International Journal for Multidisciplinary Research (IJFMR)

E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Breakeven Estimation of Solar Energy: A Machine Learning and Time Series Analysis Approach

Argula Sujith[1], Amancha Ashwith[2], Kallem Sai Kiran Reddy[3], Ms. Shruthi Kansal[4]

[1,2,3,4]Department of Computer Science and Engineering, Vardhaman College of Engineering, Hyderabad, India

Abstract:
Utilizing a comprehensive dataset featuring daily information such as sunrise time, sunset time, minimum and maximum temperatures, this study's primary aim is to ascertain the duration necessary for a solar panel installation to become profitable and reach the breakeven point. Users provide input parameters, including the solar panel area to be installed, location, cost per unit of current, and installation charges. Employing time series analysis and forecasting techniques, the system takes into account various environmental factors, energy generation potential, and local energy prices. The ultimate goal is to predict when the cumulative income from the solar panel installation will surpass the initial installation costs, facilitating an estimation of the point at which the solar panel system will break even and commence generating a net profit.

This solar panel profitability analysis holds considerable significance for individuals and organizations contemplating investments in renewable energy. It equips stakeholders with the information required to make informed decisions regarding the feasibility and financial viability of solar panel installations at specific locations. This, in turn, promotes a sustainable and cost-effective transition to renewable energy sources, aligning with the broader goal of reducing our reliance on non-renewable energy and fostering environmentally responsible practices.0

Keywords: Solar radiation, daily energy generation estimation, machine learning, random forest, time series analysis, ARIMA, SARIMA, renewable energy, energy price forecasting

1. 1.Introduction

In today's world, fundamental necessities for human survival include food, clothing, housing, and even power. It is evident from examining. the process of producing power that non-renewable resources are used to produce the vast bulk of our energy. But because of the pressing environmental problems of our day, these non-renewable fuels are running out very quickly. Therefore, reliable and ecologically beneficial alternatives are desperately needed to meet our energy needs [1]. It is generally accepted that using solar energy is a healthy and sustainable form of energy. One of its most obvious features is its widespread availability. The sun is like a massive nuclear fusion reactor, constantly producing large amounts of radiation. This light energy is carried by photons.

The precise evaluation of a solar power plant's energy output is the primary focus of this work. With its vast potential, solar energy is an important area of study that has gained popularity recently. In this context, machine learning emerges as a very helpful technique for creating time series forecasts [2]. This research investigates the modeling and computational strategies employed in the simulation and evaluation of irregular and variable shadows, including those caused by moving clouds, on solar photovoltaic (PV) array energy.

Researchers most likely employ sophisticated mathematical models and computer simulations to accurately predict how these occasional shadows would impact the performance of solar photovoltaic systems. Understanding the dynamic behavior of shadows on solar panels is necessary to maximize the efficiency and energy production of these renewable energy sources.

By using complex algorithms and simulations, this research aims to shed light on how to lessen power variations caused by shadowing occurrences. This knowledge can help with the design and administration of solar PV systems, improving their overall reliability and energy production. [3] discusses the output power of solar PV arrays. This paper proposes a new technique for one-day forecasting photovoltaic (PV) power generation. This model makes use of the unique features of weather classification to enhance its prediction abilities. It is based on the Support Vector Machine (SVM) theory. The integration of machine learning and weather classification criteria offers a promising approach to address the problems associated with renewable energy production and grid stability [4].

Previous research, as noted in [16], investigated machine learning methods in great detail to forecast the output power of photovoltaic (PV) systems. These methods included Support Vector Machines (SVM), Artificial Neural Networks (ANN), and Multiple Linear Regression. Although PV system performance predictions has received interest, energy management has received the majority of the attention. It's interesting to note that insufficient focus has been placed on PV system performance ratio prediction, despite its critical role in guaranteeing energy reliability. In order to close this gap, this study uses SVM and Multiple Linear Regression to predict performance ratios over a range of time horizons, including one day, one week, two weeks, three weeks, and one month. The goal of the research is to improve and maintain PV systems' energy dependability while offering insightful information that can be used to forecast and manage energy more efficiently.

The world urgently needs to switch to sustainable energy sources in order to fight climate change and cut greenhouse gas emissions. Given the circumstances, solar energy has demonstrated potential as an environmentally responsible way to meet the growing need for electricity. By anticipating solar radiation and estimating daily energy generation using weather data and machine learning techniques, this study aims to boost solar energy consumption and optimization.

1.1 MOTIVATION

Apart from being intriguing, it can also have significant practical uses in estimating daily energy production and forecasting solar radiation. The following are some justifications for initiating this kind of research:

Sunlight energy is a clean and renewable energy source. Accurately projecting solar radiation and energy production helps improve the development of environmentally friendly and sustainable energy solutions. With precise estimates of solar radiation, energy output can be maximized for individuals, businesses, and utilities, with potentially large financial gains. It provides cost-effective energy requirements solutions and helps reduce energy costs. As a teaching tool, this research study might be employed. Increasing people's awareness of the benefits of solar energy could prompt them to investigate other sources..

1.2 CONTRIBUTION

A crucial component of managing and planning solar energy is predicting solar radiation and estimating daily energy production from solar panels. There are various significant contributions and factors in this process, including:

- Gathering comprehensive and trustworthy weather data is essential. This includes specifics regarding temperature, humidity, cloud cover, and sun irradiance. With historical data, patterns and trends can be found.
- Gathering accurate and comprehensive weather data is essential. This contains information on temperature, humidity, cloud cover, and sun irradiance. Historical data can be used to identify patterns and trends.
- Real-time weather data integration is essential for accurate daily forecasts. This requires access to weather APIs or other data sources.
- Accounting for the cost of the solar array installation and the revenue generated by electricity generation. Return on investment (ROI) and payback time calculations are necessary.
- To forecast solar radiation and calculate daily energy generation, a combination of data science, engineering, and domain-specific expertise in the field of solar energy is required. For solar energy to be managed efficiently and economically, these forecasts must be accurate.

2. Literature Survey

It is observed that precise projections lead to large savings by preventing needless fuel purchases and emergency electricity purchases, and by promoting energy storage, electricity trading, maintenance, and other activities. In addition to outlining a novel mathematical technique for precise hourly electricity generation estimation from rooftop solar plants, the work highlights the significance of solar energy in industrial settings. In order to satisfy the evolving energy needs in a sustainable manner, it emphasizes the significance of different estimation approaches used to anticipate global solar irradiance. When compared to PV Syst software and actual data, this method shows potential for accurate and extensive integration of solar electricity.[6] uses a range of methods, including the Hammerstein-winner, Transfer function, and non-linear ARX models, to estimate the output power of solar PV systems in comparison to the Kalman filter. By contrasting various approaches, it provides insights into effective methods for estimating solar power.[7] The model computes the mean and standard deviation parameters (95% confidence interval) using a normal distribution based on historical data. By guiding the design of photovoltaic systems and energy policy decisions, this technique supports the assessment of solar energy potential. [8] This study looks into how hard it is to get the most electrical energy out of solar panels, which are usually fixed in place. An array of solar cells with a 50 m2 roof area and a 6 kWp output may produce around 10,006.7 kWh annually. This grid-connected system costs Rp 445,453,328 without batteries and will cost you maintenance costs for 25 years.[9]

Artificial intelligence approaches like as artificial neural networks (ANN) and adaptive neuro-fuzzy inference systems (ANFIS) are used to precisely anticipate ground observations through software models. Based on the study, ANFIS outperforms ANN in terms of accuracy. Using atmospheric data and surface solar radiation values, this study provides significant insights for the design and optimization of solar energy systems.[10] In addition to providing significant new information about the techniques employed for accurate solar radiation estimation, it highlights the improved performance of models based on Artificial Neural Networks (ANNs). The study addresses the main concerns and difficulties associated with Turkey's feed-in tariff system, as well as the financial implications of it. In order to estimate global solar radiation (GSR), solar radiation estimating models are created, taking into account geographical and climatic parameters due to the constraints of current monitoring sites.[11] Turkey's growing solar energy potential was demonstrated in 2011 when the country's solar installations reached 442.13 MW thanks to a government support program. For industrial applications, the study calculates costs, payback times, and scenarios. Using the PVGIS tool and testing against data from 1985 to 2006, it evaluates daily power generation and global irradiation per square meter.[12] batteries to guarantee best possible energy consumption. The main goal of this work is to forecast off-grid solar PV system reliability, which is crucial given concerns about climate change. When there is a scarcity of grid electricity, excess power produced during the day is stored in batteries for use during peak load hours.[13] [A model that might be used to comparable systems with climate-related changes is produced by analysis using mathematical and computational techniques.[14] This study examines the solar and wind energy capacities of Maharashtra, India. The study compares and contrasts the use of solar and wind energy across multiple districts. This work estimates and predicts the solar cell performance ratio using Principal Component Analysis - Support Vector Machine (PCA-SVM) using three years of meteorological data (2015-2018). The findings shed light on how renewable energy sources might alleviate the burden on traditional power plants at the state and district levels.[15] SVM parameters are optimized via grid search, and data trend analysis improves prediction precision. [16] Grid search is used to optimize SVM parameters, and data trend analysis enhances prediction accuracy. [16] This study integrates temperature and sun energy data to examine solar power estimation in Goa. PCA-SVM achieves a dependable performance ratio estimation with an RMSE of 0.11 and an R2 of 0.44, according to the results. With a focus on inverter design using PVsyst and MATLAB/Simulink and estimation at specific sites, this work explores solar power integration at the distribution level. The study demonstrates a direct correlation between solar power output and ambient temperature, with April and November showing the highest results.[17] It addresses challenges in integrating solar electricity with conventional grids by offering insights through loss diagrams for grid-connected situations and supporting the design of active network protection.[18] This study used a Seasonal Autoregressive Integrated Moving Average model to anticipate solar radiation values in different locations across India. By projecting prospective cost savings based on current electricity bills, users might save over 70% on electricity expenditures, providing insights into the financial benefits and breakeven lines for switching to solar energy.[19] The Dirashe Woreda solar photovoltaic power systems in Ethiopia are assessed in this study. Using empirical models and NASA data, it assesses solar radiation and offers information on the region's solar energy potential. The study's contribution of an Angstrom-Prescott Model equation customized for the research location advances our understanding of the solar energy resources in this region.[20] The]. In order to estimate solar power, this study proposes an approach that makes use of algebraic combinators and ensemble learning, analyzing models such as the Nearest Centroid Classifier, Support Vector Machine, and Decision Tree Classifier. When assessing solar output in relation to the construction of new power plants, the recommended methodology—which uses the Weighted Sum Rule—achieves exceptional accuracy (99.915%).[21]

3. Proposed System

By using state-of-the-art approaches and techniques, the proposed plan aims to address the deficiencies of the existing systems and make them better. Our approach is underpinned by our commitment to use state-of-the-art techniques and protocols to enhance the accuracy and efficacy of weather-related forecasts. To do this, we use a vast array of specific weather data, including temperature, humidity, wind speed, and other relevant variables, to train machine learning models. Because it can handle complex interactions and identify nonlinear patterns in the data, the random forest technique is the one we specifically chose. We also employ time series analysis techniques to improve the accuracy even more. the capacity to record regional differences in solar radiation patterns and the resilience in handling missing data and outliers.

All things considered; our recommended method represents a significant advancement in the field of weather prediction systems. By adopting cutting-edge approaches, making use of copious amounts of data, and building a robust and adaptable framework, we are ready to surmount the challenges that have historically made accurate weather prediction challenging. This method not only promises to increase the accuracy of weather forecasts, but it also offers an invaluable tool for a variety of applications that depend on accurate weather data, such as agriculture, energy management, disaster preparation, and more.

3.1 Methodology

This section outlines the methodology we employed in our study, which covers time series analysis, dataset construction, model training and selection, and data preparation.

3.2 Data Preprocessing

Preparing the data is an essential step in assuring the accuracy and dependability of the dataset. We describe the steps involved in data preprocessing in this subtopic, which include:

- We extract relevant columns from the gathered dataset, including those related to temperature, humidity, wind direction, speed, and sun radiation. These columns provide the input features for our models.
- Extra columns and parameters that aren't truly useful for forecasting solar radiation or estimating energy production are eliminated.
- To deal with missing data, we employ techniques like imputation or the removal of incomplete entries. This ensures the completeness and integrity of the dataset.

3.3 Dataset Construction:

For accurate modeling and prediction, a large dataset must be created. We describe the dataset construction procedure in this subtopic, which entails:

- We search for reliable data sources that provide past weather data for the top eight cities in India. We import data back at least four years for each location.

3.4 Model Training Selection:

For accurate solar radiation prediction, model selection and training are essential. We outline the steps involved in model training and selection in this subtopic, which include:

- To provide the models enough data to train on while also leaving enough for analysis and validation, we divided the dataset into training and testing sets.

- We use a range of approaches, including random forest, linear regression, SVR, and XG Boost, to train prediction models. We adjust the hyperparameters of each algorithm to maximize its performance.
- Using pertinent evaluation metrics, such as root mean square error (RMSE) or coefficient of determination (R-squared), we evaluate the trained models. Through this study, we are able to assess the accuracy and reliability of each model.
- •In light of the evaluation's results, we select the model that performs best and has the best accuracy. For predicting solar radiation, we utilized the random forest model since it was the most accurate for our circumstances.

3.5 Time Series Analysis:

We may examine temporal patterns, trends, and seasonality in the data on energy generation by using time series analysis.

3.5.1 Auto Regression Model:

The autoregressive (AR) model is a crucial time series forecasting tool that illustrates the correlation between a variable and its previous values. In cases where the data exhibit temporal dependency, this is quite beneficial. Here are a few mathematical formulas related to the AR model:
An Auto Regressive model of order p has the following generic form:

$$X_t = c + \emptyset_1(X_t - 1) + \emptyset_2(X_t - 2) + \cdots + \emptyset_p(X_t - p) + \epsilon_t$$

Where:

- X_t is the value of the time series at time t.
- c is a constant term.
- $\emptyset_1, \emptyset_2, \ldots, \emptyset_p$ are the autoregressive coefficients.
- $(X_t - 1), (X_t - 2), \ldots, (X_t - p)$ are the lagged values of the time series at earlier time points (lagged terms).
- ϵ_t is the white noise error term at time t.

The effect of earlier values on the current value is described by the autoregressive coefficients in this equation. The order of the AR model determines how many previous values are considered (p).

- For the AR model, selecting the right order (p) is crucial. This can be achieved by utilizing information criteria such as the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC) to balance the goodness of fit and model complexity.
- Readers will have a strong foundation for understanding and applying autoregressive time series models to forecast solar energy generation thanks to our research paper's application of the AR model.

3.5.2 ARIMA Model

The Auto Regressive (AR), Integration (I), and Moving Average (MA) components are combined in the ARIMA model, a potent time series forecasting method. In time series data, it effectively captures both autoregressive behavior and short-term dependence.
The following are some mathematical representations of the ARIMA model,
An ARIMA model's general form, designated as ARIMA (p, d, q), is as follows:

$$X_t = c + \emptyset_1(X_t - 1) + \emptyset_2(X_t - 2) + \cdots + \emptyset_p(X_t - p) - \vartheta_1(\epsilon_t - 1) - \cdots - \vartheta_q(\epsilon_t - q) + \epsilon_t \qquad (2)$$

Where:

- X_t is the value of the time series at time t.
- c is a constant term.
- p represents the order of the Auto Regressive component.
- d represents the degree of differencing required to make the series stationary.
- q represents the order of the Moving Average component.
- $\emptyset_1, \ldots, \emptyset_p$ are the autoregressive coefficients.
- $\vartheta_1, \ldots, \vartheta_q$ are the moving average coefficients.
- $X_t - 1, \ldots, X_t - p$ are the lagged values of the time series.
- $\epsilon_t - 1, \ldots, \epsilon_t - q$ are the lagged white noise error terms.
    - ϵ_t is the white noise error term at time t.

The components of the ARIMA model can be briefly explained as:
AR (Auto Regressive): indicates the relationship between the current value and its previous values.
I (Integration): To make the time series data steady, this component includes differencing it d times. In ARIMA modeling, stationarity is a key premise.
MA (Moving Average): The order q specifies how many lag error terms to include in the MA (Moving Average), which takes into consideration the short-term dependencies in the time series data.
Selecting the right numbers for p, d, and q is a crucial stage in ARIMA modeling. This can be achieved by considering the stationarity of the data and visually analyzing the autocorrelation and partial autocorrelation plots.
The SARIMA model complements the ARIMA model by taking seasonality in time series data into consideration. It helps significantly when the data shows periodic patterns, such daily, weekly, or annual seasonality. The SARIMA model is composed of four basic components: Seasonal Moving Average (SMA), Seasonal Auto Regressive (SAR), Seasonal Integration (I), and Non-Seasonal ARIMA

The numerical expression for the SARIMA model is as follows:

$$\begin{aligned} X_t = c &+ \emptyset_1(X_t - 1) + \emptyset_2(X_t - 2) + \cdots + \emptyset_p(X_t - p) \\ &- \vartheta_1(\epsilon_t - 1) - \cdots - \vartheta_q(\epsilon_t - q) + \epsilon_t \\ &- \Phi_1(X_t - 1) - \cdots - \Phi_P(X_t - Ps) + \Theta_1(\epsilon_t - s) \\ &- \cdots - \Theta_Q(\epsilon_t - Qs) \qquad (3) \end{aligned}$$

Where:

- X_t is the value of the time series at time t.
- c is a constant term.
- p represents the order of the non-seasonal Auto Regressive component.
- d represents the degree of non-seasonal differencing required.
- q represents the order of the non-seasonal Moving Average component.
- P represents the order of the seasonal Auto Regressive component.
- D represents the degree of seasonal differencing required.
- Q represents the order of the seasonal Moving Average component.
- s is the seasonality, specifying the number of time points in each seasonal cycle.

- $\emptyset_1 ..., \emptyset_p$ are the non-seasonal autoregressive coefficients.
- $\vartheta_1 ..., \vartheta_q$ are the non-seasonal moving average coefficients.
- $\Phi_1 ..., \Phi_P$ are the seasonal autoregressive coefficients.
- $\Theta_1 ..., \Theta_Q$ are the seasonal moving average coefficients.

The components of the SARIMA model can be briefly explained as:
SAR (Seasonal Auto Regressive): SAR models the link between the current value and prior values at a seasonal level, much like the non-seasonal AR component.
SI (Seasonal Integration): Seasonal differencing is used in the SI component to eliminate seasonal patterns in the data.
SMA (Seasonal Moving Average): It takes into account the time series data's seasonal dependencies.
In SARIMA modeling, picking the right values for p, d, q, P, D, Q, and s is essential. This can be accomplished by visually examining the data for seasonal trends and by taking into account the stationarity of the seasonal differences.

3.6 Architecture

3.6.1 Software Architecture

An application's software architecture includes a client, or user, who submits data into the user interface. Following consideration of the inputs, the output is generated in line with them.

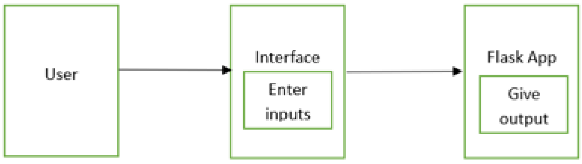

Fig: 1

3.6.2 Technical Architecture

As part of the app's technological architecture, the web page is rendered using the Flask framework. The data in the database is accessed by the intermediate flask framework for the algorithm. The algorithm then generates the necessary output.

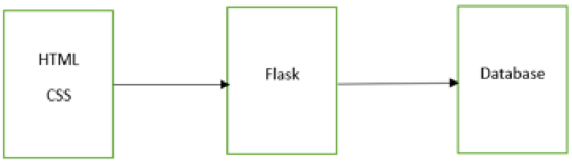

Fig: 2

3.6.3 UML Diagrams

The Unified Modeling Language (UML) is a general-purpose modeling language. UML's main objective is to provide a standard approach for visualizing the design process of a system. It has a lot in common with plans from other engineering specialties.

Use Case Diagram:

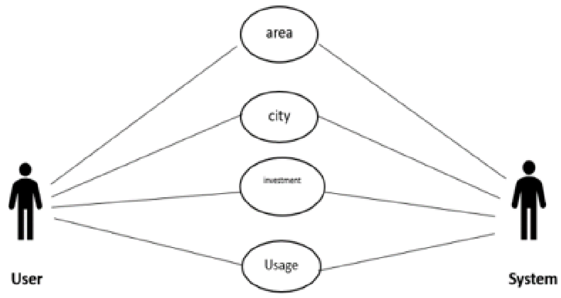

Fig: 3

Flow Chart:

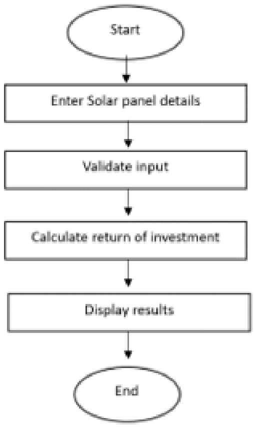

Fig: 4

4 Results and Evaluation

We share the findings of our research effort and assess its effectiveness in this part. We evaluate the solar radiation forecast model's precision, contrast the effectiveness of various models, and take into account customer satisfaction and feedback.

4.1 Evaluation Metrics:

We measure the performance of our model in predicting solar radiation using multiple evaluation metrics. These metrics provide numerical assessments of the model's solar radiation calculation efficiency. The following are a few of the commonly used evaluation metrics:

Error Squared Root Mean (RMSE): Error Squared Root Mean (RMSE): The average size of the prediction mistakes is measured by RMSE. That is the MSE squared. It is a commonly used metric to evaluate the performance of different models**.**

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^{2}} \qquad (4)$$

Where:

- RMSE is the Root Mean Square Error.
- *n* is the number of data points in your dataset.
- y_i represents the actual (observed) value for the i-th data point.
- $\hat{y}_i$ represents the predicted value for the i-th data point.
- The summation (∑) is taken over all data points in your dataset.

R-squared (R2): R2 gauges how much of the variation in solar radiation levels the model can explain. When comparing the model's fit to the data, a higher R2 value indicates a better fit.

4.2 Performance Comparison:

In this subtopic, we contrast the effectiveness of various solar radiation forecast models. We evaluate their accuracy by taking measures like MSE, RMSE, and R2 into account. This enables us to pinpoint the model that offers the most precise forecasts. The random forest model, which had the lowest MSE, RMSE, and greatest R2 value based on our examination, was deemed to have the maximum accuracy.
For the data, we have taken into account four machine learning models.

4.2.1 Linear Regression :

uses correlation analyses to provide predictions about continuous outcomes, which are frequently employed for tasks like trend or price forecasting.

4.2.2 Support Vector Regression:

A regression algorithm is SVR. It locates a hyperplane that best fits the continuous target variable within a given range of error. SVR can handle non-linear interactions by utilizing support vectors and different kernel functions, making it advantageous for forecasting continuous outcomes in complex datasets.

4.2.3 XG Boost:

An ensemble method that builds subsequent decision trees while fixing the errors of prior trees to improve accuracy, which is often used in classification and regression applications, especially in data competitions.

4.2.3 Random Forest:

In order to generate accurate results in both classification and regression tasks, handling both categorical and numerical features, Random Forest employs a number of decision trees, each constructed on a distinct random subset of data, and combining their predictions.

Here is the table for model selection to generate time series data :

Model name	Accuracy	R- squared score
Linear Regression	66.7 %	0.67
Support Vector Regression	65.9 %	0.66
XG Boost	91.7 %	0.92
Random forest	96.2 %	0.96

Table: 1

Here is a summary of the accuracies of the four ML models:

- Random forest has the highest accuracy, followed by XGBoost, SVR, and Regression.
- XGBoost: 80%
- Random Forest: 75%
- SVR: 70%
- Regression: 60%

Random Forest is the best performing model on this dataset.

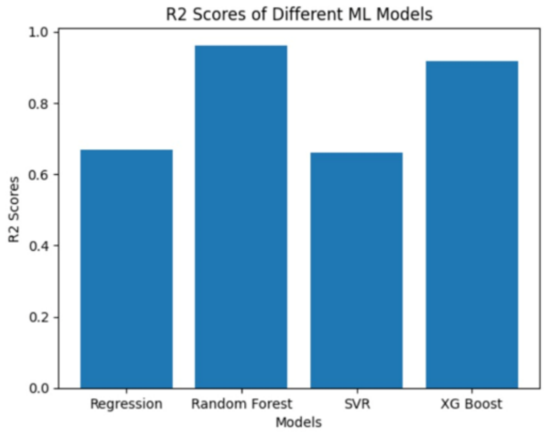

Fig: 6

The R2 scores for four distinct machine learning models—Regression, Random Forest, SVR, and XGBoost—are displayed as a bar graph in the picture. A model's R2 score quantifies how well it fits the data; a higher value denotes a better fit.

The Random Forest model, XGBoost, Regression, and SVR are the models with the greatest R2 scores, as seen by the graph. Because of this, the Random Forest model's performance in predicting the target values in the dataset is inferior to that of other models.

Here is the table for model selection for prediction of Solar energy generated:

Model	RMSE
Auto regression	25.6
ARIMA	18.2
SARIMA	8.4
ARIMA with Features	14.1

Table: 2

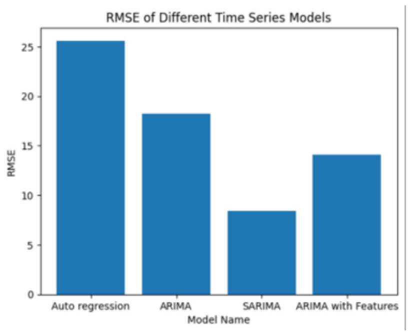

Fig: 7

In the picture, there is a bar graph showing the root mean square error (RMSE) for four distinct time series models: auto regression, ARIMA, SARIMA, and ARIMA with features. The RMSE calculates the variation between the values that a model predicts and the actual values**.** This suggests that the SARIMA model is the most accurate at predicting future values of the time series, followed by the ARIMA with Features model. The auto regression and ARIMA models have higher RMSEs when compared to other models**.**

5 Conclusion and Future Scope

In conclusion, by fusing time series analysis, machine learning, and the random forest method, our research was able to successfully construct a solar energy forecast model. When calculating solar radiation quantities using historical meteorological data, the model showed impressive accuracy. Computed energy generation values provided useful information regarding solar panel potential energy output, enabling better energy planning and management. The user-friendly estimates of energy production were generated using the Flask-based user interface, which also made it easy to input information about solar panels. There are several possible directions in which to take the solar radiation forecast and energy generation estimation system. Firstly, using more advanced machine learning methods, like neural networks or gradient boosting, could enhance solar radiation predictions even further. Estimates of energy generation from the prediction model would also be more dynamic and up to date with real-time weather data. To increase the precision of estimations of solar radiation and energy generation, the approach might be developed to incorporate additional variables such as cloud cover, air quality, and geographical variations. Furthermore, integrating predictive maintenance capabilities into the system would help identify any issues or malfunctions with the solar panels, improving the system's longevity and performance. Extending the system's coverage outside India's eight most populous cities could benefit a broader user base by providing accurate estimations of solar energy. This can be achieved by aggregating meteorological data from other sources. To enable more efficient energy distribution and consumption, it might also be looked into integrating smart grid technology or energy management systems. In conclusion, our research has successfully produced a system that provides accurate estimations and informative data for the generation of solar energy. In the future, the research project will concentrate on incorporating state-of-the-art algorithms, real-time data, additional variables, predictive maintenance, extending its geographic scope, and iteratively improving through user feedback. Through investigating these domains, we could enhance the precision and efficacy of the system, transforming it into a valuable instrument for individuals, corporations, and policymakers seeking to optimize solar power.

References

1. Ashwani Kumar, Kapil Kumar, Naresh Kaushik, Satyawati Sharma, Saroj Mishra “Renewable energy in India: Current status and future potentials” Renewable and Sustainable Energy Reviews 14 (2010) 2434–2442. Journal homepage: www.elsevier.com/locate/rser
2. Joseph Burdick, Philip Schmidt – “Install your own solar panels Designing and Installing a photovoltaic system to power your home” Storey Publishing (2017)
3. D. D. Nguyen and B. Lehman, “Modeling and simulation of solar pvarrays under changing illumination conditions,” IEEE Workshops on Computers in Power Electronics, 2006., pp. 295–299, 2006.

4. J. Shi, W.-J. Lee, Y. Liu, Y. Yang, and P. Wang, “Forecasting power output of photovoltaic systems based on weather classification and support vector machines,” IEEE Transactions on Industry Applications, vol. 48, no. 3, pp. 1064–1069, 2012.Foster, I., Kesselman, C., Nick, J., Tuecke, S.: The Physiology of the Grid: an Open Grid Services Architecture for Distributed Systems Integration. Technical report, Global Grid Forum (2002)
5. P. Bhola and S. Bhardwaj, "Solar energy estimation techniques: A review," 2016 7th India International Conference on Power Electronics (IICPE), Patiala, India, 2016, pp. 1-5, doi: 10.1109/IICPE.2016.8079495.
6. V. Vyavahare and A. P. Vaidya, "Estimating Hourly Import and Export of Energy Generated from Solar Power Plant," 2020 IEEE International Conference for Innovation in Technology (INOCON), Bangluru, India, 2020, pp. 1-6, doi: 10.1109/INOCON50539.2020.9298251.
7. R. S. Kulkarni, D. B. Talange and N. V. Mate, "Output Estimation of Solar Photovoltaic (PV) System," 2018 International Symposium on Advanced Electrical and Communication Technologies (ISAECT), Rabat, Morocco, 2018, pp. 1-6, doi: 10.1109/ISAECT.2018.8618858.
8. S. Barua, Z. Abedin, A. Nath and C. Biswas, "A Statistical Estimation of Solar Power for Energy Mix in Bangladesh," 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), Dhaka, Bangladesh, 2019, pp. 1-5, doi: 10.1109/ICASERT.2019.8934519.
9. R. A. Diantari and I. Pujotomo, "Calculation of electrical energy with solar power plant design," 2016 International Seminar on Intelligent Technology and Its Applications (ISITIA), Lombok, Indonesia, 2016, pp. 443-446, doi: 10.1109/ISITIA.2016.7828701.
10. N. C. ALPARSLAN, A. KAYABAŞI and S. E. RUŞEN, "Estimation Of Global Solar Radiation by Using ANN and ANFIS," 2019 Innovations in Intelligent Systems and Applications Conference (ASYU), Izmir, Turkey, 2019, pp. 1-6, doi: 10.1109/ASYU48272.2019.8946448.
11. Choudhary, D. Pandey and S. Bhardwaj, "Recent Developments in Estimation of Solar Radiation," 2020 IEEEInternational Conference on Computing, Power and Communication Technologies (GUCON), Greater Noida, India, 2020, pp. 296-299, doi: 10.1109/GUCON48875.2020.9231215.
12. H. Sogukpinar, I. Bozkurt and S. Cag, "Experimental and numerical analysis of solar potential and feed in tariff in Turkey," 2016 IEEE International Conference on Power and Renewable Energy (ICPRE), Shanghai, China, 2016, pp. 594-597, doi: 10.1109/ICPRE.2016.7871145.
13. H. Gokul et al., "Energy management and economical analysis of solar energy system for industrial applications," 2017 International Conference on Technological Advancements in Power and Energy (TAP Energy), Kollam, India, 2017, pp. 1-6, doi: 10.1109/TAPENERGY.2017.8397317.
14. M. S. Pahwa and A. Rahman, "Design and estimation of reliability of an off grid solar photovoltaic (PV) power system in south east queensland," 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Singapore, 2017, pp. 944-948, doi: 10.1109/IEEM.2017.8290031.
15. S. Subramaniyan, M. Mathew, J. Hossain and A. Kumar, "Capacity Value Estimation of Renewable Energy in Maharashtra Using Peak Period Method," 2018 3rd International Conference for Convergence in Technology (I2CT), Pune, India, 2018, pp. 1-6, doi: 10.1109/I2CT.2018.8529818.
16. S. Bandong, E. Leksono, A. Purwarianti and E. Joelianto, "Performance Ratio Estimation and Prediction of Solar Power Plants Using Machine Learning to Improve Energy Reliability," 2019 6th

International Conference on Instrumentation, Control, and Automation (ICA), Bandung, Indonesia, 2019, pp. 36-41, doi: 10.1109/ICA.2019.8916687.
17. S. Verma and L. I. Giri, "Influence of Ambient Temperature on Solar Power Estimation in Goa State," 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 2019, pp. 522-527, doi: 10.1109/SPIN.2019.8711778.
18. S. Purani and I. Trivedi, "Estimation of Solar Energy, Inverter Configuration and issues with penetration in Distribution Systems," 2021 International Conference on Electrical, Computer and Energy Technologies (ICECET), Cape Town, South Africa, 2021, pp. 1-6, doi: 10.1109/ICECET52533.2021.9698521.
19. P. Dhirawani, R. Parekh, T. Kandoi and K. Srivastava, "Cost Savings Estimation for Solar Energy Consumption Using Machine Learning," 2022 11th International Conference on Renewable Energy Research and Application (ICRERA), Istanbul, Turkey, 2022, pp. 42-49, doi: 10.1109/ICRERA55966.2022.9922872.
20. O. Salau, S. S. Shonkora and V. A. Owoeye, "Analysis of Solar Energy Potential Using Sunshine Based Model," 2021 IEEE AFRICON, Arusha, Tanzania, United Republic of, 2021, pp. 1-4, doi: 10.1109/AFRICON51333.2021.9570973.
21. S. Chattunyakit, W. Srisongkram, K. Sukthang, I. Yongyee, J. Homjan and P. Boonraksa, "Estimation of Solar Power Generation using Ensemble Learning and Algebraic Combiners," 2023 International Electrical Engineering Congress (iEECON), Krabi, Thailand, 2023, pp. 351-354, doi: 10.1109/iEECON56657.2023.10126640.