iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL
FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 Issue: VII Month of publication: July 2024

DOI: https://doi.org/10.22214/ijraset.2024.63720

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue VII July 2024- Available at www.ijraset.com

Renewable Energy Prediction Using Different ML Algorithms

Azka Ihtesham Uddin Ahmed[1], Dr. Balasubbareddy Mallala[2]
[1]M.E Student EEE Department, Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana
[2]Professor EEE Department, Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana

Abstract: The global community is grappling with numerous obstacles in fulfilling its energy requirements, stemming from the limited availability of fossil fuel reserves. It is, at present, rapidly shifting its focus towards Renewable Energy Sources to overcome the limitations of fossil fuels and generate environmentally friendly energy. Earlier, their unpredictable nature made them less appealing, but recent progress in Artificial Intelligence (AI) and Machine Learning (ML) has allowed for accurate forecasts of their energy production. With a well-planned approach to balancing energy loads, renewable energy sources can satisfy energy needs without the necessity for extra storage. In this piece, we conducted a machine learning analysis on a dataset obtained from the Kaggle Platform by utilizing various algorithms, such as Random Forest (RF), Linear Support Vector Regression (SVM), Decision Tree (DT), and Linear Regression (LR).

The performance of the system was assessed, and it was observed that the feature selection method known as Mutual Information Regression (FS) by choosing 11 features with the highest correlation to the target variable, the Linear Regression method resulted in the lowest error rates. The system's error rates were measured as MSE=0.105, MAE=0.2, MAPE=0.214, and R2 score=0.999.

Keywords: Renewable sources, CO2, AI, ML, FS-feature selection, Forecasting, RF, Linear SVM, DT, LR, MSE, MAE, MAPE, R2

I. INTRODUCTION

The global community is grappling with numerous problems to satisfy energy demands, stemming from the limited availability of fossil fuel reserves. The production of energy from traditional sources is leading to a variety of environmental changes and is a significant factor in the depletion of the ozone layer, which poses a threat to our existence. In particular, the energy produced by thermal power plants is the largest contributor to CO2 emissions into the atmosphere.

Achieving carbon neutrality necessitates a decrease in greenhouse gas emissions. Additionally, it's essential for [1]governments and legislative bodies to implement appropriate measures, including policies, regulations, and rules, to address the escalating issues related to CO2. Renewable energy sources[2] have been identified as the most effective solution to meet consumer needs. Previously, their intermittent nature made them less attractive, but advancements in Artificial Intelligence (AI) and Machine Learning (ML) [3] have enabled precise predictions of their output. With an optimized strategy for load balancing, renewable energy sources can meet energy demands without the need for additional storage. Renewable energy is defined by its ability to be replenished and its widespread availability. The various forms renewable energies that can be harnessed to generate electricity include Solar, Wind, Hydro, Geothermal, and Biogas, among others [4]. While it's true that renewable energy cannot produce power as instantaneously as traditional power plants, and the upfront costs are high, its long-term advantages are unparalleled.

In this article, [5] we explore the future prospects of renewable energy by employing different machine learning algorithms for prediction and conducting a comparative analysis on the dataset obtained from the Kaggle platform..

II. LITERATURE SURVEY

The authors in [6] discusses the potential shift from traditional, old-fashioned electricity production to newer, renewable energy sources. They also propose different machine learning techniques that could be utilized in the model to make predictions.

[7], [8] suggested the transformation of traditional grids into advanced smart grids, and combining them with Internet of Things (IoT), Machine Learning (ML), and other bidirectional communication technologies can enhance the efficiency of managing power and balancing supply and demand. Moreover, ML enables smart grids to handle crises by utilizing its extensive database, which is derived from historical data.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue VII July 2024- Available at www.ijraset.com

Writers in [9] mentioned that by combining renewable energy sources, battery energy storage systems, and distribution generating systems into the grid, the reliability of the power supply can be enhanced. This ensures that power is available even during grid failures. Additionally, it lowers the cost of producing electricity, which in turn decreases the amount consumers have to pay for their electricity bills.
[10] As renewable energy sources are being more widely incorporated into the primary power network, the task of precise forecasting through traditional methods has become more intricate, resulting in reduced accuracy. Consequently, Machine Learning (ML) and Deep Learning (DL) algorithms have gained popularity for their capacity to identify more intricate patterns within the data and provide reliable forecasts.
[11] This study examines whether Solar PV (photovoltaic) technology can economically match coal-fired power in Chinese urban areas without government subsidies. Furthermore, this research also explores the financial feasibility of investing in solar PV projects. Initially, the research evaluates how much distributed solar power can replace local coal-fired power stations throughout China. Additionally, it predicts that with solar power, there is potential for moderate to high financial returns.

III.MATERIAL & METHODS

The information was gathered from the Kaggle Platform. It includes data on a monthly basis spanning from 1973 to 2024, sourced from 13 various power sources, including Hydroelectric, Geothermal, Solar, Wind, Wood energy, Waste energy, Ethanol, Biomass energy, Renewable diesel fuel, other biofuels, Conventional hydroelectric power, and Bio diesel. The monthly consumption is categorized according to its use, including Commercial, Industrial, Residential, Electric power, and Transportation.

A. Data Preprocessing

The data acquired from Kaggle contained numerous inconsistencies and varied structures. It was subsequently transformed into a format suitable for machine learning, including converting entries in the dataset from objects to floats. Following this, the dataset was divided into a training set (20%) and a testing set (80%). The "Total Energy from Renewables" was designated as the target variable.

B. Feature selection

Upon thorough examination, it has been noted that not every parameter is connected to one another. Consequently, it becomes necessary to select the most relevant features to the target variable. Two distinct feature selection techniques were utilized in the model: the correlation coefficient method and the mutual info regression method.
In the correlation coefficient method, features that are most closely related to the selected target variable are pinpointed, while the others are discarded. A specific condition is set: the correlation coefficient must be greater than 0.8 and less than 1 to eliminate the remaining features.

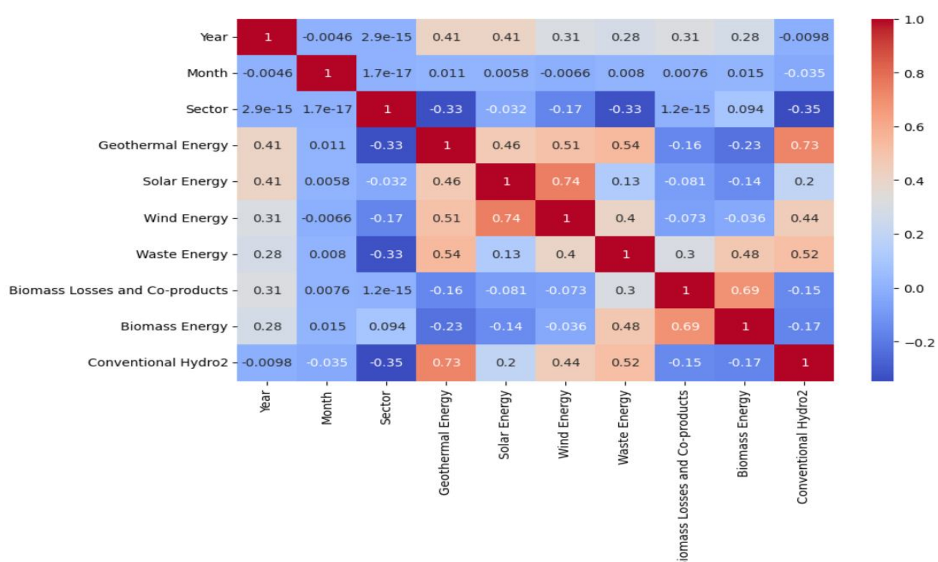


Fig. 1. Shows heatmap of Coefficient Correlation FS method

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue VII July 2024- Available at www.ijraset.com

The mutual info regression method involves ranking of features in relation with the target variable. This method is particularly effective for tackling non-linear regression issues or in situations where the connection between the attributes and the outcome is unclear.

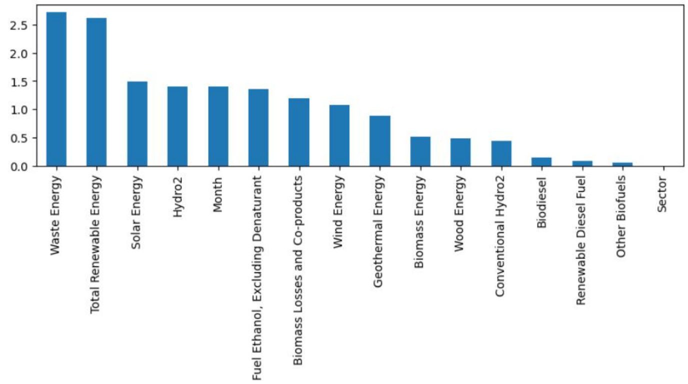


Fig. 2. Shows bar graph of Mutual Info Regression FS method

C. Machine learning algorithms

To generate most accurate predictions different machine learning algorithms are taken into consideration.

1) *Random Forest Regression:* A Random Forest Regression model merges several decision trees to form a unified model. Every tree within the forest is constructed from a unique subset of the data and generates its own separate prediction. The ultimate prediction for the given input is determined by the mean or the weighted mean of all the predictions from the individual trees.
2) *Linear SVR:* Support vector regression (SVR) belongs to the category of support vector machines (SVM) designed for tasks involving predicting continuous output values based on input values. It aims to discover the optimal function for making these predictions.
3) *Decision Tree Regressor:* Decision Tree Regression analysis, a robust method for examining data, assists companies and scholars in making well-informed choices by forecasting results from past records. It supports in predicting future trends, evaluating risks, and spotting patterns, serving a crucial function across various areas. Decision trees are capable of dealing with both numerical and categorical data without requiring one-hot encoding or additional preprocessing steps. This feature makes them suitable for datasets containing a variety of data types.
4) *Linear Regression:* The aim of a linear regression method is to forecast the value of a dependent variable using an independent variable. The stronger the linear connection between the independent and dependent variables, the more precise the forecast becomes. Companies gather vast quantities of data, and linear regression assists them in utilizing this data to enhance their understanding of the world rather than depending solely on experience and gut feelings. You can process extensive raw data and convert it into useful insights.

D. Implementation

This research employs various machine learning models, such as Random Forest Regression, Linear Support Vector Regression, Decision Tree Regression, and Linear Regression, to develop prediction models.

```
X = data.drop(columns = ['Total Renewable Energy'], axis=1)
y = data['Total Renewable Energy']
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.20, shuffle = True)

from sklearn.ensemble import RandomForestRegressor
model=RandomForestRegressor()
model.fit(X_train, y_train)
Y_pred = model.predict(X_test)

from sklearn.metrics import mean_squared_error, r2_score, mean_absolute_error, mean_absolute_percentage_error

print("Mean_Squared_Error: ", mean_squared_error(Y_pred, y_test))
print('Mean Absolute Error: ', mean_absolute_error(Y_pred, y_test))
print('Mean Absolute percentage Error: ', mean_absolute_percentage_error(Y_pred, y_test))
#print('Root Mean Squared Error:', root_mean_squared_error(Y_pred, y_test))
print("R2 Score: ", r2_score(Y_pred, y_test))
```

Fig. 3. Shows target variable selection, train test split & modelling

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue VII July 2024- Available at www.ijraset.com

IV. RESULT & DISCUSSION

The system was trained using various machine learning techniques (Random Forest, Linear Support Vector Regression, Decision Tree Regression, and Linear Regression). This resulted in an average R2 score of 99.95%. To enhance the performance, the system underwent retraining with various feature selection techniques. To assess the system, various methods for calculating errors were utilized, including Mean Squared Error, Mean Absolute Error, Mean Absolute Percentage Error, and R-squared score, as detailed in table below:

TABLE I
Table shows comparative analysis of output.

S.No	Description	MSE	MAE	MAPE	R2
Before feature selection					
i	RF Method	7.67	0.82	0.0088	0.998
ii	SVR Method	0.46	0.57	0.219	0.999
iii	DT Method	0.114	0.202	0.226	0.999
iv	LR Method	0.121	0.208	0.217	0.999
After feature selection					
1	Correlation Coefficient Method				
a.	RF Method	4.32	0.74	0.0074	0.999
b.	SVR Method	0.634	0.564	0.225	0.999
c.	DT Method	0.352	0.368	0.233	0.999
d.	LR Method	0.11	0.202	0.208	0.999
2	Mutual Regression method				
2.1	With 13 highest ranked features				
a.	RF Method	3.72	0.66	0.0071	0.999
b.	SVR Method	0.412	0.427	0.233	0.999
c.	DT Method	0.333	0.391	0.217	0.999
d.	LR Method	0.107	0.207	0.233	0.999
2.2	With 12 highest ranked features				
a.	RF Method	8.21	1.02	0.009	0.998
b.	SVR Method	0.321	0.38	0.233	0.99
c.	DT Method	0.38	0.418	0.214	0.99
d.	LR Method	0.105	0.2	0.214	0.999
2.3	With 11 highest ranked features				
a.	RF Method	17.5	1.7	0.01	0.996
b.	SVR Method	480	12.8	1.1	0.89
c.	DT Method	532.1	14.1	0.873	0.884
d.	LR Method	195	9.78	0.42	0.961

The system's performance was evaluated, and it was noted that by using mutual info regression FS method by selecting 11 features with the highest correlation to the target variable Linear Regression method achieved lowest errors. The system error rates were MSE= 0.105, MAE=0.2, MAPE=0.214 and R2 score=0.999.

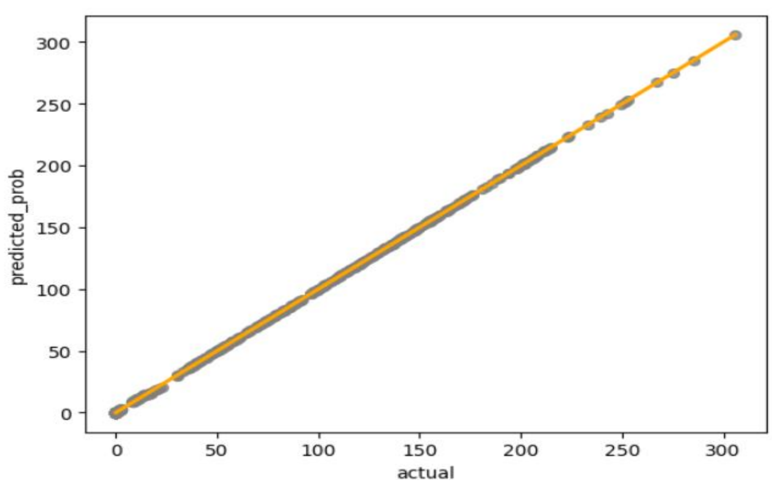


Fig 4: Shows scatter plot between Predicted V/S Actual values

V. CONCLUSIONS

The In this article, we performed machine learning analysis on a dataset sourced from Kaggle Platform by applying different algorithms, including Random Forest, Linear Support Vector Regression, Decision Tree Regression, and Linear Regression. We assessed the system's performance, finding that the Linear Regression Method resulted in the lowest error rates.

Therefore, it can be concluded that the Linear Regression method provides the most accurate level of forecasting.

Lastly, a major goal of these results is to apply this method on a global level, which means merging data from various regions to enhance the adaptability and accessibility of future models.

VI. ACKNOWLEDGMENT

The authors declare that every piece of information provided is totally our own creation and has not been imitated from any other source.

REFERENCES

[1] H. Han, J. Park, S. Kim, S. Lee, M. I. Choi, and S. Park, "Carbon Reduction Method for Intelligent Energy Transformation Based on Energy Data Analysis," in SCEMS 2022 - 2022 IEEE 5th Student Conference on Electric Machines and Systems, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/SCEMS56272.2022.9990807.

[2] T. Ahmad, H. Zhang, and B. Yan, "A review on renewable energy and electricity requirement forecasting models for smart grid and buildings," Sustain. Cities Soc., vol. 55, Apr. 2020, doi: 10.1016/j.scs.2020.102052.

[3] J. J. Daniel Raj, R. Mohan Das, S. Vinod Kumar, M. Jayanthi, A. Sujin Jose, and V. Tejas, "Electricity Demand Forecasting Using ML," in Proceedings - 2023 3rd International Conference on Pervasive Computing and Social Networking, ICPCSN 2023, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 547–551. doi: 10.1109/ICPCSN58827.2023.00095.

[4] M. Abdul Baseer, A. Almunif, I. Alsaduni, and N. Tazeen, "Electrical Power Generation Forecasting from Renewable Energy Systems Using Artificial Intelligence Techniques," Energies, vol. 16, no. 18, Sep. 2023, doi: 10.3390/en16186414.

[5] A. Bhansali, N. Narasimhulu, R. Pérez de Prado, P. B. Divakarachari, and D. L. Narayan, "A Review on Sustainable Energy Sources Using Machine Learning and Deep Learning Models," Energies, vol. 16, no. 17, Sep. 2023, doi: 10.3390/en16176236.

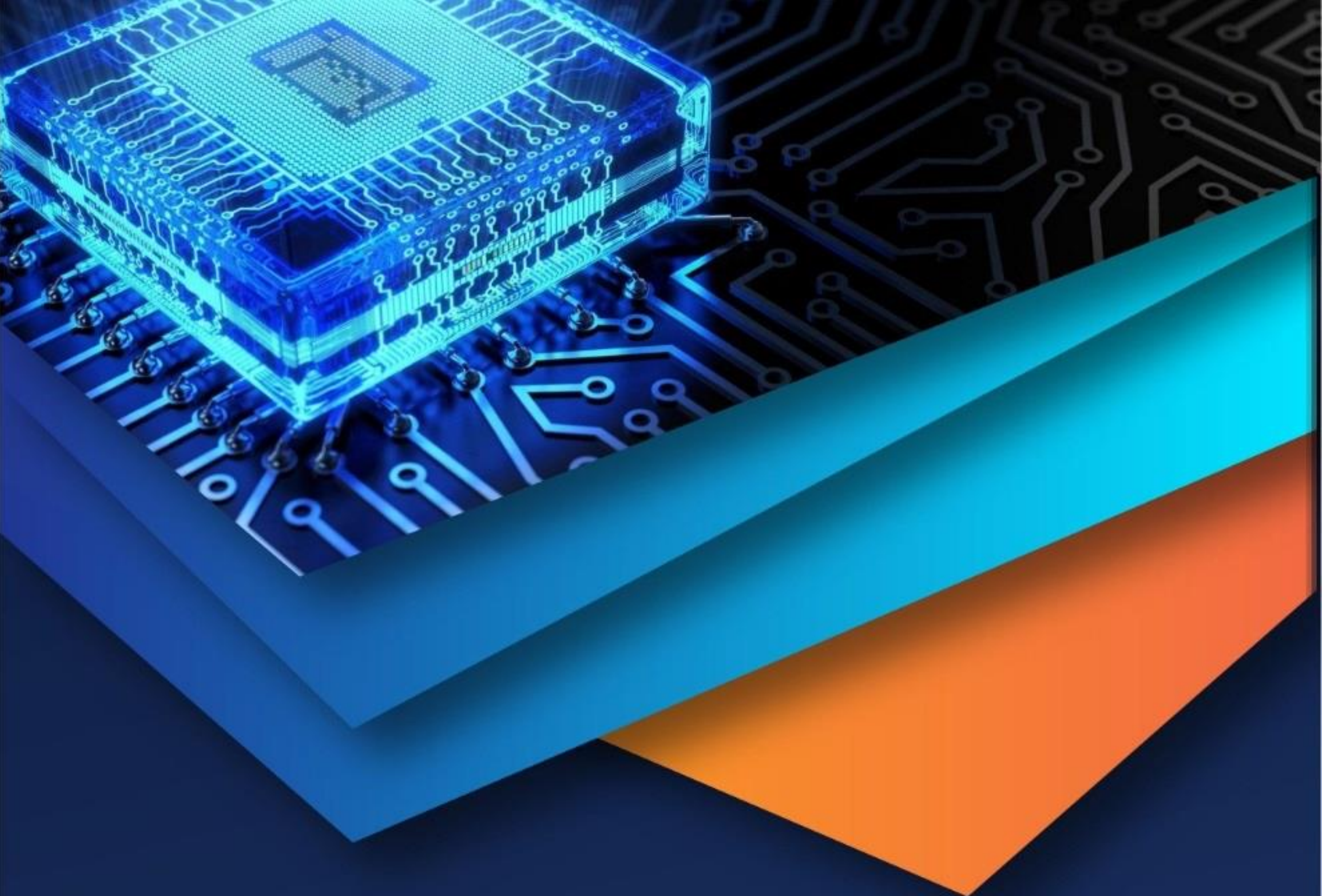
[6] A. Yağmur, M. Kayakuş, and M. Terzioğlu, "Predicting renewable energy production by machine learning methods: The case of Turkey," Environ. Prog. Sustain. Energy, vol. Volume 43, no. Issue 3, 2023, doi: https://doi.org/10.1002/ep.14077.

[7] J. Managre and N. Khatri, "A Review on IoT and ML Enabled Smart Grid for Futurestic and Sustainable Energy Management," in 2022 International Conference for Advancement in Technology, ICONAT 2022, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/ICONAT53423.2022.9725932.

[8] I. Demir, O. Ileri, and O. F. Ertuurul, "The Core of a Smart Grid: Internet of Energy and Machine Learning," in IEEE Global Energy Conference, GEC 2022, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 357–360. doi: 10.1109/GEC55014.2022.9987139.

[9] B. Mallala, A. I. Uddin Ahmed, P. V. Prasad, and P. Kowstubha, "Development of Renewable Energy System for Enhancing Reliability of Power," in Procedia Computer Science, Elsevier B.V., 2023, pp. 1–10. doi: 10.1016/j.procs.2023.12.055.

[10] N. E. Benti, M. D. Chaka, and A. G. Semie, "Forecasting Renewable Energy Generation with Machine Learning and Deep Learning: Current Advances and Future Prospects," Sustain. Switz., vol. 15, no. 9, May 2023, doi: 10.3390/su15097087.



10.22214/IJRASET

INDEX COPERNICUS
45.98

IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL
IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY