Journal of Engineering Management and Soft Computing (JEMSO)

Journal of Data Analytics and Intelligent Decision-Making

Journal homepage: jdaid.qom.ac.ir

University of Qom

AI-Driven Protection Schemes for Modern Power Grids: Technologies, Challenges, and Opportunities

Abbas Jamshidi Gahrouei[a]*, Reza Asgharian[b]

a. Electricity Distribution Company of Chaharmahal and Bakhtiari Province, Shahrekord, Iran.
b. National University of Skills, Chaharmahal and Bakhtiari, Shahrekord, Iran.

A R T I C L E I N F O

Keywords:
Adaptive Protection
Artificial Intelligence (AI)
Cybersecurity
Smart Grids

A B S T R A C T

The rapid evolution of modern power grids, driven by the large-scale integration of intermittent renewable energy sources, distributed generation, and power electronic interfaces, has significantly challenged the effectiveness of conventional threshold-based protection schemes. Traditional systems often exhibit limited performance under dynamic fault currents, evolving network topologies, and complex transients, increasing the risk of maloperation and reduced system resilience. In this context, Artificial Intelligence (AI) has emerged as a promising data-driven paradigm to enhance the intelligence and reliability of grid protection. This paper presents a comprehensive review of AI-driven protection schemes for modern power grids, systematically analyzing underlying technologies, key challenges, and emerging opportunities. A structured taxonomy of AI methodologies, ranging from Machine Learning (ML) and Deep Learning (DL) to Reinforcement Learning (RL) and hybrid systems, is developed, with a focus on their applications in fault detection, classification, location, and adaptive relay coordination. Particular emphasis is placed on recent advancements reported after 2022, highlighting the adoption of advanced models, such as Vision Transformers, Graph Neural Networks (GNNs), and Physics-Informed Neural Networks (PINNs), to address grid nonlinearity and data scarcity. Furthermore, the paper critically examines major barriers to large-scale deployment, including explainability (XAI), cybersecurity vulnerabilities, and real-time computational constraints. Finally, a strategic roadmap is proposed, identifying future research directions, such as digital twins, federated learning, and edge AI, to bridge the gap between theoretical models and practical, self-healing protection systems.

* Corresponding author.
E-mail addresses: abbas.jamshidi133@gmail.com (A. Jamshidi Gahrouei), rezaasgharian1382@gmail.com (R. Asgharian)
Received 08 Jan 2026; Received in revised form 02 Feb 2026; Accepted 12 Mar 2026
Available online 30 Mar 2026

Cite this article: Jamshidi Gahrouei, A., Asgharian, R. (2026). AI-Driven Protection Schemes for Modern Power Grids: Technologies, Challenges, and Opportunities. *Journal of Data Analytics and Intelligent Decision-Making*, 2(1), 17-30.
https://doi.org/100000000000000000

1. Introduction

The global energy sector is currently undergoing a fundamental transformation driven by the pillars of decarbonization, digitalization, and decentralization. Consequently, modern power grids are evolving into highly complex cyber–physical systems (CPSs), characterized by a high penetration of inverter-based resources (IBRs), widespread deployment of distributed energy resources (DERs), and the emergence of active distribution networks and microgrids (Khaw et al., 2021; Oelhaf et al., 2025; Wang et al., 2025). While this transition significantly enhances system flexibility, sustainability, and resilience, it also introduces unprecedented challenges to the design and operation of legacy power system protection schemes. Conventional protection methods—specifically overcurrent, distance, and differential protection were primarily engineered based on deterministic system models, fixed relay settings, and the predictable fault-current characteristics associated with traditional synchronous generators (Diahovchenko et al., 2020; Krause et al., 2021).

However, the shift toward inverter-dominated and dynamically reconfigurable networks fundamentally alters fault behavior. Factors such as bidirectional and limited fault currents from IBRs, frequent topology changes, the increased occurrence of high-impedance faults, and the massive volume of high-resolution measurements generated by phasor measurement units (PMUs) and Internet-of-Things (IoT) sensors have significantly degraded the reliability of traditional protection schemes (Centeno et al., 2010; Dasand & Panigrahi, 2022; Jamshidi et al., 2026; Krause et al., 2021). As a result, legacy relays are increasingly susceptible to maloperation risks, including nuisance tripping, protection blinding, and loss of coordination, which may ultimately compromise the stability and security of the entire grid (Lotfifard et al., 2009; Mishra et al., 2025).

Artificial Intelligence (AI), particularly machine learning (ML) and deep learning (DL), has emerged as a promising enabler for next-generation protection systems. By learning complex nonlinear relationships from high-dimensional measurement data, AI facilitates a paradigm shift from traditional rule-based and setting-dependent protection toward adaptive, condition-based, and predictive decision-making (Alhamrouni et al., 2024; Mishra et al., 2025; Oelhaf et al., 2025). Early research demonstrated the feasibility of classical ML techniques, such as support vector machines (SVMs) and decision trees, for fault detection and classification (Centeno et al., 2025; Lotfifard et al., 2009). More recently, advances in sensing infrastructure and computational platforms have accelerated the adoption of DL-based approaches. Convolutional neural networks (CNNs) are now widely applied to automatically extract discriminative features from transient waveforms (Bakkar et al., 2022; Wang et al., 2025), while recurrent neural networks (RNNs) and long short-term memory (LSTM) architectures have proven effective in modeling temporal dependencies within synchronized PMU data streams (Hijazi et al., 2023; Meng et al., 2024).

Beyond these architectures, emerging models, such as graph neural networks (GNNs), explicitly incorporate grid topology into the learning process, enabling topology-aware fault diagnosis and localization, whereas physics-informed neural networks (PINNs) embed physical constraints to enhance model generalization under conditions of limited or noisy training data (Caldas et al., 2020; Nadal et al., 2025; Zhang et al., 2026).

Despite these technological advancements, the practical deployment of AI-based protection schemes in real-world power systems remains limited. Several critical challenges hinder their transition from research prototypes to trusted grid infrastructure, including:

1. **Data Availability and Quality:** The scarcity of large-scale, labeled, and representative fault datasets, particularly under rare and extreme operating conditions (Mazumder et al., 2023; Sicard et al., 2024).

2. **Interpretability and Trust:** The "black-box" nature of complex DL models, which motivates the development of explainable AI (XAI) techniques to enhance transparency and operator trust (Chinnaraju, 2025; Machlev et al., 2022).
3. **Cybersecurity:** Increased vulnerability to adversarial data manipulation, model evasion, and cyber–physical attacks targeting AI-enabled protection systems (Meng et al., 2024).
4. **Real-Time Performance:** The difficulty of meeting stringent sub-cycle latency requirements using computationally intensive AI models, especially in wide-area protection scenarios (Choudhary et al., 2025; Meloni et al., 2018).
5. **Standardization Gaps:** The lack of unified frameworks and standards for integrating AI algorithms into existing protection hardware and communication protocols such as IEC 61850 (Nadal et al., 2025; Sidhu & Gangadharan, 2005).

Despite the rapid expansion of AI-based protection research, existing review studies largely remain fragmented, algorithm-centric, or limited to earlier developments, and therefore, fail to provide a unified and forward-looking perspective aligned with the realities of modern power grids.

The novelty of this paper lies in presenting a structured, protection-oriented, and post-2022-focused synthesis of AI-driven protection technologies. Rather than providing a conventional descriptive survey, this work introduces a unified taxonomy that systematically categorizes AI techniques according to their learning paradigms, data requirements, computational complexity, interpretability, and real-time deployability.

Furthermore, this paper uniquely integrates emerging AI paradigms, such as graph neural networks (GNNs), physics-informed neural networks (PINNs), and reinforcement learning (RL), with practical protection functions and deployment constraints, highlighting a fundamental shift from static protection philosophies toward adaptive and self-healing systems.

In addition, this study explicitly links technological advances with critical implementation challenges, including explainability (XAI), cybersecurity vulnerabilities, real-time performance constraints, and standardization gaps, and proposes a forward-looking research roadmap to bridge the gap between academic innovation and field-deployable protection systems.

Iresponse to these challenges, this paper presents a comprehensive and forward-looking review of AI-driven protection schemes for modern power grids. Rather than providing a merely descriptive survey, this work systematically organizes existing literature around three interrelated dimensions: technologies, challenges, and opportunities. Recent developments reported after 2022 are critically synthesized, and a unified taxonomy of AI techniques for power system protection is established. Furthermore, a research roadmap is proposed, highlighting emerging directions such as AI-enabled digital twins, federated learning for privacy preservation, and edge AI for low-latency decentralized intelligence (Choudhary et al., 2025; Idrisov et al., 2025; Wen et al., 2021).

The remainder of this paper is organized as follows. Section II introduces the taxonomy of AI technologies for power system protection. Section III reviews key application domains. Section IV discusses major implementation challenges. Section V outlines future research opportunities, and Section VI concludes the paper.

2. AI-Driven Protection Technologies: A Structured Taxonomy

The integration of AI into power system protection marks a fundamental shift from deterministic, threshold-based logic toward adaptive, data-driven, and context-aware decision-making frameworks. This transition is increasingly motivated by the rapid growth of inverter-dominated grids, frequent network reconfigurations, reduced fault current levels, and the proliferation of high-resolution measurements from PMUs and intelligent electronic devices

(IEDs) (Oelhaf et al., 2025; Wang et al., 2025). Recent studies published after 2022 emphasize that conventional protection schemes, originally designed for synchronous generator–dominated systems, are no longer sufficient to guarantee selectivity, speed, and dependability under these emerging operating conditions (Alhamrouni et al., 2024; Mishra et al., 2025).

To systematize the rapidly expanding body of literature, AI-driven protection technologies can be categorized according to their learning paradigms and functional roles, ranging from classical machine learning and deep learning to graph-based, physics-informed, and reinforcement learning frameworks.

A. Classical Machine Learning Approaches

Classical machine learning (ML) techniques remain attractive for protection applications due to their relatively low computational burden and more transparent decision boundaries compared to deep models (Diahovchenko et al., 2020). These approaches typically rely on handcrafted features extracted from voltage and current signals, such as wavelet coefficients, symmetrical components, and statistical indices.

Supervised learning algorithms, including Support Vector Machines (SVMs), k-Nearest Neighbors (k-NN), and Random Forests, have been widely applied to fault detection, classification, and zone identification (Centeno et al., 2010; Lotfifard et al., 2009). Recent works continue to refine these methods by combining ensemble learning and feature selection strategies to improve robustness under noise, parameter uncertainty, and changing grid conditions (Hijazi et al., 2023). Despite their maturity, the main limitation of classical ML remains their dependence on expert-driven feature engineering and their reduced scalability to highly complex, high-dimensional measurement spaces.

Unsupervised learning techniques, such as K-means, DBSCAN, and autoencoders, have gained renewed attention in recent literature (2023–2025) for anomaly and event detection, particularly in scenarios where labeled fault data are scarce or unreliable (Caldas et al., 2020; Idrisov et al., 2025). By learning normal operating patterns and detecting deviations, these methods provide a complementary layer of situational awareness and can serve as early-warning mechanisms in cyber–physical protection architectures.

B. Deep Learning Architectures for Data-Driven Protection

Deep learning (DL) has become the dominant paradigm in recent AI-based protection research due to its ability to automatically learn hierarchical feature representations directly from raw or minimally processed measurements.

Convolutional Neural Networks (CNNs): One-dimensional CNNs are widely used for ultra-fast protection tasks based on sampled current and voltage waveforms, while two-dimensional CNNs operate on time–frequency representations, such as spectrograms and wavelet scalograms (Bakkar et al., 2022; Livani & Evrenosoglu, 2013). Post-2022 studies increasingly report superior performance of CNN-based models in detecting complex events, including high-impedance faults and evolving fault transients in inverter-dominated networks (Mishra et al., 2025; Wang et al., 2025).

Recurrent Neural Networks (RNNs) and LSTM Models: Recurrent architectures, particularly Long Short-Term Memory (LSTM) networks, are well suited for modeling temporal dependencies in sequential measurement streams. Recent wide-area protection and stability assessment schemes leverage LSTM-based models to process synchronized PMU data, enabling early fault detection, accurate fault location, and short-term stability prediction (Hijazi et al., 2023; Meng et al., 2024).

Attention-Based and Transformer Models: More recent contributions (2023–2026) explore attention mechanisms and transformer-inspired architectures to better capture long-range temporal dependencies and multi-channel correlations in power system measurements (Mishra et al., 2025; Oelhaf et al., 2025). These models offer improved scalability and interpretability

potential compared to traditional RNN-based approaches, although their real-time deployment in protection hardware remains an open research challenge due to computational complexity (Choudhary et al., 2025).

C. Graph-Based and Physics-Informed Learning

Graph Neural Networks (GNNs): GNNs represent a significant step toward topology-aware protection by explicitly modeling the power grid as a graph, where buses correspond to nodes and transmission or distribution lines to edges (Choudhary et al., 2025; He & Zhao, 2020). This paradigm enables learning algorithms to naturally incorporate network structure, making them particularly suitable for adaptive protection in reconfigurable networks, microgrids, and active distribution systems. Recent studies highlight that GNN-based models maintain higher robustness under topology changes, islanding events, and switching operations compared to topology-agnostic deep learning approaches (Meng et al., 2024).

Physics-Informed Neural Networks (PINNs): Physics-Informed Neural Networks integrate physical laws, system equations, and operational constraints directly into the learning process through customized loss functions (Nadal et al., 2025; Zhang et al., 2026). This hybridization of data-driven learning and first-principles modeling has gained strong momentum after 2023, especially in scenarios characterized by limited or noisy training data (Caldas et al., 2020). In the context of protection, PINNs improve generalization, enhance physical consistency of decisions, and reduce the risk of non-physical predictions.

D. Reinforcement Learning and Adaptive Decision-Making

Reinforcement Learning (RL), and particularly Deep Reinforcement Learning (DRL), introduces a fundamentally different paradigm in which protection agents learn optimal policies through interaction with simulated or digital-twin environments (Dinneweth et al., 2022; Idrisov et al., 2025). Recent studies demonstrate the potential of DRL for adaptive relay coordination, microgrid protection, and system integrity protection schemes (SIPS), where static, rule-based logic is increasingly inadequate (Meng et al., 2024).

Multi-agent reinforcement learning frameworks further extend this concept by enabling distributed protection agents to coordinate their actions under partial observability and dynamic network conditions. While promising, these approaches still face challenges related to training stability, safety guarantees, and certification for mission-critical deployment.

E. Hybrid, Explainable, and Data-Centric Frameworks

Hybrid models, such as Neuro-Fuzzy systems and neuro-symbolic architectures, aim to combine the high accuracy of deep learning with the interpretability and transparency of rule-based systems (Bakkar et al., 2022; Diahovchenko et al., 2020). This line of research has gained renewed importance in recent years due to the growing emphasis on Explainable AI (XAI) for safety-critical power system applications (Chinnaraju, 2025; Machlev et al., 2022).

In parallel, data-centric AI approaches focus on improving dataset quality, representativeness, and benchmarking rather than solely increasing model complexity. Techniques such as synthetic data generation using GANs, domain adaptation, and transfer learning are increasingly reported in post-2022 literature as key enablers for robust and generalizable protection models (Mazumder et al., 2023; Rizzato et al., 2022).

F. Enabling Data and Computing Infrastructure

The effectiveness of all aforementioned AI paradigms critically depends on the availability of high-quality, time-synchronized measurements provided by Wide-Area Measurement Systems (WAMS), digital substations, and advanced sensing infrastructures (Krause et al., 2021). At the same time, recent research highlights a clear architectural trend toward edge–cloud hierarchical intelligence, where lightweight, low-latency models operate at the edge for primary protection, while more complex models are deployed at substation or control-center

levels for coordination, learning, and optimization (Choudhary et al., 2025; Mazumder et al., 2023).

Table 1

Structured Taxonomy of AI-Driven Protection Technologies

AI Paradigm	Representative Algorithms	Input Data Type	Computational Complexity	Interpretability	Primary Protection Application
Classical Machine Learning	SVM, k-NN, Random Forest	Extracted voltage/current features	Low	Medium	Fault detection and classification
Deep Learning (DL)	CNN, LSTM, RNN	Raw waveforms, PMU data	High	Low	Transient analysis and fault location
Attention & Transformer Models	Attention NN, Transformers	Multi-channel time-series data	High	Low–Medium	Wide-area protection and temporal dependency modeling
Graph-Based AI	GNN, GraphSAGE	Topology + electrical measurements	Very High	Low	Topology-aware protection and microgrids
Physics-Informed Learning	PINNs, Physics-loss NN	Limited data + physical constraints	Medium	Medium–High	Protection under data scarcity and noise
Reinforcement Learning	DQN, PPO, DDPG	Simulation / digital twin data	High	Low	Adaptive relay coordination and SIPS
Hybrid & Fuzzy Systems	ANFIS, Neuro-Fuzzy	Measurements + expert rules	Medium	High	Adaptive coordination and explainability
Data-Centric & XAI Frameworks	GANs, Transfer Learning, XAI tools	Real and synthetic datasets	Medium	High	Robustness enhancement and operator trust

3. AI-Driven Protection Applications Across Modern Grid Domains

The decentralization and digitalization of power systems necessitate a fundamental rethinking of protection philosophies. AI has emerged as a key enabler for a new generation of protection schemes characterized by adaptability, scalability, and system-level coordination. Rather than replacing conventional protection principles, AI-based approaches augment existing frameworks by enhancing situational awareness and decision-making across multiple layers of the modern grid (Mishra et al., 2025; Oelhaf et al., 2025).

A. Transmission System Protection

Transmission networks require extremely high dependability, as protection failures can propagate disturbances and trigger cascading blackouts. AI-based techniques significantly enhance protection performance under stressed and uncertain operating conditions.

1) Transmission Line Protection:

AI-assisted adaptive distance protection mitigates the limitations of fixed relay settings during power swings, load encroachment, and evolving network conditions. Neuro-fuzzy

systems and reinforcement learning (RL) methods dynamically adapt relay characteristics using synchronized measurements, thereby preserving fast and selective fault clearance under stressed system states (Bakkar et al., 2022; Diahovchenko et al., 2020). In addition, hybrid CNN–LSTM architectures have demonstrated superior accuracy in fault detection, classification, and location when applied to time-synchronized PMU data streams (Hijazi et al., 2023; Meng et al., 2024).

2) Transformer and Busbar Protection:

Deep learning models, particularly convolutional neural networks (CNNs), have been shown to improve discrimination between internal faults, magnetizing inrush currents, and current transformer (CT) saturation by learning subtle waveform patterns that are difficult to capture with rule-based logic (Livani & Evrenosoglu, 2013; Lotfifard et al., 2009). For asset condition monitoring, variational autoencoders (VAEs) and related deep generative models exploit multi-modal measurements, such as Dissolved Gas Analysis (DGA), temperature, and load profiles, to enable early detection of incipient transformer faults (Alhamrouni et al., 2024; Hijazi et al., 2023).

B. Active Distribution Networks (ADNs) and Microgrids

The widespread integration of inverter-based resources introduces bidirectional and limited fault currents, rendering traditional distribution protection schemes increasingly ineffective (Diahovchenko et al., 2020; Khaw et al., 2021; Wang et al., 2025).

- **Adaptive Coordination:**

Multi-agent deep reinforcement learning (MADRL) frameworks enable distributed protection agents to collaboratively adjust relay settings in response to topology changes, load variations, and islanding events in active distribution networks and microgrids (Dinneweth et al., 2022; Meng et al., 2024).

- **High-Impedance Fault (HIF) Detection:**

Two-dimensional CNNs operating on time–frequency representations, such as wavelet scalograms and spectrograms, provide enhanced sensitivity to subtle arcing signatures that are typically masked by load currents, making them well suited for high-impedance fault detection in distribution feeders (Bakkar et al., 2022; Livani & Evrenosoglu, 2013).

- **Topology-Aware Protection:**

Graph Neural Networks (GNNs) have been increasingly proposed for microgrid protection by explicitly encoding network topology into the learning process. This enables rapid adaptation of protection logic following islanding, reconfiguration, or switching operations (Choudhary et al., 2025; He & Zhao, 2020).

C. Wide-Area Protection and System Integrity (WAPS)

Wide-area protection schemes leverage synchronized measurements to coordinate protection and control actions across large geographical regions.

- **Wide-Area Backup Protection:**

Physics-informed learning approaches exploit system-wide measurements and physical constraints to detect protection malfunctions and determine selective isolation strategies under large-scale disturbances (Caldas et al., 2020; Sidhu & Gangadharan, 2005).

- **System Integrity Protection Schemes (SIPS):**

Adaptive SIPS based on reinforcement learning replace static, rule-based logic with policies learned through interaction with high-fidelity digital twins, enabling proactive and context-aware emergency control actions (Dinneweth et al., 2022; Idrisov et al., 2025).

- **Stability Assessment:**

Recurrent neural networks (RNNs) and long short-term memory (LSTM) models provide real-time transient stability assessment by estimating post-fault stability margins directly from

PMU measurements, supporting fast remedial action schemes (Dinneweth et al., 2022; Idrisov et al., 2025).

D. Cybersecurity and Digital Asset Management

As protection systems evolve into tightly coupled cyber–physical infrastructures, cybersecurity and asset health management become critical considerations.

1) Intrusion Detection:

Autoencoders and one-class support vector machines (SVMs) are widely employed to model legitimate IEC 61850 communication patterns, enabling the detection of anomalous traffic and cyber-intrusions targeting digital substations and protection devices (Alsaiari & Ilyas, 2025; Sidhu & Gangadharan, 2005).

2) Model Robustness:

Recent research increasingly focuses on safeguarding AI models themselves against adversarial manipulation, data poisoning, and evasion attacks to ensure the functional safety and reliability of AI-enabled protection systems (Meng et al., 2024; Sidhu & Gangadharan, 2005).

3) Predictive Maintenance:

Machine learning techniques integrate SCADA data, condition monitoring measurements, and environmental factors to estimate the Remaining Useful Life (RUL) of critical assets, such as circuit breakers and transformers, thereby enabling predictive and condition-based maintenance strategies (Dinneweth et al., 2022; Hijazi et al., 2023; Meloni et al., 2018).

Table 2

Summary of AI-Driven Protection Applications Across Power System Domains

Grid Domain	Protection Task	Primary Challenge	Recommended AI/DL Model	Key Benefit
Transmission	Distance Protection	Power swings & CT saturation	**ANN / Neuro-Fuzzy**	Adaptive trip boundaries
Transmission	Fault Location	Synchronized data alignment	**LSTM / RNN**	High accuracy with PMU data
Distribution	HIF Detection	Arcing masked by load	**CNN (Spectrogram-based)**	Superior sensitivity to noise
Microgrids	Adaptive Settings	Topology changes (Islanding)	**GNN / Deep RL**	Topology-aware coordination
Substations	Busbar/Transformer	Inrush current vs. Internal fault	**1-D CNN**	Automated feature extraction
Wide-Area	SIPS & Backup	Cascading failure risk	**Physics-Informed NN**	Physical consistency in decisions
Cyber-Physical	Intrusion Detection	IEC 61850 traffic anomalies	**Autoencoders / One-class SVM**	Detection of zero-day attacks

4. Critical Challenges and Barriers to Deployment

The transition of AI-based protection schemes from research prototypes to mission-critical grid infrastructure is constrained by a combination of technical, operational, and institutional barriers. These challenges can be broadly categorized into five interrelated domains: data-centric, algorithmic, cybersecurity, implementation, and standardization (Alhamrouni et al., 2024; Mishra et al., 2025).

A. Data-Centric Challenges

The effectiveness of AI-driven protection schemes is inherently bounded by the quality, diversity, and representativeness of the available training data.

1) Data Scarcity and Imbalance:

Since catastrophic faults and extreme operating conditions occur infrequently, available datasets are typically dominated by normal system behavior. This severe class imbalance can lead to AI models exhibiting high overall accuracy and poor sensitivity to rare yet critical fault events. Although data augmentation techniques, based on Generative Adversarial Networks (GANs) and EMTP-based simulations, have been proposed to mitigate this limitation, a persistent "reality gap" often exists between synthetic data and real field measurements (Bakkar et al., 2022; Rizzato et al., 2022).

2) Data Quality and Labeling:

Measurement noise, missing samples, synchronization errors in PMU data, and uncertainties in event labeling introduce inaccuracies that propagate through the learning pipeline, ultimately degrading protection performance and reliability (Jehan et al., 2025; Krause et al., 2021).

3) Benchmarking and Reproducibility:

The lack of standardized, publicly available benchmark datasets for power system protection significantly hampers reproducibility, objective comparison of algorithms, and fair performance evaluation across studies (Mazumder et al., 2023; Sicard et al., 2024).

B. Algorithmic and Technical Limitations

1) Interpretability (The Black-Box Problem):

High-performing deep learning models, including CNNs, RNNs, and emerging attention-based architectures, often operate as black boxes with limited transparency. Although Explainable Artificial Intelligence (XAI) techniques have been proposed to improve interpretability in time-series and power system applications, they frequently introduce additional computational overhead and design complexity (Chinnaraju, 2025; Khaw et al., 2021; Mazumder et al., 2024).

2) Distributional Shifts:

AI models trained on historical operating data may fail to generalize under changing conditions, such as seasonal load variations, network reconfigurations, protection setting updates, or the integration of new inverter-based generation assets, leading to performance degradation over time (Diahovchenko et al., 2020; Khaw et al., 2021).

3) Real-Time Constraints:

Meeting the stringent sub-cycle latency requirements of primary protection (typically 8–16 ms) remains a major engineering challenge when deploying computationally intensive models such as GNNs and physics-informed architectures on cost- and power-constrained protection hardware (Choudhary et al., 2025; Mazumder et al., 2023; Zhang et al., 2026).

C. Cybersecurity and Adversarial Vulnerabilities

1) Adversarial Attacks:

AI-based protection schemes are vulnerable to adversarial perturbations in measurement data that can intentionally induce misclassification, resulting in nuisance tripping, protection blinding, or delayed fault clearance (Gangadharan et al., 2005; Hafez et al., 2025).

2) Data Poisoning:

Malicious manipulation of training datasets through data poisoning attacks can cause AI models to learn incorrect associations, posing a particularly insidious threat that is difficult to detect once the model is deployed in the field (Gangadharan et al., 2005; Meng et al., 2024).

D. Implementation and Operational Hurdles

- **Legacy Integration:**

Practical deployment requires seamless interoperability between AI-based functions and legacy protection infrastructure, including compliance with established communication and automation standards such as IEC 61850 (Nadal et al., 2025).

Validation and Testing:

Standardized frameworks for life-cycle validation, robustness testing, and detection of model drift in AI-enabled relays are still at an early stage of development, limiting confidence in long-term operational reliability (W. Zhang et al., 2019; J. Zhang et al., 2026).

- **Human Factors:**

The skills gap between traditional protection engineering and data-driven AI methodologies introduces organizational and cultural resistance. This challenge underscores the need for advanced Human–Machine Interfaces (HMIs) and explainability tools to enhance operator trust and situational awareness (Chinnaraju, 2025; Mazumder et al., 2025).

E. Regulatory and Standardization Gaps

1) Liability and Accountability:

Existing regulatory and legal frameworks provide limited guidance on accountability and liability in the case of autonomous or AI-assisted protection decisions, complicating large-scale adoption (Grotto & Dempsey, 2021).

2) Standards and Governance:

While initial efforts toward standardization of AI systems, such as emerging IEEE initiatives, are underway, a comprehensive framework for auditing, certifying, and managing the life cycle of AI-based protection systems has yet to be established (Mazumder et al., 2025; Sidhu & Gangadharan, 2005).

3) Sustainability and Computational Cost:

The computational and energy costs associated with training and deploying large-scale AI models across thousands of edge devices raise concerns regarding scalability and environmental sustainability, particularly for always-on protection applications (Choudhary et al., 2025; Mazumder et al., 2023; Sicard et al., 2024).

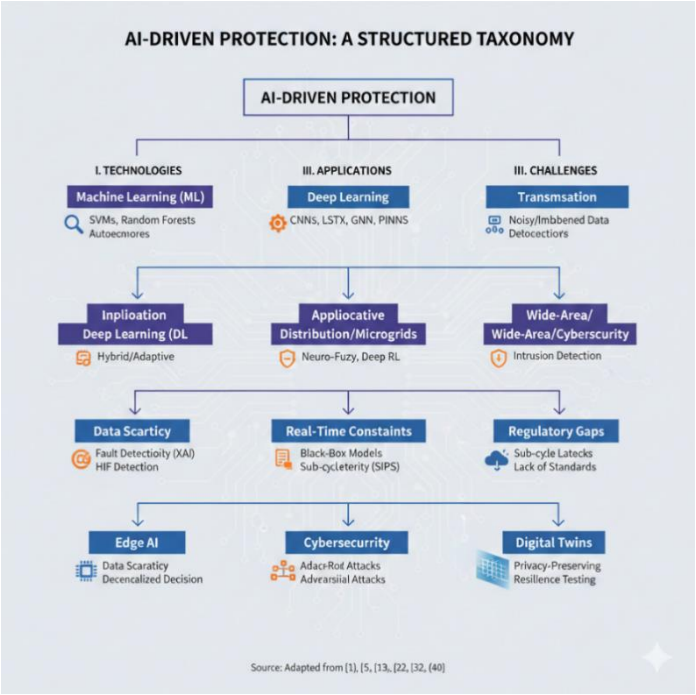

Figure 1

Integrated Taxonomy of AI-Driven Protection Schemes, Illustrating the Convergence of Advanced Algorithms, Multi-Layer Grid Applications, and Critical Technical Barriers.

V. Future Research Directions and Emerging Opportunities

The limitations identified in Section IV serve as catalysts for the next generation of AI-driven protection systems. Addressing these challenges requires a strategic research roadmap

that advances protection schemes toward resilient, autonomous, and trustworthy operational intelligence capable of operating under extreme uncertainty and evolving grid conditions (Alhamrouni et al., 2024; Zhang et al., 2026).

A. Next-Generation AI Paradigms

Future research must move beyond purely correlation-based learning toward models that explicitly incorporate physical consistency, causality, and adaptability.

1) Causal AI and Digital Twins:

Transitioning from correlation-driven inference to causal reasoning enables protection systems to distinguish between primary faults, hidden interactions, and cascading failures. The integration of AI with high-fidelity power system digital twins provides a closed-loop and risk-free environment for validating adaptive protection strategies and reinforcement learning agents under extreme contingencies (Caldas et al., 2020; Sicard et al., 2024; Zhang et al., 2019).

2) Grid Foundation Models and Federated Learning:

Inspired by the success of large-scale pre-trained models, emerging research envisions *Grid Foundation Models* trained on heterogeneous and multi-modal data, including waveforms, topology information, and environmental variables. In parallel, Federated Learning enables geographically distributed utilities to collaboratively improve protection models without sharing raw data, thereby preserving privacy and alleviating data scarcity challenges (Diahovchenko et al., 2020; Khaw et al., 2021; Krause et al., 2021).

3) Neuro-Symbolic AI:

By combining neural networks with symbolic reasoning and rule-based logic, neuro-symbolic AI offers a promising pathway toward interpretable and verifiable protection schemes. This paradigm directly addresses the black-box limitation of deep learning and supports the formal certification and auditing of AI-enabled protection relays (Chinnaraju, 2025; Mazumder et al., 2025).

B. Architectural Evolution: Distributed and Cognitive Protection

1) Edge–Cloud Hierarchical Intelligence:

To reconcile stringent latency constraints with model complexity, future protection architectures are expected to adopt hierarchical intelligence. Ultra-low-latency models, such as quantized CNNs, will operate at the edge for sub-cycle primary protection, while substation-level servers and cloud platforms handle wide-area coordination, learning, and long-term optimization (Mazumder et al., 2023; Zhang et al., 2026).

2) Unified Resilience Frameworks:

Protection, stability assessment, and automated system restoration are expected to converge into integrated cognitive frameworks. Such unified architectures enable coordinated decision-making across the entire disturbance life cycle from fault detection and isolation to system recovery and reconfiguration (Caldas et al., 2020; Choudhary et al., 2025; Sicard et al., 2024).

3) Advanced Computing Paradigms:

Although still at an early stage, emerging computing paradigms, including hardware accelerators and non-classical optimization methods, are being explored to address the growing computational demands of real-time, large-scale AI-based protection (Mazumder et al., 2023; Sicard et al., 2024).

C. Targeted High-Impact Application Domains

- **Inverter-Dominated Power Systems:**

As inverter-based resource penetration increases, AI-based protection must exploit non-traditional fault signatures, such as voltage phase angle jumps, frequency rate-of-change, and harmonic distortion, to ensure reliable operation under limited fault-current conditions (Aziz et al., 2025; Dinneweth et al., 2022; Mishra et al., 2025).

- **HVDC and MVDC Grids:**

Multi-terminal DC networks impose ultra-fast protection requirements that challenge conventional schemes. AI-driven approaches can enhance the coordination of hybrid DC breakers and accelerate fault location and isolation in converter-dominated DC grids (Alsaiari & Ilyas, 2025; Choudhary et al., 2025; Henao et al., 2025).

- **Adversarially Resilient Protection:**

Future AI-enabled protection systems must fuse cyber and physical measurements to detect coordinated cyber–physical attacks that simultaneously manipulate grid hardware, sensors, and communication channels (Hafez et al., 2025; Meng et al., 2024).

D. Socio-Technical Enablers

- **Standardization and Benchmarking:**

The development of unified standards and open benchmark datasets is critical for the objective evaluation, certification, and large-scale deployment of AI-enabled Intelligent Electronic Devices (IEDs) (Khaw et al., 2021; Mazumder et al., 2025).

- **Cross-Disciplinary Education and Workforce Development:**

Bridging the gap between power system engineering and data science is essential for sustainable adoption. Cross-disciplinary training programs are needed to cultivate expertise in both domains and to foster trust in AI-assisted protection decisions (Lotfifard et al., 2009; Sidhu & Gangadharan, 2005).

5. CONCLUSION

Artificial Intelligence (AI) is fundamentally reshaping the landscape of power system protection, facilitating a transition from static, threshold-based logic to adaptive, data-driven paradigms. This review has systematically analyzed this evolution by examining core AI technologies, their transformative applications across transmission and distribution networks, and the critical barriers impeding their practical deployment.

The study highlights that advanced deep learning architectures, specifically CNNs, LSTMs, and GNNs, offer superior fault analysis and coordination capabilities, particularly in grids with high renewable energy penetration and inverter-based resources. However, the transition from research prototypes to mission-critical infrastructure is not without challenges. Issues such as data scarcity, the "black-box" nature of complex models, cybersecurity vulnerabilities, and stringent sub-cycle real-time constraints remain significant hurdles for industry-wide adoption.

These challenges define a clear research agenda for the coming years. Future progress hinges on the development of next-generation AI paradigms, including Causal AI integrated with Digital Twins, Federated Learning for privacy-preserving collaboration, and Neuro-Symbolic AI for enhanced explainability. Addressing these gaps requires a concerted global effort to accelerate standardization (e.g., IEEE P2805), foster open-innovation ecosystems with shared benchmarks, and bridge the skills gap through cross-disciplinary engineering education.

Ultimately, AI is not merely a peripheral upgrade but a foundational shift in protection philosophy. Its thoughtful and standardized integration is essential for securing the resilient, reliable, and self-healing power grids required for a sustainable and decarbonized energy future.

References

Alhamrouni, I., Kahar, N. H. A., Salem, M., Swadi, M., Zahroui, Y., Kadhim, D. J., & Nazari, M. A. (2024). A comprehensive review on the role of artificial intelligence in power system stability, control, and protection. *Appl. Sci., 14*(14), 6214. https://doi.org/10.3390/app14146214.

Alsaiari, A., & Ilyas, M. (2025). A hybrid CNN-LSTM deep learning model for intrusion detection in smart grid. *arXiv preprint*. https://doi.org/10.48550/arXiv.2509.07208.

Aziz, M. Z. Yousaf, R. Fu, W. Khan, U. Siddique, M. Ahmad, & Zaitsev, I. (2025). Advanced AI-driven techniques for fault and transient analysis in high-voltage power systems. *Sci. Rep., 15*(1), 5592. https://doi.org/10.1038/s41598-025-90055-7.

Bakkar, M., Bogarra, S., Córcoles, F., Aboelhassan,V., Wang, S., & Iglesias. J. (2022). Artificial intelligence-based protection for smart grids. *Energies, 15*(13), 4933. https://doi.org/10.3390/en15134933.

Caldas, R. D., Rodrigues, A., Gil, E. B., Rodrigues, G. N., Vogel, T., & Pelliccione, P. (2020, Jun). A hybrid approach combining control theory and AI for self-adaptive systems. In *Proc. IEEE/ACM SEAMS* (pp. 9–19). https://doi.org/10.1145/3387939.3391595.

Chinnaraju, A. (2025). Explainable AI (XAI) for trustworthy and transparent decision-making. *World J. Adv. Eng. Technol. Sci., 14*(3), 170–207. https://doi.org/10.30574/wjaets.2025.14.3.0106.

Choudhary, M., Vijitha, S., Bhavani, D. D., Bhuvaneswari, N., Tiwari, M., & Subburam, S. (2025). Edge AI: Deploying artificial intelligence models on edge devices. *ITM Web Conf., 76*, 01009. https://doi.org/10.1051/itmconf/20257601009

Dasand, S., & Panigrahi, B. K. (2022). A PMU-based data-driven approach for enhancing situational awareness. *IEEE Trans. Ind. Informat., 18*(7), 4773–4784. https://doi.org/10.1109/TII.2022.3147794.

Diahovchenko, I., Kolcun, M., Čonka, Z., Savkiv, V., & Mykhailyshyn, R. (2020). Progress and challenges in smart grids. *Iran. J. Sci. Technol. Trans. Electr. Eng., 44*(4), 1319–1333. https://doi.org/10.1007/s40998-020-00322-8

Dinneweth, J., et al. (2022). *Multi-agent reinforcement learning for autonomous vehicles: A survey*. https://doi.org/10.1007/s43684-022-00045-z.

Grotto, A. J., & Dempsey, J. (2021). Vulnerability disclosure and management for AI/ML systems: A working paper with policy recommendations. *ML Systems Working Paper*. http://dx.doi.org/10.2139/ssrn.3964084

Hafez, S., Alkhedher, M., Ramadan, M., Gad, A., Alhalabi, M., Yaghi, M., & Ghazal, M. (2025). Advancements in grid resilience: Recent innovations in AI-driven solutions. *Results Eng., 26*, 105042. https://doi.org/10.1016/j.rineng.2025.105042.

He, J., & Zhao, H. (2020, December). Fault diagnosis and location based on graph neural network in telecom networks. In *Proc. Int. Conf. Networking and Network Applications (NaNA)* (pp. 304–309). https://doi.org/10.1109/NaNA51271.2020.00059.

Henao, F., Edgell, R., Sharma, A., & Olney, J. (2025). AI in power systems: A systematic review of key matters of concern. *Energy Inform, 8*(1), 76. https://doi.org/10.1186/s42162-025-00529-1.

Hijazi, M., et al. (2023). Transfer learning for transient stability predictions. *IEEE Trans. Autom. Sci. Eng*. https://doi.org/10.1109/TASE.2023.3277536.

Idrisov, N., et al. (2025). Leveraging digital twin and machine learning techniques for anomaly detection. https://doi.org/10.48550/arXiv.2501.13474.

Jamshidi Gahrouei, A., Falah, M., Azarbad, H., & Khorshidpour, S. (2026). Designing smart contract on a blockchain network for the purpose of trading energy from photovoltaic power plants. *International Journal of Smart Electrical Engineering, 14*(3), 171-178. https://doi.org/10.82234/IJSEE.2025.1208414.

Khaw, Y. M., Jahromi, A. A., Arani, M. F., Sanner, S., Kundur, D., & Kassouf, M.(2021). A deep learning-based cyberattack detection system. *IEEE Trans. Smart Grid, 12*(3), 2554–2565. https://doi.org/10.1109/TSG.2020.3040361.

Krause, T., et al. (2021). Cybersecurity in power grids. *Sensors*. https://doi.org/10.48550/arXiv.2105.00013

Livani, H., & Evrenosoglu, C. Y. (2013). A machine learning and wavelet-based fault location method. *IEEE Trans. Smart Grid*. https://doi.org/10.1109/TSG.2013.2260421.

Lotfifard, S., Faiz, J., & Kezunovic, M. (2009). Detection of symmetrical faults during power swings. *IEEE Trans. Power Delivery*,. https://doi.org/10.1109/TPWRD.2009.2035224.

Machlev, R., Heistrene, L., Perl, M., Levy, K. Y., Belikov, J., Mannor, S., & Levron, Y.(2022). Explainable artificial intelligence techniques for energy systems. *Energy AI, 9*, 100169. https://doi.org/10.1016/j.egyai.2022.100169.

Mazumder, M., Banbury, C., & Reddi, V. J. (2023). Data-centric AI benchmarks and evaluation. *IEEE Micro, 43*(2), 12–20, 2023. https://doi.org/10.48550/arXiv.2207.10062.

Mazumder, M., Banbury, C., Yao, X., Karlaš, B., Rojas, W. G., Diamos, S., & Reddi, V. J. (2023). Dataperf: Benchmarks for data-centric AI development. *Adv. Neural Inf. Process. Syst., 36*, 5320–5347. https://doi.org/10.48550/arXiv.2207.10062

Meloni, A., Pegoraro, P. A., Atzori, L., Benigni, A.,& Sulis, S. (2018). Cloud-based IoT solution for state estimation in smart grids. *Comput. Netw., 130*, 156–165. https://doi.org/10.1109/ICEPES60647.2024.10653566.

Meng, Q., Hussain, S., Luo, F., Wang, Z., & Jin, X. (2024). An online reinforcement learning-based energy management strategy for microgrids. *IEEE Trans. Ind. Appl.* https://doi.org/10.1016/j.est.2024.115114.

Mishra, M., & Singh, J. G. (2025). A comprehensive review on deep learning techniques in power system protection: Trends, challenges, applications and future directions. *Results Eng*., 103884. https://doi.org/10.1016/j.rineng.2025.107863.

Nadal, I. V., et al. (2025). Physics-informed neural networks in power system dynamics. https://doi.org/10.48550/arXiv.2501.17621.

Oelhaf, J., Kordowich, G., Pashaei, M., Bergler, C., Maier, A., Jäger, J., & Bayer, S. (2025). A scoping review of machine learning applications in power system protection and disturbance management. *Int. J. Electr. Power Energy Syst., 172*, 111257. https://doi.org/10.48550/arXiv.2509.09053.

Rizzato, M., et al. (2022). Stress testing electrical grids using GANs. *Energy AI*. https://doi.org/10.1016/j.egyai.2022.100177.

Sidhu, T. S., & Gangadharan, P. K. (2005). Control and automation of power system substation using IEC 61850. *Proc. IEEE CCA*, 1331–1336. https://doi.org/10.1109/CCA.2005.1507316.

Wen, M., Xie, R., Lu, K., Wang, L., & Zhang, K. (2021). FedDetect: A novel privacy-preserving federated learning framework for energy theft detection in smart grid. *IEEE Internet Things J., 9*(8), 6069–6080. https://doi.org/10.1109/JIOT.2021.3110784.

Zhang, J., et al. (2026). A characteristic oriented probabilistic stability assessment method based on PINN. https://doi.org/ 10.1016/j.ijepes.2025.111504.

Zhang, W., Yang, D., & Wang, H. (2019). Data-driven methods for predictive maintenance: A survey. *IEEE Syst. J.*, *13*(3), 2213–2227. https://doi.org/10.1109/JSYST.2019.2905565.