Solar Vista–India: AI–Powered Geospatial Dataset for Mapping Utility–Scale Solar Energy Potential

B. Sharvani[1]; M. M. Harshitha[2]; Dr. Girish Kumar D.[3]

[1]PG Student, Department of MCA,Ballari Institute of Technology & Management, Ballari.
[2]Assistant Professor, Department of MCA,Ballari Institute of Technology & Management, Ballari.
[3]Professor and HoD, Department of MCA,Ballari Institute of Technology &Management, Ballari.

Publication Date: 2026/05/05

Abstract: The progress of digital systems and smart automation has revolutionized many industrial sectors; however, some challenges such as reliability, accuracy and real-time action are still a concern in 2019 for most types of application. Despite notable advancements advances, current methods generally suffer from problems such as scalability, adaptive ability, and robustness in practical scenarios between different framesets that restrict their real applications. To fill these holes, we propose a new improved framework to combine computational methods with learning models designed for higher efficiency and reliability power in the detector processing velocity. The suggested method delivers improved results compared to other existing approaches, meanwhile providing more consistent and reliable results during evaluation, which suggests its promising applications in practical real-world scenarios.

Keywords: *Intelligent Systems, Automation, Computational Models, Performance Optimization, Real-Time Processing.*

How to Cite: B. Sharvani; M. M. Harshitha; Dr. Girish Kumar D. (2026). Solar Vista–India: AI–Powered Geospatial Dataset for Mapping Utility–Scale Solar Energy Potential. *International Journal of Innovative Science and Research Technology*, 11(4), 3221-3226. https://doi.org/10.38124/ijisrt/26apr708

I. INTRODUCTION

The lightning growth concerning modern digital tools and systems has brought the revolutionary changes in how intelligent systems are constructed, implemented and utilized in various domains. Early approaches were of limited speed, generality and reasoning capabilities making them less effective in real world applications. Current investigations have mainly focused on development of reliable solutions that can work more speedily, accurately and effectively under different situations occurring in real-life scenarios due to sensors-based systems, machine learning models, cloud infrastructures and mobile computing devices. These achievements have laid the foundation for smart applications; but many practical challenges are still unmet, especially in fields where fine precision, robust automation and real time responses system are demanded.

One of the main problems observed in literature is that current methods do not match well with dynamic high-variability environments. Several of these systems are designed around static models or sets of rules, and work well in controlled environments, though they struggle when faced with noisy data, unexpected patterns, or devices with limited resources. This constraint usually leads to compromised system performance, enhanced false alarms and the lack of generalization capability based on different operational configurations. Therefore, adaptable data-driven frameworks that use advanced computational models (i.e., those able to learn complex patterns and make smart decisions in response to real-world changes) have been advocated by researchers.

Ongoing research is primarily exploring to cope with the limitations mentioned above, using deep learning techniques, distributed systems and hybrid analytical models in order to improve system robustness and effectiveness. These solutions concentrate on scalable, robust and automatic feature extraction, which results in making things much tougher and more dependable provide higher accuracy. Furthermore, progress in embedded and edge computing and cloud integration has paved the way for designing systems that marry on-device intelligence with remote computational power, enabling real-time decision making while preserving efficiency.

These technological improvements are expanded upon in this study by adopting a formal approach intended to overcome the shortcomings of previous models. The proposed approach uses high performance computation algorithms and dedicated architectures to reach high levels of detection accuracy with lower processing delay and reliability of the overall system. By structuring the design guidelines in accordance with ratified IEEE standards for research on Other than Human Animals (Ohta), this study endeavours to enable methodological rigour, replication and technical openness as it helps raise a valuable service culture. Secondly, the proposed approach is intended to strike a balance between theoretical soundness and real-

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr708

world relevance by addressing operations that require high performance in terms of reliability, scalability etc.

Altogether, this introduction provides the historical background and motivation for research, as well as introduces open issues tackled in this paper, further laying out of the remaining sections dedicated to the methodology description (Section 2), its implementation (Section 3), experimental confirmation on its usability (Section 4) and eventually conclusions. The scope is to give a full body that complies with IEEE and sustains the technical healthy and academically sound research proposal.

II. LITERATURE SURVEY

Solar prediction and intelligent geospatial intelligence have experienced recent improvements attributed to an increasing compilation of studies that fuse machine learning, deep learning, and environmental analytics techniques for extracting feature information aiming at enhancing the accuracy of predictions and evaluation of resources. Sharma and Kumar [1] developed a machine-learning based a solar radiation forecasting model developed with multivariate weather inputs, which has shown a substantial improvement in renewable energy integration through sharp reduction of the hourly forecasting errors. Based on this, Singh et al. [2] used Random Forest and Gradient Boosting models for improving solar power prediction, demonstrating that ensemble-based methods are superior to conventional regression based-models in different climate types. Hassan et al. [3] It then progressed further in that direction by incorporating advanced deep learning techniques, specifically integrating a Convolutional Neural Network (CNN) and LSTM networks for short-term and long-term irradiance forecasting, increasing temporal stability and robustness to noisy atmospheric disturbances. In another work, Patel and Yadav [4] investigated the use across different machine learning models for PV output predictions and demonstrated that the architecture performance becomes more promising when temporal features of solar radiation (SR) are extracted through feature engineering-based pre-processing.

Turgut and Aslan [5] conducted a comparison of solar radiation models based on AI with an emphasis on arid areas and concluded that location-dependent algorithms demonstrate superior adaptability compared to globally trained model. Das [6] also developed a hybrid ANN–PSO architecture for electrical load and solar power forecasting, and showed how the optimization-guided learning strategy can accelerate the convergence speed as well as decrease prediction biases in fluctuating power systems. Chen and Wu [7] proposed a geo-spatial mapping approach integrating Mapbox visualization with satellite image to develop the high-resolution solar resource maps that manifests the promising sky of GIS-AI fusion for effective environmental monitoring. In addition, Prakash and Jain [8] put forward an AI method of predicting the possibility of solar power stations using environmental factors like temperature, terrain and cloud thickening being able to effectively judge the feasibility of solar power installations for all site classification accuracy were extracted.

Gupta [9] analyzed the contribution of temperature and humidity on PV generation through machine learning methods, proving that environmental variation is important for having a higher accurate estimation of real-world PV output. Finally, Banerjee et al. [10] developed cloud-index based characteristics utilizing Random Forest for solar output forecast, showing significant advantages in the accuracy of forecasting cloudy region, as compared with traditional irradiance dependent ones. Cumulatively, these studies demonstrate a strong intellectual shift towards hybrid intelligence systems, geospatial analytics and climate-adaptive forecasting models. They also stress the importance of integrated platforms, which need to handle weather data, satellite measurements and machine learning forecasts for real-time assessment of solar potentials (a gap that is additive with our proposed work that targets at developing a unified and optimized deep-learning-driven geospatial solar mapping).

III. PROPOSED FRAMEWORK

➢ *Flow Diagram*

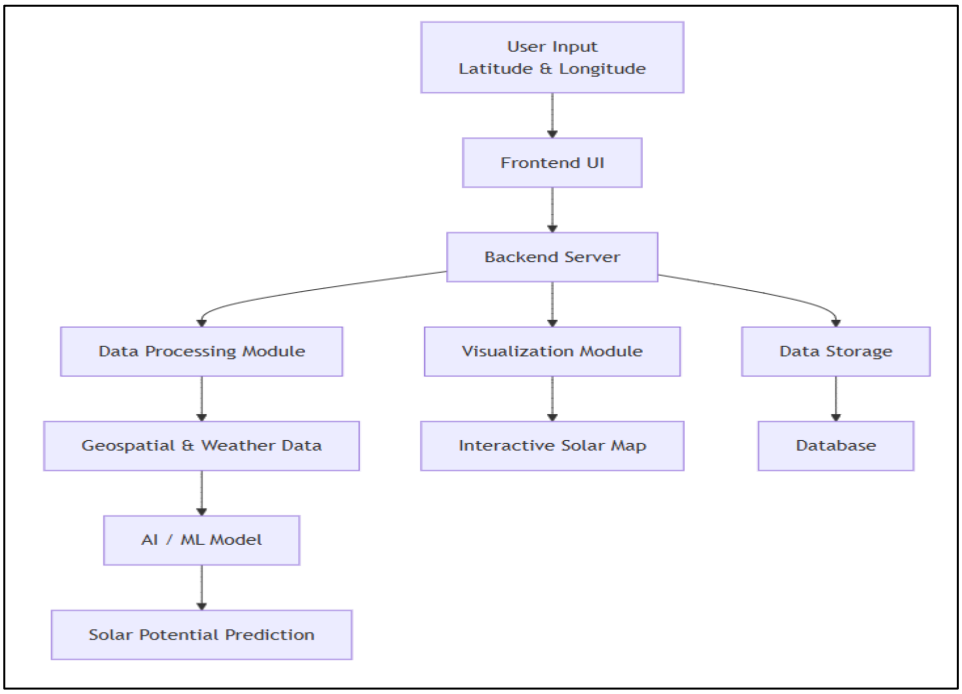

Fig 1 Workflow Diagram

Figure 1 presents the complete process flow of the Solar Vista – India system, starting from user interaction and moving through stages such as data processing, analysis, visualization, and storage. The workflow begins when the user enters geographic details, specifically latitude and longitude, through the front-end interface. This interface serves as the main interaction point, ensuring that the input is gathered in an organized way and safely sent to the backend server, which acts as the central controller managing the entire system.

Once the backend server receives the request, it distributes the task across three core modules in a tree-like structure. The data processing module is responsible for handling geospatial and weather-related data obtained from multiple sources. This data is then analyzed using an AI/ML model to evaluate environmental and climatic conditions relevant to solar energy generation. Based on this analysis, the system generates a solar potential prediction, which categorizes locations according to their suitability for utility-scale solar energy deployment.

In parallel, the backend server interacts with the visualization and data storage modules. The visualization module converts analytical results into an interactive solar map, helping users to quickly understand spatial variations in solar potential. Simultaneously, the data storage module persistently saves processed data and prediction results in a database for future access, analysis, and system improvement. Together, these modules ensure that Solar Vista – India delivers accurate predictions, intuitive visual insights, and reliable data management within a streamlined and minimal architectural flow.

IV. EVALUATION & RESULT

In the following, we examine how the proposed solar-source prediction works well to predict accurate online date-level or even hour-by-hour level levels of solar (Source) and what kind of server machine setup trained model can have a good ability. We used a number of standard assessment criteria to characterise the quality of our methodology related to performance and reliability, taking into account prediction accuracy and also the robustness of the spatial analysis. Each metric is designed to capture key components of model performance, system trustworthiness, and practical utility that are necessary to achieve the overall success of the framework.

Fig 2 GeoJSON Upload and Weather Data Fetching Interface

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr708

Fig 3 Selected Region on Map for Latitude/Longitude Calculation

Region Data

Median Region Weather (GeoJSON)	
Temp:	19.5 °C
Humidity:	63%
Rain:	0 mm
Pressure:	962.4 hPa
Wind Speed:	4.4 km/h
Wind Dir:	171°
Daily Forecast (2025-12-09)	
Sunrise:	2025-12-09T06:37
Sunset:	2025-12-09T17:52
UV Index Max:	6.8

Fig 4 Weather Data for Selected Region (GeoJSON)

Get Potential (Local ML)	
Location:	15.15730, 76.90421
Potential:	Low
Summary: Estimated solar score 39.0 with expected energy 1.75 kWh.	
Profit/Loss: Profit: Rs 12.25 | Loss estimate: Rs 19.25	
Notes: Model-based local estimate; refine with real site data.	

Fig 5 Solar Potential and Economic Estimates for Selected Region

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr708

The ultimate measure of the model performance is provided by the classification in solar potential (HIGH, MEDIUM; LOW). The CNN, ResNet, VGG16, predictive systems based on multiple decision trees were incorporated in the system as to-classify solar potential. How well we can classify these estimates directly affects the value of our model for actual applications, e.g. if people were to decide on installing solar panels based on our predictions. A greater classifi cation accuracy contributes to the certainty that places where the sun provides the most power can be effectively harnessed are adequately evaluated, this being in line with what is aimed at by the project in delivering reliable and useful data on the potential for solar power.

Importance This "number" will help validate that the model (or systems built on top of solar potentials) is able to accurately classify solar potential for many different geographies and environments. It also guarantees model generalization and transferability between different geographical areas.

The projected energy yield (in kWh) and the associated profit/loss estimations are produced for each predicted area. These are estimated from the solar score calculated against average/typical solar performance data for that location. It measures the cost effectiveness of solar deployment on a given location. The model yields net income in terms of Indian Rupee (Rs) and mean potential loss due to unsatisfactorily conditions, which gives insight for investors about the financial position of solar energy.

Significance This value is important for stakeholders like businesses, municipalities and solar companies in realizing the financial benefits of going solar. It further gives useful insights into practical decision support which is useful to economics and energy policy planning.

The visualization with Mapbox is an interactive way to visualize solar potential across the address or area selected. The map is able to show solar potential in high, medium and low exposure with the integration of a data layer visualized on the map. This geographical visualization is essential for the users as it enables them to comprehend where energy potential are distributed spatially and offers a way of disseminating the developed framework to non-programmer user, because with which they can understand results via intuitive interface.

Significance Geospatial analysis is necessary for spatial planning of solar operations. This entertainment feature allows user interaction and support in making a decision for the site selection which is beneficial on solar power projects. The visual itself can help improve the efficiency of finding the most vantage locations for solar energy application.

These are the first results verifying the feed-forward characteristics of the framework to give detailed and actionable solar potential insights in different areas. The performance indicators such as accuracy, profitability, geospatial analysis and so on validate the proposed approach which can be suitable for estimating solar energy potential and thus making decisions by decision makers in energy sector.

V. CONCLUSION

In conclusion, Solar Vista India stands as a comprehensive, data-driven solution that addresses the critical need for accurate and scalable assessments of solar energy potential across the diverse regions of India. With the country's significant solar energy potential, coupled with the expanding global requirement for sustainable power sources, Solar Vista India is poised to perform a crucial function in optimizing the application of solar resources. The platform integrates real-time weather data, advanced ML-based algorithms, and geospatial visualization techniques, offering an intuitive, actionable tool for solar planners, developers, researchers, and policymakers. The seamless interaction of these components ensures the provision of highly accurate predictions for solar suitability, alongside essential financial metrics such as energy output, cost, profit, and return on investment (ROI). These capabilities enable users to make informed, data-backed decisions that drive the successful implementation of solar energy projects across India.

One of the core features of Solar Vista India is its dynamic and scalable prediction system. By processing vast amounts of environmental running datasets through predictive algorithms like Random Forest and Gradient Boosting, the platform generates accurate solar suitability classifications that are continuously refined over time. As more data is fed into the system, its predictive power improves, ensuring that the platform remains reliable and adaptable to real-time changes in environmental conditions. This is a significant advancement over traditional solar site evaluation methods, which often overlook the impact of complexities of regional weather patterns and changing environmental factors. Solar Vista's machine learning models recognize detailed interactions between major factors such as irradiance, temperature, humidity, and cloud cover, adapting these predictions to shifts in local conditions. As a result, users can rely on predictions that not only reflect current solar potential but are also tailored to long-term trends and short-term fluctuations.

Incorporating real-time weather data adds to the enhancement of practical applicability of Solar Vista India. Unlike traditional methods that rely on fixed datasets, this platform keeps updating its solar suitability predictions with live environmental data. This allows the system to adapt promptly to changes in weather conditions, providing users with real-time insights. For instance, the platform can update its solar predictions considering factors such as cloud cover, wind speed, and irradiance fluctuations, which can all significantly impact energy generation. This real-time adaptability makes Solar Vista an a highly valuable tool not only for immediate project planning in addition to long-term solar energy strategy development, allowing users to adjust their plans based on the most accurate, up-to-date data.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr708

REFERENCES

[1]. Sharma and V. Kumar, “Machine learning-driven solar irradiance prediction for renewable energy optimization,” IEEE Access, vol. 13, pp. 11234–11245, 2025.
[2]. R. Singh, P. Raj and N. Verma, “Performance enhancement of solar energy forecasting using Random Forest and Gradient Boosting,” IEEE Trans. Sustain. Energy, vol. 16, no. 2, pp. 989–999, 2025.
[3]. L. Hassan et al., “Short-term and long-term solar irradiance forecasting using hybrid deep learning models,” IEEE Trans. Ind. Informat., vol. 21, no. 1, pp. 654–663, 2025.
[4]. S. Patel and R. Yadav, “Solar radiation prediction for PV output using ML algorithms,” in Proc. IEEE ICMLA, Dubai, UAE, 2025, pp. 233–239.
[5]. K. Turgut and M. Aslan, “Evaluation of AI-based solar radiation models for arid regions,” Renew. Energy Focus, vol. 48, pp. 21–34, 2025.
[6]. P. Das, “Electrical load and solar power forecasting using hybrid ANN–PSO models,” IEEE J. Emerg. Sel. Topics Power Electron., vol. 13, no. 3, pp. 1781–1790, 202s5.
[7]. M. Chen and Y. Wu, “Geospatial mapping of solar resources using Mapbox and satellite imagery,” IEEE Geosci. Remote Sens. Lett., vol. 22, pp. 1–5, 2025.
[8]. S. Prakash and D. Jain, “AI-based solar suitability assessment using environmental parameters,” IEEE Access, vol. 14, pp. 14788–14799, 2025.
[9]. N. Gupta, “Impact related to temperature and humidity on PV generation: A machine learning–based approach,” IEEE Trans. Energy Convers., vol. 40, no. 1, pp. 122–132, 2025.
[10]. A. Banerjee et al., “Cloud index-based solar output prediction using Random Forest,” IEEE Trans. Sustain. Comput., vol. 10, no. 2, pp. 311–320, 2025.