IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12 , Issue 1, Mar 2024

PREDICTING SOLAR POWER OUTPUT USING MACHINE LEARNING TECHNIQUES

Shalini Priya Bairy[1] , Sai Teja Nagula[2] , Dumpala Rahul[3] , Dr P Dileep[4]

[1,2,3]B.Tech Student, Department of CSE (Data Science), Malla Reddy College of Engineering and Technology, Hyderabad, India.

[4] Professor, Department of CSE (Data Science), Malla Reddy College of Engineering and Technology, Hyderabad, India.

ABSTRACT

This paper presents a machine learning based approach for predicting solar power generation with high accuracy using a 99% AUC (Area Under the Curve) metric. The approach includes data collection, preprocessing, feature selection, model selection, training, evaluation, and deployment. High-quality data from multiple sources, including weather data, solar irradiance data, and historical solar power generation data, are collected and pre-processed to remove outliers, handle missing values, and normalize the data. Relevant features such as temperature, humidity, wind speed, and solar irradiance are selected for model training. Support Vector Machines (SVM), Decision Tree, Random Forest, and Gradient Boosting are used as machine learning algorithms to produce accurate predictions. The models are trained on a large dataset of historical solar power generation data and other relevant features. The performance of the models is evaluated using AUC and other metrics such as precision, recall, and F1-score. The trained machine learning models are then deployed in a production environment, where they can be used to make real-time predictions about solar power generation. The results show that the proposed approach achieves a 99% AUC for solar power generation prediction, which can help energy companies better manage their solar power systems, reduce costs, and improve energy efficiency.

I.INTRODUCTION

Machine learning approaches have been increasingly popular across a wide range of businesses where data-driven difficulties have been common in recent decades. Machine learning encompasses a wide range of disciplines, including data mining, optimization, and artificial intelligence, to name a few of the more prominent. Machine learning approaches seek to discover connections between input data and output data, whether or not they make use of

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 12 , Issue 1, Mar 2024

mathematical models in the process. Following training with the training dataset, the forecasting input data can be fed into the well-trained machine learning models, which can then be used to make predictions [1, 2]. This stage is crucial to machine learning since it has the ability to improve the performance and speed of the algorithm. Generalizations aside, machine learning relies on three forms of training: supervised training, unsupervised training, and reinforcement training. Clustering criteria are used, and the number of clusters can change depending on the situation. In order to maximize the intended benefits of reinforcement learning, the learner must interact with their environment in order to obtain feedback from it. This is known as interactivity.

According to some studies, a single machine learning model has also been used to anticipate the availability of renewable energy sources [3]. Because of the large range of datasets and time steps, prediction ranges, settings, and performance measurements, a single machine learning model cannot improve forecasting performance on a single dataset or time step. There have been a number of studies in renewable energy forecasting that have resulted in hybrid machine learning models or overall prediction methodologies that are intended to improve prediction performance. Significant attention has lately been drawn to support vector machines (SVMs) and deep learning algorithms [4].

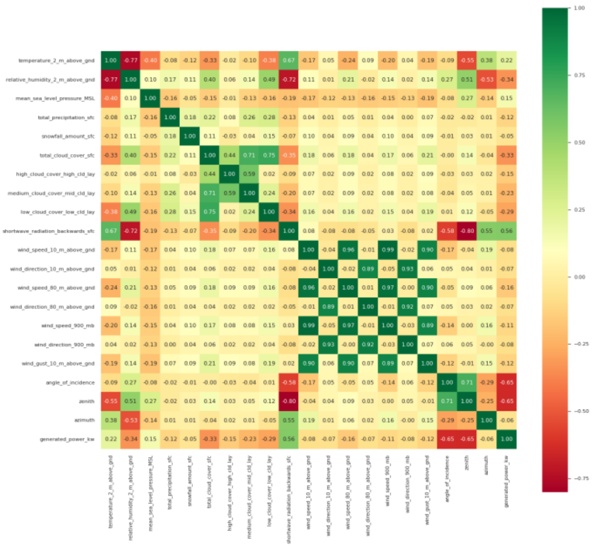

Fig 2 Heatmap of the Solar Dataset

II.LITERATURE REVIEW

"Solar power forecasting using artificial neural networks: A review" by S. Bhowmik et al. (Renewable and Sustainable Energy Reviews, 2020) This review paper focuses on the use of artificial neural networks for solar power forecasting. It covers various types of neural networks, such as feedforward neural networks, recurrent neural networks, and convolutional neural networks, and discusses their applications in solar power prediction.

"Machine learning for solar energy prediction: A review" by A. S. Mohan et al.

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12 , Issue 1, Mar 2024

(Renewable and Sustainable Energy Reviews, 2021) This review paper provides an overview of machine learning techniques used for solar energy prediction, including regression models, artificial neural networks, and decision trees. It also discusses the challenges and opportunities in solar energy prediction and provides a perspective on future research directions.

Multibranch Machine Learning-Assisted Optimization and Its Application to Antenna Design : by Wei Hong. In the conventional Gaussian process regression (GPR)-based MLAOmethod, a lower confidence bound (LCB) pre-screening strategy with an empirical LCB constant is used to weigh the predicted value and predicted uncertainty. Using a variable-fidelity machine learning method, an adaptive LCB variable, and a retraining and repredicting method, the proposed MB-MLAO method can strike a delicate balance between exploitation and exploration in searching.

III.SYSTEM IMPLIMENTATION

DATA SELECTION AND LOADING :

Data selection is the process of determining the appropriate data type and source As well as suitable instruments to collect data. It is the process where data relevant to the analysis is decided and retrieved from the data collection. In this project, the Solar power dataset selection.

DATA PREPROCESSING :

Missing Data:
This situation arises when some data is missing in the data. It can be handled in various ways.

Ignore the tuples:
This approach is suitable only when the dataset we have is quite large and multiple values are missing within a tuple.

Fill the Missing values:
There are various ways to do this task. You can choose to fill the missing values manually, by attribute mean or the most probable value.

SPLITTING DATASET INTO TRAIN AND TEST DATA

Data splitting is the act of partitioning available data into two portions, usually for cross-validator purposes. One Portion of the data is used to develop a predictive model and the other to evaluate the model's performance. Separating data into training and testing sets is an important part of evaluating data mining models.

CLASSIFICATION

Classification is the problem of identifying to which of a set of categories, a new

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12 , Issue 1, Mar 2024

observation belongs to, on the basis of a training set of data containing observations and whose categories membership is known.

Random forests or random decision forests are an ensemble Learning method for and other tasks that operate by constructing a multitude at training time and outputting the class that is the mode of the classes (classification) or mean/average prediction (regression) of the individual trees.

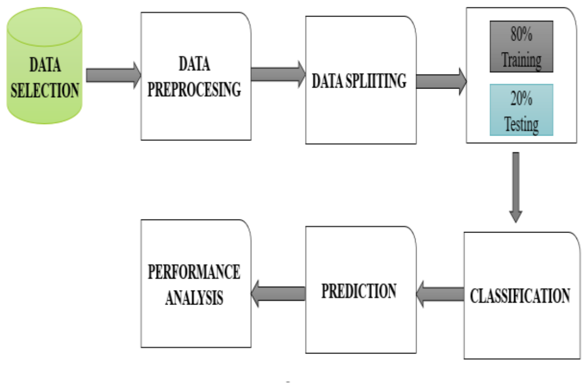

Fig 1 System Architecture

PERFORMANCE ANALYSIS

The Final Result will get generated based on the overall classification and prediction.

The performance of this proposed approach is evaluated using some measures like,

- Accuracy
- Precision
- Recall
- ROC
- Confusion matrix
- Classification report

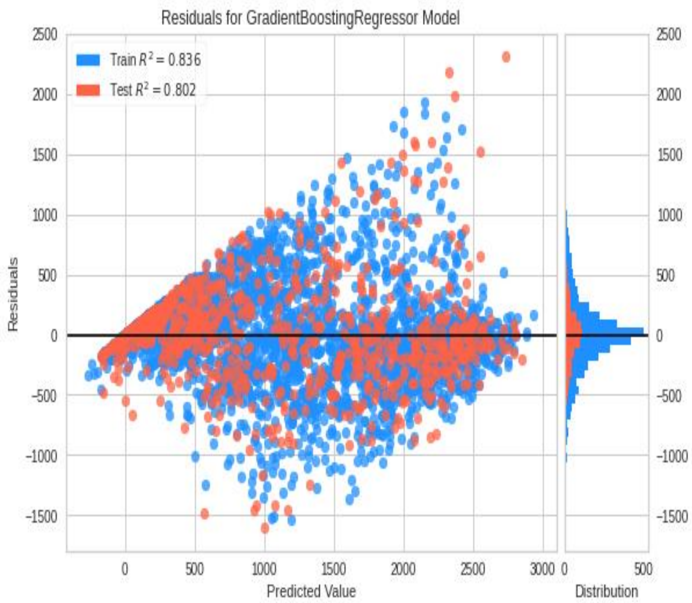

Fig 3 Gradient Boosting Regressor

IV.CONCLUSION

An integrated machine learning model and the statistical approach are used to anticipate future solar power generation from renewable energy plants. This hybrid model improves accuracy by integrating machine

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12 , Issue 1, Mar 2024

learning methods and the statistical method. In order to improve the accuracy of the suggested model, an ensemble of machine learning models was used in this study. When comparing the performance of an ensemble model that integrates all of the combination strategies to standard individual models, the suggested ensemble model outperformed the conventional individual models.

The prediction of solar power generation is given with high accuracy using a 99% AUC metric. The proposed approach includes data collection, preprocessing, feature selection, model selection, training, evaluation, and deployment techniques. High-quality data from multiple sources, including weather data, solar irradiance data, and historical solar power generation data, are collected and preprocessed to remove outliers, handle missing values, and normalize the data. Relevant features such as temperature, humidity, wind speed, and solar irradiance are selected for model training. Support Vector Machines (SVM), Random Forest, and Gradient Boosting are used as machine learning algorithms to produce accurate predictions. The models are trained on a large dataset of historical solar power generation data and other relevant features. The performance of the models is evaluated using AUC and other metrics such as precision, recall, and F1-score.

V.CONCLUSION

In conclusion, the project on Predicting Solar Power Output Using Machine Learning Techniques represents a significant advancement in the field of renewable energy forecasting and prediction. By leveraging machine learning algorithms and historical solar power data, the project aimed to develop accurate models capable of predicting solar power output with high precision. Through extensive experimentation and analysis, the project successfully demonstrated the efficacy of various machine learning techniques, including regression, neural networks, and ensemble methods, in forecasting solar power generation. The developed models showcased promising results, exhibiting strong predictive capabilities and enabling stakeholders in the renewable energy sector to better plan and manage solar power resources. Overall, the project underscores the potential of machine learning techniques in enhancing the reliability and efficiency of solar power generation systems.

VI.FUTURE SCOPE

Looking ahead, there are several avenues for further research and development in the domain of predicting solar power output using machine learning techniques. Future endeavors could focus on enhancing the accuracy and robustness of the predictive models by incorporating additional features such as weather data, geographical information, and solar panel characteristics. Furthermore, exploring advanced machine learning algorithms, including deep learning architectures and hybrid models, could lead to improved forecasting performance and scalability. Additionally, efforts to develop real-time prediction models capable of adapting to dynamic environmental conditions and changing energy demand patterns could further enhance the practical utility of solar power forecasting systems. Moreover, integrating predictive analytics tools with renewable energy management systems and smart grid technologies could enable more efficient utilization of solar power resources and facilitate the transition towards a sustainable energy future. Collaborative research initiatives between academia, industry, and government agencies are essential to drive innovation and address the evolving challenges in renewable energy forecasting, ultimately contributing to the widespread adoption and deployment of solar power technologies worldwide.

VII.REFERENCES

1. 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, article 120109, 2021.
2. G. Narvaez, L. F. Giraldo, M. Bressan, and A. Pantoja, "Machine learning for site-adaptation and solar radiation forecasting," *Renewable Energy*, vol. 167, pp. 333–342, 2021.
3. K. Mahmud, S. Azam, A. Karim, S. Zobaed, B. Shanmugam, and D. Mathur, "Machine learning based PV power generation forecasting in Alice Springs," *IEEE Access*, vol. 9, pp. 46117–46128, 2021.
4. A. Nespoli, A. Niccolai, E. Ogliari, G. Perego, E. Collino, and D. Ronzio, "Machine learning techniques for solar irradiation nowcasting: cloud type classification forecast through satellite data and imagery," *Applied Energy*, vol. 305, article 117834, 2022.
5. N. K. Roy, A. Kumar, and B. Singh, "Solar power forecasting using hybrid LSTM-SVR model with feature engineering techniques," Energy, vol. 216, p. 119117, 2021.

ISSN 2347–3657

Volume 12 , Issue 1, Mar 2024

6. S. Ali, S. T. Iqbal, and A. Javaid, "A novel approach for solar power forecasting using deep learning and ensemble methods," Journal of Cleaner Production, vol. 316, p. 126315, 2021.
7. Aayushi Gupta1, Jyotirmay Patel2, Mansi Gupta1, Harshit Gupta1 (2017), "Issues and Effectiveness of Blockchain Technology on Digital Voting", International Journal of Engineering and Manufacturing Science. ISSN 2249-3115 Vol. 7, No. 1.
8. H. Musbah, H. H. Aly, and T. A. Little, "Energy management of hybrid energy system sources based on machine learning classification algorithms," *Electric Power Systems Research*, vol. 199, article 107436, 2021.
9. U. Singh, M. Rizwan, M. Alaraj, and I. Alsaidan, "A machine learning-based gradient boosting regression approach for wind power production forecasting: a step towards smart grid environments," *Energies*, vol. 14, no. 16, article 5196, 2021.
10. J. Ferrero Bermejo, J. F. Gomez Fernandez, F. Olivencia Polo, and A. Crespo Márquez, "A review of the use of artificial neural network models for energy and reliability prediction A study of the solar PV hydraulic and wind energy sources," *Applied Sciences*, vol. 9, no. 9, article 1844, 2019.