VIVEKANAND EDUCATION SOCIETY'S INSTITUTE OF TECHNOLOGY

An Autonomous Institute Affiliated to University of Mumbai

Department of Computer Engineering

Project Report on

"SolarSense" Predictive Analytics for Solar Energy : The smart solar system approach

In partial fulfillment of the Fourth Year, Bachelor of Engineering (B.E.) Degree in Computer Engineering at the University of Mumbai

Academic Year 2023-24

Submitted by

Anishkumar Iyer (D17C , Roll no - 03)

Swapnil Thatte (D17C , Roll no - 58)

Yash Brid (D17B , Roll no - 11)

Yash Narkhede (D17B , Roll no - 46)

Project Mentor

Dr. Mr. Dashrath Mane

(2023-24)

VIVEKANAND EDUCATION SOCIETY'S INSTITUTE OF TECHNOLOGY

An Autonomous Institute Affiliated to University of Mumbai

Department of Computer Engineering

Certificate

This is to certify that ***Anishkumar Iyer, Swapnil Thatte, Yash Brid, Yash Narkhede*** of Fourth Year Computer Engineering studying under the University of Mumbai have satisfactorily completed the project on "***"SolarSense" Predictive Analytics for Solar Energy The smart solar system approach***" as a part of their coursework of PROJECT-II for Semester-VIII under the guidance of their mentor ***Dr. Dashrath Mane*** in the year 2023-24 .

This thesis/dissertation/project report entitled ***"SolarSense" Predictive Analytics for Solar Energy The smart solar system approach*** by Anishkumar Iyer, Swapnil Thatte, Yash Brid, Yash Narkhede is approved for the degree of **B.E. Computer Engineering**.

Programme Outcomes	Grade
PO1,PO2,PO3,PO4,PO5,PO6,PO7,PO8, PO9, PO10, PO11, PO12, PSO1, PSO2	

Date:

Project Guide:

Project Report Approval

For

B. E (Computer Engineering)

This project report entitled ***"SolarSense" Predictive Analytics for Solar Energy The smart solar system approach*** by ***Anishkumar Iyer, Swapnil Thatte, Yash Brid, Yash Narkhede*** is approved for the degree of **B.E. Computer Engineering**.

Internal Examiner

External Examiner

Head of the Department

Principal

Date:

Place:

Declaration

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

--	--
Anishkumar Iyer - 03	Swapnil Thatte - 58
--	--
Yash Brid - 11	Yash Narkhede - 46

Date:

ACKNOWLEDGEMENT

We are thankful to our college Vivekanand Education Society's Institute of Technology for considering our project and extending help at all stages needed during our work of collecting information regarding the project.

It gives us immense pleasure to express our deep and sincere gratitude to Assistant Professor **Dr. Dashrath Mane** (Project Guide) for her kind help and valuable advice during the development of project synopsis and for her guidance and suggestions.

We are deeply indebted to Head of the Computer Department **Dr.(Mrs.) Nupur Giri** and our Principal **Dr. (Mrs.) J.M. Nair,** for giving us this valuable opportunity to do this project.

We express our hearty thanks to them for their assistance without which it would have been difficult to finish this project synopsis and project review successfully.

We convey our deep sense of gratitude to all teaching and non-teaching staff for their constant encouragement, support and selfless help throughout the project work. It is a great pleasure to acknowledge the help and suggestions, which we received from the Department of Computer Engineering.

We wish to express our profound thanks to all those who helped us in gathering information about the project. Our families too have provided moral support and encouragement several times.

Computer Engineering Department

COURSE OUTCOMES FOR B.E PROJECT

Learners will be to,

Course Outcome	**Description of the Course Outcome**
CO 1	Able to apply the relevant engineering concepts, knowledge and skills towards the project.
CO2	Able to identify, formulate and interpret the various relevant research papers and to determine the problem.
CO 3	Able to apply the engineering concepts towards designing solutions for the problem.
CO 4	Able to interpret the data and datasets to be utilized.
CO 5	Able to create, select and apply appropriate technologies, techniques, resources and tools for the project.
CO 6	Able to apply ethical, professional policies and principles towards societal, environmental, safety and cultural benefit.
CO 7	Able to function effectively as an individual, and as a member of a team, allocating roles with clear lines of responsibility and accountability.
CO 8	Able to write effective reports, design documents and make effective presentations.
CO 9	Able to apply engineering and management principles to the project as a team member.
CO 10	Able to apply the project domain knowledge to sharpen one’s competency.
CO 11	Able to develop professional, presentational, balanced and structured approach towards project development.
CO 12	Able to adopt skills, languages, environment and platforms for creating innovative solutions for the project.

Index

List of Figures:

Abstract

As solar energy adoption accelerates, ensuring the optimal performance and reliability of solar systems becomes paramount. The SolarSense project presents a groundbreaking predictive analytics system designed to revolutionize solar energy management. At its core, SolarSense employs advanced machine learning algorithms to analyze historical solar inverter data. This enables the system to proactively predict and detect potential faults, minimizing downtime and significantly reducing maintenance costs. By identifying early warning signs, SolarSense facilitates timely interventions, safeguarding the overall reliability of solar energy installations.

Furthermore, SolarSense incorporates comprehensive rooftop solar panel analysis capabilities. The system evaluates key performance indicators, empowering users to implement targeted maintenance and optimization strategies. This data-driven approach maximizes energy production and extends the lifespan of solar panels. The project meticulously outlines its technical architecture, detailing the specific machine learning techniques employed for both fault detection and rooftop analysis.

Demonstrating its real-world applicability, SolarSense has been successfully tested in both residential and commercial solar energy settings. The results underscore its potential to streamline solar panel maintenance and enhance the overall efficiency of solar energy systems. The SolarSense project signifies a major advancement in intelligent solar energy management, offering a powerful and practical solution for the growing solar energy industry.

Chapter 1: Introduction

1.1 Introduction

The increasing global demand for renewable energy necessitates innovative solutions to maximize solar energy utilization. The Smart Solar System is a strategic response to this demand, revolutionizing solar energy utilization with data-driven methodologies for enhanced efficiency, reliability, and sustainability.

This intelligent system addresses the intermittency of solar power through accurate Solar Power Generation Forecasting, allowing optimal resource allocation and grid stability.

The Rooftop Analysis feature ensures efficient solar panel placement, considering factors such as sunlight exposure and shade.

Predictive Maintenance plays a crucial role in identifying potential issues, extending the overall lifespan of the solar infrastructure and contributing to a more sustainable and long-term solution for renewable energy generation.

By reducing reliance on non-renewable sources and mitigating carbon emissions, the Smart Solar System aligns with the global shift towards clean and environmentally responsible energy solutions. It stands as a beacon of innovation, offering a comprehensive and intelligent approach to meeting the increasing demand for clean and sustainable energy.

1.2 Motivation

The motivation behind the development and implementation of Solar Sense is rooted in a collective responsibility to combat environmental challenges. Climate change, with its far-reaching consequences, underscores the urgency to transition from conventional energy sources to sustainable alternatives. The rising global temperatures, extreme weather events, and alarming sea-level increases are clear indicators of the ecological toll exacted by human activities. The conventional power industry, dominated by fossil fuels, is a major contributor to the environmental predicament. The extraction and combustion of fossil fuels release copious amounts of greenhouse gasses, intensifying the greenhouse effect and accelerating climate change. Additionally, the extraction processes often result in ecological degradation, contributing to habitat loss and biodiversity decline. The reliance on non-renewable resources also poses geopolitical challenges, with nations vying for control over fossil fuel reserves.

In essence, solar energy is not just a source of power; it symbolizes a paradigm shift towards a sustainable, resilient, and environmentally responsible energy future. Solar Sense embodies this vision, offering a tangible and intelligent solution to address both the challenges posed by climate change and the drawbacks of conventional power systems.

1.3 Problem Definition

In light of these challenges, the development of a Smart Solar System that leverages advanced features becomes imperative. This system aims to optimize solar energy utilization through data-driven methodologies, integrating Power Generation Forecasting, Predictive Maintenance of Solar Panels, and Rooftop solar area analysis. The problem addresses the need to design and implement a comprehensive Smart Solar System that effectively addresses the technical intricacies and operational complexities associated with solar energy utilization.

1.4 Existing Systems

In the realm of solar energy solutions, two prominent players stand out: Tesla Solar and Google. Formerly known as SolarCity, Tesla Solar offers the Tesla Solar Roof Calculator, a sophisticated platform renowned for its advanced features. Employing cutting-edge algorithms, the Tesla Solar Roof Calculator conducts comprehensive rooftop analyses, considering variables such as orientation, shading, and available space to optimize solar panel placement. Furthermore, it harnesses weather data and solar modelling to forecast energy generation, empowering users with insights into their potential solar output over time. Notably, the platform seamlessly integrates with other Tesla products, providing users with a holistic approach to exploring clean energy solutions.

Meanwhile, Google's Project Sunroof sets the standard with its innovative features designed to simplify the solar adoption process. By leveraging the Google Maps API, Project Sunroof offers precise location input, visually illustrating a roof's solar potential through intuitive map interfaces. Employing sophisticated 3D modelling techniques, the platform accurately assesses shading and sun exposure, ensuring precise calculations of solar potential. Additionally, Project Sunroof provides valuable insights into potential energy savings, considering local utility rates and available incentives, thus empowering users to make informed decisions regarding solar energy adoption.

1.5 Lacuna of the existing systems

1. Lack of precise localized data compromises accuracy in solar potential assessments.

2. Insufficient data on Indian climate and Indian rooftops.
3. Difficulty in seamless integration into existing infrastructures hampers widespread adoption.
4. Over-reliance on historical data limits adaptability to emerging climate patterns.

1.6 Relevance of the Project

The Smart Solar System project represents a crucial intersection of environmental urgency, technological innovation, and energy transformation. Amid escalating climate change, this initiative serves as a beacon, offering a practical solution to environmental challenges. At its core, the project addresses the pressing need to mitigate climate change by optimizing solar energy production, diminishing reliance on fossil fuels, and actively combating greenhouse gas emissions. Through advanced analytics, smart rooftop analysis, and predictive maintenance, the project tackles inefficiencies inherent in conventional power systems.

Beyond climate concerns, the project strategically responds to challenges embedded in the traditional power industry, championing renewable energy and a smarter approach to power generation. This aligns with the global shift toward sustainable technologies. Importantly, the project promotes environmental stewardship, exemplifying a commitment to harmonizing human activities with the natural world, in stark contrast to destructive conventional energy extraction methods.

On a larger scale, the initiative represents a pivotal step toward a solar energy future. Integration of smart technologies positions solar energy as a reliable, efficient, and intelligent power source, crucial for a clean energy transition. With decreasing costs and increasing efficiency, the Smart Solar System project is not just promising but also scalable and economically viable. In essence, its relevance lies in a holistic response to climate change, a smart approach to power generation, a commitment to environmental stewardship, and its role in steering us towards a sustainable solar-powered future.

Chapter 2: Literature Survey

2.1 Literature Referred

1) **Spatial-Temporal Solar Power Forecasting for Smart Grids(Add citations)**

Abstract: The solar power penetration in distribution grids has grown fast during the last few years, particularly at the low-voltage (LV) level, which introduces new challenges when operating distribution grids. Across the world, distribution system operators (DSO) are developing the smart grid concept, and one key tool for this new paradigm is solar power forecasting. This paper presents a new spatial-temporal forecasting method based on the vector autoregression framework, which combines observations of solar generation collected by smart meters and distribution transformer controllers. The scope is 6-h-ahead forecasts at the residential solar photovoltaic and medium-voltage (MV)/LV substation levels. This framework has been tested in the smart grid pilot of Évora, Portugal, and using data from 44 microgeneration units and 10 MV/LV substations. A benchmark comparison was made with the autoregressive forecasting model (AR-univariate model) leading to an improvement on average between 8% and 10%.

Inference Drawn: The paper presents a novel solar power forecasting method using spatial-temporal modelling with Recursive Least Squares and Gradient Boosting, leveraging smart grid infrastructure. It reduces forecast errors by 8-12% in critical lead times, emphasising RLS's superiority, and suggests future research avenues such as weather data integration and probabilistic forecasting.

2) **LSTM Networks for Overcoming the Challenges Associated with Photovoltaic Module Maintenance in Smart Cities**

Abstract: Predictive maintenance is a field of research that has emerged from the need to improve the systems in place. This research focuses on controlling the degradation of photovoltaic (PV) modules in outdoor solar panels, which are exposed to a variety of climatic loads. Improved reliability, operation, and performance can be achieved through monitoring. In this study, a system capable of predicting the output power of a solar module was implemented. It monitors different parameters and uses automatic learning techniques for prediction. Its use improved reliability, operation, and performance. On the other hand, automatic learning algorithms were evaluated with different metrics in order to optimize and find the best configuration that provides an optimal solution to the problem. With the aim of

increasing the share of renewable energy penetration, an architectural proposal based on Edge Computing was included to implement the proposed model into a system. The proposed model is designated for outdoor predictions and offers many advantages, such as monitoring of individual panels, optimization of system response, and speed of communication with the Cloud. The final objective of the work was to contribute to the smart Energy system concept, providing solutions for planning the entire energy system together with the identification of suitable energy infrastructure designs and operational strategies.

Inference Drawn: The paper conducts a systematic review of predictive maintenance technology for solar PV modules, introducing a neural network that predicts module performance based on weather data, outperforming existing models by up to 25.4%. It also proposes an Edge Computing architecture for IoT device connectivity and data storage, with plans to further test the system's performance under adverse weather conditions and explore alternative prediction algorithms like ESNs.

3) Solar Panels Dirt Monitoring and Cleaning for Performance Improvement: A Systematic Review on Smart Systems

Abstract: The advancement in technology to manage energy generation using solar panels has proved vital for increased reliability and reduced cost. Solar panels emit no pollution while producing electricity as a renewable energy source. However, the solar panel is adversely affected by dirt, a major environmental factor affecting energy production. The intensity of light falling on the solar panel is reduced when dirt accumulates on the surface. This, in turn, lowers the output of electrical energy generated by the solar panel. Since cleansing the solar panel is essential, constant monitoring and evaluation of these processes are necessary to optimize them. This emphasizes the importance of using smart systems to monitor dirt and clean solar panels to improve their performance. The paper tries to verify the existence and the degree of research interest in this topic and seeks to evaluate the impact of smart systems to detect dirt conditions and clean solar panels compared to autonomous and manual technology. Research on smart systems for addressing dirt accumulation on solar panels was conducted taking into account efficiency, accuracy, complexity, and reliability, initial and running cost. Overall, real-time monitoring and cleaning of the solar panel improved its output power with integrated smart systems. It helps users get real-time updates of the solar panel’s condition and control actions from distant locations. A critical limitation of this research is the insufficient empirical analysis of existing smart systems, which should be thoroughly examined to allow further generalization of theoretical findings.

Inference Drawn: The systematic review highlights the effectiveness of smart systems in enhancing solar panel cleaning and monitoring, focusing on dirt detection, cleaning methods, wireless communication tech, and cloud platforms. While evidence supports improved performance and reduced maintenance costs, further research is needed to address knowledge gaps and determine optimal cleaning frequency and costs.

4) Mitigated cutting force and surface roughness in titanium Alloy-Multiple effective guided chaotic multi-objective Teaching learning-based optimization

Abstract: Titanium alloys have significance in engineering applications owing to their enhanced properties and their ability to retain their shape at elevated temperatures. A new Teaching Learning Based Optimization (TLBO) variant was developed with multiple search features to mitigate cutting force and surface irregularities in Titanium samples. This assists in achieving the best quality of the product at minimal cutting energy. Experiments were conducted and the significance of machining parameters on cutting force and surface finish were analyzed. It is ascertained that the best surface is attained at a lower tool feed rate with a higher cutting speed. The increase in nose radius has more influence on the surface quality. Chaotic multiobjective TLBO with multiple effective guidance was applied in both single objective and multiobjective optimization, where useful information of other non-fittest learners is leveraged for effective more population search. The performance of the new algorithm was evaluated and comprehensively discussed. The minimum cutting force F z ¼ 65:06N and R a ¼ 1:41lm can be achieved with v ¼ 130m=min, f ¼ 0:051mm=rev, nr ¼ 0:4mm and ap ¼ 0:5mm. The predicted results were validated experimentally and verified with other existing optimizers. It is concluded that this new algorithm can be applied in machining and production wastage can be greatly minimized.

Inference Drawn: A new meta-heuristic algorithm, CMTLBO-MEG, is introduced for predicting optimal machining parameters to enhance dimensional accuracy and reduce energy consumption in manufacturing.

Key findings from experiments include:

1. Lower machining speeds lead to less curled and more discontinuous chips, while higher tool feed rates increase chip width and cutting speed improves chip curling.
2. Higher machining speeds result in a rapid rise in machining force, especially with increased tool feed rate. Smoother surface finishes are achieved with larger tool radii and

lower tool feed rates at high machining speeds, with a maximum surface roughness of 2.82 μm observed.

5) An Interpretable Probabilistic Model for Short-Term Solar Power Forecasting Using Natural Gradient Boosting

Abstract: PV power forecasting models are predominantly based on machine learning algorithms which do not provide any insight into or explanation about their predictions (black boxes). Therefore, their direct implementation in environments where transparency is required, and the trust associated with their predictions may be questioned. To this end, we propose a two-stage probabilistic forecasting framework able to generate highly accurate, reliable, and sharp forecasts yet offering full transparency on both the point forecasts and the prediction intervals (PIs). In the first stage, we exploit natural gradient boosting (NGBoost) for yielding probabilistic forecasts, while in the second stage, we calculate the Shapley additive explanation (SHAP) values in order to comprehend why a prediction was made fully. Real data from two PV parks located in Southern Germany are employed to highlight the performance and applicability of the proposed framework. Comparative results with two state-of-the-art algorithms, namely the Gaussian process and lower upper bound estimation, manifest a significant increase in the point forecast accuracy and in the overall probabilistic performance. Most importantly, a detailed analysis of the model's complex nonlinear relationships and interaction effects between the various features is presented. This allows interpretation of the model, identifying some learned physical properties, explaining individual predictions, reducing the computational requirements for the training without jeopardizing the model's accuracy, detecting possible bugs, and gaining trust in the model. Finally, we conclude that the model was able to develop complex nonlinear relationships which follow known physical properties as well as human logic and intuition.

Inference Drawn: The proposed transparent model in machine learning shows no evident bugs or biases, enhancing interpretability and allowing for better feature selection, resulting in a 6% reduction in RMSE and 10% improvement in CRPS. This approach is particularly useful for stakeholders in fields with financial risks, offering both complex model capabilities and transparency. Future work will involve analyzing feature values contributing to large prediction errors and applying the method to other power system forecasting problems like wind power and load forecasting.

2.2 Patent Search

1) Solar Power Forecasting (WO2017193172A1)

Inventor: Sam West, Ryan Lagerstrom, Changming Sun, Li RONGXIN, Joel WONG

This method utilizes a distributed network of digital cameras to capture multiple sky images, extracting sun location parameters. A three-dimensional (3D) sky model is then generated, incorporating 3D object data representing sky elements and position data for accurate spatial representation. Leveraging this model, the technique calculates the solar radiation level at a specified Point of Interest (POI) by considering the 3D object data, position data, and sun location parameters. This approach provides a nuanced understanding of solar exposure at the POI, facilitating precise analysis for optimal utilization of solar energy resources.

Link: https://patents.google.com/patent/WO2017193172A1/en

2) Solar Energy Forecasting (US20170031056A1)

Inventor : Rolando Vega-Avila ,Hariharan Krish-Naswami, Jaro Nummikoski

This solar energy forecasting method involves masking sky images containing clouds, transforming them geometrically to flat sky images based on cloud base height. Identifying and tracking cloud motion over time, the technique traces solar irradiance rays through identified clouds. By ray tracing the sun's irradiance on a geographic location according to cloud motion, a solar energy forecast is generated. This approach enhances accuracy by dynamically considering cloud movement, providing a more nuanced understanding of solar energy availability for efficient energy planning and utilization.

Link: https://patents.google.com/patent/US20170031056A1

2.3 Comparison with the existing system

The proposed smart solar system demonstrates significant advantages over existing systems in the areas of fault detection, anomaly detection, rooftop analysis, and power generation prediction:

Fault Detection: Traditional fault detection in solar energy systems often relies on manual inspections and basic sensor-based detection. These methods can be time-consuming, labor-intensive, and prone to error. In contrast, the proposed system's model vision transformer approach offers several advantages:

- **Automation**: Automates the detection process, significantly reducing time and labor costs.
- **Accuracy**: Demonstrates superior accuracy in identifying a wide range of faults compared to manual and sensor-based methods.
- **Scalability**: Can be efficiently scaled to monitor large-scale solar installations.

Anomaly Detection: While many existing systems focus on fault detection, our system's two-stage approach offers distinct improvements:

- **Proactive Detection**: Employs Isolation Forest to identify potential anomalies earlier, facilitating proactive maintenance strategies and preventing costly failures.
- **Enhanced Classification**: Utilizes a Random Forest classifier informed by SHAP analysis, ensuring accurate classification of anomalies. This granular approach enables targeted corrective actions, further optimizing maintenance efficiency.

Rooftop Analysis: The system's rooftop analysis capability distinguishes itself from those in predominant use:

- **Precision**: Allows precise selection of areas of interest using polygons, enabling highly accurate assessment of solar power potential specific to targeted zones.
- **Customization**: Provides flexibility to analyze diverse rooftops and their potential, while many existing solutions focus on generic, broad estimations.

Power Generation Prediction: The proposed approach incorporates advanced techniques that outperform conventional forecasting models:

- **Leveraging Weather and Power Yield Patterns**: Captures intricate relationships between weather data and power output, leading to more reliable forecasts.
- **Time Series Analysis**: Models temporal trends and seasonality, enhancing prediction accuracy, especially for short and medium-term forecasts crucial for smart solar systems.

By integrating these advanced techniques, the proposed smart solar system delivers a more comprehensive, intelligent, and tailored approach to optimizing solar energy generation and maintenance.

Chapter 3: Requirement Gathering for the Proposed System

In this chapter, we are going to discuss the resources we have used and how we analyzed what the user actually needs and what we can provide. We will also discuss the functional and non-functional requirements and finally the software and hardware used.

3.1 Introduction to Requirement Gathering

The Requirement Gathering is a process of requirements discovery or generating list of requirements or collecting as many requirements as possible by end users. It is also called as requirements elicitation or requirement capture.

The requirements-gathering process consists of six steps :

- Identify the relevant stakeholders
- Establish project goals and objectives
- Elicit requirements from stakeholders
- Document the requirements
- Confirm the requirements
- Prioritize the requirements

3.2 Functional Requirements

- The system should employ Time Series Regression, LSTM, and SARIMA techniques for accurate solar power predictions.
- The system should utilize sophisticated algorithms to consider factors like orientation, shading, and available space for optimal solar panel placement.
- It must differentiate between regular maintenance needs and critical issues based on maintenance frequency, panel age, and environmental conditions.

3.3 Non-Functional Requirements

- The system should be scalable to accommodate a growing number of users and expanding datasets over time.

- The system must provide clear and easily understandable reports and visualizations for solar power predictions, rooftop analyses, and maintenance alerts.
- The system should ensure high availability, minimizing downtime for critical functionalities such as power generation forecasting and predictive maintenance.
- The system must implement robust security measures to protect sensitive user data, including location information and energy usage patterns.
- It must be optimized for efficient data processing, ensuring timely updates and calculations even with large datasets.

3.4 Hardware, Software, Technology and tools utilized

- **Python**:- Python is a high-level, general-purpose programming language. Its design philosophy emphasizes code readability with the use of significant indentation.
- **Reactjs**:- React (also known as React.js or ReactJS) is a free and open-source front-end JavaScript library for building user interfaces based on UI components. It is maintained by Meta (formerly Facebook) and a community of individual developers and companies. React can be used as a base in the development of single-page, mobile, or server-rendered applications with frameworks like Next.js.
- **Scikit-learn**:- Scikit-learn is a popular machine learning library in Python, offering a simple and efficient tool for data analysis and modelling. It provides a wide range of supervised and unsupervised learning algorithms for tasks like classification, regression, clustering, and dimensionality reduction.
- **Tensorflow**:- It offers a flexible ecosystem for deep learning projects, enabling tasks like image and speech recognition, natural language processing, and reinforcement learning through its comprehensive API and computational graph abstraction.
- **Tailwind**:- Tailwind CSS is a utility-first CSS framework that streamlines the process of building modern, responsive web interfaces. It provides a set of pre-designed utility classes that you can apply directly in your HTML markup, enabling rapid development without writing custom CSS.
- **Vscode**:-Visual Studio Code is a streamlined code editor with support for development operations like debugging, task running, and version control. It aims to provide just the tools a developer needs for a quick code-build-debug cycle and leaves more complex workflows to fuller featured IDEs, such as Visual Studio IDE.

- **Google Colab**:- Colab allows anybody to write and execute arbitrary Python code through the browser and is especially well suited to machine learning, data analysis and education. More technically, Colab is a hosted Jupyter notebook service that requires no setup to use, while providing access free of charge to computing resources including GPUs.

3.5 Constraints

- Internet Access is required.
- The forecasting data needs to be put manually
- The images need to be uploaded in fault detection

Chapter 4: Proposed Design

4.1 Block diagram of the system

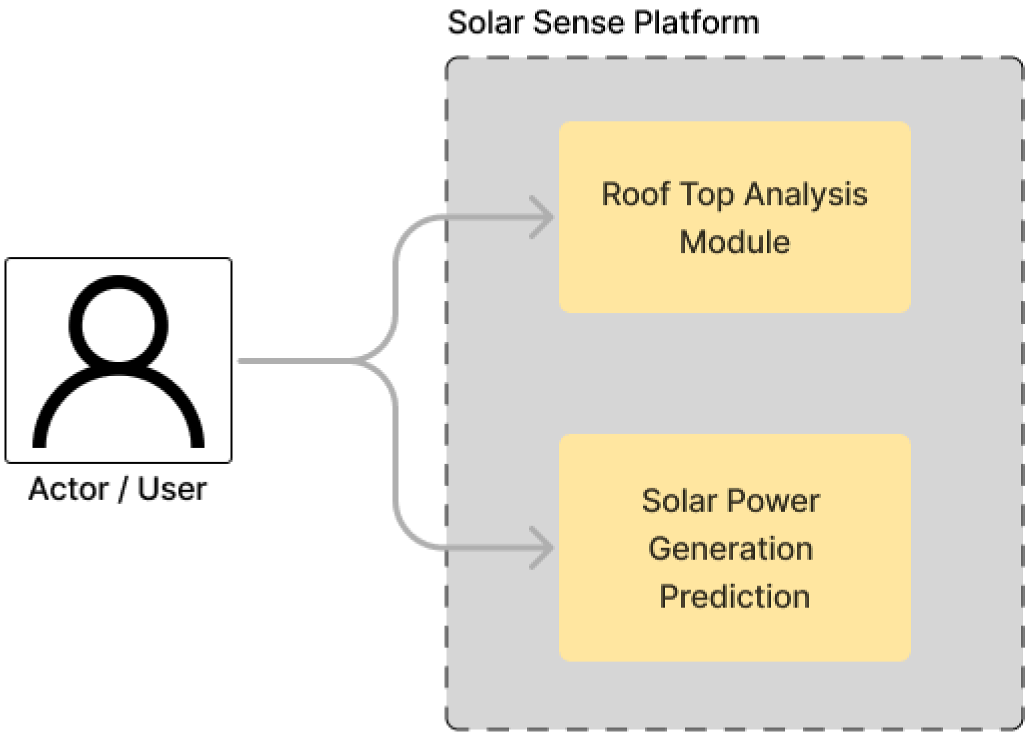


Fig 4.1: Block Diagram

- Rooftop Analysis Module: This block likely analyzes data about a physical rooftop, possibly including its dimensions, geographic location, and materials. Using this information, the module might estimate how much solar energy the rooftop could generate.
- Solar Power Generation Prediction Module: This block likely uses weather data and, potentially, rooftop data from the Rooftop Analysis Module to predict how much solar energy the rooftop will generate in the future.
- Actor/User: This block refers to the person or system that interacts with the Solar Sense Platform. An actor might be a homeowner or business looking to install solar panels, or it could be a utility company managing a solar energy grid.

Overall, the Solar Sense Platform appears designed to help users estimate how much solar energya rooftop can produce and predict future generation. This information could be valuable for making informed decisions about solar panel installation and energy usage.

4.2 Modular design of the system

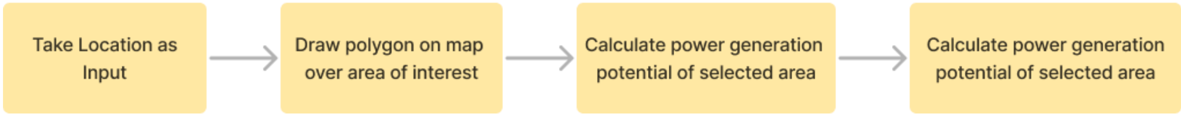


Fig 4.2 Modular Design

- Take Location as Input: The first step involves entering the location of the area of interest. This could be done by specifying coordinates or an address.
- Draw Polygon on Map: The user defines a specific area on a map by outlining a polygon. This polygon likely represents the area where solar panels would be installed.
- Calculate Potential of Selected Area: The system calculates the potential solar power generation for the chosen area. Factors influencing this calculation might include the area of the polygon and historical or real-time data on solar irradiance in that location.

Overall, the flowchart outlines a simplified method for estimating how much solar energy a specific location could generate. This information could be useful in determining the feasibility and potential benefits of installing a solar energy system at a particular site.

4.3 Detailed Design

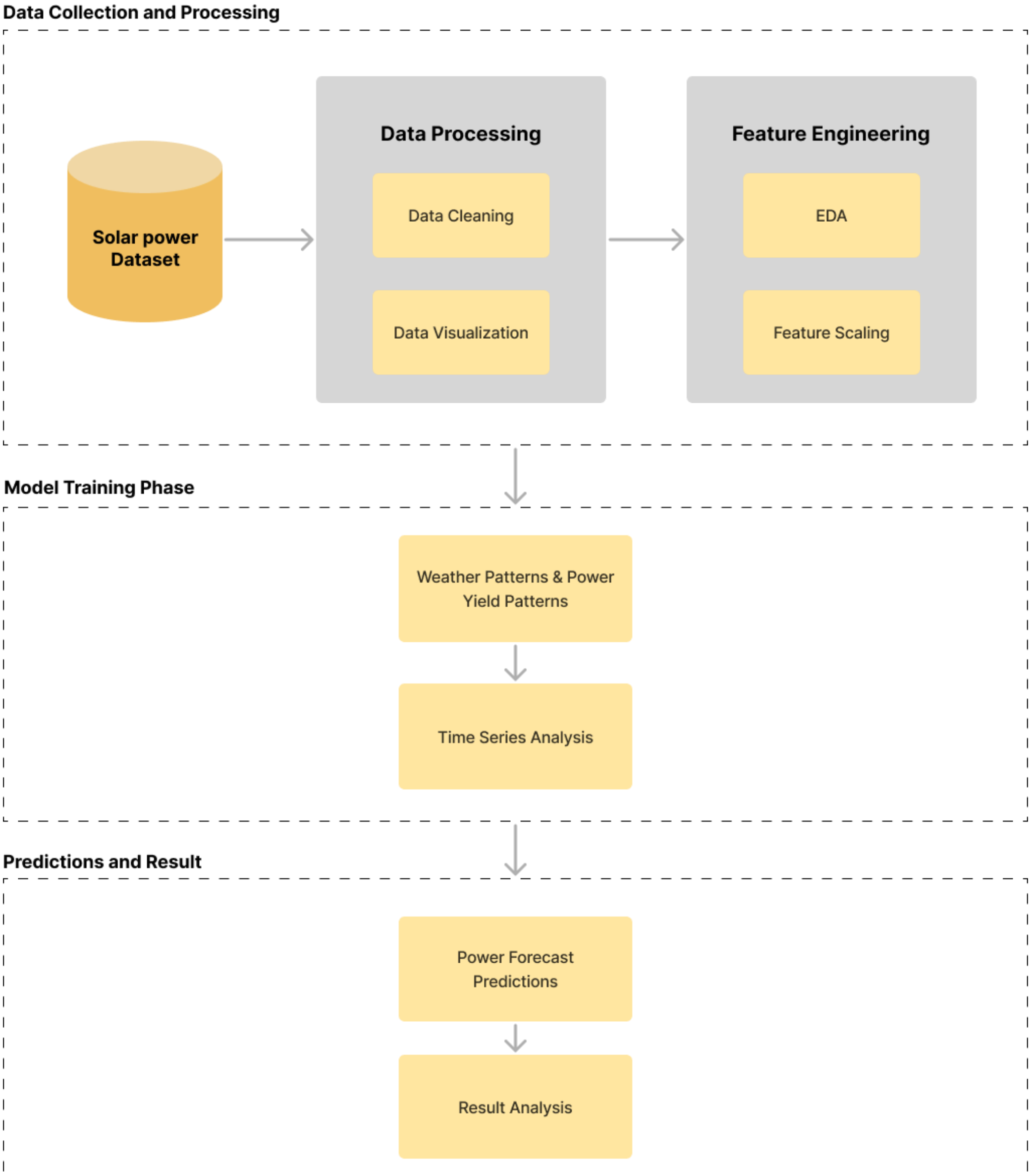


Fig 4.3 Design outlining a solar power data collection and processing system

- Solar Power Dataset: This block represents a repository of historical or real-time solar power generation data. The source of this data could be solar panels, weather stations, or a combination of both.
- Data Cleaning: This block suggests a step to remove errors or inconsistencies from the raw solar power data. This might involve handling missing values, outliers, or invalid data points.
- Data Preprocessing: This block likely encompasses a series of transformations applied to the cleaned data to prepare it for further analysis. Common preprocessing

techniques include normalization, scaling, and feature engineering (creating new features from existing ones).

- Exploratory Data Analysis (EDA): This block signifies a stage where the data is visualized and analyzed to understand its characteristics, identify patterns, trends, and potential relationships between variables.
- Feature Engineering: This block might involve creating new features based on existing ones in the data. For instance, extracting features like time of day, day of week, or season from timestamps could be beneficial for analysis.
- Model Training: This block indicates the stage where a machine learning model is trained on the preprocessed data. The specific model type (e.g., decision tree, neural network) would depend on the intended purpose of the system.
- Weather Patterns & Power Yield Patterns: This block likely refers to the model's ability to identify relationships between weather patterns and solar power output. This knowledge could be used for tasks like anomaly detection or short-term forecasting.

4.4 Project Scheduling & Tracking using Timeline / Gantt Chart

The Gantt chart of our project where we worked for the whole semester to create this model is shown in a timeline pattern. It is the most important part of thinking and designing the planning of your topic and so we planned our work like the Gantt chart shown.

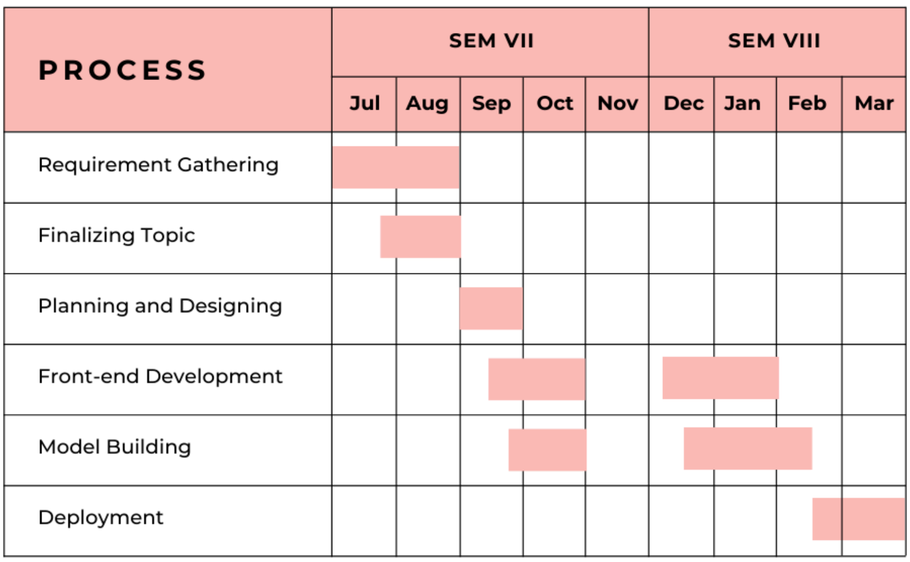


Fig 4.4 Gantt Chart

Chapter 5: Implementation of the Proposed System

5.1. Methodology employed for development

1) Solar panel fault detection:

Our approach to solar panel fault detection involves a comprehensive investigation leveraging machine learning techniques. Recognizing the criticality of maintaining optimal solar panel performance, we aim to develop an effective procedure for identifying common faults such as dust accumulation, snow cover, bird droppings, and physical or electrical damage on solar panel surfaces. To achieve this objective, we employ a model vision transformer, a state-of-the-art deep learning architecture renowned for its ability to handle image data effectively.

The methodology encompasses several key steps. Firstly, we gather a diverse dataset comprising images categorizing various fault types, including clean panels, dusty panels, bird droppings, electrical damage, physical damage, and snow-covered panels. This dataset serves as the foundation for training and evaluating our fault detection model.
Next, we preprocess the dataset to ensure uniformity and prepare it for model training. This involves tasks such as data cleaning, augmentation, and normalization to enhance the model's robustness and generalization capabilities.

Subsequently, we employ the model vision transformer architecture to train a deep learning model capable of accurately detecting and classifying different fault types on solar panel surfaces. The vision transformer model is chosen for its superior performance in handling image data and its ability to capture intricate patterns and features essential for fault detection.

During the training phase, the model learns to distinguish between various fault categories by analyzing the visual characteristics of the input images. Through iterative optimization and validation processes, the model refines its parameters to maximize accuracy and minimize errors in fault detection.

Upon model convergence, we evaluate its performance using a separate validation dataset to assess its ability to generalize to unseen data. This evaluation involves metrics such as precision, recall, and F1-score to quantify the model's effectiveness in accurately identifying different fault types.

In conclusion, our methodology encompasses the utilization of a model vision transformer architecture for solar panel fault detection, aiming to enhance maintenance efficiency and optimize solar panel performance. Through rigorous experimentation and evaluation, we aim to develop a robust and reliable fault detection system capable of supporting sustainable and efficient solar energy utilization.

2) Anomaly Detection in photovoltaic (PV) systems

The Isolation Forest algorithm consists of the following steps:

- Tree Construction: Randomly select a feature and split the data along that feature's range.
- Isolation: Continue partitioning the data recursively until isolation is achieved.
- Anomaly Score Calculation: Compute the anomaly score for each data point based on its average path length in the trees.
- Anomaly Detection: Identify anomalies as data points with lower anomaly scores.

Anomaly Classification Using Random Forest Model:

After detecting anomalies using Isolation Forest, the subsequent phase involves classifying the type of anomaly utilizing a Random Forest classifier.

Feature Selection: For effective anomaly classification, it is imperative to select relevant features that contribute significantly to the classification task. This is typically accomplished through various techniques, including:

Recursive Feature Elimination (RFE): RFE is a feature selection method that iteratively removes features from the dataset based on their importance to the model's performance. The Random Forest classifier is trained multiple times, each time eliminating the least important features until the optimal subset of features is determined.

Feature Importance from the Isolation Forest model: The Isolation Forest algorithm used for anomaly detection can provide insights into the importance of each feature in distinguishing between normal and anomalous instances. Features with higher importance scores from the Isolation Forest model are likely to be more informative for anomaly classification and are thus prioritized for inclusion in the Random Forest classifier.

SHAP (Shapley Additive exPlanations): SHAP is a method used to explain the output of machine learning models by quantifying the contribution of each feature to the model's predictions. By analyzing SHAP values, we can identify the most influential features in the Random Forest model and use them for anomaly classification. In this study, SHAP analysis reveals the following features as the most important for anomaly classification: POA, POA_CUM, MODULE_TEMP, EAE_DAY_PLANT, and EAE_DAY.

Random Forest Model Training: Once the relevant features are selected, the Random Forest classifier is trained on the preprocessed dataset to classify the detected anomalies. The training process involves the following steps:

By incorporating SHAP analysis and Random Forest classification, we can effectively identify and classify anomalies in the dataset, providing valuable insights for anomaly detection and mitigation strategies.

3) Rooftop Analysis:

The rooftop analysis process begins by obtaining a geographical location as input. This location may be specified using latitude and longitude coordinates or an address. Once the location is acquired, the subsequent step involves drawing a polygon on a map to delineate the area of interest within that location. This allows for precise selection of a specific region or boundary within the given geographical area.

Following the drawing of the polygon, the total area enclosed by the polygon is calculated. This calculation provides insights into the size or extent of the selected area within the geographical location. Utilizing this calculated area, the next step involves assessing the power generation potential of the selected area. This assessment incorporates factors such as solar exposure, shading, and solar panel efficiency to estimate the amount of solar power that can be generated within the chosen region.

By following this methodology, a comprehensive analysis of the rooftop's solar energy generation potential can be conducted. This analysis lays the foundation for the development and implementation of an efficient smart solar system, crucial for sustainable energy management and utilization.

4) Solar Power Generation Prediction:

The analysis of solar power generation commences with the Solar Power Dataset, a crucial starting point containing historical or real-time data on solar irradiance, temperature, time of day, and power output. This dataset serves as the cornerstone for subsequent phases of analysis.

The initial phase, Data Processing, involves several critical sub-processes. Data Cleaning addresses missing data, outliers, and inconsistencies, ensuring the dataset's reliability. Data Visualization techniques are then applied to gain insights into data patterns, trends, and relationships, utilizing graphical tools such as graphs and charts. Feature Engineering follows, where new features are created or existing ones modified to enhance the predictive power of the model. This includes Exploratory Data Analysis (EDA) for deeper insights and Feature Scaling to standardize data, optimizing machine learning model performance.

Transitioning to the Model Training Phase, machine learning models are trained using the pre-processed and engineered dataset. Notably, special emphasis is placed on capturing Weather Patterns and Power Yield Patterns, crucial for accurate power generation predictions. Time Series Analysis techniques are employed to understand and model temporal dependencies within the solar power generation data.

In the final phase, Predictions and Results are derived. Leveraging the trained models and current weather data, Power Forecast Predictions are generated, facilitating effective planning and optimization of solar energy utilization. This comprehensive methodology provides a systematic approach to harnessing solar power efficiently within the smart solar system framework.

5.2. Algorithms and Flowcharts for the respective modules developed:

This section details the algorithms used in each module and their corresponding flowcharts for better understanding.

Fault Detection (VGG16):

VGG16 is a convolutional neural network architecture known for its simplicity and effectiveness in image classification tasks. Transfer learning is a technique where a pre-trained model is fine-tuned on a new dataset to perform a specific task.

- **VGG16 Architecture**: VGG16 consists of 16 convolutional layers followed by fully connected layers and softmax activation for classification. The architecture comprises small (3x3) convolutional filters with max-pooling layers in between to reduce spatial dimensions.

- **Transfer Learning**: Transfer learning involves leveraging the knowledge gained from training on a large dataset (e.g., ImageNet) and applying it to a related task with a smaller dataset (e.g., fault detection). In this approach, the pre-trained VGG16 model is used as a feature extractor, and only the top layers are fine-tuned on the new dataset.
- **Training Process**: The training process involves loading the pre-trained VGG16 model, replacing the top layers with new ones for the specific task, freezing the weights of the pre-trained layers, and training the model using the new dataset. This process allows the model to learn features relevant to fault detection while retaining the knowledge gained from the pre-trained weights.

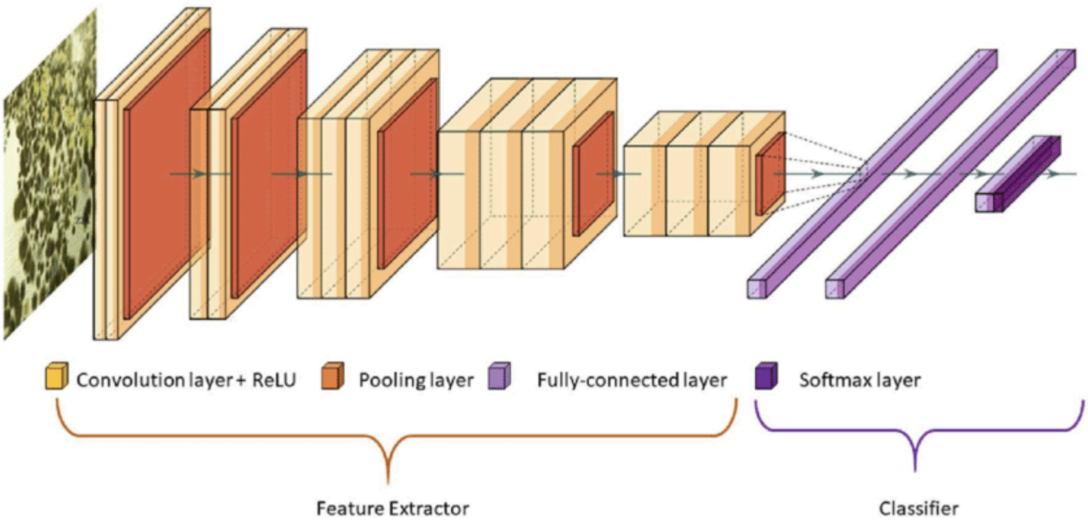


Fig 5.2.1 VGG16 Architecture for fault detection

Anomaly Detection (ResNet50):

- ResNet50 is a deep convolutional neural network architecture known for its ability to handle very deep networks effectively. In anomaly detection, ResNet50 can be used for feature extraction followed by anomaly detection techniques such as reconstruction error or density estimation.
- **ResNet50 Architecture**: ResNet50 comprises 50 layers with skip connections (residual connections) to mitigate the vanishing gradient problem. It uses a bottleneck architecture with 1x1, 3x3, and 1x1 convolutional layers in each residual block.

- **Transfer Learning Approach**: Similar to VGG16, transfer learning with ResNet50 involves utilizing a pre-trained model (often trained on ImageNet) and fine-tuning it for anomaly detection. The pre-trained ResNet50 model serves as a feature extractor, capturing hierarchical features from the input images.
- **Anomaly Detection Technique**: After feature extraction using ResNet50, various anomaly detection techniques can be applied depending on the nature of the problem. This may include methods such as reconstruction error calculation using autoencoders, density estimation using Gaussian mixture models, or threshold-based anomaly detection.

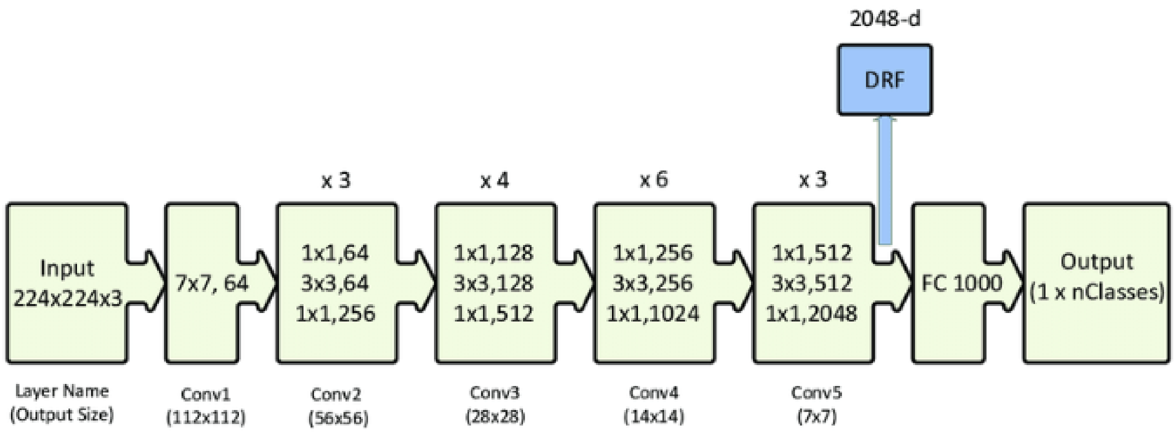


Fig 5.2.2 ResNet50 Architecture for anomaly detection

5.3. Datasets source and utilization:

We have used the “Solar Panel Images Clean and Faulty Images” dataset from Kaggle. The objective of this dataset is to explore the efficacy of different machine learning classifiers in detecting dust, snow, bird droppings, physical damage, and electrical damage on solar panel surfaces with the utmost accuracy.

Directory Overview

Directory comprises six distinct class folders designed for classification purposes. Due to the nature of data collection from internet sources, there exists a slight imbalance in the number of images collected for each class. The folders are organized as follows:

- **Clean**: Contains images of clean solar panels.
- **Dusty**: Contains images of dusty solar panels.
- **Bird-drop**: Contains images of bird droppings on solar panels.
- **Electrical-damage**: Contains images of solar panels afflicted with electrical damage.

- **Physical-Damage**: Contains images of solar panels damaged physically.
- **Snow-Covered**: Contains images of snow-covered solar panels.

Fig 5.3.1: Bird-drop

Fig 5.3.2: Clean

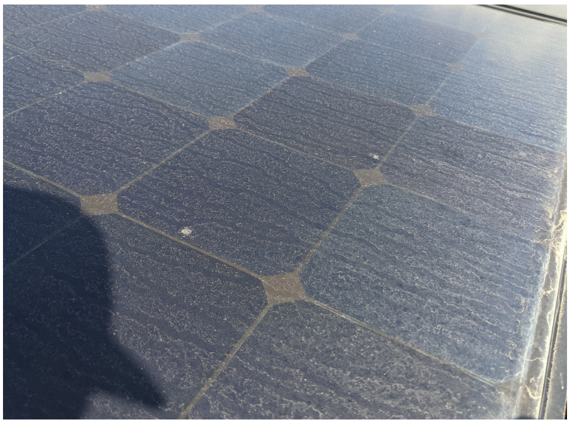


Fig 5.3.3: Dusty

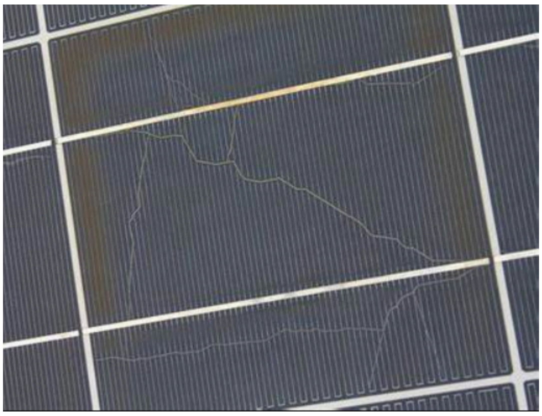


Fig 5.3.4: Electrical Damage

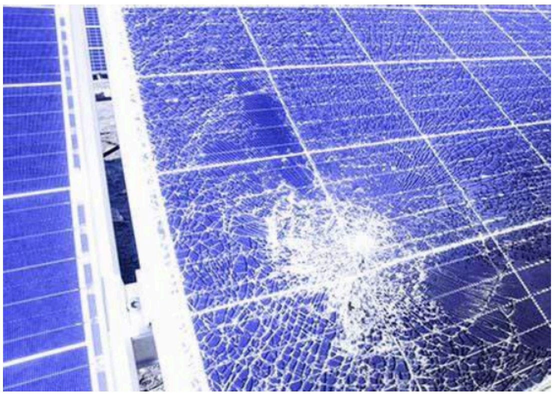


Fig 5.3.5: Physical Damage

Fig 5.3.6: Snow-covered

Chapter 6: Results and Discussions

6.1. Screenshot of Use Interface(UI) for the system:

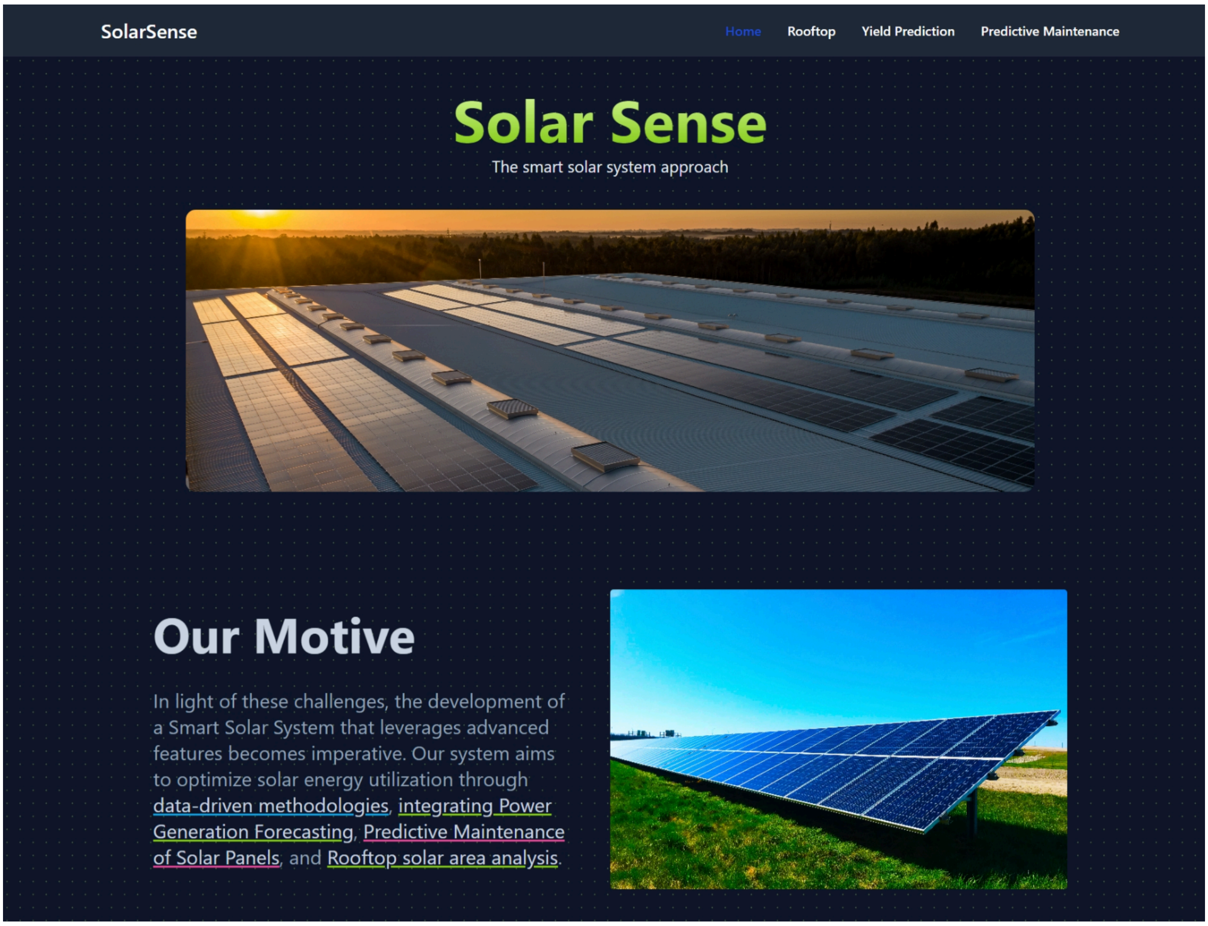


Fig 6.1.1 Home Page

This is the Home Page of our website from where users can navigate to different modules.

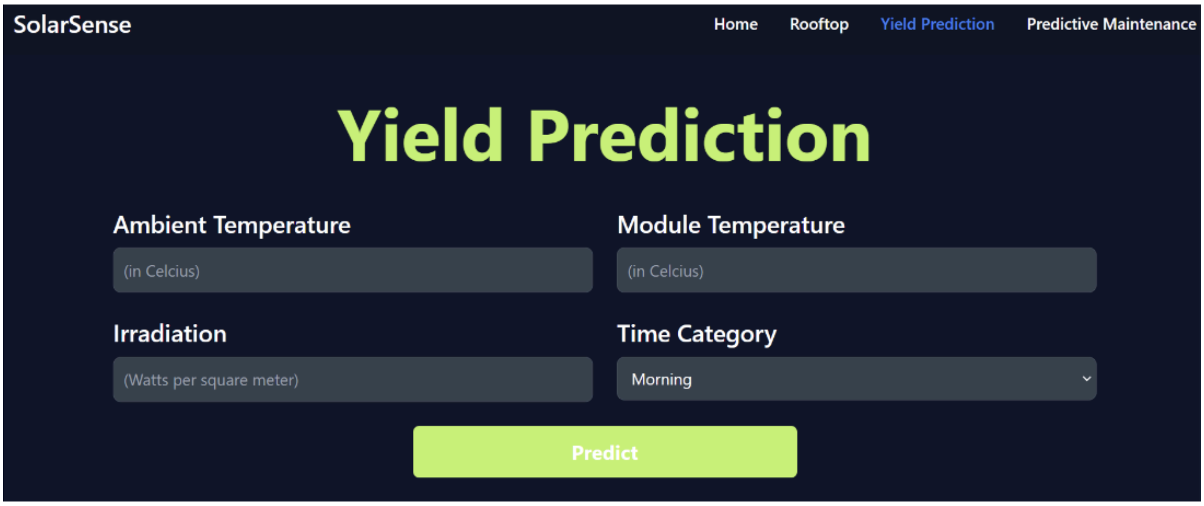


Fig 6.1.2: Yield Prediction page

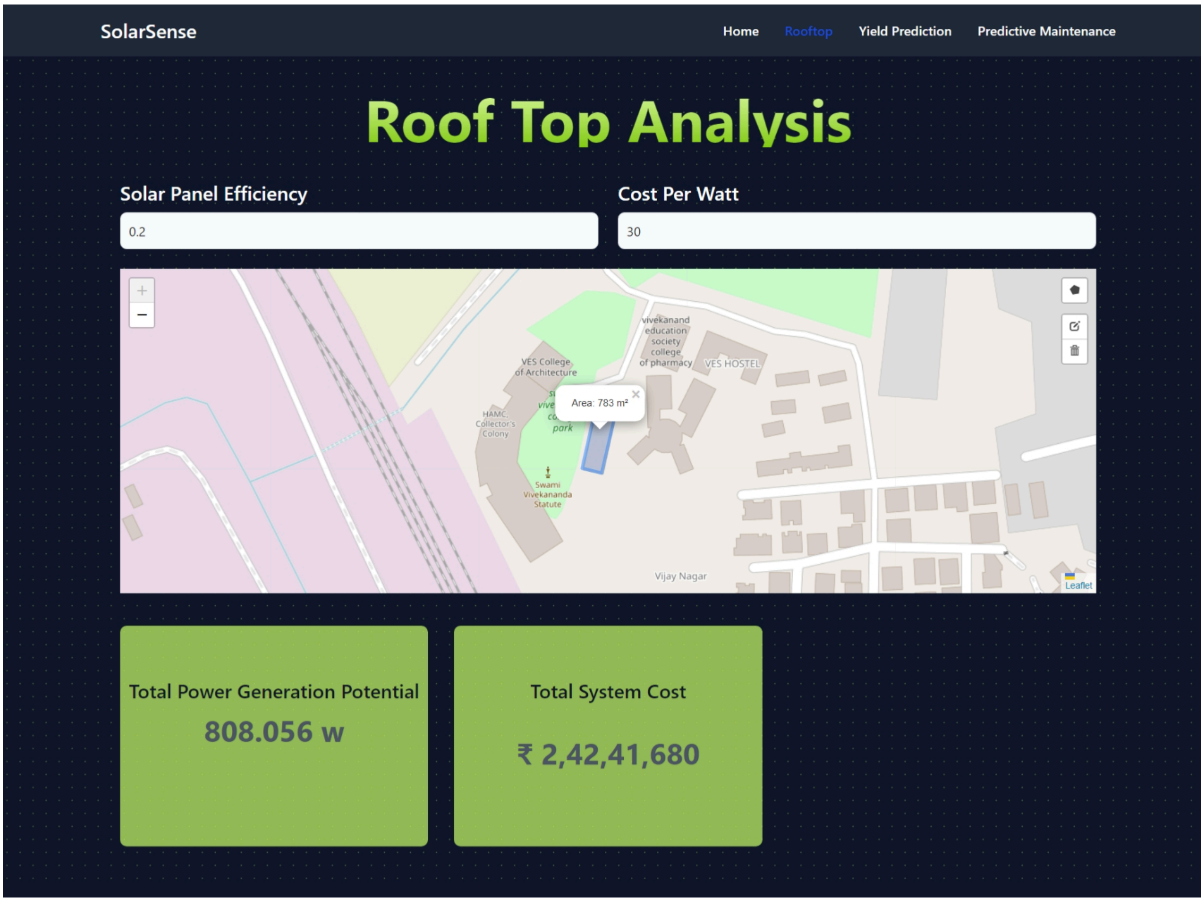


Fig 6.1.3: Rooftop analysis page

This is the Rooftop analysis module where users can get the approximate total power generation potential of the solar system and total system cost. Users first select the area that they want to install solar panels at and then input the features of the solar panel like solar panel efficiency and cost per watt.

Fig 6.1.4: Predictive Maintenance page

6.2. Model Training and Performance

These are the "**Training v/s Validation Loss**" and "**Training v/s Validation accuracies**" graphs for the proposed model.

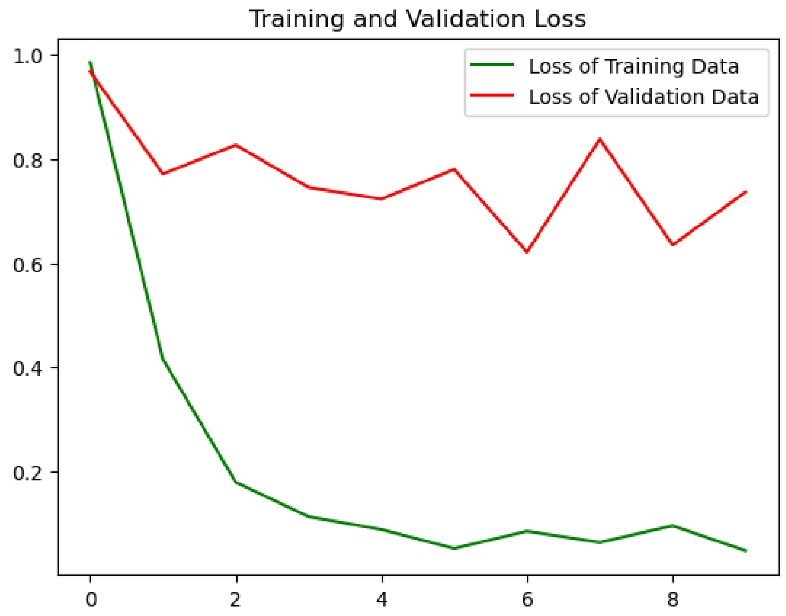


Fig 6.2.1 Training and Validation Loss

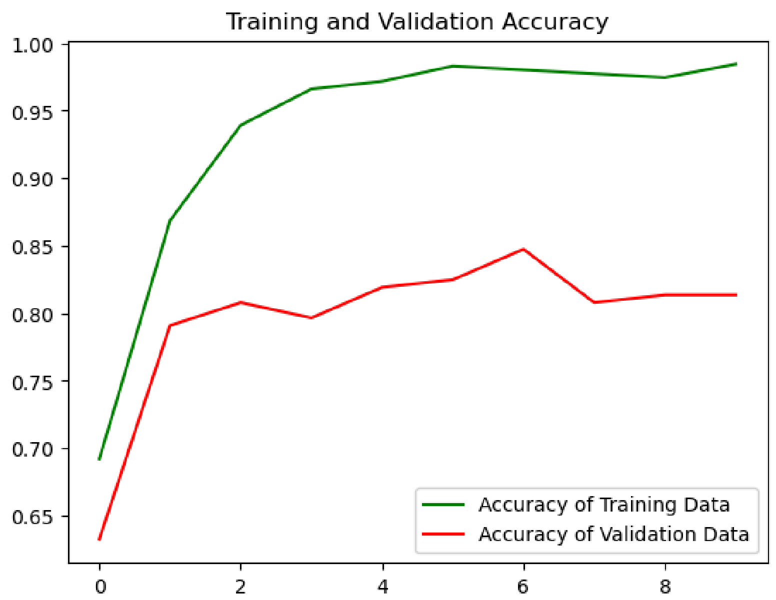


Fig 6.2.2 Training and Validation Accuracy

6.3. Performance Evaluation Measures:

1. **Precision**: Precision is one indicator of a machine learning model's performance – the quality of a positive prediction made by the model. Precision refers to the number of true positives divided by the total number of positive predictions (i.e., the number of

true positives plus the number of false positives).

The formula is:

Precision = TP / (TP + FP)

where:

TP = True Positives,

FP = False Positives.

2. **Recall**: The recall is calculated as the ratio between the number of Positive samples correctly classified as Positive to the total number of Positive samples. The recall measures the model's ability to detect positive samples. The higher the recall, the more positive samples detected.

   The formula is:

   Recall = TP / (TP + FN)

   where:

   TP = True Positives,

   FN = False Negatives.

3. **F-Score**: The F-score (also known as the F1 score or F-measure) is a metric used to evaluate the performance of a Machine Learning model. It combines precision and recall into a single score. The

   formula is:

   F-score = 2*(Precision * Recall) / (Precision + Recall)

6.4. Input Parameters/Features considered:

This section details the specific data elements used as inputs for the various modules of the smart solar system.

1. Fault Detection

Image Data:

- Image Type: Input whether RGB or thermal images are analyzed for fault detection.
- Resolution: Indicate the image dimensions used, if they have a significant bearing on fault detection accuracy.
- Fault Examples: Briefly list the types of faults the system is designed to detect (e.g., dust, bird droppings, physical damage).

2. Anomaly Detection

Solar Power Data:

- Parameters: The key parameters are irradiance, temperature, power output, voltage, current.
- Time Resolution: The temporal granularity of the data (e.g., hourly, daily, etc.).

3. Rooftop Analysis

Geographic Data:

- Coordinates or Address: The type of location data used.
- Mapping Precision: If relevant, mention the desired level of precision for rooftop boundary delineation.

4. Power Generation Prediction

Historical Solar & Weather Data:

- Parameters: Parameters such as irradiance, solar panel efficiency, cost per watt and area was considered.
- Area: Calculated rooftop area from the Rooftop Analysis module.

Chapter 7: Conclusion

7.1. Limitations:

As part of our system's current limitations, we acknowledge that it lacks integration with real-time weather data, which hampers its ability to accurately predict solar power generation based on prevailing weather conditions. Furthermore, our predictive maintenance feature solely relies on capturing images of solar panels, neglecting to monitor essential parameters of inverters crucial for maintenance prediction. This limitation prevents us from providing comprehensive and proactive maintenance scheduling based on actual equipment conditions, potentially resulting in inefficiencies and increased downtime. Addressing these limitations by integrating real-time weather data and expanding predictive maintenance capabilities to include inverter parameters will be crucial for optimizing the performance and reliability of our system in the future.

7.2. Conclusion:

The Smart Solar System represents a revolutionary and state-of-the-art solution that harnesses the power of data-driven intelligence to achieve the pinnacle of solar energy utilization. By seamlessly integrating precise Power Generation Forecasting, meticulous Rooftop Analysis, and proactive Predictive Maintenance, this system stands as a pioneering force in the pursuit of heightened efficiency, unwavering reliability, and unwavering sustainability.

The innovative approach taken by the Smart Solar System not only sets new standards in the world of solar energy but also plays a pivotal role in catalyzing our global shift toward a greener, more sustainable, and profoundly efficient energy landscape. Our future endeavors encompass an unwavering commitment to advancing Predictive Maintenance through the incorporation of cutting-edge machine learning techniques, ensuring that we remain at the forefront of the solar energy revolution. In doing so, we continue to fuel the momentum behind the adoption of solar energy and expedite the trajectory toward a cleaner and more efficient energy future on a global scale.

7.3. Future Scope:

In the future scope of our project, we aim to enhance the efficiency and accuracy of our system by integrating real-time weather data. By incorporating this data into our platform, we will be able to predict solar power generation in real time, optimizing energy production based on current weather conditions. Additionally, we plan to improve the predictive maintenance module by implementing advanced anomaly detection algorithms to identify and address issues such as current fluctuations and system tripping by taking into consideration the parameters/data of inverters promptly. This integration will not only improve the reliability of our system but also contribute to more sustainable and efficient energy management practices.

References

1. R. J. Bessa, A. Trindade, and V. Miranda, "Spatial-Temporal Solar Power Forecasting for Smart Grids," IEEE Transactions on Industrial Informatics, vol. 11, no. 1, pp. 232-241, 2015.
2. J. Vicente-Gabriel, A.-B. Gil-González, A. Luis-Reboredo, P. Chamoso, and J. M. Corchado, "LSTM Networks for Overcoming the Challenges Associated with Photovoltaic Module Maintenance in Smart Cities," Electronics, vol. 10, no. 1, p. 78, Jan. 2021, doi: 10.3390/electronics10010078.
3. B. O. Olorunfemi, O. A. Ogbolumani, and N. Nwulu, "Solar Panels Dirt Monitoring and Cleaning for Performance Improvement: A Systematic Review on Smart Systems," Sustainability, vol. 14, no. 17, p. 10920, Sep. 2022
4. R. K. Sahu, B. Shaw, J. R. Nayak, and S. Shashikant, "Short/medium-term solar power forecasting of Chhattisgarh state of India using modified TLBO optimized ELM," Engineering Science and Technology, an International Journal, vol. 24, no. 5, pp. 1180-1200, 2021. ISSN 2215-0986.
5. G. Mitrentsis and H. Lens, "An interpretable probabilistic model for short-term solar power forecasting using natural gradient boosting," Applied Energy, vol. 309, p. 118473, 2022. ISSN 0306-2619.
6. N. Elizabeth Michael, M. Mishra, S. Hasan, and A. Al-Durra, "Short-Term Solar Power Predicting Model Based on Multi-Step CNN Stacked LSTM Technique," Energies, vol. 15, no. 6, p. 2150, Mar. 2022, doi: 10.3390/en15062150.
7. S. Theocharides, G. Makrides, A. Livera, M. Theristis, P. Kaimakis, and G. E. Georghiou, "Day-ahead photovoltaic power production forecasting methodology based on machine learning and statistical post-processing," Applied Energy, vol. 268, p. 115023, 2020. ISSN 0306-2619.
8. I. Jebli, F.-Z. Belouadha, M. I. Kabbaj, and A. Tilioua, "Prediction of solar energy guided by Pearson correlation using machine learning," Energy, vol. 224, p. 120109, 2021. ISSN 0360-5442.
9. A. Gensler, J. Henze, B. Sick and N. Raabe, "Deep Learning for solar power forecasting — An approach using AutoEncoder and LSTM Neural Networks," 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Budapest, Hungary, 2016, pp. 002858-002865, doi: 10.1109/SMC.2016.7844673.
10. A. S. Shihavuddin et al., "Image based surface damage detection of renewable energy

installations using a unified deep learning approach," Energy Reports, vol. 7, pp. 4566-4576, 2021. ISSN 2352-4847.
11. Shadab, A., Ahmad, S., & Said, S. (2020). Spatial forecasting of solar radiation using ARIMA model. Remote Sensing Applications: Society and Environment, 20. https://doi.org/10.1016/j.rsase.2020.100427
12. Colak, I., Yesilbudak, M., Genc, N., & Bayindir, R. (2016). Multi-period prediction of solar radiation using ARMA and ARIMA models. Proceedings - 2015 IEEE 14th International Conference on Machine Learning and Applications, ICMLA 2015. https://doi.org/10.1109/ICMLA.2015.33
13. Sharadga, H., Hajimirza, S., & Balog, R. S. (2020). Time series forecasting of solar power generation for large-scale photovoltaic plants. Renewable Energy, 150. https://doi.org/10.1016/j.renene.2019.12.131
14. Maroua Haddad, Jean Nicod, Yacouba Boubacar Mainassara, Landy Rabehasaina, Zeina Al Masry, et al.. Wind and Solar Forecasting for Renewable Energy System using SARIMA-based Model. International conference on Time Series and Forecasting, Sep 2019, Gran Canaria, Spain.
15. Vrettos, E., & Gehbauer, C. (2019). A hybrid approach for short-term PV power forecasting in predictive control applications. 2019 IEEE Milan PowerTech, PowerTech 2019. https://doi.org/10.1109/PTC.2019.8810672
16. Bouzerdoum, M., Mellit, A., & Massi Pavan, A. (2013). A hybrid model (SARIMA-SVM) for short-term power forecasting of a small-scale grid-connected photovoltaic plant. Solar Energy, 98(PC). https://doi.org/10.1016/j.solener.2013.10.002
17. Vagropoulos, S. I., Chouliaras, G. I., Kardakos, E. G., Simoglou, C. K., & Bakirtzis, A. G. (2016). Comparison of SARIMAX, SARIMA, modified SARIMA and ANN-based models for short-term PV generation forecasting. 2016 IEEE International Energy Conference, ENERGYCON 2016. https://doi.org/10.1109/ENERGYCON.2016.7514029
18. Silva, V. L. G. da, Oliveira Filho, D., Carlo, J. C., & Vaz, P. N. (2022). An Approach to Solar Radiation Prediction Using ARX and ARMAX Models. Frontiers in Energy Research, 10. https://doi.org/10.3389/fenrg.2022.822555
19. Geetha, A., Santhakumar, J., Sundaram, K. M., Usha, S., Thentral, T. M. T., Boopathi, C. S., Ramya, R., & Sathyamurthy, R. (2022). Prediction of hourly solar radiation in Tamil Nadu using ANN model with different learning algorithms. Energy Reports, 8.

https://doi.org/10.1016/j.egyr.2021.11.190

20. Abuella, M., & Chowdhury, B. (2015). Solar power forecasting using artificial neural networks. 2015 North American Power Symposium, NAPS 2015. https://doi.org/10.1109/NAPS.2015.7335176
21. Srivastava, R., Tiwari, A. N., & Giri, V. K. (2018). Forecasting of Solar Radiation in India Using Various ANN Models. 2018 5th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering, UPCON 2018. https://doi.org/10.1109/UPCON.2018.8597170
22. Alomari, M. H., Adeeb, J., & Younis, O. (2018). Solar photovoltaic power forecasting in Jordan using artificial neural networks. International Journal of Electrical and Computer Engineering, 8(1). https://doi.org/10.11591/ijece.v8i1.pp497-504
23. Pattanaik, D., Mishra, S., Khuntia, G. P., Dash, R., & Swain, S. C. (2020). An innovative learning approach for solar power forecasting using genetic algorithms and artificial neural networks. Open Engineering, 10(1). https://doi.org/10.1515/eng-2020-0073

Appendix

a. Paper I

A Survey of Photovoltaic power forecasting approches in Solar energy Landscape

Anishkumar Mahalingam
Department of Computer Engineering
VESIT
Mumbai, India
2020.anishkumar.mahalingam@ves.ac.in

Swapnil Thatte
Department of Computer Engineering
VESIT
Mumbai, India
2020.swapnil.thatte@ves.ac.in

Yash Narkhede
Department of Computer Engineering
VESIT
Mumbai, India
2020.yash.narkhede@ves.ac.in

Yash Brid
Department of Computer Engineering
VESIT
Mumbai, India
d2020.yash.brid@ves.ac.in

Dr.Dashrath Mane
Deptartment of Computer Engineering
VESIT
Mumbai, India
dashrath.mane@ves.ac.in

***Abstract*—The integration of solar energy into the modern power grid faces challenges due to its inherent variability. Accurate solar forecasting plays a vital role in mitigating these challenges and enabling the smooth integration of solar resources. This paper presents a comprehensive overview of solar forecasting techniques, classifying them according to their prediction horizons (short-term, medium-term, and long-term). The evolution of forecasting methodologies is explored, ranging from NWP models and statistical approaches to advanced deep learning techniques like LSTM networks and autoencoders. The suitability of these techniques is discussed in the context of lead time, accuracy requirements, and computational resources. This review paper aims to provide a structured analysis of the present landscape of solar power estimation, highlight strengths and weaknesses of various methods, and guide future research efforts towards even more refined and robust forecasting solutions for promoting renewable energy integration.**

***Index Terms*—Solar forecasting, Time-series analysis, LSTM, ARIMA/SARIMA**

I. Introduction

Renewable energy sources, particularly solar power, are essential for transitioning to a more sustainable energy grid. But grid managers have a great difficulty because of the intrinsic unpredictability of solar power output. Precise solar forecasting techniques are crucial in mitigating these challenges by providing advanced predictions of solar power production. These forecasts enable informed decision-making, optimized scheduling, and effective grid integration of solar resources. This paper explores the evolution of solar forecasting, its different classifications, and the cutting-edge techniques used for reliable and accurate predictions.

Solar forecasting methodologies are broadly categorized based on their prediction horizons. Short-term forecasts spanning minutes to hours ahead support near-real-time decisions for maintaining grid stability, while medium-term forecasts over months or seasons assist in utility planning. Long-term forecasts guide decisions about future capacity expansion and project development. Forecasting methodologies have steadily advanced in complexity, including weather-driven Numerical Weather Prediction models, persistence-based approaches, statistical time-series techniques like ARIMA models, and recent deep learning architectures such as LSTMs and autoencoders.

The choice of optimal forecasting techniques depends on several factors, including the lead time desired, the required level of accuracy, and the availability of computational resources. In order to propose possible directions for future research, this study will critically assess the benefits and drawbacks of the most recent solar forecasting methods. This systematic investigation provides a comprehensive foundation for continued research and advancement toward even more precise and adaptable solar forecasting models to support the growth of renewable energy and reliable grid operation.

II. Classification of Solar Forecasts

A. Short-Term Forecasts

1) Intra-day Forecasts: Intra-day forecasting provides continuous solar power predictions from minutes to hours ahead, crucial for balancing supply and demand. These near-real-time forecasts on 5 minute to 1 hour timescales use weather modeling, machine learning, and live data to capture short term PV output fluctuations. By factoring in latest cloud cover, irradiance, and panel temperature impacts, intra-day models help operators smoothly ramp other generators to offset solar changes. Intra-day models utilize high resolution, localized weather and PV installation data, unlike day-ahead forecasts relying on regional models. With frequent model updates and granular forecasts, intra-day forecasting enables smooth solar integration despite intermittent output from passing clouds.

2) Day-ahead Forecasts: Day-ahead forecasting predicts solar output 24 hours in advance, supporting energy operations like market scheduling and utility load planning. Made each morning, day-ahead models use weather prediction data, statistics, and machine learning to estimate next day hourly PV production. Models incorporate parameters like temperature, humidity, clouds, and wind. Historical data trains machine learning algorithms on daily/seasonal patterns. On-site sensor insights provide current conditions. More accurate day-ahead solar forecasts allow traders to better schedule operations and transmission. Utilities can smooth variable solar generation in next day load forecasts. Improved forecasts also help plant operators schedule maintenance and bidding.

B. Medium-term Forecasts

1) Seasonal Forecasts: Seasonal forecasting predicts solar output over months or a whole season, crucial for utility planning, revenue estimation, and maintenance scheduling. Seasonal models use medium-term weather forecasts, historical averages, and machine learning. Inputs include expected temperature, precipitation, humidity, and cloud cover monthly or seasonally. Historical data trains algorithms on seasonal solar patterns. Understanding expected generation by season helps utilities plan for electricity demand. Grid operators can ramp up other sources if a cloudy/rainy season reduces solar output. Plant owners depend on seasonal forecasts to estimate revenue and schedule maintenance downtime during seasons with lower expected output.

C. Long-term Forecasts

1) Year-ahead forecasts: Year-ahead forecasting predicts solar generation one or more years ahead, supporting capacity planning, project development, and strategic decisions. Models use historical averages for typical annual generation, long-term weather projections for irradiation and temperature, and projected solar deployment. Understanding estimated solar generation a year out helps utilities plan long-term capacity needs. Knowing future solar growth helps transmission planners accommodate extra generation. Developers can assess project viability. However, substantial uncertainties in multi-year weather/climate models and solar deployment limit accuracy. Still, even imprecise year-ahead forecasts provide useful information for high-level planning purposes.

III. Forecasting Techniques

A. Numerical Weather Prediction Model

This type of models are computer programs that use current weather conditions as inputs to forecast the weather ahead based on mathematical equations of atmospheric physics and motion. NWP models are a key component of many solar forecasting approaches. They provide predictions of cloud cover, temperature, precipitation and other weather parameters that influence solar irradiance at the earth's surface. By modeling weather systems and patterns, NWP can estimate the solar energy available days or weeks in advance[5].

1) Weather Forecasting Model: The WRF is a widely employed mesoscale numerical weather prediction (NWP) system renowned for its versatility and accuracy. It provides various configurations for solar forecasting like precise representation of cloud microphysics, radiation transfer, and aerosol-cloud interactions. National Center for Atmospheric Research's (NCRA) WRF-SOLAR is an augmented version of standard WRF model, a comprehensive solar forecasting system developed with special emphasis on DNI, DIF, GHI[1]. Modifications were made to the radiation schemes to allow input of aerosol optical properties. This enables representing aerosol-radiation interactions and direct effects. Aerosol datasets can be specified, including climatologies or time-varying reanalysis products. This provides atmosphere aerosol representation. Cloud-aerosol feedbacks were introduced by coupling aerosols in microphysics to radiation. This allows aerosol indirect effects on clouds. Shallow cumulus clouds now provide sub-grid cloud fraction feedbacks to radiation. This reduces positive bias in irradiance. Clear-sky validation showed large improvements in irradiance accuracy from the aerosol effects, especially using reanalysis aerosol data. WRF-Solar significantly reduced errors in clear-sky GHI, DNI and DIF compared to standard WRF.

2) North American Mesoscale Forecast System: The NAM model provides short-range weather forecasts out to 84 hours over North America at 12km resolution. Developed by the National Centers for Environmental Prediction, it ingests observational data assimilated into initial conditions. With high resolution and frequent updates, the NAM better resolves mesoscale features and incorporates latest observations for improved short-term prediction.

The authors utilized the NAM weather model to generate forecasts of GHI and DNI[2]. They analyzed the NAM model forecasts using FANOVA for a specific weather station location. According to the study, the NAM model's forecast inaccuracy is influenced by the interplay of several physical elements, including solar zenith angle and GHI. Machine learning techniques were used to combine and correct the NAM model's biases differently in each weather category or subspace. By accounting for localized differences in the NAM model's biases through categorization parameters, the machine learning approach achieved over 30% improvement in forecast accuracy compared to the raw NAM model.

3) European Centre for Medium-Range Weather Forecasts: Global atmospheric variable predictions, both deterministic and ensemble, can be procured from the ECMWF model at different resolutions. It solves prognostic equations on a spherical grid with physical parameterizations. The ensemble system perturbs initial conditions to generate forecast scenarios and quantify uncertainty. ECMWF provides guidance worldwide and its ensemble is a leading tool for probabilistic forecasting from short-range to seasonal timescales.

In one such study, a short-term probabilistic solar power forecast using the ECMWF Ensemble Prediction System (EPS) is presented[3]. The authors test the approach on three solar farms in Italy with distinct climatic conditions. The EPS ensemble of weather forecasts provides uncertainty estimates that feed into a radiative transfer model after bias correction with a neural network. These are compared to a persistence ensemble baseline. Results across the test cases show the calibrated EPS ensembles provide skillful probabilistic forecasts, outperforming persistence.

B. Persistence Forecasts

Persistence forecasts exploit the autocorrelation in time series data by assuming a meteorological variable will remain unchanged over a future period. For solar forecasting, simple persistence assumes irradiance persists, while smart persistence assumes clear sky index persists. Physics-informed variants incorporate specific cloud properties, like cloud fraction or albedo, and assume they persist. Persistence models perform well for short-term solar forecasts under an hour but degrade with longer lead times as persistence assumptions breakdown from sun angle shifts and temporal cloud variations.

A study proposed a new approach called '*Physics based Smart Persistence*' for intra-hour solar power forecasting using only GHI measurements[4]. Firstly, it calculates the extraterrestrial radiation and solar zenith angle, which are fundamental parameters determined by astronomical calculations. Secondly, it employs advanced techniques to predict cloud albedo and cloud fraction, which significantly influence the amount of solar radiation reaching the Earth's surface. The cloud albedo is estimated using a sophisticated two-stream approximation model, while the cloud fraction is forecasted through a weighted moving average method over a 5-minute window, assuming that cloud structures exhibit a certain degree of persistence over short time scales. Experiments using 11 years of 1-min GHI data show PSPI outperforms the baseline models for 5-60 min forecasts, especially under cloudy conditions, without needing additional atmospheric data.

The authors of employed tests using physics-based models and data imputation techniques with data from 15 solar measurement stations across India[5]. Nine data imputation methods are evaluated to fill missing data gaps. The Kalman method is found to perform best overall. The imputed data is used to test a solar forecasting model namely Physics-based Smart Persistence model for Intra-hour forecasting (PSPI). PSPI decomposes solar forecasting into computing solar position and predicting cloud properties. Results show PSPI outperforms Smart Persistence for forecasts up to 150 minutes ahead, with increasing errors at longer time horizons. They utilized MAE for analysis of forecast accuracy and hourly performance plots for hour-hour comparisons

This article developed a new solar forecasting methodology called "stochastic persistence"[6]. It models measured solar irradiation as having a deterministic trend and stochastic noise. Data from two stations used: Ajaccio, Corsica (hourly) and Tilos, Greece (15 min). The additive scheme models the difference between clear sky index and measurements, while the multiplicative scheme models the ratio. Stochastic persistence outperforms classical persistence in accuracy. For hourly data, multiplicative works best with low error. For 15 min data, additive is better, likely due to challenges modeling clear sky at higher frequencies. Normalized RMSE evaluated prediction accuracy against ground truth measurements.

A new approach that incorporates correlations between solar radiation and cloud properties was created and assessed to anticipate global, direct, and diffuse irradiance using four new physics-informed persistence models[7]. Using 15-minute data from 1998-2014, they improved performance over benchmarks, especially at lead times beyond 1.25 hours. The 4th level models performed best overall. Model accuracy related to temporal variability of assumed persistent cloud predictors. Contrasting radiation dependencies on cloud fraction versus albedo explained relative model performance. Various statistical performance metrics were used, including skill scores, modified Taylor diagrams, and distribution comparisons.

C. Regressive Techniques

1) Auto-Regressive Integrated Moving Average models: ARIMA, a statistical analysis model which used for time-series based analysis. It leverages historical data on solar power output, meticulously accounting for trends, seasonal variations, and random fluctuations. This model essentially dissects past power generation patterns, adjusts for any long-term trends or seasonal effects, and integrates short-term variations to generate forecasts of future solar power production.

The study developed ARIMA models using 34 years of NASA POWER insolation data to forecast monthly average solar radiation over an area in India[8]. The models were validated with RMSE, MAE, MAPE metrics. Building upon their models, the researchers utilized the marching square algorithm to generate four years' worth of insolation forecasts in the form of visually descriptive contour plots, enabling spatial analysis of solar radiation patterns. Based on the forecast contours, suitable regions were identified for solar projects having maximum and consistent annual average insolation.

This paper investigates the use of ARMA and ARIMA models for predicting solar radiation one, two, and three hours ahead[9]. The proposed model ARIMA(2,2,2) outperforms both ARMA(1,2) and persistence models in all prediction horizons, with the lowest MAPE and highest improvement percentages compared to the persistence model. The paper concludes that ARIMA models are more accurate for multi-period solar radiation prediction than ARMA or persistence models, and suggests incorporating additional parameters like air temperature and sunshine duration in future studies. Evaluation metrics like MAE, MAPE and improvement percentage are used in this experiments.

This article employed models like ARMA, ARIMA, and various NN architectures, the study analyzed their performance in predicting power output from large-scale photovoltaic (PV) plants[10]. Results demonstrated NNs' superiority in accuracy and computational efficiency over statistical models for one-hour ahead prediction without direct access to weather data. However, both NNs and statistical models proved reliable only for short-term forecasts. Metrics such as MAE and RMSE were used to evaluate model performance and provide recommendations for the effective grid integration of photovoltaic facilities.

2) Seasonal Auto-Regressive Moving Average models: The SARIMA, a time series forecasting technique, is a tailored model to handle seasonal variations in data. The approach combines autoregressive, differencing, moving average models and their seasonal variants to model complex patterns. SARIMA parameters include non-seasonal and seasonal orders for AR, differencing, and MA terms. By incorporating past observations and seasonal fluctuations, SARIMA can generate accurate forecasts for time series data with periodic patterns, making it particularly useful for applications like predicting photovoltaic (PV) power generation, where seasonal variations due to weather conditions play a significant role in output fluctuations.

A study, which found the optimal SARIMA model selected for solar radiation forecasting was SARIMA(4,0,18)(0,1,1) based on comparing AIC values[11]. This means a seasonal MA term was included to capture seasonality. The model was fit on 9 years of weekly solar radiation data. It was then used to forecast for 2 years and compared to actual values. The SARIMA model performed well for solar forecasting, with low MAPE values around 7% for the LA data. This indicates it accurately captured the seasonal patterns. The model fit the data better than for wind speed forecasting. The confidence intervals were wider, indicating higher variability/uncertainty in wind predictions.

The paper suggests using a hybrid SARIMA-SVM model to predict a small-scale grid-connected solar plant's electricity output[12]. The model uses a blend of SARIMA method to estimate the linear components and SVM model to find nonlinear patterns. The authors used experimental data from a 20 kW plant to develop and validate the model. Statistical tests showed the hybrid model outperformed SARIMA and SVM alone for hourly power forecasts over several days. The approach does not require forecasted weather data and utilizes accessible MATLAB functions. The hybrid approach exhibits satisfactory precision for forecasting photovoltaic power output over short time horizons.

The paper presents a combined approach for short-term forecasting of solar power output, utilizing both a SARIMA model and an artificial neural network model[13]. The models run in parallel and their outputs are combined using optimized weighting factors. The authors evaluated various SARIMA and ANN model structures using actual PV system data. The hybrid approach reduced forecast error by 10% during periods of high PV power volatility compared to the individual models. Overall, the parallel architecture and periodic recomputation of weighting factors allows the hybrid model to dynamically adapt based on forecast accuracy of the individual models. The method shows good performance for prediction time frame from 15 minutes to 1 day.

3) Exogenous variable models: Exogenous variable models like ARIMAX, SARIMAX, and ARMAX extend traditional time series models by integrating external factors (exogenous variables) into the forecasting process. In these models, alongside the endogenous variables (the variables being forecasted), exogenous variables, such as weather data or economic indicators, are included. By incorporating these external factors, the models capture additional information that influences the behavior of the time series. This integration enables more accurate and robust forecasts, as the models can account for external influences beyond the historical data alone.

The study compared SARIMA and SARIMAX models for PV generation forecasting[14]. SARIMA struggled in variable weather but performed better in stable conditions, while SARIMAX, incorporating solar radiation forecasts, improved accuracy during irregular patterns. SARIMA adapted well in intra-day forecasting, capturing short-term variations effectively. However, SARIMAX's inclusion of exogenous factors significantly enhanced day-ahead predictions overall. The findings highlight the importance of considering external factors like solar radiation for robust PV generation forecasts, particularly in fluctuating weather. While SARIMA showed promise in certain conditions, SARIMAX's use of additional variables proved essential for optimizing accuracy across diverse scenarios.

The paper describes a method based on system identification approaches for predicting solar radiation components using linear ARX and ARMAX models[15]. Hourly meteorological data from 2005-2015 was used, including variables like air temperature, humidity, radiation components, etc. Performance was evaluated using RMSE and FIT metrics. The proposed models demonstrated robust performance in forecasting extraterrestrial normal radiation, infrared horizontal radiation, and extraterrestrial horizontal radiation, exhibiting low root mean square error (RMSE) and high goodness-of-fit (FIT) values. However, the models had very high errors and low FIT for direct normal and diffuse horizontal radiation. The authors conclude linear modeling works better for cyclical/sinusoidal data, which was not the case for direct and diffuse radiation.

D. Deep learning techniques

1) ANN Approach: Artificial Neural Networks are a category of machine learning techniques that draw inspiration from the intricate neural networks present in the human brain. These networks are composed of interconnected processing units, analogous to neurons, systematically arranged into layered structures. Each neuron processes input data by applying a weighted sum and an activation function, producing an output signal that serves as input to subsequent neurons. During training, ANNs tune connection weights to reduce prediction errors, allowing them to learn complex data patterns and relationships. ANNs are versatile tools employed for classification, pattern recognition and regression across various fields like image processing, natural language, and financial forecasting.

This study develops an ANN model to predict hourly solar radiation in Tamil Nadu, India using meteorological data from 6 locations[16]. Different ANN architectures with 3 learning algorithms (Levenberg-Marquardt, Resilient Propagation, Scaled Conjugate Gradient) were compared. The model with Levenberg-Marquardt algorithm and 12 hidden neurons showed the best performance based on statistical error analysis. With the MAPE of 0.48% also shows the high accuracy of the model, with the predictions deviating less than 0.5% from the real values on average.

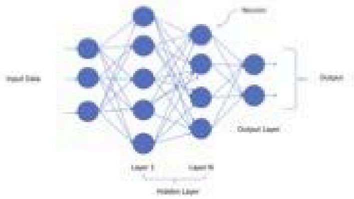


Fig. 1: Artificial Neural Network[36]

Another article presents an artificial neural network model for solar power estimation[17]. A sensitivity analysis identified the most relevant input variables, including solar irradiance and temperature. The ANN model outperformed multiple linear regression and persistence models in two test cases, with lower RMSE and higher correlation between forecasts and actual values. The results show ANNs can effectively forecast solar power using weather variables. The ANN approach acheived a RMSE score of 0.0554 which is highest compared to multi-linear regression and persistent approaches.

This paper forecasts solar radiation in India using feedforward, backpropagation, deep learning on 1 year of Gorakhpur data[18]. The model averaged network achieved the best performance with an average RMSE of 75.36% across 12 months. Backpropagation performed worst at 313.64% RMSE. Nine weather and time inputs were used. Data preprocessing converted minute to hourly averages. The model averaging approach combines multiple separately trained models, boosting performance over individual models.

This article proposes Artificial Neural Networks (ANNs) to forecast solar photovoltaic (PV) power output in Jordan on historical solar irradiance data[19]. The model uses 5 inputs of past irradiance to predict next day PV output. Trained on over 19,000 samples of irradiance and PV power from a 264KWp system, an optimized 22-layer ANN achieved excellent test results with Root Mean Square Error of 0.0721 in 24 hour PV power prediction. This demonstrates machine learning's promise for accurately forecasting solar power based on weather, helping integrate solar PV systems. The model uses real-world irradiance and power data, with preprocessing to ensure matched data pairs.

The article presents a novel solar power forecasting methodology integrating genetic algorithms and artificial neural networks (ANNs), crucial for optimizing solar photovoltaic (PV) systems' performance under varying weather conditions[20]. Utilizing real-time data from Odisha, India, the model exhibited significant advancements over traditional methods. By employing ANN-based temperature forecasting and Genetic Algorithm optimization, it achieved remarkable accuracy in predicting solar PV output. The study highlights the importance of accurate forecasting for grid stability and renewable energy integration. Additionally, statistical analyses such as regression and ANOVA underscored the method's robustness.

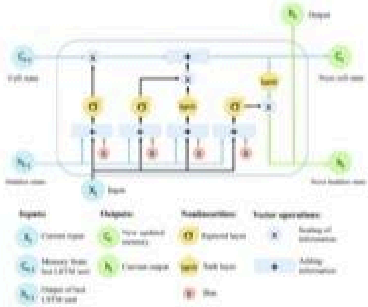


Fig. 2: Long Short Term Memory[37]

*2) **LSTM**:* Long Short Term Memory is a type of recurrent neural network design specifically developed to tackle the issue of the vanishing gradient problem commonly encountered in conventional RNNs. It uses memory cells with self-connections to maintain information over long sequences via selective retaining or forgetting. LSTM cells contain input, forget, and output gates controlled by sigmoid functions that regulate information flow. A tanh function also regulates cell state updates. The inclusion of gating mechanisms in LSTMs allows them to effectively capture long-range dependencies and manage diverse time lags. Consequently, LSTMs are highly suitable for tasks such as speech recognition, language modeling, and predicting time series data.

This article proposes a simplified LSTM approach for one day ahead solar power forecasting[21]. It compares LSTM and MLP models trained on datasets from PV systems in Thailand and Taiwan. The LSTM model with 50 hidden nodes and Tanh activation function performed best, achieving a test RMSE of 0.497 on the Thailand dataset. Data preprocessing included fixing missing values, feature selection using correlation analysis, normalization and dimensionality reduction. The LSTM model successfully captured intra-hour ramping and performed well on different weather conditions like sunny, cloudy and rainy days.

This article proposes a forecasting mechanism using LSTM neural networks with the Nadam optimizer[22]. Tested on a 250kW India plant with 366 days of data, LSTM-Nadam achieved significantly lower RMSE of 86.19 compared to 164.12 for SARIMA and 124.126 for ARIMA. It outperformed LSTM with other optimizers. This demonstrates LSTM-Nadam's efficacy for long-term forecasting, crucial for grid integration and planning of large solar PV systems. Preprocessing involved data cleaning, imputing missing values, and normalization. The model uses meteorological factors like temperature and solar radiation as inputs.

In this article, two LSTM models for short-term solar power forecasting are presented. Time of day, month, weather, and historical power statistics are used as input elements in these models.[23]. Data is from a Florida solar farm. The single-step model predicts the next time step, with performance degrading beyond 12 hour input sequences. It achieved 0.5-0.75 MW MAE across seasons. The multi-step model predicts up to 24 hour horizons. Adding time and month indices improved intraday performance. It achieved 0.58-0.99 MW MAE. The LSTM models outperformed regression and NARX networks and captured weather effects well with half-hourly data across multiple years.

The article proposes an LSTM-fully connected layer for short-term solar power generation forecasting[24]. The model combines LSTM layers to capture time continuity and periodicity and FC layers to learn feature mappings. Experiments on real PV power data show the model with all three input types achieves the best performance. Compared to GRNN, GBDT, SVM, FFNN and LSTM models, the proposed LSTM-FC model optimizes and lowers the RMSE by 11.79%, 7.77%, 9.1%, 11.33%, and 9.16% respectively. The model demonstrates effectively capturing temporal correlations and incorporating multiple relevant data sources can improve photovoltaic power forecasting accuracy.

The article proposes a hybrid LSTM and Gaussian process regression (GPR) model for short-term solar power forecasting[25]. It uses weather data and time features as inputs. The results on two datasets show LSTM outperforms a basic neural network for point forecasting accuracy. The novel LSTM-GPR model achieves slightly better point forecast accuracy than LSTM alone and significantly outperforms GPR alone. The hybrid model also provides reliable prediction intervals, unlike LSTM alone. Overall, the article demonstrates the LSTM-GPR hybrid model achieves high performance on both point and interval forecasting for solar power prediction using standard datasets. The methodology combines the strengths of deep learning via LSTM and probabilistic modeling via GPR.

The study employs LSTM network to forecast hourly solar irradiance in Johannesburg[26]. Utilizing historical meteorological data from Meteoblue, the LSTM model outperforms Support Vector Regression (SVR), achieving a nRMSE of 3.2%, significantly better than SVR's performance with the same dataset. The LSTM architecture, designed to address vanishing and exploding gradient issues, exhibits superior accuracy in predicting solar irradiance, crucial for efficient power generation. The study underscores LSTM's potential for enhancing solar energy forecasting and suggests its application for stable power generation in solar farms.

3) Auto-encoders: Autoencoders are used for dimensionality reduction and unsupervised learning. They are made up of a decoder that reconstructs the original input and an encoder that compresses input data into a latent form. By minimizing reconstruction error, autoencoders learn to capture salient data features in the latent space. As a result, high-dimensional data may be represented compactly, which is advantageous for processes like anomaly detection, denoising, and data compression. Variants like variational and denoising autoencoders offer additional capabilities. Autoencoders have diverse applications including image processing, language understanding, and recommendation systems.

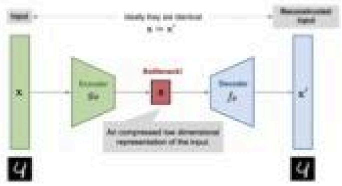


Fig. 3: Auto Encoder[35]

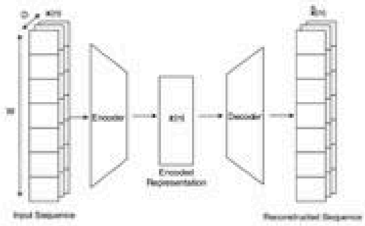


Fig. 4: LSTM Autoencoder[38]

The article proposes a variational autoencoder (VAE) approach for short-term forecasting of photovoltaic (PV) solar power production[27]. The method is evaluated on two real-world PV system datasets. The VAE model consistently outperforms seven other deep learning models (RNN, LSTM, GRU, etc.) and two machine learning baseline models (logistic regression and SVM) in terms of statistical performance metrics like RMSE, MAE, R2. In the PV power time series data, the VAE model may identify both linear and nonlinear characteristics. The research also looks at the models' ability to foresee one or more steps ahead. As a preprocessing step, the PV power data is standardized using min-max normalization between 0 and 1.

The article proposes a LSTM-AE model for photovoltaic (PV) power forecasting[28]. Using a real-world PV system dataset from Australia, the LSTM-AE model is shown to outperform benchmark methods like BPNN, RNN, LSTM and GRU in terms of RMSE, MAE and R2 performance metrics. Specifically, the LSTM-AE model achieved average RMSE of 0.0762 kW, MAE of 0.0389 kW, and R2 of 0.9993 across different test periods. Data preprocessing included missing value imputation, outlier removal, and normalization. The methodology uses encoder-decoder LSTM architecture for feature learning and reconstruction.

This article introduces a novel approach for forecasting day-ahead photovoltaic (PV) power, which merges a generative adversarial network (GAN) with a convolutional autoencoder (CAE)[29]. Real PV plant data and weather data from China are used. The data is first categorized into sunny, cloudy and rainy days using self-organizing map for weather clustering. The CAE-GAN model is then trained for each weather type.

Testing on 73 test days shows the model achieves a mean absolute error of 0.92MW, 16.73% MAPE and 19.87% RMSE across weather types, outperforming LSTM, CNN, and other GAN variants. The CAE enhances feature extraction from inputs while GAN leverages adversarial training for accuracy.

This article proposes a multi-task temporal convolutional autoencoder approach for day-ahead wind and solar power forecasting using numerical weather prediction data[30]. The method combines multi-task learning to leverage data from multiple parks and temporal convolutions to capture time dependencies. Compared to single-task and MLP-based autoencoders, the proposed approach reduces trainable parameters by up to 202 times while improving reconstruction error by up to 300% on 117 PV parks and 445 wind parks. For power forecasting, fine-tuning two encoder layers yields up to 18.3% and 1.5% error reductions for PV and wind, respectively.

4) Hybrid Models: Hybrid solar forecasting models like CNN-LSTM blend distinct deep learning components to exploit their respective advantages. CNNs extract spatial features while LSTMs capture temporal dynamics, handling complex interactions between weather, geography, and other factors affecting solar generation. This enhances accuracy. Other hybrids could integrate RNN variants like GRUs or Bi-LSTMs with CNNs, autoencoders, or sequence-to-sequence models, allowing comprehensive data processing. By capturing diverse information, these hybrid approaches enable robust, reliable solar predictions.

This research suggests a hybrid model, combining CNN and LSTM architectures within an autoencoder framework, for forecasting short-term photovoltaic (PV) power generation[31]. They used CNN for spatial feature extraction and LSTM autoencoders for temporal feature learning. Using real PV farm data from the UK along with weather data, the model is tested for 0.5h, 1h and 2h ahead forecasting. Compared to LSTM, CNN-LSTM, GRU, CNN-GRU models, the proposed hybrid model reduces errors by 5-25% in RMSE and MAE metrics while significantly decreasing training time by 70%. It also outperforms recent methods in literature by 40-80% across time horizons. Preprocessing includes normalization of the PV generation data.

The authors proposed LSTM, Bi-LSTM, and GRU for very short-term solar energy forecasting using weather data from Amherst station[32]. It compares machine learning regressors like Ridge, LASSO, ElasticNet, and SVR as baselines, with SVR giving best results. The deep learning models outperform these with Bi-LSTM giving lowest error metrics - RMSE of 0.0356, MSE of 0.0012, MAE of 0.0124 and MAPE of 12.2%. The models are robust as evident from consistent performance of variants. Hybrid models of Bi-LSTM, LSTM and GRU are also tested for multi-step ahead predictions. LSTMs and GRU prove effective for accurate and reliable solar forecasting based on the weather dataset with minimal preprocessing like feature correlation analysis.

The paper introduces a new deep learning model named Bi-LSTM-based deep stacked sequence-to-sequence autoencoder (S2SAE) for forecasting solar irradiation and wind speed[33]. Real-world datasets from NREL were used for training and testing. Data preprocessing included handling missing values and normalization. The suggested model combines the dimensionality reduction of stacked autoencoders with the capabilities of bidirectional long short-term memory (Bi-LSTM) networks. It was compared to LSTM, GRU, and shallow S2SAE models. The Bi-LSTM S2SAE model achieved the highest accuracy of 99.7% for solar forecasting and 98.42% for wind speed, the lowest MAPE of 0.2763% and 1.58%, and lowest RMSE of 0.0358, outperforming the other methods.

This literature proposes a novel GRU-CNN for short-term solar power forecasting[34]. The model combines a GRU to extract temporal features and a CNN to extract spatial features from the PV dataset. The dataset covers 5 years with a 5-minute resolution from the Desert Knowledge Australia Solar Centre. Results show the GRU-CNN model achieves superior performance over individual GRU, LSTM, and CNN models, with average error metrics of 0.0813 for MAE, 0.0194 for MSE and 0.1359 for RMSE. The key novelty is fusing temporal and spatial features through the integrated GRU-CNN architecture.

Lastly, the authors developed an hour-ahead photovoltaic (PV) power forecasting method using RNN-LSTM model[35]. It uses actual field data from three PV plants over 2016-2019, including weather data like temperature, solar irradiation and wind speed. Data preprocessing techniques like normalization are used. The proposed RNN-LSTM model is compared to regression, machine learning and hybrid models. It achieves superior test performance with the lowest RMSE of 26.85 W/m2, 19.78 W/m2, 39.2 W/m2 and highest r of 0.998 for monocrystalline, polycrystalline film PV plants respectively, demonstrating robustness across PV technologies. Different LSTM architectures are also analyzed, with the single-layered RNN-LSTM performing the best.

IV. Summary of Studies

With thorough review of existing approaches, we found that most physical methods like NAM, ECMWF, WRF, perform best when combined with machine learning deep learning approaches. Most of the studies carried out had the time-ahead approaches. Most of the studies carried out had the time-ahead duration of 24-hours, with some experiments in long-term forecasting. Commonly used evalutation metrics include RMSE(Root Mean Square Error), MAE(Mean Absolute Error), MAD(Mean Absolute Deviation), MAPE(Mean Absolute Percentage Error), r2(Coefficient of determination). By observations, we found that Auto-regressive methods like ARIMA SARIMA exceeded the performance of the physical methods. By capturing analysis of the historical data, they were accurately able to generate power generation forecasts.

A. Key parameters

The task of the power prediction is dependent on the region of study, therefore data pertained to a region such as the weather parameters like solar irradiance, temperature, cloud

TABLE I: Summary of Investigated Studies

Ref	Year	Methodology	Output	Metrics
[1]	2016	Augmented WRF + aerosol-cloud-radiation feedbacks	GHI, DIF, DNI	RMSE (Improvements) GHI: 46% DNI: 60% DIF: 70%
[2]	2016	Blending of NAM and ML models	GHI, DNI	MAE: 119 W/m2 MAPE: 11.93% RMSE: 173 W/m2 NMAE%: 14.3%
[3]	2016	ECMWF EPS forecasts inputted to NN, Post processing- VD, EMOS Persistence Ensemble (PE)	Power output(kW)	MRE (PE): (8.91%, 7.65%, 4.99%) (VD): (8.02% , 8.92% , 10.72%) EMOS: (9.41% , 7.60% , 7.47%)
[4]	2018	Physics-based Smart persistence model for Intra-hour Solar forecasting (PSPI)	GHI	MBE = 9.03 W/m$^2 FS = 0.05 MAE = 105.8 W/m^2 r = 0.78, RMSD = 189.1 W/m^2$
[5]	2021	PSPI + Kalman Data Imputation	GHI	For 15 min Ahead MAE: 65.92 for 150 min Ahead MAE : 130.1
[6]	2018	Stochastic persistence	GHI	Ajaccio: Additive nRMSE: 0.3873 Multiplicative nRMSE: 0.3844 Tilos: Additive nRMSE: 0.2902 Multiplicative nRMSE: 0.3194
[7]	2020	4 Level Hierarchical - PSPI	GHI, DNI, DIF	Percent Error (PE) for GHI: 7% at 1 hour ahead
[8]	2020	ARIMA	Avg Solar Insolation	R2 = 0.9293 RMSE = 0.3529 MAE = 0.2659 MAPE = 6.556
[9]	2015	ARMA & ARIMA (Persistence forecast as baseline)	Solar Radiation	Imrovement percentage 1- hour ahead: MAPE : 18.11% ARMA & 7.87% for ARIMA 3-hour ahead MAPE : 32.07% for ARMA and 71.67% for ARIMA
[10]	2019	ARIMA & Bi-LSTM	Power output(kW)	ARIMA: R=0.912 RMSE=1.318 (Poor) Bi-LSTM: R=0.98 RMSE=0.791(Best)
[11]	2020	SARIMA + Akaike Information crtieria for model selection	Solar radiation	MAPE = 7% for Los Angeles and 15.6% for Chicago
[12]	2019	Hybrid Model : SARIMA + ANN	Solar power output	Hybrid model - 96 step ahead: 233.3 W (10% improvement over SARIMA) Sarima : 259.0 W
[13]	2013	SARIMA + SVM	Power Output(kW)	Hybrid: NRMSE 9.4%, NMBE 0.18%, MPE 2.74%, R 0.99 &SARIMA: NRMSE 9.57%, NMBE -0.36%, MPE 5.2%, R 0.99
[14]	2016	SARIMAX - exogenous from solar radiation forecasts	Power Output(kW)	Day-Ahead NRMSE of 10.93%. Intra-day NRMSE of 9.11%.
[15]	2022	ARX and ARMAX	DNI,DHI, ETR,ETH,IHR	DNI (ARMAX) RMSE = 59.43 FIT = 51.22% EHR (ARX): RMSE = 7.15% FIT = 91.34%
[16]	2021	ANN + Levenberg-Marquardt Algo	Solar Radiation	MAD: 0.0456 MSE: 0.00578 RMSE: 0.0763 MAPE: 0.4846
[17]	2015	ANN	Solar Power output	RMSE = 0.08-0.1 and R = 0.94 - 0.96
[18]	2018	ANN, Back Propogation, Model Averaged ANN	Global Solar radiation	RMSE 75.36%
[19]	2017	ANN	Solar Power output	RMSE = 0.0721 and R values of 0.9983 and 0.9965
[20]	2020	ANN + Genetic Algorithm	Solar power output(kW)	RMSE = 1.11952 to 2.282671 ((12 Months) MAE = 0.428703704 to 1.829259259 (12Months) MAPE = 1% to 5% (12 Months)
[21]	2021	LSTM - 200 nodes + tanh activation	Solar Power output	RMSE = 0.497
[22]	2022	LSTM neural network with Nadam optimizer	Solar Power output	RMSE = 86.19 MAE = 69.976
[23]	2020	LSTM	Solar Power output (kW)	MAE = 1.14, RMSE = 1.92, MBE = -0.09
[24]	2021	LSTM- Fully Connected	Solar Power output	$R^2 = 0.96, nRMSE = 0.1743, RMSE =$ 2.5605.
[25]	2021	LSTM and Gaussian Process Regression (GPR)	Solar Power output	RMSE = 264.98, MAE= 201.77, MAPE =9.43%
[26]	2020	LSTM	Solar Irradiance	nRMSE 4-month data =0.04 (1-year data)=0.036 (10 year data)= 0.032
[27]	2020	VAE	Solar Power output	R2= 0.995 RMSE= 199.645 MAE = 99.838 EV= 0.995
[28]	2023	LSTM-AE	Solar Power output	MAE: 0.0389 kW MSE: 0.0064 RMSE: 0.0762 kW R2: 0.9993
[29]	2023	GAN + Convolutional Autoencoder (CAE)	Solar Power output	MAE = 0.9215, avg MAPE = 16.73% and avg RMSE = 19.87%.
[30]	2022	Temporal Convolution based multi-task autoencoders	Solar Power output	nRMSE = 0.098(18.3% reduction from Baseline)
[31]	2023	CNN- LSTM	Solar Power output	RMSE= 0.081 MAE=0.038
[32]	2023	LSTM, Bi-LSTM, and GRU	Solar Power output	Bi-LSTM model: MSE: 0.0012 MAE: 0.013 RMSE: 0.0359 MAPE: 12.050% GRU model: MSE: 0.0012 MAE: 0.0138 RMSE: 0.0354 MAPE: 12.564%
[33]	2023	Bi-LSTM-based deep stacked Seq2Seq Auto Encoder	Solar Power output	MAPE: 0.2763 RMSE: 0.99 R2: 3.03
[34]	2022	RNN-LSTM	Solar Power output	(MAE, RMSE for 3 plants) (19.43, 32.72) (17.13, 25.4) (49.54, 63.93)
[35]	2021	GRU-CNN	Solar Power output	MAE: 0.0813 MSE: 0.0194 RMSE: 0.1359 R2: 0.9989

cover, wind speed, humidity and are all considered in these studies. Seasonal and diurnal patterns allows for the identification of trends and patterns that can be used to train forecasting models. Most of the forecasted variables include DIF(Diffuse Horizontal radiance), DNI(Direct Normal Irradiation), GHI(Global Horizontal Irradiance)[29].

B. *Best performing approaches*

Hybrid models like CNN-LSTM, Bi-LSTM, and GRU-CNN outperform traditional methods in solar power forecasting due to their ability to capture both spatial and temporal dependencies. CNN-LSTM model trained on historical data, acheived and RMSE score of 0.081 in 1-h forecast duration[31]. The Bi-LSTM based approach also achieved higher RMSE score of 0.99 among other architectures[33]. The GRU-CNN model outperformed vanilla LSTM with a high score of RMSE 0.1359.

C. *Challenges*

Solar power forecasting faces challenges including weather uncertainty, data limitations, modeling diffuse/direct irradiance, optimizing model structure and parameters, capturing short-term variability, effectively handling multiple timescales, integrating real-time plant data, accounting for spatial variations, and managing computational complexity.

D. *Future Works*

As the adoption of solar energy continues to grow rapidly, the demand for more accurate and reliable forecasting will intensify. Moreover, as the solar energy industry expands globally, region-specific forecasting models that account for local weather patterns and environmental factors will become increasingly important. Ultimately, continued innovation in solar forecasting will not only enhance the efficiency and reliability of solar power systems but also accelerate the transition towards a sustainable energy future.

V. Conclusions

In summary, this study underscores the critical importance of diverse solar forecasting methodologies, spanning short-term to long-term predictions, and underscores the necessity of harnessing sophisticated models and data-driven strategies to enhance forecast precision and dependability. By employing numerical weather prediction models like the Numerical Weather Prediction models, in conjunction with innovative approaches such as smart persistence and physics-based models, substantial progress has been achieved in forecasting solar irradiance and power generation. These methodologies furnish invaluable insights for energy planning, grid management, and operational decision-making, thereby facilitating the seamless integration of solar energy into the broader energy framework. With solar power assuming an increasingly pivotal role in the transition toward sustainable energy systems, sustained research and development efforts in solar forecasting techniques stand poised to maximize its efficacy and ensure the realization of a dependable and resilient renewable energy landscape.

References

[1] Jimenez, P. A., and Coauthors, 2016: WRF-Solar: Description and Clear-Sky Assessment of an Augmented NWP Model for Solar Power Prediction. Bull. Amer. Meteor. Soc., 97, 1249–1264, https://doi.org/10.1175/BAMS-D-14-00279.1.

[2] X. Shao, S. Lu and H. F. Hamann, "Solar radiation forecast with machine learning," 2016 23rd International Workshop on Active-Matrix Flatpanel Displays and Devices (AM-FPD), Kyoto, Japan, 2016, pp. 19-22, https://doi: 10.1109/AM-FPD.2016.7543604.

[3] Simone Sperati, Stefano Alessandrini, Luca Delle Monache, An application of the ECMWF Ensemble Prediction System for short- term solar power forecasting, Solar Energy, Volume 133, 2016, 437-450, https://doi.org/10.1016/j.solener.2016.04.016.

[4] Kumler, Andrew. (2019). A Physics-based Smart Persistence model for Intra-hour forecasting of solar radiation (PSPI) using GHI measurements and a cloud retrieval technique. Solar Energy. 177. 494-500. 10.1016/j.solener.2018.11.046.

[5] Yelchuri, Srinath, Rangaraj, A. G., Xie, Yu, Habte, Aron, Joshi, Mohit Chandra, Boopathi, K., Sengupta, Manajit, and Balaraman, K.. 2021. "A Short-Term Solar Forecasting Platform Using a Physics-Based Smart Persistence Model and Data Imputation Method". United States. https://doi.org/10.2172/1837967. https://www.osti.gov/servlets/purl/1837967.

[6] Cyril Voyant, Gilles Notton, Solar irradiation nowcasting by stochastic persistence: A new parsimonious, simple and efficient forecasting tool, Renewable and Sustainable Energy Reviews, Volume 92, 2018, Pages 343-352, ISSN 1364-0321, https://doi.org/10.1016/j.rser.2018.04.116.

[7] Weijia Liu, Yangang Liu, Xin Zhou, Yu Xie, Yongxiang Han, Shinjae Yoo, Manajit Sengupta, Use of physics to improve solar forecast: Physics-informed persistence models for simultaneously forecasting GHI, DNI, and DHI, Solar Energy, Volume 215, 2021, Pages 252-265, https://doi.org/10.1016/j.solener.2020.12.045.

[8] Shadab, A., Ahmad, S., Said, S. (2020). Spatial forecasting of solar radiation using ARIMA model. Remote Sensing Applications: Society and Environment, 20. https://doi.org/10.1016/j.rsase.2020.100427

[9] Colak, I., Yesilbudak, M., Genc, N., Bayindir, R. (2016). Multi-period prediction of solar radiation using ARMA and ARIMA models. Proceedings - 2015 IEEE 14th International Conference on Machine Learning and Applications, ICMLA 2015. https://doi.org/10.1109/ICMLA.2015.33

[10] Sharadga, H., Hajimirza, S., Balog, R. S. (2020). Time series forecasting of solar power generation for large-scale photovoltaic plants. Renewable Energy, 150. https://doi.org/10.1016/j.renene.2019.12.131

[11] Maroua Haddad, Jean Nicod, Yacouba Boubacar Mainassara, Landy Rabehasaina, Zeina Al Masry, et al.. Wind and Solar Forecasting for Renewable Energy System using SARIMA-based Model. International conference on Time Series and Forecasting, Sep 2019, Gran Canaria, Spain.

[12] Vrettos, E., Gehbauer, C. (2019). A hybrid approach for short-term PV power forecasting in predictive control applications. 2019 IEEE Milan PowerTech, PowerTech 2019. https://doi.org/10.1109/PTC.2019.8810672

[13] Bouzerdoum, M., Mellit, A., Massi Pavan, A. (2013). A hybrid model (SARIMA-SVM) for short-term power forecasting of a small-scale grid-connected photovoltaic plant. Solar Energy, 98(PC). https://doi.org/10.1016/j.solener.2013.10.002

[14] Vagropoulos, S. I., Chouliaras, G. I., Kardakos, E. G., Simoglou, C. K., Bakirtzis, A. G. (2016). Comparison of SARIMAX, SARIMA, modified SARIMA and ANN-based models for short-term PV generation forecasting. 2016 IEEE International Energy Conference, ENERGYCON 2016. https://doi.org/10.1109/ENERGYCON.2016.7514029

[15] Silva, V. L. G. da, Oliveira Filho, D., Carlo, J. C., Vaz, P. N. (2022). An Approach to Solar Radiation Prediction Using ARX and ARMAX Models. Frontiers in Energy Research, 10. https://doi.org/10.3389/fenrg.2022.822555

[16] Geetha, A., Santhakumar, J., Sundaram, K. M., Usha, S., Thentral, T. M. T., Boopathi, C. S., Ramya, R., Sathyamurthy, R. (2022). Prediction of hourly solar radiation in Tamil Nadu using ANN model with different learning algorithms. Energy Reports, 8. https://doi.org/10.1016/j.egyr.2021.11.190

[17] Abuella, M., Chowdhury, B. (2015). Solar power forecasting using artificial neural networks. 2015 North American Power Symposium, NAPS 2015. https://doi.org/10.1109/NAPS.2015.7335176

b. Plagiarism Report of Paper I

A Survey of Photovoltaic power forecasting approches in Solar energy Landscape

ORIGINALITY REPORT

7%	8%	7%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

c. Paper II details

'SolarSense'- A Smart Solar System Approach

Anishkumar Mahalingam
Department of Computer Engineering
VESIT
Mumbai, India
2020.anishkumar.mahalingam@ves.ac.in

Swapnil Thatte
Department of Computer Engineering
VESIT
Mumbai, India
2020.swapnil.thatte@ves.ac.in

Yash Narkhede
Department of Computer Engineering
VESIT
Mumbai, India
2020.yash.narkhede@ves.ac.in

Yash Brid
Department of Computer Engineering
VESIT
Mumbai, India
d2020.yash.brid@ves.ac.in

Dr. Dashrath Mane
Deptartment of Computer Engineering
VESIT
Mumbai, India
dashrath.mane@ves.ac.in

***Abstract*—With the rapid expansion of solar energy systems worldwide, the maintenance of solar panels has emerged as a critical aspect to ensure optimal performance and longevity of photovoltaic installations. Traditional maintenance methods often rely on periodic inspections and manual assessments, which can be time-consuming, labor-intensive, and prone to human error. In response to these challenges, this research proposes a novel approach for predictive maintenance of solar panels by leveraging deep learning techniques and computer vision algorithms.**

By employing convolutional neural networks (CNNs) and advanced image processing algorithms, the system aims to classify the condition of solar panels into various categories such as dusty, damaged, clean, affected by bird droppings, and others. The proposed framework utilizes a large dataset of annotated images to train and validate the deep learning models, enabling them to accurately identify and classify different types of panel conditions

***Index Terms*—Solar forecasting, Time-series analysis, Rooftop analysis**

I. Introduction

THIS The proliferation of solar energy systems has ushered in a new era of sustainable power generation, offering a promising solution to the world's growing energy demands while mitigating environmental impact. However, ensuring the optimal performance and longevity of solar panel installations remains a significant challenge. Maintenance of solar panels is critical to maximizing energy output and safeguarding investments in renewable energy infrastructure. Traditional maintenance approaches often rely on periodic inspections and manual assessments, which are time-consuming, resource-intensive, and susceptible to oversights. Consequently, there is a pressing need for innovative solutions that can streamline maintenance processes, enhance operational efficiency, and minimize downtime.

In response to these challenges, this research proposes a novel approach to predictive maintenance for solar panels, leveraging the synergistic capabilities of deep learning techniques and computer vision algorithms. The proposed framework aims to automate the analysis of images captured by drones or cameras installed within solar farms, enabling real-time assessment of panel conditions. By harnessing the power of convolutional neural networks (CNNs) and advanced image processing algorithms, the system categorizes the condition of solar panels into distinct classes such as dusty, damaged, clean, and affected by bird droppings. This holistic approach to maintenance not only improves the reliability and performance of solar energy systems but also reduces operational costs and enhances environmental sustainability.

II. Related Work

In another study focusing on hourly solar irradiance forecasting in Johannesburg, Long Short-Term Memory (LSTM) networks are employed using ten years of historical meteorological data from Meteoblue. The LSTM model surpasses Support Vector Regression (SVR), achieving a normalized Root Mean Square Error (nRMSE) of 3.2%, which is significantly better than SVR's performance with the same dataset. LSTM's architecture, designed to mitigate vanishing and exploding gradient issues, demonstrates superior accuracy in predicting solar irradiance, which is critical for efficient power generation. This research underscores the potential of LSTM in improving solar energy forecasting and suggests its deployment for ensuring stable power generation in solar farms.

The article introduces a novel approach combining LSTM and Gaussian process regression (GPR) for short-term solar power forecasting, using weather data and time features as inputs. Results from two datasets demonstrate that LSTM surpasses a basic neural network in point forecasting accuracy. Furthermore, the hybrid LSTM-GPR model achieves slightly better point forecast accuracy than LSTM alone and significantly outperforms GPR alone. Notably, the hybrid model also offers reliable prediction intervals, a capability lacking in LSTM alone. Overall, the study illustrates that the LSTM-GPR hybrid model achieves state-of-the-art performance in both point and interval forecasting for solar power prediction using standard datasets, leveraging the strengths of deep learning through LSTM and probabilistic modeling via GPR.

The article introduces a novel day-ahead photovoltaic (PV) power forecasting approach that combines a generative adversarial network (GAN) with a convolutional autoencoder (CAE). Real PV plant data and weather data from China are utilized in this method. Initially, the data is categorized into different weather types (sunny, cloudy, rainy) using a self-organizing map for weather clustering. Subsequently, the CAE-GAN model is trained separately for each weather type. Testing on 73 days demonstrates that the model achieves a mean absolute error of 0.92MW, a mean absolute percentage error (MAPE) of 16.73%, and a root mean square error (RMSE) of 19.87% across different weather types, surpassing the performance of LSTM, CNN, and other GAN variants. The CAE contributes by improving feature extraction from the inputs, while the GAN utilizes adversarial training to enhance accuracy.

This study introduces a novel hybrid CNN-LSTM autoencoder model designed for short-term forecasting of photovoltaic (PV) power generation. The model integrates convolutional neural networks (CNN) to extract spatial features and long short-term memory (LSTM) autoencoders for learning temporal features. Using real PV farm data from the UK along with weather information, the model is evaluated for forecasting horizons of 0.5 hours, 1 hour, and 2 hours ahead. Compared to LSTM, GRU, CNN-LSTM, and CNN-GRU models, the proposed hybrid model demonstrates reductions in errors by 5-25% based on RMSE and MAE metrics, while notably reducing training time by 70%. Additionally, it surpasses recent methods in the literature by 40-80% across different time horizons. Data preprocessing includes normalizing the PV generation data.

III. Methodology

A. Solar panel fault detection

Our approach to solar panel fault detection involves a comprehensive investigation leveraging machine learning techniques. Recognizing the criticality of maintaining optimal solar panel performance, we aim to develop an effective procedure for identifying common faults such as dust accumulation, snow cover, bird droppings, and physical or electrical damage on solar panel surfaces. To achieve this objective, we employ a model vision transformer, a state-of-the-art deep learning architecture renowned for its ability to handle image data effectively.

The methodology encompasses several key steps. Firstly, we gather a diverse dataset comprising images categorizing various fault types, including clean panels, dusty panels, bird droppings, electrical damage, physical damage, and snow-covered panels. This dataset serves as the foundation for training and evaluating our fault detection model.

Next, we preprocess the dataset to ensure uniformity and prepare it for model training. This involves tasks such as data cleaning, augmentation, and normalization to enhance the model's robustness and generalization capabilities.

Subsequently, we employ the model vision transformer architecture to train a deep learning model capable of accurately detecting and classifying different fault types on solar panel surfaces. The vision transformer model is chosen for its superior performance in handling image data and its ability to capture intricate patterns and features essential for fault detection.

During the training phase, the model learns to distinguish between various fault categories by analyzing the visual characteristics of the input images. Through iterative optimization and validation processes, the model refines its parameters to maximize accuracy and minimize errors in fault detection.

Upon model convergence, we evaluate its performance using a separate validation dataset to assess its ability to generalize to unseen data. This evaluation involves metrics such as precision, recall, and F1-score to quantify the model's effectiveness in accurately identifying different fault types.

In conclusion, our methodology encompasses the utilization of a model vision transformer architecture for solar panel fault detection, aiming to enhance maintenance efficiency and optimize solar panel performance. Through rigorous experimentation and evaluation, we aim to develop a robust and reliable fault detection system capable of supporting sustainable and efficient solar energy utilization.

B. Anomaly Detection in photovoltaic (PV) systems

The Isolation Forest algorithm consists of the following steps:

Tree Construction: Randomly select a feature and split the data along that feature's range. Isolation: Continue partitioning the data recursively until isolation is achieved. Anomaly Score Calculation: Compute the anomaly score for each data point based on its average path length in the trees. Anomaly Detection: Identify anomalies as data points with lower anomaly scores. The anomaly score s(x) for a data point x is calculated as:

Isolation Forest Algorithm:

$$s(x) = 2^{-\frac{E(h(x))}{c(n)}}$$

Where: - $s(x)$ is the anomaly score for data point x. - $E(h(x))$ is the average path length of data point x in the ensemble of trees. - $c(n)$ is a normalization factor calculated as $2 \cdot H(n-1) - \frac{2(n-1)}{n}$, where $H(I)$ is the harmonic number.

Anomaly Classification using Random Forest Model

After detecting anomalies using Isolation Forest, the subsequent phase involves classifying the type of anomaly utilizing a Random Forest classifier.

3.1. Feature Selection

For effective anomaly classification, it is imperative to select relevant features that contribute significantly to the classification task. This is typically accomplished through various techniques, including:

Recursive Feature Elimination (RFE): RFE is a feature selection method that iteratively removes features from the dataset based on their importance to the model's performance. The Random Forest classifier is trained multiple times, each time eliminating the least important features until the optimal subset of features is determined.

Feature Importance from the Isolation Forest model: The Isolation Forest algorithm used for anomaly detection can provide insights into the importance of each feature in distinguishing between normal and anomalous instances. Features with higher importance scores from the Isolation Forest model are likely to be more informative for anomaly classification and are thus prioritized for inclusion in the Random Forest classifier.

SHAP (SHapley Additive exPlanations): SHAP is a method used to explain the output of machine learning models by quantifying the contribution of each feature to the model's predictions. By analyzing SHAP values, we can identify the most influential features in the Random Forest model and use them for anomaly classification. In this study, SHAP analysis reveals the following features as the most important for anomaly classification: POA, POA_CUM, MODULE_TEMP, EAE_DAY_PLANT, and EAE_DAY.

Random Forest Model Training

Once the relevant features are selected, the Random Forest classifier is trained on the preprocessed dataset to classify the detected anomalies. The training process involves the following steps:

Randomly selecting subsets of features and data points: At each node of the decision trees in the Random Forest, a random subset of features is considered for splitting, and a random subset of data points is used to train each tree. This randomness helps to reduce overfitting and improve the generalization ability of the model.

Constructing decision trees based on these subsets: Multiple decision trees are constructed using the selected subsets of features and data points. Each decision tree in the Random Forest is trained independently, utilizing a bootstrapped sample of the original dataset.

Aggregating predictions from individual trees to make final predictions: Once all decision trees are trained, predictions are made for each anomaly instance by aggregating the predictions from individual trees. In a classification task, the final prediction is typically determined by a majority voting mechanism, where the class with the most votes among all decision trees is assigned to the anomaly instance.

By incorporating SHAP analysis and Random Forest classification, we can effectively identify and classify anomalies in the dataset, providing valuable insights for anomaly detection and mitigation strategies.

C. *Rooftop Analysis*

The rooftop analysis process begins by obtaining a geographical location as input. This location may be specified using latitude and longitude coordinates or an address. Once the location is acquired, the subsequent step involves drawing a polygon on a map to delineate the area of interest within that location. This allows for precise selection of a specific region or boundary within the given geographical area.

Following the drawing of the polygon, the total area enclosed by the polygon is calculated. This calculation provides insights into the size or extent of the selected area within the geographical location. Utilizing this calculated area, the next step involves assessing the power generation potential of the selected area. This assessment incorporates factors such as solar exposure, shading, and solar panel efficiency to estimate the amount of solar power that can be generated within the chosen region.

By following this methodology, a comprehensive analysis of the rooftop's solar energy generation potential can be conducted. This analysis lays the foundation for the development and implementation of an efficient smart solar system, crucial for sustainable energy management and utilization.

D. *Solar Power Generation Prediction*

The analysis of solar power generation commences with the Solar Power Dataset, a crucial starting point containing historical or real-time data on solar irradiance, temperature, time of day, and power output. This dataset serves as the cornerstone for subsequent phases of analysis.

The initial phase, Data Processing, involves several critical sub-processes. Data Cleaning addresses missing data, outliers, and inconsistencies, ensuring the dataset's reliability. Data Visualization techniques are then applied to gain insights into data patterns, trends, and relationships, utilizing graphical tools such as graphs and charts. Feature Engineering follows, where new features are created or existing ones modified to enhance the predictive power of the model. This includes Exploratory Data Analysis (EDA) for deeper insights and Feature Scaling to standardize data, optimizing machine learning model performance.

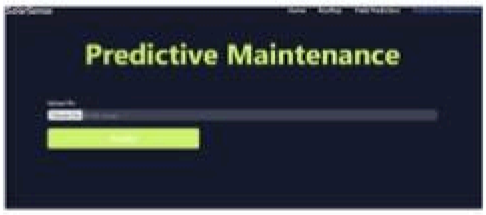
Transitioning to the Model Training Phase, machine learning models are trained using the pre-processed and engineered dataset. Notably, special emphasis is placed on capturing Weather Patterns and Power Yield Patterns, crucial for accurate power generation predictions. Time Series Analysis techniques are employed to understand and model temporal dependencies within the solar power generation data.

In the final phase, Predictions and Results are derived. Leveraging the trained models and current weather data, Power Forecast Predictions are generated, facilitating effective planning and optimization of solar energy utilization. This comprehensive methodology provides a systematic approach to harnessing solar power efficiently within the smart solar system framework.

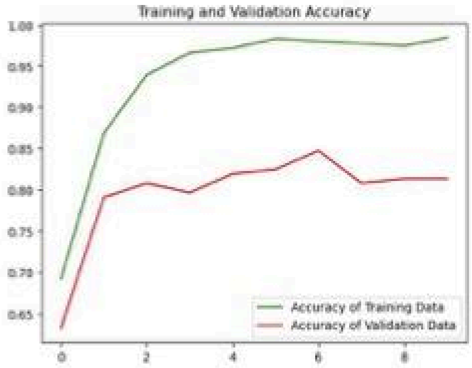
IV. Results and Implementation

This is the Rooftop analysis module where users can get the approximate total power generation potential of the solar system and total system cost. Users first select the area that they want to install solar panels at and then input the features of the solar panel like solar panel efficiency and cost per watt. The analysis of solar power generation commences with the Solar Power Dataset, a crucial starting point containing historical or real-time data on solar irradiance, temperature, time of day, and power output.

Our approach to solar panel fault detection involves a comprehensive investigation leveraging machine learning techniques. Recognizing the criticality of maintaining optimal solar panel performance, we aim to develop an effective procedure for identifying common faults such as dust accumulation, snow cover, bird droppings, and physical or electrical damage on solar panel surfaces. To achieve this objective, we employ a model vision transformer, a state-of-the-art deep learning architecture renowned for its ability to handle image data effectively.



In the final phase, Predictions and Results are derived. Leveraging the trained models and current weather data, Power Forecast Predictions are generated, facilitating effective planning and optimization of solar energy utilization. This comprehensive methodology provides a systematic approach to harnessing solar power efficiently within the smart solar system framework. Through rigorous experimentation and evaluation, a robust fault detection system is developed to optimize solar panel performance and maintenance efficiency. Similarly, in anomaly detection in PV systems, the Isolation Forest algorithm is utilized to detect anomalies followed by classification using a Random Forest model. By calculating the total area and evaluating factors such as solar exposure and shading, the power generation potential of rooftops can be estimated, laying the foundation for the development of efficient smart solar systems



V. Conclusion

The research conducted on solar panel fault detection, anomaly detection in photovoltaic (PV) systems, rooftop analysis, and solar power generation prediction presents a comprehensive methodology aimed at enhancing the efficiency and sustainability of solar energy utilization.

In solar panel fault detection, a model vision transformer architecture is employed to accurately identify and classify various fault types such as dust accumulation, snow cover, bird droppings, and physical or electrical damage. By leveraging a diverse dataset and employing preprocessing techniques such as data augmentation and normalization, the model is trained to effectively detect faults on solar panel surfaces. Through rigorous experimentation and evaluation, a robust fault detection system is developed to optimize solar panel performance and maintenance efficiency. Similarly, in anomaly detection in PV systems, the Isolation Forest algorithm is utilized to detect anomalies followed by classification using a Random Forest model. Through feature selection techniques such as recursive feature elimination and SHAP analysis, relevant features are identified to improve anomaly classification accuracy. By incorporating these methodologies, anomalies in PV systems can be effectively identified and classified, providing valuable insights for anomaly detection and mitigation strategies. The rooftop analysis methodology involves delineating specific regions of interest within a geographical location to assess the solar energy generation potential. By calculating the total area and evaluating factors such as solar exposure and shading, the power generation potential of rooftops can be estimated, laying the foundation for the development of efficient smart solar systems. Furthermore, the solar power generation prediction methodology employs machine learning models trained on historical or real-time solar irradiance, temperature, and power output data. Through data processing, feature engineering, and time series analysis techniques, accurate power generation predictions are generated, facilitating effective planning and optimization of solar energy utilization. In conclusion, the methodologies presented in this research provide systematic approaches to various aspects of solar energy utilization, including fault detection, anomaly detection, rooftop analysis, and power generation prediction. By leveraging machine learning techniques and comprehensive datasets, these methodologies contribute to enhancing the efficiency, reliability, and sustainability of solar energy systems, ultimately advancing the transition towards renewable energy sources.

ACKNOWLEDGMENT

REFERENCES

[1] Jimenez, P. A., and Coauthors, 2016: WRF-Solar: Description and Clear-Sky Assessment of an Augmented NWP Model for Solar Power Prediction. Bull. Amer. Meteor. Soc., 97, 1249–1264, https://doi.org/10.1175/BAMS-D-14-00279.1.

[2] X. Shao, S. Lu and H. F. Hamann, "Solar radiation forecast with machine learning," 2016 23rd International Workshop on Active-Matrix Flatpanel Displays and Devices (AM-FPD), Kyoto, Japan, 2016, pp. 19-22, https://doi: 10.1109/AM-FPD.2016.7543604.

[3] Simone Sperati, Stefano Alessandrini, Luca Delle Monache, An application of the ECMWF Ensemble Prediction System for short- term solar power forecasting, Solar Energy, Volume 133, 2016, 437-450, https://doi.org/10.1016/j.solener.2016.04.016.

[4] Kumler, Andrew. (2019). A Physics-based Smart Persistence model for Intra-hour forecasting of solar radiation (PSPI) using GHI measurements and a cloud retrieval technique. Solar Energy. 177. 494-500. 10.1016/j.solener.2018.11.046.

[5] Yelchuri, Srinath, Rangaraj, A. G., Xie, Yu, Habte, Aron, Joshi, Mohit Chandra, Boopathi, K., Sengupta, Manajit, and Balaraman, K.. 2021. "A Short-Term Solar Forecasting Platform Using a Physics-Based Smart Persistence Model and Data Imputation Method". United States. https://doi.org/10.2172/1837967. https://www.osti.gov/servlets/purl/1837967.

[6] Cyril Voyant, Gilles Notton, Solar irradiation nowcasting by stochastic persistence: A new parsimonious, simple and efficient forecasting tool, Renewable and Sustainable Energy Reviews, Volume 92, 2018, Pages 343-352, ISSN 1364-0321, https://doi.org/10.1016/j.rser.2018.04.116.

[7] Weijia Liu, Yangang Liu, Xin Zhou, Yu Xie, Yongxiang Han, Shinjae Yoo, Manajit Sengupta, Use of physics to improve solar forecast: Physics-informed persistence models for simultaneously forecasting GHI, DNI, and DHI, Solar Energy, Volume 215, 2021, Pages 252-265, https://doi.org/10.1016/j.solener.2020.12.045.

[8] Shadab, A., Ahmad, S., Said, S. (2020). Spatial forecasting of solar radiation using ARIMA model. Remote Sensing Applications: Society and Environment, 20. https://doi.org/10.1016/j.rsase.2020.100427

[9] Colak, I., Yesilbudak, M., Genc, N., Bayindir, R. (2016). Multi-period prediction of solar radiation using ARMA and ARIMA models. Proceedings - 2015 IEEE 14th International Conference on Machine Learning and Applications, ICMLA 2015. https://doi.org/10.1109/ICMLA.2015.33

[10] Sharadga, H., Hajimirza, S., Balog, R. S. (2020). Time series forecasting of solar power generation for large-scale photovoltaic plants. Renewable Energy, 150. https://doi.org/10.1016/j.renene.2019.12.131

[11] Maroua Haddad, Jean Nicod, Yacouba Boubacar Mainassara, Landy Rabehasaina, Zeina Al Masry, et al.. Wind and Solar Forecasting for Renewable Energy System using SARIMA-based Model. International conference on Time Series and Forecasting, Sep 2019, Gran Canaria, Spain.

[12] Vrettos, E., Gehbauer, C. (2019). A hybrid approach for short-term PV power forecasting in predictive control applications. 2019 IEEE Milan PowerTech, PowerTech 2019. https://doi.org/10.1109/PTC.2019.8810672

[13] Bouzerdoum, M., Mellit, A., Massi Pavan, A. (2013). A hybrid model (SARIMA-SVM) for short-term power forecasting of a small-scale grid-connected photovoltaic plant. Solar Energy, 98(PC). https://doi.org/10.1016/j.solener.2013.10.002

[14] Vagropoulos, S. I., Chouliaras, G. I., Kardakos, E. G., Simoglou, C. K., Bakirtzis, A. G. (2016). Comparison of SARIMAX, SARIMA, modified SARIMA and ANN-based models for short-term PV generation forecasting. 2016 IEEE International Energy Conference, ENERGYCON 2016. https://doi.org/10.1109/ENERGYCON.2016.7514029

[15] Silva, V. L. G. da, Oliveira Filho, D., Carlo, J. C., Vaz, P. N. (2022). An Approach to Solar Radiation Prediction Using ARX and ARMAX Models. Frontiers in Energy Research, 10. https://doi.org/10.3389/fenrg.2022.822555

[16] Geetha, A., Santhakumar, J., Sundaram, K. M., Usha, S., Thentral, T. M. T., Boopathi, C. S., Ramya, R., Sathyamurthy, R. (2022). Prediction of hourly solar radiation in Tamil Nadu using ANN model with different learning algorithms. Energy Reports, 8. https://doi.org/10.1016/j.egyr.2021.11.190

[17] Abuella, M., Chowdhury, B. (2015). Solar power forecasting using artificial neural networks. 2015 North American Power Symposium, NAPS 2015. https://doi.org/10.1109/NAPS.2015.7335176

[18] Srivastava, R., Tiwari, A. N., Giri, V. K. (2018). Forecasting of Solar Radiation in India Using Various ANN Models. 2018 5th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering, UPCON 2018. https://doi.org/10.1109/UPCON.2018.8597170

[19] Alomari, M. H., Adeeb, J., Younis, O. (2018). Solar photovoltaic power forecasting in Jordan using artificial neural networks. International Journal of Electrical and Computer Engineering, 8(1). https://doi.org/10.11591/ijece.v8i1.pp497-504

[20] Pattanaik, D., Mishra, S., Khuntia, G. P., Dash, R., Swain, S. C. (2020). An innovative learning approach for solar power forecasting using genetic algorithm and artificial neural network. Open Engineering, 10(1). https://doi.org/10.1515/eng-2020-0073

[21] Liu, C. H., Gu, J. C., Yang, M. T. (2021). A Simplified LSTM Neural Networks for One Day-Ahead Solar Power Forecasting. IEEE Access, 9. https://doi.org/10.1109/ACCESS.2021.3053638

[22] Sharma, Jatin Soni, Sameer Paliwal, Priyanka Shaik, Saboor Chaurasiya, Prem Sharifpur, Mohsen Khalilpoor, Nima Afzal, Asif. (2022). A novel long term solar photovoltaic power forecasting approach using LSTM with Nadam optimizer: A case study of India. Energy Science Engineering. 10. 10.1002/ese3.1178.

[23] Hossain, M. S., Mahmood, H. (2020). Short-term photovoltaic power forecasting using an LSTM neural network and synthetic weather forecast. IEEE Access, 8. https://doi.org/10.1109/ACCESS.2020.3024901

[24] Li, Y., Ye, F., Liu, Z., Wang, Z., Mao, Y. (2021). A Short-Term Photovoltaic Power Generation Forecast Method Based on LSTM. Mathematical Problems in Engineering, 2021. https://doi.org/10.1155/2021/6613123

[25] Wang, Y., Feng, B., Hua, Q. S., Sun, L. (2021). Short-term solar power forecasting: A combined long short-term memory and gaussian process regression method. Sustainability (Switzerland), 13(7). https://doi.org/10.3390/su13073665

[26] Obiora, C. N., Ali, A., Hasan, A. N. (2020). Forecasting Hourly Solar Irradiance Using Long Short-Term Memory (LSTM) Network. 11th International Renewable Energy Congress, IREC 2020. https://doi.org/10.1109/IREC48820.2020.9310449

[27] Dairi, A., Harrou, F., Sun, Y., Khadraoui, S. (2020). Short-term forecasting of photovoltaic solar power production using variational auto-encoder driven deep learning approach. Applied Sciences (Switzerland), 10(23). https://doi.org/10.3390/app10238400

[28] Sabri, M., el Hassouni, M. (2023). Photovoltaic power forecasting with a long short-term memory autoencoder networks. Soft Computing, 27(15). https://doi.org/10.1007/s00500-023-08497-y

[29] Pan, X., Zhou, J., Sun, X., Cao, Y., Cheng, X., Farahmand, H. (2023). A hybrid method for day-ahead photovoltaic power forecasting based on generative adversarial network combined with convolutional autoencoder. IET Renewable Power Generation, 17(3). https://doi.org/10.1049/rpg2.12619

[30] Schreiber, J., Sick, B. (2022). Multi-Task Autoencoders and Transfer Learning for Day-Ahead Wind and Photovoltaic Power Forecasts. Energies, 15(21). https://doi.org/10.3390/en15218062

[31] Ibrahim, M.S., Gharghory, S.M. Kamal, H.A. A hybrid model of CNN and LSTM autoencoder-based short-term PV power generation forecasting. Electr Eng (2024). https://doi.org/10.1007/s00202-023-02220-

[32] Zameer, A., Jaffar, F., Shahid, F., Muneeb, M., Khan, R., Nasir, R. (2023). Short-term solar energy forecasting: Integrated computational intelligence of LSTMs and GRU. PLoS ONE, 18(10 October).

https://doi.org/10.1371/journal.pone.0285410

[33] Mughees, N., Jaffery, M. H., Mughees, A., Mughees, A., Ejsmont, K. (2023). Bi-LSTM-Based Deep Stacked Sequence-to-Sequence Autoencoder for Forecasting Solar Irradiation and Wind Speed. Computers, Materials and Continua, 75(3). https://doi.org/10.32604/cmc.2023.038564

[34] Akhter, M. N., Mekhilef, S., Mokhlis, H., Almohaimeed, Z. M., Muhammad, M. A., Khairuddin, A. S. M., Akram, R., Hussain, M.
M. (2022). An Hour-Ahead PV Power Forecasting Method Based on an RNN-LSTM Model for Three Different PV Plants. Energies, 15(6). https://doi.org/10.3390/en15062243

[35] Sabri, M., el Hassouni, M. (2022). A Novel Deep Learning Approach for Short Term Photovoltaic Power Forecasting Based on GRU-CNN Model. E3S Web of Conferences, 336. https://doi.org/10.1051/e3sconf/202233600064

[36] Lesort, Timothe´e. (2020). Continual Learning: Tackling Catastrophic Forgetting in Deep Neural Networks with Replay Processes.

d. Project Review Sheet

Project review sheet 1:

Inhouse/ Industry _Innovation/Research:

Sustainable Goal:

Class: D17 A/B/C

Project Evaluation Sheet 2023 - 24

Group No.: 22

Title of Project: Solar-Sense- A Smart Solar System Approach.

Group Members: Anishkumar Iyer (D17C 03), Swapnil Thatte (D17C 58), Yash Narkhede (D17B, 46), Yash Brid (D17B ")

Engineering Concepts & Knowledge	Interpretation of Problem & Analysis	Design / Prototype	Interpretation of Data & Dataset	Modern Tool Usage	Societal Benefit, Safety Consideration	Environment Friendly	Ethics	Team work	Presentation Skills	Applied Engg&Mgmt principles	Life - long learning	Professional Skills	Innovative Approach	Research Paper	Total Marks
(5)	(5)	(5)	(3)	(5)	(2)	(2)	(2)	(2)	(2)	(3)	(3)	(3)	(3)	(5)	(50)
03	03	03	03	03	02	02	02	02	02	02	02	02	03	02	38

Comments: publications due b4re next review

Dr. D. G. Mane

Name & Signature Reviewer1

Engineering Concepts & Knowledge	Interpretation of Problem & Analysis	Design / Prototype	Interpretation of Data & Dataset	Modern Tool Usage	Societal Benefit, Safety Consideration	Environment Friendly	Ethics	Team work	Presentation Skills	Applied Engg&Mgmt principles	Life - long learning	Professional Skills	Innovative Approach	Research Paper	Total Marks
(5)	(5)	(5)	(3)	(5)	(2)	(2)	(2)	(2)	(2)	(3)	(3)	(3)	(3)	(5)	(50)

Comments:

Date: 10th february, 2024

Dr. Rohini Temkar

Name & Signature Reviewer 2

Project review sheet 2:

Project Evaluation Sheet 2023 - 24

22

Title of Project: 'Solar-Sense' - A Smart Solar System Approach.

Group Members: Anishkumar Iyer (D17C 03), Swapnil Thatte (D17C 58), Yash Narkhede (D17B 46), Yash Brid (D17B ")

Engineering Concepts & Knowledge	Interpretation of Problem & Analysis	Design / Prototype	Interpretation of Data & Dataset	Modern Tool Usage	Societal Benefit, Safety Consideration	Environment Friendly	Ethics	Team work	Presentation Skills	Applied Engg&Mgmt principles	Life - long learning	Professional Skills	Innovative Approach	Research Paper	Total Marks
(5)	(5)	(5)	(3)	(5)	(2)	(2)	(2)	(2)	(2)	(3)	(3)	(3)	(3)	(5)	(50)
04	04	04	03	04	02	02	02	02	2	02	02	02	02	04	41

Comments: publication is awaited.

Dr. D. G. Mane

Name & Signature Reviewer1

Inhouse/ Industry _Innovation/Research:

Engineering Concepts & Knowledge	Interpretation of Problem & Analysis	Design / Prototype	Interpretation of Data & Dataset	Modern Tool Usage	Societal Benefit, Safety Consideration	Environment Friendly	Ethics	Team work	Presentation Skills	Applied Engg&Mgmt principles	Life - long learning	Professional Skills	Innovative Approach	Research Paper	Total Marks
(5)	(5)	(5)	(3)	(5)	(2)	(2)	(2)	(2)	(2)	(3)	(3)	(3)	(3)	(5)	(50)
04	04	04	03	04	02	02	02	02	02	03	02	2	03	03	40

Comments:

Date: 9th March, 2024

Name & Signature Reviewer 2

Dr. Rohini Temkar.