© April 2026 | IJIRT | Volume 12 Issue 11 | ISSN: 2349-6002

Climate Based Energy Consumption Prediction System For University Campuses Using Machine Learning

Gayatri Deepak Malani[1], Bhushan Rajendra Sawant[2], Koushal Ravindra Khichade[3], Harshad Rajesh Kharat[4] and Prof. U.A. Patil[5]

1,2,3,4Department of Computer Science Engineering, D.Y. Patil Technical Campus Talsande, Faculty of Engineering and Faculty of Management, Talsande, Kolhapur, Maharashtra, India
5Guide, Department of Computer Science Engineering, D.Y. Patil Technical Campus Talsande, Faculty of Engineering and Faculty of Management, Talsande, Kolhapur, Maharashtra, India

Abstract—This paper presents a climate-based energy consumption prediction system designed for university campuses using machine learning techniques. Energy usage in universities is highly dynamic due to varying environmental conditions and time-based patterns, making accurate forecasting a challenging task. The proposed system integrates historical energy consumption data with real-time weather parameters such as temperature, humidity, wind speed, rainfall, and solar radiation to improve prediction accuracy. Machine learning models including Random Forest Regressor and XGBoost Regressor are implemented to perform hourly and time-range energy predictions. The system also provides an interactive web-based dashboard for visualization and analysis of energy trends. Experimental results demonstrate high prediction accuracy, with the Random Forest model achieving an R^2 score of 0.9799 and XGBoost achieving 0.9609. The proposed system supports efficient energy management, reduces operational costs, and contributes to sustainable energy utilization in university environments.

Index Terms—Climate data, Energy consumption prediction, Machine learning, Random Forest, Time-series forecasting, XGBoost.

I. INTRODUCTION

Energy consumption in large institutions such as university campuses has become a critical concern due to the increasing demand for electricity and the need for efficient resource management. Universities consist of multiple facilities including classrooms, laboratories, hostels, and administrative buildings, all of which contribute to complex and dynamic energy usage patterns. These patterns are significantly influenced by environmental factors such as temperature, humidity, wind speed, rainfall, and solar radiation, as well as time-based variations like daily schedules and seasonal changes.

Traditional energy forecasting methods mainly rely on historical data and basic statistical techniques, which are often insufficient to capture the nonlinear relationships between climate conditions and energy consumption. With the advancement of machine learning techniques, it is now possible to develop intelligent systems that can analyze large datasets and generate accurate predictions. In this paper, a climate-based energy consumption prediction system is proposed, which integrates historical energy data with real-time weather parameters. The system utilizes Random Forest Regressor for hourly prediction and XGBoost Regressor for time-range forecasting. The proposed approach improves prediction accuracy and supports efficient energy management, cost reduction, and sustainable utilization of energy resources in university environments.

II. METHODOLOGY

The proposed system follows a structured methodology to predict energy consumption using climate data and machine learning techniques. The process consists of multiple stages including data collection, preprocessing, feature engineering, model development, and prediction.

In the first stage, historical energy consumption data and real-time weather data are collected from datasets and APIs such as OpenWeatherMap. The collected data includes environmental parameters such as temperature, humidity, wind speed, rainfall, and solar radiation, which significantly influence energy usage patterns in university campuses.

The collected data is then preprocessed to improve data quality and ensure accurate predictions. This includes handling missing values, removing inconsistencies, normalization, and transformation of data into a suitable format. Time-based features such as hour, day, and month are also extracted to capture temporal variations in energy consumption.

In the next stage, feature engineering is performed to select relevant attributes that contribute to energy prediction. These features are used to train machine learning models. The system utilizes Random Forest Regressor for real-time and hourly energy prediction, as it effectively handles nonlinear relationships between environmental parameters and energy usage. Additionally, XGBoost Regressor is used for time-range forecasting based on historical time-series patterns, improving long-term prediction accuracy.

Finally, the trained models generate energy consumption predictions, which are displayed through an interactive dashboard. The system also evaluates model performance using metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R^2 score to ensure reliability and accuracy of predictions.

III. SYSTEM ARCHITECTURE

The system architecture of the proposed climate-based energy consumption prediction system is designed to efficiently process both historical and real-time data to generate accurate energy forecasts. The system consists of multiple interconnected modules including data collection, data preprocessing, machine learning models, and visualization.

Initially, the system collects historical energy consumption data along with real-time weather data from external sources such as the OpenWeatherMap API. The collected data includes important environmental parameters such as temperature, humidity, wind speed, rainfall, and solar radiation, which directly influence energy usage patterns in university campuses.

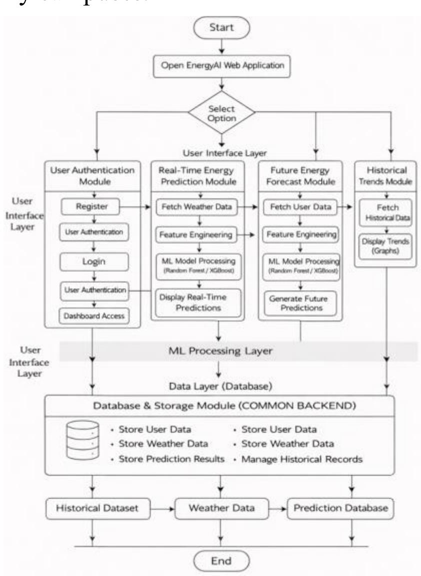

Fig. 1. System Architecture of Proposed Climate-Based Energy Consumption Prediction System.

The collected data is then passed to the preprocessing module, where data cleaning, normalization, and transformation are performed to ensure data quality and consistency. Time-based features such as hour, day, and month are also extracted to capture temporal variations in energy consumption.

After preprocessing, the data is fed into the machine learning module, where models such as Random Forest Regressor and XGBoost Regressor are applied. The Random Forest model is used for real-time and hourly prediction, while the XGBoost model is used for forecasting energy consumption over a selected time period. These models analyze complex relationships between environmental parameters and energy usage to generate accurate predictions.

The predicted results are then stored in the database and displayed to the user through an interactive web-based dashboard. The visualization module presents the output in the form of graphs and charts, allowing users to analyze trends and make informed decisions. The overall system ensures smooth data flow from input to prediction and output, providing an efficient and reliable solution for energy management.

IV. DATA FLOW DIAGRAM

The data flow diagram (DFD) represents the flow of data within the proposed system and illustrates how input data is processed through different modules to generate energy consumption predictions.

A. DFD Level 0

The Level 0 DFD represents the overall system as a single process and shows the interaction between the user, system, and external data sources such as weather APIs.

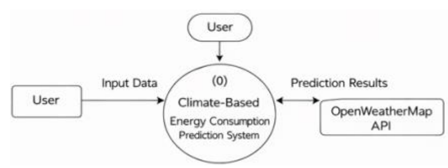

Fig. 2. Data Flow Diagram Level 0 of Proposed System.

B. DFD Level 1

The Level 1 DFD provides a detailed view of the system by breaking it into major processes such as data collection, preprocessing, machine learning, and visualization.

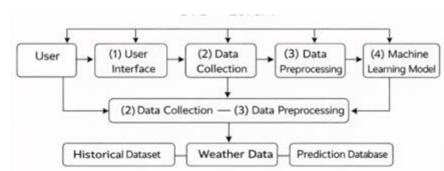

Fig. 3. Data Flow Diagram Level 1 of Proposed System.

C. DFD Level 2

The Level 2 DFD further explains the internal working of each module, including feature engineering, model processing, prediction generation, and data storage.

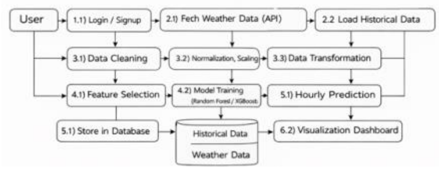

Fig. 4. Data Flow Diagram Level 2 of Proposed System.

V. RESULTS AND DISCUSSION

The performance of the proposed climate-based energy consumption prediction system is evaluated using machine learning models including Random Forest Regressor and XGBoost Regressor. The evaluation is carried out using standard performance metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R^2 score.

Table I: Model Performance Comparison

Model	MAE	MSE	RMSE	R^2 Score	Prediction Type
Random Forest Regressor	13.35	288.28	16.98	0.9799	Hourly Prediction
XGBoost Regressor	18.64	562.21	23.71	0.9609	Time-Range Prediction

The Random Forest model achieved a high R^2 score of 0.9799 with low error values, indicating excellent accuracy in predicting hourly energy consumption. The low MAE and RMSE values demonstrate that the model effectively captures the relationship between environmental parameters and energy usage patterns. This makes it highly suitable for real-time and short-term predictions.

The XGBoost model achieved an R^2 score of 0.9609, showing strong performance in forecasting energy consumption over a selected time period. Although the error values are slightly higher compared to Random Forest, the model performs efficiently for time-range prediction due to its ability to learn complex patterns from historical time-series data.

From the results, it can be observed that Random Forest performs better for real-time and hourly predictions, while XGBoost is more effective for long-term forecasting. The integration of both models enhances the flexibility and accuracy of the system. Additionally, the use of climate parameters such as temperature, humidity, wind speed, rainfall, and solar radiation significantly improves prediction performance.

VI. CONCLUSION

This paper presented a climate-based energy consumption prediction system for university campuses using machine learning techniques. The system integrates historical energy data with real-time environmental parameters such as temperature, humidity, wind speed, rainfall, and solar radiation to improve prediction accuracy. The implementation of Random Forest Regressor for hourly prediction and XGBoost Regressor for time-range forecasting enables the system to capture both short-term and long-term energy consumption patterns effectively.

The experimental results demonstrate that the proposed models achieve high prediction accuracy, with Random Forest showing slightly better performance for real-time predictions, while XGBoost performs efficiently for time-based forecasting. The system also provides an interactive dashboard for visualization, which enhances user understanding and supports informed decision-making. Overall, the proposed approach offers a reliable and scalable solution for efficient energy management, cost reduction, and sustainable utilization of resources in university environments.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to Prof. U. A. Patil for their valuable guidance, continuous support, and encouragement throughout the development of this project. The authors also thank D.Y. Patil, Faculty of Engineering and Faculty of Management, Talsande for providing the necessary resources and infrastructure to carry out this work successfully.

REFERENCES

[1] J. Dong, Z. Wang, and H. Liu, "Short-term energy consumption prediction using machine learning techniques," IEEE Access, vol. 8, pp. 123456–123467, 2020.

[2] S. Fan and R. J. Hyndman, "Short-term load forecasting based on a semi-parametric additive model," IEEE Transactions on Power Systems, vol. 27, no. 1, pp. 134–141, 2012.

[3] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proc. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 785–794.

[4] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[5] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.

[6] A. Jain and B. Satish, "Energy consumption prediction using machine learning for smart grids," International Journal of Energy Research, 2021.

[7] H. Zou, Y. Wang, and J. Liu, "Energy consumption prediction using deep learning approach," Applied Energy, vol. 240, pp. 101–112, 2019.

[8] A. Ahmad, N. Javaid, and Z. A. Khan, "A data-driven approach for energy consumption prediction in smart grids," Sustainable Cities and Society, vol. 41, pp. 695–706, 2018.

[9] S. P. Wang, Z. Gao, and L. Wang, "Short-term load forecasting using hybrid machine learning models," Energy Reports, vol. 6, pp. 100–108, 2020.

[10] OpenWeatherMap, "Weather API," [Online]. Available: https://openweathermap.org/api