Predictive Modeling of Solar Energy Output Using Big Data Analytics in Saudi Arabia

M Giri[1], C Venkatesh[2], A Dhanasekhar Reddy[3]

[1]P.G Scholar, Department of MCA, Sri Venkatesa Perumal College of Engineering & Technology, Puttur, E-mail: mudurugiri@gmail.com , ORC-ID: https://orcid.org/0009-0008-5380-034X

[2] Assistant Professor, Department of CSE(AI & ML), Sri Venkatesa Perumal College of Engineering & Technology, Puttur, E-mail: chevireddyvenkatesh22@gmail.com, ORC-ID: https://orcid.org/0009-0007-1038-5978

[3] Assistant Professor, Department of MCA, Sri Venkatesa Perumal College of Engineering & Technology, Puttur, E-mail: dhanasekhar918@gmail.com, ORC-ID: https://orcid.org/0009-0008-6256-0405

Abstract: The rising energy demand in Saudi Arabia, propelled by swift urbanization and industrial expansion, requires sustainable alternatives to traditional power generation, which substantially contributes to CO_2 emissions. A distributed Big Data analytics system leveraging Apache Spark is applied to predict solar energy production based on several environmental parameters, including wind speed, temperature, and humidity. The dataset, derived from records of the King Abdullah City for Atomic and Renewable Energy and augmented by a Kaggle repository, undergoes comprehensive preprocessing that includes normalization, missing value imputation, and feature selection via the Chi-square approach. Machine learning methods, including Linear Regression, Decision Tree, Random Forest, and Gradient Boosted Tree (GBT), are taught and assessed using performance measures such as RMSE, MAE, and R^2 score. Experimental results indicate that the improved Gradient Boosted Tree model provides exceptional prediction accuracy, obtaining a R^2 value of 99%, signifying excellent reliability in forecasting solar energy output. This analytical methodology improves renewable energy planning and facilitates Saudi Arabia's shift to carbon-neutral power generation via data-driven insights.

*"**Index Terms:** Renewable energy, big data, machine learning, solar energy, photovoltaic cells, Saudi Arabia".*

1. INTRODUCTION

In the last twenty years, renewable energy technologies have significantly progressed, especially in solar and wind power, revolutionizing global electricity generation into a more sustainable and economically feasible system [1]. The rising efficiency and decreasing costs of renewable technologies have established them as important elements in attaining energy transition and environmental sustainability. Saudi Arabia, possessing significant solar potential, has proactively initiated the incorporation of renewable energy into its national energy portfolio to diversify resources and diminish reliance on fossil fuels. In accordance with the national Vision 2030 plan, the Kingdom seeks to utilize renewable energy to satisfy its increasing electrical demand while reducing environmental repercussions. The region, receiving over 3,000 hours of annual sunlight, possesses considerable potential for extensive solar energy generation and optimization.

Notwithstanding these encouraging advancements, Saudi Arabia's electricity generation sector continues to be a significant source of carbon dioxide emissions, intensified by increasing energy demand stemming from population increase and industrial expansion. The variability and intermittency of renewable energy sources, especially solar power, present operational issues for maintaining grid stability and assuring a steady electricity supply. Current research has examined the implementation of renewable energy and hybrid energy systems in the Kingdom; however, it frequently lacks thorough evaluations that incorporate climatic diversity, regional solar potential, and temporal variability across various geographical areas [7]. Moreover, a significant research gap persists in large-scale predictive modeling and analytical frameworks specifically designed for Saudi Arabia's unique environmental circumstances and comprehensive solar dataset availability [8].

This study seeks to tackle these difficulties by examining the prediction of solar energy output at several solar plants in Saudi Arabia. This project aims to improve the comprehension of solar energy variability and its effects on the nation's renewable infrastructure by utilizing comprehensive data from national monitoring stations. The research classifies solar energy facilities according to regional and climatic attributes to facilitate more precise and location-specific evaluations of solar energy production. This method enables a comprehensive analysis of solar potential distribution throughout Saudi Arabia and aids in the enhancement of energy planning, policy development, and resource distribution [9].

This research is significant for its potential to enhance Saudi Arabia's transition to a sustainable and diverse energy system. Precise solar energy forecasting improves grid reliability, diminishes operational uncertainty, and facilitates decision-making for renewable energy projects. This study establishes a data-driven basis for future solar energy projects, thereby supporting the national purpose of significantly reducing greenhouse gas emissions and enhancing the nation's energy sustainability goals [10].

2. LITERATURE REVIEW

Numerous studies have investigated the potential and viability of solar energy implementation in Saudi Arabia, indicating the country's increasing dedication to renewable energy advancement. Shaher et al. [11] evaluated the technical feasibility of rooftop solar photovoltaic (PV) systems in residential and commercial sectors throughout the Kingdom. Their research indicated that rooftop photovoltaic installations might satisfy a significant fraction of urban electricity consumption, especially in cities with elevated sun irradiation. This research provided useful insights on urban solar adoption but mostly concentrated on localized applications, neglecting large-scale forecasts and regional variability in solar energy production.

Subsequent inquiries into the incorporation of solar photovoltaic systems into particular sectors have produced promising outcomes. Bakheet [12] developed and simulated a solar energy system for hospital electrification in Saudi Arabia, demonstrating its potential to decrease emissions and operational expenses. Bouzguenda et al. [13] performed a techno-economic feasibility analysis of photovoltaic systems with energy storage for healthcare facilities, determining that solar PV may fulfill up to 15% of hospital energy requirements at competitive pricing. While these research highlighted sustainability within specific sectors, their constrained geographical and chronological scopes limited the applicability of findings to the wider energy infrastructure. Shervani [14] offered a design for a utility-scale solar farm that could mitigate approximately 1.68 million tons of CO_2 emissions over 25 years, highlighting the enduring environmental advantages of extensive renewable energy implementation. The lack of predictive modeling in extensive installations constrains comprehension of solar production variability and operational optimization.

On a larger scale, Saudi Arabia's national initiatives have showcased the aspiration to excel in renewable innovation. Yusuf and Abdulmohsen [15] examined the NEOM project as a paradigm for economically sustainable urban planning, using solar and wind energy for hydrogen generation. This extensive renewable program underscores the Kingdom's deliberate transition to green energy, however it lacks detailed monitoring of solar production trends across various locations. Conversely, Al-Sharafi et al. [16] assessed rooftop solar PV legislative frameworks and global best practices, providing recommendations for the regulatory landscape in Saudi Arabia. Their policy-focused methodology offered significant governance insights yet failed to tackle the technical difficulties related to solar energy forecasts and grid stability.

The technical integration of renewable energy sources into Saudi Arabia's electricity infrastructure has garnered significant interest. Yasin et al. [17] performed a technical evaluation of grid performance with heightened renewable integration, highlighting the necessity for enhanced forecasting and load management. Boretti [18] investigated dispatchable renewable energy with concentrated solar power (CSP) systems integrated with advanced geothermal storage, revealing the capability for a continuous energy supply. Nevertheless, both studies predominantly concentrated on system-level viability, neglecting data-driven modeling and predictive analysis of solar energy generation.

Notwithstanding considerable progress, substantial obstacles remain. Aldhubaib [19] examined the future of Saudi Arabia's energy sector, pinpointing cost, site selection, and fluctuation as significant obstacles to extensive solar adoption. Mian et al. [20] conducted a further investigation into optimal site selection methods for photovoltaic deployment, emphasizing the significance of sustainability and environmental appropriateness. However, prior studies have inadequately utilized extensive datasets to examine spatial and temporal solar energy trends throughout the Kingdom. This study employs extensive solar resource monitoring data to improve forecast accuracy and regional comprehension of solar production. This research enhances the reliability and efficiency of solar energy forecasting in Saudi Arabia's renewable energy transition by systematically categorizing solar stations and analyzing their varied meteorological circumstances.

3. MATERIALS AND METHODS

The proposed system introduces a distributed Big Data analytics framework based on Apache Spark, designed to predict solar energy production throughout Saudi Arabia utilizing the Solar Resource Monitoring Stations dataset. The framework effectively analyzes extensive environmental characteristics, such as temperature, humidity, wind speed, and Global Horizontal Irradiance (GHI), utilizing Spark's scalable design. Following extensive data preprocessing and feature selection via the Chi-square approach to improve input significance, the system utilizes machine learning methods like Linear Regression, Decision Tree, Random Forest, and Gradient Boosted Tree for predictive modeling. Hyperparameter optimization enhances model performance and minimizes forecasting error, while ensemble learning techniques improve resilience and generalization. Additionally, the use of a Flask-based web interface facilitates real-time user engagement, data submission, and presentation of predictive results. This integrated solution guarantees superior computational efficiency, scalability, and practical accessibility, facilitating data-driven decision-making for sustainable solar energy planning and management.

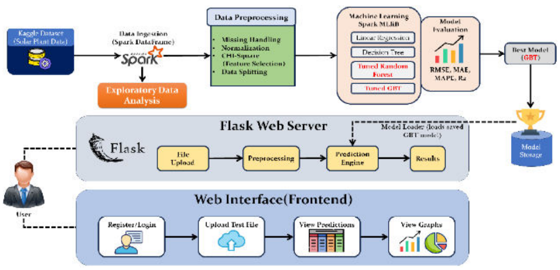


Fig.1 Proposed Architecture

The figure depicts the system architecture of the proposed solar energy forecasting framework. The procedure commences with user engagement via a Flask-based web application, facilitating data upload and result display. The input data undergoes processing in an Apache Spark environment, where Big Data analytics and machine learning models conduct predictive analysis. The processed outputs are archived in a database for future reference, while prediction results are represented graphically, facilitating efficient analysis, scalability, and real-time accessibility for informed energy decision-making.

a) Dataset Collection:

The dataset utilized in this study, Solar electricity Station Data, was sourced from the Kaggle repository and consists of 13,057 samples with 14 variables including environmental and electricity generation parameters. The features comprise temperature, humidity, wind speed, wind direction, barometric pressure, and Global Horizontal Irradiance (GHI), whilst the objective variable denotes solar power generation output. The collection has hourly recordings, documenting temporal fluctuations crucial for forecasting precision. Its varied environmental scope renders it appropriate for assessing and improving predictive models in solar energy analytics.

year	month	day	hour	generation	temperature	humidity	DNI	peakwind	windspeed	winddir	barometric_pressure	DHI	GHI
2022	6	22	0	0.0	12.4	80.83	0.0	41.8	14.4	290.0	1007.9	0	0.0
2022	6	22	1	0.0	12.5	79.23	0.0	43.9	24.1	281.1	1007.0	0	0.0
2022	6	22	2	0.0	12.7	78.19	0.0	49.3	27.0	289.8	1007.0	0	0.0
2022	6	22	3	0.0	10.9	82.85	0.0	49.3	10.8	300.0	1009.8	0	0.0
2022	6	22	4	0.0	12.2	79.72	0.0	50.4	27.4	293.7	1007.0	0	0.0
2022	6	22	5	0.0	11.9	80.76	0.0	51.5	27.0	294.7	1007.0	0	0.0
2022	6	22	6	1.0	11.5	86.35	0.0	49.0	18.4	306.0	1008.3	1	0.0
2022	6	22	7	8.0	12.0	81.32	0.0	47.9	25.2	298.9	1007.0	7	0.0
2022	6	22	8	8.0	12.3	81.91	0.0	50.8	27.4	303.6	1008.0	19	0.1
2022	6	22	9	127.0	13.9	76.01	100.0	54.0	14.4	311.0	1009.4	23	0.1
2022	6	22	10	293.0	14.3	76.37	0.0	53.6	25.9	316.4	1009.0	28	0.1
2022	6	22	11	622.0	16.0	67.1	0.0	50.4	25.6	322.1	1009.0	52	0.2
2022	6	22	12	642.0	18.5	59.08	0.0	56.9	19.6	304.0	1009.4	562	2.0
2022	6	22	13	563.0	18.9	57.77	0.0	56.2	27.0	317.3	1009.0	877	3.2
2022	6	22	14	508.0	20.2	51.54	0.0	50.8	26.6	323.5	1009.0	459	1.7
2022	6	22	15	438.0	21.1	46.97	0.0	52.6	19.6	326.0	1009.5	766	2.8
2022	6	22	16	289.0	20.9	45.53	0.0	50.8	26.3	322.2	1010.0	475	1.7
2022	6	22	17	309.0	20.9	43.42	0.0	48.2	23.8	327.7	1009.0	458	1.6
2022	6	22	18	318.0	22.3	40.14	0.0	45.0	14.4	320.0	1009.8	544	2.0
2022	6	22	19	295.0	20.0	48.78	0.0	41.4	17.6	325.9	1010.0	391	1.4

only showing top 20 rows

Fig.2 Dataset Collection

b) Pre-Processing:

The preprocessing phase guarantees data quality, consistency, and preparedness for machine learning. The process encompasses cleaning, normalization, feature selection, and data partitioning to improve model performance, scalability, and forecasting precision in solar energy prediction.

Data Cleaning and Missing Value Handling: The dataset was cleaned to rectify missing or inconsistent records that could undermine analytical precision. Missing values were addressed using suitable imputation methods to guarantee dataset integrity. This technique eradicated noise, redundancy, and anomalies, guaranteeing that solely pertinent and precise data were utilized for model training. Data cleansing markedly raises input quality, diminishes model bias, and bolsters the overall dependability of predictive results.

Data Normalization: Normalization was utilized to standardize environmental factors, including temperature, humidity, and wind speed, to a consistent range. This technique reduces the influence of attributes with greater numerical ranges, guaranteeing equitable contribution from all features during model training. Normalized data enhances convergence velocity and model stability for regression algorithms. This step improves the performance, accuracy, and interpretability of machine learning models in the Spark environment by normalizing input magnitudes.

Feature Selection: The Chi-Square feature selection method was utilized to determine the most statistically significant factors of solar energy production. This method assesses the relationship between input features and the target variable, identifying the five traits having the most significance. Feature selection diminishes dimensionality, decreases computational burden, and mitigates overfitting. Thus, it improves the learning efficiency and predictive accuracy of machine learning models utilized in the suggested system.

Data Visualization: A thorough data visualization was performed to analyze patterns and correlations among environmental variables influencing solar power generation. Correlation analysis was utilized to identify feature interdependencies and exclude highly associated variables that could compromise model learning. Graphical representations, including GHI distribution, temperature fluctuations, and wind speed patterns, facilitated a more profound comprehension of data dynamics. This phase enhanced interpretability, facilitating feature engineering and augmenting the explainability of the predictive model.

c) Training and Testing:

The preprocessed dataset was divided into training and testing subsets to enable model training and performance assessment. Generally, 80% of the data was allocated for training and 20% for testing, facilitating equitable learning and impartial evaluation. This division enables the algorithms to generalize proficiently on unfamiliar data, assisting in the validation of their predicted resilience. The procedure guarantees dependable model assessment by metrics including RMSE, MAE, and R^2 score.

d) Algorithms:

Linear Regression: Linear Regression is a supervised machine learning approach intended to forecast continuous outcomes by defining a linear connection between independent and dependent variables. It identifies the ideal regression line that minimizes prediction error, therefore quantifying the influence of each factor on the target variable. This technique enhances system performance by offering a straightforward, interpretable, and computationally efficient baseline for forecasting, allowing researchers to examine correlations and discern critical environmental elements affecting solar energy production.

Decision Tree: The Decision Tree technique employs a hierarchical, rule-based framework to execute classification and regression tasks by recursively partitioning data into subsets according to feature criteria. Each branch signifies a decision rule that culminates in a predicted conclusion, providing interpretability and adaptability for both categorical and continuous data. It improves prediction accuracy by identifying complicated, non-linear interdependencies among variables, facilitating the effective modeling of intricate interactions in renewable energy forecasting and decision-making contexts.

$$I\ (i) = 1 - \sum_{i=1}^{k} p_i^2 \qquad (1)$$

Tuned Random Forest: Tuned Random Forest is an ensemble learning method that builds numerous decision trees and consolidates their predictions to enhance model robustness and generalization. Parameter tuning, including the optimization of the number of estimators, depth, and feature selection, reduces variance and overfitting while ensuring predictive stability. This approach improves performance by utilizing feature variety and averaging several models, resulting in reduced prediction errors and increased accuracy in energy forecasting and environmental data analysis.

$$Gini = 1 - \sum_{i=1}^{C} (Pi)^2 \qquad (2)$$

Tuned Gradient Boosting: Uned Gradient Boosting is a sophisticated ensemble technique that constructs predictive models in a sequential manner, with each tree addressing the residuals of its predecessors. By optimizing hyperparameters like learning rate, number of estimators, and maximum depth, it attains enhanced accuracy and expedited convergence. The technique improves system performance by accurately modeling complicated, non-linear interactions in extensive data, thus enhancing prediction accuracy, reducing errors, and facilitating superior decision-making in renewable energy forecasting applications.

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) \qquad (3)$$

e) Spark API and Flask Framework:

The Apache Spark API functions as the foundation of the proposed Big Data analytical system, facilitating distributed and parallel analysis of extensive solar energy information. Spark's in-memory compute abilities markedly enhance data processing, model training, and evaluation workloads across numerous nodes. Utilizing Spark MLlib, diverse machine learning techniques including Linear Regression, Decision Tree, Random Forest, and Gradient Boosted Trees are effectively employed to predict solar power generation with enhanced scalability and precision.

The Flask framework enhances this design by offering an interactive web interface for real-time prediction and visualization. It enables users to submit test data, activate learned models, and promptly observe expected energy results. Flask connects analytical computation with user accessibility, providing a seamless, scalable, and practical deployment environment for renewable energy forecasting applications.

4. EXPERIMENTAL RESULTS

RMSE: The root mean square error (RMSE) quantifies the average discrepancy between a statistical model's predicted values and the observed values. It is the standard deviation of the residuals, mathematically. Residuals signify the deviation between the regression line and the data points.

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

MAE: Absolute mistake refers to the magnitude of mistake in measurements. It is the discrepancy between the measured value and the "true" value. If a scale indicates 90 pounds while your actual weight is 89 pounds, the absolute inaccuracy of the scale is calculated as 90 lbs – 89 lbs = 1 lb.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad (5)$$

R2 Score: The sum of squared regression is the aggregate of the squared residuals, whereas the total sum of squares represents the sum of the squared distances of the data from the mean.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \qquad (6)$$

Table.1 Performance Evaluation Table

Algorithm Name	RMSE	MAE	R2Score
Linear Regression	78.974022	58.117344	0.877895
Decision Tree	89.705544	64.782807	0.842456
Extension Tuned Random Forest	33.342067	19.444256	0.978236
Extension Tuned GBT	**12.599621**	**6.356185**	**0.996892**

Table 1 presents a performance evaluation of models utilizing RMSE, MAE, and R² Score. The Extension Tuned GBT model surpasses all others, attaining the minimal errors and the best accuracy score.

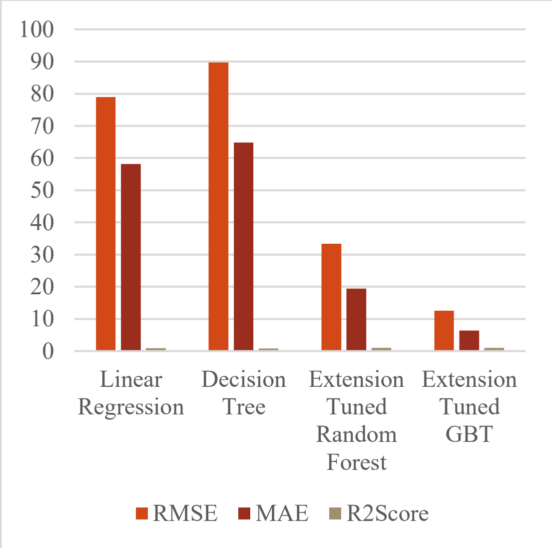


Fig.3 Comparison Graph

Figure 3 presents a comparative graph that visually depicts model performance in terms of RMSE, MAE, and R² Score. Extension Tuned GBT has enhanced accuracy with little mistakes, surpassing all other models in prediction efficacy and dependability.

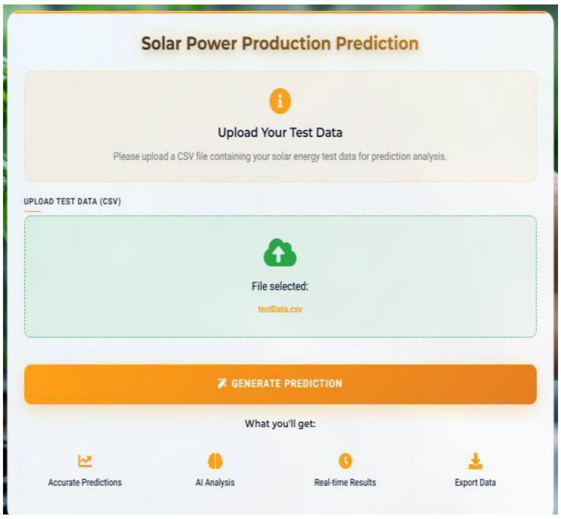
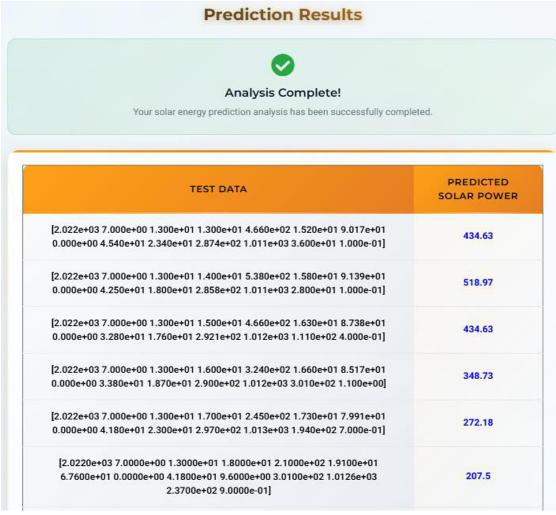


Fig.4 Upload Test Data


Prediction Results

Analysis Complete!

Your solar energy prediction analysis has been successfully completed.

TEST DATA	PREDICTED SOLAR POWER
[2.022e+03 7.000e+00 1.300e+01 1.300e+01 4.660e+02 1.520e+01 9.017e+01 0.000e+00 4.540e+01 2.340e+01 2.874e+02 1.011e+03 3.600e+01 1.000e-01]	434.63
[2.022e+03 7.000e+00 1.300e+01 1.400e+01 5.380e+02 1.580e+01 9.139e+01 0.000e+00 4.250e+01 1.800e+01 2.858e+02 1.011e+03 2.800e+01 1.000e-01]	518.97
[2.022e+03 7.000e+00 1.300e+01 1.500e+01 4.660e+02 1.630e+01 8.738e+01 0.000e+00 3.280e+01 1.760e+01 2.921e+02 1.012e+03 1.110e+02 4.000e-01]	434.63
[2.022e+03 7.000e+00 1.300e+01 1.600e+01 3.240e+02 1.660e+01 8.517e+01 0.000e+00 3.380e+01 1.870e+01 2.900e+02 1.012e+03 3.010e+02 1.100e+00]	348.73
[2.022e+03 7.000e+00 1.300e+01 1.700e+01 2.450e+02 1.730e+01 7.991e+01 0.000e+00 4.180e+01 2.300e+01 2.970e+02 1.013e+03 1.940e+02 7.000e-01]	272.18
[2.0220e+03 7.0000e+00 1.3000e+01 1.8000e+01 2.1000e+02 1.9100e+01 6.7600e+01 0.0000e+00 4.1800e+01 9.6000e+00 3.0100e+02 1.0126e+03 2.3700e+02 9.0000e-01]	207.5

Fig.5 Predicted Results

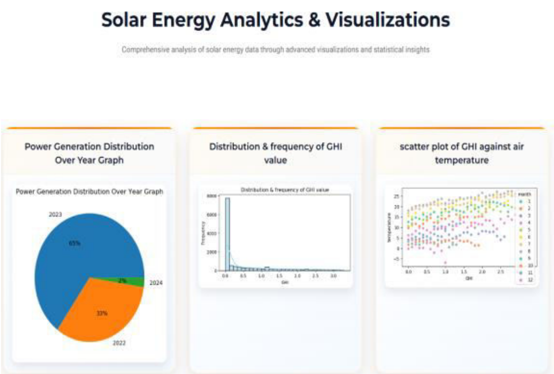


Fig.6 Analytics & Visualizations

5. CONCLUSION

This work effectively illustrates the capacity of Apache Spark-based Big Data analytics to predict solar energy generation by utilizing extensive environmental datasets and sophisticated machine learning algorithms. The dataset was optimized for efficient model training by rigorous preprocessing procedures, including normalization, missing value management, and Chi-square feature selection. A comparative analysis of various regression algorithms demonstrated that the optimized Gradient Boosted Tree model attained the highest predictive performance, achieving a R^2 score of 99%, indicating remarkable accuracy in modeling the relationship between environmental parameters and solar power output. This great precision underscores the model's resilience in managing intricate, real-world data amidst fluctuating weather conditions. The decentralized architecture of Apache Spark markedly enhanced data processing and computation, facilitating scalable and efficient analysis appropriate for extensive energy datasets. The incorporation of refined learning methodologies illustrates the capabilities of Big Data frameworks in improving renewable energy forecasting systems. These findings enhance decision-making in energy management, facilitate effective resource allocation, and bolster Saudi Arabia's long-term goal of transitioning to sustainable, low-carbon energy solutions using intelligent, data-driven predictive models.

Future developments may concentrate on incorporating real-time environmental data streams using IoT-enabled sensors to improve the forecasting model's responsiveness and adaptability. Integrating deep learning architectures like Long Short-Term Memory (LSTM) or Convolutional Neural Networks (CNN) can enhance forecast accuracy by capturing temporal and spatial relationships in solar energy patterns. Incorporating geographical segmentation and satellite imaging into the dataset will facilitate regional energy optimization. Moreover, adopting cloud-based deployment for remote training and creating an interactive analytics dashboard can enhance real-time monitoring and decision-making. These enhancements will augment model scalability, predictive capability, and practical utility in extensive renewable energy management systems.

REFERENCES

[1] Kim, E., & Hau, Y. S. (2025). Advancing photovoltaic solar power forecasting by the hybrid

model of CNN and DNN and the XAI with SHAP. International Journal of Green Energy, 1-12.

[2] Nadeem, T. B., Ali, S. U., Asif, M., & Suberi, H. K. (2024). Forecasting daily solar radiation: An evaluation and comparison of machine learning algorithms. AIP Advances, 14(7).

[3] Baloch, M., Honnurvali, M. S., Kabbani, A., Jumani, T. A., & Chauhdary, S. T. (2024). An intelligent SARIMAX-based machine learning framework for long-term solar irradiance forecasting at Muscat, Oman. Energies, 17(23), 6118.

[4] Rashid, L. B., Shuja, S. Z., & Rehman, S. (2025). Machine Learning Forecasting of Direct Solar Radiation: A Multi-Model Evaluation with Trigonometric Cyclical Encoding. Forecasting, 7(4), 58.

[5] Tandon, A., Awasthi, A., Pattnayak, K. C., Tandon, A., Choudhury, T., & Kotecha, K. (2025). Machine learning-driven solar irradiance prediction: advancing renewable energy in Rajasthan. Discover Applied Sciences, 7(2), 107.

[6] F. S. Al-Ismail, M. S. Alam, M. Shafiullah, M. I. Hossain, and S. M. Rahman, ''Impacts of renewable energy generation on greenhouse gas emissions in Saudi arabia: A comprehensive review,'' Sustainability, vol. 15, no. 6, p. 5069, Mar. 2023, doi: 10.3390/su15065069.

[7] Y. H. A. Amran, Y. H. M. Amran, R. Alyousef, and H. Alabduljabbar, ''Renewable and sustainable energy production in Saudi Arabia according to Saudi vision 2030; current status and future prospects,'' J. Cleaner Prod., vol. 247, Feb. 2020, Art. no. 119602, doi: 10.1016/j.jclepro.2019. 119602.

[8] A. Elshurafa, F. Felder, and N. Alhaidari, ''Achieving renewable energy targets without compromising the power sectors reliability,'' King AbdullahPetroleumStudiesRes.Center,Riyadh,Saud iArabia,Tech.Rep. KS-2021-DP023-ARA, Mar. 2022, doi: 10.30573/ks–2021-dp23.

[9] Swapna, G., Sreenivasulu, K., Deepika, M., Baseer, K. K., Neerugatti, V., & Viswanath, G. (2025). Brain tumour detection using MRI images in CNN. Advances in Science, Engineering and Technology, 6.

[10] O. S. Alzaid, B. Salim, J. Orfi, S. Khan, and H. Alshehri, ''Hybrid solar andwindpowergenerationinSaudiArabia,''EnergyE nviron.Res.,vol.10, no. 2, p. 25, Nov. 2020, doi: 10.5539/eer.v10n2p25.

[11] A. Shaher, S. Alqahtani, A. Garada, and L. Cipcigan, ''Technical potential for rooftop solar photovoltaic in commercial and residential areas in Saudi Arabia,'' in Proc. 57th Int. Universities Power Eng. Conf. (UPEC), Aug. 2022, pp. 1–6, doi: 10.1109/UPEC55022.2022.9917795.

[12] M. A. Bakheet, ''Design and modelling of solar energy system for electrification for a hospital in Saudi Arabia,'' Eur. J. Energy Res., vol. 3, no. 2, pp. 1–12, Jul. 2023, doi: 10.24018/ejenergy.2023.3.2.106.

[13] M. Bouzguenda, A. Almulhim, A. Shawki, A. Al-Baadani, I. Ahmed, M.Alkulaib, and M. Al-Aqil, ''Technical and economic feasibility of solar pv systems supported by energy storage in hospitals in KSA,'' Mater. Res. Proc., vol. 2023, p. 31, Aug. 2023, doi: 10.21741/9781644902592-55.

[14] S. Shervani, ''Design of a utility scale solar farm in Saudi Ara bia,'' Trends Tech. Sci. Res., vol.

4, no. 3, p. 14, 2020, doi: 10.19080/ttsr.2020.04.555639.

[15] N. Yusuf and D. Abdulmohsen, ''Saudi Arabia's NEOM project as a testing ground for economically feasible planned cities: Case study,'' Sustainability, vol. 15, no. 1, p. 608, Dec. 2022, doi: 10.3390/su15010608.

[16] Viswanath, G., Prasad, K. K., Lakshmi, J. M., & Swapna, G. (2025). Diabetes Diagnosis Using Machine Learning with Cloud Security. Cuestiones De Fisioterapia, 54(2), 417-431. https://doi.org/10.48047/r2mhn978

[17] J. Yasin, A. Ali, M. T. Hussain, M. Al-Hajji, M. Farhan, H. Qazi, S. Temtem, S. Stepanecsu, G. Hardeep, F. Georges, and S. Minić, ''Technical assessment of KSA power system with large participation of renewables,'' in Proc. Saudi Arabia Smart Grid (SASG), Dec. 2022, pp. 1–9, doi: 10.1109/sasg57022.2022.10201226.

[18] A. Boretti, ''Dispatchable renewable energy from CSP and CSP+EGS in the kingdom of Saudi Arabia,'' IOP Conf. Ser., Earth Environ. Sci., vol. 813, no. 1, Jul. 2021, Art. no. 012003, doi: 10.1088/1755 1315/813/1/012003.

[19] H. A. Aldhubaib, ''Electrical energy future of Saudi Arabia: Challenges and opportunities,'' Frontiers Energy Res., vol. 10, pp. 2–4, Dec. 2022, doi: 10.3389/fenrg.2022.1005081.

[20] S. H. Mian, K. Moiduddin, H. Alkhalefah, M. H. Abidi, F. Ahmed, and F. H. Hashmi, ''Mechanisms for choosing PV locations that allow for the most sustainable usage of solar energy,'' Sustainability, vol. 15, no. 4, p. 3284, doi: 10.3390/su15043284.

[21] A.-N. Buturache and S. Stancu, ''The use of artificial neural networks and big data infrastructure for predictive analytics in solar energy,'' in Proc. Int. Conf. Bus. Excellence, 2021, vol. 15, no. 1, pp. 292–301, doi: 10.2478/picbe-2021-0028.

[22] D. Abdulai, S. Gyamfi, F. A. Diawuo, and P. Acheampong, ''Dataanalytics for prediction of solar PV power generation and system performance: A real case of bui solar generating station, Ghana,'' Sci. Afr., vol. 21, Sep. 2023, Art. no. e01894, doi: 10.1016/j.sciaf.2023.e01894.

[23] R. Ahmed, A. Das Gupta, R. M. Krishnamurthy, M. Goyal, K. S. Kumar, and D. Gangodkar, ''The role of smart grid data analytics in enhancing the paradigm of energy management for sustainable develop ment,'' in Proc. 2nd Int. Conf. Advance Comput. Innov. Technol. Eng. (ICACITE), Apr. 2022, pp. 198–201, doi: 10.1109/ICACITE53722.2022. 9823542.

[24] M. Sharma and R. Sikka, ''An impact on structural model upon renewable energy data science and continuing to improve the energy conservation mindset for environmental sustainability,'' in Proc. 11th Int. Conf. Syst. Model. Advancement Res. Trends (SMART), Aug. 2022, pp. 1098–1103, doi: 10.1109/smart55829.2022.10046940.

[25] Viswanath G., Krishna Prasad K ., Dr. J Maha Lakshmi., Dr.G.Swapna (2024). Health Prediction Using Machine Learning with Drive HQ Cloud Security. Frontiers in HealthInformatics, 13(8), 2755-2761, https://doi.org/10.5281/zenodo.19128870

[26] M. Bowman. (Oct. 2, 2019). EIA Projects That Renewables Will Provide Nearly Half of World Electricity By 2050—Today in Energy U.S. Energy Information Administration (EIA). [Online].

Available:
https://www.eia.gov/todayinenergy/detail.php?id=41533

[27] National Transformation Program. Accessed: Mar. 24, 2025. [Online]. Available: https://www.vision2030.gov.sa/en/explore/projects/neom

[28] The Future of Energy. Accessed: Nov. 6, 2024. [Online]. Available: https://www.neom.com/en-us/our-business/sectors/energy.

[29] Global Solar Atlas. Accessed: Nov. 6, 2024. [Online]. Available: https://globalsolaratlas.info/download/saudi-arabia

[30] Lakshmi, J. M., Prasad, K. K., & Viswanath, G. (2025). Proactive Security in Multi-Cloud Environments: A Blockchain Integrated Real-Time Anomaly Detection and Mitigation Framework. Cuestiones De Fisioterapia, 54(2), 392-417.