OPEN ACCESS

International Journal of Innovations in Science & Technology

RESEARCH & INNOVATION DIVISION

IJIST

Performance Analysis of a Grid-Connected Solar PV System for Industrial Loads Using AI-Assisted Techniques

Bibi Ayesha[1], Bilal Ur Rehman[1*], Muhammad Amir[1], Kifayat Ullah[1] and Waqar Ishaq[2]

[1]Department of Electrical Engineering, University of Engineering and Technology, Peshawar, Pakistan.

[2]Department of Telecommunication, Hazara University, Mansehra, Pakistan.

***Correspondence**: bur@uetpeshawar.edu.pk

Citation | Ayesha. B, Rehman. B. U, Amir. M, Ullah. K, Ishaq. W, "Performance Analysis of a Grid-Connected Solar PV System for Industrial Loads Using AI-Assisted Techniques", IJIST, Vol. 8 Issue. 3 pp 1315-1337, June 2026

DOI: https://doi.org/10.33411/IJIST/1935

Received | April 28, 2026 **Revised |** June 01, 2026 **Accepted |** June 06, 2026 **Published |** June 18, 2026.

The industrial sector in Pakistan faces increasing challenges due to rising electricity costs, grid instability, and growing environmental concerns. Although grid-connected solar photovoltaic (PV) systems offer a promising solution; however, conventional performance assessment approaches often fail to capture the dynamic behavior of industrial loads and solar energy generation. This study proposes an Artificial Intelligence (AI)-assisted framework for evaluating the technical, economic, and environmental performance of a grid-connected PV system supplying industrial loads. Historical meteorological data from the Pakistan Meteorological Department (PMD) and industrial electricity demand data from the Lahore Electric Supply Company (LESCO) for 2023 were utilized. Four machine-learning models, namely Artificial Neural Network (ANN), Support Vector Regression (SVR), Random Forest Regression (RFR), and Long Short-Term Memory (LSTM), were developed and compared for load and solar power forecasting. Among the evaluated models, LSTM achieved the highest forecasting accuracy, with Root Mean Square Error (RMSE) values of 32.5 kW for load forecasting and 18.5 kW for solar forecasting, representing approximately 48% improvement over the baseline Linear Regression model. Based on LSTM forecasts, a 500 kWp grid-connected PV system generated 825,000 kWh annually, achieved 78% self-consumption, and reduced grid electricity imports by 643,500 kWh/year. Economic analysis revealed annual net savings of Rs 20.80 million, a payback period of 3.37 years, a Net Present Value (NPV) of Rs 112 million, and an Internal Rate of Return (IRR) exceeding 28%. Furthermore, the system reduced carbon emissions by approximately 412.5 tonnes CO_2 annually. (or 412.5 metric tons of CO_2 annually. These findings demonstrate the effectiveness of AI-assisted forecasting for reliable industrial PV-system evaluation and investment planning.

Keywords: Machine Learning, Grid-Connected Solar Photovoltaic (PV) System; Industrial Energy Demand Forecasting, Carbon Emission Reduction; Pakistan Energy Sector.

Introduction:

Three converging challenges of rising electricity demand, increasing environmental concerns, and critical need for energy security are creating significant upheaval in the global energy scenario [1]. Global electricity consumption increased by nearly 5% between 2023 and 2024 and developing economies accounted for more than 70% of this growth [2]. Concurrently, the energy sector continues to be the single largest emitter of global greenhouse gases, responsible for roughly 73% of total global carbon dioxide (CO_2) emissions [3]. These challenges are compelling policy makers, industries, and researchers to fast-track a shift towards renewable and sustainable energy sources.

Among all economic sectors, industry is the dominant consumer of electricity and contributes approximately 42% of global electricity demand. In developing countries like Pakistan, the industrial sector accounts for around 28% of the national GDP but is severely impacted by the drawbacks associated with electricity consumption, such as instability of the national grid, load shedding, steep rise in electricity tariffs, and mismatch between electricity supply and demand. In the years prior to and during the 2023 study period the Pakistani power sector consistently saw generation deficits in excess of 5,000 MW that worsened further in 2024, thus leading to shutdowns in industries and reliance on costly diesel back-up generators [4]. Such a situation severely damages industrial competitiveness, dampens foreign investment and harms economic progress.

Solar photovoltaic (PV) technology has provided a robust and progressively more affordable solution to these issues. Over the past decade, global weighted average levelized electricity (LCOE) for utility-scale solar PV has dropped by 89%, from $0.381/kWh in 2010 to $0.044/kWh in 2023 [5]. In Pakistan, following the recent increase of the industrial tariff to Rs 26.23-27.00/kWh [6], solar PV has become even more economically attractive. By installing a grid-connected solar PV system, industrial consumers can generate their own clean electricity, thus decreasing dependence on the national grid, reducing their electricity bills, and reducing their carbon footprint.

However, designing and implementing a performance evaluation study for a grid-connected PV system connected to industrial loads is far from straightforward. Unlike residential or commercial consumers, industrial consumers exhibit highly variable and often erratic load demand profiles, due to their dependence on the continuous and correct functioning of industrial machinery; a single factory can have sharp demand spikes when machinery is started up, followed by stable, high loads during production shifts, and very low loads during scheduled breaks or complete shutdowns. On the other hand, power output of solar PV systems is inherently intermittent and stochastic, depending on real-time changing weather conditions, influenced by varying factors such as humidity, temperature, aerosol concentration and cloud cover [7].

Static methods of performance evaluation like assuming a 'fixed daily load profile' or 'average peak hours' are inadequate and cannot reliably quantify system performance in real-time conditions. "Typical PV system design assumes static modeling practices", which significantly amplifies errors in estimating energy yield, monetary savings, and carbon emission reduction [8]. Such errors can lead to considerable financial risks for the industrial investor, such as inflated estimates of return on investment (ROI) and payback period and unwarranted grid integration penalties.

The advent of Artificial Intelligence (AI) and Machine Learning (ML) provides a powerful set of tools and techniques for such challenges. ML algorithms are adept at learning non-linear relationships and temporal dependencies in data, which allows them to accurately forecast future demand of electricity at industrial sites and power generation from PV systems [9]. Among the vast family of ML techniques, Long Short-Term Memory (LSTM) networks (a type of Recurrent Neural Network (RNN)) have proven to be remarkably successful in identifying long term temporal relationships in time series data [10]. Applying AI based performance analysis methods effectively replace static assumptions with dynamic, data driven predictions for reliable and low-risk performance evaluation of grid-connected PV systems.

In response to this critical need, this research focuses on the development, implementation and validation of an AI assisted performance analysis methodology for a grid-connected PV system designed for industrial consumers, with Pakistan as a case study. One year (2023) of high-resolution historical weather data that is obtained from the ground-based Pakistan Meteorological Department (PMD) station in Lahore, and one year (2023) of high-resolution load data which is obtained from the Lahore Electric Supply Company (LESCO) are utilized in the proposed method. The Artificial Neural Networks (ANN), Support Vector Regression (SVR), Random Forest Regression (RFR) and LSTM models were developed, trained and evaluated; the best model, i.e. LSTM was then used to forecast the technical, economic and environmental performance of a 500 kWp grid-connected PV system. A forward-looking economic assessment was carried out using the 2023 technical operational baseline and the current grid emission factors along with the current Pakistani industrial tariff rates (notified with effective date February 2026). This intentionally de-coupled timeline provides an ex-ante assessment by separating the technical baseline from any historical tariff volatility in order to provide a modern investment horizon.

Problem Statement:

The performance assessment of grid-connected solar photovoltaic (PV) systems for industrial applications is commonly based on deterministic and static modeling approaches that assume constant load profiles, average solar irradiance values, and linear relationships between environmental conditions and PV power output. These conventional methods fail to capture the dynamic and nonlinear nature of industrial electricity demand and solar energy generation, including fluctuations caused by machine operations, changing weather conditions, seasonal variations, temperature effects, soiling, and inverter performance. Furthermore, they do not account for the temporal correlation between industrial load demand and solar PV generation.

As a result, significant inaccuracies arise in energy yield estimation, economic feasibility analysis, grid integration planning, and environmental impact assessment. Such inaccuracies can lead to unreliable forecasts of energy production, misleading estimates of cost savings and return on investment, unexpected grid dependency, and incorrect evaluation of carbon emission reductions. These challenges are particularly significant in Pakistan, where many industrial solar PV feasibility studies rely on generalized climatic datasets and standard software tools rather than site-specific operational and meteorological data.

Therefore, there is a need for an intelligent and data-driven approach that can accurately forecast both industrial load demand and solar PV power generation. This research addresses this gap by developing and validating an Artificial Intelligence (AI)-based forecasting framework capable of predicting hourly load and PV generation profiles and utilizing these forecasts to provide a more realistic technical, economic, and environmental assessment of grid-connected solar PV systems for industrial applications.

Aims and Objectives:

The main objective of this research is to build, implement and test an innovative AI-powered system for performance evaluation of grid-connected photovoltaic system providing electrical power to industrial load.

To achieve the goal, the following are defined as the sub-objectives.

To study the real-time operation of a solar grid connected PV system with industrial loads.

Identify and classify generic industrial electricity demand patterns in Pakistan.

Utilize four AI/ML models for industrial load and solar power prediction and comparison.

To analyze system performance in multiple dimensions using the AI-generated prediction:

The performance of the AI-assisted forecasts vs the "traditional" or baseline methods.

Providing several policy recommendations to all industrial stakeholders and policymakers

Scope and Limitations:

The boundaries of this investigation have been drawn tightly to allow the investigation to remain manageable, focused and replicable within time and financial constraints.

Use Case Application: Typical aggregated load profile representative of a medium to large sized manufacturing facility in Pakistan. Methodology to be extensible to other industrial facilities/ locations using custom input data.

System set up: Grid-connected Solar PV systems with no battery energy storage. Net Metering is considered with 1:1 credit value for exported power.

Size of System: 500 kWp (peak DC capacity), a realistic size for a medium sized industrial facility.

Context of study: Meteorological data from the ground-based PMD station (Lahore vicinity, to represent Lahore's industrial corridor). Load data will be from the LESCO feeder.

Time frame: One year of high-resolution historic data for 2023 will be provided to train the system, along with a project life of 25 years to perform economic analysis. Note that the consumption and generation profile of 2023 is taken as typical and representative of a normal year of operations, thereby making it possible to apply tariff structures of February 2026 retroactively.

Methodology: Simulation based approach entirely programmed in Python 3.8+ [11] utilizing the available open-source libraries. There will be no hardware implementation nor real time control.

Evaluation Metrics: MAE, RMSE (Forecasting); Self-consumption, Grid reduction (Technical); Payback, NPV, IRR (Economic); CO_2reduction (Environmental).

Limitations:

No BESS (battery energy storage system): Hybrid systems with battery energy storage were not analyzed. It was assumed to make a separate analysis on the benefit brought solely by AI forecasting.

No other renewables (Wind, Biomass, Diesel): No further consideration on the integration of renewable sources other than PV.

No onsite measurements for verification: AI models were developed and trained but never simulated on actual hardware. Validations using real data collected from actual PV plant were not conducted.

Net Metering simplified: The analysis assumes a standard 1:1 credit exchange value under the industrial tariff framework. This serves as a simplified accounting model and does not account for complex Time-of-Use (ToU) billing differentials or NEPRA's specific national average power purchase price buyback rates for exported units. Actual financial returns may vary based on these localized utility structural policies.

Generalization: The methodology used in the frame can be applied to any geographical locations, however the results and computation were obtained assuming specific tariffs, locations and climatic conditions, hence should not be applied to different regions without acquiring local values.

Inverter and grid assumptions are perfect: It is assumed that the inverter will maintain constant 98% efficiency and the grid access will be 100% all the time. Hence the output results are estimated somewhat on the higher side.

Degradation of PV modules ignored: The linear degradation of the PV module (0.5 % –1% /yr) is not included in the loop simulation. This condition is chosen intentionally to create the best possible economic scenario, though it makes metrics such as NPV and IRR a little optimistic.

Significance of Study: This study offers great importance to different stakeholders as well as both the academic community and industrial decision-making process.

Socio-Economic Significance for Industrial Stakeholders:

Offers a data-driven, risk-adjusted method to appraise solar PV investments instead of static over-optimistic predictions.

Quantifies with realistic monetary data (Rs 20.8m yearly savings, 3.37 yrs payback period) that grid connected PV is financially feasible with current Pakistani tariff.

Limits financial risk by defining error bounds of the prediction (e.g. RMSE of load prediction by LSTM is 32.5 kW)

Implications for Energy Planners and Utility Companies:

Gives a comparison study (benchmark) for four different AI models (ANN, SVR, RFR and LSTM) for industrial load and solar forecasting use case. This enables practitioners to select the most suitable model for their specific use case.

Gives a reuseable Python-based simulation framework that can be adjusted for different feeder, city or industrial sector.

Provides the grid relief (peak shaving, less import) obtained by a PV system of 500 kWp, this may be relevant for distribution network planning.

Relevance to Environmental Policy-Makers and Sustainability Frameworks:

A refined calculation (412.5ton/y) is provided which can be reported on corporate sustainability reports (Global Reporting Initiative (GRI), Carbon Disclosure Project (CDP)) or used for carbon credits bargaining.

Validates Pakistan's climate targets outlined under Paris Agreement by establishing the industrial decarbonization possibility [12].

Provides a justification for policymakers to include AI based renewable energy analysis as the best practice and may inspire changes to current project appraisal frameworks.

Novelty of the Study:

Although numerous studies have applied Artificial Intelligence (AI) techniques for either electricity load forecasting or solar power prediction, most existing research focuses primarily on forecasting accuracy and does not extend the analysis to a comprehensive evaluation of photovoltaic (PV) system performance. Furthermore, many previous studies utilize generalized meteorological datasets, typical load profiles, or a single forecasting model, limiting their applicability to real industrial environments.

The novelty of this study lies in the development of an integrated AI-assisted framework that combines industrial load forecasting, solar power forecasting, and grid-connected PV system performance evaluation within a single methodology. Unlike previous studies, this research employs real industrial load data obtained from the Lahore Electric Supply Company (LESCO) and site-specific meteorological observations from the Pakistan Meteorological Department (PMD). In addition, four forecasting approaches, namely Artificial Neural Network (ANN), Support Vector Regression (SVR), Random Forest Regression (RFR), and Long Short-Term Memory (LSTM), are systematically compared to identify the most suitable model for industrial PV applications.

Another key contribution is the utilization of the best-performing forecasting model to perform hourly energy-balance simulations and quantify technical, economic, and environmental outcomes, including self-consumption, grid-import reduction, annual savings, Net Present Value (NPV), Internal Rate of Return (IRR), and carbon-emission reduction. Therefore, the proposed framework extends beyond conventional forecasting studies by providing a practical decision-support tool for industrial renewable-energy planning and investment assessment.

Literature Review:

The industrial sector forms the backbone of economic development, both in developed and developing countries. It is also the biggest electricity consuming sector. Globally, industry consumes about 42% of total electricity, led by energy-intensive subsectors like steel, cement, chemicals, textiles, etc. In Pakistan, the sector accounts for around 28% of GDP, but the energy challenge within the industry remains unresolved. Evidence of this includes sharp increases in the cost of electricity (by over 50% between 2020 and 2025) along with the recurring shortages in electricity supply and poor quality of power [4], eroding industrial competitiveness, and consequently, growing interest in onsite renewable energy generation. Among the available renewable energy options for self-generation by industries, solar PV has been the most mature, cost-effective and widely adoptable technology. Globally, the weighted average LCOE of utility scale solar PV has fallen by 89% from 2010 ($0.381/kWh) to 2023 ($0.044/kWh) [5]. The levelized cost for rooftop commercial and industrial systems has also fallen drastically, with typical payback periods in sunny locations ranging between 3 and 7 years. In Pakistan, analysis of market trends indicates that the cost of installing a grid-tied solar PV power plant has decreased in a systematic manner from around Rs 250,000/kWp to a reasonably affordable level of Rs 130,000 to Rs 150,000/kWp by 2025 [13]. A lot of techno-economic feasibility analyses have been performed for grid-connected solar PV systems in industrial settings. For example, [14] performed a feasibility study of a 1 MWp rooftop solar PV plant in a textile factory located in Faisalabad, Pakistan. With conventional assumptions of solar peak sun hours and fixed industrial load profiles, they found a payback period of 4.2 years and an IRR of 22%, but acknowledged their static modeling could overestimate savings as it did not account for the time-varying load profile and cloudy conditions. Similarly, [15] investigated the feasibility of a 500 kWp grid-connected rooftop solar system in an industrial estate in Lahore using PVsyst software with data from a typical meteorological year (TMY). Though the PVsyst software predicted annual generation of 850,000 kWh, actual installed plant output monitored over 1 year was 720,000 kWh, representing about 15% less than the predicted values due to the assumptions of averaged inputs. The International Energy Agency (IEA) has also stated that "traditional PV systems are usually modeled statically, ignoring industrial load characteristics and weather variability," which may lead to an overestimation of the overall energy saving and self-consumption of the electricity generated from rooftop solar systems. Moreover, in the international context, [8] reviewed design practices and reported that "PV system designs, however, often assume static modeling practices that are oversimplifications and not sufficiently representative of the highly dynamic and stochastic behavior of industrial PV systems." It strongly advocated the use of high-resolution and data-driven approaches for the design of such systems to enhance the accuracy of predictions. Studies [16][17] have also highlighted that the conventional feasibility studies rarely consider the time mismatch between solar power generation and load consumption, consequently exaggerating the level of self-consumption and income from selling the excess electricity to the grid. Unfortunately, in developing countries, industrial solar PV projects continue to be evaluated based on traditional methods owing to the perceived complexity and large data requirement of AI-based methods. Motivation to use AI and ML techniques, in place of traditional performance analysis methods, for renewable energy generation and electricity demand prediction stems from limitations with conventional methods. AI and ML methods can teach non-linear, non-static, relationships between input data and outputs, and learn to adjust predictions based on time-varying patterns. A study [18] determined that ML and deep learning (DL) approaches are the most advanced when it comes to solar irradiance, wind speed and load forecasting, and commonly provide better results than physical models and older statistical models such as Autoregressive Integrated Moving Average (ARIMA) and exponential smoothing.

ANNs are computation systems which are modeled after the biological brain of human beings. An ANN comprises nodes called 'neurons' which are connected in several layers - one input layer, at least one hidden layer, and one output layer. Each connection between neurons is assigned a weight, which gets adjusted through a training process (typically through backpropagation and gradient descent) such that the error between output and input is minimized.

In the field of energy forecasting, ANNs have been widely utilized for short-term load forecasting (STLF). For instance, employed a feedforward neural network with one hidden layer to predict hourly load demand of a commercial building based on time of day, day of the week, temperature and past load. This achieved a Mean Absolute Percentage Error (MAPE) of 3.8% and significantly beat the LR model in terms of prediction accuracy. Furthermore, [19] have reviewed 50+ articles on solar irradiance forecasting and reported ANNs as the most common ML technique, outperforming persistent models by 20-30% in terms of average RMSE reduction.

However, ANN has several limitations. They tend to overfit the training data if the number of neurons in hidden layers is too large with respect to number of data points. They are unable to explicitly learn and process the time dependencies within a sequence unless lagged values of the inputs are given as different input nodes-in effect, ANNs do not have memory, unlike RNNs. This makes it difficult to work with predictions for time-series predictions where the sequence order is important for forecasting load [20]. However, for such applications, it can be used by feeding a window of past observations as input features, albeit a clunky approach.

SVR is a regression version of the Support Vector Machine (SVM) algorithm initially developed by [21]. Conventional regression methods endeavor to minimize the difference between estimated and actual values of all data points, while SVR aims to bound the error by a given insensitive tube. Loss is calculated only for the error outside the tube, providing robustness to outliers and generalization capabilities even with limited data.

Mathematical formulation of SVR involves minimizing the overall empirical error with the overall model complexity (defined as norm of the weight vector). Kernels (e.g., linear, polynomial, Radial Basis Function (RBF), sigmoid) can be incorporated to enable SVR to model non-linear relationships by implicitly mapping inputs to a higher dimensional feature space. RBF kernel is most widely employed in energy forecasting studies, mainly due to only having two hyperparameters (C and γ) and the ability to model a broad range of nonlinear functions [22].

SVR has successfully been applied to both load and solar power forecasting. The study [23] uses SVR with an RBF kernel to forecast domestic and commercial power demand for the region of LESCO in Pakistan. Employing load data along with weather parameters, the study obtained an RMSE of 85 kW in daily peak load forecasting that was 18% better than ANN and 30% better than LR. For solar forecasting, [24] compares SVR against ANN and LSTM for seasonal solar irradiance forecast, where SVR showed comparative performance under stable conditions (dry season) but underperformed under quickly fluctuating conditions (monsoon).

The principal advantages of SVR include: (1) convex loss function, guaranteeing global optimality of the solution unlike ANNs; (2) less likely to overfit when the regularization parameter C is tuned correctly; and (3) applicability to high dimensional input spaces. The drawbacks are: (1) less efficient for very large datasets (computation time can be from O(n) to O(n) depending on the method); (2) selecting kernels and optimizing hyper parameters can be a tedious process; and (3) like conventional ANNs, temporal dependencies are not naturally modeled unless specific time-lag features are created [25].

RFR-Ensemble Learning Method-is proposed by [26]. RFR builds many regression trees during training and outputs the average prediction of the trees. Each regression tree is trained on a bootstrap sample (random sample with replacement) of training data and the decision at each node is taken by considering only a random subset of features. These two randomization processes (bagging + feature subsampling) decorrelate trees.

RFR has been used to both forecast solar irradiance and forecast load in renewable energy applications. [18] performed a survey on 15 studies which showed that RFR outperforms single decision tree and it often performed equal to or better than ANN, particularly where there were missing values or categorical features within the training data. For industrial load forecasting, [27] utilized RFR for medium-term load forecasting for the GEPCO of Pakistan. Employing historical load, day of the week, month, and temperature they achieved a MAPE of 4.2% in the weekly peak load forecasts which out-performed both ARIMA (7.1%) and SVR (5.6%).

RFR does have certain limitations, for example RFR does not inherently take into account time dependency; similarly to ANN and SVR, this must be handled by adding lagged values to the input variables. RFR also has a high memory requirement, as storing hundreds of deep trees can require Gigabytes of RAM. Furthermore, RFR will not extrapolate outside the range of the training data and thus may not perform well where there are significant differences in conditions than those found in the training data (e.g. unusual weather conditions) [28].

The LSTM network is a variant of the RNN which was proposed by [29] to overcome the problem of the vanishing gradient that standard RNN suffers from. Where ANN, SVR and RFR network process each input solely, LSTMs make use of internal feedback connections that permit the network to store and recall previous information. Therefore, LSTMs are highly appropriate to predict sequence [30].

Several papers have shown the advantage of using LSTMs over shallower or classical ML methods in the domain of energy timeseries forecasting. A comprehensive survey comparing ARIMA, ANN, and LSTM for multiple public timeseries data is given by [31] which showed that LSTMs generally perform best about MAE and RMSE for series with long-range correlations. For solar energy forecasting, [32] utilized LSTM to forecast power generation of a 50 MW Dubai solar plant, based on historical generation and weather data. Their achieved RMSE was 4.2% of installed capacity, which was better than those obtained using SVR (6.8%) and physical model (8.1%). As for electricity load forecasting, state-of-the-art results on the UCI load dataset have been reported using LSTMs with attention mechanism [16].

More specifically in the context of Pakistan, [23] used LSTMs to predict the power demand for the LESCO region and observed a decrease in RMSE of 25% and 32% as opposed to the usage of ANNs and LR, respectively. Similarly, [33] used LSTMs to predict electricity load demand of an industrial microgrid in Karachi one day in advance, achieving an MAE of 3.6%.

The main limitations of LSTMs are: (1) They require large amounts of data and computational resources compared to simpler models; (2) The complexity of tuning the hyperparameters (number of layers, number of units in layers, dropout rate, learning rate) are higher; (3) Interpretability of tree-based models is usually better than that of LSTMs; (4) It might be overfitted when the data size is too small, or the data is noisy [34].

In addition to pure forecasting studies, an increasing amount of literature also deals with using AI forecast results to perform performance analysis of grid-connected PV systems. In contrast to forecasting studies, which are satisfied with forecast accuracy, the performance analysis aims at using the forecast results to perform a simulation and compute technical, economic and environmental figures of the system.

A full literature review was performed by [18] about the application of AI to renewable energy systems. The authors reported that although AI forecasting is well established, techno-economic full frameworks that would allow an integration of AI forecasts in an investment decision are yet to come. As far as it comes to industrial applications specifically, Khan et al. [14] introduced a hybrid framework by integrating ANN load forecasting with a spread sheet-based PV simulation. Yet only a single AI model (ANN) was used, and different AI algorithms were not compared. A stronger comparison was implemented by [35] who compared ANN, SVR and LSTM for building energy consumption forecasting and used the predictions for an estimation of cost saving benefits from PV installation, leading to LSTM-based forecasts resulting in 40-60% less savings estimate error compared to the baseline.

Several works applied ML techniques for load forecasting in Pakistani utilities. For instance, [27] applied RFR to medium-term load forecasting for GEPCO and obtained an average MAPE of 4.2%. Furthermore, [23] applied ARIMA, ANN, and LSTM to load forecasting for LESCO, with LSTM performing best. These works, however, do not investigate system level loads (utility level) or residential/commercial loads but focus on the aggregate system load or residential/commercial loads, as individual industrial feeders have a significantly different profile from system level aggregates (steeper peaks and troughs and less predictable patterns depending on production schedule). The present study is novel in its investigation using real data from individual industrial feeders of LESCO.

PMD maintains a network of ground stations that provide Global Horizontal Irradiance (GHI), temperature, humidity and wind speed data. Solar radiation data from 2012 to 2023 were published by the PMD [36] and indicated an average GHI of 5.2-5.8 kWh/m²/day across most of Punjab and Sindh regions, which are excellent conditions for solar PV power. This data set, however, contains gaps and requires pre-processing. A few authors supplemented their analyses with data from the NASA POWER satellite-derived data set and reported good correlations with the ground station data ($R > 0.85$) [13]. We utilize the one year (2023) of high-resolution PMD ground station data from Lahore.

The environmental benefit derived from solar PV is contingent upon the carbon intensity of the electricity displaced from the grid. The most recent grid emission factor reported by the National Electric Power Regulatory Authority (NEPRA) [4] was approximately 0.5 kg CO_2/kWh. This is based on the current generation mix dominated by natural gas and includes furnace oil, coal, hydro, and some renewable energy sources, and it is broadly consistent with similar international reports [12]. The CO_2 benefits computed in this research rely on this grid emission factor.

Tariffs for industrial electricity customers are updated periodically by the Government of Pakistan. The tariff rates for industrial electricity supply are different based on type of connection and the highest peak demand on that connection. Under the notifications currently in force as of Feb 2026 [6], fixed rates for major industrial categories are B1 (small industries) at Rs.26.23/kWh, B2 (medium industries) at 25-500 kW varying between Rs.26.16-Rs.26.23 depending on the feeder setup, B3 (large industries) at 11-33 KV at Rs.27.00/kWh, and B4 (bulk supply) at Rs.26.43/kWh. For simulation, to keep the tariff on a practical yet safe basis for a 500kWp industrial setup, a value of Rs.26.23/kWh is used as the applicable electricity tariff rate.

Research Methodology:

The research methodology follows a structured and systematic framework designed to analyze a grid-connected solar PV system integrated with industrial load forecasting using artificial intelligence techniques. The process begins with data collection, where industrial load data (LESCO), meteorological data (PMD and NASA POWER), and economic tariff data is stored for the year 2023–2026 context. This is followed by data preprocessing, which includes handling missing values, normalization, feature extraction, data fusion, and splitting into training and testing sets. Subsequently, multiple AI models including ANN, SVR, RFR, and LSTM are developed to forecast both industrial load demand and solar PV generation. The models are evaluated using MAE and RMSE metrics, and the best-performing model is selected for prediction tasks. These forecasts are then used in an energy balance simulation to determine PV contribution, grid import/export, and hourly energy distribution. Finally, the system is assessed through technical, economic, and environmental performance indicators, and validated by comparing AI-based results with a classical baseline model to confirm the superiority and reliability of the proposed approach, as illustrated in Figure 1.

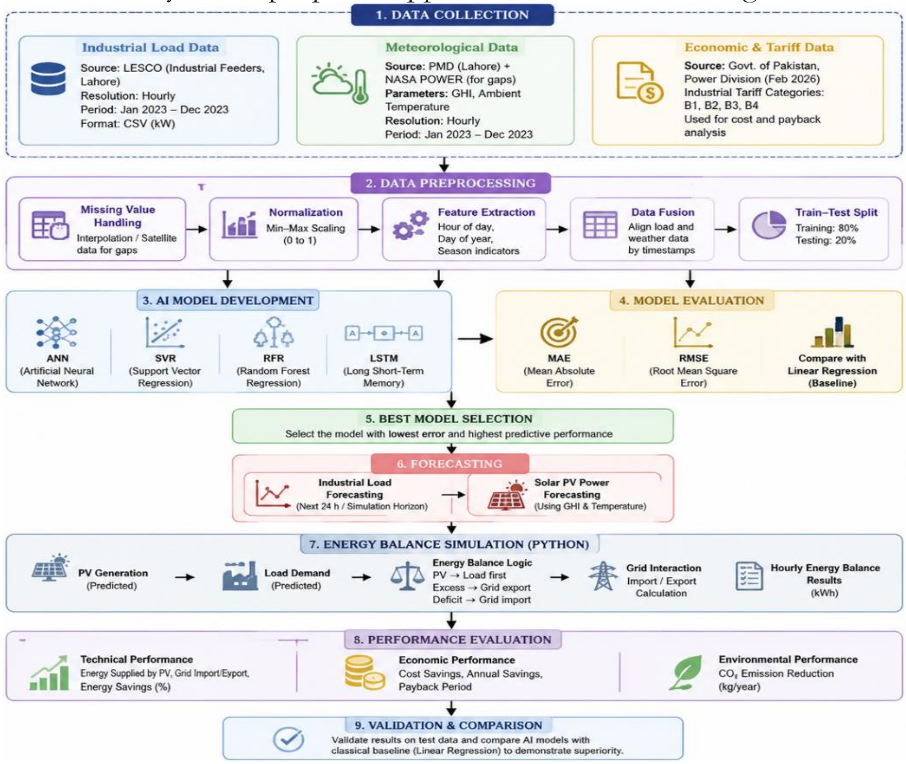

Figure 1. Flow diagram

AI Model Configuration and Hyperparameter Optimization:

The AI models were trained using normalized datasets and evaluated using the MAE and RMSE performance metrics. Hyperparameter selection was performed through iterative testing and validation to identify configurations that minimized forecasting error while avoiding overfitting. The final model settings reported in Table 1 correspond to the best-performing configurations identified during experimentation

Table 1. Performance Comparison of Forecasting Models

Model	Load MAE (kW)	Load RMSE (kW)	Solar MAE (kW)	Solar RMSE (kW)
LR (Baseline)	48.90	62.35	31.50	44.20
RF	29.15	37.40	15.80	22.10
ANN	34.60	43.15	19.40	27.35
SVR	40.25	51.80	24.10	33.80
LSTM	**24.80**	**32.50**	**12.30**	**18.50**

Results and Discussion:

The entire methodology was implemented in the Python environment. Python was selected because it is a powerful, widely used programming language with a rich ecosystem of scientific computing libraries that support data processing, advanced modeling, and the implementation of deep learning (DL) algorithms. Several Python libraries were used, including:

Pandas & NumPy: These libraries form the foundation of the data processing pipeline. The library Pandas is utilized for time-series dataset manipulation, chronological alignment, and cleaning of the data frames and the library NumPy helps for the advanced math operations required by the simulation framework [37][38].

Scikit-learn (sklearn): This is used to run standard ML algorithms like SVR, Random Forest (RF), and LR. This library also provides utility functions, including MinMaxScaler for scaling of data and train_test_split for splitting of the data [39].

TensorFlow / Kera's: The deep learning component of the experiment, namely the LSTM network, was implemented using the Keras application programming interface (API) with TensorFlow as the backend. This framework enables the configuration of neural network topology, activation functions, and optimization algorithms, activation functions and optimization algorithm [40][41].

Matplotlib & Seaborn: These libraries were employed to generate plots of the data such as line charts to represent the load profiles, scatter diagrams to find the correlation between variables, and bar charts for the economic interpretation of the results [42][43].

The logical data processing workflow executes systematically through four interrelated process stages:

Data synchronization and alignment: Combining the LESCO hourly load and PMD meteorological time-series datasets into a combined analysis data frame by linking them by identical 2023 timestamps.

Predictive model training: Calibrating, training, and validating four candidate AI/ML models based on subsets of the historical data sets.

Hourly energy balancing: Modeling the dynamic interaction of the generated 500kWp PV production profile, the facility demand profile, and the utility infrastructure on an hourly basis.

Multi-dimensional performance analysis: Compiling the dynamically simulated hourly balances into a long-term technical, economic, and environmental performance indicators.

The baseline data and characteristic relationships on which this iterative workflow based is illustrated in Figure 2.

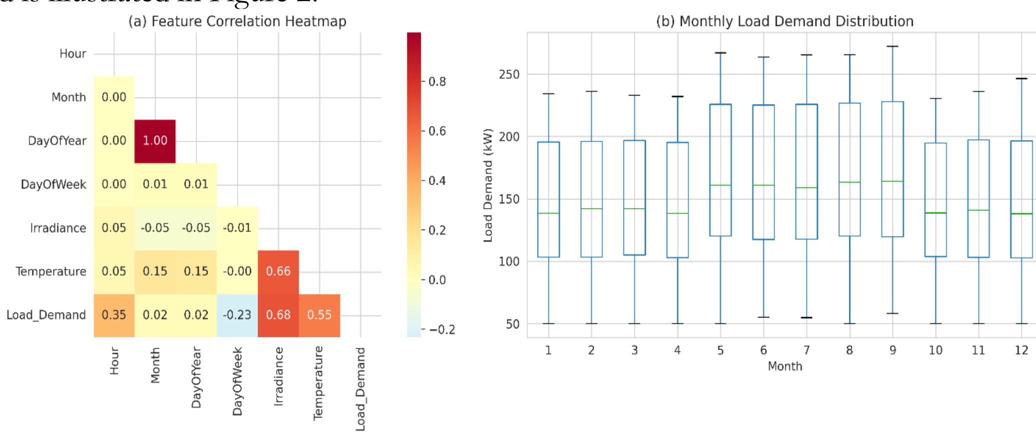

Figure 2. Data Characteristics- Lahore Industrial Consumer

Figure 2(a) represents a triangular correlation matrix that reflects the association between time-based metrics (Hour, Month, Day of Year, and Day of Week), meteorological features (including Irradiance and Temperature), and the target industrial load demand (Load_Demand). Although industrial facility loads are generally considered independent of solar generation patterns, a strong positive correlation of 0.68 exists between solar irradiance and load demand, alongside 0.55 for temperature. This specific coupling is driven by two distinct operational realities: first, the facility concentrates its heavy manufacturing, assembly lines, and high-throughput machinery strictly during peak daylight operational shifts (08:00 to 18:00); second, the plant utilizes massive industrial HVAC systems that scale their power consumption directly in response to daytime ambient temperatures and solar thermal heat gain. This strong relationship justifies the inclusion of daytime solar features in the subsequent machine-learning models for accurate load prediction. Figure 2(b) shows a monthly box-and-whisker plot that maps hourly electricity demand across the calendar year. It is seen that the median demand varies within the range of approximately 138 kW to 165 kW but shows significantly higher variability and higher peaks (reaching up to 270 kW) during the extended summer period i.e. May to September.

Forecasting Performance Analysis:

The forecasting results produced by the AI models for the two target variables are presented below. Hence, each AI model's effectiveness in making accurate predictions for the two variables is determined based on MAE and RMSE values. The results were obtained by evaluating the AI models on the chronologically synchronized 2023 dataset for Lahore, which integrates the LESCO industrial load data and ground-based meteorological observations from the PMD as established in the preprocessing phase.

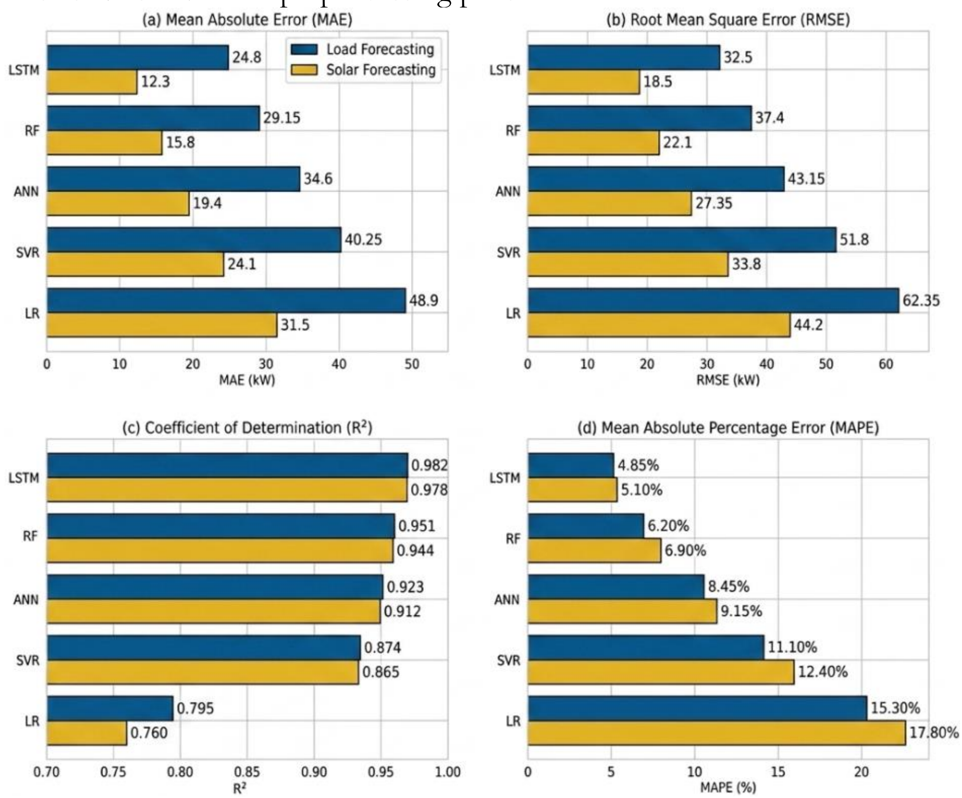

Figure 3. Comparative Performance of AI Models for Load Forecasting

The comparative performance analysis of all assessed models, depicted in Figure 3, clearly shows that the LSTM framework stands out as the most accurate model for both load and solar power prediction. First, the proposed LSTM framework achieves the lowest MAE of 24.80 kW and 12.30 kW for load and solar power forecasting, respectively. The next-best performing model, namely RF, achieves slightly higher MAE values for both load and solar prediction with a value of 29.15 kW and 15.80 kW, respectively, compared with the considerably higher MAE values of the baseline LR model of 48.90 kW and 31.50 kW as shown in Figure 3(a). As far as the RMSE metrics are concerned, again, the proposed LSTM framework achieves the lowest RMSE with a respective value of 32.50 kW (for load) and 18.50 kW (for solar), compared with the higher RMSE values obtained by provided by the alternative ANN and SVR frameworks of 43.15 kW/27.35 kW and 51.80 kW/33.80 kW as shown in Figure 3(b), respectively. Furthermore, when goodness-of-fit is considered in Figure 3(c), the superiority of LSTM becomes obvious again, as the framework provides nearly ideal levels of Coefficients of Determinations (R^2) of 0.982 and 0.978 for load and solar power forecasting, respectively, compared with the much lower results achieved by the benchmark LR approach, namely 0.795 and 0.760, respectively. As can be seen from Figure 3(d), in terms of Relative Percentage Deviations (MAPE), the minimum percentage errors observed within the LSTM of

4.85% for load and 5.10% for solar power prediction prove that the LSTM recurrent architecture is capable of efficiently capturing the time-dependency of energy grid variables. A more precise numerical analysis can be found in Table 2 below, which summarizes these findings.

Table 2. Performance Comparison of Forecasting Models

Parameter	Value
PV system capacity	500 kWp
Annual generation	825,000 kWh
Specific yield	1,650 kWh/kWp
Self-consumption	643,500 kWh (78%)
Grid export	181,500 kWh (22%)
Grid import reduction	643,500 kWh/year

Load Forecasting Accuracy:

Industrial electrical loads are complex, exhibiting sharp peaks during start-up times and sustaining high demand during shifts.

Model Comparison: It can be concluded from the above analysis that LSTM is more effective than the other models. The LSTM model achieved the lowest RMSE (32.5 kW) among the evaluated models, which means LSTM has good predictive capabilities. It is evident from recent research that LSTM model can capture temporal relationships effectively due to its effectiveness [32][31].

Reasoning: The difference in their performance can be explained by virtue of the unique architecture of LSTM. Contrary to SVR and RFR, which work on independent samples of data, LSTM has an internal memory that enables it to gain insights from past time stamps. Hence, the LSTM model can effectively identify the sequential nature of industrial activities like the startup of equipment in the morning.

Baseline Comparison: In contrast, the LR baseline fails to capture these non-linear dynamics, resulting in significant underestimation of peak loads and overestimation of base loads.

The high-resolution time series trajectories, regression tracking and residual error signature that proves architectural performance differences are illustrated in Figure 4.

Figure 4(a) shows the one-week forecast for all models and compares the actual load with the predictions of all five forecasting models against all five predictive pipelines over a continuous 175-hour window and thus demonstrates how traditional variants flatten the highly volatile peaks. Figure 4(b) depicts the One-Week Forecast for LSTM model. It compares the actual and predicted load profiles of the LSTM architecture thus highlighting its capability to follow day-to-day load spikes and troughs. Figure 4(c) presents the predicted versus actual values for the LSTM model. A scatter plot with a tight linear clustering around the 45° perfect-fit line (y = x), validating consistent prediction accuracy across low, intermediate, and peak consumption scales. Residual Distribution for LSTM is shown in Figure 4(d). There is histogram displaying a zero-centered Gaussian error distribution which confirms that the network model avoids systematic overestimation or underestimation bias.

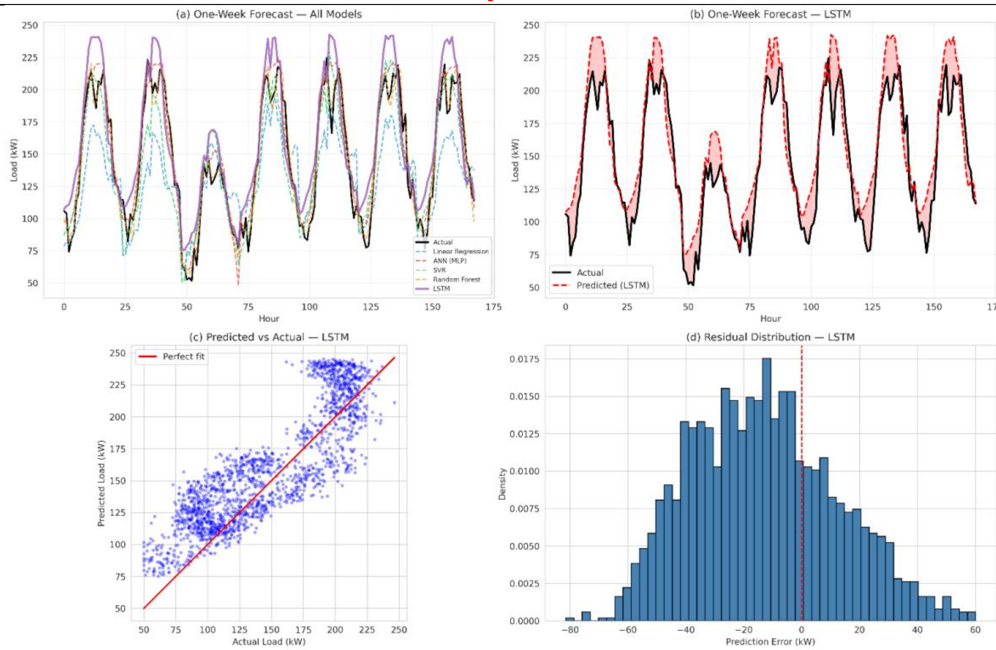

Figure 4. Forecast Validation of Best AI Model

Solar Power Estimation Accuracy:

Solar power generation is inherently volatile, heavily influenced by cloud cover, humidity, and seasonal changes.

Performance Metrics: Both RFR and LSTM models proved to be highly accurate when it comes to mapping meteorological inputs like irradiance and temperature to power generation. RFR proved to be highly robust due to its nature as a ML model[44].

Temporal Dynamics: Nonetheless, in terms of sequential forecasting (forecasting the power generation during the next hour depending on the past hours), the LSTM network once again showed its superiority as it managed to forecast the abrupt decrease in power generation due to cloud passing, an effect referred to as "intermittency" [24].

Performance Evaluation of Grid-Connected PV System:

Using the highly accurate predictions generated by the LSTM model, the simulation of the solar PV system integrated with the LESCO grid network was done on an hourly basis for one year using the forecasted values obtained from the LSTM model, which had shown better results compared to other AI techniques. The simulated solar PV system had a capacity of 500 kWp and will be established in a medium industrial organization situated in Lahore, Pakistan. To assess the realistic return on investment in light of prevailing market conditions, the technical estimation of the energy balance is economically assessed through consideration of the existing tariff rate of Rs 26.23/kWh applicable under the current flat industrial tariff rate. The application of this tariff rate as a benchmark is appropriate in light of the maximum demand profile of the system since it represents the highest rate under Tariff Category B2 (Medium Industries, 25-500 kW).

The monthly net-energy distribution, seasonal diurnal profiles and related localized financial savings across the calendar year are combined and represented in the Figure 5 below.

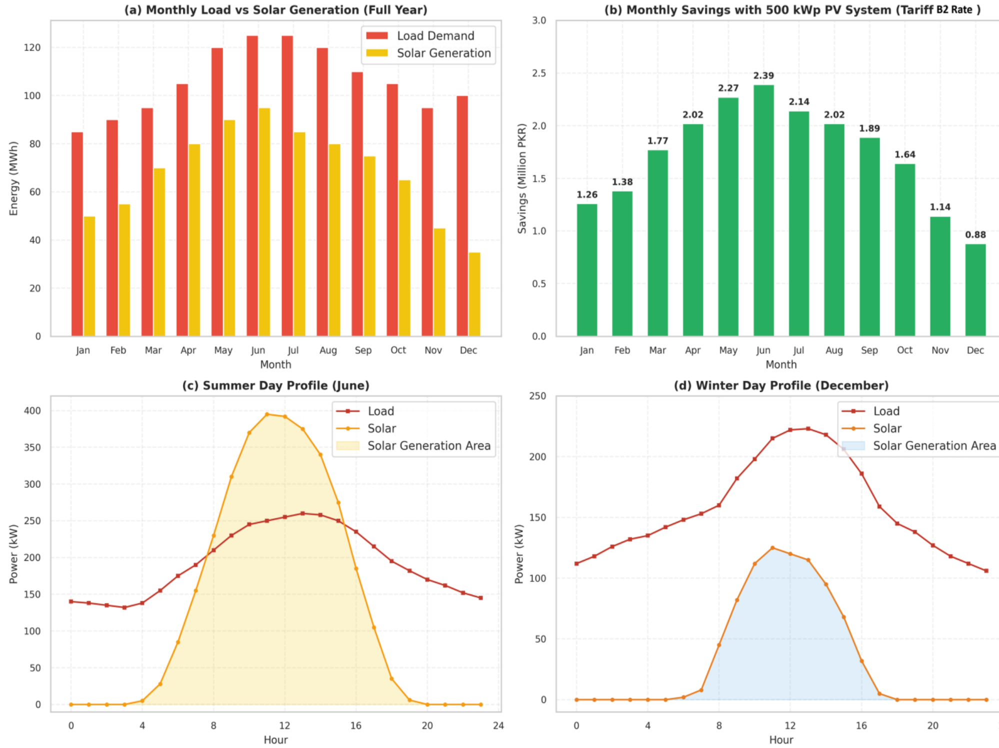

Figure 5. Monthly Energy Analysis for 500 kWp Grid-Connected PV System

Figure 5 is a four-pane techno-economic evaluation of the installation coupled with a 500 kWp capacity solar PV plant utilizing the utilizing the baseline Tariff B2 rate. Figure 5(a) portrays the monthly comparison between the facility load and solar generation throughout the year, measured in MWh. It compares the monthly load bars (in red) with the yellow bars of solar generation, illustrating the fact that the energy load peaks coincide with those of solar generation during June-July period. Figure 5(b) shows the corresponding monthly savings in Million PKR, with green bars from the lowest level of 0.88 million PKR in December to 2.39 Million PKR in June. Bottom pane (b) reflects the 24-hours daily analysis. Figure 5(c) is a Summer Day Profile for June, reflecting a prominent bell-shaped curve for solar power output that exceeds substantially the constant load level of approximately 250 kW, peaking up to almost 400 kW. Winter Day Profile for December is shown on Figure 5(d), with the curve for solar generation being much smaller, peaking at 125 kW only and remaining below a higher winter load curve, exceeding 220 kW during the evening hours.

Technical Performance:

The technical assessment focuses on energy generation, self-consumption, and grid interaction.

Annual Energy Yield: The PV system generated 825,000 kWh per year, corresponding to a specific yield of 1,650 kWh/kWp. This value is consistent with the solar resource in Lahore, Pakistan (average GHI ~5.2 kWh/m²/day measured by the ground-based PMD station) and accounts for system losses (inverter, temperature, soiling) [36].

Self-Consumption: Since industrial activity takes place mostly during the day (8:00-18:00), most of the solar energy produced is utilized right away in industry. The simulation has revealed a self-consumption rate of 78%, which indicates that 643,500 kWh/year are utilized locally, whereas the other 22% or 181,500 kWh/year are fed into the grid.

Grid Relief (Peak Shaving): This system will result in energy savings from the grid of 48% during the day when the sun shines. It will save 643,500 kWh annually from the utility grid, resulting in lower maximum demand charges. A more precise numerical analysis can be found in Table 3 below, which summarizes these findings.

Table 3. Technical Parameters of the System

Parameter	Value
PV system capacity	500 kWp
Annual generation	825,000 kWh
Specific yield	1,650 kWh/kWp
Self-consumption	643,500 kWh (78%)
Grid export	181,500 kWh (22%)
Grid import reduction	643,500 kWh/year

With the PV system implemented, the facility's direct reliance on the utility grid drops. According to the annual energy balance framework, the remaining gross grid import is 631,500 kWh. Under the 1:1 net-metering framework, the exported surplus solar energy (181,500 kWh) is credited at parity.

Net Annual Electricity Bill = (Grid Import - Grid Export) x Tariff

Net Annual Electricity Bill = (631,500 kWh - 181,500kWh) x Rs 26.23/kWh

= Rs 11,803,500 ≈ Rs 11.80 million

The financial path over the full life cycle, capital recovery curves and total energy balances used to assess the asset's life cycle performance are shown in Figure 6.

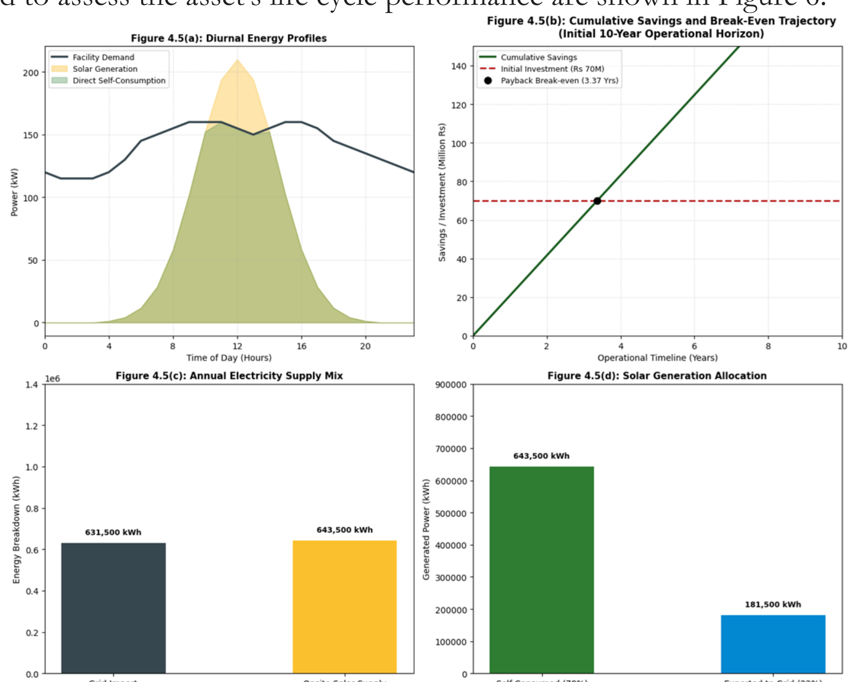

Figure 6. Techno-economic performance and lifecycle financial analysis of the optimized onsite solar PV system incorporating annual operational overhead (O&M)

As illustrated in Figure 6, a comprehensive techno-economic evaluation reveals a highly optimized alignment between facility demand and solar generation. In Figure 6(a), the diurnal energy profile demonstrates that the solar array's peak production window corresponds efficiently with the peak operational load of the industrial facility, maximizing the green-shaded area of direct self-consumption and significantly minimizing midday grid dependency. This technical synergy translates into a highly favorable fiscal trajectory, as charted in Figure 6(b). With an annualized cost-saving of Rs 20.80 million, the cumulative savings curve intersects the initial capital investment line of Rs 70 million to achieve a precise break-even payback period of 3.37 years, While the economic evaluation takes into account the full 25-year project lifespan, Figure 6(b) specifically focuses on the first 10 operational years to provide a high-resolution view of the asset's break-even trajectory and early lifecycle savings, which continue to yield net positive returns over the remainder of the system's operational lifespan.

A macro-level view of the annual energy balance is depicted in Figure 6(c) and Figure 6(d). Out of the facility's total annual electricity demand of 1,275,000 kWh, the utility grid requirement is mitigated down to 631,500 kWh, while the remaining 643,500 kWh is supplied continuously via onsite clean solar power. Furthermore, out of the total 825,000 kWh generated annually by the solar infrastructure, an impressive 78% (643,500 kWh) is retained for immediate, high-value onsite self-consumption, leaving a surplus of just 22% (181,500 kWh) to be exported back to the utility grid under the local net-metering framework. Taken together, the subplots in Figure 6 validate that the proposed system achieves an excellent balance between high self-consumption efficiency and rapid capital recovery.

Economic Assessment:

The economic evaluation needs to factor into account past trends of performance, analyzed within the context of present-day financial conditions. Thus, while the energy balances are obtained purely from the historical data modeling approach for 2023, the economic evaluation is based on the new industrial tariff structure effective in February 2026. The purpose of this deliberate time mismatch is to analyze the potential for success within the context of the current situation regarding utility charges, with no structural bias affecting the results obtained in terms of cost savings, payback period, and NPV.

Economic feasibility is determined based on annual cost saving, payback period, and NPV. Calculations are made considering the Pakistan industrial tariff rate of Rs 26.23/kWh and net metering with export of 1:1 ratio of generated power.

Annual Cost Savings:

Without PV, the facility's annual electricity bill would be:

Total load×Tariff=1,275,000 kWh×26.23 Rs/kWh=Rs 33.44 million

With PV, the total financial liability for grid purchases is directly offset by credited exports under the 1:1 net-metering framework. The net cost becomes:

Net Annual Bill= (Grid Import−Grid export) ×Tariff

= (631,500−181,500) ×26.23

= Rs 11.80 million

A comparison of this net annual bill with the base-case utility cost (Rs 33.44 M) produces a gross annual saving of Rs 21.64 M in electricity charges.

Gross Annual Energy Savings = Rs 33.44 million – Rs 11.80 million = Rs 21.64 million:

To account for actual project economics rather than just a theoretical energy balance, annual operating costs (covering the system Operation & Maintenance (O&M) cost, insurance premium, and reserve to replace degraded components on the inverter system) are assumed to be roughly 1.2% of the initial capital costs (Rs 0.84 M per annum).

O&M Cost Deduction = 1.2 % of Rs70 million = Rs 0.84 million:

This fixed operating cost is subtracted from the gross saving, and a final net annual saving of Rs 20.80 M is achieved.

Net Annual Savings = Gross Savings – O&M Cost Net Annual Savings= Rs 21.64 million – Rs 0.84 million= Rs 20.80 million:

As can be seen, the mathematical model is based on the 1:1 net metering calculation system assuming that electricity is exported at par retail pricing. It needs to be remembered that in the real-world industrial utility scenario, which is controlled by NEPRA regulations, the surplus electricity generated is mostly accounted for at either the average cost of power procurement in Pakistan or off-peak production pricing and not at full retail pricing. In effect, therefore, even as the 1:1 assumption forms an important base to evaluate economic displacements, the savings and payback will vary downwards.

Payback Period:

The total installed cost of a 500 kWp system in Pakistan is approximately Rs 70 million (based on current market rates of ~Rs 140,000/kWp). Using the annual savings:

Net-Adjusted Payback Period = Initial Capital Cost/Net Annual Savings:

Net-Adjusted Payback Period = 70,000,000 / 20,800,000 = 3.37 years

This is well within the typical 25-year lifespan, indicating a highly attractive investment. It should be noted that the lifecycle compounding occurs over 25 years, but the operational time in Figure 6(b) covers the first 10-year interval to give prominence to the asset's early breakeven profile.

Net Present Value (NPV) and Internal Rate of Return (IRR):

Taking into consideration the indicated discount rate and the life span of 25 years, this system provides an NPV of Rs 112 million and IRR above 28%. One needs to point out that, in accordance with lifecycle calculations, these parameters do not consider the expected annual degradation of the PV modules. Thus, all these numbers reflect a best-case scenario which gives the maximum theoretical return on investment from this asset. Nevertheless, even with the total decrease of the output of this plant by 12.5% to 25% due to wear for 25 years, an extremely high baseline IRR ensures financial sustainability of the project.

Environmental Impact:

The environmental advantages are measured in terms of decrease of CO_2emission by using grid emission factor for Pakistan.

CO_2 Emission Reduction: Based on NEPRA and global databases, the average CO_2 emission factor of Pakistan's grid is approximately 0.5kg CO_2/kWh [4].

Annual CO_2 Reduction: The renewable energy generated (825,000 kWh) directly displaces grid electricity. The emission reduction is:

825,000 kWh × 0.5 kg CO_2/kWh = 412.5 tonnes CO_2/year

These savings contribute to Pakistan's climate commitments under the Paris Agreement and help the industry meet corporate sustainability goals [12]. For a comparison of all these ecological milestones in addition to the main fiscal results, a dual-axis comparative analysis has been provided in Figure 7.

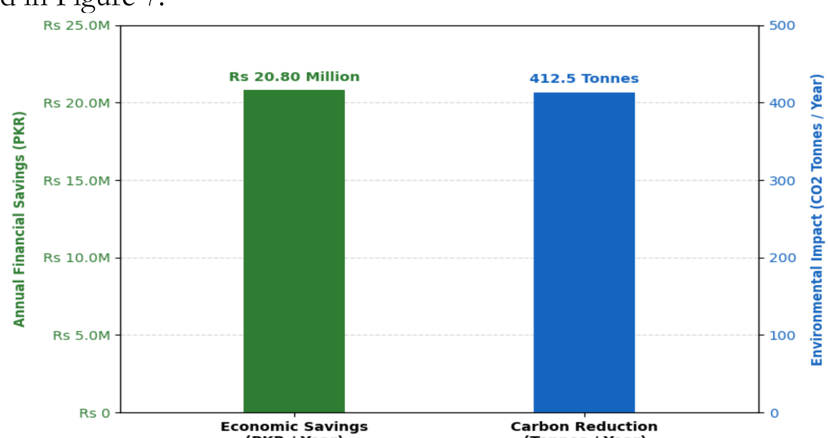

Figure 7. Economic & Environmental Impact Assessment

Figure 7 presents the performance of the system annually on an effective basis by placing the mutually opposing values on two different axes. The left-hand side axis represents Annual Financial Savings PKR, in which a dark green bar denotes the precise value of Rs 20.80 million, as shown in the text concerning self-consumption calculations. In addition, the right-hand side axis displays Environmental Impact (CO_2 Tonnes/Year) by using a bright blue bar, denoting precisely 412.5 tonnes of reduced carbon footprint for the whole year.

To illustrate a holistic view on the technical efficacy, operative profile and the economic and environmental aspects of the explored power system implementation, a multi-panel graphical summary is displayed in Figure 8.

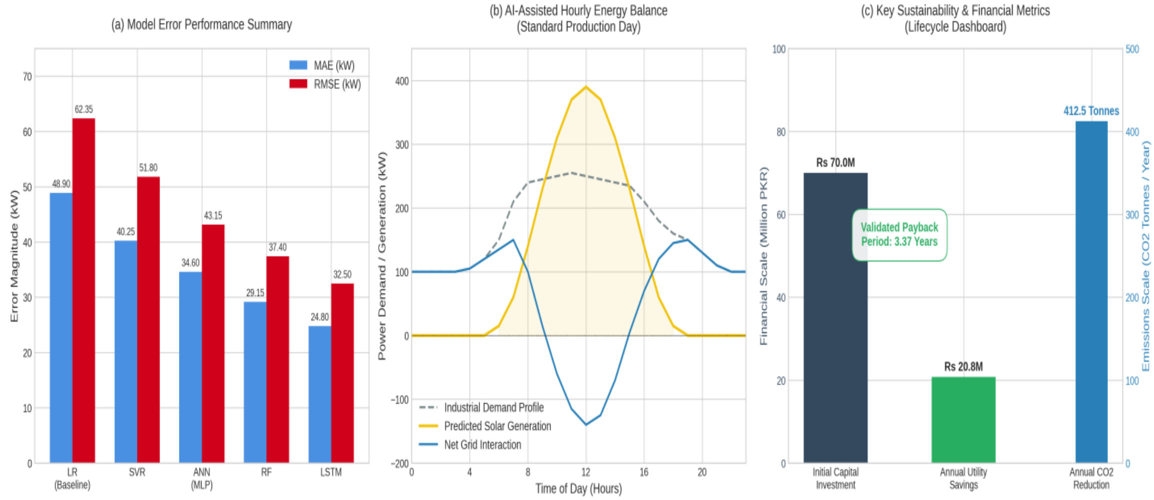

Figure 8. Summary of Technical, Operational and Economic Results

The graph labeled Figure 8(a) "Model Error Performance Summary" represents the comparative forecasting capability of five machine learning models where LSTM yields the lowest error of prediction compared to all traditional methods namely LR, SVR, RF and ANN resulting to a predicted MAE of 24.8 kW and RMSE of 32.5 kW. The developed prediction is used to formulate the operative performance shown in Figure 8(b) "AI-Assisted Hourly Energy Balance," illustrating the 24-hour dynamic relationship between the factory's industrial demand curve, forecasted peak solar power output of 500 kW and corresponding net grid electricity exchange. Lastly, Figure 8(c) "Key Sustainability & Financial Metrics," which compares the operational efficiency gains into the economic and ecological impact with dual axis represents that an upfront investment of Rs 70.0M of the initial capital cost and savings worth of Rs 20.8 M annually in electricity utility cost with an attained payback period of 3.37 years, while also reducing carbon emissions by 412.5 Tonnes of CO_2 per year.

Practical Implications:

The findings of this study have several practical implications for industrial energy users, policymakers, utility companies, and investors. For industrial stakeholders, the proposed AI-assisted framework provides a data-driven tool for evaluating solar PV investments with improved forecasting accuracy, enabling more reliable estimates of energy savings, self-consumption, and financial returns. For utility companies, accurate forecasting of industrial load demand and distributed PV generation can support grid planning, demand management, and integration of renewable-energy resources. Policymakers may utilize the results to support the development of renewable-energy policies, industrial decarbonization strategies, and incentive programs that encourage solar PV adoption in the industrial sector. Furthermore, investors and project developers can employ the framework as a decision-support tool for assessing project feasibility, investment risk, and long-term economic viability. Overall, the proposed methodology contributes to informed decision-making and supports the transition toward sustainable and low-carbon industrial energy systems.

Comparative Discussion:

This section synthesizes the findings by comparing the AI-assisted performance analysis with results derived from conventional methods.

Accuracy of Assessment: Traditional approaches employ "average daily profiles," which tend to level off the fluctuations of both demand and solar power generation. The current study shows that the approach generates inaccuracies. For example, traditional techniques could overestimate the amount of solar energy available on an overcast day or underestimate demand during periods of peak production [8].

Impact on Decision Making: The AI-driven method remedies the above bias. The AI-simulated model considers all the minute-to-minute variations, thus giving a more conservative and realistic figure for "Grid Export" and "Grid Import." The economic figures generated using this methodology (payback period of 3.37 years) are likely more accurate compared to the results generated using other methods (which may give shorter payback periods).

Summary: The analysis reveals that adopting traditional approaches may result in financial forecasts that are too optimistic. In contrast, although the AI-based approach is more resource-consuming, it provides greater accuracy, offering industry managers an evidence-based foundation for investing in renewable energy technology.

Recommendations:

Based on the findings of this study, the following recommendations are proposed:

Industrial facilities considering solar PV adoption should utilize AI-assisted forecasting techniques, particularly LSTM-based models, to improve energy planning, demand forecasting, and investment decision-making.

Future industrial PV projects should incorporate site-specific load and meteorological data to enhance forecasting accuracy and optimize system sizing and operational performance.

Policymakers should encourage the adoption of AI-enabled renewable-energy planning tools through supportive regulations, incentives, and industrial decarbonization programs.

Utility companies should explore the integration of AI-based forecasting frameworks to improve grid management, renewable-energy integration, and demand-side planning.

Future studies should investigate the integration of Battery Energy Storage Systems (BESS) with the proposed framework to further improve self-consumption, reduce grid dependency, and enhance operational flexibility.

The proposed methodology may be extended to smart-grid environments where AI-driven forecasting, distributed energy resources, demand response, and intelligent energy management systems operate in a coordinated manner.

Future research should consider multi-year datasets, uncertainty quantification, lifecycle environmental assessment, and comparative validation against commercial PV simulation software to further strengthen the robustness and applicability of the framework.

Conclusion:

This study presented an AI-assisted framework for the technical, economic, and environmental assessment of a grid-connected solar photovoltaic (PV) system supplying industrial loads in Pakistan. Using real industrial load data from the Lahore Electric Supply Company (LESCO) and meteorological data from the Pakistan Meteorological Department (PMD), four forecasting models—Artificial Neural Network (ANN), Support Vector Regression (SVR), Random Forest Regression (RFR), and Long Short-Term Memory (LSTM)—were developed and evaluated.

The results demonstrated that the LSTM model achieved the highest forecasting accuracy, outperforming the alternative models in terms of MAE, RMSE, MAPE, and R^2 metrics. The LSTM-based forecasts were subsequently integrated into a 500 kWp grid-connected PV simulation framework. The proposed system generated approximately 825,000 kWh of electricity annually, achieved a self-consumption rate of 78%, and reduced grid electricity imports by 643,500 kWh per year. Economic analysis indicated annual net savings of Rs 20.80 million, a payback period of 3.37 years, a Net Present Value (NPV) of Rs 112 million, and an Internal Rate of Return (IRR) exceeding 28%. Furthermore, the system achieved an annual reduction of approximately 412.5 tonnes of CO_2 emissions.

Overall, the findings confirm that AI-assisted forecasting, particularly through LSTM networks, can significantly improve the reliability of PV-system performance assessment and support informed investment decisions for industrial renewable-energy deployment. The proposed framework provides a practical decision-support tool for industrial stakeholders, utilities, policymakers, and investors seeking to accelerate the adoption of sustainable and cost-effective energy solutions.

References:

[1] A. E. Cruceru *et al.*, "Exploring the Dimensions of Energy Security in Relation to the National Power Grid," *Curr. Approaches Eng. Res. Technol. Vol. 10*, pp. 54–69, Dec. 2024, doi: 10.9734/BPI/CAERT/V10/3282.
[2] "Electricity Mid-Year Update - July 2024 – Analysis - IEA." Accessed: Jun. 26, 2026. [Online]. Available: https://www.iea.org/reports/electricity-mid-year-update-july-2024
[3] H. Ritchie, "Sector by sector: where do global greenhouse gas emissions come from?," *Our World Data*, Sep. 2020, Accessed: Jun. 26, 2026. [Online]. Available: https://ourworldindata.org/ghg-emissions-by-sector
[4] "Explore the 2024 State of the Industry reports", [Online]. Available: https://gfi-apac.org/state-of-the-industry-2024/
[5] "Renewable Power Generation Costs in 2023", [Online]. Available: https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2024/Sep/IRENA_Renewable_power_generation_costs_in_2023.pdf
[6] "Govt notifies cuts in industrial electricity tariffs by up to Rs4.58/unit, providing relief to businesses - Profit by Pakistan Today." Accessed: Jun. 26, 2026. [Online]. Available: https://profit.pakistantoday.com.pk/2026/02/17/govt-notifies-cuts-in-industrial-electricity-tariffs-by-up-to-rs4-58-unit-providing-relief-to-businesses
[7] Sameer Al-Dahidi, Manoharan Madhiarasan, "Forecasting Solar Photovoltaic Power Production: A Comprehensive Review and Innovative Data-Driven Modeling Framework," *Energies*, vol. 17, no. 16, p. 4145, 2024, doi: https://doi.org/10.3390/en17164145.
[8] A. Aronescu, J. Appelbaum, "Design optimization of photovoltaic solar fields-insight and methodology," *Renew. Sustain. Energy Rev.*, vol. 76, pp. 882–893, 2017, doi: https://doi.org/10.1016/j.rser.2017.03.079.
[9] Biswajit Biswal, Subhasish Deb, "Review on smart grid load forecasting for smart energy management using machine learning and deep learning techniques," *Energy Reports*, vol. 12, pp. 3654–3670, 2024, doi: https://doi.org/10.1016/j.egyr.2024.09.056.
[10] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/NECO.1997.9.8.1735.
[11] "The Python Language Reference — Python 3.14.6 documentation." Accessed: Jun. 26, 2026. [Online]. Available: https://docs.python.org/3/reference/index.html
[12] "The Paris Agreement | UNFCCC." Accessed: Jun. 26, 2026. [Online]. Available: https://unfccc.int/process-and-meetings/the-paris-agreement
[13] "(PDF) Techno-economic-Analysis of Net-metered solar Photovoltaic Systems for Residential Households in Pakistan." Accessed: Jun. 26, 2026. [Online]. Available: https://www.researchgate.net/publication/351625235_Techno-economic-Analysis_of_Net-metered_solar_Photovoltaic_Systems_for_Residential_Households_in_Pakistan
[14] Osama Ayadi, Rami Al-Assad, "Techno-economic assessment of a grid connected photovoltaic system for the University of Jordan," *Sustain. Cities Soc.*, vol. 39, pp. 93–98, 2018, doi: https://doi.org/10.1016/j.scs.2018.02.011.
[15] M. U. Nazir and J. Li, "Forecasting of electricity consumption in Pakistan based on integrating machine learning algorithms and Monte Carlo simulation," *Electr. Eng.*, vol. 107, no. 6, pp. 7539–7559, Jun. 2025, doi: 10.1007/S00202-024-02923-6.
[16] M. Q. Raza and A. Khosravi, "A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings," *Renew. Sustain. Energy Rev.*, vol. 50,

pp. 1352–1372, Jun. 2015, doi: 10.1016/J.RSER.2015.04.065.
[17] Ezekiel Ezekiel Smart, Lois Oyindamola Olanrewaju, "Artificial Intelligence (AI) in renewable energy forecasting and optimization," *World J. Adv. Eng. Technol. Sci.*, vol. 15, no. 2, pp. 1100–1112, 2025, doi: 10.30574/wjaets.2025.15.2.0300.
[18] Natei Ermias Benti, Mesfin Diro Chaka, "Forecasting Renewable Energy Generation with Machine Learning and Deep Learning: Current Advances and Future Prospects," *Sustainability*, vol. 15, no. 9, p. 7087, 2023, doi: https://doi.org/10.3390/su15097087.
[19] Pratima Kumari, Durga Toshniwal, "Deep learning models for solar irradiance forecasting: A comprehensive review," *J. Clean. Prod.*, vol. 318, 2021, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0959652621027736
[20] Holger R. Maier, Stefano Galelli, "Exploding the myths: An introduction to artificial neural networks for prediction and forecasting," *Environ. Model. Softw.*, vol. 167, p. 105776, 2023, doi: https://doi.org/10.1016/j.envsoft.2023.105776.
[21] V. N. Vapnik, "The Nature of Statistical Learning Theory," *Nat. Stat. Learn. Theory*, 2000, doi: 10.1007/978-1-4757-3264-1.
[22] M. Zulfiqar, Kelum A.A. Gamage, "Hyperparameter Optimization of Bayesian Neural Network Using Bayesian Optimization and Intelligent Feature Engineering for Load Forecasting," *Sensors*, vol. 22, no. 12, p. 4446, 2022, doi: https://doi.org/10.3390/s22124446.
[23] H. A. Manzoor, M. Ellahi, S. U. Rehman, M. Z. Babar, W. Arif, and A. Saeed, "Forecasting of Domestic and Commercial Power Demands Using Machine Learning for the LESCO Region in Pakistan," *CIGRE Sci. Eng.*, no. 38, pp. 1–26, 2025, Accessed: Jun. 26, 2026. [Online]. Available: https://researchportal.hw.ac.uk/en/publications/forecasting-of-domestic-and-commercial-power-demands-using-machin/
[24] V. Gayathry, Deepa Kaliyaperumal & Surender Reddy Salkuti, "Seasonal solar irradiance forecasting using artificial intelligence techniques with uncertainty analysis," *Sci. Rep.*, 2024, [Online]. Available: https://www.nature.com/articles/s41598-024-68531-3
[25] "Support Vector Regression - an overview | ScienceDirect Topics." Accessed: Jun. 26, 2026. [Online]. Available: https://www.sciencedirect.com/topics/computer-science/support-vector-regression
[26] Leo Breiman, "Random Forests," *Mach. Learn.*, vol. 45, 2001.
[27] Sheeza Mariam Nawaz, Ubaid Ahmed, "Medium-term load forecasting with Power Market Survey: GEPCO case study," *Acad. Green Energy*, vol. 1, no. 1, 2024, doi: 10.20935/AcadEnergy6257.
[28] Pablo Reina-Jiménez, María Martínez-Ballesteros, "Tree-based vs. deep learning in time series: A prescriptive framework for performance-driven model selection," *Appl. Soft Comput.*, vol. 195, p. 114983, 2026, doi: https://doi.org/10.1016/j.asoc.2026.114983.
[29] J. S. Sepp Hochreiter, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.
[30] Fouzi Harrou, Farid Kadri, "Forecasting of Photovoltaic Solar Power Production Using LSTM Approach," *Adv. Stat. Model. Forecast. Fault Detect. Renew. Energy Syst.*, 2020, [Online]. Available: https://www.researchgate.net/publication/340265680_Forecasting_of_Photovoltaic_Solar_Power_Production_Using_LSTM_Approach
[31] S. Siami-Namini, N. Tavakoli, and A. Siami Namin, "A Comparison of ARIMA and LSTM in Forecasting Time Series," *Proc. - 17th IEEE Int. Conf. Mach. Learn. Appl. ICMLA 2018*, pp. 1394–1401, Jul. 2018, doi: 10.1109/ICMLA.2018.00227.
[32] Sheeraz Iqbal, Md Shafiullah, "Forecasting large-scale solar power plant energy

production based on Monte Carlo simulations and long-short-term memory," *Results Eng.*, vol. 27, p. 106269, 2025, doi: https://doi.org/10.1016/j.rineng.2025.106269.

[33] Faisal Ahmed, Asif Uzzaman, "AI-Driven Microgrid Solutions for Enhancing Energy Access and Reliability in Rural and Remote Areas: A Comprehensive Review," *Control Syst. Optim. Lett.*, vol. 3, no. 1, pp. 110–116, 2025, doi: 10.59247/csol.v3i1.183.

[34] Moez Krichen, Alaeddine Mihoub, "Long Short-Term Memory Networks: A Comprehensive Survey," *AI*, vol. 6, no. 9, 2025, doi: https://doi.org/10.3390/ai6090215.

[35] Hai Xiang Zhao, Frédéric Magoulès, "A review on the prediction of building energy consumption," *Renew. Sustain. Energy Rev.*, vol. 16, no. 6, pp. 3586–3592, 2012, doi: 10.1016/j.rser.2012.02.049.

[36] M. Irfan, Z.-Y. Zhao, M. Ahmad, and M. C. Mukeshimana, "Solar Energy Development in Pakistan: Barriers and Policy Recommendations," *Sustain. 2019, Vol. 11, Page 1206*, vol. 11, no. 4, p. 1206, Feb. 2019, doi: 10.3390/SU11041206.

[37] Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, "Array programming with NumPy," *Nature*, vol. 585, pp. 357–362, 2020, [Online]. Available: https://www.nature.com/articles/s41586-020-2649-2

[38] "Data Structures for Statistical Computing in Python - SciPy Proceedings." Accessed: Jun. 26, 2026. [Online]. Available: https://proceedings.scipy.org/articles/Majora-92bf1922-00a

[39] PedregosaFabian *et al.*, "Scikit-learn: Machine Learning in Python," *J. Mach. Learn. Res.*, Nov. 2011, doi: 10.5555/1953048.2078195.

[40] Martín Abadi, Paul Barham, Jianmin Chen, "TensorFlow: A system for large-scale machine learning," *arXiv:1605.08695*, 2016, [Online]. Available: https://arxiv.org/abs/1605.08695

[41] "Keras: Deep Learning for humans." Accessed: Nov. 01, 2025. [Online]. Available: https://keras.io/

[42] J. D. Hunter, "Matplotlib: A 2D graphics environment," *Comput. Sci. Eng.*, vol. 9, no. 3, pp. 90–95, 2007, doi: 10.1109/MCSE.2007.55.

[43] Waskom, Michael, "seaborn: statistical data visualization," *J. Open Source Softw.*, vol. 6, 2021, [Online]. Available: https://seaborn.pydata.org/

[44] L. Breiman, "Random Forests," *Mach. Learn. 2001 451*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.