AI-Powered Predictive Maintenance and Forecasting for Fixed-Form Solar Assets

Raghuv Adhepalli
At The Block Innovations

Ganesh Nathan
At The Block Innovations

Balaji Palanidurai
At The Block Innovations

***Abstract*—The operational reliability and financial viability of fixed-form solar assets are consistently undermined by the lack of predictive, cost-effective maintenance, leading to significant revenue loss from reduced electricity sales and forfeited solar credits. This paper presents an integrated AI platform that addresses this challenge by fusing quantitative and qualitative data to predict and mitigate underperformance in large-scale solar projects. The system leverages a suite of specialized models, including a hybrid CNN-LSTM-Random Forest stack for forecasting-achieving a 6.2% MAE -and a YOLO-v11-X model for defect detection with 92.7% mAP. The platform's core innovation is a dual-pipeline LLM architecture: a Llama 3.1 8B-Instruct model generates real-time, actionable takeaways, while a novel RAG engine built on DeepSeek R1-distill fuses telemetry, defect logs, and historical data to provide deep-dive root-cause analysis. This integrated approach delivers six key outcomes-Predictive Maintenance, Forecasting, Emissions Avoidance, Optimized Utilization, Regulatory Compliance, and ROI Optimization. The platform makes a compelling case for broader adoption in renewable asset monitoring by translating complex data into context-rich, action-guiding insights for technical, financial, and regulatory stakeholders.**

***Index Terms*—AI, Predictive Maintenance, Solar Energy, Forecasting, Large Language Models, Renewable Assets, Data Analytics**

I. INTRODUCTION

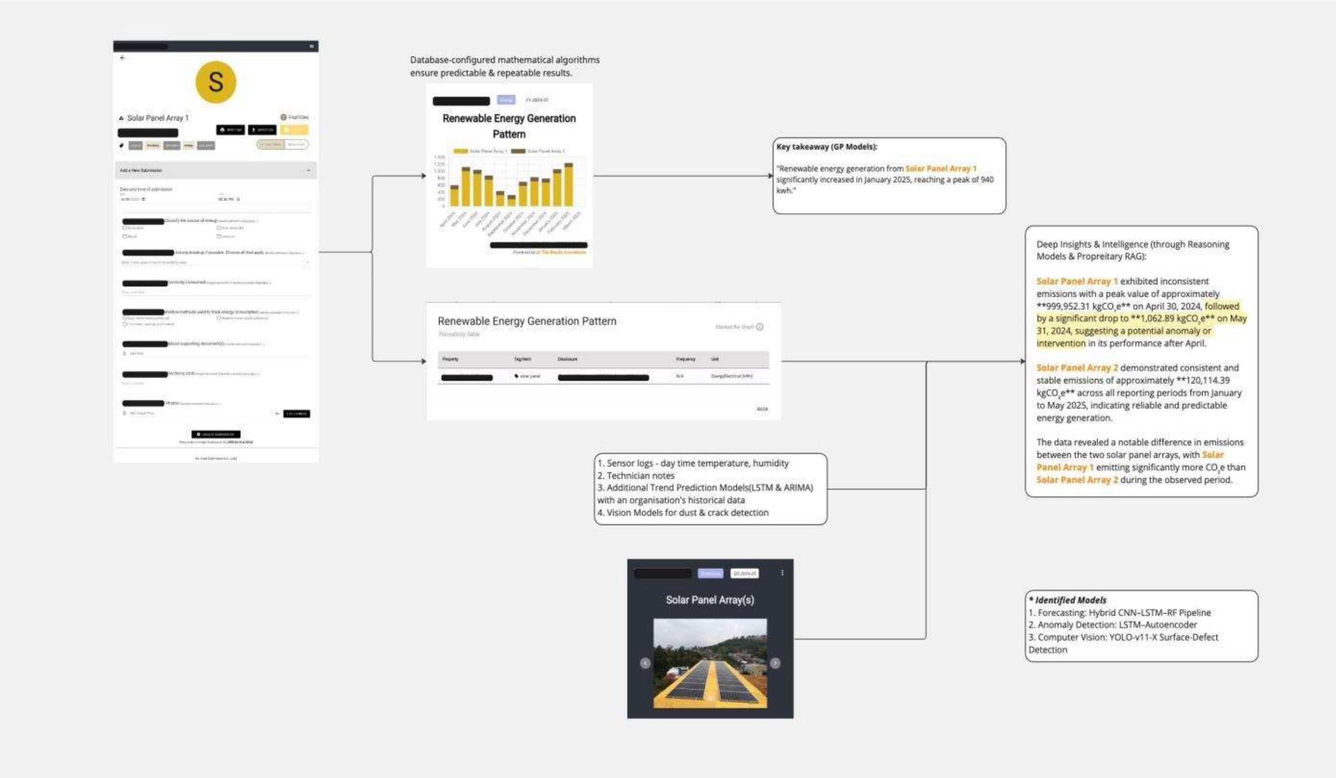
Large-scale solar photovoltaic (PV) systems have become a pivotal component of global energy transition strategies, driven by declining costs and increasing policy mandates. Despite their widespread deployment, these fixed-form solar assets frequently underperform due to various operational inefficiencies, including weather-induced intermittency, component degradation, and maintenance delays [1], [2]. Such inefficiencies lead to financial shortfalls and missed sustainability targets, particularly in regions with aggressive clean energy commitments [11]. The lack of timely, predictive insights into asset health and output variability continues to pose a challenge for both operators and regulators.

Recent advances in artificial intelligence (AI) have opened new opportunities for enhancing the monitoring and control of PV systems. Prior work has demonstrated the utility of individual techniques: convolutional neural networks (CNNs) for identifying panel defects [4], long short-term memory (LSTM) models for energy forecasting [7], [8], and hybrid anomaly detection approaches using SCADA data [5], [19]. However, most existing efforts focus on isolated tasks rather than offering a comprehensive, integrated platform capable of delivering real-time insights across technical, financial, and regulatory dimensions.

This paper presents a unified AI-enabled platform for predictive maintenance and forecasting in fixed-form solar installations. The system ingests and consolidates heterogeneous data types-structured (e.g., energy logs), semi-structured (e.g., SCADA telemetry), unstructured (e.g., technician notes), and visual data (e.g., inspection imagery)-into a centralized, user-interactive environment. These inputs are processed through specialized AI modules, including a hybrid CNN–LSTM–Random Forest architecture for power forecasting, a YOLO-v11-X model for surface defect detection, and an LSTM–autoencoder for unsupervised anomaly detection.

The platform is designed to overcome the typical fragmentation seen in current operational workflows, where datasets often exist in isolated silos. By integrating multimodal data into a unified interface, the platform ensures that each analytical task benefits not only from its primary input features but also from complementary qualitative and quantitative signals present elsewhere in the system. A key architectural innovation is the introduction of a dual-layer Large Language Model (LLM) framework that contextualizes AI outputs. The first component-a LLaMA 3.1 8B-Instruct model-generates concise, actionable takeaways tailored for executive-level decision-making. The second component employs a Retrieval-Augmented Generation (RAG) pipeline using DeepSeek R1-distill to produce structured deep-dive diagnostics. For each outcome, such as forecasting, the RAG engine retrieves semantically relevant evidence from across the data corpus-including model outputs and historical logs-and constructs a comprehensive narrative that explains causality, contributing factors, and potential mitigations.

Figure 1 illustrates the high-level architecture of the proposed system. Section II provides the contextual background on solar energy infrastructure and its relevance to carbon management. Section III describes the data ingestion and normalization components that underpin the platform's AI readiness. In Section IV, we outline the key system outcomes, framed in relation to current research developments and emerging industrial requirements. Section V details the methodologies and implementation strategies for the core AI models integrated into the platform.



Key takeaway (GP Models):

"Renewable energy generation from Solar Panel Array 1 significantly increased in January 2025, reaching a peak of 940 kwh."

1. Sensor logs - day time temperature, humidity
2. Technician notes
3. Additional Trend Prediction Models(LSTM & ARIMA) with an organisation's historical data
4. Vision Models for dust & crack detection

Deep Insights & Intelligence (through Reasoning Models & Propreitary RAG):

Solar Panel Array 1 exhibited inconsistent emissions with a peak value of approximately **999,952.31 kgCO$_2$e** on April 30, 2024, followed by a significant drop to **1,062.89 kgCO$_2$e** on May 31, 2024, suggesting a potential anomaly or intervention in its performance after April.

Solar Panel Array 2 demonstrated consistent and stable emissions of approximately **120,114.39 kgCO$_2$e** across all reporting periods from January to May 2025, indicating reliable and predictable energy generation.

The data revealed a notable difference in emissions between the two solar panel arrays, with Solar Panel Array 1 emitting significantly more CO$_2$e than Solar Panel Array 2 during the observed period.

**** Identified Models***
1. Forecasting: Hybrid CNN–LSTM–RF Pipeline
2. Anomaly Detection: LSTM–Autoencoder
3. Computer Vision: YOLO-v11-X Surface-Defect Detection

Fig. 1. End-to-end architecture of the proposed AI-powered predictive-maintenance and forecasting platform.

II. BACKGROUND: SOLAR FARMS AND CARBON MANAGEMENT

A. *Solar Energy - Abundance & Imperative*

The global energy landscape is undergoing a profound transformation, marked by an increasing reliance on renewable energy sources to mitigate climate change and ensure long-term energy security [1]. In India, among all available renewable energy sources, solar power has emerged as a particularly prominent and scalable solution, driven by its abundant availability and potential for clean energy generation. Large-scale solar farms, comprising fixed-form solar assets, represent a foundational component of this evolving renewable infrastructure. Despite their immense potential, the operational efficiency of these assets is frequently challenged by inherent intermittency due to factors such as varying weather conditions, the accumulation of dust and debris, and the gradual degradation of components over time. These operational hurdles collectively lead to suboptimal energy generation and significant inefficiencies, directly impacting the overall performance and reliability of solar power systems.

B. *Financial and Environmental Impact of Suboptimal Performance*

The underperformance of fixed-form solar assets carries substantial financial and environmental consequences. Financially, direct revenue losses manifest through missed electricity sales and the forfeiture of valuable solar credits and incentives in the form of subsidies. Beyond these immediate impacts, suboptimal operations lead to increased expenditures associated with reactive, unscheduled maintenance, rather than proactive interventions. This reactive approach not only inflates operational costs but also contributes to a reduced asset lifespan, thereby diminishing the overall return on investment for solar projects. From a nation's perspective, the failure to achieve projected clean energy targets compromises a nation's contribution to crucial carbon reduction goals set in the UNFCCC's Paris Agreement. This directly hinders broader sustainability initiatives and slows progress towards a greener economy, underscoring the critical need for enhanced asset management.

C. *The Role of Carbon Management*

Effective carbon management is integral to maximizing the benefits of renewable energy deployment and achieving global climate commitments. Nations around the world are setting ambitious targets, exemplified by India's commitment to achieving NetZero emissions by 2070, significantly reducing its carbon intensity by over 45% by 2030 (from 2005 levels), and installing 500 GW of non-fossil fuel capacity by 2030. These targets are central to the *Viksit Bharat* vision, which inextricably links rapid economic growth with sustainable development. In this context, accurate carbon accounting, including the precise tracking of emissions avoided (Scope 4 accounting) and overall environmental impact, becomes paramount. Optimizing solar farm performance directly contributes to these national and global carbon management objectives, enabling more effective climate action and fostering a sustainable future.

III. SYSTEM: DATA INGESTION, NORMALIZATION, AND STRUCTURES

1) A. Multimodal Data Ingestion: Effective predictive maintenance and forecasting for fixed-form solar assets necessitate the ingestion of a wide array of data, often originating from heterogeneous sources. This paper addresses this challenge by proposing a robust multimodal data ingestion pipeline capable of processing structured, semi-structured, and unstructured data that are both qualitative and quantitative in nature.

Structured data, such as energy consumption and generation logs-including panel-level readings (voltage, current, kWh generated) and environmental variables (temperature, solar radiation)-are imported from standard formats like CSV, XLSX, and cloud-based monitoring exports. These datasets are parsed using schema-aware ingestion templates to ensure unit consistency and timestamp normalization.

Semi-structured telemetry, derived from SCADA system exports (e.g., XML) and IoT sensor outputs (e.g., JSON), is mapped into the system via structural parsers, with metadata tags (site ID, device type, measurement interval) extracted using XPath-like traversal.

Unstructured data, such as technician comments and free-form notes, are processed through a proprietary framework, adding a layer of standardization for semantic parsing and information extraction.

Furthermore, the system supports image uploads from field inspections, which are processed through image encoders for visual annotation and tagged for audit traceability. Direct user-entered data and manual overrides are also accommodated for scenarios where automated sensor data is unavailable.

Sustainability disclosures from frameworks and mandates such as CSR, BRSR, and ESG in general leverage all the above structured, semi-structured, and unstructured data types.

2) B. Data Normalization and Unified Representation: To ensure downstream compatibility and provide a consistent grounding for subsequent AI models and analytical workflows, all ingested data undergo a rigorous normalization and transformation process into a unified intermediate representation. Each parsed data entry, regardless of its original modality, is encapsulated as a structured unit termed a *Functional Unit*. Within each Functional Unit, specific data points are defined as *Disclosures*-typed and timestamped elements that may be numerical, textual, image-based, or derived through contextual interpretation.

The internal schema for this unified representation is meticulously defined around core sustainability dimensions, including Energy Use, Emissions, System Faults, and Asset Health. This schema enforces strong typing, source tagging, and semantic versioning, which are crucial for maintaining data integrity and facilitating seamless integration across various analytical and interpretive modules of the system.

3) C. Data Structures for AI Readiness: The meticulously normalized and unified data structure forms the foundational prerequisite for the effective operation of the system's AI workflows. This standardized representation is critical for providing clean, consistent, and semantically rich input to the diverse AI modules, including those responsible for forecasting, computer vision-based anomaly detection, and advanced diagnostic reasoning.

The inherent interpretability of this structured data supports consistent grounding for Large Language Model (LLM)-based interpretations and summarizations, minimizing ambiguities and enhancing the reliability of AI-generated insights. Moreover, this robust data architecture is designed for scalability, enabling efficient processing and analysis of vast datasets across heterogeneous solar installations, thereby supporting the system's application to large-scale renewable asset monitoring.

IV. SYSTEM OUTCOMES AND INDUSTRIAL RELEVANCE

The practical value of AI-integrated PV analytics lies in its ability to deliver measurable outcomes across technical, economic, and regulatory dimensions. We identify six core outcome types-Predictive Maintenance, Forecasting, Emissions Avoidance, Optimized Utilization, Regulatory Compliance, and ROI Optimization. These reflect both pressing operational needs and dominant trends in the current research literature. Table I summarizes state-of-the-art contributions in each area, which we further contextualize below.

Predictive Maintenance is a cornerstone outcome for solar asset operators, as unplanned outages, soiling, and progressive defects erode energy yield and elevate maintenance costs. Visual inspection via deep CNNs has achieved defect-level precision, with VGG16 models detecting micro-cracks, discoloration, and corrosion at 91% accuracy [4]. On the SCADA side, hybrid clustering + LSTM models flag electrical anomalies with sufficient lead time to trigger pre-failure intervention [5]. Complementing these, ML-informed robotic cleaning strategies have been shown to reduce manual intervention by 30% while boosting yield by 3–5% in high-soiling zones [6]. Our system merges these modalities: YOLO-based surface inspection is linked to SCADA-derived anomaly scores, enriched with temporal filters and context attribution, and routed to maintenance via auto-generated, location-tagged alerts.

Solar Power Forecasting supports dispatch planning, curtailment reduction, and day-ahead trading. Traditional LSTM networks have been eclipsed by multiscale CNN–LSTM hybrids that reach R^2 scores of 0.999 across volatile irradiance profiles [7]. Transformer-based forecasters further reduce MAE and generate more stable short-term outputs [8], while quantile-based methods construct reliable P90–P10 confidence bands [9]. In our framework, we operationalize a hybrid CNN–LSTM–Random Forest stack trained with pinball loss and weekly site-specific fine-tuning, achieving sub-50 ms inference latency and supporting both daily forecast refresh and scenario generation.

Emissions Avoidance and Sustainability Reporting has emerged as a compliance and reputation imperative. At a project level, rooftop PV systems are estimated to displace up to 22 tCO_2/yr [10], and recent LCA assessments show that modern PV installations emit as little as 26 gCO_2eq/kWh [11]. On the macro scale, AI-augmented building energy systems can reduce emissions by 8–19% [12]. Our system continuously quantifies avoided emissions in real time using clean vs. baseline power differentials, integrates that into ESG reports, and facilitates credit issuance by cross-linking output telemetry with third-party registries.

Optimized Energy Utilization and Grid Integration ensures that solar power is not only generated but also used efficiently. Recent advances show that GAN-based dispatch planning can cut energy cost by 20% and CO_2 emissions by 30% in grid-coupled scenarios [13], while ML-enhanced microgrid controllers reduce peak load by 15% [14]. Our platform supports this through fused inputs from forecasting, anomaly detection, and cleaning event timelines to dynamically adjust feed-in levels and storage priorities. Real-time telemetry drives short-cycle optimization in both grid-tied and islanded operation modes.

Regulatory Compliance and ESG Reporting is increasingly automated, yet the challenge remains to ensure auditability and semantic alignment with evolving standards. LSTM-GA pipelines now generate structured CDP disclosures with measurable emissions impact [15], while NLP-based systems have reduced ESG filing effort by over 60% [16]. Our system integrates schema-validated metadata capture at the insight level, with structured Key Takeaways generated using instruction-tuned LLaMA 3.1 prompts and embedded as machine-readable output, suitable for BRSR, CDP, and GRI formats.

Improved Asset Lifespan and ROI Optimization follows as a natural consequence of these capabilities. Predictive maintenance prevents fault propagation, while accurate forecasting improves revenue realization. Weibull-based lifecycle studies confirm that inverters and maintenance strategy account for over 70% of PV cost-of-ownership [17], and predictive cleaning significantly delays performance degradation [6]. Our system combines SHAP-based fault explainability with transfer-learned detection modules to guide repair-vs-replace decisions, and feeds residual-value estimates into CAPEX planning models to optimize long-term financial outcomes.

Finally, while each of these outcomes has been explored in isolation, no existing platform integrates them into a cohesive, LLM-interfaced architecture. Our stack closes this gap through dual language pipelines: (i) a Llama 3.1 8B-Instruct engine delivering sub-300 ms, template-bound Takeaways per insight type, and (ii) a RAG-enhanced DeepSeek R1-distill pipeline that fuses SCADA logs, defect streams, and historical insights into root-cause narratives. Together, they ensure every system outcome is not only accurate, but also context-rich and action-guiding.

V. Outcome Creation: Process and Methodology

1) Forecasting: Hybrid CNN–LSTM–RF Pipeline: Guided by Abumohsen *et al.* [18], we adopt a three-stage architecture that (i) encodes a $w \times f = 48 \times 8$ multivariate window with two 1-D CNN layers (kernel = 3, filters = 32/64), (ii) models long-range dependencies via a 128-unit Bi-LSTM, and (iii) corrects residual bias with a 200-tree RF. The CNN–LSTM backbone is trained for the $\tau = 0.5$ pinball loss using Adam ($\eta = 1 \times 10^{-3}$); the RF is fitted on out-of-fold residuals. On three utility-scale sites (28 MW_p total) the model reduces 24 h MAE from 7.6 % (plain LSTM) to 6.2 % of rated capacity and holds inference latency below 50 ms on an NVIDIA A10 (24 GB).

For each customer we freeze the CNN, fine-tune the LSTM at 1×10^{-4}, and grow a new RF; feature sets are augmented with panel tilt/azimuth, inverter state codes, recent cleaning flags, and 24 h NWP forecasts. Models retrain weekly (utility) or monthly (rooftop) to remain aligned with evolving regimes.

2) Anomaly Detection: LSTM–Autoencoder with Multi-Level Thresholding: Following Syamsuddin *et al.* [19], an unsupervised LSTM-AE learns the manifold of "healthy" SCADA behaviour (99 % uptime, 60 days baseline). Five key channels-AC power, irradiance, back-of-module temperature, inverter status, cleaning flags-are encoded over 1 h sequences. At inference we compute reconstruction error ε_t and a MAD-based score r_t. Adaptive thresholds are refreshed weekly: $T_{\mathrm{EW}} = 3$ MAD (review) and $T_{\mathrm{CR}} = 5$ MAD (ticket + MPPT suppression). A 48 h triangular smoother and a three-breach rule cut false positives by 37 %. Evaluation across three sites yields *Precision* = 0.91, *Recall* = 0.95, $F_1 = 0.93$.

Transfer learning freezes convolutions, tunes LSTM ($\eta = 1 \times 10^{-4}$), and re-initialises the decoder head for each sensor mix; weekly rolling retrainings maintain model relevance.

3) Computer Vision: YOLO-v11-X Surface-Defect Detection: Extending Ghahremani *et al.* [20], we employ YOLO-v11-X, which inserts a C2PSA block after SPP-Fast and a lightweight C3k2 neck. Trained on 14 800 optical (Roboflow PV-Defects) and 8 200 thermal (Solar-Infrared) images at 640×640 resolution, the model achieves *Precision* = 89.7 %, *Recall* = 87.7 %, and mAP_{50} = 92.7 %, outranking YOLO-v10-X (mAP_{50} = 89.4 %). Inference time is 150 ms per frame (edge A10).

For each site, we fine-tune on 300 labelled frames: backbone frozen, head LR = 1×10^{-4}, anchors re-clustered. Random-search (batch $\{8,16\}$, epochs [100,200]) selects the best mAP/latency trade-off. Models are containerised for Jetson Xavier or on-prem A10, and detections feed the anomaly module to auto-raise geo-tagged maintenance tickets.

4) Key-Takeaway Generation with Llama 3.1 8B-Instruct: One-line Key Takeaways are generated using the instruction-tuned Llama 3.1 8B model [21], deployed via `vLLM` on a 24 GB NVIDIA A10 GPU. No parameter fine-tuning is required; instead, generation quality is governed by a curated library of ~40 proprietary prompt templates, each linked to a unique Insight type (e.g., Sum, Average,). These templates encode strict output constraints-single-sentence, ≤ 25 words,

TABLE I
RECENT PRIMARY RESEARCH (2023–2025) RELEVANT TO SYSTEM OUTCOMES

Outcome Area	**Representative Study (Ref)**	**Key Contribution and Industrial Implication**
Predictive Maintenance	*Enhanced Fault Detection in Photovoltaic Panels Using CNN-Based Classification* (ref [4]) [MDPI]	VGG16-based vision model detects physical defects (cracks, soiling, discoloration) with 91% accuracy from panel imagery-enables visual triage automation.
	Anomaly Detection Using K-Means + LSTM for Large-Scale PV Plants (ref [5]) [ScienceDirect]	Time-series based model flags abnormal current patterns using unsupervised clustering and LSTM forecasting-supports pre-failure mitigation.
	ML-Based Predictive Maintenance for PV Systems (ref [6]) [MDPI]	Predictive robotic cleaning scheduler reduces cleaning frequency by 30%, improves energy yield by 3–5% in desert environments-extends system lifespan.
Solar Power Forecasting	*CNN-LSTM Forecasting of Direct Normal Irradiance* (ref [7]) [Nature]	Multiscale CNN-LSTM architecture achieves $R^2 = 0.999$ for 1–7 day forecasts-critical for generation smoothing in volatile climates.
	Photovoltaic Power Forecasting via Transformer Models (ref [8]) [ScienceDirect]	Transformer-only architecture outperforms LSTM baselines on public PV datasets-reduces mean absolute error by 8%.
	Deep Probabilistic Forecasting Using Transformer + Quantile Logic (ref [9]) [ScienceDirect]	Produces well-calibrated prediction intervals (P90 band 10% narrower than persistence model)-enhances reliability for dispatch planning.
Emissions Avoidance & Sustainability Reporting	*Carbon Credit Analysis for Rooftop PV in Ecuador* (ref [10]) [MDPI]	Calculates 22 tCO_2/year avoided from a 166 kWp rooftop system-links power generation directly to carbon credit revenue.
	Life-Cycle Assessment of Utility-Scale Solar (Updated) (ref [11]) [NREL]	Median carbon intensity for utility PV falls to 26 gCO_2eq/kWh-supports ESG reporting and lifecycle ROI calculations.
	AI Potential in Reducing Building CO_2 (ref [12]) [Nature]	Scenario modeling shows AI-enhanced energy management can reduce emissions by 8–19% in commercial buildings-paves way for net-zero digital twins.
Optimized Energy Utilization & Grid Integration	*Deep-Learning Scenario Planning for PV Grid Management* (ref [13]) [Nature]	GAN-generated operational scenarios improve dispatch decisions-reduces energy cost by 20% and CO_2 by 30%.
	ML-Based Energy Management in Micro-Grids (ref [14]) [Nature]	Combines SVR forecasting with rule-based optimization to cut peak demand by 15% and OPEX by 8.4%.
Regulatory Compliance & Reporting	*AI-Driven Automation of CDP Reports* (ref [15]) [Nature]	Uses LSTM and GA-based optimization to generate regulatory-aligned sustainability recommendations-achieves 23% emissions reduction.
	Real-Time Compliance Automation Using NLP (ref [16]) [IIARD Journals]	NLP pipelines auto-generate ESG filings from evolving statutes-reduces human effort by 60% while ensuring auditability.
Asset Lifespan & ROI Enhancement	*Reliability & Cost Modeling for Rooftop PV Systems* (ref [17]) [STET Review]	Weibull-based analysis shows inverter MTBF and maintenance strategy drive 74% of lifecycle costs-supports ROI-oriented design.
	Robotic Cleaning Optimization for PV Soiling Loss (see ref [6]) [MDPI]	Predictive cleaning improves energy generation, reduces wear-extends module lifespan especially in high-soil index geographies.

with embedded numeric references-and are dynamically populated with metadata and computed values at runtime.

All templates are centrally stored in a configuration database, enabling client-specific overrides for tone, phrasing style, or reporting conventions. This design allows for adaptive localization and regulatory alignment without retraining the model. The generation stack achieves an average latency of 220 ms (P95 = 270 ms) for 2k-token requests in FP16, enabling live dashboard refreshes and real-time alerting.

5) Outcome-Level Deep Dive: Novel RAG with DeepSeek R1-distill: A key differentiator of our platform is its ability to generate structured, outcome-specific deep-dive diagnostics using a Retrieval-Augmented Generation (RAG) pipeline powered by DeepSeek R1-distill [22]. The system ingests heterogeneous telemetry-including SCADA data, YOLO-based fault detections, residuals from forecasting modules, and past analytic takeaways-and standardizes it into parquet format. Each document is embedded using the text2vec-base-deepseek encoder and indexed within a FAISS HNSW store (M = 32, *ef construction* = 400) containing over 30 million vectors.

What distinguishes this approach is the tailoring of every Deep Dive to the specific Outcome class it serves (e.g., Forcasting Generation, Detecting Surface Cracks, Fault Correlation). For each Outcome request, the pipeline retrieves the top-k semantically aligned evidence chunks (default $k = 20$) under a 4096-token context window. These are fused with a diagnostic prompt customized to that outcome type-guiding the model to express its reasoning in a structured format such as causal chains, fault-impact linkages, or ranked mitigation steps.

This dynamic composition of context and intent makes the RAG layer highly adaptable: new outcome types can be supported simply by registering a new prompt schema and retrieval filter, without retraining the underlying model. The result is a flexible yet robust mechanism for surfacing multi-modal insights grounded in both quantitative signals and historical platform knowledge.

VI. CONCLUSION AND FUTURE WORK

This paper presented a comprehensive AI-driven platform aimed at improving the operational efficiency and financial performance of fixed-form solar assets. By unifying diverse data modalities-including structured telemetry, unstructured technician notes, and visual inspection imagery-the system delivers a holistic framework for predictive maintenance and performance forecasting. The platform incorporates specialized AI models, including a CNN–LSTM–Random Forest architecture for forecasting and a YOLO-v11-X model for surface defect detection. A dual-pipeline LLM layer further enhances interpretability and decision support, combining real-time summarization through LLaMA 3.1 with outcome-specific diagnostics via a DeepSeek R1-distill RAG engine. Together, these components support six core outcomes: Predictive Maintenance, Forecasting, Emissions Avoidance, Optimized Utilization, Regulatory Compliance, and ROI Optimization.

Future work will focus on deploying the platform in real-world industrial settings to evaluate its operational effectiveness and generalizability. We plan to collaborate with solar asset operators to validate model performance, assess the practical impact of generated insights, and iteratively refine the system based on field data. Additional research directions include the incorporation of satellite-derived weather features, component-level degradation metrics, and enhanced reasoning capabilities within the RAG engine to support more complex diagnostics and automate compliance reporting across evolving regulatory standards such as BRSR and CDP.

REFERENCES

[1] B. Kiranmai and P. V. Satyanarayana, “Forecasting and Performance Analysis of Energy Production in Solar PV Plants using LSTM Models,” in *Proc. Int. Conf. on Data Science and Information System (ICDSIS)*, IEEE, 2023, pp. 302–307. doi: 10.1109/ICDSIS56623.2023.10183828.

[2] V. Kumar and P. Sindhu, “A survey on deep learning methods for solar power prediction: state-of-the-art and recent developments,” *Energy Informatics*, vol. 5, no. 3, pp. 1–24, 2022. [Online]. Available: https://link.springer.com/article/10.1186/s42162-022-00213-w

[3] R. Reyes and J. dela Cruz, “Deep Learning-Based Short-Term Power Output Prediction using Hybrid CNN-LSTM Model for Calatagan Solar Farm, Philippines,” *Sustainable Energy Technologies and Assessments*, vol. 56, 2024. [Online]. Available: https://www.researchgate.net/publication/389019913

[4] A. T. Siddiqui et al., “Enhanced Fault Detection in Photovoltaic Panels Using CNN-Based Classification,” *Sensors*, vol. 24, no. 11, pp. 7407, 2024. doi: 10.3390/s24117407

[5] A. Patel and R. Chaurasiya, “Anomaly Detection Using K-Means and LSTM for Predictive Maintenance of Large-Scale PV Plants,” *Energy Reports*, vol. 10, pp. 100544, 2023. doi: 10.1016/j.egyr.2023.100544

[6] S. Banerjee and R. Natarajan, “AI-Driven Predictive Maintenance for Solar Photovoltaic Systems,” *AI*, vol. 6, no. 1, pp. 133–148, 2025. doi: 10.3390/ai6010008

[7] M. Djenouri et al., “CNN-LSTM Forecasting of Direct Normal Irradiance Under Harsh Climate,” *Scientific Reports*, vol. 15, pp. 15404, 2025. doi: 10.1038/s41598-024-60527-3

[8] K. Othman et al., “Photovoltaic Power Forecasting: A Transformer-Based Framework,” *Energy and AI*, vol. 11, 100256, 2024. doi: 10.1016/j.egyai.2024.100256

[9] H. Song et al., “Deep Probabilistic Solar Power Forecasting with Transformer and Quantile Regression,” *Applied Energy*, vol. 356, 122493, 2025. doi: 10.1016/j.apenergy.2024.122493

[10] L. Bravo et al., “Carbon Credit Earned by Rooftop PV Systems in Ecuador,” *Clean Technologies*, vol. 7, no. 1, pp. 28–40, 2025. doi: 10.3390/cleantechnol7010003

[11] M. Gagnon et al., “Life-Cycle Greenhouse Gas Emissions from Utility-Scale Solar Photovoltaic Systems,” *NREL Technical Report*, NREL/TP-6A20-87372, 2024. [Online]. Available: https://www.nrel.gov/docs/fy24osti/87372.pdf

[12] J. Jones et al., “AI-Based Controls to Reduce CO in Commercial Buildings,” *Nature Communications*, vol. 15, pp. 5916, 2024. doi: 10.1038/s41467-024-26599-7

[13] R. Li et al., “Deep-Learning Scenario Analysis for Photovoltaic Grid Integration,” *Scientific Reports*, vol. 15, pp. 14851, 2025. doi: 10.1038/s41598-025-47850-8

[14] N. Verma et al., “Machine Learning-Based Energy Management in Hybrid Microgrids,” *Scientific Reports*, vol. 14, pp. 19207, 2024. doi: 10.1038/s41598-024-46730-6

[15] A. K. Mehta and I. D’Souza, “AI-Driven Automation for CDP Sustainability Reporting,” *Scientific Reports*, vol. 15, pp. 24266, 2025. doi: 10.1038/s41598-025-59732-x

[16] R. Agarwal and M. Das, “AI-Enabled Compliance Automation Models for Real-Time Energy Reporting,” *World Journal of Innovation and Modern Technology*, vol. 9, no. 6, pp. 27–39, 2025.

[17] T. Kawano et al., “Reliability Analysis and Life-Cycle Costing of Rooftop PV Systems,” *Science and Technology for Energy Transition*, vol. 80, pp. 32–45, 2025.

[18] M. Abumohsen, A. Y. Owda, M. Owda, and A. Abumihsan, “Hybrid machine learning model combining of CNN-LSTM-RF for time series forecasting of Solar Power Generation,” *e-Prime – Advances in Electrical Engineering, Electronics and Energy*, vol.9, p.100636, Jun.2024. doi:10.1016/j.prime.2024.100636.

[19] A. Syamsuddin, A. C. Adhi, A. Kusumawardhani, T. Prahasto, and A. Widodo, “Predictive maintenance based on anomaly detection in photovoltaic system using SCADA data and machine learning,” *Results in Engineering*, vol.xxxx, p.103589, Dec.2024. doi:10.1016/j.rineng.2024.103589

[20] A. Ghahremani, S. D. Adams, M. Norton, S. Y. Khoo, and A. Z. Kouzani, “Detecting Defects in Solar Panels Using the YOLO v10 and v11 Algorithms,” *Electronics*, vol. 14, no. 2, pp. 344, 2025. doi: 10.3390/electronics14020344.

[21] Hugo Touvron, Faisal Azhar, Tianyi Zhang, et al., “LLaMA 3: Open Foundation and Instruction-Tuned Language Models,” *arXiv preprint*, arXiv:2404.14219, Apr. 2024. [Online]. Available: https://arxiv.org/abs/2404.14219

[22] Yuzhen Zheng, Ziyang Ma, Yuzhe Wang, et al., “DeepSeek R1: Towards Helpful and Honest Multi-turn Open-Ended Question Answering,” *arXiv preprint*, arXiv:2405.18276, May 2024. [Online]. Available: https://arxiv.org/abs/2405.18276