Islamic University Journal of Applied Sciences (IUJAS)
https://journals.iu.edu.sa/jesc

Volume VIII, Issue I, July 2026, Pages 269-279

Machine Learning Approaches for Indoor Temperature Prediction in Northern Saudi Arabia

Yazeed Basheer O. Alshammari [1, *]

[1] Department of AI & Data Science, College of Computer Science and Engineering, University of Hail, Hail, Saudi Arabia

*Corresponding author: (Yazeed Basheer O. Alshammari), Email: s202309381@uoh.edu.sa

Abstract

This paper compares ten different machine learning algorithms for estimating indoor temperatures using empirical sensor readings from a real-world environment in Hail, Saudi Arabia. The data was collected during July 2023, and includes a total sample size of 8,966 records with five minutes between each reading. Five distinct environmental factors were recorded: light level, humidity, pressure, altitude and temporal features. The machine learning algorithms evaluated were Linear Regression, Ridge Regression, Lasso Regression, Elastic Net, K-Nearest Neighbors, SVR, Decision Tree, Random Forest, Gradient Boosting, and XGBoost. Each compared using R-squared, RMSE, MSE, and correlation metrics. Random Forest algorithm has achieved the highest level of performance with an R-squared value of 0.9307 and RMSE value of 0.4370 after hyperparameter tuning compared to traditional linear algorithms which have an R-squared value of below 0.33. As well, the Feature Importance analysis showed that the atmospheric pressure and humidity were the primary predictors. These findings validate tree based ensemble methods for modeling nonlinear relationships in indoor environmental systems and demonstrate practical applicability for smart building implementations in harsh climatic conditions.

Keywords: Indoor temperature prediction; Machine learning; Ensemble methods; Environmental sensors; Smart buildings

https://doi.org/10.63070/jesc.2026.027
Received 06 February 2026; Revised 20 April 2026; Accepted 25 April 2026.
Available online 30 April 2026.
Published by Islamic University of Madinah on behalf of *Islamic University Journal of Applied Sciences.*
This is an open access article published under the Creative Commons Attribution 4.0 International license (CC BY 4.0):https://creativecommons.org/licenses/by/4.0/

1. Introduction

Building upon the technologies of smart buildings, it is seen that unprecedented opportunities exist for a data-centric approach to environmental management. Indoor temperature is a key aspect that affects energy usage, as well as the productivity of occupants [1]. Indoor temperature can therefore be predicted using machine learning models to take proactive measures in terms of maintaining a proper atmosphere. This is more significant in regions of extreme climates, like Northern Saudi Arabia, where buildings are subjected to severe thermal conditions and consume up to 70% of residential electric energy for cooling in peak summer seasons [2]. Conventional reactive control systems adapt HVAC system control based purely on current temperature deviations [3]. Such an approach leads to an uncontrolled usage of energy and discomfort during transitions. Machine learning techniques can help achieve predictive control by learning complex nonlinear relationships between environmental factors without mathematical formulations [4]. Advances in low-cost sensors and computer capabilities have made it economically and technically possible to utilize sophisticated prediction models in real-time environments [5]. Despite considerable research into machine learning in building energy prediction, knowledge gaps still persist in this field of research. Most of the existing research has been based on temperate climate zones, with insufficient consideration of extreme conditions, which prevail in the Middle Eastern countries. Similarly, comparative studies tend to be based on a small subset of algorithms rather than on broader modeling approaches. Research results are mostly based on artificial or unavailable databases that cannot be verified elsewhere. This research fills the existing gaps by comparatively and comprehensively evaluating ten machine learning algorithms using real-world data collected by the Hail city sensors during peak summer. Some of the primary objectives of this research are identifying the optimal approaches for handling temperature prediction problems under extreme climate conditions, quantifying the variations between the standard statistical approaches and machine learning techniques, determining the significance of environment sensor variables, and formulating guidelines for practitioners of building management systems.

2. Literature Review

The traditional methodologies undertaken for indoor temperature prediction involved classical statistical techniques such as ARIMA models, exponential smoothing, and multiple linear regression techniques [6]. These methods have limited forecasting accuracy, especially when environmental interactions are involved. The drawback of these methods for indoor climate predictions lies with their inability to determine interaction effects among multiple environmental parameters. Machine learning applications in building energy systems have seen considerable progress in recent times [7]. Research

has shown that artificial neural network models have shown superior efficiency for energy consumption prediction tasks and temperature prediction tasks compared with other models. Support Vector Machines have been adequately used for HVAC energy prediction tasks, although sensitivity to hyperparameters remains a concern [8]. Tree-based ensemble techniques have been identified as excellent candidates for various tasks, considering the ability of these techniques for handling nonlinear relationships, the management of heterogeneous data, lack of overfitting with trees, and interpretability of these techniques [9]. Random Forest algorithms have been recorded as the best-performing algorithms for building energy system task benchmarks [10]. Gradient Boosting algorithms, including XGBoost, have shown superior accuracy with more computing power required for these techniques compared with the other algorithms. Even with progress, gaps persist across current research. While some work emphasizes lab settings or data gathered in mild weather, real-world use in scorching, dry zones like northern Saudi Arabia still lacks support. Additionally, numerous analyses test just a few machine learning methods - typically sticking to basic linear strategies or only a couple of combined models - which weakens how thoroughly different systems can be judged. Though tree-based models like Random Forest and Gradient Boosting excel at forecasting building energy use and indoor conditions - thanks to their skill with complex patterns and mixed sensor data - their tested success often ties closely to particular locations or datasets. Because of this, moving predictions across different environments stays a challenge, despite shared methodological roots. What often gets overlooked is how uneven the application of hyperparameter adjustments remains across studies, which may introduce bias. Because of this inconsistency, one algorithm may appear stronger just by setup - not skill. A quiet problem then emerges: comparisons lose balance when tuning effort varies without acknowledgment. Without equal refinement, claims about model superiority rest on shaky ground. Fair tests demand equal care in preparation - something still missing in much of the existing research. This work tackles existing gaps by comparing ten machine learning methods within one standardized setup, relying on actual sensor readings from a hot, dry region. Such an approach supports fairer judgments about how models function, especially when weather grows harsh - conditions rarely covered well in past studies. Recent studies have demonstrated strong performance of deep learning models for indoor temperature prediction in HVAC systems [11]. Additional research has validated machine learning approaches for indoor temperature prediction across different building environments [12].

3. Methodology

3.1 Dataset Description

The dataset was collected by University of Hail faculty and made publicly available through Ploms Science Press [13]. Data collection occurred during July 2023 within a residential building in Hail, Northern Saudi Arabia. The study period corresponds to peak summer when daily maximum temperatures frequently exceed 45 degrees Celsius. Environmental measurements were captured using IoT sensors at five-minute intervals, recording timestamp, light intensity measured in lux, temperature in degrees Celsius, humidity as a percentage, atmospheric pressure in Pascals, and altitude in meters. The raw dataset comprises 9,811 records; after removing 56 observations with missing values, 8,966 complete observations remained for analysis.

3.2 Data Preprocessing

With the help of extensive preprocessing, data quality was assured. The data present in the columns was represented uniformly by removing spaces from the column identifiers, also known as variables. Similarly, the variable ‘timestamp’ was converted from its original representation of time, where each time entry is specified with hours, minutes, and seconds, to a numerical form with all hours being normalized to an integer value from 0 to 86,399, representing the total time elapsed since the start of the day, which is captured by algorithms with diurnal patterns. The target variable used is the data on temperature, while the other five variables are considered as feature variables. Scale functions are not applied on the variables as they show invariance properties on addressed data, and experiments suggested that standardization did not impact the results.

3.3 Machine Learning Models

Ten regression algorithms were compared: Linear Regression, Ridge Regression with L2 regularization, Lasso with L1 regularization, Elastic Net combining L1 and L2 regularizations, K-Nearest Neighbors, Support Vector Regression with RBF kernel, Decision Tree, Random Forest, Gradient Boosting, and XGBoost. All the algorithms in this were implemented using scikit-learn version 1.1 and higher and XGBoost version 1.7 and higher [14][15]. Initial training was done on default hyperparameters to define baseline performance. This allows a fair assessment of intrinsic model behavior prior to optimization. After baseline evaluation, hyperparameter optimization was performed for the best algorithm based on the baseline: the randomized search with fivefold cross-validation evaluating 50 configurations.

3.4 Evaluation Protocol

The dataset was split into a training set comprising 80% of observations (7,173 samples) and a testing set comprising 20% of observations (1,793 samples). To preserve the temporal structure of the sensor data and avoid potential data leakage, the dataset was sorted chronologically based on the time feature prior to splitting. The first 80% of observations were used for training, while the remaining 20% were reserved for testing. Quantification of outcomes was performed using R-squared, which estimates the proportion of variance explained, Mean Squared Error (MSE), which measures average squared error, Root Mean Squared Error (RMSE), which is interpretable in degrees Celsius, Pearson correlation, and Spearman correlation. R-squared and RMSE were considered the primary performance indicators due to their interpretability and relevance to regression accuracy.

4. Results and Analysis

4.1 Model Performance Comparison

Performance metrics show significant variations of the modelling techniques. It can be observed that the highest performance of the baseline model was achieved by the Random Forest model with an R-squared value of 0.9233 and an RMSE of 0.4599 degrees. XGBoost model ranked second, achieving an R-squared value of 0.9071 and an RMSE of 0.5061 degrees. Decision Tree model ranked third, with an R-squared of 0.8660 and an RMSE of 0.6080 degrees, while the Gradient Boosting model obtained an R-squared of 0.7941 and an RMSE of 0.7534 degrees. High Pearson correlation of 0.9609 and Spearman correlation of 0.9381 for the Random Forest model indicate very close agreement between the predicted values and the actual temperature values. The linear models showed a performance that was substantially worse than others. Ridge and Linear Regression had almost identical performance, with an R-squared of around 0.32 and an RMSE of around 1.37 degrees. Lasso and Elastic Net models performed slightly worse than this. The poor performance of linear models with an R-squared of less than 0.33 reinforces the supposition of the non-linear relationships between sensor readings and temperature. The distance-based models had the poorest performance overall, with an R-squared of 0.0561 and an RMSE of 1.6133 degrees for SVR, and an R-squared of 0.0409 and an RMSE of 1.6262 degrees for KNN, which likely resulted from the curse of dimensionality and the use of default hyperparameters. Overall, the results demonstrate clear and consistent performance differences across model families, confirming that nonlinear ensemble methods are better suited for indoor temperature prediction under real-world sensor conditions. Similar performance advantages of ensemble learning methods have been reported in recent building energy prediction studies [16]. Deep learning hybrid models have also shown strong capability in capturing temporal dependencies in indoor temperature data [17]. Figure 1 provides a visualization of the performance of the models.

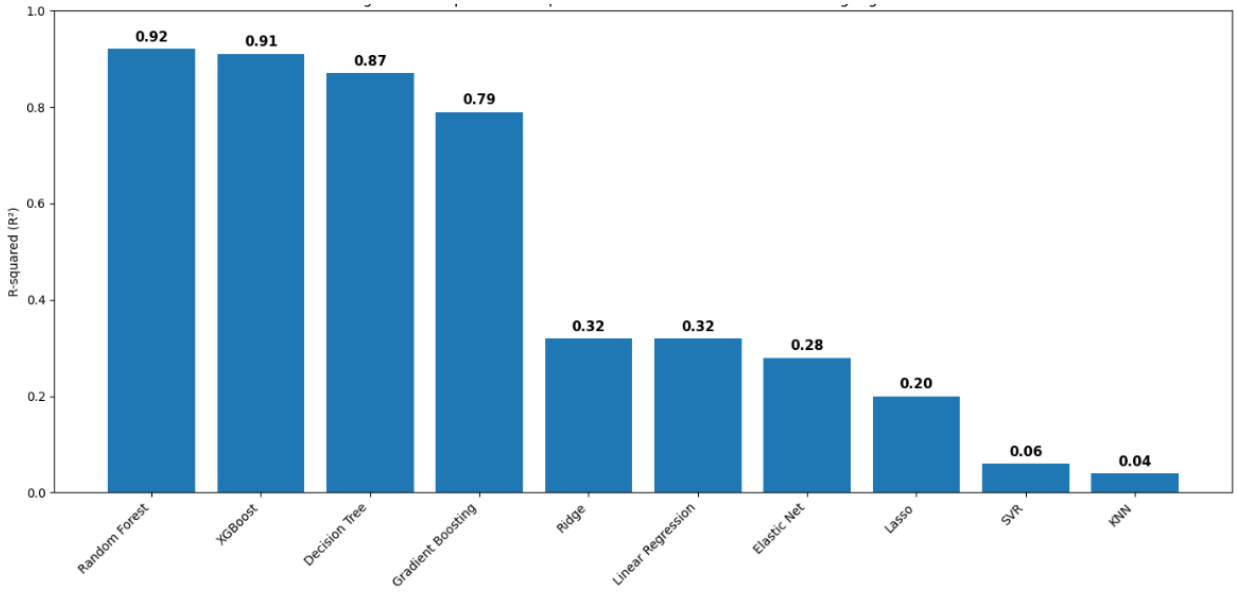


Figure 1. Comparative R-squared Performance of Machine Learning Algorithms

4.2 Hyperparameter Optimization

Comprehensive hyperparameter tuning of Random Forest using randomized search evaluated 50 configurations. Optimal configuration consisted of 300 estimators, unlimited depth, minimum 5 samples for splitting, minimum 1 sample per leaf, square root feature selection, and disabled bootstrap. This achieved cross-validation R-squared of 0.9013 and test performance R-squared of 0.9307 with RMSE of 0.4370 degrees. Optimization yielded meaningful improvements with R-squared increasing by 0.74 percentage points from 0.9233 to 0.9307, and RMSE decreasing 0.023 degrees from 0.4599 to 0.4370, representing 9.6% reduction in unexplained variance and 5% reduction in error magnitude with tangible operational benefits for building management.

4.3 Environmental Parameter Analysis

Analysis from the optimized Random Forest reveals relative contribution of environmental measurements to prediction accuracy. Atmospheric pressure emerged most influential at 38%, followed by humidity at 27%, time-of-day at 22%, light intensity at 11%, and altitude at 2%. Pressure dominance aligns with physical expectations as variations are associated with weather systems influencing outdoor conditions and building thermal loads. Humidity importance reflects the relationship with sensible heat transfer and HVAC latent cooling demands. Temporal feature captures systematic diurnal patterns from solar radiation, occupancy schedules, and HVAC modes. Light serves as a proxy for solar heat gain and occupancy-related internal heat. Minimal altitude importance is

expected for fixed sensor location. Figure 2 illustrates the relative contribution of environmental parameters.

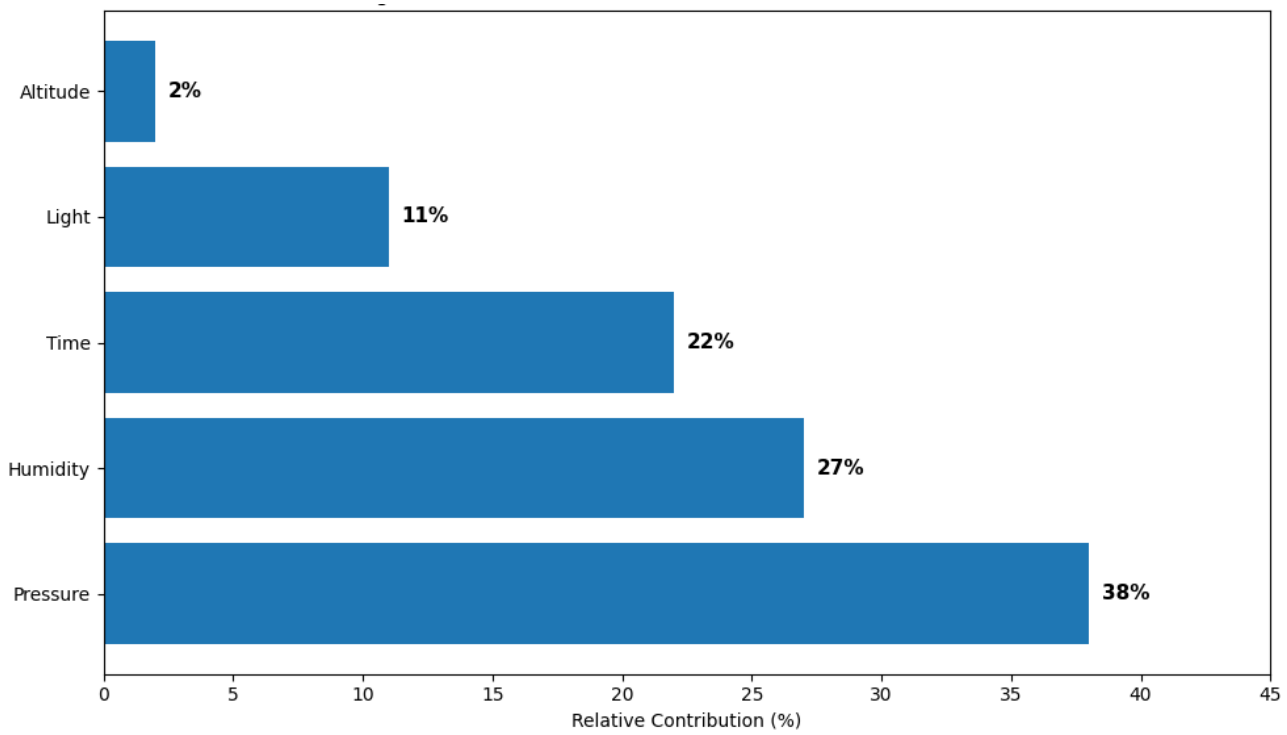


Figure 2. Relative Contribution of Environmental Parameters to Temperature Prediction

5. Discussion

5.1 Performance Interpretation

Despite their simplicity, linear approaches fall short when forecasting indoor temperatures, mainly because building thermal patterns involve complex shifts across multiple factors. Results show that ensemble techniques handle such complexity far better than traditional distance-driven or straight-line models. Nonlinear relationships - like those linking humidity, atmospheric pressure, and hour-by-hour changes - are too complex for basic methods to grasp accurately. As a result, more adaptive strategies emerge stronger in these conditions. Despite their differences, tree ensembles like Random Forest, XGBoost, and Gradient Boosting often outperform simpler models by splitting data dynamically while capturing complex patterns naturally. What distinguishes Random Forest apart is steady behavior across datasets - its strength comes from combining multiple trees built on resampled training cases, a method that smooths out errors and supports reliable predictions. Unlike more flexible approaches, linear methods - such as Linear, Ridge, Lasso, and Elastic Net - perform poorly, which suggests the process at play isn't governed by straight-line patterns. Because these models miss how environmental factors influence one another, they consistently fail to fit the data well. Performance issues with KNN and SVR reveal how strongly they react to unscaled features, tuning choices, and noisy, high-dimensional inputs. Without intense scaling or fine-tuned kernels, their results lag - ensemble

approaches handle such conditions better when left near defaults. Across different models, outcomes remain stable - a sign the method holds up under variation - while also showing ensemble techniques with nonlinear traits handle indoor environment forecasts more effectively. Results do not waver much, suggesting the setup works well, given how these predictive systems perform similarly when conditions change slightly. Predictions stay accurate because the approach adapts naturally to complex patterns found indoors, unlike simpler linear alternatives. Stability in performance implies trust in design, especially since varied setups lead to similar conclusions about forecasting ability. These findings align with broader research indicating that nonlinear models outperform linear approaches in complex environmental systems [18].

5.2 Practical Implications

The optimized Random Forest model, with an RMSE of 0.4370°C, demonstrates strong potential for supporting predictive control in HVAC systems by enabling short-term forecasting of indoor temperature trends. This allows building management systems to shift from purely reactive control to anticipatory adjustments that can improve thermal stability and operational efficiency. The model relies only on standard environmental sensors commonly available in modern building monitoring systems, eliminating the need for additional specialized hardware. This improves its feasibility for deployment in existing infrastructure. In terms of implementation, the model is computationally lightweight at inference time, making it suitable for real-time use in embedded or edge-based controllers. This supports integration into smart building systems where low-latency decision-making is required. Additionally, the use of feature importance analysis improves interpretability by identifying the most influential environmental variables, which can assist in system diagnostics and provide insight into sensor or environmental behaviour under different operating conditions. Overall, the results indicate that tree-based ensemble models can be effectively applied in smart building environments to support data-driven temperature prediction and enhance HVAC decision-making under real-world constraints. Data-driven HVAC control strategies have been widely studied for improving energy efficiency and system responsiveness [19].

5.3 Limitation and Future Work

Despite strong results, several limitations remain. The dataset is restricted to a single building over one summer month, limiting exposure to diverse weather conditions and usage patterns. This reduces the generalizability of the findings to other buildings, seasons, and climatic regions. The evaluation relies on a single chronological train–test split, which preserves temporal order but does not capture broader temporal variability. More robust approaches such as rolling or sliding window validation would better

reflect model performance over time. Model comparison is also slightly affected by uneven hyperparameter tuning, as only the Random Forest model was optimized while others used default configurations. Although baseline comparisons remain informative, this introduces potential bias in relative performance assessment. Finally, the study focuses on short-term prediction and does not evaluate integration with real HVAC control systems or energy consumption impact, leaving practical deployment outcomes untested. Future work should expand the dataset across multiple buildings and seasons, apply consistent hyperparameter optimization across all models, and adopt more rigorous time-series validation techniques. Integration with real HVAC systems would also allow evaluation of energy efficiency and operational performance in real-world conditions. Advanced forecasting techniques such as sequence-based deep learning models have shown promise for improving long term prediction accuracy [20]. Comparative studies of machine learning models further highlight the importance of robust validation in time-series prediction tasks [21].

6. Conclusion

From a residential structure in northern Saudi Arabia, sensors gathered real-world climate readings during intense summer periods, forming the foundation for assessing ten predictive algorithms. Tree ensembles like Random Forest and XGBoost lead in performance, outperforming simpler linear approaches along with proximity-driven ones. Every scoring measure aligns with this trend, hinting at complex, nonlinear interactions within indoor thermal patterns. Such outcomes suggest conventional models may lack sensitivity when handling erratic household environment shifts. Although data stems from one location, it still captures realistic indoor thermal dynamics under extreme climatic conditions. This work primarily offers a side-by-side analysis of various regression methods, tested uniformly through actual sensor readings taken in severe weather. While earlier research often examines only a few models or idealized settings, here a wider selection of algorithms undergoes assessment under identical conditions - making contrasts in results more reliable. What stands out is the coherence of testing applied across diverse approaches, something frequently absent in past efforts. Moreover, atmospheric pressure, along with humidity and time-based trends, emerges clearly as central to forecasting temperature, reinforcing their role in simulating indoor conditions. Evidence shows methods rooted in data are capable of reflecting intricate heat dynamics within built spaces. Despite these findings, conclusions face constraints due to reliance on one structure and a narrow seasonal window, limiting broader applicability. Moving ahead, research might cover diverse structures and annual cycles instead; stronger temporal analysis techniques could be applied alongside testing within actual heating and cooling operations to examine effects on power use and daily function.

Funding

The authors declare that this research received no external funding.

Data Availability Statement

Not applicable.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] H. X. Zhao and F. Magoulès, "A review on the prediction of building energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 6, pp. 3586–3592, 2012. https://doi.org/10.1016/j.rser.2012.02.049

[2] A. Gassar and S. Cha, "Assessment of current energy consumption in residential buildings in Jeddah, Saudi Arabia," *Buildings*, vol. 9, no. 7, p. 163, 2019. https://doi.org/10.3390/buildings9070163

[3] A. Afram and F. Janabi-Sharifi, "Theory and applications of HVAC control systems: A review of model predictive control (MPC)," *Building and Environment*, vol. 72, pp. 343–355, 2014. https://doi.org/10.1016/j.buildenv.2013.11.016

[4] K. Amasyali and N. M. El-Gohary, "A review of data-driven building energy consumption prediction studies," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 1192–1205, 2018. https://doi.org/10.1016/j.rser.2017.04.043

[5] C. Deb, F. Zhang, J. Yang, S. E. Lee, and K. W. Shah, "A review on time series forecasting techniques for building energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 74, pp. 902–924, 2017. https://doi.org/10.1016/j.rser.2017.02.028

[6] C. Voyant et al., "Machine learning methods for solar radiation forecasting: A review," *Renewable Energy*, vol. 105, pp. 569–582, 2017. https://doi.org/10.1016/j.renene.2017.05.038

[7] T. Ahmad et al., "Machine learning in building energy prediction: A review," *Renewable and Sustainable Energy Reviews*, vol. 144, p. 110611, 2021. https://doi.org/10.1016/j.rser.2020.110611

[8] Y. Zhang, Z. O'Neill, B. Dong, and G. Augenbroe, "Comparisons of inverse modeling approaches for predicting building energy performance," *Building and Environment*, vol. 86, pp. 177–190, 2015. https://doi.org/10.1016/j.buildenv.2015.01.021

[9] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001. https://link.springer.com/article/10.1023/A:1010933404324

[10] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.
https://projecteuclid.org/euclid.aos/1013203451

[11] R. Zhuang et al., "Deep learning-based room temperature prediction for HVAC systems," *Sensors*, vol. 23, no. 5, p. 2732, 2023.
https://doi.org/10.3390/s23052732

[12] H. Golshanfard et al., "Machine learning-based indoor temperature prediction in buildings," *Journal of Building Engineering*, vol. 52, p. 104432, 2022.
https://doi.org/10.1016/j.jobe.2022.104432

[13] University of Hail Faculty, "Temporal Variations Dataset for Indoor Environmental Parameters in Northern Saudi Arabia," 2023.
https://www.researchgate.net/publication/371764282

[14] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
https://jmlr.org/papers/v12/pedregosa11a.html

[15] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794.
https://doi.org/10.1145/2939672.2939785

[16] F. Elmaz et al., "CNN-LSTM architecture for predictive indoor temperature modeling," *Building and Environment*, vol. 206, p. 108327, 2021.
https://doi.org/10.1016/j.buildenv.2021.108327

[17] N. Norouzi et al., "Applicability of deep learning algorithms for predicting indoor temperatures: Towards the development of digital twin HVAC systems," *Buildings*, vol. 13, no. 6, p. 1542, 2023.
https://doi.org/10.3390/buildings13061542

[18] J. Wei et al., "Deep learning for building energy prediction: A review," *Applied Energy*, vol. 272, p. 114683, 2020.
https://doi.org/10.1016/j.apenergy.2019.114683

[19] C. Fan et al., "Short-term building load prediction using machine learning methods," *Energy and Buildings*, vol. 236, p. 110373, 2021.
https://doi.org/10.1016/j.enbuild.2020.110373

[20] Y. Chen et al., "Machine learning for building energy prediction: Methods and applications," *Energy*, vol. 196, p. 117560, 2020.
https://doi.org/10.1016/j.energy.2020.117560

[21] G. Li et al., "Performance evaluation of sequence-to-sequence-attention model for short-term building energy prediction," *Energy*, vol. 259, p. 124915, 2022.
https://doi.org/10.1016/j.energy.2022.124915