iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: III Month of publication: March 2026

DOI: https://doi.org/10.22214/ijraset.2026.78229

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

An AI-Powered Solar Management Suite Integrating Precision Forecasting, Smart Scheduling, and Financial Analytics

Siddhant Gangavane[1], Omkar Dhoke[2], Rohan Ghadage[3], Sahil Ghodvinde[4], Prof. Monica Shetty[5]
[1, 2, 3, 4]Student, [5]Professor, Atharva College of Engineering Malad, Mumbai

Abstract: The increasing global energy demand and environmental concerns have accelerated the adoption of renewable power sources, particularly solar energy. However, solar power generation is highly dependent on fluctuating atmospheric conditions, which makes accurate forecasting challenging. This study presents an intelligent solar prediction system that estimates energy generation using weather-based parameters and location-specific data. The proposed model employs Linear Regression trained on historical meteorological and solar datasets to generate reliable forecasts. Additionally, the framework features a Solar Payback Calculator, Smart Appliance Scheduling, and Automated Maintenance Alerts. By integrating automated location detection, electricity bill analysis, and PDF report generation, the framework simplifies user interaction while improving overall energy management efficiency.

Keywords: Solar Energy, Renewable Energy Forecasting, Linear Regression, Machine Learning, CO2 Emission Reduction

I. INTRODUCTION

The global energy landscape is undergoing a significant paradigm shift, driven by the dual pressures of rapidly increasing electricity demand and the continuous depletion of finite fossil fuel reserves. As countries strive to achieve carbon neutrality and reduce greenhouse gas emissions, the transition toward renewable energy sources has accelerated. Among these sources, solar energy has emerged as a pivotal component of a sustainable energy future due to its abundance, renewability, and environmentally benign nature. Solar power offers a viable pathway to reduce dependence on conventional power grids while minimizing carbon footprints.

Despite its advantages, the effective integration of solar energy into residential and commercial infrastructures is constrained by its inherent variability. Unlike conventional fuel-based power generation, solar energy production is highly stochastic and depends on dynamic meteorological parameters such as cloud cover, ambient temperature, solar irradiation, humidity, and atmospheric dust. These parameters fluctuate continuously, making accurate prediction of solar energy output a complex and challenging task. Inaccurate forecasting can lead to inefficient energy planning, unexpected power shortages, financial losses for prosumers, and increased reliance on non-renewable backup energy sources.

Traditionally, solar energy forecasting has relied on static and rule-based methods, including manual mathematical models and fixed assumptions derived from historical averages. While these conventional approaches may provide approximate estimates under stable environmental conditions, they lack the adaptability required to respond to real-time weather variations, such as sudden cloud formation or temperature changes. Consequently, their predictions often become out dated and unreliable. Additionally, many existing forecasting tools impose a significant burden on end users by requiring manual entry of complex parameters, including geographical coordinates, panel specifications, efficiency values, and electricity tariffs. This process is time-consuming and highly prone to human error, which further compromises the accuracy of energy output and financial savings estimations.

To overcome these limitations, there is a pressing need for a dynamic, intelligent, and user-centric forecasting system capable of learning from environmental patterns while minimizing manual intervention. In this paper, we propose “Heliocast,” an advanced solar energy forecasting framework that integrates Machine Learning (ML) with automated data acquisition technologies. The proposed system employs a Linear Regression model trained on extensive historical weather and solar generation datasets to identify meaningful relationships between meteorological variables and solar energy output.

Distinct from traditional systems, the proposed approach introduces several novel features aimed at enhancing accuracy, usability, and accessibility. Live Geolocation integration is used to automatically fetch precise user coordinates through Geolocation APIs, ensuring hyper-localized weather data and eliminating inaccuracies associated with manual location entry.

Furthermore, Optical Character Recognition (OCR) is incorporated for automated electricity bill analysis, allowing users to upload their bills and enabling the system to extract electricity consumption and tariff details without manual input. This significantly reduces user effort and input-related errors.

In addition to forecasting solar energy generation, the system provides comprehensive reporting, including estimated energy output, potential cost savings, and carbon dioxide (CO_2) emission reductions. These insights are presented through an enhanced web interface and can be exported as downloadable PDF reports for offline analysis and decision-making. By combining machine learning, real-time data integration, and intelligent automation, Heliocast delivers an end-to-end solution for efficient residential and commercial solar energy planning while promoting sustainable energy adoption.

II. LITERATURE SURVEY

Solar energy forecasting has been extensively studied using a variety of traditional and modern approaches. Early forecasting techniques primarily relied on **statistical models** and **historical averages** of meteorological parameters such as temperature, sunlight hours, and solar irradiation. Although these methods provide acceptable accuracy under stable weather conditions, they are largely **static in nature** and often fail to adapt to sudden environmental changes, resulting in unreliable predictions.

1) Manual Calculation Methods:
Conventional forecasting techniques based on solar angles, geographical parameters, and empirical formulas require extensive manual computation. These methods are time-consuming, lack scalability, and are highly susceptible to human error, limiting their practical applicability.

2) IoT-Based Monitoring Systems:
Internet of Things (IoT) solutions utilize sensors installed on solar panels to collect real-time performance and environmental data. While effective for monitoring and fault detection, these systems primarily focus on current system behavior rather than accurate future energy forecasting.

3) Satellite-Based Forecasting Models:
Satellite imagery and cloud movement analysis have been employed to improve solar energy prediction accuracy. Although these approaches offer better forecasting capabilities, they require sophisticated infrastructure, high computational resources, and expert handling, making them largely inaccessible to average residential and small-scale commercial users.

4) AI and Machine Learning Models:
Recent research highlights the growing adoption of Machine Learning (ML) techniques for solar energy forecasting. These models improve prediction accuracy by learning complex patterns from historical and real-time data. However, many such systems remain confined to academic research, lacking practical deployment, automation, and intuitive user interfaces for non-technical users.

III. METHODOLOGY

The Heliocast system was developed using a structured, data-oriented workflow to ensure accurate forecasting and seamless integration between analytical and application components. The process consisted of four stages: data collection, feature evaluation, model development, and deployment.

Historical weather records were gathered, including solar irradiance, temperature, UV index, humidity, and cloud cover. The dataset was cleaned to remove inconsistencies, missing entries, and outliers to improve reliability. Pearson correlation analysis was then performed to examine relationships between weather variables and solar output. Solar irradiance and UV index showed strong positive correlations, indicating their significant impact on energy generation.

Three regression algorithms—Linear Regression, Random Forest, and Gradient Boosting—were tested for performance comparison. The dataset was divided into training and testing subsets to validate predictive capability. Model accuracy was evaluated using MAE, RMSE, and R^2 metrics. Linear Regression achieved the best balance between simplicity and performance, with an R^2 value of 0.995.

The selected model was deployed using a Flask backend integrated with live weather APIs, geolocation services, and OCR-based billing analysis, delivering forecasts, cost estimates, and emission reduction insights through a web interface.

IV. PROPOSED FRAMEWORK

The Heliocast system follows a modular client–server architecture designed to support scalability, responsiveness, and user accessibility. The framework is divided into four functional layers: Data Acquisition, Intelligent Processing, Forecasting Engine, and Presentation & Reporting. Each layer performs a specific task, ensuring structured data movement and reliable solar energy prediction.

A. Data Acquisition Layer

This layer gathers essential inputs from users and external environmental sources while minimizing manual effort and potential inaccuracies.

1) Live Geolocation Module:

Rather than requiring users to manually enter a city or region, the system utilizes browser-based geolocation services to obtain precise latitude and longitude coordinates after user consent. This enables retrieval of hyper-local weather data, accounting for small climatic variations that can influence solar energy output. By using exact installation coordinates instead of generalized regional data, forecasting accuracy is significantly improved.

2) *OCR-Based Bill Processing:*

To streamline financial analysis, an optical character recognition module is incorporated. Users upload an image of their electricity bill, which undergoes preprocessing steps such as grayscale conversion and noise filtering. The OCR engine then extracts critical information, including the electricity tariff rate (₹/kWh) and average monthly consumption. Automating this process reduces manual input errors and enhances user convenience.

B. Intelligent Processing Layer

Once data is collected, it is refined and prepared for machine learning inference within the Intelligent Processing Layer.

1) Weather Data Integration:

The system connects to real-time weather APIs to obtain updated environmental parameters such as solar radiation, temperature, humidity, UV index, and cloud cover.

2) Data Preparation:

Incoming data is cleaned to handle missing values and irregular entries. Temporal alignment ensures that forecast intervals match the model's requirements, maintaining consistency and improving prediction reliability.

C. AI Forecasting Engine

The AI Forecasting Engine serves as the analytical core of the Heliocast system, responsible for estimating solar power generation based on environmental inputs.

1) Algorithm Selection:

A Linear Regression approach was selected due to its computational efficiency, transparency, and suitability for continuous data modelling. The algorithm establishes a mathematical relationship between meteorological variables—such as solar radiation, temperature, humidity, and UV index—and the target variable, solar energy output (kWh). Its simplicity allows faster execution while maintaining high predictive reliability.

2) Training and Validation:

The model was trained using historical weather and solar production data. To evaluate performance objectively, the dataset was divided into separate learning and validation portions. The optimization objective focused on reducing Mean Squared Error (MSE), ensuring minimal deviation between predicted and actual values. This approach enables stable forecasting even under varying atmospheric conditions.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

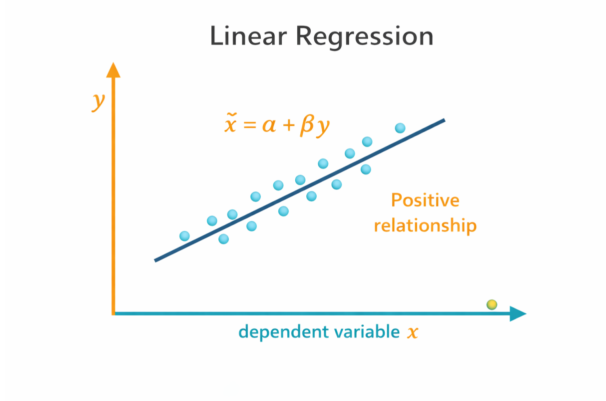

Figure 1 Linear Regression

D. Presentation and Reporting Layer

This layer converts technical predictions into user-friendly insights.

1) Automated PDF Reporting:

The system generates downloadable PDF reports that summarize localized solar potential, consumption trends, and estimated CO2 reductions for offline decision-making.

2) Solar Payback Calculator:

This module utilizes extracted tariff rates and estimated installation costs to provide users with a projected timeline for breaking even on their investment.

E. Smart Energy Management

1) Smart Scheduling Tips:

Based on predicted peak generation hours, the system provides actionable recommendations for scheduling high-load appliances (e.g., washing machines, EV chargers) to maximize self-consumption.

2) Maintenance & Cleaning Alerts:

The system monitors performance efficiency and environmental factors like dust accumulation or seasonal changes to trigger automated alerts, ensuring optimal panel output through timely maintenance.

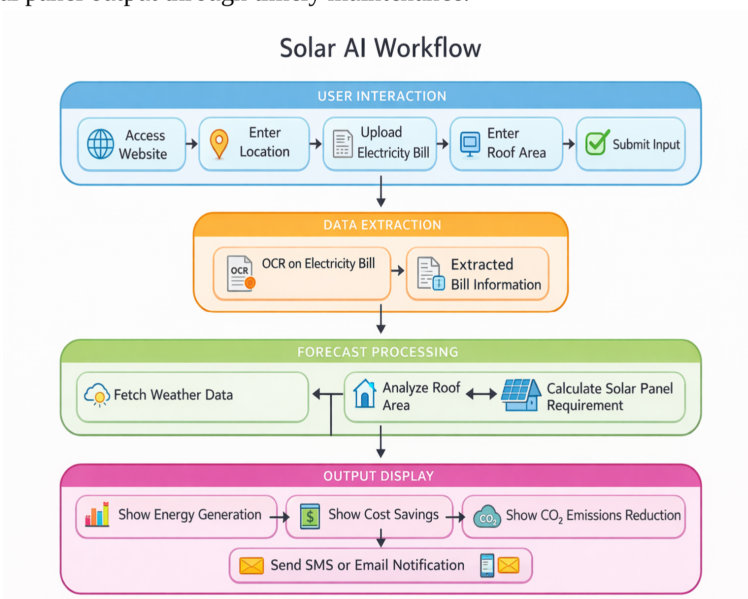

Figure 2 Solar Forecasting System

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

V. RESULTS AND DISCUSSION

The effectiveness of the Heliocast system was assessed from two primary dimensions: the predictive capability of the implemented machine learning models and the operational performance of the deployed web application.

A. Statistical Analysis and Model Validation

To examine forecasting reliability, a comparative study was conducted using three regression-based algorithms: Linear Regression, Random Forest, and Gradient Boosting. Each model was trained and evaluated using the same historical weather dataset to maintain consistency in comparison. Performance was measured using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2).

Table 1: Comparative Performance of Machine Learning Models

Model	MAE	RMSE	R^2 Score
Linear Regression	0.085	0.105	0.995
Random Forest	0.097	0.131	0.999
Gradient Boosting	0.099	0.123	0.997

Among the evaluated techniques, Linear Regression demonstrated the lowest MAE value (0.085), indicating minimal average deviation between predicted and actual solar output values. Although Random Forest achieved a slightly higher R^2 score, it produced larger error values and required greater computational resources. Considering prediction accuracy, processing efficiency, and model simplicity, Linear Regression proved to be the most practical solution for real-time implementation.

B. Feature Correlation Analysis

To better understand the relationship between weather conditions and photovoltaic output, a Pearson correlation analysis was performed. The results revealed that solar radiation had an almost perfect positive correlation (0.999) with generated energy, confirming its primary role in solar power production. Similarly, the UV index displayed a strong positive relationship (0.944), reinforcing its significance as an indicator of solar intensity.

In contrast, cloud cover (-0.403) and precipitation (-0.345) showed negative correlations with energy generation, aligning with practical observations that increased cloud density and rainfall reduce solar exposure. The strong linear dependency between major features and energy output further justified the selection of Linear Regression as the forecasting model.

C. Application Performance and Usability Evaluation

Beyond statistical accuracy, the system was evaluated for responsiveness and user experience. Automatic coordinate detection significantly reduced manual setup time, allowing predictions to be generated almost instantly. The OCR module accurately extracted tariff rates and consumption values from sample electricity bills, minimizing user input errors.

The interactive dashboard displayed daily solar generation trends, capturing fluctuations between 18.53 MJ/m² and 25.62 MJ/m² during the testing phase. This dynamic visualization offers more detailed insights compared to traditional static monthly estimates.

D. Socio-Economic and Environmental Impact

The system translated predictive outputs into practical sustainability metrics. Estimated daily savings ranged from ₹109.37 to ₹143.50, providing financial clarity for residential users. Additionally, the projected reduction in carbon emissions ranged between 14 kg and 18 kg of CO_2 per day, highlighting the environmental benefits of solar adoption supported by accurate forecasting. Beyond financial clarity through daily savings, the Payback Calculator allows users to visualize long-term ROI. Furthermore, the Smart Scheduling tips assist users in shifting their load to peak solar hours, potentially increasing actual savings beyond the initial 109.37 to 143.50 range

VI. CONCLUSIONS

This study introduced Heliocast, a machine learning–driven solar forecasting platform developed to address the drawbacks of conventional estimation approaches. By integrating a Linear Regression model with automated geolocation, OCR-enabled billing, and intelligent maintenance alerts, the system provides a precise and user-centric energy suite.

The platform converts forecasts into actionable metrics, including appliance scheduling tips and investment payback periods, enabling holistic planning for residential and commercial solar installations.

VII. ACKNOWNLEDMENT

The authors sincerely thank their project supervisor and faculty members for their constant guidance, support, and encouragement throughout this work. They also acknowledge the institution for providing essential resources and infrastructure that enabled successful project completion. Appreciation is further extended to researchers and academic communities whose published studies and reference materials contributed to the development and refinement of this research.

REFERENCES

[1] P. Kumari and A. Kumar, "Solar power forecasting using ARIMA model," International Journal of Renewable Energy Research, 2018.
[2] P. Sharma et al., "Enhancing and optimizing solar power forecasting in Dhar district of India using machine learning," Smart Grids and Sustainable Energy, 2024.
[3] F. Wang et al., "A review of deep learning for renewable energy forecasting," Energy Conversion and Management.
[4] A. K. Yadav et al., Intelligent Data Analytics for Solar Energy Prediction and Forecasting. Elsevier, 2021.
[5] J. Antonanzas et al., "Review of photovoltaic power forecasting," Solar Energy, 2016.
[6] V. Shingne et al., "AI-based solar energy forecasting," International Research Journal of Engineering and Technology (IRJET), vol. 12, no. 04, Apr. 2025.