a manner consistent with earlier cyber–physical security theory that stresses joint evidence for reliable OT intrusion detection. The stability of fusion gains under imbalance mitigation and transfer checks had further supported the interpretation that fusion captures fundamental cross-layer structures rather than merely overfitting a single dataset (Lee, 2020).

Attack-type inference had offered additional depth by clarifying where models succeeded most strongly and where industrial stealth remained challenging, again aligning with established OT attack literature. Earlier industrial studies had often reported that abrupt cyber-layer disruptions like denial-of-service and explicit command injection are easier to detect because they perturb network rhythms or protocol semantics clearly (Yeboah-Ofori et al., 2021). This study had observed the highest recall and F1 for denial-of-service and command tampering, confirming the detectability of attacks that generate overt cyber or process shocks. Prior work had also emphasized that replay, false data injection, and stealth drift are more difficult, because adversaries preserve plausible cyber syntax or spread deviations gradually across variables to remain inside normal margins. This study had found significantly lower recall for these categories, with stealth drift yielding the lowest values, which had been consistent with earlier demonstrations that low-and-slow manipulation reduces statistical separability. The discussion had interpreted this hierarchy as a direct expression of industrial attack design: the more an intrusion mimics normal cyclic communication and physically feasible trajectories, the more evidence streams must be integrated and temporally modeled to surface inconsistencies. Fused deep and hybrid models had reduced, but not eliminated, the detection gap for stealth drift, a finding that had paralleled earlier results where deep temporal learning improves low-and-slow detection but still contends with benign process noise and regime transitions (Pramanik et al., 2019). The robustness checks had provided a realistic qualifier to these attack-type findings by showing that all models experienced some performance reduction under cross-dataset transfer. Earlier OT generalization studies have routinely warned that distribution shift across plants is a persistent barrier, driven by differences in protocol mix, sensor density, and operational regimes. This study had reproduced that challenge by documenting transfer drops, while also showing smaller declines for fusion-capable deep and hybrid families. The discussion had interpreted this resilience as evidence that fused representations capture more universal cyber–physical patterns compared with single-stream or label-bound classifiers. The attack-type results had therefore contributed a nuanced understanding to the literature: detection superiority is not uniform across threats, and model benefits are most pronounced where adversarial stealth exploits the cyber–physical split. By grounding attack sensitivity in measured recall differences and by linking these differences to known industrial attack mechanics, the discussion had placed the study's conclusions squarely within the established taxonomy of OT threat detectability (Tsakalidis et al., 2019).

Overall, the discussion had integrated descriptive patterns, measurement diagnostics, inferential comparisons, and robustness outcomes into a cohesive explanation of how machine learning–based

cybersecurity models safeguarded industrial automation and critical infrastructure in the tested contexts. This study had confirmed the central narrative emphasized in earlier OT security research: industrial data are highly structured in normal operation, severely imbalanced in label distribution, and vulnerable to adversaries who exploit determinism and cross-layer dependencies (Montasari et al., 2020). Classical supervised methods had remained valuable for detecting known attacks with high precision, reflecting their established role in industrial IDS baselines. Semi-supervised and unsupervised models had shown the expected advantage in novelty sensitivity coupled with threshold-driven false-alarm risk. Deep temporal and reconstruction models, especially when paired with cyber–physical fusion, had delivered the strongest balance of recall and nuisance control, aligning with a growing body of industrial evidence that views deep sequence learning and multiteam integration as optimal for CPS intrusion detection. The statistical interaction between model class and fusion strategy had strengthened that interpretation by showing that richer architectures derived greater value from cross-layer evidence. Attack-type differences had reiterated established detectability hierarchies in industrial settings, with abrupt cyber disruptions remaining easiest to surface and stealthy cross-layer manipulations remaining hardest (Foschini et al., 2021). Measurement and collinearity diagnostics had shown that reliable inference required domain-aligned constructs and redundancy control, a methodological stance that earlier industrial ML security work has advocated consistently. The robustness results had underscored that distribution shift and imbalance remain active constraints in OT intrusion detection, yet fusion-based deep and hybrid models had shown comparatively higher stability under those stresses. In sum, the comparative evidence had positioned cyber–physical fusion and deep/hybrid learning as the most empirically reliable safeguard against modern industrial threats in the benchmark cases evaluated. The convergence of these findings with earlier research streams had indicated that the study had extended, clarified, and quantitatively reinforced established OT cybersecurity knowledge while maintaining a clear focus on operationally meaningful outcomes such as recall, false-positive control, and latency under cyber–physical attack conditions (Välja et al., 2020).

CONCLUSION

Machine Learning–Based Cybersecurity Models for Safeguarding Industrial Automation and Critical Infrastructure Systems had been positioned as a decisive response to the cyber–physical risk profile of modern operational technology, where digital intrusions can translate into process disruption, equipment damage, and cascading service failures. Industrial automation environments had operated through layered control architectures that linked field sensors and actuators to PLC/RTU controllers, supervisory SCADA and DCS platforms, and human–machine interfaces connected through enterprise gateways, producing continuous cyber traffic and multivariate physical telemetry. These systems had generated highly periodic network patterns, deterministic command sequences, and tightly coupled sensor trajectories during normal operation, while attacks had introduced deviations that were often subtle, coordinated, and designed to remain plausible within single domains. Traditional rule-based defenses had remained limited by signature dependence, low sensitivity to novel or polymorphic threats, and high nuisance alarms during benign regime shifts, which had been especially costly in safety-critical processes requiring real-time reliability. The operational data realities of critical infrastructure—severe label scarcity, extreme class imbalance, and nonstationary yet structured process dynamics—had therefore made machine learning a natural analytical foundation for intrusion detection. Supervised models had shown strong precision and accuracy on known attacks but had weakened when labels were sparse or when threats were engineered for stealth. Semi-supervised and unsupervised anomaly detectors had addressed novelty by modeling normal baselines directly from benign telemetry, although their false-positive control had depended heavily on threshold calibration and mode awareness. Deep temporal predictors and reconstruction-based models had advanced detection by learning multivariate dependencies and long-horizon coherence embedded in control cycles, enabling stronger recognition of replay, false data injection, and drift manipulation that preserved normal packet syntax. Cyber–physical fusion strategies had further strengthened safeguards by integrating cyber evidence of intent with physical evidence of consequence, allowing models to detect cross-layer inconsistencies that single-stream methods often missed, and to reduce nuisance alarms by discounting benign cyber deviations that did not violate physical feasibility. Hybrid architectures that stacked classical and deep learners, embedded rule-guided feasibility constraints, or

used attention to weight cross-stream relevance had stabilized performance across heterogeneous plants and attack types. Quantitative evaluations across representative infrastructure datasets had consistently indicated that fused deep and hybrid families produced the highest recall and F1 scores while maintaining the lowest false-positive rates, with modest latency trade-offs acceptable within industrial monitoring windows. Attack-type comparisons had shown that abrupt cyber disruptions like denial-of-service and direct command tampering were detected more easily than stealthy low-and-slow or coordinated multi-point manipulations, reinforcing the need for temporal modeling and cross-layer evidence. Overall, machine learning–based cybersecurity models had served as statistically grounded, process-aware safeguards that matched the scale, structure, and adversarial complexity of industrial automation and critical infrastructure, providing measurable improvements in detecting cyber–physical threats within high-availability environments.

RECOMMENDATIONS

Recommendations for Machine Learning–Based Cybersecurity Models for Safeguarding Industrial Automation and Critical Infrastructure Systems should be framed around practical deployment realism, measurement discipline, and cyber–physical alignment, because industrial environments impose stricter constraints than conventional IT security. First, model selection should prioritize cyber–physical fusion rather than single-stream detection, since industrial attacks frequently preserve normal-looking network syntax while altering physical trajectories or, conversely, disrupt traffic timing without immediately shifting process variables. Fusion should be implemented with complementary feature design, avoiding redundant indicators that inflate multicollinearity; early fusion can be used when cyber and process telemetry are tightly synchronized and sampled comparably, whereas late fusion should be favored when streams differ in sampling rate, noise profile, or dimensionality. Second, deep temporal and reconstruction-based architectures should be adopted as primary detection engines for high-risk plants, because these models capture multivariate dependency and control-cycle coherence that classical models often miss; however, deployment should include lightweight variants or edge-optimized distillation to meet real-time constraints and limited compute at OT gateways. Third, supervised models should remain in the security stack as high-precision classifiers for known attack families, but they should be paired with semi-supervised or unsupervised anomaly detectors to broaden coverage for novelty and polymorphic threats, especially in sites where attack labels are scarce. Fourth, thresholding and alarm calibration should be treated as an operational control problem rather than a one-time tuning step; adaptive thresholds, mode-aware baselines, and temporal smoothing should be used so that legitimate regime transitions—startup, shutdown, maintenance, load shifting, and recalibration—do not trigger nuisance alarms that erode operator trust. Fifth, model maintenance should incorporate drift management as a formal requirement. Baselines should be refreshed on a scheduled cadence aligned with production cycles, and drift indicators should be monitored so that retraining occurs before performance decay becomes safety-relevant. Sixth, robustness validation should be extended beyond single-dataset benchmarking. Cross-site testing and transfer trials should be included prior to deployment, and imbalance-sensitive training strategies should be standard practice, since normal operations dominate OT logs and rare attacks can be suppressed without mitigation. Seventh, interpretability support should be integrated into detection dashboards through feature-attribution summaries, cross-layer inconsistency indicators, and process-feasibility explanations, because industrial response depends on rapid human verification under stress. Finally, governance and architecture should ensure that ML detection sits within defense-in-depth rather than replacing traditional safeguards. Network segmentation, protocol allow listing, secure remote access, backup safety interlocks, and incident playbooks should operate alongside ML intrusion detection to prevent single-point reliance on learned models. Taken together, these recommendations emphasize that effective ML-based industrial cybersecurity requires fused evidence, temporally aware learning, continuous calibration, drift control, robustness checks, operator-centered explanations, and layered security architecture, enabling measurable protection gains without compromising industrial continuity and safety.

LIMITATIONS

Limitations of the study on Machine Learning–Based Cybersecurity Models for Safeguarding Industrial Automation and Critical Infrastructure Systems had arisen from the nature of available industrial data,

the experimental framing required for quantitative comparison, and the inherent diversity of real-world operational technology environments. The first limitation had been dataset representativeness. Although benchmark cyber–physical datasets had provided synchronized network traffic, control-command logs, and process telemetry, they had still reflected limited plant configurations compared with the global variety of critical infrastructure. Differences in vendor implementations, protocol customizations, sensor densities, and control philosophies across real facilities had not been fully captured, so generalization beyond the evaluated domains had remained partially constrained. A second limitation had involved label realism. Attack windows in benchmark cases had been either simulated or staged under controlled conditions, which had supported repeatable evaluation but had not perfectly mirrored the improvisational behavior, timing irregularities, and blended tactics seen in live adversarial campaigns. This limitation had been amplified by the scarcity of multi-stage hybrid attacks in the sampled population, which had reduced statistical power for fine-grained inference on complex coordinated intrusions. Third, the study had relied on fixed windowing aligned to control cycles; while this had standardized learning inputs, it had imposed a discretization that may not match all industrial timing regimes, especially in plants with variable cycle lengths or asynchronous event-driven control. Fourth, feature engineering had been optimized for the selected datasets and protocols, meaning that some engineered indicators—such as specific function-code distributions or residual structures—might require redesign in sites that use different industrial protocols, encrypted channels, or alternative process topologies. Fifth, model comparison had focused on performance metrics derived from historical replay, so detection latency and computational overhead had been measured under offline conditions. Real-time deployment can introduce additional delays from edge processing, network jitter, and resource contention, which were not fully reproduced in the experimental environment. Sixth, robustness testing had included imbalance mitigation and cross-dataset transfer, yet the transfer scope had remained bounded to available benchmark pairs; broader cross-sector transfer, especially across geopolitical threat contexts or highly specialized industrial processes, had not been directly measured. Seventh, adversarial machine learning threats had not been exhaustively explored, so the evaluated models' resistance to poisoning, evasion, or mimicry strategies crafted specifically against learned detectors had remained only partially assessed. Finally, interpretability had been treated as an enabling consideration rather than a quantified outcome, leaving limited empirical insight into how well different model families supported operator trust and rapid incident triage under real operational stress. Collectively, these limitations had indicated that while the study provided strong quantitative evidence for comparative detection behavior in representative cyber–physical benchmarks, caution had been required when extending results to all industrial automation contexts without site-specific recalibration, expanded adversarial realism, and deeper real-time validation.

REFERENCES

[1]. Abbas, S. G., Hashmat, F., & Shah, G. A. (2020). A multi-layer industrial-IoT attack taxonomy: Layers, dimensions, techniques and application. 2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom),

[2]. Abdul, H. (2023). Artificial Intelligence in Product Marketing: Transforming Customer Experience And Market Segmentation. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *3*(1), 132–159. https://doi.org/10.63125/58npbx97

[3]. Abdulla, M., & Md. Wahid Zaman, R. (2023). Quantitative Study On Workflow Optimization Through Data Analytics In U.S. Digital Enterprises. *American Journal of Interdisciplinary Studies*, *4*(03), 136–165. https://doi.org/10.63125/y2qshd31

[4]. Abhale, A. B., & Manivannan, S. (2020). Supervised machine learning classification algorithmic approach for finding anomaly type of intrusion detection in wireless sensor network. *Optical Memory and Neural Networks*, *29*(3), 244-256.

[5]. Abikoye, O. C., Bajeh, A. O., Awotunde, J. B., Ameen, A. O., Mojeed, H. A., Abdulraheem, M., Oladipo, I. D., & Salihu, S. A. (2021). Application of internet of thing and cyber physical system in Industry 4.0 smart manufacturing. In *Emergence of Cyber Physical System and IoT in Smart Automation and Robotics: Computer Engineering in Automation* (pp. 203-217). Springer.

[6]. Aboueata, N., Alrasbi, S., Erbad, A., Kassler, A., & Bhamare, D. (2019). Supervised machine learning techniques for efficient network intrusion detection. 2019 28th international conference on computer communication and networks (ICCCN),

[7]. Afaq, A., Haider, N., Baig, M. Z., Khan, K. S., Imran, M., & Razzak, I. (2021). Machine learning for 5G security: Architecture, recent advances, and challenges. *Ad Hoc Networks*, *123*, 102667.

[8]. Ahanger, A. S., Khan, S. M., & Masoodi, F. (2021). An effective intrusion detection system using supervised machine learning techniques. 2021 5th International Conference on Computing Methodologies and Communication (ICCMC),
[9]. Ahmad, S., Lavin, A., Purdy, S., & Agha, Z. (2017). Unsupervised real-time anomaly detection for streaming data. *Neurocomputing*, *262*, 134-147.
[10]. Ahmadi-Assalemi, G., Al-Khateeb, H., Epiphaniou, G., & Maple, C. (2020). Cyber resilience and incident response in smart cities: A systematic literature review. *Smart Cities*, *3*(3), 894-927.
[11]. Ainam, J.-P., Qin, K., Liu, G., & Luo, G. (2019). Sparse label smoothing regularization for person re-identification. *Ieee access*, *7*, 27899-27910.
[12]. Aiyanyo, I. D., Samuel, H., & Lim, H. (2020). A systematic review of defensive and offensive cybersecurity with machine learning. *Applied sciences*, *10*(17), 5811.
[13]. Al-Abassi, A., Karimipour, H., Dehghantanha, A., & Parizi, R. M. (2020). An ensemble deep learning-based cyber-attack detection in industrial control system. *Ieee access*, *8*, 83965-83973.
[14]. Al-Jarrah, O. Y., Al-Hammdi, Y., Yoo, P. D., Muhaidat, S., & Al-Qutayri, M. (2018). Semi-supervised multi-layered clustering model for intrusion detection. *Digital Communications and Networks*, *4*(4), 277-286.
[15]. Al-Mhiqani, M. N., Ahmad, R., Zainal Abidin, Z., Yassin, W., Hassan, A., Abdulkareem, K. H., Ali, N. S., & Yunos, Z. (2020). A review of insider threat detection: Classification, machine learning techniques, datasets, open challenges, and recommendations. *Applied sciences*, *10*(15), 5208.
[16]. Alazab, M., & Tang, M. (2019). *Deep learning applications for cyber security*. Springer.
[17]. Alem, S., Espes, D., Martin, E., Nana, L., & de Lamotte, F. (2020). New dataset for industry 4.0 to address the change in threat landscape. International Conference on Risks and Security of Internet and Systems,
[18]. Ali, S., Al Balushi, T., Nadir, Z., & Hussain, O. K. (2018). *Cyber security for cyber physical systems* (Vol. 768). Springer.
[19]. Alimi, O. A., Ouahada, K., Abu-Mahfouz, A. M., Rimer, S., & Alimi, K. O. A. (2021). A review of research works on supervised learning algorithms for SCADA intrusion detection and classification. *Sustainability*, *13*(17), 9597.
[20]. Almeaibed, S., Al-Rubaye, S., Tsourdos, A., & Avdelidis, N. P. (2021). Digital twin analysis to promote safety and security in autonomous vehicles. *IEEE Communications Standards Magazine*, *5*(1), 40-46.
[21]. Althar, R. R., & Samanta, D. (2021). The realist approach for evaluation of computational intelligence in software engineering. *Innovations in Systems and Software Engineering*, *17*(1), 17-27.
[22]. Andrade, R. O., Yoo, S. G., Tello-Oquendo, L., & Ortiz-Garcés, I. (2020). A comprehensive study of the IoT cybersecurity in smart cities. *Ieee access*, *8*, 228922-228941.
[23]. Ani, U. P. D., He, H., & Tiwari, A. (2017). Review of cybersecurity issues in industrial critical infrastructure: manufacturing in perspective. *Journal of Cyber Security Technology*, *1*(1), 32-74.
[24]. Ani, U. P. D., Watson, J. M., Green, B., Craggs, B., & Nurse, J. R. (2021). Design considerations for building credible security testbeds: Perspectives from industrial control system use cases. *Journal of Cyber Security Technology*, *5*(2), 71-119.
[25]. Anthi, E., Williams, L., Rhode, M., Burnap, P., & Wedgbury, A. (2021). Adversarial attacks on machine learning cybersecurity defences in industrial control systems. *Journal of Information Security and Applications*, *58*, 102717.
[26]. Anton, S. D. D., Sinha, S., & Schotten, H. D. (2019). Anomaly-based intrusion detection in industrial data with SVM and random forests. 2019 International conference on software, telecommunications and computer networks (SoftCOM),
[27]. Arfan, U., Sai Praveen, K., & Alifa Majumder, N. (2021). Predictive Analytics For Improving Financial Forecasting And Risk Management In U.S. Capital Markets. *American Journal of Interdisciplinary Studies*, *2*(04), 69–100. https://doi.org/10.63125/tbw49w69
[28]. Arfan, U., Tahsina, A., Md Mostafizur, R., & Md, W. (2023). Impact Of GFMIS-Driven Financial Transparency On Strategic Marketing Decisions In Government Agencies. *Review of Applied Science and Technology*, *2*(01), 85-112. https://doi.org/10.63125/8nqhhm56
[29]. Bergmann, P., Batzner, K., Fauser, M., Sattlegger, D., & Steger, C. (2021). The MVTec anomaly detection dataset: a comprehensive real-world dataset for unsupervised anomaly detection. *International Journal of Computer Vision*, *129*(4), 1038-1059.
[30]. Bhamare, D., Zolanvari, M., Erbad, A., Jain, R., Khan, K., & Meskin, N. (2020). Cybersecurity for industrial control systems: A survey. *computers & security*, *89*, 101677.
[31]. Bo, X., Chen, X., Li, H., Dong, Y., Qu, Z., Wang, L., & Li, Y. (2021). Modeling method for the coupling relations of microgrid cyber-physical systems driven by hybrid spatiotemporal events. *Ieee access*, *9*, 19619-19631.
[32]. Bruzgiene, R., & Jurgilas, K. (2021). Securing remote access to information systems of critical infrastructure using two-factor authentication. *Electronics*, *10*(15), 1819.
[33]. Chen, C.-Y., Hasan, M., & Mohan, S. (2018). Securing real-time internet-of-things. *Sensors*, *18*(12), 4356.
[34]. Chen, G., Wang, P., Feng, B., Li, Y., & Liu, D. (2020). The framework design of smart factory in discrete manufacturing industry based on cyber-physical system. *International Journal of Computer Integrated Manufacturing*, *33*(1), 79-101.
[35]. Chen, S., Wang, J., Li, H., Wang, Z., Liu, F., & Li, S. (2020). Top-down human-cyber-physical data fusion based on reinforcement learning. *Ieee access*, *8*, 134233-134245.
[36]. Chen, X., Deng, L., Huang, F., Zhang, C., Zhang, Z., Zhao, Y., & Zheng, K. (2021). Daemon: Unsupervised anomaly detection and interpretation for multivariate time series. 2021 IEEE 37th International Conference on Data Engineering (ICDE),

[37]. Conti, M., Donadel, D., & Turrin, F. (2021). A survey on industrial control system testbeds and datasets for security research. *IEEE Communications Surveys & Tutorials*, *23*(4), 2248-2294.
[38]. Dai, M., Guo, W., & Feng, X. (2020). Over-smoothing algorithm and its application to gcn semi-supervised classification. International Conference of Pioneering Computer Scientists, Engineers and Educators,
[39]. Damianov, D., & Demirova, S. (2018). Principles of designing automated logistics systems-hybrid component of cyber-physical systems. 2018 International Conference on High Technology for Sustainable Development (HiTech),
[40]. Derbyshire, R., Green, B., Prince, D., Mauthe, A., & Hutchison, D. (2018). An analysis of cyber security attack taxonomies. 2018 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW),
[41]. Dhirani, L. L., Armstrong, E., & Newe, T. (2021). Industrial IoT, cyber threats, and standards landscape: Evaluation and roadmap. *Sensors*, *21*(11), 3901.
[42]. Dixit, P., & Silakari, S. (2021). Deep learning algorithms for cybersecurity applications: A technological and status review. *Computer science review*, *39*, 100317.
[43]. Djenna, A., Harous, S., & Saidouni, D. E. (2021). Internet of things meet internet of threats: New concern cyber security issues of critical cyber infrastructure. *Applied sciences*, *11*(10), 4580.
[44]. Dobaj, J., Krisper, M., & Macher, G. (2019). Towards cyber-physical infrastructure as-a-service (CPIaaS) in the era of industry 4.0. European Conference on Software Process Improvement,
[45]. Echeverría, A., Cevallos, C., Ortiz-Garces, I., & Andrade, R. O. (2021). Cybersecurity model based on hardening for secure internet of things implementation. *Applied sciences*, *11*(7), 3260.
[46]. Farzad, A., & Gulliver, T. A. (2020). Unsupervised log message anomaly detection. *ICT Express*, *6*(3), 229-237.
[47]. Fausto, A., Gaggero, G. B., Patrone, F., Girdinio, P., & Marchese, M. (2021). Toward the integration of cyber and physical security monitoring systems for critical infrastructures. *Sensors*, *21*(21), 6970.
[48]. Ferdous Ara, A. (2021). Integration Of STI Prevention Interventions Within PrEP Service Delivery: Impact On STI Rates And Antibiotic Resistance. *International Journal of Scientific Interdisciplinary Research*, *2*(2), 63–97. https://doi.org/10.63125/65143m72
[49]. Ferdous Ara, A., & Beatrice Onyinyechi, M. (2023). Long-Term Epidemiologic Trends Of STIs PRE- and POST-PrEP Introduction: A National Time-Series Analysis. *American Journal of Health and Medical Sciences*, *4*(02), 01–35. https://doi.org/10.63125/mp153d97
[50]. Foschini, L., Mignardi, V., Montanari, R., & Scotece, D. (2021). An SDN-enabled architecture for IT/OT converged networks: A proposal and qualitative analysis under DDoS attacks. *Future internet*, *13*(10), 258.
[51]. Fredriksson, T., Mattos, D. I., Bosch, J., & Olsson, H. H. (2021). Assessing the suitability of semi-supervised learning datasets using item response theory. 2021 47th Euromicro conference on software engineering and advanced applications (SEAA),
[52]. Georgescu, T.-M. (2020). Natural language processing model for automatic analysis of cybersecurity-related documents. *Symmetry*, *12*(3), 354.
[53]. González-Granadillo, G., González-Zarzosa, S., & Diaz, R. (2021). Security information and event management (SIEM): analysis, trends, and usage in critical infrastructures. *Sensors*, *21*(14), 4759.
[54]. Groshev, M., Guimaraes, C., Martín-Pérez, J., & de la Oliva, A. (2021). Toward intelligent cyber-physical systems: Digital twin meets artificial intelligence. *IEEE Communications Magazine*, *59*(8), 14-20.
[55]. Hamouda, D., Ferrag, M. A., Benhamida, N., & Seridi, H. (2021). Intrusion detection systems for industrial internet of things: A survey. 2021 International Conference on Theoretical and Applicative Aspects of Computer Science (ICTAACS),
[56]. Hao, W., Yang, T., & Yang, Q. (2021). Hybrid statistical-machine learning for real-time anomaly detection in industrial cyber–physical systems. *IEEE Transactions on Automation Science and Engineering*, *20*(1), 32-46.
[57]. Hatzivasilis, G., Ioannidis, S., Smyrlis, M., Spanoudakis, G., Frati, F., Goeke, L., Hildebrandt, T., Tsakirakis, G., Oikonomou, F., & Leftheriotis, G. (2020). Modern aspects of cyber-security training and continuous adaptation of programmes to trainees. *Applied sciences*, *10*(16), 5702.
[58]. Hoffmann, M. W., Malakuti, S., Grüner, S., Finster, S., Gebhardt, J., Tan, R., Schindler, T., & Gamer, T. (2021). Developing industrial cps: A multi-disciplinary challenge. *Sensors*, *21*(6), 1991.
[59]. Hossain, E., Khan, I., Un-Noor, F., Sikander, S. S., & Sunny, M. S. H. (2019). Application of big data and machine learning in smart grid, and associated security concerns: A review. *Ieee access*, *7*, 13960-13988.
[60]. Hwang, R.-H., Peng, M.-C., Huang, C.-W., Lin, P.-C., & Nguyen, V.-L. (2020). An unsupervised deep learning model for early network traffic anomaly detection. *Ieee access*, *8*, 30387-30399.
[61]. Jadidi, Z., & Lu, Y. (2021). A threat hunting framework for industrial control systems. *Ieee access*, *9*, 164118-164130.
[62]. Jahid, M. K. A. S. R. (2021). Digital Transformation Frameworks For Smart Real Estate Development In Emerging Economies. *Review of Applied Science and Technology*, *6*(1), 139–182. https://doi.org/10.63125/cd09ne09
[63]. Jayalaxmi, P., Saha, R., Kumar, G., Kumar, N., & Kim, T.-H. (2021). A taxonomy of security issues in Industrial Internet-of-Things: Scoping review for existing solutions, future implications, and research challenges. *Ieee access*, *9*, 25344-25359.
[64]. Karabiyik, U., & Akkaya, K. (2019). Digital forensics for IoT and WSNS. In *Mission-Oriented Sensor Networks and Systems: Art and Science: Volume 2: Advances* (pp. 171-207). Springer.
[65]. Khan, I. A., Moustafa, N., Pi, D., Sallam, K. M., Zomaya, A. Y., & Li, B. (2021). A new explainable deep learning framework for cyber threat discovery in industrial IoT networks. *IEEE Internet of Things Journal*, *9*(13), 11604-11613.
[66]. Krishna, R. R., Priyadarshini, A., Jha, A. V., Appasani, B., Srinivasulu, A., & Bizon, N. (2021). State-of-the-art review on IoT threats and attacks: Taxonomy, challenges and solutions. *Sustainability*, *13*(16), 9463.

[67]. Lee, I. (2020). Internet of Things (IoT) cybersecurity: Literature review and IoT cyber risk management. *Future internet*, *12*(9), 157.
[68]. Li, B., Wu, Y., Song, J., Lu, R., Li, T., & Zhao, L. (2020). DeepFed: Federated deep learning for intrusion detection in industrial cyber–physical systems. *IEEE Transactions on Industrial Informatics*, *17*(8), 5615-5624.
[69]. Li, M., Xue, Y., Ni, M., & Li, X. (2020). Modeling and hybrid calculation architecture for cyber physical power systems. *Ieee access*, *8*, 138251-138263.
[70]. Li, X., Wen, C., Cao, Q., Du, Y., & Fang, Y. (2021). RETRACTED: A novel semi-supervised method for airborne LiDAR point cloud classification. In: Elsevier.
[71]. Liang, W., Li, K.-C., Long, J., Kui, X., & Zomaya, A. Y. (2019). An industrial network intrusion detection algorithm based on multifeature data clustering optimization model. *IEEE Transactions on Industrial Informatics*, *16*(3), 2063-2071.
[72]. Lopez, J., Liefer, N. C., Busho, C. R., & Temple, M. A. (2017). Enhancing critical infrastructure and key resources (CIKR) level-0 physical process security using field device distinct native attribute features. *IEEE Transactions on Information Forensics and Security*, *13*(5), 1215-1229.
[73]. Makrakis, G. M., Kolias, C., Kambourakis, G., Rieger, C., & Benjamin, J. (2021). Industrial and critical infrastructure security: Technical analysis of real-life security incidents. *Ieee access*, *9*, 165295-165325.
[74]. Markopoulou, D., & Papakonstantinou, V. (2021). The regulatory framework for the protection of critical infrastructures against cyberthreats: Identifying shortcomings and addressing future challenges: The case of the health sector in particular. *Computer law & security review*, *41*, 105502.
[75]. McKee, D. W., Clement, S. J., Almutairi, J., & Xu, J. (2017). Massive-scale automation in cyber-physical systems: Vision & challenges. 2017 IEEE 13th International Symposium on Autonomous Decentralized System (ISADS),
[76]. Md Al Amin, K., & Md Mesbaul, H. (2023). Smart Hybrid Manufacturing: A Combination Of Additive, Subtractive, And Lean Techniques For Agile Production Systems. *Journal of Sustainable Development and Policy*, 2(04), 174-217. https://doi.org/10.63125/7rb1zz78
[77]. Md Ariful, I., & Efat Ara, H. (2022). Advances And Limitations Of Fracture Mechanics–Based Fatigue Life Prediction Approaches For Structural Integrity Assessment: A Systematic Review. *American Journal of Interdisciplinary Studies*, *3*(03), 68-98. https://doi.org/10.63125/fg8ae957
[78]. Md Arman, H., & Md.Kamrul, K. (2022). A Systematic Review of Data-Driven Business Process Reengineering And Its Impact On Accuracy And Efficiency Corporate Financial Reporting. *International Journal of Business and Economics Insights*, *2*(4), 01–41. https://doi.org/10.63125/btx52a36
[79]. Md Foysal, H., & Aditya, D. (2023). Smart Continuous Improvement With Artificial Intelligence, Big Data, And Lean Tools For Zero Defect Manufacturing Systems. *American Journal of Scholarly Research and Innovation*, *2*(01), 254–282. https://doi.org/10.63125/6cak0s21
[80]. Md Hamidur, R. (2023). Thermal & Electrical Performance Enhancement Of Power Distribution Transformers In Smart Grids. *American Journal of Scholarly Research and Innovation*, *2*(01), 283–313. https://doi.org/10.63125/n2p6y628
[81]. Md Harun-Or-Rashid, M., Mst. Shahrin, S., & Sai Praveen, K. (2023). Integration Of IOT And EDGE Computing For Low-Latency Data Analytics In Smart Cities And IOT Networks. *Journal of Sustainable Development and Policy*, *2*(03), 01-33. https://doi.org/10.63125/004h7m29
[82]. Md Mesbaul, H., & Md. Tahmid Farabe, S. (2022). Implementing Sustainable Supply Chain Practices In Global Apparel Retail: A Systematic Review Of Current Trends. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *2*(1), 332–363. https://doi.org/10.63125/nen7vd57
[83]. Md Musfiqur, R., & Md.Kamrul, K. (2023). Mechanisms By Which AI-Enabled Crm Systems Influence Customer Retention And Overall Business Performance: A Systematic Literature Review Of Empirical Findings. *International Journal of Business and Economics Insights*, *3*(1), 31-67. https://doi.org/10.63125/qqe2bm11
[84]. Md Muzahidul, I., & Md Mohaiminul, H. (2023). Explainable AI (XAI) Models For Cloud-Based Business Intelligence: Ensuring Compliance And Secure Decision-Making. *American Journal of Interdisciplinary Studies*, *4*(03), 208–249. https://doi.org/10.63125/5etfhh77
[85]. Md Nahid, H. (2022). Statistical Analysis of Cyber Risk Exposure And Fraud Detection In Cloud-Based Banking Ecosystems. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *2*(1), 289–331. https://doi.org/10.63125/9wf91068
[86]. Md Sarwar Hossain, S., & Md Milon, M. (2022). Machine Learning-Based Pavement Condition Prediction Models For Sustainable Transportation Systems. *American Journal of Interdisciplinary Studies*, *3*(01), 31–64. https://doi.org/10.63125/1jsmkg92
[87]. Md. Abdur, R., & Zamal Haider, S. (2022). Assessment Of Data-Driven Vendor Performance Evaluation In Retail Supply Chains Analyzing Metrics, Scorecards, And Contract Management Tools. *Journal of Sustainable Development and Policy*, *1*(04), 71-116. https://doi.org/10.63125/2a641k35
[88]. Md. Al Amin, K., & Sai Praveen, K. (2023). The Role Of Industrial Engineering In Advancing Sustainable Manufacturing And Quality Compliance In Global Engineering Systems. *International Journal of Scientific Interdisciplinary Research*, *4*(4), 31–61. https://doi.org/10.63125/8w1vk676
[89]. Md. Hasan, I., & Ashraful, I. (2023). The Effect Of Production Planning Efficiency On Delivery Timelines In U.S. Apparel Imports. *Journal of Sustainable Development and Policy*, 2(04), 35-73. https://doi.org/10.63125/sg472m51

[90]. Md. Jobayer Ibne, S., & Md. Kamrul, K. (2023). Automating NIST 800-53 Control Implementation: A Cross-Sector Review Of Enterprise Security Toolkits. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *3*(1), 160–195. https://doi.org/10.63125/prkw8r07

[91]. Md.Akbar, H., & Farzana, A. (2021). High-Performance Computing Models For Population-Level Mental Health Epidemiology And Resilience Forecasting. *American Journal of Health and Medical Sciences*, *2*(02), 01–33. https://doi.org/10.63125/k9d5h638

[92]. Men, J., Lv, Z., Zhou, X., Han, Z., Xian, H., & Song, Y.-N. (2020). Machine learning methods for industrial protocol security analysis: Issues, taxonomy, and directions. *Ieee access*, *8*, 83842-83857.

[93]. Merkle, L., Segura, A. S., Grummel, J. T., & Lienkamp, M. (2019). Architecture of a digital twin for enabling digital services for battery systems. 2019 IEEE international conference on industrial cyber physical systems (ICPS),

[94]. Mohammad Mushfequr, R., & Ashraful, I. (2023). Automation And Risk Mitigation in Healthcare Claims: Policy And Compliance Implications. *Review of Applied Science and Technology*, *2*(04), 124–157. https://doi.org/10.63125/v73gyg14

[95]. Mohammad Mushfequr, R., & Sai Praveen, K. (2022). Quantitative Investigation Of Information Security Challenges In U.S. Healthcare Payment Ecosystems. *International Journal of Business and Economics Insights*, *2*(4), 42–73. https://doi.org/10.63125/gcg0fs06

[96]. Mokhtari, S., Abbaspour, A., Yen, K. K., & Sargolzaei, A. (2021). A machine learning approach for anomaly detection in industrial control systems based on measurement data. *Electronics*, *10*(4), 407.

[97]. Montasari, R., Carroll, F., Macdonald, S., Jahankhani, H., Hosseinian-Far, A., & Daneshkhah, A. (2020). Application of artificial intelligence and machine learning in producing actionable cyber threat intelligence. In *Digital forensic investigation of internet of things (IoT) devices* (pp. 47-64). Springer.

[98]. Mortuza, M. M. G., & Rauf, M. A. (2022). Industry 4.0: An Empirical Analysis of Sustainable Business Performance Model Of Bangladeshi Electronic Organisations. *International Journal of Economy and Innovation*. https://gospodarkainnowacje.pl/index.php/issue_view_32/article/view/826

[99]. Moustafa, N., Adi, E., Turnbull, B., & Hu, J. (2018). A new threat intelligence scheme for safeguarding industry 4.0 systems. *Ieee access*, *6*, 32910-32924.

[100]. MR, G. R., Ahmed, C. M., & Mathur, A. (2021). Machine learning for intrusion detection in industrial control systems: challenges and lessons from experimental evaluation. *Cybersecurity*, *4*(1), 27.

[101]. Muna, A.-H., Moustafa, N., & Sitnikova, E. (2018). Identification of malicious activities in industrial internet of things based on deep learning models. *Journal of Information Security and Applications*, *41*, 1-11.

[102]. Munir, M., Siddiqui, S. A., Chattha, M. A., Dengel, A., & Ahmed, S. (2019). FuseAD: unsupervised anomaly detection in streaming sensors data by fusing statistical and deep learning models. *Sensors*, *19*(11), 2451.

[103]. Munir, M., Siddiqui, S. A., Dengel, A., & Ahmed, S. (2018). DeepAnT: A deep learning approach for unsupervised anomaly detection in time series. *Ieee access*, *7*, 1991-2005.

[104]. Nassif, A. B., Talib, M. A., Nasir, Q., & Dakalbab, F. M. (2021). Machine learning for anomaly detection: A systematic review. *Ieee access*, *9*, 78658-78700.

[105]. Nguyen, T. N. (2018). The challenges in ml-based security for sdn. 2018 2nd Cyber Security in Networking Conference (CSNet),

[106]. Noorizadeh, M., Shakerpour, M., Meskin, N., Unal, D., & Khorasani, K. (2021). A cyber-security methodology for a cyber-physical industrial control system testbed. *Ieee access*, *9*, 16239-16253.

[107]. Pankaz Roy, S., & Md. Kamrul, K. (2023). HACCP and ISO Frameworks For Enhancing Biosecurity In Global Food Distribution Chains. *American Journal of Scholarly Research and Innovation*, *2*(01), 314–356. https://doi.org/10.63125/9pbp4h37

[108]. Pereira, J., & Silveira, M. (2019). Learning representations from healthcare time series data for unsupervised anomaly detection. 2019 IEEE international conference on big data and smart computing (BigComp),

[109]. Peres, R. S., Jia, X., Lee, J., Sun, K., Colombo, A. W., & Barata, J. (2020). Industrial artificial intelligence in industry 4.0-systematic review, challenges and outlook. *Ieee access*, *8*, 220121-220139.

[110]. Podgorski, D., Majchrzycka, K., Dąbrowska, A., Gralewicz, G., & Okrasa, M. (2017). Towards a conceptual framework of OSH risk management in smart working environments based on smart PPE, ambient intelligence and the Internet of Things technologies. *International Journal of Occupational Safety and Ergonomics*, *23*(1), 1-20.

[111]. Pordelkhaki, M., Fouad, S., & Josephs, M. (2021). Intrusion detection for industrial control systems by machine learning using privileged information. 2021 IEEE International Conference on Intelligence and Security Informatics (ISI),

[112]. Pramanik, H. S., Kirtania, M., & Pani, A. K. (2019). Essence of digital transformation—Manifestations at large financial institutions from North America. *Future Generation Computer Systems*, *95*, 323-343.

[113]. Qu, Y., Ming, X., Liu, Z., Zhang, X., & Hou, Z. (2019). Smart manufacturing systems: state of the art and future trends. *The international journal of advanced manufacturing technology*, *103*(9), 3751-3768.

[114]. Radanliev, P., De Roure, D., Page, K., Nurse, J. R., Mantilla Montalvo, R., Santos, O., Maddox, L. T., & Burnap, P. (2020). Cyber risk at the edge: current and future trends on cyber risk analytics and artificial intelligence in the industrial internet of things and industry 4.0 supply chains. *Cybersecurity*, *3*(1), 13.

[115]. Rai, R., & Sahu, C. K. (2020). Driven by data or derived through physics? a review of hybrid physics guided machine learning techniques with cyber-physical system (cps) focus. *Ieee access*, *8*, 71050-71073.

[116]. Rakibul, H., & Samia, A. (2022). Information System-Based Decision Support Tools: A Systematic Review Of Strategic Applications In Service-Oriented Enterprises. *Review of Applied Science and Technology*, *1*(04), 26-65. https://doi.org/10.63125/w3cevz78
[117]. Rathore, S., & Park, J. H. (2020). A blockchain-based deep learning approach for cyber security in next generation industrial cyber-physical systems. *IEEE Transactions on Industrial Informatics*, *17*(8), 5522-5532.
[118]. Rawindaran, N., Jayal, A., & Prakash, E. (2021). Machine learning cybersecurity adoption in small and medium enterprises in developed countries. *Computers*, *10*(11), 150.
[119]. Reza, M., Vorobyova, K., & Rauf, M. (2021). The effect of total rewards system on the performance of employees with a moderating effect of psychological empowerment and the mediation of motivation in the leather industry of Bangladesh. *Engineering Letters*, *29*, 1-29.
[120]. Rodofile, N. R., Radke, K., & Foo, E. (2019). Extending the cyber-attack landscape for SCADA-based critical infrastructure. *International Journal of Critical Infrastructure Protection*, *25*, 14-35.
[121]. Rony, M. A., & Ashraful, I. (2022). Big Data And Engineering Analytics Pipelines For Smart Manufacturing: Enhancing Efficiency, Quality, And Predictive Maintenance. *American Journal of Scholarly Research and Innovation*, *1*(02), 59–85. https://doi.org/10.63125/rze0my79
[122]. Saba, A., Shaikat, B., & Tonoy Kanti, C. (2023). Integration Of Artificial Intelligence And Advanced Computing To Develop Resilient Cyber Defense Systems. *Journal of Sustainable Development and Policy*, *2*(04), 74-107. https://doi.org/10.63125/rxyc6y88
[123]. Saba, A., & Tonoy Kanti, C. (2023). Explainable Artificial Intelligence (XAI) Approaches For Cyber Risk Assessment In Financial Services. *American Journal of Interdisciplinary Studies*, *4*(03), 96-135. https://doi.org/10.63125/3gjcb322
[124]. Saikat, S. (2021). Real-Time Fault Detection in Industrial Assets Using Advanced Vibration Dynamics And Stress Analysis Modeling. *American Journal of Interdisciplinary Studies*, *2*(04), 39–68. https://doi.org/10.63125/0h163429
[125]. Saikat, S. (2022). CFD-Based Investigation of Heat Transfer Efficiency In Renewable Energy Systems. *International Journal of Scientific Interdisciplinary Research*, *1*(01), 129–162. https://doi.org/10.63125/ttw40456
[126]. Salloum, S. A., Alshurideh, M., Elnagar, A., & Shaalan, K. (2020). Machine learning and deep learning techniques for cybersecurity: a review. The International Conference on Artificial Intelligence and Computer Vision,
[127]. Sarker, I. H. (2021). Deep cybersecurity: a comprehensive overview from neural network and deep learning perspective. *SN Computer Science*, *2*(3), 154.
[128]. Schlegl, T., Seeböck, P., Waldstein, S. M., Langs, G., & Schmidt-Erfurth, U. (2019). f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks. *Medical image analysis*, *54*, 30-44.
[129]. Schlette, D., Menges, F., Baumer, T., & Pernul, G. (2020). Security enumerations for cyber-physical systems. IFIP Annual Conference on Data and Applications Security and Privacy,
[130]. Serpanos, D. (2018). The cyber-physical systems revolution. *Computer*, *51*(3), 70-73.
[131]. Shaikh, S., & Aditya, D. (2021). Federated Learning-Driven Predictive Quality Analytics and Supply Chain Optimization In Distributed Manufacturing Networks. *Review of Applied Science and Technology*, *6*(1), 74-107. https://doi.org/10.63125/k18cbz55
[132]. Shaikh, S., & Md. Tahmid Farabe, S. (2023). Digital Twin-Driven Process Modeling For Energy Efficiency And Lifecycle Optimization In Industrial Facilities. *American Journal of Interdisciplinary Studies*, *4*(03), 65–95. https://doi.org/10.63125/e4q64869
[133]. Shaikh, S., & Sudipto, R. (2022). Multi-Objective Thermo-Economic and Supply Chain Optimization Modeling For Hydrogen Energy Integration In Smart Factories. *International Journal of Scientific Interdisciplinary Research*, *1*(01), 163–193. https://doi.org/10.63125/p9y8p705
[134]. Shaukat, K., Luo, S., Varadharajan, V., Hameed, I. A., & Xu, M. (2020). A survey on machine learning techniques for cyber security in the last decade. *Ieee access*, *8*, 222310-222354.
[135]. Sobb, T., Turnbull, B., & Moustafa, N. (2020). Supply chain 4.0: A survey of cyber security challenges, solutions and future directions. *Electronics*, *9*(11), 1864.
[136]. Taher, K. A., Jisan, B. M. Y., & Rahman, M. M. (2019). Network intrusion detection using supervised machine learning technique with feature selection. 2019 International conference on robotics, electrical and signal processing techniques (ICREST),
[137]. Taylor, J. M., & Sharif, H. R. (2017). Security challenges and methods for protecting critical infrastructure cyber-physical systems. 2017 International conference on selected topics in mobile and wireless networking (MoWNeT),
[138]. Tsakalidis, G., Vergidis, K., Petridou, S., & Vlachopoulou, M. (2019). A cybercrime incident architecture with adaptive response policy. *computers & security*, *83*, 22-37.
[139]. Tschuchnig, M. E., & Gadermayr, M. (2021). Anomaly detection in medical imaging-a mini review. International Data Science Conference,
[140]. Tsiknas, K., Taketzis, D., Demertzis, K., & Skianis, C. (2021). Cyber threats to industrial IoT: a survey on attacks and countermeasures. *IoT*, *2*(1), 163-186.
[141]. Ukwandu, E., Farah, M. A. B., Hindy, H., Brosset, D., Kavallieros, D., Atkinson, R., Tachtatzis, C., Bures, M., Andonovic, I., & Bellekens, X. (2020). A review of cyber-ranges and test-beds: Current and future trends. *Sensors*, *20*(24), 7148.
[142]. Välja, M., Heiding, F., Franke, U., & Lagerström, R. (2020). Automating threat modeling using an ontology framework. *Cybersecurity*, *3*(1), 1-20.
[143]. Váncza, J., & Monostori, L. (2017). Cyber-physical manufacturing in the light of Professor Kanji Ueda's legacy. *Procedia Cirp*, *63*, 631-638.

[144]. Varma, R., Lee, H., Kovačević, J., & Chi, Y. (2019). Vector-valued graph trend filtering with non-convex penalties. *IEEE transactions on signal and information processing over networks*, *6*, 48-62.
[145]. Vikram, A. (2020). Anomaly detection in network traffic using unsupervised machine learning approach. 2020 5th International Conference on Communication and Electronics Systems (ICCES),
[146]. Wolf, M., & Serpanos, D. (2020). *Safe and secure cyber-physical systems and internet-of-things systems*. Springer.
[147]. Xenofontos, C., Zografopoulos, I., Konstantinou, C., Jolfaei, A., Khan, M. K., & Choo, K.-K. R. (2021). Consumer, commercial, and industrial iot (in) security: Attack taxonomy and case studies. *IEEE Internet of Things Journal*, *9*(1), 199-221.
[148]. Xu, H., Yu, W., Griffith, D., & Golmie, N. (2018). A survey on industrial Internet of Things: A cyber-physical systems perspective. *Ieee access*, *6*, 78238-78259.
[149]. Yadykin, V., Barykin, S., Badenko, V., Bolshakov, N., De la Poza, E., & Fedotov, A. (2021). Global challenges of digital transformation of markets: Collaboration and digital assets. *Sustainability*, *13*(19), 10619.
[150]. Yakimov, P. I., Asparuhova, K. K., Grigorova, T. G., & Shehova, D. A. (2020). Industry 4.0 and the challenges faced by stem education. 2020 XXIX International Scientific Conference Electronics (ET),
[151]. Yeboah-Ofori, A., & Islam, S. (2019). Cyber security threat modeling for supply chain organizational environments. *Future internet*, *11*(3), 63.
[152]. Yeboah-Ofori, A., Islam, S., Lee, S. W., Shamszaman, Z. U., Muhammad, K., Altaf, M., & Al-Rakhami, M. S. (2021). Cyber threat predictive analytics for improving cyber supply chain security. *Ieee access*, *9*, 94318-94337.
[153]. Zamal Haider, S., & Hozyfa, S. (2023). A Quantitative Study On IT-Enabled ERP Systems And Their Role In Operational Efficiency. *International Journal of Scientific Interdisciplinary Research*, *4*(4), 62–99. https://doi.org/10.63125/nbpyce10
[154]. Zhou, X., Xu, X., Liang, W., Zeng, Z., Shimizu, S., Yang, L. T., & Jin, Q. (2021). Intelligent small object detection for digital twin in smart manufacturing with industrial cyber-physical systems. *IEEE Transactions on Industrial Informatics*, *18*(2), 1377-1386.
[155]. Zürn, J., Burgard, W., & Valada, A. (2020). Self-supervised visual terrain classification from unsupervised acoustic feature learning. *IEEE Transactions on Robotics*, *37*(2), 466-481.