NIPES - Journal of Science and Technology Research
www.nipes.org

Machine Learning-Driven Cybersecurity Solutions for Enhanced Smart Grids and Critical Infrastructure: A Review

Idowu Afe[1]*, Ismaila Idris[1], Joseph Adebayo Ojeniyi[1], Sikiru O Subairu[1] and Moses Dogonyaro Noel[1]

[1] Cyber Security Science Department, Federal University of Technology Minna, Nigeria.
*Corresponding Author: idowuafe9@gmail.com

Article Info

Keywords: *Machine Learning, Cybersecurity, Smart Grids, Critical Infrastructure, Anomaly Detection, Intrusion Detection Systems*

Received 22 April 2025
Revised 25 May 2025
Accepted 23 June 2025
Available online 08 July 2025

https://doi.org/10.37933/nipes/7.3.2025.11

eISSN-2682-5821, pISSN-2734-2352

Abstract

Distributed Denial of Service (DDoS) attacks have emerged as one of the most pervasive and damaging threats to network security, disrupting services and incurring substantial financial costs. Machine learning (ML) has been widely explored as a potential solution to enhance DDoS detection and mitigation. This systematic review evaluates the effectiveness of ML techniques in detecting DDoS attacks, synthesizing findings from studies published between 2004 and 2024. The review analyzes models such as Random Forest, Support Vector Machines (SVM), and K-Nearest Neighbors (K-NN) based on key performance metrics like accuracy, precision, recall, and F1-score. A comprehensive search of multiple databases, including Web of Science, IEEE, Scopus, ScienceDirect, and Google Scholar, resulted in 19 studies that met inclusion criteria. The findings show that ensemble methods, particularly Random Forest, consistently outperformed other models in terms of detection rates, mainly due to their ability to handle large feature sets and reduce overfitting. Support Vector Machines also performed well in specific scenarios. However, their effectiveness was sometimes limited by computational complexity and the dataset size. K-Nearest Neighbors showed mixed results, depending on the nature of the attack patterns. This review emphasizes the potential of ensemble learning approaches for DDoS detection, demonstrating their robustness in dynamic environments. The review identifies key gaps in the existing research, including the need for better feature selection and exploring deep learning techniques to enhance DDoS detection accuracy and adaptability further. This study contributes valuable insights into the strengths and limitations of ML-based DDoS detection models, offering a foundation for future advancements in the field. It underscores the importance of continued research into hybrid and deep learning models to address the evolving and increasingly sophisticated nature of DDoS attacks in real-world applications.

1.0. Introduction

Cybersecurity in smart grids and critical infrastructure has become a focal point due to the increasing reliance on Information and Communication Technology (ICT). This dependency enhances efficiency and operational control and introduces significant cyber-attack vulnerabilities [1]. Smart grids, which integrate digital technology with the traditional power grid, enable real-time monitoring and automated control but expose critical infrastructure to new cyber

threats [2]. Recent high-profile cyber-attacks, such as the 2015 Ukraine power grid attack that left over 230,000 people without electricity and the 2021 Colonial Pipeline ransomware attack, which disrupted fuel supplies across the Eastern United States, underscore the potential for widespread societal and economic disruption [3]–[5].

The sophistication and frequency of these attacks have escalated dramatically. According to recent reports, Distributed Denial of Service (DDoS) attacks increased by 31% in 2022 compared to the previous year, with attacks growing in number, complexity, and scale [6]. The proliferation of Internet of Things (IoT) devices, combined with increasingly sophisticated malware, has expanded the attack surface for malicious actors targeting critical infrastructure [7]. These developments necessitate advanced detection and mitigation strategies capable of adapting to evolving threats.

Machine Learning (ML) has emerged as a transformative tool in cybersecurity, offering capabilities to identify patterns, predict potential threats, and automate response mechanisms [8]. Unlike traditional rule-based systems, ML algorithms can analyze vast network traffic data, detect subtle anomalies, and adapt to novel attack vectors in real time. To illustrate the growing academic and technological interest in this field, a Google Ngram analysis was conducted to track the frequency of key terms such as "Machine Learning," "Cybersecurity," and "Smart Grids" in published literature over time (Figure 1).

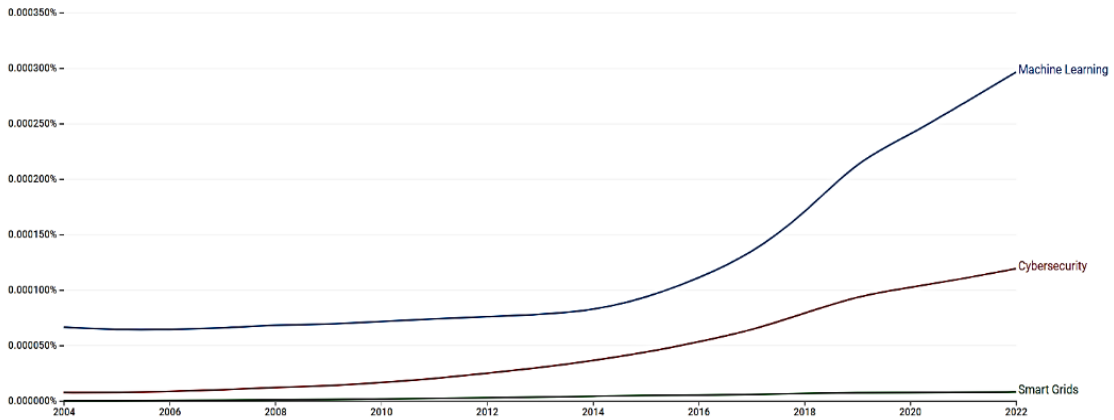

Figure 1: *Google Ngram Analysis Showing the Rising Frequency of "Machine Learning," "Cybersecurity," and "Smart Grids" in Published Literature (2000–2022)*

The Google Ngram graph illustrates distinct trends in the prominence of these terms between 2004 and 2022, reflecting broader shifts in technological focus and research priorities. *Machine Learning* shows a significant rise beginning around 2014, with a sharp increase continuing through 2022, indicating its rapid adoption across various fields, including healthcare, finance, and cybersecurity. Similarly, *cybersecurity* has steadily grown since 2008, accelerating further after 2016, mirroring the growing awareness of digital threats and the escalating need for robust security solutions amidst the rise of IoT devices and cloud computing. In contrast, *Smart Grids* exhibit a modest increase, with relatively flat growth. This suggests that while smart grids are crucial in the energy sector, they have not garnered the same attention in academic literature as machine learning and cybersecurity. These trends highlight the increasing relevance of ML and cybersecurity and a potential research gap in smart grid security, emphasizing the need for integrating advanced ML techniques into smart grid cybersecurity frameworks. ML algorithms can process vast amounts of data, detect anomalies, and adapt to new attack vectors, making them ideal for securing dynamic and complex smart grid systems [9].

Detecting and mitigating cyber threats in smart grids presents unique challenges. The interconnected nature of these systems means that a vulnerability in one component can compromise the entire network [1]. Traditional cybersecurity approaches, such as rule-based systems and signature-based detection, are often insufficient against sophisticated and evolving cyber threats [10]. These methods struggle with real-time data processing and fail to detect novel or complex attack patterns, leading to delayed responses and increased risk of system compromise [11].

Furthermore, the heterogeneity of smart grid components, which include various sensors, communication protocols, and control devices, complicates the development of a unified security framework [12]. The critical nature of these infrastructures amplifies the impact of successful attacks, highlighting the urgent need for more robust, adaptive, and intelligent security solutions [13].

Despite significant progress in applying machine learning to cybersecurity, several gaps and limitations remain. Current ML-based solutions often struggle with high computational complexity, limited scalability, and the inability to handle real-time data processing in dynamic environments [14], [15]. Furthermore, many existing models are not adaptable enough to detect emerging and sophisticated attack patterns, and the application of hybrid models combining supervised, unsupervised, and reinforcement learning approaches remains underexplored [16]–[18].

The novelty of this research lies in its comprehensive review and synthesis of current machine learning techniques, identifying the gaps in their real-world applications for smart grids and critical infrastructure. This review also proposes novel hybrid ML techniques and strategies to address the challenges of scalability, adaptability, and real-time threat detection. This research provides a roadmap for developing more efficient, scalable, and resilient ML-driven cybersecurity solutions tailored to smart grids and critical infrastructure needs by examining case studies and practical applications.

1.1. Objectives of the Review

This systematic review aims to comprehensively analyze the application of ML techniques in enhancing cybersecurity for smart grids and critical infrastructure. The objectives are to:

1. Evaluate the effectiveness of supervised, unsupervised, and hybrid ML techniques in detecting and mitigating cyber threats.
2. Identify gaps in the existing literature, focusing on the limitations of current ML applications and the challenges faced in real-world implementations.
3. Propose future research directions to address identified gaps and improve the resilience of smart grids against cyber threats.

1.2. Scope of the Review

The review focuses on various ML techniques, including supervised, unsupervised, and hybrid approaches, and their applications in cybersecurity for smart grids and critical infrastructure. It includes case studies and practical applications that demonstrate the real-world effectiveness of these techniques. The review covers advancements in anomaly detection, intrusion detection systems (IDS), and the integration of ML with other emerging technologies to enhance cybersecurity.

2.0. Methodology

2.1. Systematic Review Protocol

This review follows the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines to ensure a transparent, reproducible, and comprehensive literature analysis [19]. Figure 2 shows the PRISMA diagram for this study.

2.2. Data Sources and Search Strategy

The literature search for this review covered publications from 2004 to 2024. Searches were conducted using Boolean operators across multiple databases, including Google Scholar, IEEE Xplore, Scopus, ScienceDirect and Web of Science. The search terms included ("Machine Learning" AND "Cybersecurity" AND "Smart Grids" AND "Critical Infrastructure" AND ("Anomaly Detection" OR "Intrusion Detection Systems"))

2.3. Inclusion and Exclusion Criteria

The inclusion criteria for this review focused on peer-reviewed articles, conference papers, and case studies published in the last 20 years, specifically addressing the application of ML in enhancing cybersecurity for smart grids and critical infrastructure. Only studies that explore the use of ML techniques, such as anomaly detection, intrusion detection systems, and other relevant cybersecurity applications within the context of smart grids, were considered. This ensured that the review captured the most recent advancements and examined the evolving challenges and solutions that address modern cybersecurity threats.

On the other hand, articles excluded from this review include non-English papers, as language barriers could hinder accurate interpretation and comparison of methodologies and results. Additionally, studies without full-text access were excluded to ensure a comprehensive evaluation of the content and methodology. Articles that did not directly relate to cybersecurity within the context of smart grids and critical infrastructure were also excluded, ensuring that the focus remained explicitly on the intersection of ML and cybersecurity.

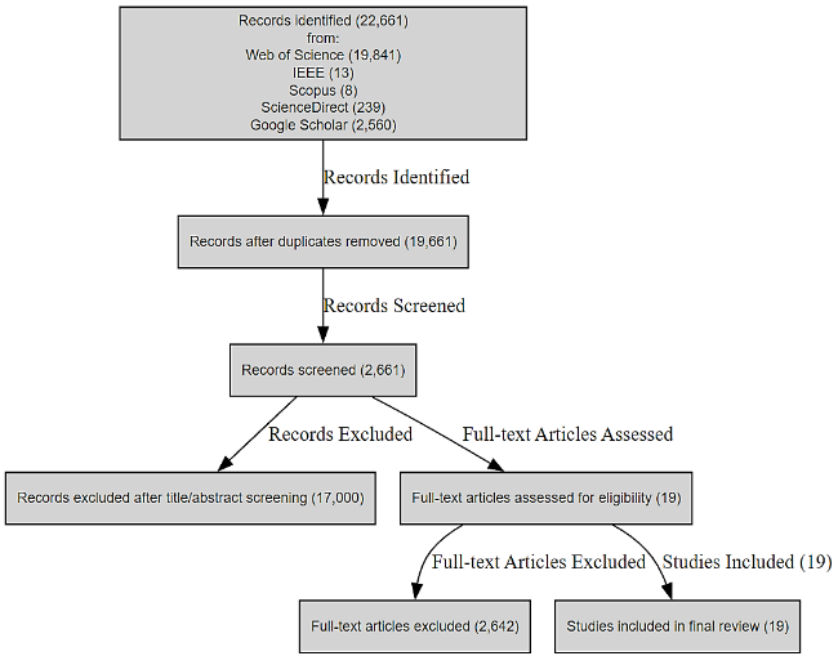

Figure 2: PRISMA Diagram

2.4. Quality Assessment

The quality assessment of the included studies in this review evaluates key factors relevant to machine learning-driven cybersecurity solutions for smart grids and critical infrastructure. The evaluation consists of four main criteria: clear research objectives, relevance to smart grid cybersecurity, appropriateness of the methodology, and the comprehensiveness of the results evaluation. Studies that clearly defined their objectives/hypotheses demonstrated a strong connection to the field of smart grid cybersecurity, employed appropriate machine learning techniques, and provided thorough evaluations of their results, which received higher quality ratings. This assessment ensures that the review captures studies that meet the essential criteria for advancing the understanding of cybersecurity challenges and solutions within the context of smart grids.

2.5. Data Extraction and Synthesis

The extracted key information included the ML techniques, datasets, performance metrics, and results. Qualitative and quantitative synthesis methods were employed to analyze the findings and draw comprehensive conclusions.

3. Results

3.1 Study Characteristics

In the systematic review process, a total of 22,661 records were initially identified from multiple databases, including Web of Science (19,841), IEEE (13), Scopus (8), ScienceDirect (239), and Google Scholar (2,560). After the removal of 3,000 duplicate records, 19,661 unique articles remained. Following title and abstract screening, 17,000 articles were excluded for not meeting the inclusion criteria, leaving 2,661 records for further evaluation. These remaining articles underwent full-text screening, during which 2,642 were excluded due to not meeting the eligibility criteria. Ultimately, 19 studies passed all inclusion requirements and were included in the final systematic review.

3.2 Overview of the Research Landscape

Cybersecurity in smart grids and cyber-physical systems (CPS) has received increasing attention in recent years due to the growing integration of communication networks, Internet of Things (IoT) devices, and control systems. This integration enhances operational efficiency and system intelligence. It introduces new vulnerabilities to sophisticated cyber threats such as DDoS attacks, False Data Injection Attacks (FDIA), and data integrity breaches. To address these challenges, ML has emerged as a key approach for detecting, classifying, and mitigating cyber-attacks. The rising frequency and complexity of such threats have exposed the shortcomings of traditional security measures, reinforcing the need for dynamic and adaptive cybersecurity solutions. ML techniques capable of processing vast amounts of data, identifying anomalies, and responding to evolving threats offer considerable potential. However, implementing ML models such as Random Forests, Support Vector Machines, and Deep Learning in real-world settings presents several challenges. These include high computational requirements, limited scalability, difficulties with real-time processing, and issues like data imbalance, overfitting, and reduced adaptability to novel attack strategies. As a result, many high-performing models in controlled environments fail to maintain the same effectiveness in practical applications. These limitations highlight the need for more robust, scalable, and efficient ML-based cybersecurity frameworks. This systematic review critically examines current ML-driven solutions for securing smart grids and critical infrastructure, identifies key gaps in the literature, and proposes future research directions to enhance the resilience and security of these essential systems. Table 1 offers a comparative overview of the selected studies, summarizing methodologies, ML techniques, datasets, performance metrics, results, and identified limitations to support the analysis and conclusions presented in this review.

Table 1: Included Studies

Author	**Methodology**	**ML Techniques Used**	**Datasets**	**Performance Metrics**	**Results**	**Gaps/Limitations**
Leligou [20]	Hybrid ML + Rule-based IDS	Supervised & Unsupervised ML (8 techniques)	Real-time F2F supply chain traffic (via CIC-FLOWMETER)	Accuracy, TPR	Accuracy: 99.97%, TPR: 99.96%, Real-time deployment accuracy: 98.71%	Unrealistic datasets in prior work; lack of flow-level summary
Balta [21]	Digital Twin for anomaly detection	Data-driven ML, Physics-based models, SME knowledge	Experimental data from 3D printers	Detection accuracy, false alarms	Effective in transient response; detects abnormality vs. expected anomaly	Existing models not extensible to practical CPMS
Yuancheng [22]	Reconstructive ML for DDoS	Deep and shallow reconstructive models	Two DDoS-specific benchmark datasets	Accuracy	High accuracy; no retraining for new attacks	Retraining needed in existing ML models; practical disruption
Alharthi [23]	Literature survey	ML & DL techniques	Various cited studies	Not specified	Broad overview of ML/DL in IoT-DDoS	No empirical validation or model proposal
Swarup [24]	MVCC ensemble detection of FDIA	Majority voting-based ensemble learning	IEEE 24 & 39 bus systems	Detection accuracy	Outperformed state-of-the-art methods	The comparison scope could be expanded
Cai [25]	Adaptive detection (ADAM) in SD-CPS	Info entropy + Unsupervised ML	Real-world DDoS traffic	Mitigation accuracy	Accuracy: 99.13%, Reduced FPR by 35–59%	Does not rely on predefined attack features

Vinod [26]	Multilabel classification of FDIA	Binary Relevance, Classifier Chain	IEEE 14-bus system	Accuracy	Binary Relevance: 95.1% accuracy	Focused on FDIA, not generalizable to other attacks
Dhaou [27]	ML + Quantum Computing	Quantum SVM (QSVM)	Real DDoS dataset	Detection effectiveness	Effective DDoS detection on smart microgrid	Still early in QC-ML integration for cybersecurity
Kumari [28]	AI-Driven Detection + Prevention	XGBoost, SHA-512 cryptography	Custom testbed	Accuracy, cycles/byte	99.12% accuracy, better than existing	Integration challenges in real systems
Raza [29]	Review + comparative analysis	SVM, DBN, RNN, CNN, anomaly detection	Public datasets (such as NSL-KDD)	Various performance metrics	Comprehensive coverage of ML in CPS security	Not an empirical implementation
Gyawali [30]	ML-based cyberattack detection	Decision, tree, bagging and random forest	IEEE-34 bus with attack simulation	Recall, precision and accuracy	Identified false-data injection attacks	Needs clarification on ML models
Aribisala [31]	Hybrid feed-forward ANN (SEQ-FFNN)	ANN, PCA, Hyperparameter tuning	NSL-KDD	Accuracy	99.59% with Sigmoid activation	Model generalization to other datasets untested
Saghezchi [32]	ML-based anomaly detection in CPPS	11 ML algorithms, incl. Decision Tree	Real factory data (semiconductor)	Accuracy, FPR	Decision Tree: 99.9% accuracy, 0.001 FPR	Previous methods based on synthetic data
Meriaux [26]	Performance comparison	DT, RF, QDA, SVM, NB, XGBoost	KDDCup'99, CICIDS'17	Accuracy, time, storage	Compared ML performance on smart grids	Limited to the classification of DDoS only
Singh [33]	Cyber-Physical Anomaly Detection (CPADS)	VMD + Decision Tree	IEEE 39 bus system	Accuracy, Recall, F-Measure	High performance in noisy/noise-free settings	Processing overhead
Farrukh [34]	Hierarchical 2-layer ML	Supervised ML	Custom testbed	Accuracy	95.44% accuracy	Comparison with only a few recent models
Wang [35]	DoS detection via ML	SVM, DT, Naive Bayes	KDD99	Accuracy	SVM best performer	Older dataset, limited modern attack types
Chengming [36]	SDAE + Ensemble ML	SDAE, XGBoost	Co-simulated HIL testbed	Classification accuracy	>90%, 8% improvement vs. state-of-the-art	Complexity in heterogeneous data fusion
Junejo [37]	Behavior-based detection	Supervised ML	Water treatment CPS testbed	FP rate, precision, recall	Fast, robust, low FP	Complex system modeling is needed

3.3 Publication Trends

The evolution of ML applications in CPS security reveals several noteworthy trends, as reflected in the selected studies (Table 2). Over the past few years, a discernible shift has been toward more advanced, context-aware, and practically deployable ML-based solutions, addressing detection effectiveness and operational constraints.

1. Emergence of Hybrid and Ensemble Approaches

Recent contributions, such as those by Leligou [20] and Swarup [24], highlight the growing preference for hybrid and ensemble methodologies that integrate supervised, unsupervised, and rule-based learning paradigms. These models leverage the strengths of multiple techniques to enhance detection accuracy and robustness, particularly in complex environments. For instance, Leligou's hybrid intrusion detection system (IDS) achieved a reported accuracy of 99.97%, illustrating the efficacy of combining diverse ML strategies.

2. Transition Toward Realistic and Real-time Datasets

There is a notable trend toward utilizing real-world or operational datasets, in contrast to earlier reliance on synthetic or legacy benchmarks such as KDD99 or NSL-KDD. Studies by Saghezchi [32] and Leligou [20] employed real factory and real-time flow-level traffic data, respectively, marking a critical move toward improving external validity and deployment readiness. This shift addresses longstanding concerns about the overfitting and limited generalizability of models trained solely on simulated data.

3. Context-specific and Domain-informed Modeling

Several recent studies emphasize the importance of domain-specific adaptations and context-aware modeling. For example, Balta [21] integrates data-driven ML with physics-based models and expert knowledge in a digital twin framework for industrial anomaly detection. Similarly, Yuancheng [22] adopts reconstructive learning techniques for DDoS detection, emphasizing models that require minimal retraining for new attack variants. These approaches reflect a broader movement toward solutions tailored to the operational semantics of specific CPS domains.

4. Focus on Complex Threat Vectors: FDIA and DDoS

There is increasing attention on advanced threats such as False Data Injection Attacks (FDIA) and Distributed Denial of Service (DDoS), which are particularly disruptive in critical infrastructure. Studies such as those by Vinod [26] and Cai [25] explore nuanced detection techniques, including multilabel classification and entropy-based unsupervised learning, respectively. These efforts aim to improve detection specificity and reduce false positive rates, crucial for maintaining trust in CPS environments.

5. Integration of Emerging and Interdisciplinary Technologies

Several recent works demonstrate the incorporation of emerging technologies, including quantum computing [23] and cryptographic techniques [28], into ML-based CPS security frameworks. Though still in the exploratory phase, these integrations suggest a growing interest in leveraging cross-disciplinary innovations to overcome limitations such as scalability and resistance to adversarial attacks.

6. Benchmarking and Comparative Performance Evaluation

Authors such as Meriaux [26] and Raza [29] emphasize the importance of benchmarking ML algorithms across various datasets and performance metrics. Such comparative analyses facilitate the identification of optimal algorithms for specific tasks. However, many of these studies remain confined to classification scenarios and often lack real-time constraints or operational deployment considerations.

7. Persistent Challenges in Generalization and Practical Deployment

Despite methodological advances, several studies acknowledge ongoing challenges, particularly for model generalization and real-world implementation. For instance, [31]and Vinod [26] report strong performance on benchmark datasets but also note limitations in cross-domain applicability. Additionally, integration issues, computational overhead, and limited scalability impede the practical deployment of ML models in live CPS environments.

Table 2: Trends and Strengths Across Studies

Trend	Highlights
Hybrid Approaches	Many studies, such as Leligou [20], Hu [36] and Kumari [28], combine ML with rule-based systems, cryptographic schemes, or digital twins to improve detection robustness.

Use of Realistic & Domain-Specific Data	There is an evident shift toward using real-world datasets such as Saghezchi [32], Balta [21], and Yuancheng [22] over synthetic or generic benchmarks like KDDCup'99.
Advanced Models	Deep learning (Such as SDAE, ANN, CNN), ensemble methods (XGBoost, MVCC), and even quantum ML (QSVM) are explored for more accurate detection, Dhaou [27], Swarup [24], and Hu [36].
Performance Emphasis	Most studies focus heavily on accuracy, true positive rate (TPR), and false positive rate (FPR), with several models achieving over 99% accuracy such as, Saghezchi [32], Kumari [28], and Aribisala [31].

3.4 Categorization by year of publication

The distribution of publications over time reveals a clear upward trajectory in research activity related to ML applications for CPS security. Out of 19 studies reviewed, 2022 emerged as the most prolific, with six published works reflecting a peak in scholarly attention. This surge is part of a broader trend of steady growth observed from 2021 through 2024, suggesting increasing academic and practical interest in leveraging ML to address evolving CPS security challenges (Figure 3). The sustained rise in publications during this period underscores the field's maturation and the urgency to develop robust, real-world solutions to emerging threats across critical infrastructure systems.

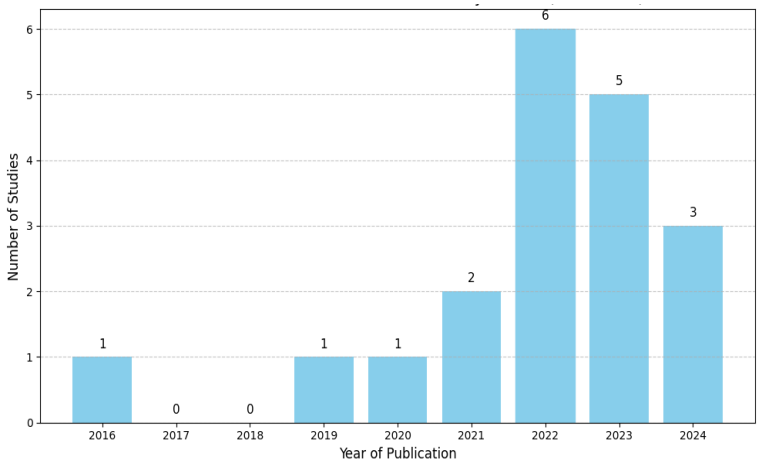

Figure 3: Year of Publication

3.5. Categorization by Machine Learning Techniques

ML techniques used in cybersecurity for smart grids and critical infrastructure vary significantly, each addressing specific security challenges. In supervised learning, a Random Forest Classifier (RFC) is employed three times, reflecting its common use for classification tasks. Decision Trees and Multi-layer Perceptron (MLP) are each applied twice, demonstrating their effectiveness in classification and prediction. K-Nearest Neighbors (KNN) is also used twice, indicating its value in pattern recognition. Support Vector Machines (LSVM) and statistical ML models are applied less frequently, but these methods are more specialized. In the unsupervised learning category, techniques like Variational Mode Decomposition (VMD) and Stacked Denoising Autoencoders are utilized for anomaly detection. Information Divergence is explicitly used to detect DDoS attacks. Deep learning models, such as Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, indicate the growing shift toward more complex architectures capable of capturing intricate patterns in cybersecurity data.

Additionally, hybrid and ensemble techniques, such as Gradient Boosting and models combining multiple classifiers or even integrating Game Theory, demonstrate innovative methods that leverage the strengths of different models to enhance security. These diverse machine-learning techniques collectively contribute to building robust cybersecurity solutions for dynamic and evolving threats that smart grids and critical infrastructure face. Figure 4 shows the categorization by machine learning techniques.

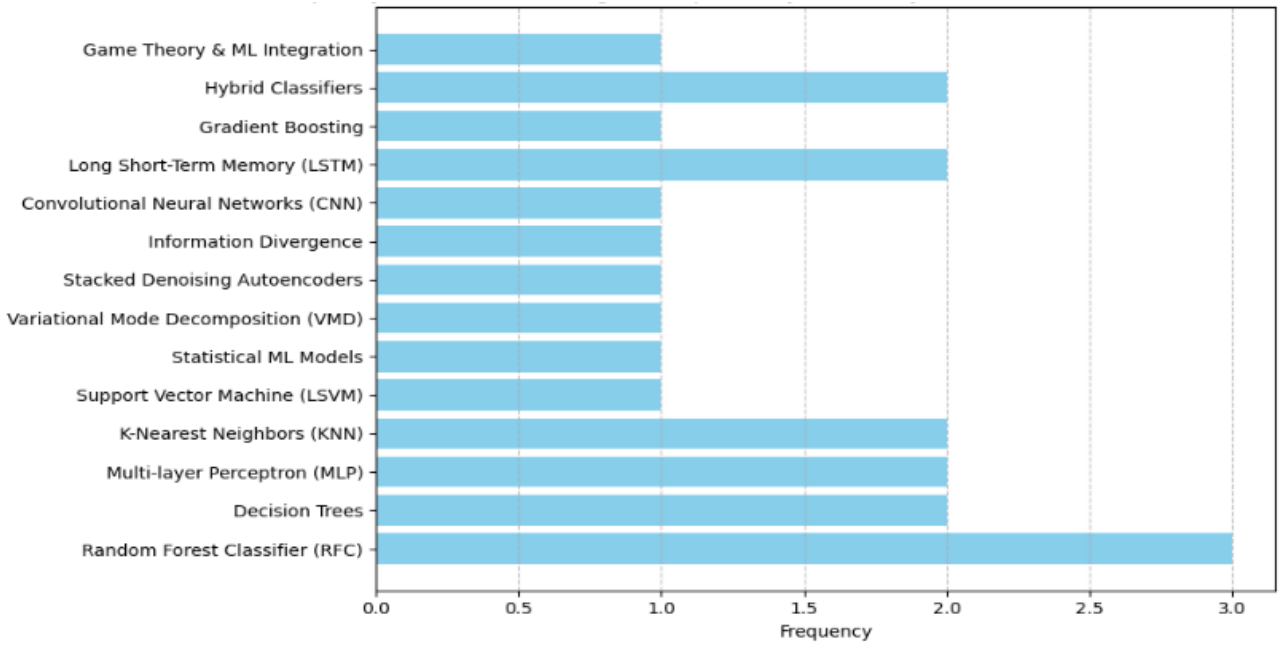

Figure 4: Frequency of Machine Learning Techniques in Cybersecurity for Smart Grids and Critical Infrastructure

3.6. Categorization by Application Domain

The application domains in the reviewed studies span a wide range of cybersecurity areas within critical infrastructure. Power system security is addressed in several studies, focusing on grid disturbances and power system datasets, highlighting the importance of safeguarding electrical grids. DDoS detection emerges as a primary concern, with multiple studies targeting various DDoS attacks and mitigation aspects. This includes approaches to detecting and mitigating attacks in systems like IoT and cloud environments, emphasizing the need for robust defences. IoT security and intrusion detection are also discussed in several studies, underscoring the growing need to secure interconnected devices. SCADA systems are specifically addressed, emphasizing the importance of intrusion detection in industrial control systems. Cyber-physical systems and smart grid environments are each the focus of individual studies, with techniques designed to detect attacks and safeguard simulation data. The review also covers cloud security, particularly SYN flood detection, and the role of SDN in cybersecurity, with studies exploring its use for attack detection in software-defined networks. A limited number of studies discuss grid networks and botnet detection, further showcasing the diversity of application domains covered in the literature. Feature extraction for IDS and DDoS mitigation in grids round out the domains explored, each with a singular study contributing to the broader field. Figure 5 shows the application domains in cybersecurity for smart grids and critical infrastructure.

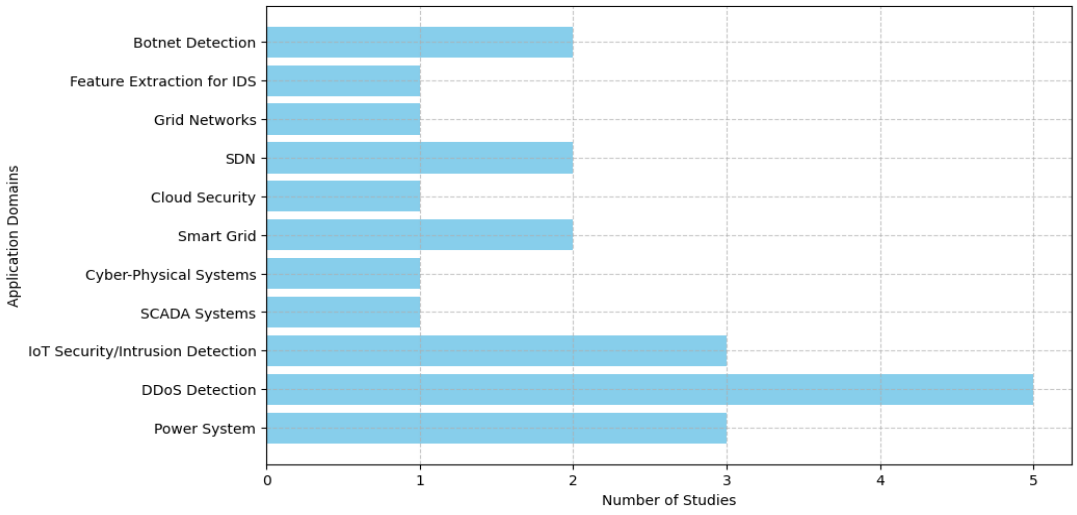

Figure 5: Application Domains in Cybersecurity for Smart Grids and Critical Infrastructure

3.7. Categorization by Datasets Used

The literature reviewed reveals the diverse use of public and custom datasets in machine learning-driven cybersecurity applications. Public datasets like CICIDS2017 and UNSW-NB15 are frequently used, with CICIDS2017 appearing in two studies and UNSW-NB15 in three. These datasets are crucial for evaluating attack detection methods, especially in the context of DDoS attacks and other security threats. In addition, the NSL-KDD dataset is utilized for intrusion detection. At the same time, the IEEE 39-bus system is applied for power grid disturbance detection. Alongside these public datasets, custom and simulated datasets are commonly employed to address domain-specific challenges. For instance, custom datasets related to AC/DC attack vectors and SDN environments are used in studies focused on specific types of attacks. In contrast, simulated datasets like Smart Grid Simulation Data and Simulated Grid Datasets cater to smart grid security evaluations. Botnet DDoS datasets are used to test DDoS mitigation techniques in IoT environments.

Furthermore, SCADA and industrial data are integral to understanding the vulnerabilities in industrial control systems, with datasets from SCADA systems, power system disturbances, and power grid data used in the reviewed studies. The focus on IoT-specific datasets emphasizes the growing importance of securing interconnected devices, as seen in studies addressing IoT security and network traffic data. Additionally, TCP/IP header data is applied in studies focusing on SYN flood detection and other network-level threats. This variety of datasets highlights the wide-ranging scope of research, demonstrating the tailored approaches used to address specific cybersecurity challenges within smart grids and critical infrastructure. Figure 6 shows the frequency of dataset usage in machine learning-driven cybersecurity applications.

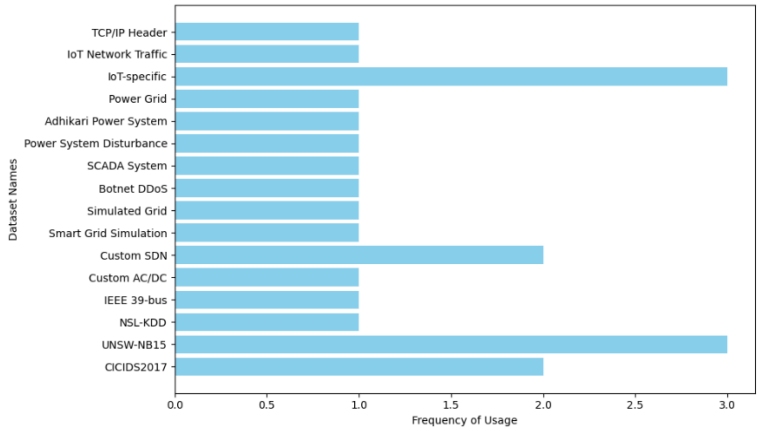

Figure 6: Frequency of Dataset Usage in Machine Learning-Driven Cybersecurity Applications

3.8. Performance Metrics and Analysis

The performance metrics employed in the studies encompass a variety of measures commonly used to evaluate the effectiveness of machine learning models in cybersecurity applications. Accuracy emerges as the most frequently utilized metric, appearing in 12 studies, with several achieving high accuracy scores, particularly those above 95%. Models in this range are often considered excellent, as they demonstrate reliable performance in distinguishing between different classes or predicting outcomes, making them suitable for deployment in critical applications like cybersecurity. In contrast, moderate accuracy (90%-94%) is also common, with studies showing solid performance, though there is still room for improvement. These models are generally effective but may experience occasional errors, such as false positives or negatives, which can be critical in sensitive environments. Lower accuracy models (below 90%) face challenges like data imbalance or inadequate training, making them less suitable for high-stakes deployment. However, they still hold value for research and optimization purposes. Metrics like precision, recall, and F1-score are frequently used to assess model performance, providing insights into how well the models perform regarding true positive and negative rates. Detection and false positive rates are included in some studies, highlighting the importance of minimizing errors in attack detection scenarios. Other metrics, such as ROC-AUC and detection success, further complement the performance evaluations and help gauge the overall robustness of the models. This categorization of accuracy ranges high, moderate, and lower serves as a benchmark, offering valuable insights into the models' performance in different application domains, such as DDoS detection, IoT security, and power system disturbances. Figure 7 shows the performance matrix accuracy categories.

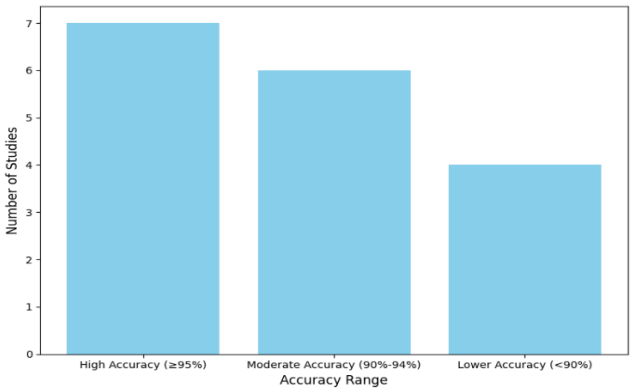

Figure 7: Performance Metrics Accuracy Categories

3.9. Identified Gaps and Limitations

The gaps and limitations identified in the literature on ML applications for cybersecurity in smart grids and critical infrastructure highlight several key challenges that need to be addressed in future research. One of the most common issues is high computational complexity and resource demands, which pose a significant barrier to real-time implementations of ML-based cybersecurity solutions. These deep learning-based, resource-intensive models are complex to deploy in environments with limited computational resources. Exploring hybrid ML models and optimization strategies that balance efficiency with detection effectiveness is essential. Another recurring issue is limited scalability, especially when dealing with large or dynamic datasets in systems like smart grids and IoT networks. To address this, the review investigates advanced techniques such as ensemble learning, distributed learning, and cloud-based approaches to enhance scalability and ensure ML models can handle modern infrastructure's increasing size and complexity. Real-time data processing and adaptability to evolving threats are major concerns, with many models struggling to detect new, unseen attack patterns. This highlights the need for unsupervised learning techniques to detect anomalies autonomously without relying heavily on labeled data and strategies for continuous learning and model retraining. Dataset-specific constraints and lack of generalizability are further limitations, as many models perform well on specific datasets but fail to generalize across different systems. The review evaluates various ML techniques' generalizability. It explores approaches like transfer learning to improve adaptability across diverse real-world applications. In high-traffic or dynamic environments, some ML models experience performance drops, making them unsuitable for deployment in critical infrastructure.

To mitigate this, the research focuses on hybrid models, data fusion, and edge computing techniques that can handle large data streams while maintaining accuracy. High false positive rates and overfitting also remain persistent issues, which can lead to unnecessary alerts or unreliable detections. The review explores strategies for reducing overfitting, such as model regularization and feature selection, while assessing advanced ensemble methods to improve the reliability of detection systems. Lastly, limited cross-domain validation restricts the applicability of some ML models, necessitating a focus on enhancing the transferability of cybersecurity solutions across different critical infrastructure sectors to create more versatile and adaptable ML-based security systems. Figure 8 shows a word cloud for the identified gaps and limitations.

Figure 8: Identified Gaps and Limitations

4.0. Overview of Cybersecurity Challenges in Smart Grids and Critical Infrastructure

4.1. Nature of Cyber Threats

Smart grids, which are increasingly central to the operation of modern infrastructure, face a wide range of cyber threats. These threats include False Data Injection (FDI) attacks, DoS attacks, malware, and phishing, all of which can significantly disrupt the normal functioning of the grid. FDI attacks, for example, can manipulate sensor data to mislead operators into making incorrect decisions. In contrast, DoS attacks aim to overwhelm grid systems with traffic, rendering them unavailable. The presence of malware in grid systems can lead to the corruption or destruction of critical data, and phishing attacks target human vulnerabilities to gain unauthorized access to the system. These types of attacks can lead to financial losses, data theft, and disruptions to grid operations, affecting not just energy distribution but also the economic stability of entire regions. Cyber-attacks on critical infrastructure pose significant risks due to the integration of IT systems and the interconnected nature of these infrastructures. The increasing sophistication of attacks necessitates a comprehensive approach to cybersecurity, including implementing advanced defense technologies and a risk-based strategy to enhance resilience. These incidents demonstrate the sophisticated nature of modern cyber threats and the need for robust, adaptive security mechanisms in smart grids.

4.2. Limitations of Traditional Cybersecurity Methods

Traditional cybersecurity approaches, such as rule-based and signature-based detection systems, have long been employed to safeguard digital systems. However, they exhibit notable limitations when applied to smart grids. These systems rely on predefined rules and known attack signatures to identify threats. However, reliance on these fixed patterns significantly hinders their ability to detect new, emerging, or sophisticated attacks. As cyber attackers continually evolve their techniques, rule-based systems often fail to recognize previously unseen attack vectors. Moreover, smart grids generate massive amounts of real-time data from various sources, making it challenging for traditional methods to process and analyze this information swiftly. The result is often delayed detection, which can exacerbate the damage caused by cyber-attacks. Conventional methods also lack the flexibility to adapt to new threats, especially in dynamic and complex environments like smart grids. As grid systems grow increasingly interconnected and automated, the need for more intelligent, adaptive, and scalable cybersecurity methods becomes even more apparent. Consequently, these methods have been effective in specific contexts. However, their limitations in handling the complex, real-time, and evolving nature of cyber threats in smart grids call for exploring and implementing more advanced, machine learning-driven solutions [38].

Therefore, the nature of cyber threats targeting smart grids, combined with the limitations of traditional cybersecurity approaches, highlights the critical need for more sophisticated, adaptive, and efficient methods to ensure the protection and resilience of these vital infrastructures. Integrating advanced technologies, such as machine learning and anomaly detection, can help address these challenges and offer more comprehensive, dynamic defense mechanisms in the face of evolving cyber risks.

5. Machine Learning Techniques in Cybersecurity

ML techniques have emerged as powerful tools for enhancing cybersecurity in smart grid systems, which are increasingly vulnerable to cyber threats due to their interconnected nature and reliance on digital infrastructure [39]. These methods are critical for anomaly detection, intrusion detection, and identifying complex attack patterns, particularly in scenarios where traditional methods struggle. This section discusses various ML techniques, their architectures, hyperparameter tuning strategies, and the challenges that arise, including computational constraints and overfitting, especially with high-dimensional datasets.

5.1 Supervised Learning Approaches

Supervised learning techniques, such as Decision Trees, Random Forests, Support Vector Machines (SVM), and Neural Networks, are widely employed in cybersecurity for smart grids to detect and mitigate cyber threats like false data injection attacks [40]. These models require labeled datasets where input data is paired with predefined output labels, enabling the model to distinguish between benign and malicious activities.

5.1.1 Model Architectures

i. **Decision Trees**: These trees work by recursively splitting data into branches based on feature values, aiming to create partitions that maximize purity within each partition. Although interpretable and easy to implement, decision trees tend to overfit, especially with high-dimensional data, where small variances can lead to overly complex decision boundaries. This tendency for overfitting may hinder their ability to generalize to new, unseen data [41].

ii. **Random Forests (RF)**: Random Forests are an ensemble method composed of multiple decision trees, each trained on a random subset of the data. Predictions are made by aggregating the individual tree predictions, typically using a majority vote for classification tasks. This approach reduces overfitting, offering a robust solution for handling noisy data and enhancing generalization [42]. Random Forests are particularly effective in handling high-dimensional data, reducing variance, and performing well even in noise and collinearity [43].

iii. **Support Vector Machines (SVM)**: SVMs aim to find the optimal hyperplane that separates different classes in high-dimensional space. Using kernel functions (such as linear, polynomial, and radial basis functions), SVMs effectively classify data that is not linearly separable. However, SVMs can be computationally intensive, especially when large datasets or complex kernel functions are involved (Wang et al., 2024). The complexity of training an SVM model makes it a demanding choice for large-scale systems.

iv. **Neural Networks (NN)**: Neural Networks, particularly Multi-Layer Perceptrons (MLP), are highly effective in classification tasks. MLPs apply activation functions (such as ReLU and softmax) to transform input data into output predictions. They are beneficial for detecting complex patterns in data, but they require substantial labeled datasets for training [45]. Furthermore, improper regularization may lead to overfitting, negatively impacting their ability to generalize to new data [46].

5.1.2 Hyperparameter Tuning

To optimize the performance of supervised models, techniques like grid search and random search are used for hyperparameter tuning. These methods exhaustively search for the best combination of parameters, such as tree depth in Random Forests or learning rates in Neural Networks. However, grid search can be computationally expensive, especially for models with large parameter spaces [47].

i. **Random Forest**: Hyperparameters such as the number of trees (n_estimators), depth of trees (max_depth), and the number of features considered for splits (max_features) play an essential role in determining model performance [48].

ii. **SVM**: Key parameters, including the kernel function (such as linear, RBF) and the regularization parameter (C), are essential for tuning the model to classify complex attack patterns [49].

iii. **Neural Networks**: Optimizing hyperparameters like learning rate, batch size, and dropout rates helps balance accuracy and generalization [50].

Despite their success, supervised learning faces limitations in real-time applications. They heavily rely on large labeled datasets, which may not be readily available in dynamic environments like smart grids. Furthermore, their inability to detect novel or evolving attacks without prior knowledge poses a challenge for adaptive cybersecurity solutions [51].

5.2 Unsupervised Learning Approaches

In contrast to supervised learning, unsupervised learning techniques such as K-Means clustering, Autoencoders, and Isolation Forests are valuable for cybersecurity tasks when labeled data is scarce or unavailable. These methods focus on anomaly detection, identifying unusual behavior or deviations from normal system operations without requiring prior knowledge of attack types [52].

5.2.1 Model Architectures

i. **K-Means Clustering**: K-Means groups data into clusters based on similarities, with data points that do not fit well into any cluster being flagged as anomalies. While effective for simple datasets, K-Means struggles with more complex data, especially when patterns exhibit non-spherical distributions or significant noise [53]. This model may also face challenges in determining the optimal number of clusters.

ii. **Autoencoders**: These neural networks are designed to learn a compressed representation of standard data and reconstruct it to its original form. Anomalous data points are detected based on high reconstruction errors [54].

Autoencoders are effective in high-dimensional datasets and can handle deviations from normal behavior. However, they require careful tuning to avoid overfitting and achieve optimal results.

iii. **Isolation Forest**: Unlike traditional clustering techniques, this model isolates outliers by recursively partitioning the dataset. It isolates anomalous points faster and more efficiently than methods like K-Means, making it an attractive choice for large datasets [55]. The algorithm's efficiency in detecting anomalies is one of its most significant advantages.

Unsupervised learning techniques often struggle with high false positive rates, where normal activities may be misclassified as anomalies. Moreover, selecting the optimal number of clusters in K-Means or tuning thresholds in Autoencoders can be challenging without domain-specific knowledge. Advanced techniques such as density-based clustering or anomaly scoring can help refine these models, reducing false positives and improving detection accuracy [56].

5.3 Hybrid and Ensemble Methods

Hybrid and ensemble methods combine the strengths of both supervised and unsupervised learning techniques, improving detection capabilities and robustness while adapting to various attack scenarios. These methods are essential for tackling complex threats in dynamic environments like smart grids.

5.3.1 Hybrid Models

i. **Autoencoder + SVM**: The Autoencoder detects anomalies. At the same time, the SVM refines the classification to increase precision, particularly in distinguishing normal activities from malicious behaviors [57].
ii. **Random Forest + K-Means**: Random Forests classify known attack types. At the same time, K-Means helps identify novel patterns, enhancing the system's ability to detect new and previously unseen attacks [58].

5.3.2 Ensemble Techniques

i. **Bagging**: Methods such as Random Forest combine multiple classifiers to reduce variance and prevent overfitting. Each classifier is trained on a random subset of data, and the final prediction is determined by aggregating individual predictions [59].
ii. **Boosting**: Techniques like AdaBoost and XGBoost iteratively improve weak classifiers by adjusting weights on misclassified data points. These techniques are particularly useful when dealing with imbalanced datasets, as they ensure the model focuses on hard-to-classify instances [60].

Ensemble and hybrid methods improve model accuracy and robustness but often introduce higher computational costs, complicating real-time deployment in environments with limited resources [61].

5.4 Deep Learning in Smart Grid Security

Deep learning techniques, including Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, provide powerful solutions for cybersecurity in smart grids, particularly when analyzing time-series data and detecting complex attack patterns.

5.4.1 Model Architectures

i. **Convolutional Neural Networks (CNNs)**: CNNs are designed to capture spatial dependencies and hierarchical features in data. CNNs are widely used for image processing. They are also effective in smart grid cybersecurity for time-series data [62]. They apply filters to detect local dependencies and patterns, then pool layers to reduce dimensionality. CNNs are computationally intensive, requiring substantial hardware resources to function effectively. However, various optimization techniques, such as stochastic computing, hardware-efficient dataflows, approximate computing, and FPGA implementations, can significantly reduce power consumption, memory usage, and computational complexity while maintaining accuracy [63].
ii. **Long Short-Term Memory (LSTM) Networks**: LSTMs, a Recurrent Neural Network (RNN) type, are excellent for handling sequential data with long-term dependencies. They are particularly effective in detecting evolving threats, such as Advanced Persistent Threats (APTs) and DDoS attacks [64]. The architecture includes

memory cells that retain information over time, allowing for the detection of temporal dependencies in time-series data.

5.4.2 Hyperparameter Tuning for Deep Learning

i. **CNNs**: Important hyperparameters include kernel size, learning rate, and dropout rate, and kernel size is crucial for improving the performance, accuracy, and generalization ability of CNNs while preventing overfitting [65].
ii. **LSTMs**: Hyperparameters such as sequence length (time steps), learning rate, and hidden layer size significantly affect the network's ability to detect anomalies in time-series data [66].

Deep learning models face several challenges. They are computationally demanding, requiring high-performance GPUs for training and inference. In resource-constrained smart grid environments, this can limit their practicality. Furthermore, deep learning models are prone to overfitting, especially when dealing with small or imbalanced datasets. Techniques like dropout, early stopping, and data augmentation effectively mitigate overfitting, enhance generalization, and improve model performance in neural networks [67].

6. Datasets and Benchmarking

6.1. Commonly Used Datasets

Publicly available datasets such as UNSW-NB15, NSL-KDD, and CICIDS2017 are widely used in cybersecurity for training and evaluating machine learning models. These datasets contain labeled data that can be used to develop intrusion detection systems (IDS) and to benchmark the performance of various machine learning algorithms. UNSW-NB15, for example, provides a comprehensive set of network traffic data, including both normal and attack traffic, which is essential for testing the efficacy of different detection models. Similarly, NSL-KDD is an enhanced version of the older KDD Cup 99 dataset, often used for evaluating intrusion detection systems due to its variety of attack types and real-world relevance. CICIDS2017, developed by the Canadian Institute for Cybersecurity, offers modern network traffic data with more recent attack types, making it a crucial resource for researchers working on contemporary cybersecurity problems.

In addition to these general-purpose datasets, specialized datasets for smart grids are also essential for developing and testing intrusion detection systems tailored to these critical infrastructures. These datasets often focus on Phasor Measurement Unit (PMU) data and power system events, providing unique insights into the behavior of power systems under normal and attack conditions. Such datasets are invaluable for testing cybersecurity models in scenarios that reflect smart grids' operational characteristics and vulnerabilities. They offer critical data points, such as voltage and current measurements, which can be used to identify abnormal patterns that may indicate a cyber attack on the grid. Given the complexity and specificity of smart grid systems, these specialized datasets play a crucial role in enhancing the robustness and relevance of machine learning models in smart grid systems by improving energy management, stability prediction, and system efficiency [68].

6.2. Challenges in Dataset Availability and Standardization

One of the primary challenges in obtaining high-quality datasets for smart grid cybersecurity is data privacy and access issues. Smart grid systems often involve sensitive operational data, which can be difficult to access due to privacy regulations, security concerns, and proprietary restrictions. Utilities and other stakeholders may be unwilling to release detailed data because of the risks of exposing vulnerabilities or violating privacy laws. This makes it challenging for researchers to obtain the necessary real-world data to develop and test intrusion detection models. Furthermore, the lack of standardized datasets complicates the comparability of research findings. Without common benchmarks, assessing the performance of different models or evaluating the generalizability of findings across various contexts becomes difficult.

The lack of standardized datasets also impedes the development of universally applicable models. Given that smart grids can vary significantly in terms of architecture, communication protocols, and operational conditions, a dataset that works well for one grid system may not be suitable for another. This lack of consistency in the available data can lead to highly specialized models for systems that are ineffective in broader contexts. As a result, researchers may struggle to create solutions that can be applied across a wide range of smart grid environments.

Synthetic data generation and simulation techniques are often employed to address these challenges. These methods allow researchers to create artificial datasets that mimic the characteristics of real-world smart grid systems. By simulating different attack scenarios, researchers can create diverse datasets that help to train and test machine learning models without compromising privacy or security. Simulation tools can model various aspects of grid behavior, including normal operations and potential threats, thus providing a controlled environment for evaluating intrusion detection systems. Despite their utility, synthetic datasets have limitations. They may not capture all the nuances of real-world data, and the simulated attack scenarios may lack the complexity or unpredictability of actual cyber threats. Therefore, combining synthetic data with real-world data is essential for creating robust and effective models for smart grid cybersecurity whenever possible.

Therefore, while publicly available datasets like UNSW-NB15, NSL-KDD, and CICIDS2017 provide essential resources for general cybersecurity research, specialized datasets for smart grids are critical for developing tailored intrusion detection systems for this domain. However, data privacy, access, and standardization challenges continue to hinder progress. Synthetic data generation and simulation provide valuable solutions. However, they must be carefully balanced with real-world data to ensure the effectiveness and applicability of machine learning models for smart grid cybersecurity.

7. Evaluation Metrics and Performance Analysis

7.1 Common Metrics

Evaluating the performance of ML models is crucial to ensure that the selected models meet the required standards for detecting and mitigating cyber threats effectively. Various performance metrics are employed to assess how well an ML model performs in real-world cybersecurity scenarios. Common metrics include Accuracy, Precision, Recall, F1-Score, ROC-AUC, and False Positive Rate. These metrics provide valuable insights into the strengths and weaknesses of a model in distinguishing between normal and malicious activities.

i. Accuracy measures the percentage of correct predictions the model makes [69]. However, it may not provide a complete picture, particularly in imbalanced datasets, such as benign samples significantly outnumber malicious ones. High accuracy might be misleading in these cases, as a model could predict the majority class without identifying threats.

ii. Precision refers to the percentage of true positive results from all positive predictions. This metric is essential in minimizing false positives and unwanted alarms that could overwhelm cybersecurity professionals and lead to unnecessary interventions. High precision ensures that it will likely be correct when the model predicts a threat [70].

iii. Recall measures the percentage of true positives out of all actual positives, prioritizing the model's ability to detect as many threats as possible. This might come at the cost of increased false positives, a trade-off critical in cybersecurity contexts where missing a threat could have severe consequences. In smart grid environments, false alarms could disrupt operations, so recall optimization must be carefully balanced with precision to avoid unnecessary downtime [71].

iv. F1-Score is a harmonic mean of Precision and Recall, which balances the trade-off between the two. This metric is beneficial when there is an uneven class distribution, as it avoids the bias toward the majority class that accuracy might introduce [72]. The F1-score is commonly used when both false positives and false negatives are costly.

v. **ROC-AUC (Receiver Operating Characteristic - Area Under Curve)** is another important metric that indicates the model's ability to distinguish between classes under various thresholds. A higher AUC score generally suggests better model performance [73]. ROC-AUC is particularly useful for evaluating models that need to balance different thresholds of detection sensitivity, such as those used in evolving cyberattack detection.

vi. **False Positive Rate (FPR)** measures the proportion of benign events misclassified as threats. A high FPR can be particularly problematic in cybersecurity systems as it could lead to unnecessary interventions and disrupt the grid's or infrastructure's normal functioning. In smart grids, minimizing false positives is critical to ensuring normal operations are not interrupted unnecessarily [74].

These metrics collectively provide a balanced view of a model's performance, allowing for a comprehensive evaluation of its suitability in detecting cyber threats and minimizing unnecessary alerts. The trade-offs between these metrics must be carefully considered depending on the use case, such as whether the priority is minimizing false positives or maximizing detection accuracy. In cybersecurity, particularly for smart grids, where real-time decisions must be made to protect infrastructure without disrupting operations, selecting the right evaluation metrics becomes even more vital.

7.2 Comparative Analysis of Techniques

A comparative analysis of different ML techniques offers insight into how various models balance trade-offs between accuracy, computational cost, and real-time applicability. Some models, particularly those based on deep learning techniques, provide high accuracy in detecting complex attack patterns. However, these models often require substantial computational resources for training and inference, which may not be feasible in real-time cybersecurity applications where rapid decision-making is critical [75].

In contrast, Decision Trees or Random Forests may provide faster processing times with fewer computational demands. However, they can sacrifice some degree of accuracy compared to more complex models. These models are often favored in environments where quick response times and lower resource consumption are prioritized, overachieving the highest accuracy [76].

Similarly, ensemble methods, which combine the outputs of multiple models, can achieve higher detection accuracy by leveraging the strengths of different classifiers. However, they come with increased computational complexity, making them less suitable for systems with strict real-time requirements [77]. These methods, including Boosting and Bagging, may be necessary when detection accuracy is prioritized over speed or computational efficiency.

The selection of the most appropriate model depends on the system's specific needs, such as high accuracy, the available computational resources, and the real-time demands of the environment. For instance, deep learning models may be ideal for large-scale attack detection in environments with powerful computational resources. However, they could be infeasible in a constrained environment like a smart grid. Decision Trees or Random Forests may be better suited for deployment in smart grids, where computational resources are limited and quick responses are essential [78]. This underscores the need for an informed and context-sensitive approach to model adoption in cybersecurity applications.

For effective model selection in real-world deployments, this review introduces a critical comparison table (Table 3) that highlights key trade-offs across essential performance dimensions such as accuracy, interpretability, scalability, and real-time suitability. The framework serves as a practical reference for identifying appropriate machine learning techniques aligned with the operational demands of smart grid environments. Beyond detection performance, the analysis emphasizes the importance of balancing model efficiency with deployment feasibility, particularly in resource-constrained settings were minimizing latency and maintaining system reliability are paramount.

Table 3: Critical Comparison of ML Model Trade-offs

Model	Accuracy	Interpretability	Real-Time Suitability	Scalability	Data Requirement
SVM	High	Low	Medium	Low	High
Random Forest	High	Medium	High	High	Medium
CNN	Very High	Low	Low	Medium	Very High
Decision Tree	Medium	High	High	Medium	Low
Autoencoder	High	Low	Medium	Medium	Medium
LSTM	High	Low	Medium	Low	High

7.3 Comparison Between Traditional and Modern Machine Learning Methods

Cybersecurity in smart grids and critical infrastructure benefits from various ML approaches. Understanding these techniques' strengths and limitations requires comparing traditional machine learning methods and more modern, advanced techniques. Traditional methods, such as rule-based systems, decision trees, and linear models, have laid the foundation for many applications. However, modern ML methods, particularly deep learning and hybrid models, provide advanced capabilities better suited to addressing complex cybersecurity threats in dynamic environments like smart grids. This section compares traditional and modern ML methods, emphasizing the differences in performance, adaptability, and suitability for securing smart grid infrastructures. Table 4 summarizes the key differences in performance and approach between traditional and modern machine learning methods.

7.3.1 Traditional Machine Learning Methods

Traditional ML methods generally rely on simpler algorithms and models that are easier to interpret and require fewer computational resources. These include rule-based systems, decision trees, support vector machines (SVM), and linear regression. Each method has its unique advantages and limitations when applied to cybersecurity challenges.

1. **Rule-based Systems**: These systems operate on predefined rules to detect specific malicious behaviour patterns. Though easy to implement, they often struggle to detect novel attack patterns and adapt to new threats.
2. **Decision Trees and Random Forests**: Decision trees are widely appreciated for their simplicity and interpretability. Random forests, an ensemble of decision trees, offer greater robustness and the ability to handle more complex data. Despite this, they still encounter challenges with scalability, especially when real-time threat detection is required in large-scale networks.
3. **Support Vector Machines (SVM)**: SVMs are effective for binary classification problems and excel in high-dimensional spaces. However, they often demand substantial training data and are computationally intensive, which poses limitations in time-sensitive applications.
4. **Linear Models (such as Logistic Regression)**: Linear models are efficient and easy to implement but often fail to capture the complexities of non-linear relationships inherent in cybersecurity data.

Though traditional methods offer valuable tools for specific contexts, they typically face difficulties processing large volumes of dynamic, high-dimensional data, such as the type generated by smart grids.

7.3.2 Modern Machine Learning Methods

Modern ML methods, particularly deep learning and hybrid models have become essential in cybersecurity because they manage vast data and adapt to evolving threats. These methods include neural networks, CNNs, RNNs, and ensemble learning techniques.

1. **Neural Networks and Deep Learning**: These models identify complex patterns within large datasets. They can autonomously learn features from raw data, making them well-suited for detecting novel and sophisticated attack patterns. However, they require significant computational resources and large volumes of labeled data for effective training.
2. **Convolutional Neural Networks (CNNs)**: Originally developed for image recognition, CNNs are adept at identifying spatial patterns in network traffic data. They are instrumental in tasks that involve pattern recognition in data streams, such as detecting abnormal behaviors within smart grids.
3. **Recurrent Neural Networks (RNNs)**: RNNs are highly effective for sequential data analysis, making them ideal for real-time network traffic analysis. These networks retain information from previous data points, making detecting patterns essential for identifying advanced persistent threats (APTs) over time.
4. **Hybrid Models**: Hybrid models combine multiple ML techniques, such as supervised learning integrated with unsupervised and reinforcement learning. These approaches offer enhanced scalability and flexibility, allowing them to handle known and unknown threats in dynamic environments like smart grids.
5. **Ensemble Methods**: Ensemble methods, including boosting and bagging, leverage multiple models to improve performance. These methods are often more robust than individual models, enhancing detection rates and reducing the likelihood of false positives.

Table 4: Key Differences in Performance and Approach

Feature	Traditional Methods	Modern Methods
Model Complexity	Simpler models, easier to interpret	Complex models may be difficult to interpret
Adaptability to New Threats	Limited ability to adapt to novel attack patterns	High adaptability to novel threats and unknown attack vectors
Scalability	Struggles with scalability in large-scale or dynamic environments	Highly scalable, able to handle large datasets and real-time data
Computational Resource Usage	Lower resource requirements	Higher computational demands (such as GPU usage for deep learning)
Performance with Complex Data	May struggle with non-linear relationships in data	Excellent performance in detecting complex, non-linear patterns
Real-Time Processing	Can be slow for real-time applications	High performance for real-time threat detection

Use of Labeled Data	Often requires well-labeled data for training	Can work with large datasets. Some models require less labeled data.
Interpretability	High interpretability (such as decision trees)	Lower interpretability (such as deep learning)
Common Algorithms	Decision Trees, SVM, Logistic Regression	Neural Networks, CNNs, RNNs, Hybrid models

Modern ML methods, particularly deep learning and hybrid approaches, offer significant advantages over traditional performance, adaptability, and scalability methods. While traditional methods remain useful for certain tasks, mainly where interpretability is essential, modern techniques are better suited to the challenges posed by smart grids' dynamic, complex nature and critical infrastructure cybersecurity. The ability of modern models to process large volumes of data, detect novel attack patterns, and adapt to evolving threats makes them essential for securing the next generation of smart grid systems.

7.3.3 Classification Framework

This study introduces a structured classification framework that organizes ML methods based on their alignment with specific cyberattack types, operational advantages and drawbacks, feasibility for deployment, and dataset compatibility. The framework enhances the synthesis of existing literature and offers a decision-support tool for researchers and practitioners. Each method is contextualized within its practical utility, enabling a more strategic approach to model selection as shown in Table 5.

Table 5: Classification Framework

ML Method	**Suitable Attack Types**	**Pros**	**Cons**	**Deployment Feasibility**	**Dataset Type**
Random Forest	DDoS, FDIA	Robust, scalable	Resource intensive	Medium	Public
Autoencoder + SVM	Novel/zero-day attacks	Adaptive to new threats	Requires careful tuning	Low	Custom
LSTM	APTs, sequential threats	Captures temporal patterns	High computational requirements	Low–Medium	Simulated, Custom
CNN	Complex traffic patterns	High accuracy	Low interpretability	Low	Public/Custom
K-Means	Unknown anomaly types	No labeled data required	High false positive rate	High	Public/Unlabeled

The framework, presented in Table 5, enables meaningful comparisons of model characteristics across critical dimensions such as accuracy, adaptability, resource efficiency, and applicability in real-world environments. Random Forests, for example, provide robust and scalable detection of DDoS and FDIA attacks but incur moderate resource demands, making them suitable for medium-scale deployment. In contrast, Autoencoder-SVM hybrids offer adaptability for detecting novel threats but present challenges in tuning and deployment readiness. Deep learning models such as CNNs and LSTMs capture complex temporal patterns with high accuracy but remain limited in interpretability and resource efficiency. K-Means clustering, though suitable for anomaly detection without labeled data, often suffers from high false positive rates. Therefore, this classification not only clarifies methodological trends but also supports the development of more tailored and resilient cybersecurity solutions for smart grids and critical infrastructure.

8. Challenges and Limitations

Despite the growing adoption of ML techniques in cybersecurity particularly within critical infrastructure systems such as smart grids and CPS several systemic challenges continue to hinder their practical implementation, as outlined in Table 6. These challenges span technical, operational, and deployment dimensions, including issues such as data quality and availability, limited model generalization, lack of interpretability, computational constraints, and susceptibility to adversarial attacks. Furthermore, a prevalent over-reliance on accuracy as the primary performance metric often obscures equally critical considerations such as latency, scalability, real-time feasibility, and explainability. Addressing these multifaceted limitations is essential for developing robust, trustworthy, and operationally viable ML-driven cybersecurity frameworks capable of performing effectively in dynamic and high-stakes environments.

8.1 Data Quality and Availability

The effectiveness of ML models in cybersecurity is heavily influenced by the quality, diversity, and representativeness of the training datasets. A major limitation in current research is the widespread reliance on outdated, synthetic, or narrowly scoped datasets, which significantly undermines the generalizability and robustness of ML models across diverse and dynamic operational environments, such as smart grid systems. The scarcity of comprehensive, labeled datasets that accurately capture real-world cyber-attack scenarios particularly those involving emerging or sophisticated threats further impedes model development and evaluation. This challenge is compounded by barriers related to data privacy, proprietary restrictions, and inconsistent data formats, all of which hinder access to realistic and varied training data. Without high-quality, well-annotated datasets that reflect a broad spectrum of attack types and conditions, ML models struggle to adapt effectively to new threats, limiting their practical applicability in real-world cybersecurity contexts [80].

8.2 Model Interpretability and Trust

Interpretability remains a critical barrier to the deployment of ML models in cybersecurity, particularly those based on complex deep learning architectures. These models often function as "black boxes," offering limited transparency into their decision-making processes, which poses significant challenges in regulated or safety-critical domains such as energy infrastructure. The lack of explainability can undermine stakeholder trust and impede adoption by system operators who require clear insights into model outputs for accountability, compliance, and informed decision-making. Even when such models demonstrate high accuracy, their opacity can deter practical use in high-stakes environments. Therefore, enhancing interpretability through Explainable AI (XAI) techniques is essential not only for building trust but also for ensuring responsible and effective deployment in contexts where the consequences of automated decisions are significant [81].

8.3 Real-Time Implementation Constraints

Real-time applicability is a fundamental requirement for deploying ML models in smart grid environments, yet many existing approaches fall short due to high computational demands, latency issues, and integration challenges. Smart grids generate massive volumes of streaming data that must be processed rapidly to enable timely threat detection and response. However, ML models that perform well in controlled or experimental settings often struggle in live environments, where computational resources are limited and decision-making must occur with minimal delay. Integration with legacy systems, such as SCADA, further complicates deployment due to compatibility constraints. These limitations underscore the urgent need for lightweight, scalable, and deployment-ready ML solutions specifically optimized for real-time performance in critical infrastructure settings [82].

8.4 Vulnerability to Adversarial Attacks

The increasing sophistication of adversarial attacks, such as data poisoning and input manipulation, poses a significant threat to the reliability of ML-based detection systems in cybersecurity. These attacks exploit model vulnerabilities by subtly altering inputs to trigger incorrect predictions, thereby undermining the effectiveness of security mechanisms particularly in critical infrastructure environments. Ensuring robustness against such threats is an emerging priority that requires the development of resilient ML architectures and the integration of defense mechanisms capable of detecting and mitigating adversarial behavior during both training and inference phases. Advancing techniques for adversarial resilience is essential to maintaining the integrity and trustworthiness of ML-driven cybersecurity solutions [83].

Table 6: Observed Research Gaps and Limitations

Gap Area	Explanation & Evidence
Over-reliance on Accuracy Metrics	Accuracy is often prioritized, while other real-world metrics like latency, scalability, and adaptability to evolving threats remain underexplored.
Limited Generalization	Many models are trained and tested on narrow scenarios and datasets, raising concerns about real-world deployment across varied CPS environments [26], [31].
Static ML Models	Several ML approaches require full retraining when new attack types arise, which limits real-time adaptability [26], [84]. Yuancheng [22] addressed this via reconstructive models.
Underexplored Real-Time Constraints	Despite real-time systems being critical (especially in grids), many studies don't account for processing overheads or real-time deployment feasibility [33].

Lack of Explainability	Most models are black-box, offering little transparency or interpretability. This limits operator trust and hinders adoption in regulated environments.
Scalability & Integration Challenges	Solutions like ADAM show promise, but integration into legacy SCADA or SDN-based systems at scale remains insufficiently explored [25].

A visual summary of key limitations in current machine learning applications for cybersecurity in smart grids. The identified challenges span data availability, interpretability, real-time feasibility, model generalization, and resilience against adversarial attacks (Figure 9). These limitations highlight critical gaps requiring attention to ensure reliable and scalable deployment in critical infrastructure systems.

9. Future Directions and Research Opportunities

As ML continues to shape cybersecurity strategies for critical infrastructure systems, particularly smart grids, future research must shift from conceptual development to practical, scalable deployment. Addressing current limitations while aligning with operational needs will be crucial for building adaptive and trustworthy security solutions. Key research priorities include privacy-preserving model training, interpretability, real-time responsiveness, and integration with emerging technologies. Figure 10 presents a strategic framework that illustrates how these innovations can be operationalized within real-world smart grid environments.

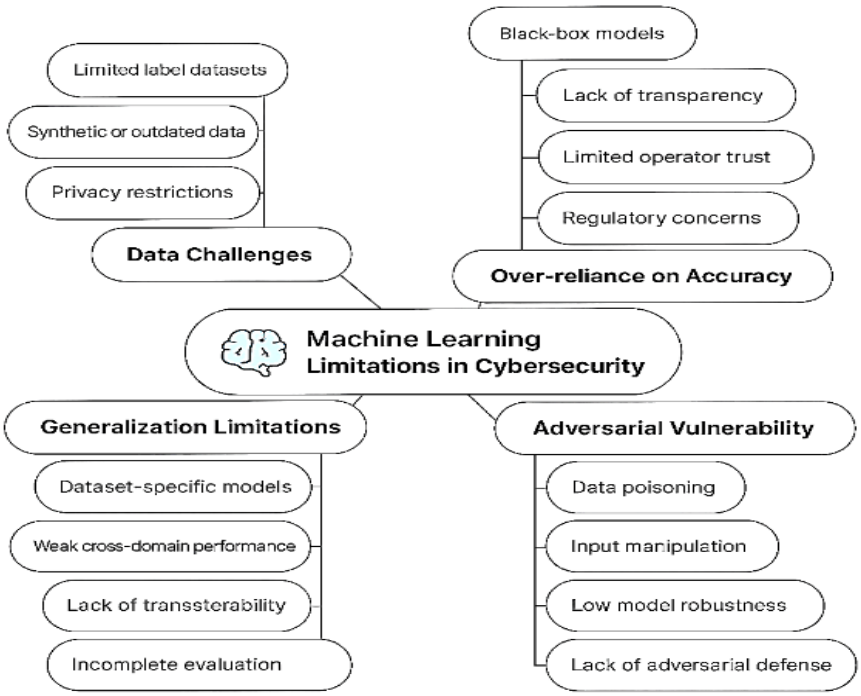

Figure 9: Identified Research Gaps and Limitations in ML-Based Cybersecurity for Smart Grids and Critical Infrastructure

9.1 Federated Learning and Privacy-Preserving Techniques

Federated Learning (FL) enables decentralized model training across distributed devices or entities without sharing raw data, thereby preserving privacy a crucial requirement in sensitive infrastructures. For example, FL could support collaborative intrusion detection across geographically dispersed smart meters, enabling threat detection while maintaining data locality and regulatory compliance. Future research should enhance FL's scalability, communication efficiency, and resilience against poisoning attacks, particularly in real-time, resource-constrained environments.

9.2 Explainable AI (XAI) in Cybersecurity

Explain ability is critical for operational trust and regulatory acceptance of ML-driven systems. In smart grid environments, XAI can be integrated into CNN-based intrusion detection systems (IDS) deployed in grid control rooms to offer interpretable alerts to operators. This transparency can facilitate informed decision-making, compliance, and rapid incident response. Continued development of domain-specific XAI methods will be essential to bridge the gap between model complexity and human understanding in critical infrastructure.

9.3 Integration with Emerging Technologies

Combining ML with technologies such as blockchain, the IoT, and edge computing can significantly enhance cybersecurity. For instance, integrating ML with blockchain can secure smart contract-based load balancing in microgrids, ensuring data integrity and preventing tampering. IoT-enabled devices in substations or homes can benefit from lightweight ML models deployed at the edge for real-time anomaly detection. Future work should explore how these technologies can jointly support autonomous, secure, and low-latency grid operations.

9.4 Development of Real-World Datasets

Robust ML models require access to high-quality, real-world datasets that capture the complexity of smart grid operations and attack behaviors. Joint efforts between academia, industry, and government are needed to develop open, standardized datasets reflecting diverse scenarios such as coordinated attacks, equipment faults, and varying load profiles. These datasets will enable more rigorous benchmarking and foster the creation of models that generalize across different system architectures and operational conditions.

9.5 Domain-Adaptive Learning

To ensure adaptability to evolving cyber threats, domain-adaptive learning methods such as transfer learning, continual learning, and online learning should be explored. These approaches can enable ML models to retain performance in new environments without full retraining, making them suitable for dynamic smart grid contexts where threats and operational patterns evolve continuously.

9.6 Benchmark Standardization

Establishing standardized evaluation protocols using real-world testbeds such as Hardware-in-the-Loop (HIL) systems and IEEE benchmark grids is critical for ensuring comparability and reproducibility across studies. Such standardization will support objective assessment of model performance under realistic constraints, including attack variability, latency, and computational cost.

9.7 End-to-End System Evaluation

Future research must extend beyond accuracy metrics to include time complexity, energy consumption, deployment cost, and fault tolerance. Evaluating ML models holistically within operational workflows will help determine their true feasibility and effectiveness in smart grid environments.

9.8 Collaborative Defense Models

Given the interconnected nature of modern power systems, collaborative defense mechanisms such as FL-based or decentralized ML frameworks can enhance situational awareness across multiple infrastructure nodes. These approaches can coordinate anomaly detection and response strategies while mitigating the risk of centralized points of failure. Figure 10 illustrates a strategic framework integrating these future directions, emphasizing how theoretical advances can be translated into deployable solutions within real-world smart grid infrastructures.

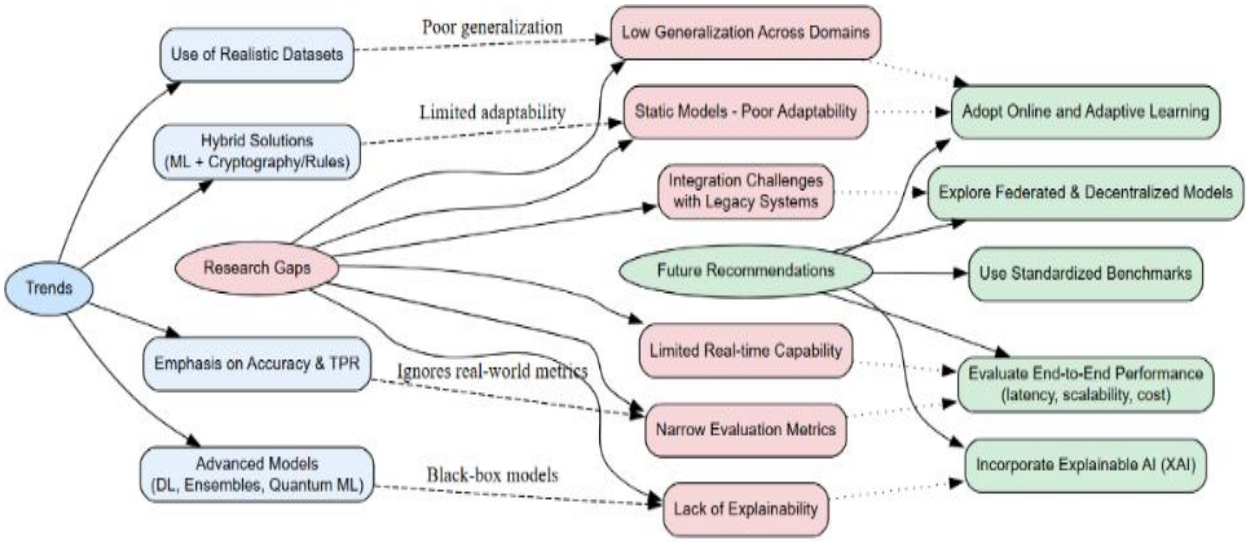

Figure 10: Strategic Framework for Advancing ML-Driven Cybersecurity for Smart Grids

10. Conclusion

This systematic review has highlighted the significant potential of ML techniques in strengthening cybersecurity for smart grids and critical infrastructure. The findings underscore the effectiveness of various ML approaches, such as supervised, unsupervised, and hybrid models, in detecting and mitigating cyber threats. However, the review also identifies several key challenges, including scalability issues, high computational demands, and difficulties in real-time data processing. These limitations have hindered ML solutions' widespread adoption and effectiveness in real-world settings.

This review provides an in-depth synthesis of the current state of ML applications in the cybersecurity of smart grids and critical infrastructure. Critically analyzing the existing literature reveals significant gaps, such as the need for more adaptive models to address emerging attack vectors and dynamic environments. Furthermore, it offers valuable insights into future research directions, particularly in developing more scalable, efficient, and real-time ML solutions. These insights enhance theoretical understanding and propose practical approaches for improving cybersecurity measures.

Integrating Machine Learning into cybersecurity frameworks for smart grids is crucial to building resilient, adaptive, and secure infrastructure. As cyber threats continue to evolve, sustained research and innovation in this field will be essential for keeping pace with new attack methods and ensuring the protection of critical systems. Future research must focus on overcoming existing limitations and improving the adaptability of ML models, ultimately leading to more robust defences against sophisticated and ever-changing cyber threats.

Acknowledgements: This study acknowledges all authors' contributions to the reviewed studies. Cyber Security Science Department, Federal University of Technology Minna, Nigeria.

Conflict of Interest: The authors have no conflict of interest.

Funding: No funding was received for this research.

Author's Contribution: Idowu Afe: Conceptualization, methodology, investigation, formal analysis, original draft writing, review, and editing. **Ismaila Idris:** Supervision, project administration, review, and critical editing. **Joseph Adebayo Ojeniyi:** Resources, visualization, supervision, review, and editing. **Sikiru O. Subairu:** Data curation, literature validation, review, and editing. **Moses Dogonyaro Noel:** Software support, referencing, proofreading, and manuscript formatting.

References

[1] S. Tufail, I. Parvez, S. Batool, and A. Sarwat, "A survey on cybersecurity challenges, detection, and mitigation techniques for the smart grid," *Energies*, vol. 14, no. 18, 2021, doi: 10.3390/en14185894.

[2] J. Ding, A. Qammar, Z. Zhang, A. Karim, and H. Ning, "Cyber Threats to Smart Grids: Review, Taxonomy, Potential Solutions, and Future

Directions," *Energies*, vol. 15, no. 18, 2022, doi: 10.3390/en15186799.

[3] G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong, "The 2015 Ukraine Blackout: Implications for False Data Injection Attacks," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3317–3318, 2017, doi: 10.1109/TPWRS.2016.2631891.

[4] L. Gjesvik and K. Szulecki, "Interpreting cyber-energy-security events: experts, social imaginaries, and policy discourses around the 2016 Ukraine blackout," *Eur. Secur.*, vol. 32, no. 1, pp. 104–124, 2023, doi: 10.1080/09662839.2022.2082838.

[5] M. Watney, "Cybersecurity Threats to and Cyberattacks on Critical Infrastructure: A Legal Perspective," *Eur. Conf. Inf. Warf. Secur. ECCWS*, vol. 2022-June, pp. 319–327, 2022, doi: 10.34190/eccws.21.1.196.

[6] A. Furfaro, P. Pace, and A. Parise, "Facing DDoS bandwidth flooding attacks," *Simul. Model. Pract. Theory*, vol. 98, 2020, doi: 10.1016/j.simpat.2019.101984.

[7] A. Djenna, S. Harous, and D. E. Saidouni, "Internet of things meet internet of threats: New concern cyber security issues of critical cyber infrastructure," *Appl. Sci.*, vol. 11, no. 10, 2021, doi: 10.3390/app11104580.

[8] A. Handa, A. Sharma, and S. K. Shukla, "Machine learning in cybersecurity: A review," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 9, no. 4, 2019, doi: 10.1002/widm.1306.

[9] S. Kolli, M. Hasan, B. Hazela, and A. A. J. Pazhani, "Secure the smart grid by machine learning," *2023 Int. Conf. Comput. Commun. Informatics, ICCCI 2023*, 2023, doi: 10.1109/ICCCI56745.2023.10128269.

[10] H. Balisane, E. I. Egho-Promise, E. Lyada, and F. Aina, "Towards Improved Threat Mitigation in Digital Environments: a Comprehensive Framework for Cybersecurity Enhancement," *Int. J. Res. -GRANTHAALAYAH*, vol. 12, no. 5, 2024, doi: 10.29121/granthaalayah.v12.i5.2024.5655.

[11] I. A. Kandhro *et al.*, "Detection of Real-Time Malicious Intrusions and Attacks in IoT Empowered Cybersecurity Infrastructures," *IEEE Access*, vol. 11, pp. 9136–9148, 2023, doi: 10.1109/ACCESS.2023.3238664.

[12] Z. Wang, D. Jiang, F. Wang, Z. Lv, and R. Nowak, "A polymorphic heterogeneous security architecture for edge-enabled smart grids," *Sustain. Cities Soc.*, vol. 67, 2021, doi: 10.1016/j.scs.2020.102661.

[13] E. Samanis, J. Gardiner, and A. Rashid, "Adaptive Cyber Security for Critical Infrastructure," *Proc. - 13th ACM/IEEE Int. Conf. Cyber-Physical Syst. ICCPS 2022*, pp. 304–305, 2022, doi: 10.1109/ICCPS54341.2022.00043.

[14] L. E. Lwakatare, A. Raj, I. Crnkovic, J. Bosch, and H. H. Olsson, "Large-scale machine learning systems in real-world industrial settings: A review of challenges and solutions," *Inf. Softw. Technol.*, vol. 127, 2020, doi: 10.1016/j.infsof.2020.106368.

[15] V. Sharma, "A Study on Data Scaling Methods for Machine Learning," *Int. J. Glob. Acad. Sci. Res.*, vol. 1, no. 1, 2022, doi: 10.55938/ijgasr.v1i1.4.

[16] A. Menon, A. Anurag, R. Ojha, H. Bhosale, and S. Oak, "DDOS Intrusion Detection Using Hybrid ML Model," *2023 3rd Int. Conf. Smart Gener. Comput. Commun. Networking, SMART GENCON 2023*, 2023, doi: 10.1109/SMARTGENCON60755.2023.10442411.

[17] S. Ponnapalli, R. R. Dornala, and K. T. Sai, "A Hybrid Learning Model for Detecting Attacks in Cloud Computing," *Proc. - 2024 3rd Int. Conf. Sentim. Anal. Deep Learn. ICSADL 2024*, pp. 318–324, 2024, doi: 10.1109/ICSADL61749.2024.00058.

[18] R. H. R. Al-Ruwaili and O. M. Ouda, "A Hybrid Classification Approach of Network Attacks using Supervised and Unsupervised Learning," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 8, pp. 818–828, 2023, doi: 10.14569/IJACSA.2023.0140890.

[19] A. Mustapha, A. M. Abdul-Rani, N. Saad, and M. Mustapha, "Advancements in traffic simulation for enhanced road safety: A review," *Simul. Model. Pract. Theory*, vol. 137, 2024, doi: 10.1016/j.simpat.2024.103017.

[20] A. Hussain, E. Marin Tordera, X. Masip-Bruin, and H. C. Leligou, "Rule-Based With Machine Learning IDS for DDoS Attack Detection in Cyber-Physical Production Systems (CPPS)," *IEEE Access*, vol. 12, no. August, pp. 114894–114911, 2024, doi: 10.1109/ACCESS.2024.3445261.

[21] E. C. Balta, M. Pease, J. Moyne, K. Barton, and D. M. Tilbury, "Digital Twin-Based Cyber-Attack Detection Framework for Cyber-Physical Manufacturing Systems," *IEEE Trans. Autom. Sci. Eng.*, vol. 21, no. 2, pp. 1695–1712, 2024, doi: 10.1109/TASE.2023.3243147.

[22] S. S. A. Naqvi, Y. Li, and M. Uzair, "DDoS attack detection in smart grid network using reconstructive machine learning models," *PeerJ Comput. Sci.*, vol. 10, pp. 1–23, 2024, doi: 10.7717/peerj-cs.1784.

[23] A. A. Alahmadi *et al.*, "DDoS Attack Detection in IoT-Based Networks Using Machine Learning Models: A Survey and Research Directions," *Electron.*, vol. 12, no. 14, pp. 1–24, 2023, doi: 10.3390/electronics12143103.

[24] H. Goyel and K. S. Swarup, "Data Integrity Attack Detection Using Ensemble-Based Learning for Cyber-Physical Power Systems," *IEEE Trans. Smart Grid*, vol. 14, no. 2, pp. 1198–1209, 2023, doi: 10.1109/TSG.2022.3199305.

[25] T. Cai, T. Jia, S. Adepu, Y. Li, and Z. Yang, "ADAM: An Adaptive DDoS Attack Mitigation Scheme in Software-Defined Cyber-Physical System," *IEEE Trans. Ind. Informatics*, vol. 19, no. 6, pp. 7802–7813, 2023, doi: 10.1109/TII.2023.3240586.

[26] E. Meriaux, D. Koehler, M. Z. Islam, V. Vokkarane, and Y. Lin, "Performance Comparison of Machine Learning Methods in DDoS Attack Detection in Smart Grids," *2022 IEEE MIT Undergrad. Res. Technol. Conf. URTC 2022*, pp. 1–5, 2022, doi: 10.1109/URTC56832.2022.10002244.

[27] D. Said, "Quantum Computing and Machine Learning for Cybersecurity: Distributed Denial of Service (DDoS) Attack Detection on Smart Micro-Grid," *Energies*, vol. 16, no. 8, Apr. 2023, doi: 10.3390/en16083572.

[28] A. Kumari *et al.*, "AI-Empowered Attack Detection and Prevention Scheme for Smart Grid System," *Mathematics*, vol. 10, no. 16, pp. 1–18, 2022, doi: 10.3390/math10162852.

[29] A. Raza, S. Memon, M. A. Nizamani, and M. Hussain Shah, "Machine Learning-Based Security Solutions for Critical Cyber-Physical Systems," *10th Int. Symp. Digit. Forensics Secur. ISDFS 2022*, pp. 1–6, 2022, doi: 10.1109/ISDFS55398.2022.9800811.

[30] S. Gyawali and O. Beg, "Cyber Attacks Detection using Machine Learning in Smart Grid Systems," *INFOCOM WKSHPS 2022 - IEEE Conf. Comput. Commun. Work.*, pp. 1–2, 2022, doi: 10.1109/INFOCOMWKSHPS54753.2022.9797941.

[31] A. Aribisala, M. S. Khan, and G. Husari, "Feed-Forward Intrusion Detection and Classification on a Smart Grid Network," *2022 IEEE 12th*

Annu. Comput. Commun. Work. Conf. CCWC 2022, pp. 99–105, 2022, doi: 10.1109/CCWC54503.2022.9720898.

[32] F. B. Saghezchi, G. Mantas, M. A. Violas, A. M. de Oliveira Duarte, and J. Rodriguez, "Machine Learning for DDoS Attack Detection in Industry 4.0 CPPSs," *Electron.*, vol. 11, no. 4, pp. 1–14, 2022, doi: 10.3390/electronics11040602.

[33] V. K. Singh and M. Govindarasu, "A Cyber-Physical Anomaly Detection for Wide-Area Protection Using Machine Learning," *IEEE Trans. Smart Grid*, vol. 12, no. 4, pp. 3514–3526, 2021, doi: 10.1109/TSG.2021.3066316.

[34] Y. A. Farrukh, Z. Ahmad, I. Khan, and R. M. Elavarasan, "A Sequential Supervised Machine Learning Approach for Cyber Attack Detection in a Smart Grid System," *2021 North Am. Power Symp. NAPS 2021*, 2021, doi: 10.1109/NAPS52732.2021.9654767.

[35] Z. Wang, W. Cheng, and C. Li, "DoS attack detection model of smart grid based on machine learning method," *Proc. 2020 IEEE Int. Conf. Power, Intell. Comput. Syst. ICPICS 2020*, pp. 735–738, 2020, doi: 10.1109/ICPICS50287.2020.9202401.

[36] C. Hu, J. Yan, and C. Wang, "Robust feature extraction and ensemble classification against cyber-physical attacks in the smart grid," *2019 IEEE Electr. Power Energy Conf. EPEC 2019*, vol. 3, 2019, doi: 10.1109/EPEC47565.2019.9074827.

[37] K. N. Junejo and J. Goh, "Behaviour-based attack detection and classification in cyber physical systems using machine learning," *CPSS 2016 - Proc. 2nd ACM Int. Work. Cyber-Physical Syst. Secur. Co-located with Asia CCS 2016*, no. Ml, pp. 34–43, 2016, doi: 10.1145/2899015.2899016.

[38] F. Mohammadi, "Emerging challenges in smart grid cybersecurity enhancement: A review," *Energies*, vol. 14, no. 5, 2021, doi: 10.3390/en14051380.

[39] S. Ness, "Adversarial Attack Detection in Smart Grids Using Deep Learning Architectures," *IEEE Access*, 2024, doi: 10.1109/ACCESS.2024.3523409.

[40] A. Shees, M. Tariq, and A. I. Sarwat, "Cybersecurity in Smart Grids: Detecting False Data Injection Attacks Utilizing Supervised Machine Learning Techniques," *Energies*, vol. 17, no. 23, 2024, doi: 10.3390/en17235870.

[41] R. Rivera-Lopez, J. Canul-Reich, E. Mezura-Montes, and M. A. Cruz-Chávez, "Induction of decision trees as classification models through metaheuristics," *Swarm Evol. Comput.*, vol. 69, 2022, doi: 10.1016/j.swevo.2021.101006.

[42] M. H. Roy and D. Larocque, "Prediction intervals with random forests," *Stat. Methods Med. Res.*, vol. 29, no. 1, pp. 205–229, 2020, doi: 10.1177/0962280219829885.

[43] K. Matsuki, V. Kuperman, and J. A. Van Dyke, "The Random Forests statistical technique: An examination of its value for the study of reading," *Sci. Stud. Read.*, vol. 20, no. 1, pp. 20–33, 2016, doi: 10.1080/10888438.2015.1107073.

[44] H. Wang, Z. Zhu, and Y. Shao, "Fast Support Vector Machine With Low-Computational Complexity for Large-Scale Classification," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 54, no. 7, pp. 4151–4163, 2024, doi: 10.1109/TSMC.2024.3375021.

[45] A. Al Bataineh, D. Kaur, and S. M. J. Jalali, "Multi-Layer Perceptron Training Optimization Using Nature Inspired Computing," *IEEE Access*, vol. 10, pp. 36963–36977, 2022, doi: 10.1109/ACCESS.2022.3164669.

[46] Z. Zhu, "Systematic Optimization of Overfitting Problem in Machine Learning," *Highlights Sci. Eng. Technol.*, vol. 111, pp. 353–359, 2024, doi: 10.54097/3tkzrj84.

[47] S. Vasishth, "Using approximate Bayesian computation for estimating parameters in the cue-based retrieval model of sentence processing," *MethodsX*, vol. 7, 2020, doi: 10.1016/j.mex.2020.100850.

[48] L. Liao, H. Li, W. Shang, and L. Ma, "An Empirical Study of the Impact of Hyperparameter Tuning and Model Optimization on the Performance Properties of Deep Neural Networks," *ACM Trans. Softw. Eng. Methodol.*, vol. 31, no. 3, 2022, doi: 10.1145/3506695.

[49] V. Ivanova, T. Tashev, and I. Draganov, "DDoS Attacks Classification using SVM," *WSEAS Trans. Inf. Sci. Appl.*, vol. 19, pp. 1–11, 2022, doi: 10.37394/23209.2022.19.1.

[50] M. A. K. Raiaan *et al.*, "A systematic review of hyperparameter optimization techniques in Convolutional Neural Networks," *Decis. Anal. J.*, vol. 11, 2024, doi: 10.1016/j.dajour.2024.100470.

[51] J. Gui *et al.*, "A Survey on Self-supervised Learning: Algorithms, Applications, and Future Trends," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2024, doi: 10.1109/TPAMI.2024.3415112.

[52] S. J. Pinto, P. Siano, and M. Parente, "Review of Cybersecurity Analysis in Smart Distribution Systems and Future Directions for Using Unsupervised Learning Methods for Cyber Detection," *Energies*, vol. 16, no. 4, 2023, doi: 10.3390/en16041651.

[53] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, and J. Heming, "K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data," *Inf. Sci. (Ny).*, vol. 622, pp. 178–210, 2023, doi: 10.1016/j.ins.2022.11.139.

[54] C. Fan, F. Xiao, Y. Zhao, and J. Wang, "Analytical investigation of autoencoder-based methods for unsupervised anomaly detection in building energy data," *Appl. Energy*, vol. 211, pp. 1123–1135, 2018, doi: 10.1016/j.apenergy.2017.12.005.

[55] H. Xu, G. Pang, Y. Wang, and Y. Wang, "Deep Isolation Forest for Anomaly Detection," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 12, pp. 12591–12604, 2023, doi: 10.1109/TKDE.2023.3270293.

[56] B. Tu, X. Yang, N. Li, C. Zhou, and D. He, "Hyperspectral anomaly detection via density peak clustering," *Pattern Recognit. Lett.*, vol. 129, pp. 144–149, 2020, doi: 10.1016/j.patrec.2019.11.022.

[57] S. M. Erfani, S. Rajasegarar, S. Karunasekera, and C. Leckie, "High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning," *Pattern Recognit.*, vol. 58, pp. 121–134, 2016, doi: 10.1016/j.patcog.2016.03.028.

[58] A. A. J. Al-Abadi, M. B. Mohamed, and A. Fakhfakh, "Enhanced Random Forest Classifier with K-Means Clustering (ERF-KMC) for Detecting and Preventing Distributed-Denial-of-Service and Man-in-the-Middle Attacks in Internet-of-Medical-Things Networks," *Computers*, vol. 12, no. 12, 2023, doi: 10.3390/computers12120262.

[59] V. Zenkov, "Bagging in the Lack of Reproducibility of Neural Network Results," in *Proceedings of 2024 17th International Conference on Management of Large-Scale System Development, MLSD 2024*, 2024. doi: 10.1109/MLSD61779.2024.10739475.

[60] N. H. N. B. M. Shahri, S. B. S. Lai, M. B. Mohamad, H. A. B. A. Rahman, and A. Bin Rambli, "Comparing the performance of adaboost, xgboost, and logistic regression for imbalanced data," *Math. Stat.*, vol. 9, no. 3, pp. 379–385, 2021, doi: 10.13189/ms.2021.090320.

[61] S. Ardabili, A. Mosavi, and A. R. Várkonyi-Kóczy, "Advances in Machine Learning Modeling Reviewing Hybrid and Ensemble Methods," *Lect. Notes Networks Syst.*, vol. 101, pp. 215–227, 2020, doi: 10.1007/978-3-030-36841-8_21.

[62] C. Benrebbouh, H. Mansouri, S. Cherbal, S. Djahel, and D. Arrar, "An Explainable CNN-based Intrusion Detection System for Enhanced Smart Grid Security," in *2024 International Conference on Information and Communication Technologies for Disaster Management, ICT-DM 2024*, 2024. doi: 10.1109/ICT-DM62768.2024.10798953.

[63] P. Thejaswini, G. Suresh, V. Chiraag, and S. Nandi, "Approximate CNN Hardware Accelerators for Resource Constrained Devices," *IEEE Access*, 2025, doi: 10.1109/ACCESS.2025.3529668.

[64] G. Francis, M. Sanan, M. Hatoum, N. Bakir, and K. Samrouth, "Detecting Advanced Persistent Threats on a Network Using Machine Learning," in *2024 IEEE International Conference on Smart Systems and Power Management, IC2SPM 2024*, 2024, pp. 89–94. doi: 10.1109/IC2SPM62723.2024.10841340.

[65] M. Shen, J. Yang, S. Li, A. Zhang, and Q. Bai, "Nonlinear hyperparameter optimization of a neural network in image processing for micromachines," *Micromachines*, vol. 12, no. 12, 2021, doi: 10.3390/mi12121504.

[66] S. Lin, R. Clark, R. Birke, S. Schonborn, N. Trigoni, and S. Roberts, "Anomaly Detection for Time Series Using VAE-LSTM Hybrid Model," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, vol. 2020-May, pp. 4322–4326, 2020, doi: 10.1109/ICASSP40776.2020.9053558.

[67] H. Naveed, S. Anwar, M. Hayat, K. Javed, and A. Mian, "Survey: Image mixing and deleting for data augmentation," *Eng. Appl. Artif. Intell.*, vol. 131, 2024, doi: 10.1016/j.engappai.2023.107791.

[68] T. Kotsiopoulos, P. Sarigiannidis, D. Ioannidis, and D. Tzovaras, "Machine Learning and Deep Learning in smart manufacturing: The Smart Grid paradigm," *Comput. Sci. Rev.*, vol. 40, 2021, doi: 10.1016/j.cosrev.2020.100341.

[69] E. Heidaryan, "A note on model selection based on the percentage of accuracy-precision," *J. Energy Resour. Technol. Trans. ASME*, vol. 141, no. 4, 2019, doi: 10.1115/1.4041844.

[70] A. Tzovara *et al.*, "High-precision magnetoencephalography for reconstructing amygdalar and hippocampal oscillations during prediction of safety and threat," *Hum. Brain Mapp.*, vol. 40, no. 14, pp. 4114–4129, 2019, doi: 10.1002/hbm.24689.

[71] P. L. Bhattar and N. M. Pindoriya, "False Data Injection Attack with Max-Min Optimization in Smart Grid," *Comput. Secur.*, vol. 140, 2024, doi: 10.1016/j.cose.2024.103761.

[72] S. Riyanto, I. S. Sitanggang, T. Djatna, and T. D. Atikah, "Comparative Analysis using Various Performance Metrics in Imbalanced Data for Multi-class Text Classification," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 6, pp. 1082–1090, 2023, doi: 10.14569/IJACSA.2023.01406116.

[73] Y.-J. Sun, S. Dey, D. Hakkani-Tur, and G. Tur, "Confidence Estimation for LLM-Based Dialogue State Tracking," 2024, doi: 10.1109/SLT61566.2024.10832237.

[74] A. Mahi-Al-rashid, F. Hossain, A. Anwar, and S. Azam, "False Data Injection Attack Detection in Smart Grid Using Energy Consumption Forecasting," *Energies*, vol. 15, no. 13, 2022, doi: 10.3390/en15134877.

[75] J. Bian *et al.*, "Machine Learning in Real-Time Internet of Things (IoT) Systems: A Survey," *IEEE Internet Things J.*, vol. 9, no. 11, pp. 8364–8386, 2022, doi: 10.1109/JIOT.2022.3161050.

[76] Y. Gu, "A Comparative Analysis Study of Stock Prediction Based on Random Forest and Decision Tree," *Proc. - 2024 Int. Conf. Electron. Devices, Comput. Sci. ICEDCS 2024*, pp. 96–100, 2024, doi: 10.1109/ICEDCS64328.2024.00022.

[77] B. A. Tama and S. Lim, "Ensemble learning for intrusion detection systems: A systematic mapping study and cross-benchmark evaluation," *Comput. Sci. Rev.*, vol. 39, p. 100357, 2021, doi: 10.1016/j.cosrev.2020.100357.

[78] B. Talekar, "A Detailed Review on Decision Tree and Random Forest," *Biosci. Biotechnol. Res. Commun.*, vol. 13, no. 14, pp. 245–248, 2020, doi: 10.21786/bbrc/13.14/57.

[79] F. Lin, S. C. Fang, X. Fang, and Z. Gao, "Distributionally robust chance-constrained kernel-based support vector machine," *Comput. Oper. Res.*, vol. 170, 2024, doi: 10.1016/j.cor.2024.106755.

[80] H. Wang and R. Liu, "Hiding outliers into crowd: Privacy-preserving data publishing with outliers," *Data Knowl. Eng.*, vol. 100, pp. 94–115, 2015, doi: 10.1016/j.datak.2015.06.012.

[81] C. Xu, Z. Liao, C. Li, X. Zhou, and R. Xie, "Review on Interpretable Machine Learning in Smart Grid," *Energies*, vol. 15, no. 12, 2022, doi: 10.3390/en15124427.

[82] H. Padmanaban, "Machine Learning Algorithms Scaling on Large-Scale Data Infrastructure," *J. Artif. Intell. Gen. Sci. ISSN3006-4023*, vol. 3, no. 1, pp. 1–26, 2024, doi: 10.60087/jaigs.vol03.issue01.p26.

[83] R. Huang and Y. Li, "Adversarial Attack Mitigation Strategy for Machine Learning-Based Network Attack Detection Model in Power System," *IEEE Trans. Smart Grid*, vol. 14, no. 3, pp. 2367–2376, 2023, doi: 10.1109/TSG.2022.3217060.

[84] S. khan, T. Mazhar, T. Shahzad, M. Amir khan, A. U. Rehman, and H. Hamam, "Integration of smart grid with Industry 5.0: Applications, challenges and solutions," *Meas. Energy*, vol. 5, p. 100031, 2025, doi: https://doi.org/10.1016/j.meaene.2024.100031.