SOLAR RADIATION LEVEL PREDICTION USING MACHINE LEARNING

V.R.B.Rohini
Assistant Professor
Department of Information Technology
S.R.K.R. Engineering College(A)
SRKR Marg, Bhimavaram.

Borra pavani
Department of Information Technology
S.R.K.R. Engineering College(A)
SRKR Marg, Bhimavaram.

Duggirala varsha sri
Department of Information Technology
S.R.K.R. Engineering College(A)
SRKR Marg, Bhimavaram.

Chekuri kusuma keerthi
Department of Information Technology
S.R.K.R. Engineering College(A)
SRKR Marg, Bhimavaram.

Chekuri durga sri sowmya
Department of Information Technology
S.R.K.R. Engineering College(A)
SRKR Marg, Bhimavaram.

Abstract— Solar energy is an abundant and sustainable source of power, making it a critical component of the global transition towards clean energy solutions. Accurate prediction of solar radiation is essential for optimizing the performance of solar energy systems, such as photovoltaic panels and solar thermal plants. This study presents a comprehensive exploration of machine learning techniques for the prediction of solar radiation. Machine learning models have been developed and trained on historical solar radiation data, incorporating various meteorological parameters, geographical factors, and timerelated features. The predictive accuracy of these models has been evaluated using real-world datasets from diverse geographic locations and climates. The results demonstrate the effectiveness of machine learning algorithms in accurately forecasting solar radiation levels. These predictions can empower energy stakeholders, grid operators, and solar energy system operators to make informed decisions regarding energy generation, distribution, and consumption. Additionally, the study highlights the significance of feature engineering, model selection, and hyper parameter tuning in enhancing prediction performance. The methodology involves data Preprocessing, feature engineering, model selection, and evaluation using metrics like Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). Additionally, model interoperability techniques, such as feature importance analysis, will be employed to enhance the transparency of predictions

Keywords— Solar radiation level, preprocessing, Machine learning, Decision Tree, Random Forest, AdaBoost, Linear Regression, KNN.

I. INTRODUCTION

Solar energy has emerged as a pivotal component of the world's transition to sustainable and clean energy sources. The harnessing of solar power heavily relies on the accurate prediction of solar radiation, which is vital for optimizing the efficiency and reliability of solar energy systems. As climate change concerns and the quest for greener energy solutions continue to grow, the development of effective solar radiation prediction models becomes increasingly imperative. This project aims to address this pressing need by leveraging the power of machine learning to predict solar radiation levels. Machine learning, with its ability to discern complex patterns from large datasets, offers a promising approach to enhance the accuracy and precision of solar radiation forecasting. By integrating historical solar radiation data with various meteorological, geographical, and time-related factors, machine learning models can provide real-time and future predictions of solar radiation levels.

The significance of this project lies in its potential to revolutionize the renewable energy sector. Accurate solar radiation predictions can assist energy grid operators, solar power plant managers, and energy policymakers in making informed decisions about energy generation, storage, and distribution. Moreover, these predictions can contribute to reducing the dependency on non-renewable energy sources and mitigating greenhouse gas emissions. This project's methodology encompasses data preprocessing, feature engineering, model selection, and performance evaluation using metrics such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). In this endeavour, we will explore various machine learning techniques, conduct rigorous data analysis, and develop predictive models to optimize solar radiation forecasting. By doing so, we aim to support the transition towards a more sustainable and environmentally friendly energy future.

II. LITERATURE SURVEY (RELATED WORK)

[1] **Hossain, M.R., O. A.M.T., Shawkat Ali, A.B.M.: Hybrid prediction method for solar power using different computational intelligence algorithms. Smart Grid and Renewable Energy 4(1), 76-87 (2013).** In the study by Hossain, M.R., et al. (2013), a hybrid prediction method for solar power generation is proposed, utilizing various computational intelligence algorithms. This research focuses on integrating machine learning techniques to forecast solar radiation levels accurately. The study explores the synergy between different computational methods to enhance the prediction accuracy of solar energy generation, contributing to the efficient utilization of solar power resources.

[2] **E. Lorenz, T. Scheid Steger, J. Hurka, D. Heinemann, and C. Kurz, Regional PV power prediction for improved grid integration, Prog. Photovoltaics Res. Appl., vol. 19, no. 7, pp. 757–771, 2011.**The paper authored by E. Lorenz et al. in 2011 discusses the application of regional photovoltaic (PV) power prediction methods aimed at enhancing the integration of PV systems into the grid. While not primarily focused on machine learning, it addresses the importance of accurate solar radiation prediction, which is a fundamental component of PV power generation. The authors likely highlight the significance of reliable solar radiation forecasting as part of their broader strategy for improving grid integration of solar energy systems.

[3] M.Abuella and B. Chowdhury, Solar Power **Probabilistic Forecasting by Using Multiple Linear Regression Analysis, in IEEE Southeastcon Proceedings, Ft. Lauderdale, FL, 2015.** In their study, Abuela and Chowdhury (2015) applied multiple linear regression analysis to probabilistically forecast solar power generation, showcasing the use of machine learning for accurate solar

radiation prediction and its relevance in optimizing renewable energy systems

[4] Friedman, J.H.: Greedy function approximation: a gradient boosting machine. Annals of Statistics pp. 1189{1232 (2001). Friedman's 2001 paper on "Greedy function approximation: a gradient boosting machine" introduces a foundational algorithm for machine learning. While not directly related to solar radiation prediction, it lays the groundwork for powerful ensemble techniques like gradient boosting, which can be effectively applied to enhance the accuracy of solar radiation prediction models.

[5] Faizan Jawiad, Khurum Nazir Junejo Predicting Daily Mean Solar Power Using Machine Learning Regression Techniques. In their research paper titled "Predicting Daily Mean Solar Power Using Machine Learning Regression Techniques," Faizan Jawiad and Khurum Nazir Junejo investigate the application of machine learning regression methods to forecast daily mean solar power output, contributing to the field of solar radiation prediction and renewable energy optimization.

[6] Paulescu, E. Paulescu, P. Gravila, V. Badescu, Weather Modelling and Forecasting of PV Systems Operation, Springer London, London, 2013. In their research paper titled "Predicting Daily Mean Solar Power Using Machine Learning Regression Techniques," Faizan Jawiad and Khurum Nazir Junejo investigate the application of machine learning regression methods to forecast daily mean solar power output, contributing to the field of solar radiation prediction and renewable energy optimization.

[7] Gangwani, P., Soni, J., Upadhyay, H., & Joshi, S. (2020), Proposed a deep learning approach for modelling geothermal energy prediction. The study aims to develop an accurate model for geothermal energy prediction, which can help in the efficient utilization of geothermal energy resources. The proposed approach involves the use of long short-term memory (LSTM) networks for predicting the geothermal energy output. The study aims to evaluate the performance of the proposed model and compare it with other existing methods for geothermal energy prediction.

[8] Erduman, A. (2020) ,The objective of this study was to develop a smart short-term solar power output prediction system using an artificial neural network (ANN) to improve the accuracy of solar power output prediction. The study aimed to use the ANN model to predict the solar power output of a photovoltaic (PV) system over a short period of time (15 minutes) and to compare the results with those of traditional methods such as persistence and linear regression. The ultimate goal was to improve the efficiency and reliability of solar power generation by accurately predicting solar power output.

III. SYSTEM IMPLEMENTATION

- ### Proposed approach

Proposed several machine learning models to Predict the Solar radiation, but none have adequately addressed this misdiagnosis problem. Also, similar studies that have proposed models for evaluation of such performance classification mostly do not consider the heterogeneity and the size of the data Therefore, we propose a KNN, Decision Tree, Random forest and AdaBoost Regression techniques.

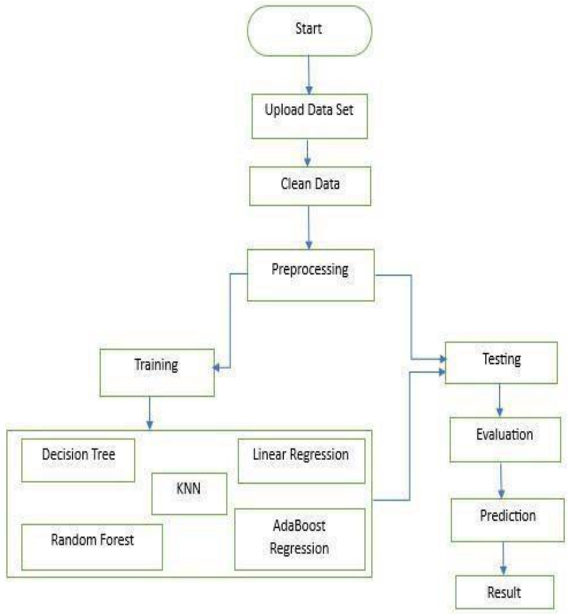
- ### System Architecture

The methodology involves data preprocessing, feature engineering, model selection, and evaluation using metrics like Mean Absolute Error (MAE) and Root Mean Squared Error.

1.Take the Dataset: The system accepts and processes the dataset provided by the user. This dataset forms the foundation for building the predictive model

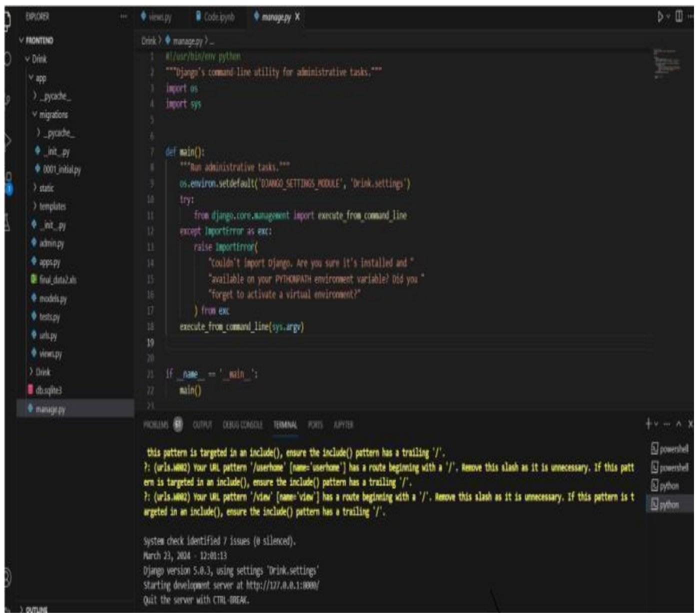
2. Preprocessing: Before training a predictive model, the system preprocesses the dataset. This includes handling missing data, data cleaning, and feature extraction. Preprocessing ensures that the data is in a suitable format for modelling.

3.Training: The system uses machine learning techniques and Python modules to train a model based on the pre-processed dataset. The model learns patterns and relationships within the data, allowing it to make predictions. .



IV. EXPERIMENTS & RESULTS

The project aimed to predict solar radiation levels using machine learning techniques. The dataset comprised 32,656 rows and 11 columns including temperature, humidity, pressure, Unix time data, and sunrise/sunset times. Various models such as linear regression, decision tree, random forest, k-nearest neighbours (KNN), and kmeans were trained and hyper-parameters were tuned to achieve high accuracy, with random forest reaching 94%. The front end of the project included an about section explaining the project, data upload functionality, data viewing with customizable splitting values, and model accuracy visualization through dropdown selection. Additionally, users could input variables except radiation for the model to predict solar radiation levels based on learned patterns from the training data.



Knn :

```
Accuracy = 0.9460569927175618
mean_absolute_error = 25.303633897613707
mean_squared_error = 5373.068763862533
```

K- Means:

```
Accuracy = -0.423800112000458
mean_absolute_error = 208.3016118881119
mean_squared_error = 142189.09505220279
```

Decision Tree:

```
Accuracy = 0.9143703546518978
mean_absolute_error = 34.808835611089265
mean_squared_error = 8529.260711615985
```

Random Forest :

```
Accuracy = 0.9412227651639078
mean_absolute_error = 27.96482933918009
mean_squared_error = 5854.588767557216
```

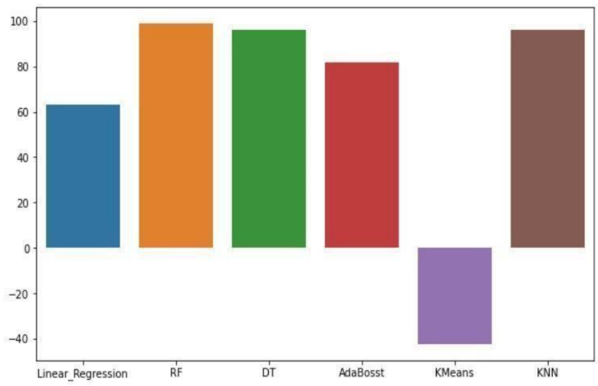
Linear Regression :

```
Accuracy = 0.629148896543411
mean_absolute_error = 145.95666790871613
mean_squared_error = 37035.38323614506
```

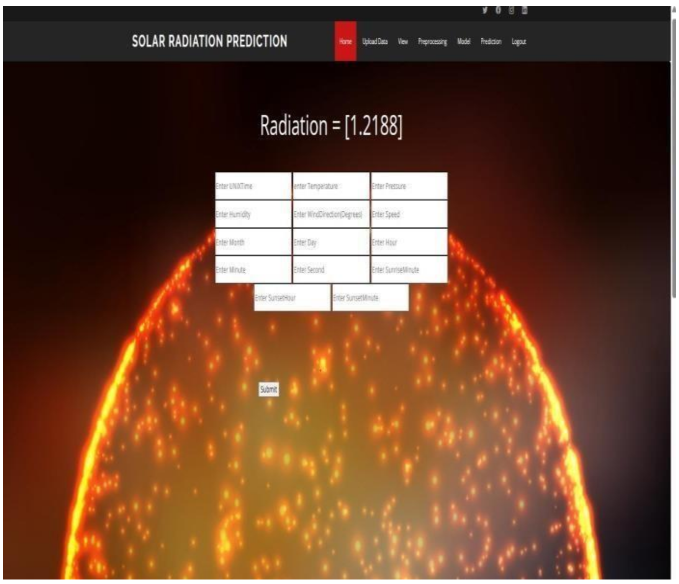
AdaBoost :

```
Accuracy = 0.8165923912038677
mean_absolute_error = 85.19432088075253
mean_squared_error = 18316.167909110332
```

Visualization of models :



Front end:



V. CONCLUSION

In conclusion, the application of machine learning techniques for solar radiation prediction represents a crucial advancement in the field of renewable energy. This technology has demonstrated its potential to significantly improve the efficiency and reliability of solar energy systems. Accurate predictions empower energy stakeholders to optimize energy generation, enhance grid integration, reduce costs, and minimize environmental impact. As we strive for a more sustainable and eco-friendly energy future, machine learning based solar radiation forecasting emerges as a vital tool, paving the way for the widespread adoption of solar power and contributing to our global efforts to combat climate change and achieve energy sustainability.

VI. FUTURE ENHANCEMENT

High-Resolution Predictions Developing models capable of providing high-resolution solar radiation predictions on a regional or even local scale will be crucial. This will allow for precise energy production estimates for specific locations, improving the efficiency of distributed solar installations. Incorporating Climate Change Data Integrating climate change data and climate models will help account for long-term shifts in weather patterns, ensuring the sustainability of solar energy projects amid changing environmental conditions. Hybrid Models Combining machine learning with physical models, such as radiative transfer models, can enhance accuracy. Hybrid models can leverage the strengths of both approaches, improving the reliability of predictions. Data Fusion Utilizing diverse data sources, including satellite imagery, ground-based sensors, and atmospheric data, can lead to more comprehensive input features for machine learning models, resulting in more robust predictions.

VII. REFERENCES

[1] Hossain, M.R., O. A.M.T., Shawkat Ali, A.B.M.: Hybrid prediction method for solar power using different computational intelligence algorithms. Smart Grid and Renewable Energy 4(1), 7687 (2013).

[2] E. Lorenz, T. Scheidsteger, J. Hurka, D. Heinemann, and C. Kurz, “Regional PV power prediction for improved grid integration,” Prog. Photovoltaics Res. Appl., vol. 19, no. 7, pp. 757– 771, 2011.

[3] M. Abuella and B. Chowdhury, “Solar Power Probabilistic Forecasting by Using Multiple Linear Regression Analysis,” in IEEE Southeastcon Proceedings, Ft. Lauderdale, FL, 2015.

[4] Friedman, J.H.: Greedy function approximation: a gradient boosting machine. Annals of Statistics pp. 1189{1232 (2001).

[5] Faizan Jawiad, Khurum Nazir Junejo “Predicting Daily Mean Solar Power Using Machine Learning Regression Techniques “.

[6] Paulescu, E. Paulescu, P. Gravila, V. Badescu, Weather Modeling and Forecasting of PV Systems Operation, Springer London, London, 2013.