Global Academic Journal of Economics and Business

Available online at https://www.gajrc.com
DOI: https://doi.org/10.36348/gajeb.2026.v08i06.024

OPEN ACCESS



ISSN: 2706-9001 (P)
ISSN: 2707-2584 (O)

Original Research Article

AI Driven Energy Efficiency Optimization and Predictive Maintenance of Electrical Infrastructure in Saudi Arabia

Mubarak Abdulla[1]*
[1]Independent Researcher

***Corresponding Author**
Mubarak Abdulla
Independent Researcher

Article History
Received: 18.09.2026
Accepted: 10.11.2026
Published: 18.11.2026

Abstract: Saudi Arabia's electrical infrastructure is expanding in scale and operational complexity while utilities, building operators and project owners face simultaneous pressure to limit energy waste and maintain service continuity. This critical review examines how artificial intelligence can connect energy efficiency optimization with predictive maintenance across transformers, switchgear, distribution feeders, motors and building electrical systems. It synthesizes peer reviewed research published between 2020 and 2025 using an explicitly bounded, thematic evidence mapping approach, distinguishing measured operational outcomes from simulations and conceptual proposals. The central finding is conditional rather than technological: successful integration depends on physically meaningful data, asset specific failure mechanisms, trustworthy uncertainty estimates and maintenance decisions linked to verifiable energy baselines. Classification accuracy alone cannot demonstrate reduced losses, avoided outages or economic benefits. A staged framework is proposed for Saudi applications, linking sensing, quality assurance, interpretable models, operational constraints and audited intervention outcomes. The review identifies priorities for locally validated datasets, cross site transfer, cybersecure integration and joint reliability efficiency evaluation.

Keywords: Artificial Intelligence, Energy Efficiency, Predictive Maintenance, Electrical Infrastructure, Condition Monitoring, Power Distribution, Saudi Arabia, Asset Management.

Copyright © 2026 The Author(s): This is an open-access article distributed under the terms of the Creative Commons Attribution **4.0 International License (CC BY-NC 4.0)** which permits unrestricted use, distribution, and reproduction in any medium for non-commercial use provided the original author and source are credited.

1 INTRODUCTION

The modernization of electricity infrastructure exposes a persistent engineering tension. Operators seek lower technical losses, improved asset utilization and more responsive demand management, but aggressive operating schedules can accelerate degradation or reduce protective margins. Conversely, excessive preventive maintenance may interrupt healthy equipment without proportionate benefits. Artificial intelligence (AI) could connect these goals, but few practical systems assess efficiency and reliability through a shared, physically grounded model. Reviews of distribution network digitalization identify substantial progress in sensor based services, alongside unresolved questions about decision integration and investment value [1-26].

That distinction is particularly important in Saudi Arabia. Large construction programmes, climate dependent cooling demand, geographically dispersed assets and ambitious energy transition commitments create different exposure patterns across substations, buildings and industrial facilities. A classifier trained on a laboratory transformer dataset cannot simply be interpreted as a validated

maintenance policy for a hot, dust exposed installation. Similarly, a forecast that reduces simulated peak demand does not prove lower feeder losses once switching limits, harmonics, voltage constraints and maintenance downtime are included. The review therefore treats Saudi relevance as a context for evaluating transferability, not as proof that international results have been replicated locally.

Research has advanced along partially disconnected tracks: transformer diagnosis from dissolved gas analysis, time series prediction from smart sensors, demand response optimization and digital twins for virtual operational testing [2–23]. A useful review needs to examine how these approaches interact. It must also ask whether proposed control policies distinguish efficiency gains from demand shifting, whether anomaly alerts trigger appropriate interventions and whether operators can challenge model recommendations. These questions motivate an integrated rather than algorithm centred synthesis.

2 AIM, OBJECTIVES AND RESEARCH QUESTIONS

The aim is to critically evaluate the conditions under which AI assisted energy optimization and predictive maintenance can jointly improve the efficiency, availability and lifecycle performance of electrical infrastructure relevant to Saudi Arabia. Rather than proposing a one size fits all architecture, the review develops an implementation approach grounded in the available evidence.

Five objectives guide the analysis: (i) classify electrical assets and measurable efficiency loss mechanisms; (ii) assess sensing and fault prognostic methods against their data and failure mode assumptions; (iii) examine the interface between maintenance scheduling and operational optimization; (iv) identify barriers to transfer into Saudi climatic, organizational and procurement settings; and (v) establish an evaluation framework incorporating energy, reliability, cost and model governance. These objectives separate diagnosis, optimization, deployment and verification because each presents different engineering challenges.

Three research questions follow. RQ1 asks which AI approaches have credible evidence for actionable fault identification and prognostics in electrical infrastructure. RQ2 asks how maintenance decisions can be coupled to energy efficiency targets without eroding system reliability or safety. RQ3 asks what evidence, institutional capabilities and performance indicators are necessary before such systems can be scaled in Saudi projects.

3 REVIEW METHODOLOGY

A structured critical literature review was undertaken with a publication boundary of January 2020 to December 2025. Searches targeted combinations of “predictive maintenance”, “condition monitoring”, “power transformer”, “distribution grid”, “energy efficiency”, “machine learning”, “digital twin”, “fault diagnosis”, “demand response” and “electrical infrastructure”. Relevant publisher records and academic index discoverability were screened by title, abstract and reported methods. The synthesis prioritised peer reviewed journal studies and technically informative conference publications that describe electrical assets, AI approaches, validation practices or connected operational decisions.

Eligibility required an identifiable engineering application and sufficient information to discuss either data provenance, modelling choice, diagnostic outcome, optimization objective or deployment constraints. Papers consisting exclusively of generic AI advocacy, unrelated mechanical maintenance, or non electrical conceptual claims were not used as core evidence. Studies outside the specified date window were excluded. Where an article described a simulation, its performance was interpreted as simulation evidence; where field measurements were supplied, the claims were linked to those conditions only. The final reference set comprises 30 distinct DOI bearing publications chosen for a balanced thematic synthesis, rather than a comprehensive census of all published studies.

Evidence was coded under five analytical themes: equipment health, distribution observability, energy optimization, intervention scheduling and implementation assurance. Methodological scrutiny considered class imbalance, measurement uncertainty, temporal leakage, external validation, operational baselines and reproducibility. Findings were compared narratively because asset populations, fault labels and loss metrics are too heterogeneous to support an ethically defensible pooled effect size. No numerical study screening totals are claimed because a prospectively logged, exhaustive database search was not performed. This limitation distinguishes the present critical review from a formal systematic review.

The selection strategy deliberately combines landmark reviews with representative empirical investigations so that claims of broad technical opportunity can be tested against concrete modelling practice. For each source the appraisal distinguished prediction target, sensor modality, training unit, testing unit, maintenance action and claimed operational benefit. Sources without an identifiable

energy baseline were not treated as evidence of energy reduction. Disagreements among studies were examined through differences in asset populations and outcome definitions rather than averaged away. This protocol makes the narrative transparent, while acknowledging that a full bibliometric search or quantitative meta analysis would require further database access and a preregistered screening log.

4 Electrical Losses and Maintenance Mechanisms

Energy inefficiency is not a single technical outcome. Feeder I^2R losses change with current, conductor resistance and configuration; transformer no load losses persist under energization; poor power factor raises apparent demand; excessive temperature increases resistance; and voltage imbalance can degrade motor performance. Fault signatures may be informative about these mechanisms without directly quantifying their energy consequences. A model that predicts insulation deterioration, for example, helps avoid catastrophic failure, but any associated energy saving must be independently measured. The broad smart grid literature emphasizes forecasting and monitoring capabilities while often separating them from causal accounting of physical losses [17-28].

Asset criticality modifies the relevance of a predicted event. In a transformer, dissolved gases and temperature excursions may indicate insulation stress, whereas an abnormal feeder waveform can indicate harmonics or incipient switching defects. A motor's thermal condition depends on loading, ventilation and mechanical interactions, and a switchgear panel may present sparse labels despite high consequences of failure. Those differences require separate signal processing pipelines and maintenance action thresholds. Uniform "health scores" may simplify dashboards but can conceal incomparable underlying risks.

Maintenance and optimization should therefore share an asset register, operating context and event history while retaining different decision outputs. An efficiency optimizer seeks improved operating set points under engineering constraints. A prognostic service estimates fault likelihood, degradation rate or remaining useful life. The supervisory decision layer must consider both and explicitly document why a recommended load transfer, service window or replacement is justified. If diagnosis remains detached from work order execution, the system is informative but not genuinely predictive in its maintenance practice [1-24].

Loss mechanisms interact in ways that complicate simple intervention claims. A reduction in electrical resistance following equipment replacement may improve efficiency under comparable current, but current itself depends on downstream demand and operational switching. Power factor correction may lower upstream reactive flow yet require harmonic studies to avoid resonance. Likewise, reducing transformer load can lower copper losses while increasing losses elsewhere if load is transferred to a less efficient feeder. The appropriate assessment boundary therefore extends beyond the monitored asset to the electrically coupled system. Intervention plans should explicitly identify whether benefit is expected from restored efficiency, avoided failure, changed dispatch or extended equipment lifetime.

Table 1: Asset specific evidence, model outputs and decision limitations

Asset	Observations	AI output	Critical verification
Power transformer	Gas, load, oil temperature	Fault class; health index	Confirmed faults; thermal context
Distribution feeder	Current, voltage, switch events	Loss forecast; anomaly risk	Topology and power flow
Motor and drives	Vibration, temperature, power	Condition trend	Operating duty; maintenance outcome
Switchgear	Partial discharge, events	Risk priority	Protection integrity; sparse labels
Building electrical plant	Metering, occupancy, weather	Demand and load forecast	Weather adjusted energy baseline

5 AI Enabled Condition Assessment

Transformer studies offer the clearest evidence base because dissolved gas analysis provides established chemical features and partially standardized diagnostic categories. Reviews map a progression from ratio rules and expert systems to support vector machines, ensembles and deep neural networks [2]. Studies using optimized neuro fuzzy inference or machine learning root cause analysis demonstrate richer nonlinear classification [9, 10]. Yet this technical improvement must be interpreted against sampling frequency, missing measurements and inconsistent fault confirmation. A laboratory cross validation score should not be reported as field reliability.

A major concern is the rarity of genuine fault events. Oversampling can balance training data but may reproduce synthetic patterns that differ from operational failures. Studies of oversampled transformer data and dense convolutional models show why performance varies with preprocessing choices [6, 7]. If samples from the same physical unit appear in training and test sets, nominally independent evaluation can leak device specific signatures. Rigorous deployment assessment should therefore partition by asset and time, retain an untouched external test set and publish precision recall trade offs for rare events.

Interpretability is equally important. Feature aware transformer approaches reveal how gas concentrations interact in a prediction [3], while methods coupling engineering rules with neural networks aim to preserve diagnostic structure [5]. Explanation charts, however, do not establish physical causality. Correlated temperature, loading and gas indicators may all respond to a common operating condition; an explanation that attributes risk to one feature can mislead if its conditioning assumptions are omitted. Such outputs should supplement, rather than replace, conventional engineering inspection.

Networked sensing adds latency, synchronization and cybersecurity risks. An internet connected transformer diagnosis architecture can classify remotely but must address sensor drift, clock misalignment, communications outages and adversarially altered inputs [4]. Multimodal models that combine sequential and cross sensor information may capture richer failure signatures [8], although they require compatible sampling and calibration. For Saudi deployment, historical fault registers, summer loading patterns, commissioning records and equipment specific specifications are particularly important inputs to validate before fitting a model.

Data labels are particularly fragile in preventive maintenance because an intervention may prevent the failure that a model was intended to predict. Treating every repaired component as a confirmed impending failure exaggerates event rates; treating its continued operation as a false alarm may penalize successful prevention. One remedy is to classify intervention outcomes by evidence strength, separating confirmed physical defects from precautionary replacements and unverified alerts. Another is to retain censoring information when modelling time to failure. The assessment should also compare predictions with simple engineering baselines, including operating hour thresholds and conventional diagnostic rules, rather than only with alternative machine learning classifiers.

6 Energy Optimization and Asset Interaction

AI based energy optimization ranges from short term load forecasts to demand flexibility, set point selection and integrated scheduling. Reviews of smart grid optimization show significant algorithmic diversity, including neural forecasting, heuristic scheduling and multi agent control [17, 18]. The key engineering choice is not which model appears most sophisticated; it is how the objective function represents actual energy losses, customer service, equipment thermal limits and uncertainty. Optimizing electricity expenditure may shift demand to cheaper periods without reducing kilowatt hours, whereas reducing demand peaks may improve transformer loading but incur operational trade offs.

Building electricity systems provide a useful comparison. Forecasting studies show that weather and occupancy variables can improve predictions of heating or cooling loads [13-16]. However, HVAC energy is affected by control logic, thermal comfort and plant maintenance, so a lower forecast error does not establish energy savings. For infrastructure operators, forecasting should feed into constrained scheduling with independently recorded consumption, baseline weather adjustment and safeguards that prevent abnormal equipment cycling.

Factory case studies challenge the assumption that digital methods always dominate physical upgrades. Comparative energy saving research found that AI driven interventions may be smaller in magnitude than hardware changes while potentially offering attractive returns owing to lower implementation cost [12]. This result supports a portfolio approach: efficient transformers, capacitor compensation, improved ventilation, targeted condition monitoring and AI scheduling should be evaluated together. Procurement should compare lifetime energy and failure costs, not acquisition cost or algorithm performance in isolation.

AI forecasting also has an energy footprint. Large models demand compute, cooling and frequent retraining; lightweight methods may offer better net value when prediction tasks are narrow. Research on green learning for regional load prediction provides a direct reason to compare modelling complexity with practical benefit [14]. Any energy efficiency claim should report the monitoring infrastructure's incremental electricity use and maintenance requirements, especially where edge devices are installed at scale.

Demand flexibility introduces another distinction between energy and cost. Scheduling industrial loads away from tariff peaks can lower bills without reducing energy use, while shifting an operation into a hotter period may increase cooling

requirements. A credible optimization study must specify the objective, the baseline and the constraints before comparing outcomes. The decision should account for equipment start stop cycles, production deadlines, voltage limits, power quality implications and the possibility that a forecast error creates an adverse operating condition. Reporting a single cost savings percentage without these assumptions makes results difficult to transfer from simulation to operational infrastructure.

7 Coupling Prognostics with Optimization

Integration is most valuable when the health state changes the feasible operating envelope. A transformer with deteriorating insulation may require a lower permitted loading level; a healthy parallel unit may accept temporary load redistribution; and maintenance should be arranged before deterioration creates disproportionate outage risk. Scheduling research on multiple electrical assets shows that fault predictions can be translated into intervention priorities and time windows [24]. However, optimality depends on reliable degradation estimates and transparent assumptions about service crew availability, spare parts and network configuration.

Digital twins may support what if analysis by representing asset states, operating constraints and anticipated maintenance interventions. Reviews of digital twin maintenance indicate their value for exploring degradation where observed failures are scarce, alongside challenges in model fidelity and virtual physical synchronization [22, 23]. Digital twin studies of integrated energy scheduling also demonstrate the possible operational gains from uncertainty aware virtual coordination [21]. These examples do not justify treating a digital model as a validated replica without repeated calibration against the physical installation.

An operational decision process can combine three distinct outputs: expected incremental energy losses, a calibrated probability of equipment failure over a specified horizon, and the cost of intervention. Importantly, a predicted risk should not automatically trigger shutdown. Human approval, protection coordination, regulatory constraints and work permit procedures remain mandatory. The appropriate output is an auditable ranked action list that records model confidence, relevant sensor evidence, consequences of delay and alternative remedies.

Figure 1 illustrates this chain as a governance structure rather than a claim of a tested proprietary architecture. The proposed integration makes a maintenance recommendation actionable only after it passes engineering constraints and creates a measurable work order. Subsequent asset data then close the loop, allowing validation of both technical forecasts and operational impacts.

Consider a simplified decision: an ageing distribution transformer is forecast to approach a thermal risk threshold during a high demand interval. The energy optimizer might recommend redistributing load, while the maintenance system might request an inspection. A defensible decision combines the expected impact on feeder losses, the receiving asset's thermal headroom, the consequences of outage and the availability of qualified personnel. If the transferred load raises system losses, the operator may nevertheless prefer it to reduce failure probability. Such trade offs demonstrate why multi objective evaluation with explicit risk tolerances is preferable to an unqualified "energy optimal" recommendation.

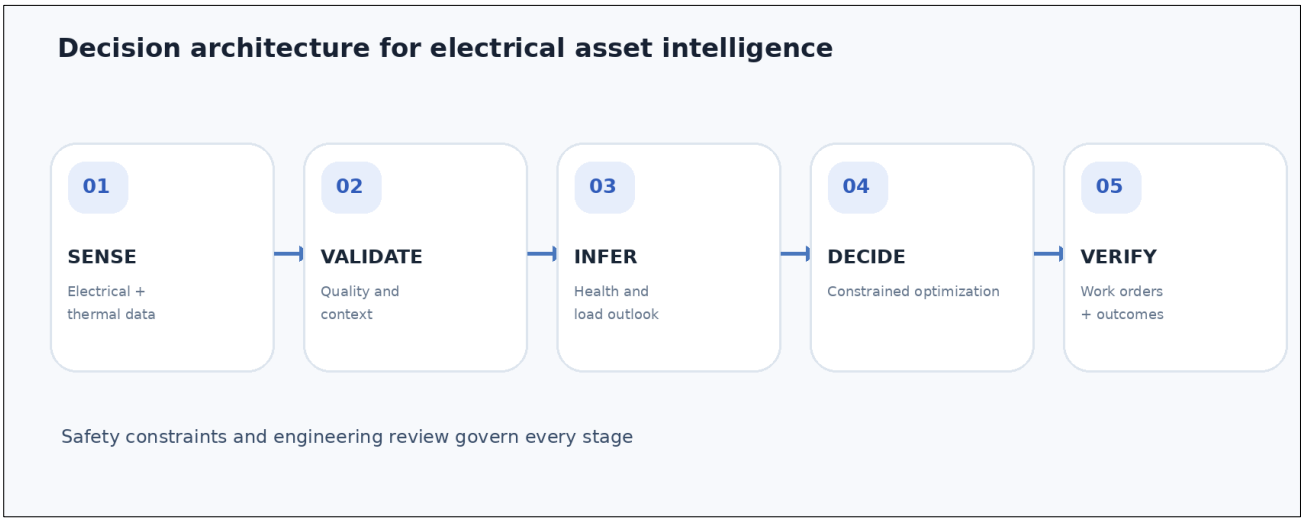


Figure 1: Proposed link between operational sensing, maintenance decisions and outcome verification

8 Evaluation, Uncertainty and Governance

The most consequential weakness across heterogeneous AI studies is the use of metrics that answer the wrong operational question. Classification accuracy may be high because most observations are normal. More useful diagnostic measures include event level recall, false alarms per asset year, detection lead time and calibration of predicted risk. Where remaining useful life is estimated, confidence intervals and late prediction penalties matter because overconfident estimates can defer essential maintenance. Industrial failure analysis surveys repeatedly emphasize the dependence of model conclusions on data structure and assessment design [11].

Energy savings need to be evaluated independently of model accuracy. Report technical losses in kilowatt hours, energy intensity with a defined denominator, peak load, service availability and emissions only when the emissions factor and calculation boundary are documented. A before and after comparison is insufficient if ambient temperature, utilization, tariff structures or network topology changed. Preferred designs use comparable untreated assets, interrupted time series or carefully matched operating windows. Prospective baselines should be archived before any optimization is activated.

Model governance is essential to safe engineering practice, rather than a paperwork exercise. Automated recommendations can fail when sensors are recalibrated, new equipment arrives, load composition changes or vendors update firmware. A controlled deployment should maintain model versioning, feature provenance, safety constraints, rollback procedures, periodic drift checks and a documented authority matrix. Reviews of grid AI deployment warn that simulations frequently bypass real protection requirements and thus overstate readiness [20]. Privacy and cyber risk analysis further indicates that expanded connectivity creates new attack surfaces [29].

Reliability and efficiency metrics must also be evaluated jointly. A marginal energy improvement that increases breaker trips or shortens transformer life is not a net infrastructure improvement. Conversely, an intervention that temporarily increases consumption may prevent a severe outage or equipment replacement. Lifecycle economic assessment should incorporate failure consequences, avoided maintenance visits, spare parts logistics, operating losses and computational cost, with uncertainty ranges rather than one headline payback figure.

The validation horizon should match the decision horizon. Minute level anomalies may justify immediate inspection, whereas replacement decisions require months or years of degradation evidence. An alarm that predicts a fault only seconds before occurrence offers little scheduling value despite excellent classification. Equally, a prognostic model with wide uncertainty may still be useful if it reliably identifies a small set of assets requiring additional testing. Publication quality reporting should include confusion matrices at the selected threshold, subgroup performance by asset type, temporal and external testing, confidence intervals and the number of actual intervention opportunities observed.

9 SAUDI APPLICATION PATHWAY

A staged Saudi implementation should begin with a bounded asset class where sensor data, failure history and operational accountability are already adequate. Transformer substations and large commercial or industrial electrical installations offer plausible pilot settings, but suitability is site specific. The initial task is not algorithm selection: it is agreement on a measurable decision problem. Operators should select assets with recurring avoidable losses or costly unplanned interventions and record their present performance under comparable loading and climatic conditions.

The second stage is data readiness. Asset IDs must connect commissioning information, manufacturer specifications, maintenance reports, procurement records, switching events, environmental measurements and sensor feeds. Correct timestamps and stable units are essential. Saudi projects often involve multiple contractors and equipment suppliers; therefore, ambiguous ownership of data and inconsistent handover documentation may be more restrictive than insufficient computing power. Standardized maintenance coding and procurement specifications for monitoring devices would strengthen continuity across project phases.

Third, asset teams should run models in advisory mode while preserving existing maintenance rules. Predictions should be compared with engineer assessments, subsequent faults and documented interventions. A pre agreed threshold for false alarms prevents a model from overwhelming teams with unnecessary work orders. Only after sustained field validation should selected outputs influence dispatch, scheduling or procurement decisions. Table 2 specifies a staged evaluation plan that distinguishes technical feasibility from verified operational value.

Fourth, scaling requires lifecycle governance. Electrical engineering teams should define protection and load constraints; procurement teams should test vendor interoperability, cybersecurity assurances and replacement lead times; facility managers should approve access windows; and financial teams should audit energy baselines. These responsibilities fit the practical intersection between electrical infrastructure and technical procurement, but they do not imply that any individual procurement engineer has personally validated the proposed AI methods.

Climatic transfer cannot be assumed merely because equipment ratings are identical. Ambient heat modifies cooling capacity and insulation ageing; dust can obstruct ventilation and increase maintenance demand; and operating peaks may coincide with extreme weather. Sites also differ in standby arrangements, criticality and access restrictions. A locally meaningful benchmark should therefore stratify validation by season, asset age, loading band and equipment vendor, and should report when the model abstains from making a recommendation. Abstention under poor data quality is preferable to an apparently precise warning that operators cannot substantiate.

Table 2: Proposed staged field evaluation for Saudi electrical infrastructure

Stage	Evidence required	Go/no go criterion
Scope	Asset criticality; loss baseline; failure history	Documented operational question
Instrument	Calibrated sensors; asset IDs; secure feeds	Traceable, sufficiently complete data
Validate	Temporal external test; false alert rates	Accepted error and drift limits
Pilot	Advisory warnings; engineer disposition	Actions feasible within safety limits
Demonstrate	Matched energy and reliability measures	Audited incremental value
Scale	Ownership, retraining, procurement interoperability	Repeatable results across sites



Figure 2: Evidence requirements increase from initial sensing to demonstrated operational improvement

10 DISCUSSION AND RESEARCH GAPS

The principal synthesis is that AI supports two distinct mechanisms: information improvement and decision improvement. Better information may diagnose degradation sooner; better decisions may reduce losses, improve maintenance timing or avoid outages. Evidence for the first mechanism is more mature, especially for transformers, than evidence that both mechanisms produce sustained benefits simultaneously. Reviews of distribution maintenance reveal similar fragmentation between data selection, predictive modelling and operational implementation [1]. Research should consequently shift from model contests toward controlled studies of complete decision pathways.

Four gaps deserve priority. First, open datasets rarely capture Saudi specific operating conditions, so model portability across seasonal loading and ambient temperatures remains uncertain. Second, health indices and optimization objectives often use inconsistent timescales, complicating coordination between hourly control and multi month degradation. Third, prospective

demonstrations seldom report both avoided energy losses and maintenance outcomes. Fourth, few papers fully expose model drift, cybersecurity controls or the human authority needed to reject unsafe recommendations. These are limitations of the available evidence, not proof that the underlying technologies cannot work.

A useful research programme would develop multi site validation protocols, test uncertainty aware maintenance scheduling and compare AI interventions with conventional engineering upgrades. It should also quantify the burden of sensors and computing, evaluate fairness across older and newer assets, and test whether actionable explanations improve decisions made by experienced technicians. Without these steps, increasingly complex algorithms risk improving numerical scores while leaving operational performance unchanged.

The interdisciplinary implication is significant. Electrical engineers define safe operating constraints, maintenance specialists judge failure plausibility, data scientists establish statistical validity, procurement teams secure comparable sensor and equipment specifications, and managers evaluate whole life value. A programme that omits any one of these functions risks a technically capable but operationally isolated platform. Future comparative trials should evaluate combinations of monitoring, conventional energy upgrades and AI based scheduling, including a control group where feasible. This would allow observed benefits to be attributed to the integrated decision process rather than to concurrent refurbishment or changing demand.

A prospective Saudi field programme should first select three comparable feeder or substation groups, document their load patterns, then reserve at least one group as an operational benchmark. The sampling design needs to capture major differences in age, manufacturer, electrical configuration, maintenance history and exposure to environmental stress, because each can alter the apparent value of monitoring. Researchers should state the failure definition before training and distinguish verified deterioration from inspections triggered by uncertain alerts. An event level report should describe whether technicians encountered insulation damage, contamination, overheating or an alternative explanation, with adequate time stamps for independent review. Where repair averts a potential breakdown, investigators must account for incomplete knowledge of the counterfactual failure rather than converting every preventive intervention into a true positive. Predictive claims should be tested against an ordinary engineering threshold rule and an existing scheduled maintenance policy, not only against a second machine learning architecture. For optimization, investigators need feeder level energy balances that separate technical losses, consumer demand, distributed generation and switching actions. An apparent fall in consumption becomes persuasive only after normalization for operational hours, season, weather conditions and changes in service delivery. Industrial facilities must also track production volumes because lower output can create superficially attractive energy indicators while masking deteriorating process efficiency. Uncertainty intervals should accompany reported benefits and be wide enough to reflect instrument precision, seasonal variability and missing observations. Maintenance staff should be able to record why they accepted, postponed or rejected each recommendation, creating a decision history that is more informative than alarm logs alone. An effective trial would measure the time from alert to assessment, the proportion of useful alerts and the costs of unnecessary site attendance. Protection engineers must retain authority over switching and derating, with clearly documented safeguards against thermal overload and loss of coordination. Training and procurement specialists should ensure that model suppliers deliver data dictionaries, accessible exports, version documentation and agreed procedures for software changes. Long term studies should verify that monitoring devices and data pipelines remain supportable after project handover and that models do not become unmaintained dependencies. Subsequent research can evaluate whether calibrated uncertainty and transparent explanations improve engineers' intervention decisions rather than simply increasing confidence in dashboards. A cross site consortium would be especially useful because pooled evaluation can reveal seasonal and equipment specific failure modes that are invisible in a single pilot. However, sharing must respect asset security: federated evaluation, restricted access and carefully aggregated outputs may be more appropriate than exposing detailed substation telemetry. The comparative economic analysis should account for the capital burden of instruments, installation access, system integration, data quality staff and cybersecurity assurance. It must also distinguish recurring software subscription costs from one time hardware improvements and report payback sensitivity to uncertain failure frequency. An intervention that improves fault anticipation but creates excessive inspection workload may not be cost effective even when its statistical ranking is excellent. Conversely, reducing high consequence incidents may justify costs that a narrow kilowatt hour assessment would overlook. Researchers should publish thresholds, denominators, exclusion rules and model update schedules to let future operators interpret the findings in their own

11 Limitations

This review is restricted to publications dated 2020–2025 and to an intentionally bounded collection of 30 DOI identified studies; it is not an exhaustive systematic evidence inventory. The literature covers diverse assets, countries, validation settings and performance definitions. As a result, the synthesis does not estimate an average energy saving percentage or forecast a Saudi return on investment. Some references supply general smart grid context rather than direct evidence of maintenance linked efficiency gains. The proposed implementation framework is analytical and has not been experimentally validated. Its suitability must be tested against actual Saudi utility, industrial or facility datasets.

12 CONCLUSION

AI enabled condition monitoring and energy optimization can form a valuable integrated capability for electrical infrastructure, but their combined benefit is neither automatic nor demonstrated by predictive accuracy alone. The strongest research evidence concerns asset specific diagnostics, while the translation into auditable maintenance and energy outcomes remains comparatively thin. For Saudi Arabia, defensible adoption requires calibrated local data, physically constrained control, safety governed intervention and independent performance measurement. The recommended path is a staged programme that evaluates reliability, energy, lifecycle cost and cyber resilience together. Future studies should demonstrate decision level results under real operating conditions rather than infer infrastructure improvement from isolated algorithms.

REFERENCES

1. Designing digitally enabled proactive maintenance systems in power distribution grids: A scoping literature review. Energy Reports (2024). https://doi.org/10.1016/j.egyr.2024.08.044
2. Fault detection and diagnosis in power transformers: A comprehensive review and classification of publications and methods. Electric Power Systems Research (2022). https://doi.org/10.1016/j.epsr.2022.107990
3. A bi level machine learning method for fault diagnosis of oil immersed transformers with feature explainability. International Journal of Electrical Power & Energy Systems (2022). https://doi.org/10.1016/j.ijepes.2021.107356
4. Effective IoT based deep learning platform for online fault diagnosis of power transformers against cyberattacks and data uncertainties. Measurement (2022). https://doi.org/10.1016/j.measurement.2021.110686
5. Learning from even a weak teacher: Bridging rule based Duval method and a deep neural network for power transformer fault diagnosis. International Journal of Electrical Power & Energy Systems (2022). https://doi.org/10.1016/j.ijepes.2021.107619
6. Incipient fault diagnosis in power transformers by data driven models with over sampled dataset. Electric Power Systems Research (2021). https://doi.org/10.1016/j.epsr.2021.107519
7. Transformer fault diagnosis based on improved deep coupled dense convolutional neural network. Electric Power Systems Research (2022). https://doi.org/10.1016/j.epsr.2022.107969
8. Multi modal information analysis for fault diagnosis with time series data from power transformer. International Journal of Electrical Power & Energy Systems (2023). https://doi.org/10.1016/j.ijepes.2022.108567
9. An intelligent system based on optimized ANFIS and association rules for power transformer fault diagnosis. ISA Transactions (2020). https://doi.org/10.1016/j.isatra.2020.03.022
10. Root cause analysis improved with machine learning for failure analysis in power transformers. Engineering Failure Analysis (2020). https://doi.org/10.1016/j.engfailanal.2020.104684
11. A comprehensive survey on applications of AI technologies to failure analysis of industrial systems. Engineering Failure Analysis (2023). https://doi.org/10.1016/j.engfailanal.2023.107172
12. Universal artificial intelligence workflow for factory energy saving: Ten case studies. Journal of Cleaner Production (2024). https://doi.org/10.1016/j.jclepro.2024.143049
13. Net Zero Energy cost Building system design based on Artificial Intelligence. Applied Energy (2024). https://doi.org/10.1016/j.apenergy.2023.122348
14. A novel green learning artificial intelligence model for regional electrical load prediction. Expert Systems with Applications (2024). https://doi.org/10.1016/j.eswa.2024.124907
15. Simulating long term energy consumption prediction in campus buildings through enhanced data augmentation and metaheuristic optimized artificial intelligence. Energy and Buildings (2024). https://doi.org/10.1016/j.enbuild.2024.114191
16. Optimizing building energy performance predictions: A comparative study of artificial intelligence models. Journal of Building

Engineering (2024). https://doi.org/10.1016/j.jobe.2024.109247
17. A comprehensive review of artificial intelligence approaches for smart grid integration and optimization. Energy Conversion and Management: X (2024). https://doi.org/10.1016/j.ecmx.2024.100724
18. Deep learning for intelligent demand response and smart grids: A comprehensive survey. Computer Science Review (2024). https://doi.org/10.1016/j.cosrev.2024.100617
19. Leveraging the machine learning techniques for demand side flexibility – A comprehensive review. Electric Power Systems Research (2024). https://doi.org/10.1016/j.epsr.2024.111185
20. Challenges and Limitations of Artificial Intelligence Implementation in Modern Power Grid. Procedia Computer Science (2024). https://doi.org/10.1016/j.procs.2024.05.008
21. Digital twins based day ahead integrated energy system scheduling under load and renewable energy uncertainties. Applied Energy (2022). https://doi.org/10.1016/j.apenergy.2021.117899
22. Predictive maintenance using digital twins: A systematic literature review. Information and Software Technology (2022). https://doi.org/10.1016/j.infsof.2022.107008
23. Advances of Digital Twins for Predictive Maintenance. Procedia Computer Science (2022). https://doi.org/10.1016/j.procs.2022.01.348
24. Predictive maintenance scheduling for multiple power equipment based on data driven fault prediction. Computers & Industrial Engineering (2022). https://doi.org/10.1016/j.cie.2021.107898
25. Artificial intelligence techniques for enabling Big Data services in distribution networks: A review. Renewable and Sustainable Energy Reviews (2021). https://doi.org/10.1016/j.rser.2021.111459
26. A systematic review of machine learning applications in the operation of smart distribution systems. Energy Reports (2022). https://doi.org/10.1016/j.egyr.2022.09.068
27. Impacts of digitalization on smart grids, renewable energy, and demand response: An updated review of current applications. Energy Conversion and Management: X (2024). https://doi.org/10.1016/j.ecmx.2024.100790
28. Smart grids: A comprehensive survey of challenges, industry applications, and future trends. Energy Reports (2024). https://doi.org/10.1016/j.egyr.2024.05.051
29. Big data analytics and artificial intelligence aspects for privacy and security concerns for demand response modelling in smart grid: A futuristic approach. Heliyon (2024). https://doi.org/10.1016/j.heliyon.2024.e35683
30. A systematic review of big data innovations in smart grids. Results in Engineering (2024). https://doi.org/10.1016/j.rineng.2024.102132.