Energy Proceedings
Vol 48, 2024
ISSN 2004-2965

Ensemble-Tree Model Based on Bayesian Optimization for Solar Energy Generation Prediction in Smart Homes#

Oluwatoyosi Bamisile[1], Dongsheng Cai[1,2,*], Chukwuebuka J. Ejiyi[5], Chiagoziem C. Ukwuoma[1,2,3], Qi Huang[1,2], Olusola Bamisile[1,2,4]

1 College of Nuclear Technology and Automation Engineering, Chengdu University of Technology, Sichuan P.R., 610059, China.

2 Sichuan Engineering Technology Research Centre for Industrial Internet Intelligent Monitoring and Application, Chengdu University of Technology, Sichuan P.R., 610059, China.

3 OBU, Sino-British Collaborative Education, Chengdu University of Technology, Sichuan P.R., 610059, China

4 Centre for Environmental Policy, Imperial College London, United Kingdom.

5 School of Information and Software Engineering, University of Electronic Science and Technology of China

Chengdu, China

(Corresponding Author: caidongsheng@cdut.edu.cn)

ABSTRACT

Smart homes use devices that automate tasks like security, lighting, and temperature control. These homes let people control appliances remotely through the Internet of Things (IoT), adjusting to their schedules for better energy use. But as energy use rises, it causes more pollution, and climate problems, and puts more strain on energy sources. Therefore, it's important to track energy use closely as the world moves into the use of renewable energy to avoid power outages, save money, and protect the environment especially because of the intermittent nature of renewable energy. This paper proposed an Ensemble-Tree Model Based on Bayesian Optimization for Solar Energy Generation Prediction in Smart Homes. First, three tree-like machine learning models training hyperparameters were optimized using the Bayesian optimization technique. Secondly, their output was concatenated based on Mean Aggregation Methods in mathematics. Lastly, the prediction was done based on k-fold cross-validation. The 'smart-home-dataset-with-weather-information dataset is used while using the R-squared (R^2), Mean Squared Error (MSE), and Mean Absolute Error (MAE) to observe how accurate the predictions are. Results show that the proposed model outperforms other machine learning models with R^2 value of 0.988 as compared in this paper.

Keywords: Solar Generation, Smart Homes, Internet of Things, Machine Learning, Prediction

NONMENCLATURE

Abbreviations	*Meaning*
RF	Random forest
DT	Decision Tree
KNN	K-Nearest Neighbors
MAE	Mean Absolute Error
MSE	Mean Squared Error
LR	Linear Regression
LGBM	Light Gradient Boosting Machine
ADABOOST	Adaptive Boosting
Symbols	
R^2	R-squared
CO^2	Carbon Dioxide

1. INTRODUCTION

Electricity consumption is one of the leading factors to the surge in global energy consumption that has been linked to an increase in carbon dioxide (CO_2) emission and adds to changes in climate. As the world unites in the pursuit of clean energy to mitigate global temperature rise, renewable sources have emerged as a promising alternative to fossil fuels for electricity generation, owing to their pollution-free and waste-free nature [1]. To embrace this shift, numerous industries and sectors have considered the installation of wind turbines, solar panels, and other renewable energy resources. Due to this increase, especially in solar power, the need for accurate and reliable solar prediction becomes sacrosanct in smart homes to allow proper electricity management to avoid light outages because all the gadgets in smart homes are solely dependent on electricity [2]. In this context, the application of Artificial Intelligence (AI) in the form of smart homes has emerged as a leading initiative, leveraging information and communication

This is a paper for the 10th Applied Energy Symposium: Low Carbon Cities & Urban Energy Systems (CUE2024), May. 11-13, 2024, Shenzhen, China.

technologies to enhance residents' convenience, security, and comfort [3].

Smart homes, which are technologically advanced domiciles, are entirely managed and monitored through internet-connected devices, offering numerous benefits such as enhanced efficiency, security, privacy, and energy management [4]. The seamless interaction between humans and computers, facilitated by innovative technologies like the Internet of Things (IoT) [5], plays a central role in the functionality of smart homes. The IoT framework enables the networking of various appliances and devices within smart homes, empowering communication and automation through sensors and actuators [6]. For example, sensors can diligently monitor and regulate water and electricity consumption, granting residents remote control capabilities via applications installed on their mobile devices or laptops. By harnessing the potential of IoT-based devices and leveraging the power of machine learning algorithms, smart homes can predict and optimize solar energy generation, which is a critical component of sustainable energy management [7].

Machine learning algorithms have made significant strides in recent years, equipping smart homes with the ability to analyze extensive weather datasets and forecast weather patterns with remarkable accuracy [8]. This predictive capability optimizes solar energy generation within smart home ecosystems, enabling more efficient energy management and cost-effective solutions. The precision of solar power prediction within smart homes relies heavily on machine learning algorithms [9]. This precision holds immense significance in ensuring the dependable, economical, and secure operation of electrical energy systems. By accurately forecasting solar power generation, homeowners can make well-informed decisions regarding their energy consumption, thereby fostering a more sustainable and environmentally conscious lifestyle.

Ensemble learning and hyperparameter optimization techniques are key techniques in machine learning for improving predictive performance. Ensemble learning uses the collective prediction of multiple models to enhance accuracy, robustness, and generalization [10]. Hyperparameter optimization techniques uses iterative selection of hyperparameters, optimize them efficiently. When combined, these methods unlock unprecedented potential for model performance, enhancing predictive accuracy and adaptability across diverse datasets and problem domains[11].

Sequel to the above, this paper proposed an ensemble-tree model based on Bayesian optimization for solar energy generation prediction in smart homes. Unlike the traditional methods, this model harnesses the power of ensemble learning and hyperparameter optimization techniques to achieve superior accuracy and reliability in predicting solar energy generation. The objective is to assess the generalization of these models in optimizing solar energy generation within smart home environments using weather variables by comparing key metrics such as R-squared, Mean Squared Error (MSE), and Mean Absolute Error (MAE). The findings of this study aim to show the strength of ensemble models and hyperparameter optimization techniques for accurately predicting solar energy generation using Machine learning approach. This, in turn, facilitates enhanced energy management, cost-effectiveness, and environmental sustainability within smart homes.

This paper is organized as follows: Section 2 reviews previous work, Section 3 describes the proposed model, machine learning algorithm compared, dataset and data preprocessing, evaluation metrics and the implementation setup. Section 4 covers the results and discussions while Section 5 summarizes the main findings of the paper.

2. RELATED WORKS

Predicting solar energy generation in smart homes using machine learning algorithms represents a critical intersection of renewable energy research and advanced computational techniques. However, to the best of my knowledge few researchers have ventured into this research area. Researchers like Huertas-Tato and Brito [12] explored the efficacy of smart persistence and random forests in predicting photovoltaic energy production. Their work emphasizes the versatility of machine learning approaches in handling the variability and unpredictability inherent in solar energy generation, thus enhancing the reliability of predictions. Al-Dahidi et al. [13] demonstrated the effectiveness of extreme learning machines in predicting solar photovoltaic power, these finding is particularly relevant for smart homes, where accurate energy predictions are crucial for optimizing energy management and storage systems. It is worth to note that, the integration of machine learning algorithms into smart home energy systems illustrates the potential for these technologies to improve energy efficiency and reduce costs [14]. M. Kutseva,[15] not only highlights the practical applications of machine learning in managing and forecasting energy in smart homes but also underscores the importance of real-time data processing and adaptive learning algorithms in achieving these goals. Moreover, the work by Dinh Van Tai [16]on predicting the output power of solar photovoltaic panels using machine learning approaches further demonstrates the utility of these algorithms in optimizing solar energy systems under varying environmental conditions.

3. MATERIAL AND METHODS

This section discusses in details the proposed method, the dataset used, the evaluation metric, and the implementation details.

3.1 Proposed method

The proposed model is an ensemble of the Random Forest (RF), Decision Tree (DT), and K-Nearest Neighbors (KNN). Ensemble-based models propose that combining multiple algorithms reduces the likelihood of errors compared to relying on a single algorithm. Implementing an ensemble model improves overall prediction performance beyond that of the baseline models within the ensemble. This is attributed to the diverse nature of the classifiers working together which tend to predict accurately compared to using non-diverse classifiers alone. We denoted their optimal hyperparameters as θ^*_{RF}, θ^*_{DT}, and θ^*_{KNN} which were obtained through Bayesian optimization. The prediction of the individual model is expressed as a function of their hyperparameters respectively:

$$\hat{y}_{RF} = RF(\theta^*_{RF,}X) \quad (1)$$

$$\hat{y}_{DT} = DT(\theta^*_{DT}, X) \quad (2)$$

$$\hat{y}_{KNN} = KNN(\theta^*_{KNN}, X) \quad (3)$$

The weight is denoted as $w_{RF}, w_{DT,} and\ w_{KNN}$ respectively. The ensemble prediction is represented mathematically in Equation (4).

$$\hat{y}_{Ensemble} = w_{RF} \cdot \hat{y}_{RF} + w_{DT} \cdot \hat{y}_{DT} + w_{KNN} \cdot \hat{y}_{KNN} \quad (4)$$

Constraint is added to ensure the total weight is up to 1 since they represent proportion, which ensures that the prediction is a convex combination of the individual model predictions as seen in Equation (5).

$$w_{RF} + w_{DT} + w_{KNN} = 1 \quad (5)$$

Each weight w_i are determined based on the performance of the corresponding model.

3.2 Compared baseline models

Four Machine learning regression algorithms were analyzed in this paper namely Linear Regression, Light Gradient Boosting Machine, Random Forest and Adaptive Boosting. The LightGBM Regressor uses tree-based ensemble, in which several weak learners, often decision trees, are merged to create a powerful predictive model.

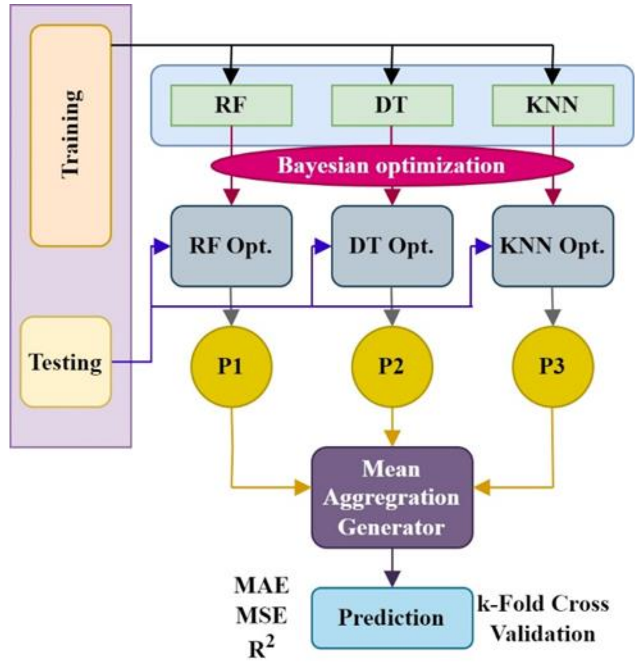

Fig. 1. Proposed methodology

The algorithm aims to minimize a loss function that is differentiable by optimizing the weights of the weak learners and the final prediction is the sum of predictions from all the trees in the ensemble. LightGBM regression model is mathematically explained as thus;

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i) \quad (6)$$

Where, $\hat{y}_i$ is the predicted value for the ith data point, $f_k(x_i)$ is the kth tree's prediction for the ith data point, K is the total number of trees in the LightGBM model, x_i represents the features of the ith data point.

Random Forest constructs several decision trees using bagging and random feature selection. Every tree is trained on a portion of the data and considers a random selection of characteristics at each division. In regression problems, the ensemble prediction is obtained by calculating the average of the predictions made by each individual tree as seen in Equation (7).

$$\hat{y} = \frac{1}{n} \sum_{i=1}^{n} T_i(x) \quad (7)$$

Where, $\hat{y}$ is the predicted value of the input data point, $T_i(x)$ denotes the prediction of the ith decision tree for the input data point x, n is the total number of decision trees in the Random Forest ensemble.

AdaBoost combines weak learners to create a powerful classifier. The AdaBoost algorithm adds weights to the training instances, giving more importance to the ones that are misclassified, in order to train weak learners in a sequential manner and enhance the overall accuracy.

$$\hat{y} = (\sum_{t=1}^{T} \alpha_t \, . \, h_t(x)) \tag{8}$$

Where $\hat{y}$ is the final prediction for a given input x, T is the total number of weak learners, α_t is the weight assigned to the t^{th} weak learner, $h_t(x))$ is the prediction made by the t^{th} weak learner for the input x.

Linear regression is used for modeling the relationship between a dependent variable and one or more independent variables. The changes in independent variables are related to the relationship of dependent variables in the form of a straight line because their relationship is assumed to be linear.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \epsilon \tag{9}$$

Where: Y is the dependent variable, $X_1, X_2, \dots, X_p$ are the independent variables, β_0 is the intercept, $\beta_1, \beta_2, \dots, \beta_p$ are the coefficient of the independent variables, ϵ is the error term.

3.3 Dataset and data preprocessing

The project utilized a dataset called the 'Smart Home Dataset with Weather Information' from Kaggle [17], which contains readings from household appliances over 350 days, recorded in kilowatts (kW) and accompanied by corresponding weather conditions. The dataset contains 503,910 data points, each with 32 features. The dataset was divided into training and testing sets, with 80% allocated for training and 20% for testing.

In the preprocessing stage of the paper, all non-value elements were eliminated. Object types were transformed into integers utilizing label encoding. Feature selection was conducted using Elastic Net, followed by grid search optimization.

3.4 Evaluation metrics

Three evaluation metrics was considered in this experiment namely the Mean Absolute Error (MAE), R-squared (R^2) and Mean Squared Error (MSE). The MAE quantifies the average absolute differences between the predicted and actual values, providing a measure of the average magnitude of errors.

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{10}$$

Where n is the number of data points, y_i is the actual value and $\hat{y}_i$ is the predicted value. R^2 quantifies the extent to which the variability in the dependent variable can be explained by the independent variables. It is expressed as a value between 0 and 1.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{11}$$

Where n is the number of data points, y_i is the actual value, $\hat{y}_i$ is the predicted value and $\bar{y}$ is the mean of the actual values. MSE is a metric that calculates the average of the squared differences between predicted and actual values, providing a measure of the average squared error.

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{12}$$

3.5 Implementation Details

The Implementation step of the deployed models include loading the dataset. Converting of some columns into float data type using label encoding as well as imputing missing values using the mean strategy. The features and target variables are defined based on the dataset, where the predictors (X) consist of all columns except "total," "Solar," and "House overall," and the target variable is "Solar." The Elastic Net is applied to the predictors (X) and the target variable to identify important features based on their coefficients. Features with non-zero coefficients are selected (top features with the highest scores). For the proposed model, the Bayesian optimization is used for hyperparameter selection of the individual models before the concatenation. Also, the k =5-fold cross validation is employed. GridSearchCV is employed in all the machine learning models to perform hyperparameter tuning using cross-validation (cv=5) on the training set.

4. RESULTS

This section focuses on the predicted results of the deployed models, First, the correlation heatmap showing the correlation coefficients between different variables of the dataset is presented in Fig. 2. It helps to identify patterns and relationships between variables in the dataset. The color intensity represents the strength and direction of the correlation: positive correlations are indicated by warmer colors (e.g., red), negative correlations by cooler colors (e.g., blue), and weaker correlations by colors closer to neutral (e.g., white).

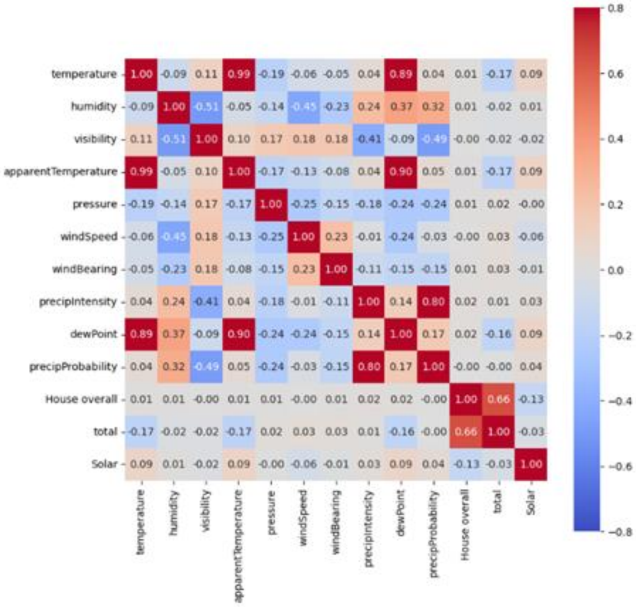

Fig. 2 Correlation heatmap

4.1 Proposed model result

The proposed model result is shown in Table 2 and Fig 3 with K=5 cross-validation. The result shows that the proposed model performed consistently well across different folds, as indicated by the low RMSE values, which range from 0.013 to 0.014.

Table 2. Proposed Model Result

K-Fold Validation	RMSE	R^2	MSE	MAE
1	0.013	0.988	0.001	0.004
2	0.014	0.986	0.001	0.005
3	0.013	0.988	0.001	0.004
4	0.013	0.989	0.000	0.004
5	0.013	0.988	0.001	0.004
Average	**0.013**	**0.988**	**0.001**	**0.004**

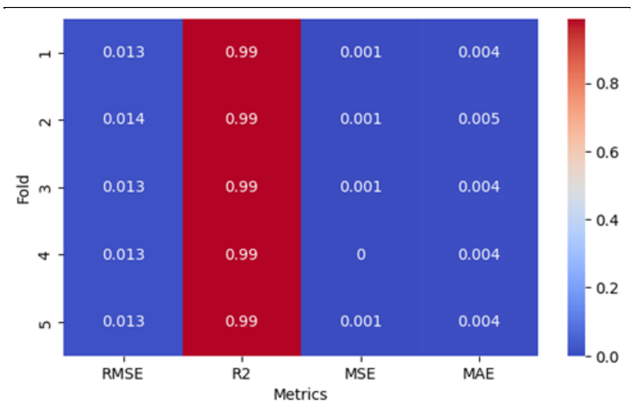

Fig. 3 Proposed model prediction visualization

Additionally, R^2 values are very high, ranging from 0.986 to 0.989, indicating that the proposed model explains a high percentage of the variance in the data. The MSE values are consistently low, all at 0.001, indicating that the average squared difference between the predicted values and the actual values is very small. Similarly, the MAE values are consistently low, all at 0.004, which means that, on average, the absolute difference between the predicted values and the actual values is very small. Overall, these results indicate that the model is accurate and reliable in predicting the target variable, as indicated by the low error metrics and high R^2 values.

4.2 Result Comparison with Machine Learning Models

The proposed model result is compared with 4 machine learning models as shown in Table 3 and Fig. 4. The Random Forest model outperforms the listed models, demonstrating the highest R^2 value and the lowest MSE and MAE. However, compared with the proposed model, the proposed model stood out indicating that the proposed model excels in accurately and precisely predicting solar energy generation in smart homes based on the given features.

Table 3. Proposed Model Result Comparison

Model	R^2	MSE	MAE
LGBM	0.9165	0.0014	0.0184
Random Forest	0.9734	0.0004	0.0057
AdaBoost	0.6097	0.0064	0.0464
Linear Regression	0.3819	0.0101	0.0748

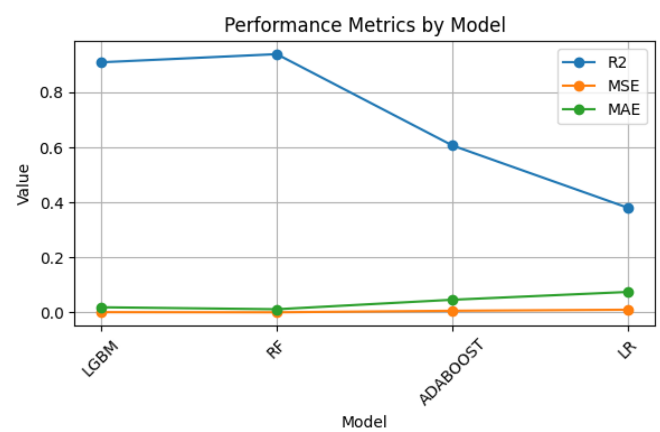

Fig. 4 Visualization of the predicted results

Although LGBM also performs well, with high R-squared and relatively low MSE and MAE, it falls slightly behind the Random Forest model in terms of predictive accuracy. On the other hand, AdaBoost and Linear Regression models exhibit lower R-squared values and higher MSE and MAE when compared to Random Forest and LGBM, indicating their lesser effectiveness in explaining and predicting solar energy generation in smart homes based on the given features.

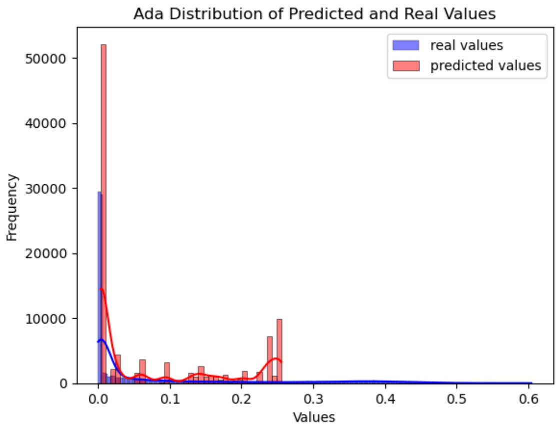

Fig. 5 AdaBoost prediction

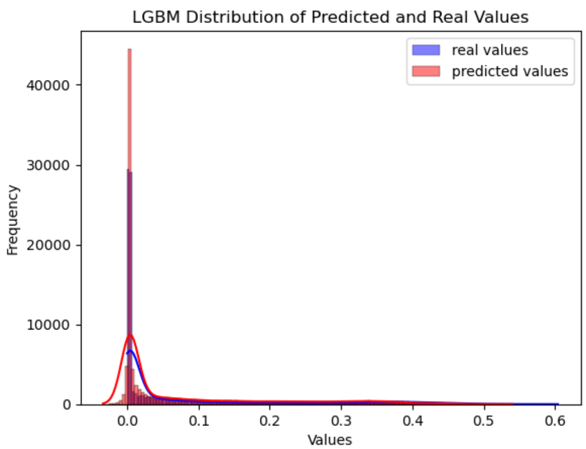

Fig. 6 LGBM prediction

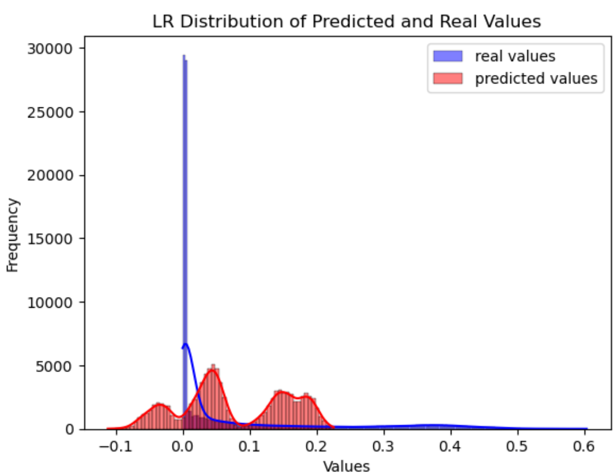

Fig. 7 Linear Regression prediction

Given that accuracy and precision are the primary concerns in predicting solar energy generation in smart homes, the Random Forest model emerges as the preferred choice based on the provided results. Fig. 5 shows the relationship between the real and predicted values of the AdaBoost model. The blue line represents the real values, while the red line represents the predicted values indicating that the predicted values generally being slightly lower than the real values. Fig. 6 shows a scatter plot of the predicted vs real values of the LGBM model with the values plotted on a two-dimensional axis. The x-axis represents the predicted values, while the y-axis represents the real values. The graph shows how closely the predicted values align with the real values.

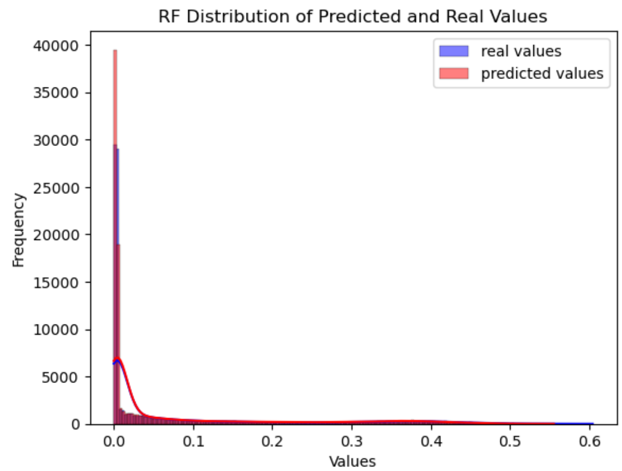

Fig. 8 Random Forest prediction

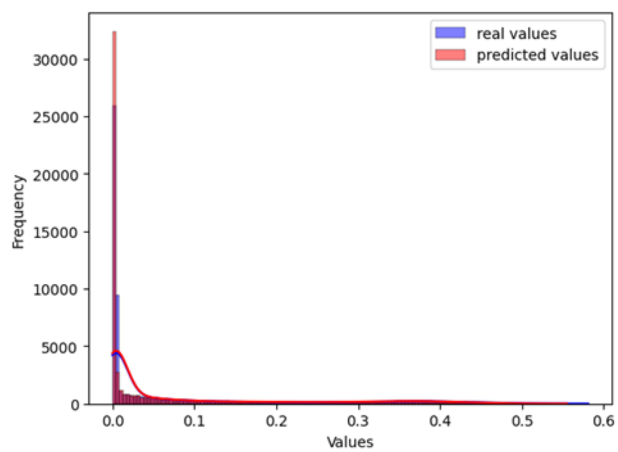

Fig. 9 Proposed model prediction

Fig. 7 shows that of the LR model. It is seen from the graph that the predicted values are slightly higher than the real values. The real values are represented by the blue line, while the predicted values are represented by the red line. This indicates that the prediction model tends to overestimate the values compared to the actual values. Fig. 8 and Fig. 9 show the graphs of the RF and the proposed model. it is evident that there is a strong correlation between the real and predicted values. The graphs show a clear relationship between the two lines, with the predicted values closely following the real values, The red and blue lines are almost parallel, indicating a high level of accuracy in the predictions. However, there is a slight gap between the two lines, indicating that there may be some discrepancies between the actual and predicted values.

Table 1. Descriptive statistics summary of the dataset

	Temp	Humidity	Visibility	App Temp	Pressure	Wind Speed	Wind Bearing	Precip Intensity	Dew Point	Precip Probability	House overall	Total	Solar
count	503910	503910	503910	503910	503910	503910	503910	503910	503910	503910	503910	503910	503910
mean	0.541	0.678	0.925	0.477	0.975	0.290	0.564	0.014	0.513	0.067	0.058	0.067	0.076
Std	0.204	0.198	0.161	0.218	0.008	0.174	0.297	0.059	0.253	0.197	0.072	0.054	0.128
Min	-0.135	0.133	0.027	-0.317	0.946	0.000	0.000	0.000	-0.361	0.000	0.000	0.000	0.000
25%	0.382	0.520	0.942	0.308	0.970	0.160	0.412	0.000	0.326	0.000	0.025	0.035	0.003
50%	0.537	0.694	1.000	0.498	0.975	0.259	0.579	0.000	0.517	0.000	0.038	0.049	0.004
75%	0.707	0.857	1.000	0.655	0.980	0.390	0.822	0.000	0.726	0.000	0.066	0.086	0.084
Max	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

5. CONCLUSION

With the rising global energy use and the urgent need for cleaner energy sources to protect the environment, accurately predicting solar energy is crucial. This study proposed an Ensemble-Tree Model Based on Bayesian Optimization for Solar Energy Generation Prediction in Smart Homes. The result is evaluated based on K-Fold cross validation while using the ‘smart-home-dataset-with-weather-information dataset as well as R-squared (R^2), Mean Squared Error (MSE), and Mean Absolute Error (MAE) to observe how accurate the predictions are. Furthermore, four machine learning methods: Light Gradient Boosting Machine (LGBM), Random Forest, AdaBoost, and Linear Regression were developed and compared in this study. We found that the Random Forest model outperforms others while the proposed model supersedes all other performances with an R^2 value of 0.988. Our findings show how the proposed model can help manage energy, save costs, and protect the environment in smart homes. Future research can explore better techniques and test these models in real-world situations. Our study adds to our understanding of renewable energy prediction and how machine learning can make smart homes more sustainable.

ACKNOWLEDGMENT

The authors gratefully acknowledge the support of the National Natural Science Foundation of China (NFSC, Grant No. 52007025, 24NSFSC1803), the Fundamental Research Funds for the Central Universities Grant No. ZYGX2019J034, Science and Technology Innovation Talent Program of Sichuan Provincial (Grant No. 22CXRC0010) and the Science and Technology Support Program of Sichuan Province (2022JDRC0025).

REFERENCE

[1] R. Sabzehgar, D. Z. Amirhosseini, and M. Rasouli, “Solar power forecast for a residential smart microgrid based on numerical weather predictions using artificial intelligence methods,” *Journal of Building Engineering*, vol. 32, p. 101629, Nov. 2020, doi: 10.1016/j.jobe.2020.101629.

[2] C. Jiménez-Xamán, J. Xamán, M. Gijón-Rivera, I. Zavala-Guillén, F. Noh-Pat, and E. Simá, “Assessing the thermal performance of a rooftop solar chimney attached to a single room,” *Journal of Building Engineering*, vol. 31, p. 101380, Sep. 2020, doi: 10.1016/j.jobe.2020.101380.

[3] D. Marikyan, S. Papagiannidis, and E. Alamanos, “Smart Home Technology Acceptance: An Empirical Investigation,” in *Digital Transformation for a Sustainable Society in the 21st Century*, I. O. Pappas, P. Mikalef, Y. K. Dwivedi, L. Jaccheri, J. Krogstie, and M. Mäntymäki, Eds., in Lecture Notes in Computer Science. Cham: Springer International Publishing, 2019, pp. 305–315. doi: 10.1007/978-3-030-29374-1_25.

[4] I. Priyadarshini, S. Sahu, R. Kumar, and D. Taniar, “A machine-learning ensemble model for predicting energy consumption in smart homes,” *Internet of Things*, vol. 20, p. 100636, 2022, Accessed: Feb. 19, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2542660522001172

[5] A. Zielonka, A. Sikora, M. Woźniak, W. Wei, Q. Ke, and Z. Bai, “Intelligent Internet of Things System for

Smart Home Optimal Convection," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 6, pp. 4308–4317, Jun. 2021, doi: 10.1109/TII.2020.3009094.

[6] O. Taiwo and A. E. Ezugwu, "Internet of Things-Based Intelligent Smart Home Control System," *Security and Communication Networks*, vol. 2021, p. e9928254, Sep. 2021, doi: 10.1155/2021/9928254.

[7] H. Klaus, K. Potter, and S. Mohamed, "Exploring the Potential of Machine Learning for Improving Energy Efficiency in IoT-Enabled Smart Homes," *Internet of Things*, Jan. 2024.

[8] "The Future of Weather Prediction: Harnessing Machine Learning Algorithms." Accessed: Mar. 25, 2024. [Online]. Available: https://www.linkedin.com/pulse/future-weather-prediction-harnessing-machine-learning-sophiya-divine-hwerc

[9] Q. Meteier, M. El Kamali, L. Angelini, and O. Abou Khaled, "A Recommender System for Increasing Energy Efficiency of Solar-Powered Smart Homes," *Sensors (Basel)*, vol. 23, no. 18, p. 7974, Sep. 2023, doi: 10.3390/s23187974.

[10] S. Bouazizi and H. Ltifi, "Enhancing accuracy and interpretability in EEG-based medical decision making using an explainable ensemble learning framework application for stroke prediction," *Decision Support Systems*, vol. 178, p. 114126, Mar. 2024, doi: 10.1016/j.dss.2023.114126.

[11] J. Wu, X.-Y. Chen, H. Zhang, L.-D. Xiong, H. Lei, and S.-H. Deng, "Hyperparameter Optimization for Machine Learning Models Based on Bayesian Optimizationb," *Journal of Electronic Science and Technology*, vol. 17, no. 1, pp. 26–40, Mar. 2019, doi: 10.11989/JEST.1674-862X.80904120.

[12] J. Huertas Tato and M. Centeno Brito, "Using Smart Persistence and Random Forests to Predict Photovoltaic Energy Production," *Energies*, vol. 12, no. 1, Art. no. 1, Jan. 2019, doi: 10.3390/en12010100.

[13] S. Al-Dahidi, O. Ayadi, J. Adeeb, M. Alrbai, and B. R. Qawasmeh, "Extreme learning machines for solar photovoltaic power predictions," *Energies*, vol. 11, no. 10, p. 2725, 2018.

[14] P. Pandiyan, S. Saravanan, K. Usha, R. Kannadasan, M. H. Alsharif, and M.-K. Kim, "Technological advancements toward smart energy management in smart cities," *Energy Reports*, vol. 10, pp. 648–677, Nov. 2023, doi: 10.1016/j.egyr.2023.07.021.

[15] M. Kutseva, "Achieving Energy Efficiency in Smart Home by Using Machine Learning Algorithm for Prediction," in *2023 International Scientific Conference on Computer Science (COMSCI)*, Sep. 2023, pp. 1–4. doi: 10.1109/COMSCI59259.2023.10315864.

[16] D. V. Tai, "Solar photovoltaic power output forecasting using machine learning technique," *J. Phys.: Conf. Ser.*, vol. 1327, no. 1, p. 012051, Oct. 2019, doi: 10.1088/1742-6596/1327/1/012051.

[17] "Smart Home Dataset with weather Information." Accessed: Apr. 01, 2024. [Online]. Available: https://www.kaggle.com/datasets/taranvee/smart-home-dataset-with-weather-information