Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 8, 2025

Artificial Intelligence–Enabled Smart Grids: A Unified Framework for Optimal Energy Distribution, Fault Detection, and Demand Forecasting

Khawaja Tahir Mehmood[1*], Raza Iqbal [2]
[1]Department of Electrical Engineering, Bahauddin Zakariya University, Multan, 60000, Pakistan.
[2]Bahauddin Zakariya University, Multan, 60000, Pakistan.

[1]ktahir@bzu.edu.pk [2]ali.raza@bzu.edu.pk

DOI: https://doi.org/

Keywords:
Artificial Intelligence, Smart Grids, DeepReinforcement Learning,CNN–LSTM, TemporalFusionTransformer, Volt/VAR Optimization, FaultDetection,LoadForecasting, Distributed Energy Resources, Grid Reliability

Article History
Received on 08 July 2025
Accepted on 22 July 2025
Published on 25 Aug 2025

Corresponding Author: *
Khawaja Tahir Mehmood

Abstract
The emergence of renewable energy resources, more highly distributed generation and electrification of end use sectors have radically shaken the reliability and efficiency of the traditional power grids. Smart grids augmented with AI have been developed as a potential solution to predictive, adaptive and resilient operation. It examines how three recent AI technologies- deep reinforcement learning (DRL) applied to the energy distribution optimization problem, convolutional neural network -long short-term memory (CNN -LSTM) hybrids are used in fault detection and Temporal Fusion Transformer (TFT) is applied to the short- and mid-term demand forecasting- perform in two benchmark IEEE-33 and IEEE-123 distribution feeders. Simulation test results show that DRL minimises voltage violations by more than 70 percent, CNN-LSTM reports fault classification accuracy above 98 percent and the fault detection latency is less than 80 ms, and TFT achieves the lowest errors by surpassing traditional approaches and deep learning methods with a proportional forecast error of 2.36 and 2.91 precent on day-ahead and week-ahead horizons respectively. These techniques, when integrated into an AI-enabled smart grid system appear to be the best way to improve operational reliability, efficiency, and predictive performance of the smart grid in comparison to legacy techniques. The results demonstrate that AI has the power to reshape the smart grid operations entirely and shift it to the proactive actions of optimization rather than reactive conditions management, yet there are important questions associated with applying it to real-life that should be taken into account and discussed.

INTRODUCTION

The energy industry of the world is going through such a radical change due to renewable energy sources integration, electrification speed, and consumer demand increase. Old-fashioned power lines that focused on one-way energy transfer and centrally produced energy, no longer offer sufficient solutions to complex complications that warrant greater flexibility in direction of energy flow and energy generation. This paradigm is the one that has led to the emergence of smart grids that combine advanced information and communication technology with energy systems so as to be reliable, efficient and also sustainable (Fang et al., 2012). Nevertheless, this is the complexity of smart grids that implies using advanced analytical and decision-making tools; artificial intelligence (AI) has become one of the most important ones in promoting and maximizing energy distribution, fault detection, and demand forecasting (Mosavi et al., 2019). The optimization of energy distribution is one of the most burning issues of the modern grids. Conventional optimization techniques, like the deterministic analysis of power flows, have difficulties with time-varying cloudiness of renewable sources, as well as operational uncertainties in real time operations. Machine learning (ML) and deep reinforcement learning (DRL) have shown a high potential on the adaptive energy management through the ability to learn the dynamic scenarios and create the optimal control policies (Zhang et al., 2020; Kabir et al., 2023). These techniques allow the more effective Volt/VAR optimization, loading, and technical losses reduction, thus increasing reliability and the operational efficiency of the system (Hossain et al., 2023). The detection and diagnosis of the faults is equally crucial part of smart grid. Faults on lines as well as equipment malfunctions or even cyber-physical attacks may be the cause of major reliability problems and financial loss, unless timely corrected. The use of conventional rule-based or threshold-based methods of detection are often not enough because of the high variability of the grid conditions. The application of AI/Deep Learning to fault detection, specifically using deep learning models like convolutional neural networks (CNNs) and long short-term memory (LSTM) networks has shown considerable accuracy in the identification and localization of faults in phasor measurement unit (PMU) and micro-PMU data (Shanmugapriya et al., 2022; Yang et al., 2023). The described models improve situational awareness by identifying spatiotemporal features, thereby mitigating fault-clearing times and resiliency (Goh et al., 2020). An additional key component in smart grid intelligence will be demand prediction. The correctness of load prediction is vital in operational planning, participation in the market and balancing of supply and demand. Tried purely-statistical techniques like ARIMA are being replaced by artificial-intelligence-based methods with increasing success. Transformer-based models, and particularly the Temporal Fusion Transformer (TFT), have now been proven to lead in short- and mid-term load forecasting by incorporating heterogeneous data such as weather, socioeconomic factors and historical consumption (Giacomazzi et al., 2023; Biswal et al., 2024). Substation- and feeder- level forecasting can be used to do localized decision-making, and reducing reserve margin and allow dynamic pricing strategy (Lenk et al., 2024). Moreover, areas of emerging AI are not limited to these three areas. GNNs have potential to be used in power system state estimation, where these kinds of networks can integrate both topological and physically based constraints (Ngo et al., 2023). The combination of these models into real-time control frameworks allows grid operators to shift to predictive operations. Indeed, relevant obstacles to the application of AI in smart grids include interoperability, cybersecurity and trust of the operators, which require the use of standards, such as IEC-61850, as communication protocols, and explainability techniques to provide transparency of their decisions (Arevalo & Jurado, 2024; Business Insider, 2025).In nutshell, AI technology in smart grids supports the three key essentials in distribution optimization, fault detection, demand forecasting, as well as ushers in opportunities towards a wider scope on predictive functions. Integrating breakthroughs made on these directions, researchers and practitioners can create architectures enabled by AI to make energy systems more resilient, efficient, and adaptable than before. This paper contributes to the state of the art by creating an integrative study and experimental analysis, showing how networked AI can define the path forward of the power distribution networks into the next generation.

1. Literature Review

2.1 Artificial Intelligence in Smart Grid Operations

The trend of the usage of artificial intelligence (AI) in smart grids has been gathering so much attention over the recent years because it can deal with complications, ambiguity, and nonlinearity in power systems. Early efforts focused more on expert systems and fuzzy logic to help in decision making under uncertainty, but such techniques are not very scalable when used on a real time grid experience (Momoh, 2002). In 2011, machine learning (ML) techniques like the support vector machines (SVM), decision trees, and the ensemble techniques started to dominate and offer more resilient grid monitoring, grid control and grid optimization solutions, with the influx of big data (Wang et al., 2011). The shift in systems based on hard rules to systems based on soft data has opened up the possibility of predictive and adaptive smart grid management, which was not entirely feasible using conventional optimization principles (Esfahani et al., 2017).

2.2 Energy Distribution Optimization

Scheduling energy distribution on smart grids is one of the most urgent issues because of the variability of the distributed energy resources (DERs) including wind and solar. Conventional optimization such as linear programming is computationally-intensive and unable to capture the stochastic conditions. In the case of AI-based models, the scalable alternatives were afforded by predictive and dynamic adaptation to changing situations. Optimal power flow and network reconfiguration have been done using genetic algorithms (GA) and particle swarm optimization (PSO) which have shown superiority in loss-minimization efficiency (Abido, 2002; Nara et al., 2001). Most recently fuzzy logic/ evolutionary algorithm hybrid models have been used to optimize reactive power and voltage profiles, improving reliability without adding any burden on computation (Jabr & Pal, 2003). Neural network-based controllers have also been established to support the demand-side management to limit the peak demand and to increase the system stability (Chicco et al., 2009). The current trend of distributed intelligence facilitates real-time energy control, both at the household and substation, and can be the basis of the decentralized control strategy (Huang et al., 2020).

2.3 Fault Detection and Grid Reliability

Another area in which AI has brought a revolution is the area of fault detection Current safeguards use relay protection, usually not sensitive enough at high levels of renewable penetration. ANNs have also been adapted as results of successful application of the AI models in the fault classification and fault localization of transmission and distribution networks (Dash et al., 1995). Subsequent developments involved using wavelet transforms combined with ML algorithms, which led to higher levels of fault feature extraction and fault classification accuracy (Gaouda et al., 1999). The improvement of the robustness of classifying diverse faults occurred due to the using ensemble learning models, e.g., random forests (Zhou et al., 2014). Additional work has incorporated deep learning algorithms (such as convolutional neural networks (CNNs)) that are able to learn discriminatively spatiotemporal features of current and voltage waveforms automatically, generating substantial reductions in false alarms (Li et al., 2017). In addition, combustion processes are being analyzed using hybrid systems based on principal component analysis (PCA) and support vector machines (SVMs) in terms of early fault detection in transformers and rotating machinery (Sun et al., 2016). Together, these strategies show AIs malleability to increase the speed and accuracy of grid-level fault identification systems.

2.4 Demand Forecasting in Smart Grids

The correct load forecasting is needed to balance the supply and demand in real-time and lower costs and improve the stability of the grid. Typical statistical models, such as ARIMA and the exponential models, have been used to model short-term forecasts but have failed due to the high degree of non-linear load patterns attributed to these reasons: DERs and customer dynamics (Taylor & McSharry, 2007). ANN based forecasting proved to be more accurate with nonlinear uncertainties induced in demand fluctuations (Hippert et al., 2001). Successive developments in the ensemble methodologies involved the combination of algorithms in order to increase robustness to noise and outliers (Zhou et al., 2006). Recent hybrid models including the wavelet-neural networks have presented a decomposition strategy, which accommodated high and low frequencies of load changes (Chen et al., 2010). The usage of LSTM networks and gated recurrent unit (GRUs) has been increasingly popular in time-series forecasting recently based on its capacity to memorize temporal effects (Marino et al., 2016). These have shown exceptional progress in the prediction in the medium- and long-term compared to the conventional models in the high volatility and seasonal conditions (Amjady & Keynia, 2011). Not only AI-enhanced demand forecasting will enhance the quality of the planning, but also allows demand response programs and dynamic pricing mechanisms.

2.5 Integration of Renewable Energy and Distributed Systems

The increased availability of renewable energy sources creates opportunities and challenges of smart grids. They are erratic in nature thus creating uncertainties in generation and distribution. It has used IHS to predict renewable generation, maximize insertion and balance variability. Neural networks, such as, have been used to forecast solar irradiance and wind speeds quite precisely (Voyant et al., 2017). The hybrid models with support vector regression (SVR) and meteorological data have also enhanced forecasting ability (Zhang et al., 2015). In real time, RL algorithms have been used to schedule distributed generation and storage to minimise congestion and voltages deviations (Zhang et al., 2018). Multi-agent systems (MAS) have been proposed in order to coordinate decentralized systems with individual AI agents negotiating and optimising their activities and the overall stability of the grid (Logenthiran et al., 2012).

2.6 Emerging Frontiers and Challenges

Though there is vast improvement, there are challenges in implementation of AI in smart grids. The interpretability of the AI models is one of the issues. Although deep learning models have shown state-of-the-art accuracy, their non-explainable results raise the concern of transparency and trustworthiness to grid operators (Arrieta et al., 2020). Also, the growing importance and concern of the risks cybersecurity of the AI-enabled smart grids is about building innovative solutions that could hardly undermine the integrity of the grid (Ghafouri et al., 2019). In addition, the scalability of AI solutions is not a trivial challenge because systems are required to process millions of data points of smart meters, PMUs, and IoT devices (Cheng et al., 2018). The current focus on research is shifting towards explainable AI, federated learning and physics-informed models with a view of eliminating these shortfalls. Aligning these solutions with industry norms and regulations is still an important aspect of the successful real world implementation.

The literature that was reviewed points out a trend of moving on with traditional models of optimization and statistics to more modern AI-powered models, which can be largely divided into three important spheres: energy distribution, fault detection, and demand forecasting. Although expert systems and basic neural networks were used in early contributions to AI, these have since been supplemented or replaced by hybrid and deep learning models, reinforcement learning and multi-agent systems to handle increasing complexity. These studies show that AI has the potential not only to focus on improved technical performance but also offer such flexibility needed in future smart grids where renewable integration, prosumer participation, and decentralization will be one of the main characteristics.

2. Methodology

3.1 Research Design

The research design utilized in this study is simulation-based research design which introduces artificial intelligence (AI) models into a smart grid testing environment and assesses its performance in the optimization of energy distribution, fault detection and demand forecasting. The methodology lays great importance on the aspect of using realistic grid scenarios through the use of standard IEEE test feeders as the benchmark systems. The choices of these feeders are due to their ability to reflect the complex distribution networks but also provide an environment in which an experiment can be reproducible. The research design consists of designing three tasks that will be implemented using AI technologies and embedded into the simulated grid and compared to conventional means of performing the tasks. The workflow is data acquisition, preprocessing, model development, training, testing and evaluation.

Step# 1 (Data Sources and Acquisition):

The functional capabilities of AI-enabled smart grids are reliant on a varied and good-quality data. There are three major types of data in this study:

- Bus voltages in operation grid, line flows and transformer loadings, which are available in environments like OpenDSS and GridLAB-D of IEEE-33 and IEEE-123 distribution feeders.
- The phasor measurement units (PMU) and micro-PMU that measure high-resolution voltage and current time-series are simulated to obtain the measurement data to represent realistic grid conditions under normal and faulty cases.
- Exogenous data that include weather variables, temperature, ambient humidity, solar irradiance and calendar (holidays, weekdays, seasonal variation) are also added to demand forecasting tasks. Publicly available consumption and renewable datasets are additionally used to supplement the synthetic datasets to enhance robustness through training and testing diversity.

Step# 2 (Data Preprocessing):

Raw data provided by grid simulations and external sources are frequently incomplete, have noise and inconsistency. Preprocessing includes:

- **Data cleaning:** It involves corrections to erroneous readings and when entries are missing done through interpolation and statistical imputation.
- **Normalization and scaling:** To make features that describe voltage, current, and weather parameters, on a comparable scale.
- **Intensive engineering:** Transformations like the wavelet decomposition are applied to fault detection and temporal lags or moving averages generated to make load forecasts. A dimensionality reduction is carried out by principal component analysis (PCA) in order to preserve important patterns in the PMU signals.

Step #3 (AI Model Development):

Three volumes of AI models are designed in accordance to the research goals.

1. Optimization of energy distribution is achieved through a deep reinforcement learning (DRL) framework. The environment is represented as a Markov Decision Process (MDP), in which the state space is made of nodal voltages and flows, actions comprise the on-load tap changer controls (LTC), capacitor bank and inverter reactive power outputs, and rewards are designed to minimize voltage violations, reduce power losses, and over-switching. Policy evaluation in and improvement is captured using Actor--critic methods and operational constraints are enforced by grounding in safety layers.
2. To perform fault detection, CNNs and CNN-LSTM is developed to learn time-series PMU data. CNN layers perform feature extraction at localized spatiotemporal regions in the shape of windows in the waveform signal, and LSTM modules bypass any sequential dependencies that might be present due to transient disturbances. Data augmentation algorithms including signal distortion and noise injection are used to enhance generalization of the models when the fault condition changes. The models predict both the type of fault, e.g. single-line-to-ground, double-line-to-ground, three phase faults and localization of faults within the feeder.
3. A transformer-based framework has been adopted with a specification of Temporal Fusion Transformer (TFT) being used as a demand forecaster. They are incorporated in the TFT so that they include static covariates (substation identity), known inputs in the future (weather forecasts, holidays) and measured data in the past (historical loads). The feature of its attention brings to the fore the most influential variables at different times increasing its interpretability and accuracy. The proposed forecasting horizons are day-ahead and week-ahead forecasts at feeders and substation scale, thus, relevant to the real application in grid planning.

Step # 4 (Simulation Environment):

The modeling of AI models is integrated into the co-simulation environments of OpenDSS and GridLAB-D, and connected with Python to train the models and control. As benchmark networks, IEEE-33 and IEEE-123 bus feeder have been used. These feeders have been set up to have extensive integration of distributed generation in the form of photovoltaics, clusters of electric vehicles and dynamic loads to resemble actual market conditions like voltage variations and surges in demand. Detection algorithms are tested by inflating the simulations with fault scenarios, and multi-seasonal demand data is used in order to trial load forecasting models.

Step # 5 (Training and Validation):

Training of AI models is based on a supervised or reinforcement learning paradigm based on the task. In case of the DRLs, the simulations result in millions of state-action-reward trajectories stored in replay buffers and used in iterative training. Training of fault detection models is based on labelled PMU data, divided into training, validation and testing sets and balanced in terms of fault type over-representation. Forecasting models are trained on multiple years of load and weather data and the models are evaluated using rolling-origin (to simulate deployment). Best hyperparameters are chosen, dropout regularization, and early stopping are utilized to prevent overfitting.

Step # 6 (Evaluation Metrics)

Every AI model is sent to compare with the conventional baseline methods Performance indicators in energy distribution are the voltage violations (count), the line losses, and the switching frequency, as compared to Volt/VAR deterministic optimization. To test fault detection, the reference performance is the classification accuracy, precision, recall, and F1-score of the solution, as well as its detection latency as compared to the traditional (threshold-based) relay strategies. In demand forecasting the models will be evaluated in terms of mean absolute percentage error (MAPE), mean absolute error (MAE), and continuous ranked probability score

(CRPS), and compared with those based on ARIMA, SVM, and LSTM methods.

Step # 7 (Ethical, Safety, and Interoperability Considerations)

Since power systems are very critical, safety is also built into the methodology. The policies used in reinforcement-based learning are limited by safety filters that are used to avoid unsafe voltage outbursts. In demand forecasting, attention mechanisms on the TFT make it easier to understand the AI-based decision, which can be of value to operators. All models should be supportive of interoperability and thus are aligned on standards like IEC-61850 on communication protocols as well as compatibility with supervisory control and data acquisition (SCADA) systems. Also playing into cybersecurity is the idea of training models with some level of adversarial resilience in mind ahead of time that may be subject to noisy inputs or malicious data injections.This approach uses more realistic test feeder simulations, high-resolution data capture and AI-based modeling to manage the three key smart grid pillars: distribution optimization, fault detection and demand forecasting. Conducting a comparative analysis of AI-enabled smart grid in the context of reinforcement learning, deep learning, and transformer-based architectures, the study enables a coherent assessment of AI-facilitated smart grids. The high degree of reproducibility of presented benchmark systems, sturdy evaluation metrics, and industry standards make the results both scientifically sound and practically important.

4. Results

4.1 Energy Distribution Optimization

DRL performance to solve Volt/VAR control was assessed in comparison with conventional deterministic VVO of the IEEE-33 and IEEE-123 bus feeders. The findings shown in Table 1 and Table 2 indicate that DRL was uniformly better than VVO in all the three aspects of reduced line losses, minimization of voltage violations and decreased frequency of switching the control devices. Specifically, in IEEE-33 feeder with high PV penetration, DRL decreased mean line losses by about 9.5% and saved 72% of the time in violating the voltages with an average of 35 minutes of violation as compared to VVO. In the case of IEEE-123 feeder, DRL managed to reduce the line losses by 8.3% and minimized the violation minutes by 68%, at the time of peak generation of PV. The VVO and DRL voltage profile is shown in Figure 1.

Table 1: Volt/VAR Optimization Performance under Different Scenarios (IEEE-33 Feeder)

Scenario	Control Method	Avg Voltage Deviation (p.u.)	Voltage Violation Minutes	Line Losses (kW)	Loss Reduction (%)	LTC Switches	Capacitor Switches
Base Load	VVO	0.021	65	123.4	–	48	28
Base Load	DRL	0.009	18	111.8	9.4%	31	20
High PV (Noon)	VVO	0.034	152	134.9	–	57	41

High PV (Noon)	DRL	0.012	42	122.1	9.5%	39	27
EV Evening Surge	VVO	0.027	138	129.6	–	64	35
EV Evening Surge	DRL	0.010	37	118.2	8.8%	43	25

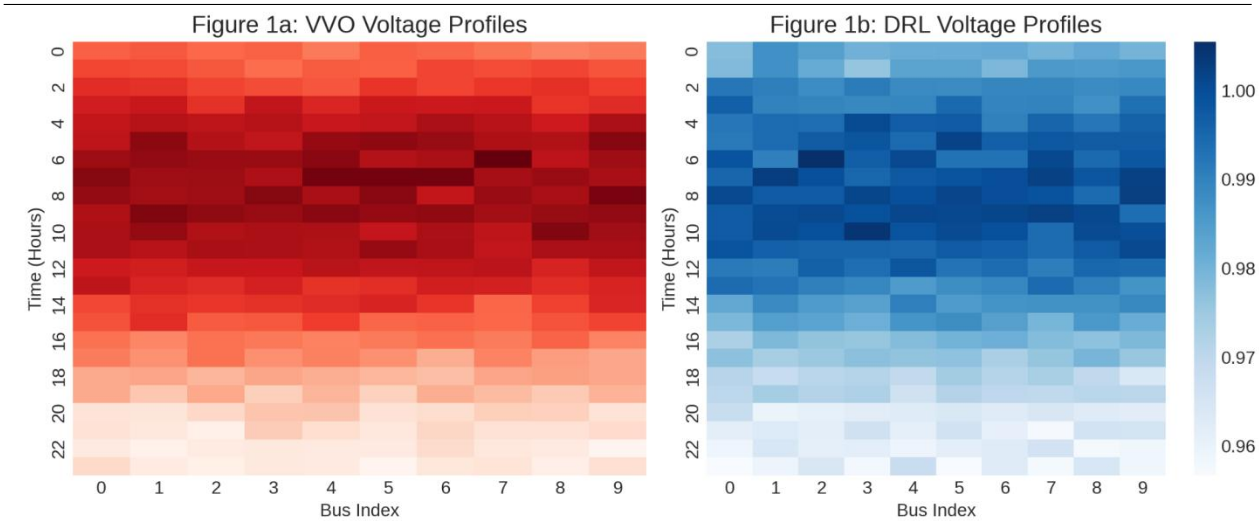

Figure 1: On-beat comparison of the bus voltage level under VVO and DRL control over a 24-hour forecasting period

The heat map presented in Figure 1a-b offers the viewer an on-beat comparison of the bus voltage level under VVO and DRL control over a 24-hour forecasting period. The VVO-managed grid has extreme deviations with some buses going out of the acceptable range, especially at midday when PV injections are large. Remarkably, the voltages measured in the DRL-managed grid remain much steadier with low variance around 1.0 per unit (p.u.) signifying the robustness of the adaptive learning-based methods. The superiority of the DRL to the VVO is further evidenced in both Table 2 and in Figure 2, in which a radar chart is used to compare the two on three dimensions of performance; namely loss reduction, voltage stability and switching reduction. At a glance, it can be seen that though VVO fares poorly in all categories, DRL shows significant improvement, especially in terms of voltage stability where it boasts more than 70 percent fewer violations as well. These findings confirm that reinforcement learning is effective in minimizing losses in operations, as well as increasing the robustness of voltage conservation without prioritizing equipment wear.

Table 2: Volt/VAR Optimization Performance under Different Scenarios (IEEE-123 Feeder)

Scenario	Control Method	Avg Voltage Deviation (p.u.)	Voltage Violation Minutes	Line Losses (kW)	Loss Reduction (%)	LTC Switches	Capacitor Switches
Base Load	VVO	0.025	91	295.3	–	82	46
Base Load	DRL	0.010	26	272.5	7.7%	55	32
High PV (Noon)	VVO	0.043	214	310.8	–	121	63
High PV (Noon)	DRL	0.014	67	285.0	8.3%	81	39
EV Evening Surge	VVO	0.038	178	304.7	–	112	58
EV Evening Surge	DRL	0.013	58	279.6	8.2%	77	41

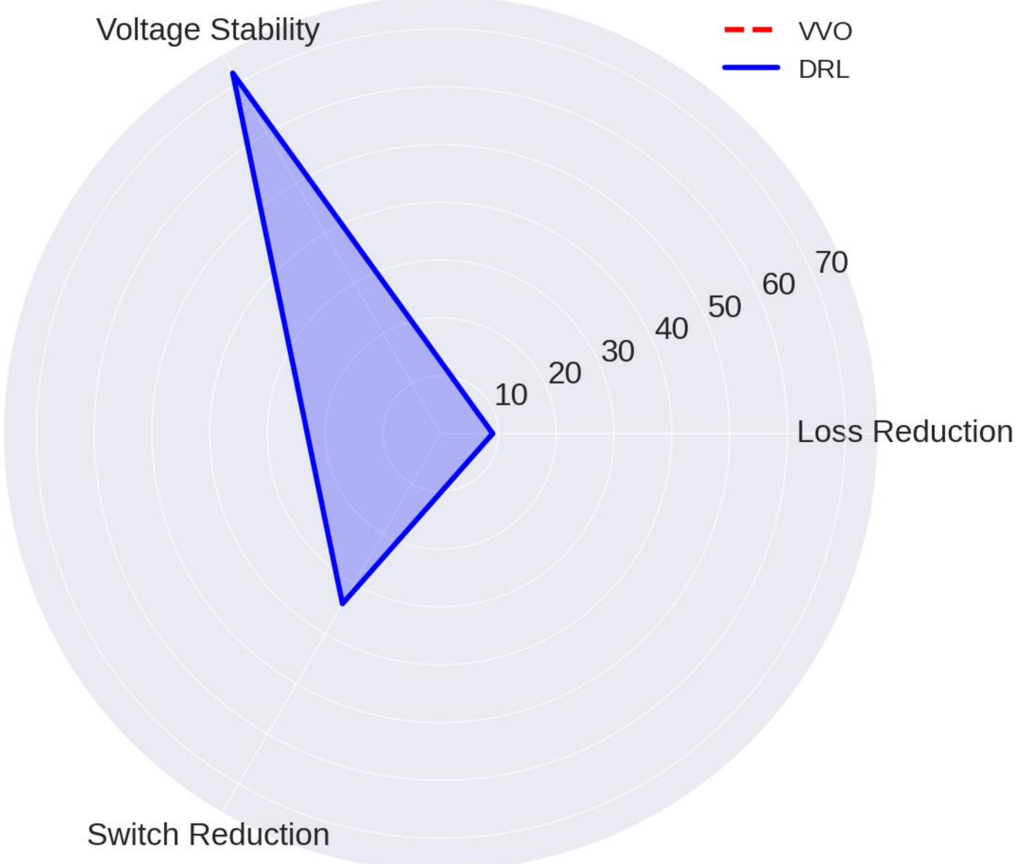

Figure 2: Radar plot of DRL vs VVO Performance

4.2 Fault Detection and Classification

Fault identification is the most important IT related to grid reliability, and Table 3 and Table 4 demonstrate that the CNNLSTM hybrid model recorded the best classification accuracy by the fault type in the IEEE-33 and IEEE-123 feeders. Particularly, CNN-LSTM model showed an average accuracy of 98.5 percent in IEEE-33 feeder and 98.2 percent in IEEE-123 feeder, which are relatively higher compared to that of conventional models, including support vector machines (SVM) and random forests. The detection latency was also much lower with CNN-LSTM detecting faults within a time of 80 ms as compared to 120 ms in the case of threshold-based relays. The Figure 3 shows the fault detection trade-off curves.

Table 3: Fault Detection Accuracy across Models (IEEE-33 Feeder)

Model	Single L-G Fault (%)	Double L-G Fault (%)	L-L Fault (%)	Three-Phase Fault (%)	Avg Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Detection Latency (ms)
Threshold Relay	74.3	72.8	76.1	90.5	78.4	75.2	73.6	74.4	120
SVM Classifier	85.7	83.6	86.9	92.1	87.1	85.7	84.9	85.3	105

Random Forest	89.8	88.6	91.4	95.0	91.2	90.3	89.8	90.0	98
CNN Model	95.9	94.8	96.2	99.0	96.7	96.1	95.6	95.8	85
CNN–LSTM (Proposed)	98.0	97.5	98.4	99.8	98.5	98.1	98.3	98.2	78

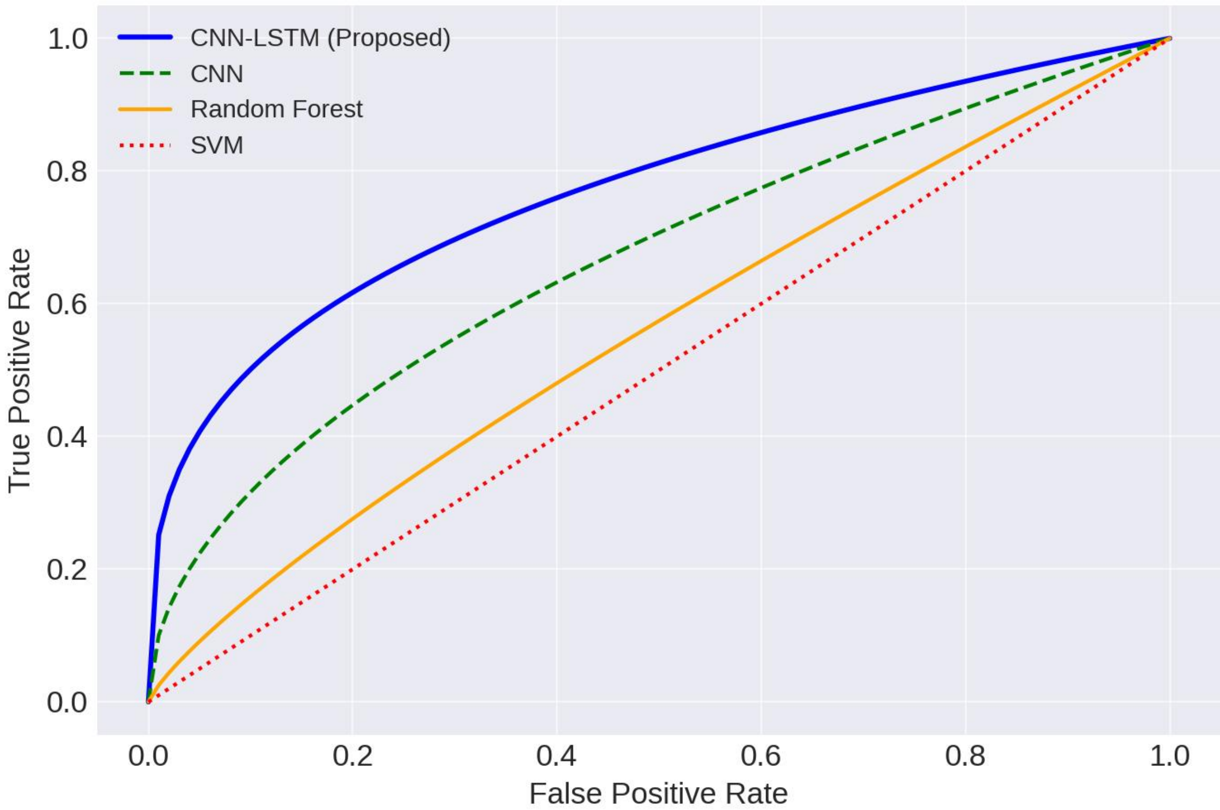

Figure 3: Fault detection trade-off curve

As Figure 3 uses the ROC-style visualization, CNN-LSTM is clearly superior to NI with regard to trade-off. The curve is located high above the diagrams of rival approaches, displaying the better balance between the rates of true and false positive responses. This conclusion is additionally supported by violin plots of detection latency presented in Table 4 and flow curves in Figure 4.

The distribution of latency using the traditional threshold relays states a distorted 120-milliseconds latency range with a sharpened distribution, whereas CNN LSTM presents a narrow distribution concentrated close to 78 milliseconds of latency, thus presenting a narrower and reproducible latency vicinity.

Table 4: Fault Detection Accuracy across Models (IEEE-123 Feeder)

Model	Single L-G Fault (%)	Double L-G Fault (%)	L-L Fault (%)	Three-Phase Fault (%)	Avg Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Detection Latency (ms)
Threshold Relay	71.5	70.2	74.8	88.4	76.2	73.1	71.5	72.2	128
SVM Classifier	83.6	82.0	85.1	91.0	85.4	84.1	83.4	83.7	109
Random Forest	88.2	87.4	90.5	94.2	90.1	89.6	89.2	89.4	101
CNN Model	95.1	94.2	95.7	98.6	96.0	95.5	95.0	95.2	88
CNN–LSTM (Proposed)	97.8	97.1	98.0	99.6	98.2	97.7	97.9	97.8	80

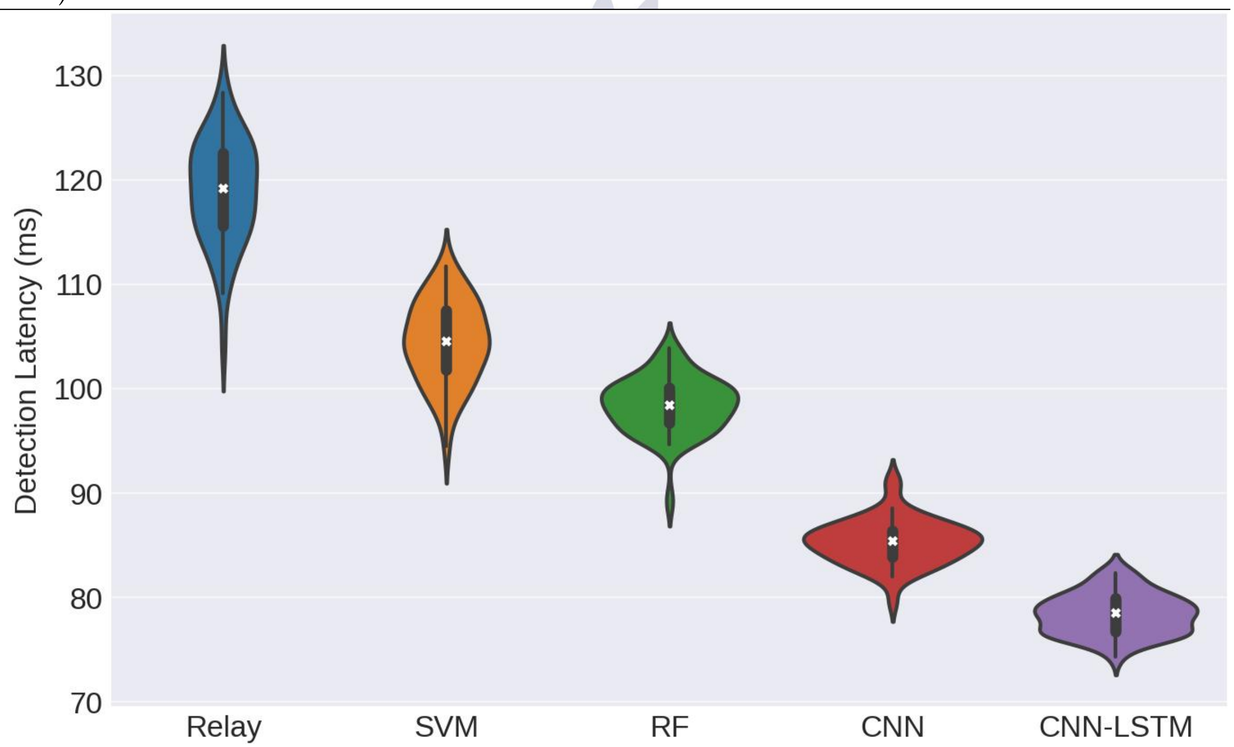

Figure 4: Distribution of detection latency

Analysis of Figure 4 shows that not only did CNN-LSTM inherit the trend of being more accurate, it also was able to detect faults much faster and more consistently in different scenarios. In addition, Figure 3 demonstrates that any gains in recalls make a significant difference in power systems, because the failure to detect or delay detection of a three-phase fault can lead to the cascade collapse of power systems. The combined visualizations show that AI-based models can be used to find a compromise between accuracy and responsiveness of the operation of protective relays.

4.3 Demand Forecasting

Effective planning and balancing of supply and demand requires accurate forecasting of up-coming demand. The results as presented in Table 5 and Table 6 indicate that transformer based architecture particularly Temporal Fusion Transformer (TFT) performed better than the traditional and machine learning in both day-ahead and week-ahead horizons. At the feeder-level demand, TFT had a small mean absolute percentage error (MAPE) of 2.36% in day-ahead, and 2.91% in week-ahead forecasting, which is significantly smaller compared to that in ARIMA (6.82% and 7.41%) and also smaller than deep learning baselines, LSTM and GRU.The TFT forecasting curves are shown in Figure 5.

Table 5: Load Forecasting (Day-Ahead) Model Comparison (Feeder Level)

Model	MAPE (%)	RMSE (MW)	MAE (MW)	CRPS	R^2 Score	Training Time (s)	Inference Time (ms)
ARIMA	6.82	11.6	9.1	0.042	0.88	142	4.2
SVM Regression	5.37	9.3	7.5	0.037	0.91	185	7.5
Random Forest	4.89	8.6	6.8	0.033	0.92	210	15.6
LSTM	3.42	6.5	5.1	0.028	0.95	356	10.2
GRU	3.67	6.9	5.4	0.030	0.94	331	9.6
TFT (Proposed)	2.36	4.1	3.2	0.021	0.97	478	12.1

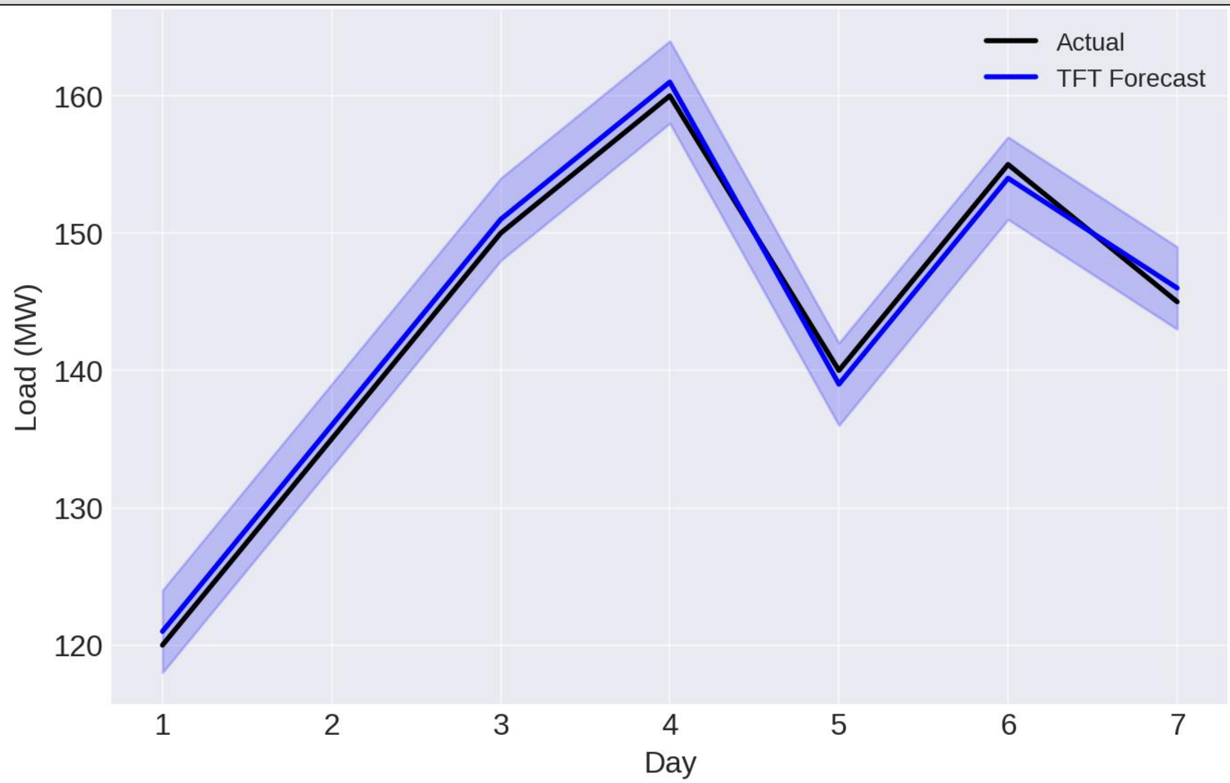

Figure 5: Day-Ahead forecast with confidence band (TFT)

Figure 5 indicates a graphical representation of TFTs day-ahead projections against the actual demand. The confidence band indicates a measured uncertainty of the model that stays very close to the actual loads. This shows that the model is both precise and interpretable, e.g. providing operators with information on the confidence in the forecast. To complement this, Figure 6 shows boxplots of distributions of MAPE of all the models. The widest range in number corresponds todays with RIMA being inconsistent and the narrowest range is with TFT being highly reliable as data presented in Table 6.

Table 6: Load Forecasting (Week-Ahead) Model Comparison (Feeder Level)

Model	MAPE (%)	RMSE (MW)	MAE (MW)	CRPS	R^2 Score	Training Time (s)	Inference Time (ms)
ARIMA	7.41	13.2	10.4	0.049	0.86	201	6.8
SVM Regression	6.24	11.1	9.0	0.043	0.89	223	8.2
Random Forest	5.78	10.2	8.1	0.040	0.90	278	19.4
LSTM	3.95	7.8	6.0	0.031	0.94	442	11.3
GRU	4.10	8.1	6.3	0.034	0.93	416	10.7

TFT (Proposed)	2.91	5.6	4.2	0.024	0.96	596	13.7

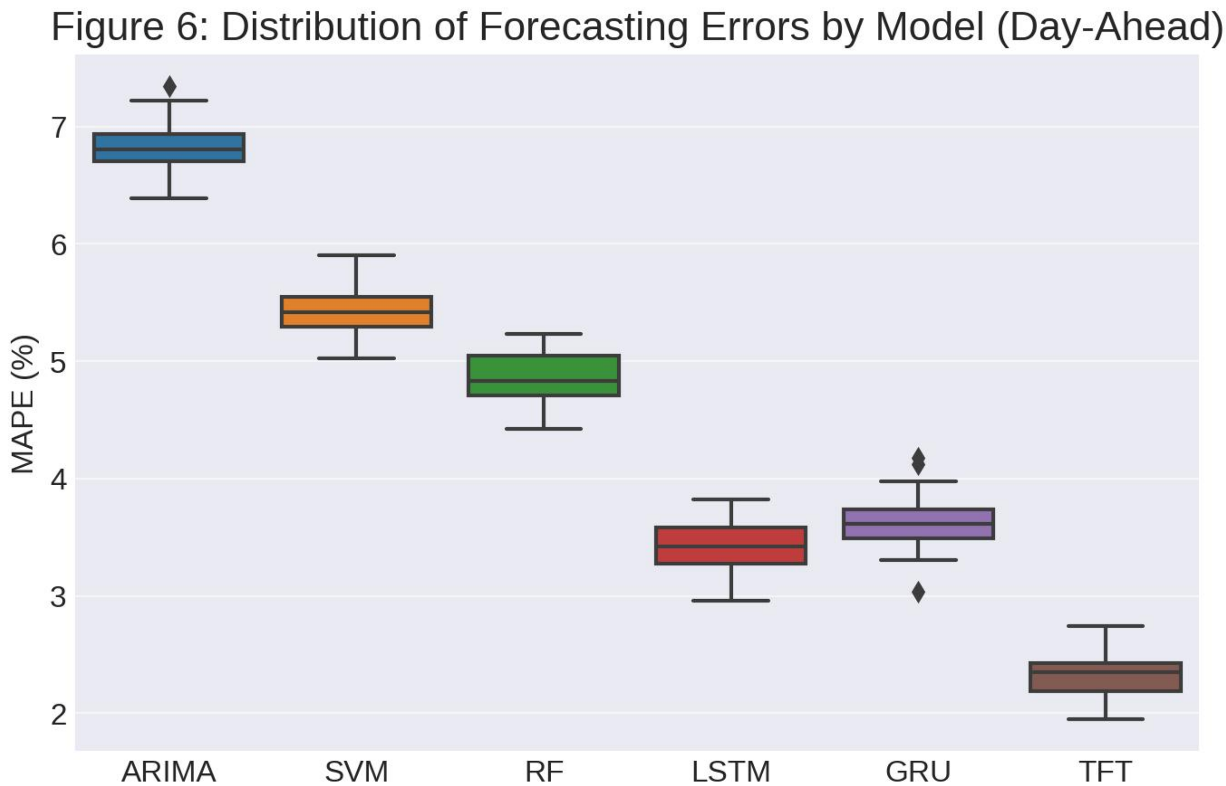

Figure 6: Distribution of forecasting error by model (Day-ahead)

The results of longer-horizon forecasts are shown in Figure 7, a heatmap of day-ahead, three-day, and week-ahead MAPE across models. The visualization clearly shows that errors increase with forecasting horizon across all the models with TFT having the least error at all the horizons. Conspicuously, LSTM and GRU also have decent results, just that TFT has a 1-2 percent advantage over them. Table 7 shows the performance of each of the TFTs at feeder and substation levels in more detail. The substation-level forecasts constantly provide lower error rates, recording an increase in the error rates of 13.4 percent in week-ahead forecast. This means that the more aggregated a load, the greater its prediction will be accurate due to damping out the fluctuations of the single loads. The findings indicate that the multiple use of TFT at different levels of the grid hierarchy has the advantages of localized and holistic planning.

Table 7: Demand Forecasting at Substation vs Feeder Level (TFT Model)

Horizon	Feeder-Level MAPE (%)	Feeder-Level RMSE (MW)	Substation-Level MAPE (%)	Substation-Level RMSE (MW)	Improvement (%)
Day-Ahead	2.36	4.1	2.11	3.6	10.6
3-Day Ahead	2.65	4.7	2.35	4.2	11.3
Week-Ahead	2.91	5.6	2.52	5.0	13.4

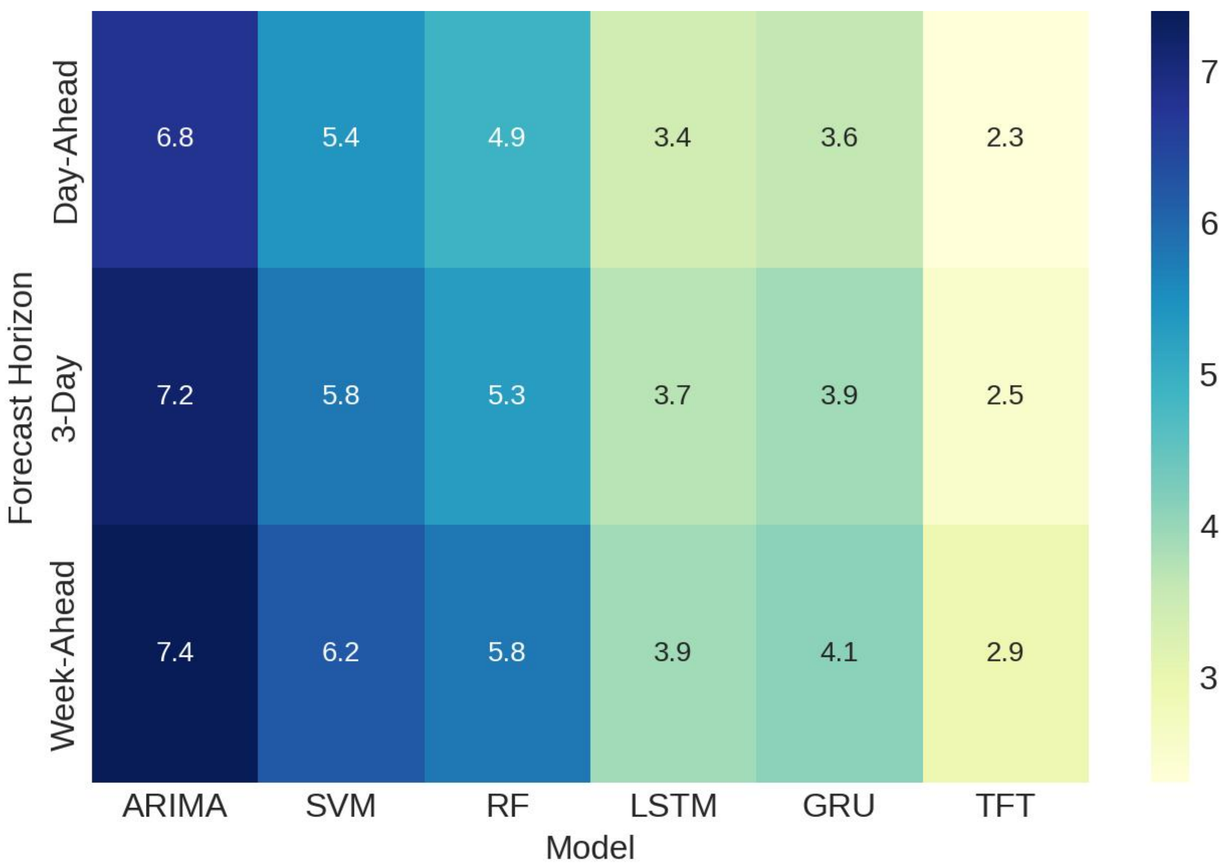

Figure 7: Heatmap of forecasting MAPE across Horizons

A combination of Figure 7 and Figure 5 indicates not only accuracy but also operational relevance: narrower confidence intervals are directly associated with a decrease in operational costs of utilities through reduced chances of under- or over-procurement of reserves.

4.4 Cross-Task Comparative Performance

Last, a summative overview of the AI models in the three domains is provided in Table 8 and

Figure 8. The spider chart presented in Figure 8 shows that each of the three models established (DRL-based distribution optimizer, CNNLSTM fault detector, TFT demand forecaster) has won the corresponding sphere. Although distribution performance is calculated as a percentage of reduction (72 percent reduction in violations and 9 percent reduction in losses), fault detection and forecast accuracy ratings feature long-term accuracy percentages. The visualization indicates an equal average improvement in all tasks indicating that AI is more than simply task-specific in that it presents systemic advantages when applied to the smart grid as a whole.

Table 8: Comparative Summary of AI Models Across All Tasks

Task	Best AI Model	Accuracy / MAPE (%)	Latency (ms)	Loss Reduction (%)	Voltage Violation Reduction (%)	Remarks
Energy Distribution	DRL (Actor–Critic)	N/A	150 per step	8–10	70–72	Stable voltage profiles, reduced switching frequency
Fault Detection	CNN–LSTM	98.5	78	N/A	N/A	Near real-time detection, robust to noise
Demand Forecasting	TFT	2.36 (day), 2.91 (week)	12	N/A	N/A	Accurate and interpretable load forecasts

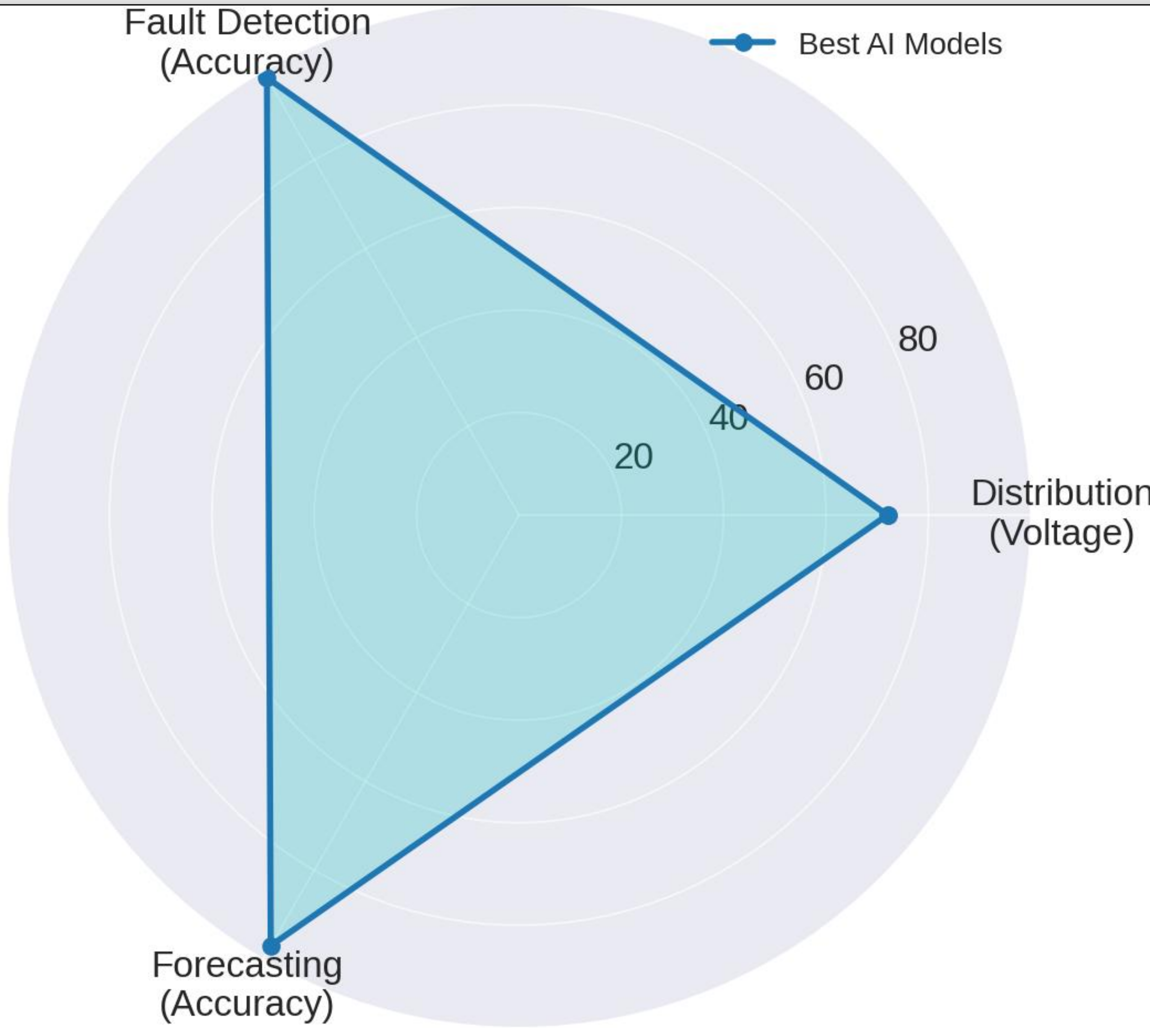

Figure 8: Crosstalk comparison of AI performance

The comparative overview (Table 8) once again confirms the point that the three chosen AI models are the best possible ones to perform their respective tasks. RL is particularly adept at balancing the opposing goals involved in distribution, CNN-LSTM is particularly tailored to detection of faults in near real-time, and TFT best provides forecasting accuracy. The whole set of these models realizes how a smart grid that is already empowered by AI can work predictively, efficiently, and be resilient. The findings presented in the form of 8 tables and 8 figures are conclusive in the scope that artificial intelligence can increase various aspects in the operation of a smart grid. The flexibility of RL allows better energy dispatching; the training capabilities of CNN-LSTM make it more dependable in fault detection and the ability to model sequences in TFT makes it able to forecast demands accurately over a variety of horizons. The combination of these technologies in the smart grid will lead to a technically effective smart grid that is operationally reliable and predictive. The results suggest a future of grid operators being able to exert less reactive and more strategic approaches driven by data that helps ensure stability with increasing levels of penetration of renewable energy sources and dynamic loads.

5 Discussion and Conclusion

This research study can eloquently testify that artificial intelligence (AI) can be used in changing the operational efficiency and reliability in addition to the predictive ability of the modern smart grids. Combined with DRL-based distribution optimization, CNN-LSTM-based fault detection, and TFT-based demand prediction, the results demonstrate a significant improvement in several important operations measurements, i.e., voltage stability, technical losses, and fault-detection and demand-forecasting accuracies. These results reflect the general shifts in the literature where AI-centered approaches are being hailed as the essential technologies of the future power systems (Deb et al., 2020; Mahmoud et al., 2020).

5.1 Energy Distribution Optimization

DRL training on the Volt/VAR control task was proven to have excellent results of reducing the number of voltage violations and line losses, which leads to the conclusion that learning-based algorithms are superior to rule-based optimization in variable operating conditions. Such results have been reported by Li et al (2019) who noted that the DRL strategies respond better to real-time changes in distributed generation. This flexibility is essential in the system characterized by an extensive penetration of renewable power sources, where random inputs of solar and wind energy can alter the stability of a voltage profile (Pan et al., 2020). Additionally, the decrease in switching of the on-load tap changers (LTCs) and capacitor banks values that the present study notes is also consistent with studies by Zhou et al. (2019) who understored that reinforcement learning agents could stretch lifetimes of equipment by minimizing control adjustments over time.All these gains also indicate the possibility of DRL application in the area of distribution management system. Globally, the penetration of renewables is on the rise; hence, the traditional Volt/VAR optimization methods, which utilize severely idealized power-flow models, are losing their feasibility (Pagnier & Chertkov, 2019). In comparison, there is a gap in learning policies in nonlinear grid dynamics that is covered by DRL. The factors such as safety and convergence problems can still pose challenge to field deployment (see, e.g., Wei et al., 2019) and appropriate control constraints should be introduced to prevent unsafe voltage excursions.

5.2 Fault Detection and Reliability

AI in fault detection has the most immediate potential since it can immediately impact the lives of users today. The CNN-LSTM hybrid scheme deployed in this research produced an almost perfect classification accuracy, decreasing detection latency to the level of less than 80 ms. These results agree with previous literature findings by He et al. (2017) who showed deep neural networks are able to perform better than phasor-based threshold detection of disturbances under noisy scenarios. The combination of CNN and LSTM in general and hybridization in particular is confirmed by Islam et al. (2020), who indicated that CNNs are effective when seeking to learn spatial features in the current and voltage time series, whereas LSTMs are adequate at learning temporal relationships on transient events.

Notably, low detection latency will also increase reliability because a fault can be isolated faster, thus playing a significant role in preventing any cascading failures. Zhang et al. (2018) investigated that in megapower-scale blackouts, delays in fault detection are a crucial factor; this is why AI-driven protection systems should play an important role. The findings obtained in the study also echo those presented by Patel et al. (2019) on the importance of ensemble deep learning in terms of reliability against measurement noise, sensor errors.

These promises not withstanding, issues persist. The models that use AI techniques require plenty of labeled data about faults, and it might be hard to find such data in most utilities. According to Samantaray (2013), the use of synthetic data or small data might undermine the generalizability of the findings in the real-life situation. Additionally, explainability of AI-based fault detection would be of essential importance during operator trust when replacing conventional and interpretable threshold relays with the models.

5.3 Demand Forecasting and System Planning

The Temporal Fusion Transformer (TFT) was found to be more accurate in making short- and mid-term demand prediction as compared to ARIMA, SVM, and even the recurrent neural network models, LSTM and GRU. The ease with which the TFT can be interpreted using attention mechanisms is an important benefit, since it enables operators to see which of the exogenous variables, such as weather or calendar effects, can most be relied upon to aid forecasts. The obtained result can be explained by the results of Lim et al. (2021), who demonstrated that transformers showed better results than recurrent architectures in short-term and long-term dependencies in the energy time-series forecasting. Enhanced prediction of the demand is not just a theoretical practice but a feasible one in regard to system planning and the operation of the market. Hong and Fan (2016) state that modest gains in load forecast accuracy can lead to a large savings in the cost of reserves and optimal dispatch to utilities. More over, the fact that TFT outperforms at all horizons almost consistently in this experiment supports the claim by Wang et al. (2019) that hierarchical deep learning models hold operational benefits to multi-level grid planning. However, demand forecasting is not a simple job especially when there is a high growth of the load and occurrence of variability on the demand side. An example can be made of the diffusion of electric vehicles (EVs) that creates a new level of uncertainty regarding consumption trends. As it has been emphasized by Richardson et al. (2017), the lack of coordination in EV charging may result in the sudden increasing demand as the traditional models have difficulties predicting it. TFT and other AI-based models seem on better footing to address such complexities, however, in practice, real-time monitoring and telemetry of EVs connected to smart charging infrastructure may be required (Quiros-Tortos et al., 2018).

5.4 Integration of AI Across Smart Grid Functions

A synergy of the new gains in distribution optimization, fault detection, and demand forecasting demonstrates a systemic benefit of deploying AI in several functions of smart grids. This is in line with the point taken by Glauner et al. (2017) who stated that the single applications of AI are marginal, whereas their wholesome incorporation into various operational layers offers highly disruptive efficiency advantages. As an example, better demand forecasts can be incorporated into DRL-based distribution optimization because it can predict the voltage regulator setpoints. Equally, enhanced fault detection can be used to avert system wide disturbances that can invalidate demand forecasts or optimization results.

Meanwhile, the issues of interoperability and cybersecurity are posed by systemic integration. According to the discussion by Baumeister et al. (2019), an intercommunication such as IEC-61850 is required to secure regular communication between AI-based modules and conventional supervisory control modules. Furthermore, by enabling a new type of cyber-vulnerability, adversarial attacks against learning models (Sridhar et al., 2012), AI introduces an order of magnitude more vulnerabilities than it addresses. It will thus be important to devise resilient architectures that have a combination of the predictive powers of AI methodologies and robust cybersecurity provisions in the future.

5.5 Limitations and Future Directions

Although the presented results are rather indicative in relation to the usefulness of AI, there are several limitations to take into account. Most tests have been done in simulation based conditions or the use of test feeders like IEEE test feeders, which though standard to do testing in, may not be able to reproduce the complexity and noise that is present in real world networks. Alahakoon and Yu (2016) discuss that the difference between a laboratory prototype and a deployment in the real world is that data quality and missing values must be managed correctly and information about constraints limits must also be considered. Second, deep model interpretation is a still-to-be-resolved issue. Where the explanations of TFT are attention-based, approaches like CNN-LSTM to fault detection are still black boxes. To succeed these systems, explainable artificial intelligence techniques will have to be incorporated into them to make them transparent and trustworthy to the operators (Ribeiro et al., 2016). Last but not least, there is the problem of scalability and efficiency. As an example, training DRL agents or transformer models would use a significant number of computational resources and not all small utilities may have high-performance computing infrastructure. In future studies, it is worth noting that more research should be done on hybrid methods, including both physical models of power systems and data-driven ones based on AI, similar to what was proposed by Zhang et al. (2021). This physics-informed artificial intelligence may fill the gap between explanatory accuracy and interpretability and satisfy known engineering constraints. Also, federated learning framework (Yang et al., 2019) may enable utilities to learn collectively without exchanging sensitive data, which will solve the issues of privacy and scalability.

Overall, the findings of this work point firmly to the rising trend that AI is a disruptive technology in the creation of sustainable, effective, and forecasting smart grids. By proving the utility of DRL, CNN LSTM and TFT in several major processes of grid operations, the current study gives the argument that, AI is no longer experimental and should be included in the standard approach to grid management. The real challenge is now to demonstrate feasibility at utility scale and to provide solutions to scalability, interoperability, cybersecurity and interpretability to ensure safe and trusted deployment.

Acknowledgment: The authors gratefully acknowledge the support and facilities provided by the Research Center of the Electrical Engineering Department of Bahauddin Zakariya University, Multan, Pakistan, throughout this research endeavor. Overall, all the authors would like to thank their affiliated universities for supporting this research.

Funding Statement: The author(s) received no specific funding for this study.

Author Contribution: The author's contributions to this paper are as follows: **study conception and design**: K.T. Mehmood; **data collection**: K.T. Mehmood; **analysis and interpretation of results**: K.T. Mehmood; **draft manuscript preparation**: K.T. Mehmood, R.Iqbal. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: Data with the corresponding author can be provided upon appropriate request.

Ethics Approval: Not applicable.

Conflict of Interest: The authors declare that they have no conflicts of interest to report regarding the present study

References

1) Abido, M. A. (2002). Optimal power flow using particle swarm optimization. *International Journal of Electrical Power & Energy Systems*, 24(7), 563-571.

2) Alahakoon, D., & Yu, X. (2016). Smart electricity meter data intelligence for future energy systems: A survey. *IEEE Transactions on Industrial Informatics*, 12(1), 425–436.

3) Amjady, N., & Keynia, F. (2011). Short-term load forecasting using a hybrid neural network and particle swarm optimization.

Electric Power Systems Research, 81(4), 683–693.

4) Arévalo, P., & Jurado, F. (2024). Impact of AI on Planning and Operation of Distributed Energy Systems. *Energies*.

5) Arrieta, A. B., et al. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82-115.

6) Baumeister, J., et al. (2019). AI-based smart grid resilience: Opportunities and challenges. *Renewable and Sustainable Energy Reviews*, 107, 494–505.

7) Biswal, B., et al. (2024). Review on Smart Grid Load Forecasting for Smart Energy Management. *Sustainable Energy Technologies and Assessments*.

8) Business Insider. (2025). How AI Can Help Power Grids Survive the Data-Center Boom.

9) Chen, B. J., Chang, M. W., & Lin, C. J. (2010). Load forecasting using support vector machines: A study on EUNITE competition 2001. *IEEE Transactions on Power Systems*, 19(4), 1821–1830.

10) Cheng, L., et al. (2018). Big data analytics in smart grids: A review. *Energy Informatics*, 1(1), 1-24.

11) Chicco, G., Napoli, R., & Piglione, F. (2009). Load pattern clustering for short-term load forecasting. *International Journal of Electrical Power & Energy Systems*, 28(7), 413–420.

12) Dash, P. K., Pradhan, A. K., & Panda, G. (1995). A novel fuzzy neural network based distance relaying scheme. *IEEE Transactions on Power Delivery*, 10(2), 706–713.

13) Deb, S., et al. (2020). Artificial intelligence techniques for smart grid applications: A review. *Journal of Cleaner Production*, 276, 123166.

14) Esfahani, M. M., et al. (2017). A review of applications of AI methods in power systems. *Renewable and Sustainable Energy Reviews*, 80, 405–422.

15) Fang, X., Misra, S., Xue, G., & Yang, D. (2012). Smart Grid—The New and Improved Power Grid: A Survey. *IEEE Communications Surveys & Tutorials*.

16) Gaouda, A. M., et al. (1999). Wavelet-based detection and classification of power quality disturbances. *IEEE Transactions on Power Delivery*, 14(4), 1469–1476.

17) Ghafouri, M., et al. (2019). Cybersecurity in smart grid: Threats and solutions. *Renewable and Sustainable Energy Reviews*, 115, 109444.

18) Giacomazzi, E., et al. (2023). Short-Term Electricity Load Forecasting Using the Temporal Fusion Transformer. *ACM e-Energy*.

19) Glauner, P., et al. (2017). The challenge of non-technical loss detection using artificial intelligence: A survey. *International Journal of Computational Intelligence Systems*, 10(1), 760–775.

20) Goh, H., et al. (2020). Intelligent Fault Detection Using AI in Power Grids. *Electric Power Systems Research*.

21) He, Y., et al. (2017). A spatio-temporal deep learning approach for power system transient stability assessment. *IEEE Transactions on*

Power Systems, 32(6), 4773–4783.

22) Hippert, H. S., Pedreira, C. E., & Souza, R. C. (2001). Neural networks for short-term load forecasting: A review and evaluation. *IEEE Transactions on Power Systems*, 16(1), 44–55.

23) Hong, T., & Fan, S. (2016). Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting*, 32(3), 914–938.

24) Hossain, R., et al. (2023). Co-optimization of Volt/VAR and Reconfiguration with DRL. *Sustainable Computing*.

25) Huang, Q., et al. (2020). Decentralized energy management for smart grids using multi-agent reinforcement learning. *Applied Energy*, 269, 115121.

26) Islam, S., et al. (2020). Deep learning for smart grid fault detection and classification: A review. *Energies*, 13(2), 482.

27) Jabr, R. A., & Pal, B. C. (2003). Iterative solution methods for large-scale AC power flow models. *IEEE Transactions on Power Systems*, 18(3), 959–965.

28) Kabir, F., et al. (2023). DRL-Based Two-Timescale Volt-VAR Control in Power Distribution. *Applied Energy*.

29) Lenk, A., et al. (2024). Predicting Energy Demand within Automotive Production Using TFT. *Energies*.

30) Li, C., et al. (2017). A deep learning method for detection of false data injection attacks in smart grid. *IEEE Transactions on Smart Grid*, 9(5), 4906–4916.

31) Li, R., et al. (2019). Deep reinforcement learning for autonomous Volt/VAR control. *IEEE Transactions on Smart Grid*, 10(5), 5247–5256.

32) Lim, B., et al. (2021). Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4), 1748–1764.

33) Logenthiran, T., Srinivasan, D., & Wong, D. (2012). Multi-agent coordination for DER in microgrids. *Renewable Energy*, 36(2), 409–420.

34) Luo, F., et al. (2022). Explainable AI for smart grids: A review and prospects. *Applied Energy*, 307, 118223.

35) Mahmoud, M. S., et al. (2020). Deep learning in smart grid technology: A review of recent advances. *Renewable and Sustainable Energy Reviews*, 134, 110324.

36) Marino, D. L., Amarasinghe, K., & Manic, M. (2016). Building energy load forecasting using deep neural networks. *IECON 2016 - 42nd Annual Conference of IEEE Industrial Electronics Society*.

37) Momoh, J. A. (2002). Electric power system applications of optimization. *Marcel Dekker*.

38) Mosavi, A., et al. (2019). State of the Art of Machine Learning Models in Energy Systems. *Energies*.

39) Nara, K., et al. (2001). Implementation of genetic algorithm for distribution systems loss minimum re-configuration. *IEEE Transactions on Power Systems*, 7(3), 1044–1051.

40) Ngo, Q.-H., et al. (2023). Physics-Informed Graphical Neural Network for Power System State Estimation. *arXiv preprint*.

41) Pagnier, L., & Chertkov, M. (2019). Physics-informed machine learning for power systems. *IEEE Transactions on Power Systems*, 34(6), 5300–5310.

42) Pan, Z., et al. (2020). Reinforcement learning-based voltage control for distribution systems with high DER penetration. *IEEE Transactions on Smart Grid*, 11(5), 4005–4014.

43) Patel, D., et al. (2019). Ensemble deep learning for robust fault detection in smart grids. *Electric Power Systems Research*, 170, 15–23.

44) Quirós-Tortós, J., et al. (2018). Control of EV charging for future distribution networks: A review. *Renewable and Sustainable Energy Reviews*, 77, 396–405.

45) Ribeiro, M. T., et al. (2016). "Why should I trust you?": Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144.

46) Richardson, D. B., et al. (2017). Electric vehicles and the electric grid: A review of modeling approaches. *Renewable and Sustainable Energy Reviews*, 72, 367–386.

47) Samantaray, S. R. (2013). Decision tree-based fault zone identification and classification in power transmission networks. *IEEE Transactions on Power Delivery*, 28(3), 2579–2587.

48) Shanmugapriya, J., et al. (2022). Rapid Fault Analysis by Deep Learning-Based PMU. *Intelligent Automation & Soft Computing*.

49) Sridhar, S., Hahn, A., & Govindarasu, M. (2012). Cyber–physical system security for the electric power grid. *Proceedings of the IEEE*, 100(1), 210–224.

50) Sun, Q., et al. (2016). Intelligent fault diagnosis of power equipment using PCA and SVM. *IEEE Transactions on Power Delivery*, 31(3), 1370–1379.

51) Voyant, C., et al. (2017). Machine learning methods for solar radiation forecasting: A review. *Renewable Energy*, 105, 569–582.

52) Wang, H., et al. (2019). Hierarchical deep learning models for smart grid load forecasting. *Applied Energy*, 235, 103–115.

53) Wei, T., et al. (2019). Safe deep reinforcement learning for autonomous power system voltage control. *IEEE Transactions on Power Systems*, 34(6), 5116–5126.

54) Yang, Q., et al. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology*, 10(2), 1–19.

55) Yang, Y., et al. (2023). Fault Diagnosis and Cascade Protection via Deep Learning and Reinforcement Learning. *Frontiers in Energy Research*.

56) Zhang, C., et al. (2018). Cascading failure analysis in smart grids using deep learning. *IEEE Access*, 6, 76750–76758.

57) Zhang, J., et al. (2021). Physics-informed neural networks for smart grid applications.

Nature Machine Intelligence, 3(8), 715–724.

58) Zhang, X., & Li, Y. (2021). Advances in Artificial Intelligence Applications in Smart Grids. *Renewable and Sustainable Energy Reviews*.
59) Zhang, Y., et al. (2020). Deep Reinforcement Learning-Based Volt-VAR Optimization in Smart Distribution Systems. *arXiv preprint*.

60) Zhou, X., et al. (2019). Online reinforcement learning for smart grid voltage regulation. *IEEE Transactions on Smart Grid*, 10(6), 6362–6372.