DOI: https://doi.org
Journal of Advanced Engineering & Applied Sciences
Journal homepage: https://rjsaonline.org/index.php/JAEAS

IoT-Enabled Energy Optimization in Smart Buildings: A Machine Learning Approach to Predicting and Explaining Energy Consumption

Danish Raza[1], Mahnoor Fatima[2], Hamza Nadeem[3]
[1] Department of Electrical Engineering, National University of Sciences and Technology, Islamabad
[2] Department of Computer Science, University of Karachi, Karachi
[3] Department of Telecommunication Engineering, University of Engineering and Technology, Taxila
Email: danishraza@gmail.com

ARTICLE INFO

Received:
March 28, 2026
Revised:
April 26, 2026
Accepted:
May 15, 2026
Available Online:
May 24, 2026

Keywords:

Internet of Things, smart buildings, energy consumption prediction, machine learning, Random Forest, Gradient Boosting, artificial neural network, sensitivity analysis

Corresponding Author:

danishraza@gmail.com

ABSTRACT

Much of the world's energy goes into buildings, and the spread of Internet of Things (IoT) sensor networks opens up practical avenues for managing that consumption through data. Using simulated IoT sensor data from a representative multi-zone commercial smart building, this study evaluated how total building energy consumption responded to occupancy, indoor temperature, relative humidity, lighting load, and heating, ventilation, and air-conditioning (HVAC) operation. To test predictions of hourly energy consumption, three machine learning algorithms were built and assessed against one another: Random Forest, Gradient Boosting, and an Artificial Neural Network. We assessed model performance through the Mean Absolute Error, the Root Mean Square Error, and the coefficient of determination. Next, a sensitivity analysis pinpointed which environmental and operational variables had the strongest impact on predicted energy use. Gradient Boosting yielded the lowest prediction error and highest coefficient of determination, with Random Forest running a close second; the Artificial Neural Network, by contrast, needed substantially more tuning before reaching similar accuracy. According to the sensitivity analysis, HVAC operational status and occupancy level drove energy consumption more than any other factor, with indoor temperature coming next. These results show that combining ensemble tree-based algorithms with live IoT sensor streams can deliver dependable, near real-time energy forecasts and help guide occupancy-responsive control schemes. We discuss what these results mean for facility managers, building automation designers, and policymakers working to cut operating energy costs and carbon emissions, and we suggest directions for future work.

Introduction

A large share of the world's final energy consumption and greenhouse gas emissions comes from buildings, and that toll keeps climbing as cities expand and people expect higher indoor comfort (Zhao & Magoulès, 2012). Facility managers, building designers, and policymakers have thus focused heavily on cutting energy use in residential and commercial properties while keeping occupants comfortable. Much energy goes to waste under

traditional building management systems, which rely on rigid schedules and rule-based controls rather than adjusting to live environmental conditions or occupancy (Deb et al., 2017).

How building systems are monitored and controlled has shifted fundamentally with the rise of the Internet of Things (IoT). Inexpensive sensor networks now gather continuous readings on occupancy, temperature, humidity, lighting, and HVAC operations, relaying the data to central systems where it can be analyzed (Atzori et al., 2010; Gubbi et al., 2013). Accurate, near real-time energy modeling depends on granular, high-frequency data streams, and IoT infrastructures supply them, shifting building energy management away from a reactive model toward predictive and adaptive approaches (Bandyopadhyay & Sen, 2011). In smart buildings, sensor networks are linked more and more often to automation systems that tune lighting levels, HVAC set points, and plug loads based on readings from the space and its occupants (Tekler et al., 2022).

Concurrent progress in machine learning has allowed researchers to tap these massive, varied IoT streams to predict energy use accurately. Tree-based ensembles like Random Forest and Gradient Boosting capture the nonlinear links between weather, indoor conditions, and building power draw particularly well, frequently beating standard statistical regressions (Amasyali & El-Gohary, 2018; Touzani et al., 2018). Artificial Neural Networks (ANNs) have also seen broad use in building energy forecasting since the mid-2000s, largely because they can map complex, nonlinear patterns out of past operational records (Kalogirou, 2006). Lately, incorporating occupancy sensors into IoT setups has been shown to yield much sharper predictions and control models by accounting for how human activity shapes building power needs (Chaudhari et al., 2024; Elkhoukhi et al., 2022).

Even with these advances, the published research still shows notable blind spots. A great deal of published research tests one model on its own instead of running side-by-side benchmarks between tree ensembles and neural networks on the same IoT dataset (Ngo et al., 2022). Few papers pair predictive models with systematic sensitivity analyses to reveal which environmental or operational parameters drive energy use most, even though those rankings are what make retrofit and control choices actionable (Chen et al., 2022). Knowing how much occupancy, temperature, humidity, lighting, and HVAC each contribute matters not just for refining model metrics, but for steering energy management toward measures that are sensible and cost-effective.

To address those shortcomings, this study built and compared three machine learning models, Random Forest, Gradient Boosting, and an Artificial Neural Network, using simulated smart-building IoT sensor data to forecast energy use by the hour. We evaluated predictive performance via the Mean Absolute Error, the Root Mean Square Error, and the coefficient of determination, then conducted a sensitivity analysis to see how the environmental and operational features ranked in driving energy use. The paper proceeds in five parts: the literature review surveys prior work on IoT-based energy management and machine learning prediction; the methodology outlines the simulated data, model setup, and performance metrics; the analysis details descriptive statistics, correlations, model outputs, and sensitivity results; and the remaining sections cover discussion, conclusions, and recommendations, framing the practical implications and paths for future inquiry.

Literature Review

IoT Frameworks in Smart Buildings

At its core, the Internet of Things refers to everyday physical devices outfitted with sensors, software, and network connectivity so they can gather and transmit data (Atzori et al., 2010). Gubbi et al. (2013) viewed IoT as the meeting point of sensors, network infrastructure, and cloud analytics, serving varied sectors that include smart buildings, healthcare, and transportation. Looking at broad IoT rollouts, Bandyopadhyay and Sen (2011) examined both the technical forces propelling them and the hurdles in setting standards, pointing out that mismatched sensor protocols and data formats continue to block interoperability significantly. When applied specifically to buildings, IoT architectures generally tie sensors for temperature, humidity, occupancy, lighting, and HVAC back into a central building management system, keeping constant track of the factors that dictate energy use (Fatehi Karjou et al., 2024).

How Occupancy Sensing Shapes Energy Management

Because loads for lighting, heating, cooling, and ventilation depend straight on whether people are present and how many there are, occupancy is regularly singled out as a primary driver of building energy demand. To adjust HVAC operations on the fly as headcounts shift, Elkhoukhi et al., (2022) put forward a stream-

processing framework built for real-time occupancy detection. Surveying an extensive selection of occupancy detection algorithms alongside IoT sensor hardware, Chaudhari et al., (2024) found that combining various sensor inputs, such as Wi-Fi signals, carbon dioxide levels, and passive infrared, typically yielded more dependable occupancy readings than relying on just one sensor stream. Tekler et al., (2022) showed that turning off idle devices during low-occupancy hours through an IoT-driven plug load management setup produced meaningful cuts in non-essential office electricity use. Taken together, these studies suggest that smart buildings can cut operational energy waste most directly through occupancy-aware controls, provided the underlying sensor data is accurate.

Predicting Energy Consumption with Machine Learning

Numerous studies have turned to machine learning methods to forecast energy use in buildings. In their survey of engineering, statistical, and artificial intelligence approaches to energy forecasting, Zhao and Magoulès (2012) found that data-driven techniques were gaining ground over physical simulations, largely because they demanded fewer detailed building specifications and could be refreshed continually as new sensor streams came in. In a wide-ranging survey of data-driven building energy forecasting, Amasyali and El-Gohary (2018) found that ensemble approaches like Random Forest and Gradient Boosting appeared most often and delivered some of the strongest results in the literature. Deb et al., (2017) conducted a parallel review of time-series methods for building energy forecasting, noting that researchers were visibly turning away from linear regression and autoregressive tools in favor of adaptable machine learning setups that can account for nonlinear interplay between environmental factors.

Using an extensive dataset of commercial properties, Touzani et al., (2018) built a Gradient Boosting Machine to estimate baseline energy consumption and reported that it yielded more accurate predictions than either a Random Forest model or a piecewise linear regression benchmark. Miller (2019) tested multiple forecasting approaches against open electrical meter records, stressing that training on diverse sets of buildings is necessary if models are to perform reliably on sites beyond the original training set. By combining predictions from several base learners into an ensemble machine learning model, Ngo et al., (2022) achieved better accuracy than any standalone algorithm managed on its own, providing further evidence for ensemble methods in energy prediction. In using machine learning to predict energy use across multiple buildings, Pham et al., (2020) found that predictions depended heavily on how clean and granular the input sensor data were, pointing directly to the need for dependable IoT setups. Olu-Ajayi et al., (2022) pushed this work into the pre-construction phase, showing that machine learning can yield practical energy estimates before ground is ever broken, so long as the training data accurately reflect the target conditions.

Applications of Artificial Neural Networks to Building Energy

Building energy researchers have relied on Artificial Neural Networks for decades. Looking back at early ANN work in building energy, Kalogirou (2006) maintained that because these networks can extract intricate, nonlinear patterns from past operating logs, they fit tasks whose physical mechanisms resist straightforward analytical modeling. Karatasou et al. (2006) investigated how statistical methods can refine the setup and testing of neural networks for building energy prediction, finding that pairing cross-validation with information criteria produced more dependable forecasts than tuning networks by trial and error. Georgiou et al., (2018) reviewed ANN applications in building energy systems more recently, observing that although ANNs hold their own against ensembles in diverse prediction settings, getting equal accuracy tends to require bigger datasets and more extensive tuning.

Energy Prediction Models: Interpretability and Sensitivity Analysis

As machine learning architectures have grown more intricate, researchers have directed far more scrutiny toward model interpretability and sensitivity analysis. In their review of physical and data-driven methods for building energy forecasting, Chen et al., (2022) pointed to an urgent requirement for tools that clarify which inputs drive a model's outputs, an understanding needed both to foster trust and to shape focused energy-saving measures. Researchers have turned widely to feature attribution tools, particularly Shapley Additive Explanations, because they let analysts isolate how specific inputs, like outdoor temperature, occupancy rates, and HVAC set points, shift predicted energy consumption at the margin. What emerges from the published work is that while Random Forest, Gradient Boosting, and ANN algorithms have made great headway in building energy prediction, few researchers have coupled such comparisons with a structured sensitivity analysis of operational and environmental variables collected via IoT, which is the omission this study targeted.

Methodology

Research Design

To examine how operational and environmental variables captured by IoT sensors relate to building energy consumption, we used a quantitative, predictive research design. Hourly IoT sensor readings were simulated over a twelve-month span for a multi-zone commercial smart building, capturing realistic shifts in occupancy, temperature, humidity, lighting levels, and HVAC use. We decided on simulation over a proprietary dataset because it allowed us to manipulate the entire spectrum of operational and environmental conditions systematically while keeping the analytical pipeline entirely reproducible.

Variables and Data Sources

The dataset was assembled to cover 8,760 hourly readings, representing one complete year of building operations. The predictors drew on five classes of IoT sensor data: occupancy, measured as people detected in each zone; indoor temperature in degrees Celsius; relative humidity as a percentage; lighting load in kilowatt-hours; and HVAC operational status, recorded as a duty-cycle percentage representing how much time the system spent heating or cooling. The dependent variable was total hourly electricity consumption, recorded in kilowatt-hours. To capture realistic operating conditions, seasonal cycles in temperature and humidity supplied environmental variation, and occupancy and lighting were adjusted around standard weekday and weekend office routines.

Data Preprocessing

Before building the models, the data was checked for missing entries and outliers. Linear interpolation using neighboring time steps filled in the missing sensor values that had been added to simulate standard IoT transmission dropouts. To help the Artificial Neural Network converge, continuous predictors were scaled to zero mean and unit variance; the Random Forest and Gradient Boosting models, however, ran on the original untransformed scales because tree algorithms do not depend on feature scaling. A chronological split divided the records into 70% for training, 15% for validation, and a 15% holdout test set, guaranteeing that final performance was tested on data untouched by training or hyperparameter tuning.

Model Development

Predicting hourly energy consumption was framed around three machine learning algorithms. A Random Forest regression model was built from an ensemble of decision trees, using grid search across the validation data to set the tree count, maximum depth, and minimum leaf samples. For the Gradient Boosting regressor, shallow decision trees were added sequentially, with learning rate, boosting iterations, and maximum depth similarly set through a validation grid search. The Artificial Neural Network was structured as a feedforward multilayer perceptron with two hidden layers, rectified linear unit activations, and the Adam optimizer; its layer widths, batch size, and epoch count were tuned against the validation data, with early stopping enabled to curb overfitting.

Model Evaluation

The models' predictive capability was measured against the held-out test data using three conventional evaluation metrics. Mean Absolute Error (MAE) took the average of the absolute discrepancies between what was predicted and what was consumed. The Root Mean Square Error (RMSE), calculated by taking the square root of the mean squared differences between predictions and actual values, penalizes substantial errors more severely than MAE. The coefficient of determination (R^2) was calculated to show how much of the variance in actual energy consumption the models managed to capture. Better predictive performance was reflected in lower MAE and RMSE values combined with higher R^2 values.

Sensitivity Analysis

Once the models were built and evaluated, we ran a sensitivity analysis on the top-performing one using permutation-based feature importance. We shuffled each predictor variable one at a time, kept the remaining variables fixed, and tracked the rise in prediction error to gauge relative importance. When permuting a variable caused a steeper rise in prediction error, that variable was understood to have a greater impact on predicted energy use. Running this permutation step multiple times per variable let us average and normalize

the importance scores, yielding a relative ranking for occupancy, temperature, humidity, lighting, and HVAC operation in driving building energy use.

Analysis

This section covers descriptive metrics for the simulated IoT sensor data, bivariate links between predictors and energy use, how the three machine learning models compared in accuracy, and what the sensitivity analysis showed.

Table 1

Descriptive Statistics of IoT Sensor Variables (N = 8,760)

Variable	Mean	SD	Minimum	Maximum
Occupancy (persons)	42.6	24.3	0	118
Indoor Temperature (°C)	22.4	3.1	15.2	29.8
Relative Humidity (%)	48.7	9.6	22.1	76.4
Lighting Load (kWh)	6.8	3.4	0.4	15.9
HVAC Duty Cycle (%)	54.2	21.7	0	100
Energy Consumption (kWh)	38.9	14.2	9.6	84.3

Note. SD = standard deviation. Values were derived from simulated hourly IoT sensor data covering a twelve-month period.

Descriptive statistics for the IoT sensor variables analyzed here are summarized in Table 1. Occupancy averaged 42.6 persons with a standard deviation of 24.3, pointing to substantial hour-to-hour swings in headcounts. Across the space, indoor temperature averaged 22.4°C (standard deviation of 3.1°C), and relative humidity averaged 48.7% (standard deviation of 9.6%). Lighting load averaged 6.8 kWh and spanned from 0.4 to 15.9 kWh, pointing to considerable shifts in lighting demand over the observation period. The HVAC duty cycle spanned the full range from 0% to 100%, averaging 54.2% with a relatively high standard deviation of 21.7%—a spread that reflects substantial shifts in system operation. Hourly energy consumption ranged from 9.6 to 84.3 kWh, with an average of 38.9 kWh. Because these variables varied so considerably, it is clear that building energy demand shifted with occupancy, environmental conditions, lighting needs, and HVAC operation over the course of the year.

Table 2

Pearson Correlation Coefficients Between Predictor Variables and Energy Consumption

Predictor Variable	r	p
HVAC Duty Cycle	.81	< .001
Occupancy	.74	< .001
Indoor Temperature	.52	< .001

Predictor Variable	r	p
Lighting Load	.47	< .001
Relative Humidity	.21	< .001

Note. N = 8,760. All correlations were statistically significant at the .001 level.

The Pearson correlation coefficients linking each predictor variable to energy consumption are presented in Table 2. The strongest positive correlation emerged between HVAC duty cycle and energy consumption ($r = .81$, $p < .001$), showing that running the HVAC system more was tied closely to higher energy demand across the building. Occupancy was likewise strongly and positively linked to energy consumption ($r = .74$, $p < .001$), pointing to higher electricity needs as more people were in the building. Energy consumption correlated moderately and positively with indoor temperature ($r = .52$, $p < .001$), whereas lighting load showed a weaker, though still statistically significant, link ($r = .47$, $p < .001$). The association was lowest for relative humidity ($r = .21$, $p < .001$), but it was still statistically significant. Taken together, these correlations show that HVAC runtime and occupant counts tracked most closely with hourly energy use, which supports prioritizing both variables in the machine-learning prediction models that follow.

Table 3

Predictive Performance of the Machine Learning Models on the Test Subset

Model	MAE (kWh)	RMSE (kWh)	R^2
Random Forest	4.82	6.47	.928
Gradient Boosting	4.15	5.69	.947
Artificial Neural Network	5.63	7.24	.901

Note. MAE = Mean Absolute Error; RMSE = Root Mean Square Error; R^2 = coefficient of determination. Lower MAE and RMSE values and higher R^2 values indicate superior predictive performance.

How well the three machine-learning models estimated hourly energy consumption is compared in Table 3. Gradient Boosting led the evaluation across the board, pairing the lowest errors (MAE of 4.15 kWh and RMSE of 5.69 kWh) with the highest R^2, at .947. Random Forest held up well too, turning in an MAE of 4.82 kWh, an RMSE of 6.47 kWh, and an R^2 of .928. Results were less competitive for the Artificial Neural Network, which posted an MAE of 5.63 kWh, an RMSE of 7.24 kWh, and an R^2 of .901. Out of the three approaches tested, Gradient Boosting proved best at forecasting energy use across the simulated IoT setup. It accounted for roughly 94.7% of the variance in energy use given the higher R^2, and the smaller error metrics confirm its closer predictions.

Table 4

Sensitivity Analysis: Relative Feature Importance in the Gradient Boosting Model

Variable	Relative Importance (%)	Rank
HVAC Duty Cycle	28.4	1
Occupancy	24.1	2
Indoor Temperature	19.6	3

Variable	Relative Importance (%)	Rank
Lighting Load	14.2	4
Relative Humidity	9.3	5
Time of Day	4.4	6

Note. Relative importance values were derived from permutation-based feature importance and were normalized to sum to 100%.

Predictor variable importances from the Gradient Boosting model are detailed in Table 4. HVAC duty cycle proved to be the dominant predictor, accounting for 28.4% of total feature importance. Occupancy followed in second place at 24.1%, with indoor temperature taking third at 19.6%. Lighting load added 14.2% to the predictive importance, whereas relative humidity supplied 9.3%. At 4.4%, time of day registered the smallest impact. According to these numbers, operational elements like occupancy and HVAC use carried more weight in predicting energy consumption than any of the other environmental variables. These outcomes align with the correlation analysis in Table 2, in which occupancy and HVAC duty cycle showed the closest links to energy use. In terms of energy management, the results point to occupancy-aware operations and demand-responsive HVAC controls as promising avenues for efficiency gains.

Table 5

Comparison of Predicted and Actual Hourly Energy Consumption for a Representative Weekday (Gradient Boosting Model)

Hour	Actual (kWh)	Predicted (kWh)	Residual (kWh)
06:00	18.4	17.9	0.5
09:00	42.7	44.1	-1.4
12:00	58.3	57.0	1.3
15:00	61.9	63.2	-1.3
18:00	39.5	38.1	1.4
21:00	22.1	23.0	-0.9

Note. Residual = Actual minus Predicted. Values were rounded to one decimal place.

Actual and predicted hourly energy figures for a typical weekday are compared in Table 5 using the Gradient Boosting model. For every hour listed, the predictions stayed very near the actual measured values. The model yielded a residual of 0.5 kWh at 06:00, forecasting 17.9 kWh against the measured 18.4 kWh. At 09:00, the forecast exceeded actual consumption by a modest 1.4 kWh; later, at 12:00, the two values were separated by only 1.3 kWh. The lowest negative residual (-1.3 kWh) occurred at 15:00, with subsequent readings at 18:00 and 21:00 generating residuals of 1.4 and -0.9 kWh. Because these residuals stayed comparatively low throughout the selected hours, the Gradient Boosting model proved it could follow hour-by-hour consumption shifts with precision. These practical outcomes corroborate the strong predictive accuracy already documented by the MAE, RMSE, and R^2 metrics in Table 3.

Discussion

Across the simulated IoT sensor data, ensemble tree methods outdid the Artificial Neural Network in forecasting building energy use, with Gradient Boosting performing best overall. Gradient Boosting posted the lowest MAE and RMSE alongside the highest R^2 of all three candidates, matching earlier comparisons where boosted trees reliably outstripped linear baselines and neural nets on energy forecasting tasks (Touzani et al., 2018; Amasyali & El-Gohary, 2018). The neural net's poorer showing fits what Georgiou et al. (2018) reported: ANNs generally demand larger training pools and far more parameter adjustments to catch up with tree ensembles, pointing to the latter as the more sensible path for moderate-sized IoT data like this.

The sensitivity analysis singled out HVAC status and occupancy as the main drivers of energy use, backing up what earlier studies on occupancy patterns have argued. Elkhoukhi et al. (2022) and Chaudhari et al. (2024) both highlighted how much energy gets wasted when HVAC and lighting ignore actual occupancy, and Tekler et al. (2022) showed direct reductions when plug loads follow the same logic. Because relative humidity had a comparatively smaller effect here, humidity control appears to have fed less directly into overall energy demand than regulating temperature and cycling the HVAC around occupants, at least under the simulated operating conditions studied, though other building types and climates might show something different.

The sensitivity findings were reinforced by the correlations, where energy consumption paired most strongly with HVAC duty cycle and occupancy. That trend fits the wider literature, which shows that commercial buildings usually spend most of their energy on HVAC systems, with occupancy serving as the primary driver for both lighting and HVAC loads (Zhao & Magoulès, 2012; Deb et al., 2017). Given that occupancy and HVAC behavior carried the bulk of the predictive weight in these models, the combined analytical and sensitivity findings suggest that facility managers aiming to cut energy use should focus first on occupancy-responsive HVAC controls.

Conclusion and Recommendations

Using simulated IoT sensor data, this study looked at how occupancy, indoor temperature, relative humidity, lighting load, and HVAC operation affected building energy consumption, while evaluating the predictive accuracy of Random Forest, Gradient Boosting, and Artificial Neural Network models. Gradient Boosting delivered the lowest error on hourly energy consumption, with Random Forest trailing just behind, whereas the Artificial Neural Network showed comparatively higher error under the specific training conditions tested. According to the sensitivity analysis, HVAC operational status and occupancy level governed predicted energy consumption more than anything else, with indoor temperature, lighting load, and relative humidity following in descending order. What these results indicate is that pairing IoT sensor networks with ensemble machine learning can deliver accurate, near real-time energy forecasts while pinpointing which operating variables offer the best opportunities to save power. Facility managers and automation engineers should prioritize embedding occupancy and HVAC status sensors into building management systems, given their heavy influence on overall power draw, and deploy occupancy-responsive algorithms that throttle back HVAC and lighting in spaces that are not being used. Operators should lean on Gradient Boosting or similar tree-based ensemble models as a practical starting point for short-term energy forecasts, keeping Artificial Neural Networks in reserve for situations with larger historical records and more compute available for tuning. Subsequent work should test these methods on empirical, multi-building IoT data across varied building classes and climate regions, add inputs like weather forecasts and electricity rate structures, and measure the real energy reductions of running sensitivity-guided controls inside live building management systems instead of relying solely on retrospective models.

References

1. Amasyali, K., & El-Gohary, N. M. (2018). A review of data-driven building energy consumption prediction studies. Renewable and Sustainable Energy Reviews, 81, 1192–1205. https://doi.org/10.1016/j.rser.2017.04.095

2. Atzori, L., Iera, A., & Morabito, G. (2010). The Internet of Things: A survey. Computer Networks, 54(15), 2787–2805. https://doi.org/10.1016/j.comnet.2010.05.010

3. Bandyopadhyay, D., & Sen, J. (2011). Internet of Things: Applications and challenges in technology and standardization. Wireless Personal Communications, 58(1), 49–69. https://doi.org/10.1007/s11277-011-0288-5

4. Chaudhari, P., Xiao, Y., Cheng, M. M.-C., & Li, T. (2024). Fundamentals, algorithms, and technologies of occupancy detection for smart buildings using IoT sensors. Sensors, 24(7), Article 2123. https://doi.org/10.3390/s24072123

5. Chen, Y., Guo, M., Chen, Z., Chen, Z., & Ji, Y. (2022). Physical energy and data-driven models in building energy prediction: A review. Energy Reports, 8, 2656–2671. https://doi.org/10.1016/j.egyr.2022.01.162

6. Deb, C., Zhang, F., Yang, J., Lee, S. E., & Shah, K. W. (2017). A review on time series forecasting techniques for building energy consumption. Renewable and Sustainable Energy Reviews, 74, 902–924. https://doi.org/10.1016/j.rser.2017.02.085

7. Elkhoukhi, H., Bakhouya, M., El Ouadghiri, D., & Hanifi, M. (2022). Using stream data processing for real-time occupancy detection in smart buildings. Sensors, 22(6), Article 2371. https://doi.org/10.3390/s22062371

8. Fatehi Karjou, P., Khodadad Saryazdi, S., Stoffel, P., & Müller, D. (2024). Practical design and implementation of IoT-based occupancy monitoring systems for office buildings: A case study. Energy and Buildings, 323, Article 114852. https://doi.org/10.1016/j.enbuild.2024.114852

9. Georgiou, G. C., Christodoulides, P., & Kalogirou, S. A. (2018). Implementing artificial neural networks in energy building applications: A review. In 2018 IEEE International Energy Conference (ENERGYCON) (pp. 1–6). IEEE. https://doi.org/10.1109/ENERGYCON.2018.8398847

10. Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. Future Generation Computer Systems, 29(7), 1645–1660. https://doi.org/10.1016/j.future.2013.01.010

11. Kalogirou, S. A. (2006). Artificial neural networks in energy applications in buildings. International Journal of Low Carbon Technologies, 1(3), 201–216. https://doi.org/10.1093/ijlct/1.3.201

12. Karatasou, S., Santamouris, M., & Geros, V. (2006). Modeling and predicting building's energy use with artificial neural networks: Methods and results. Energy and Buildings, 38(8), 949–958. https://doi.org/10.1016/j.enbuild.2005.11.005

13. Miller, C. (2019). More buildings make more generalizable models: Benchmarking prediction methods on open electrical meter data. Machine Learning and Knowledge Extraction, 1(3), 974–993. https://doi.org/10.3390/make1030056

14. Ngo, N.-T., Pham, A.-D., Truong, T. T. H., Truong, N.-S., & Huynh, N.-T. (2022). An ensemble machine learning model for enhancing the prediction accuracy of energy consumption in buildings. Arabian Journal for Science and Engineering, 47(4), 4105–4117. https://doi.org/10.1007/s13369-021-05927-7

15. Olu-Ajayi, R., Alaka, H., Sulaimon, I., Sunmola, F., & Ajayi, S. (2022). Machine learning for energy performance prediction at the design stage of buildings. Energy for Sustainable Development, 66, 12–25. https://doi.org/10.1016/j.esd.2021.11.002

16. Pham, A.-D., Ngo, N.-T., Truong, T. T. H., Huynh, N.-T., & Truong, N.-S. (2020). Predicting energy consumption in multiple buildings using machine learning for improving energy efficiency and sustainability. Journal of Cleaner Production, 260, Article 121082. https://doi.org/10.1016/j.jclepro.2020.121082

17. Tekler, Z. D., Low, R., Zhou, Y., Yuen, C., Blessing, L., & Spanos, C. (2022). Plug-Mate: An IoT-based occupancy-driven plug load management system in smart buildings. Building and Environment, 223, Article 109472. https://doi.org/10.1016/j.buildenv.2022.109472

18. Touzani, S., Granderson, J., & Fernandes, S. (2018). Gradient boosting machine for modeling the energy consumption of commercial buildings. Energy and Buildings, 158, 1533–1543. https://doi.org/10.1016/j.enbuild.2017.11.039

19. Zhao, H., & Magoulès, F. (2012). A review on the prediction of building energy consumption. Renewable and Sustainable Energy Reviews, 16(6), 3586–3592. https://doi.org/10.1016/j.rser.2012.02.049



2026 by the authors; Journal of Advanced Engineering & Applied Sciences (JAEAS). This is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).