

*Original Article*

Intelligent Air Cooling Control in Thermal HVAC Systems Using Deep Learning

Dr. Jvalant Kumar Kanaiyalal Patel

Assistant Professor, Dept of Computer of Application, Shri Manilal Kadakia College of Commerce, Management, Science and Computer Studies, Ankleshwar.

Received On: 12/01/2025**Revised On:** 14/02/2026**Accepted On:** 18/02/2026**Published On:** 21/02/2026

Abstract - Occupancy-based control of Heating, Ventilation, and Air Conditioning (HVAC) systems with the use of machine learning is described. The common HVAC systems tend to be inefficient because they are either programmed or sensor-driven and thus they waste energy and poor the indoor environmental setup. With real-time occupancy data, HVAC systems can also adjust on a real-time basis by adjusting heating, cooling, and ventilating based on the presence or absence of occupants. This paper suggests a machine learning-based model using the UCI Energy Efficiency Dataset to model and predict performance with regard to cooling. Several models were assessed and compared based on standard regression measures, such as MLP, NN, LSSVR, and Random Forest (RF). The Random Forest model provided the highest level of accuracy and strength with an R^2 of 96.4 and very low prediction errors ($MAE = 0.01$, $MSE = 0.01$, $RMSE = 0.076$). As the outcomes show, the suggested RF-based model successfully determines the nonlinear dependencies between thermal and environmental factors, which is why it can be deemed as a sound solution to air cooling control and optimization of air conditioning systems with consideration of energy efficiency.

Keywords - Intelligent HVAC Systems, Model Predictive Control (MPC), Deep Reinforcement Learning (DRL), Residential Energy Efficiency, Thermal Comfort Optimization, Smart Building Automation, Occupancy-Aware Control, Hybrid Control Strategies, AI-Driven HVAC.

1. Introduction

Heating, Ventilation, and Air Conditioning (HVAC) [1][2] systems are essential for maintaining thermal comfort and acceptable indoor air quality in residential, commercial, and industrial buildings [3][4]. By regulating temperature, humidity, and ventilation, these systems directly affect occupant comfort, health [5][6], and productivity under varying external weather conditions. However, HVAC systems [7] account for a significant portion of building energy consumption [8], making their efficient operation critical for reducing energy use [9], operational costs, and environmental impact [10][11][12][13]. The operation of HVAC systems involves multiple interdependent processes [14], including heating [15], cooling, ventilation, humidification, and dehumidification, which are carried out using components such as compressors, condensers,

evaporators, fans, and expansion valves. Accurate sensing and monitoring of environmental conditions, refrigerant states, and system performance are necessary for reliable control. Predicting HVAC [16] cooling loads remains challenging due to nonlinear interactions among weather conditions [17], building characteristics [18], system dynamics, and occupant behavior [19][20][21].

Recent advancements in sensing technologies, cloud computing, and the Internet of Things (IoT) [22][23][24][25] have enabled real-time data collection and communication across building systems. These innovations have led to the development of smart building management systems that integrate weather data, occupancy information, and operational parameters to optimize HVAC performance. Achieving an effective balance between occupant comfort and energy efficiency is particularly important in commercial buildings, where energy consumption impacts cost and indoor comfort influences space utilization and productivity [26][27][28]. In this context, machine learning [29] and deep learning [30] techniques have emerged as powerful tools for intelligent air-cooling control in thermal HVAC systems [31]. These data-driven approaches can model complex nonlinear relationships, adapt to dynamic operating conditions, and enable predictive and adaptive control strategies. By transforming conventional HVAC systems into intelligent and self-learning systems, ML [32]- and DL-based [33][34][35] methods offer significant potential for improving energy efficiency, enhancing thermal comfort [36], and supporting sustainable building operations [37][38][39][40].

1.1. Motivation and Contribution

The increasing pressure on designing energy-efficient buildings and sustainable HVAC operation has augmented the desire that the smart air-cooling control strategies must be able to incorporate complex and dynamic indoor environments. Traditional control strategies do not always represent nonlinear relationships between thermal, occupancy, and environmental control factors, and thus result in inefficient energy consumption and lower occupant comfort. Recent developments in machine learning can offer a chance to eliminate these constraints because it allows modeling data-driven and accurately predicting cooling behavior. This research was inspired by the desire to

minimize energy use, improve thermal comfort, and assist in the management of smart buildings by discussing and analyzing the best new machine learning methods to manage air-cooling systems in thermal HVAC systems. There are some significant contributions in this study as enumerated below:

- Introduced a machine learning-based air-cooling framework for thermal HVAC systems based on the UCI Energy Efficiency Dataset.
- Did intensive data analysis and preprocessing, such as handling of missing values, noise reduction, feature normalization, and correlation analysis.
- Offered information regarding the influence of the features and robustness of the models, which justify the use of ensemble learning in the optimization of energy-conscious HVAC systems.
- Illustrated the superiority of the Random Forest model, which has the highest R^2 value with the least error in prediction.
- Compared the model performance based on the standard regression variables, which are Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and the R^2 score to examine the prediction accuracy and reliability in a comprehensive manner.

1.2. Justification and Novelty

The rationale behind the proposed study is the increasing demand of precise and information-based air-cooling control systems that can be capable of controlling the energy consumption and still provide thermal comfort in the interiors in HVAC systems. Compared to traditional rule-based methods or single-model methods, this work combines extensive data pre-processing, extensive correlation analysis and a comparative analysis of various machine learning methods to determine the most trustworthy predictive model. The originality of this work is in that, having an optimized Random Forest model to control air cooling is applied using the Energy Efficiency Dataset, and the final results are better predictivity and lower error indicators. Also, the research presents the effective performance comparison with MLP, NN and LSSVR models in terms of robustness and generalization of performance of the ensemble learning to intelligent and energy-efficient HVAC control system.

1.3. Organization of the Paper

The article is structured as follows: Section II provides the literature related to air-cooling control with machine learning in thermal HVAC systems, and Section III provides a description of the dataset, preprocessing and suggested model. Section IV gives experimental findings and comparisons and Section V ends with major findings and future research direction.

2. Literature Review

The basis of the present study was reinforced by an extensive review and critical discussion of the existing studies on air cooling control in thermal HVAC systems. Table 1 presents a review of recent research by identifying the suggested models, the dataset applied, the major findings,

and noted limitations, which will give a clear research context and rationalize the creation of the proposed solution and its validation.

Boodi and Beddiar (2025) Identify thermal parameters and to predict indoor temperature and heating/cooling loads. The hybrid PIML model is applied to a building with limited information on HVAC operation and system characteristics. Results show that the hybrid PIML model captures thermal dynamics more accurately than the physics-only model, achieving a mean absolute error (MAE) of 0.454°C and a root mean square error (RMSE) of 0.547°C , compared with 0.851°C and 0.974°C for the physicsonly model. The proposed framework is lightweight, reproducible, and compatible with model-based control applications[41].

Li *et al.* (2025) used to predict future vehicle velocity over a short horizon, facilitating precise temperature tracking and minimizing variations for optimal control. Simulation results under real driving conditions indicate that the proposed strategy reduces energy consumption by 18.90%, improves computational efficiency by 83.53%, and results in a 0.055% improvement in battery state of health (SOH) over extended cycles, without compromising the cooling requirements of the passenger cabin[42].

Prema Arokia Mary *et al.* (2025) comparing Deep Deterministic Policy Gradient (DDPG), Proximal Policy Optimization (PPO) and Twin Delay DDPG (TD3) based DRL controllers and the baseline controller, the DDPG based controller reduced energy consumption by 97.3 % and operating cost by 17.7 % during the peak heating period with reference to baseline method. Then on analyzing the impact of inclusion of forecast parameters occupancy, solar irradiance, and electricity prices over the period 3, 6 and 12 hours in DDPG based controller. The prediction for 3 hours gave the greatest reduction in thermal discomfort of 99.7 % and prediction for 12 hours gave maximum reduction in cost by 30.4 % but resulted in only 82% reduction in thermal comfort when compared with baseline method indicating that longer prediction horizon is not necessarily results in better performance[43].

Singh, Das and Singh (2024) considered as inputs to the ANN. A total of 2591 input and output datasets were used, of which 80 % data were consumed for training, whereas 10% each were involved for validating and testing. The proposed ANN offered prediction of the cooling capacity with correlation coefficient of above 90 %. However, due to more irregular pattern of the heating load, its prediction accuracy was lower for this case[44].

Wange *et al.* (2024) uses smart technologies to design energy-efficient buildings. Model accuracy is improved by variational dequantization for discrete features and mode-specific normalisation for continuous features. The model was trained for greatest energy savings using RL-GAN. The model had 91.48% accuracy, proving that the proposed method reduces energy use while retaining comfort. This

study offers practical advice for creating energy-efficient, occupant-centred smart buildings[45].

Domingues, Barbosa and Bartolomeu (2023) The objective of this work is to develop an intelligent control system using ML techniques to optimize costs in HVAC systems encompassing a Heat Pump (HP). The study aims to create a tool that generates a suitable Time Program (TP) based on observed operating conditions and user needs, a model for predicting HP consumption based on different operating temperatures, and an intelligent control of the HP heating curve to minimize operational costs. Results show that implementing the proposed ML-based TP optimization leads to significant energy consumption reduction in the HP, with an average reduction of about 16.3%. The algorithm

takes into account various factors, such as adjustments to the SP temperature, to optimize energy efficiency based on outdoor temperatures and user preferences [46].

Liu *et al.* (2022) deployment with very few training steps compared with online methods. algorithm incorporates a Kullback-Leibler (KL) regularization term to penalize policies that deviate far from the previous ones. They evaluate framework on a real multi-zone, multi-floor building-it achieves 7.2% in energy reduction cf. the state-of-the-art batch RL method, and outperforms other BRL methods in occu-pants' thermal comfort, and 16.7% energy reduction compared to the default rule-based control[47].

Table 1: Recent Studies on Air Cooling Control in Thermal HVAC Systems using Machine Learning Techniques

Author	Proposed Work	Results	Key Findings	Limitations & Future Work
Boodi & Beddiar (2025)	Developed a hybrid Physics-Informed Machine Learning (PIML) model to identify thermal parameters and predict indoor temperature and HVAC heating/cooling loads under limited HVAC system information.	Hybrid PIML achieved MAE = 0.454°C and RMSE = 0.547°C, compared to physics-only model MAE = 0.851°C and RMSE = 0.974°C.	Hybrid PIML significantly improves prediction accuracy of thermal dynamics. Framework is lightweight, reproducible, and compatible with model-based control.	Requires validation for complex multi-zone buildings and diverse climatic conditions. Performance may depend on data quality.
Li et al. (2025)	Proposed short-horizon vehicle velocity prediction to improve HVAC control in electric vehicles for precise temperature tracking.	18.90% reduction in energy consumption; 83.53% improvement in computational efficiency; 0.055% improvement in battery SOH.	Predictive velocity control enhances energy efficiency without affecting cabin cooling performance.	Limited to short-horizon prediction; future work could evaluate long-term real-world deployment and robustness.
Prema Arokia Mary et al. (2025)	Compared DDPG, PPO, and TD3-based DRL HVAC controllers including forecast parameters (occupancy, solar irradiance, electricity prices) over 3, 6, and 12-hour horizons.	DDPG reduced energy use by 97.3% and cost by 17.7% (peak heating). 3-hour horizon reduced discomfort by 99.7%. 12-hour horizon reduced cost by 30.4% but lowered comfort improvement to 82%.	DDPG outperformed PPO and TD3. Longer prediction horizons do not necessarily improve overall performance due to comfort–cost trade-off.	Multi-objective optimization needed to better balance cost and comfort. Real-time implementation complexity remains.
Singh, Das & Singh (2024)	Developed ANN model using 2591 datasets (80% training, 10% validation, 10% testing) to predict cooling and heating capacity.	Cooling capacity prediction achieved correlation coefficient >90%. Lower accuracy observed for heating load prediction.	ANN effectively models cooling loads; heating loads are more irregular and harder to predict.	Future work may explore hybrid/ensemble models to improve heating load prediction accuracy.
Wange et al. (2024)	Proposed RL-GAN model for smart building energy optimization using variational dequantization and mode-specific normalization techniques.	Achieved 91.48% model accuracy with improved energy savings while maintaining occupant comfort.	Integration of RL and GAN improves energy efficiency and occupant-centered smart building design.	Computational complexity and real-time scalability challenges need further investigation.
Domingues, Barbosa & Bartolomeu (2023)	Developed ML-based intelligent HVAC control system for Heat Pump (HP), including Time Program (TP) optimization, HP consumption prediction,	Achieved average 16.3% reduction in HP energy consumption.	Adaptive TP and heating curve control optimize operational cost considering outdoor temperature and user preferences.	Further validation across different climates and integration with renewable systems recommended.

	and adaptive heating curve control.			
Liu et al. (2022)	Proposed Batch Reinforcement Learning (BRL) framework with KL-regularization for multi-zone, multi-floor building HVAC control.	7.2% energy reduction vs state-of-the-art BRL; 16.7% reduction vs rule-based control; improved thermal comfort.	KL-regularization enhances policy stability and reduces training requirements. Effective for real building deployment.	Needs testing under highly dynamic occupancy scenarios and large-scale building environments.

Research gaps: Although advancement has been made on machine learning-based air-cooling control of HVAC systems, current research is usually based on complex deep reinforcement learning or hybrid models, which require a high computation rate and a lot of understanding of the system at hand. Numerous methods are weak in terms of their use with small or extrapolated data sets and are seldom tested under a variety of building conditions. Moreover, feature influence and model interpretability are also commonly disregarded, which decreases the practical deployment capability. The existence of these gaps shows that it is crucial to have a simple, precise and interpretable machine learning model with good generalization properties to facilitate efficient air-cooling control.

3. Research Methodology

This paper uses the UCI Energy Efficiency Dataset that comprises 768 simulated building samples that contain eight input features to come up with an air-cooling control model of thermal HVAC systems. Bar plots and correlation heatmaps were used to perform exploratory data analysis to get an understanding of how data was distributed and how features related to each other. StandardScaler was used to preprocess the dataset based on the treatment of missing values, reduction of noise, time synchronization, feature engineering, and normalization. The obtained processed data were divided into training and testing sets in the proportion of 80:20 and modeled with the help of a Random Forest regression algorithm. R², MAE, MSE, and RMSE were used to determine the level of prediction and strength across models. Fig.1 presents the proposed flowchart of the Thermal HVAC System with machine learning as Air-Cooling Control.

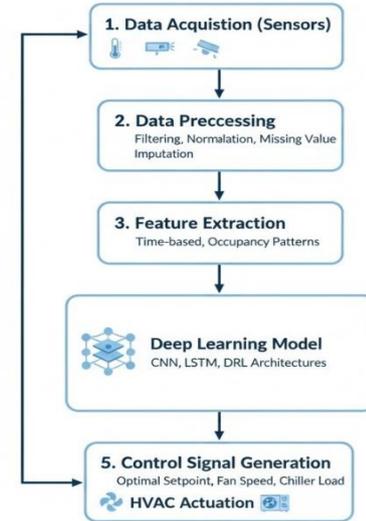


Fig 1: Proposed flowchart for Air Cooling Control in Thermal HVAC Systems using machine learning

The next paragraph is going to give a comprehensive description of every step in the proposed methodology:

3.1. Data Gathering and Analysis

The dataset that was used in the research was the Energy Efficiency Dataset of the UCI Repository. It contains 768 building samples, and the samples of a building are described by eight features. In Ecotect, to obtain 768 combinations, 12 building shapes are simulated. Data visualization in the form of bar plots and heatmaps was analyzed to display the distribution, feature correlations, etc., as shown below:

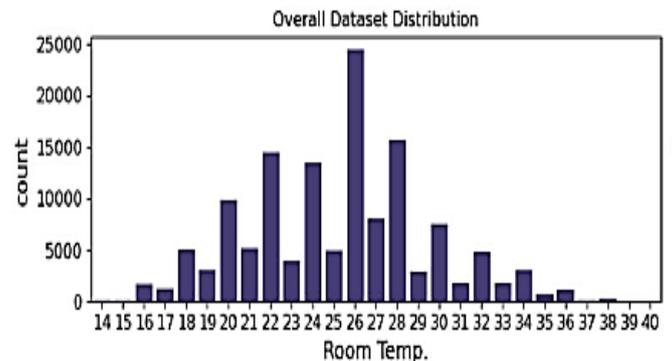


Fig 2: Data Distribution on the Energy Efficiency Dataset

The distribution in figure 2 is unimodal with the majority of values falling between 22 and 31 with the highest values around 27-28. Trends on both ends are decreasing, which means that there are fewer observations on either side.

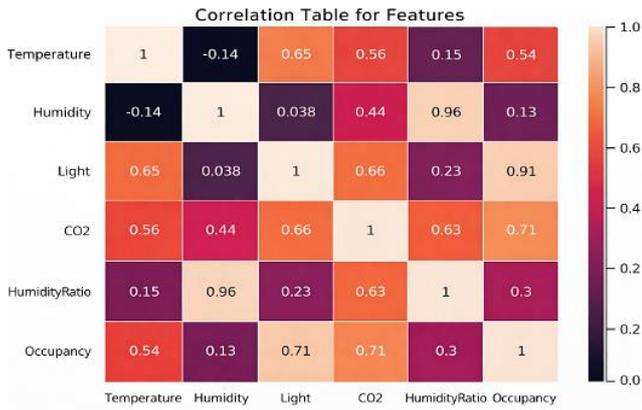


Fig 3: Correlation Matrix Heatmap on Energy Efficiency Dataset for Air Cooling Control in Thermal HVAC Systems

Figure 3 presents a heatmap of correlation in which the strong positive relationships are found between HumidityHumidityRatio (0.96), LightOccupancy (0.91), and CO2Occupancy (0.71). The correlation between Temperature and Light and CO2 is moderate and Humidity has weak correlations with most other features.

3.2. Data Pre-processing

The data was the Energy Efficiency, and the data was prepared using concatenation, cleaning, and feature engineering. The processing consisted of missing value processing, time processing, noise removal, data and normalization labeling. The main preprocessing processes are as follows:

- **Treatment of Missing Values:** The missing value processing was done to maintain data quality and guarantee a good model performance. Incomplete records were also performed with the help of proper imputation methods to enter the values that have not been available. This method limited the loss of information and minimized the possibility of bias in the data.
- **Time Synchronization:** Time synchronization assures that records of all data are synchronized at a specific time point. This procedure enhances the consistency of time-dependent features and proper analysis and modelling.
- **Noise Reduction:** Noise reduction refers to the act of reducing unwanted changes and variations in the data. It assists in enhancing signal quality, resulting in a better analysis and model projections.

3.3. Normalization using standerscaler()

Their input features have different scales thus, by the use of the StandardScaler () method, the dataset has been standardized. This step normalizes the data to make the data features mean zero and standard deviation equal to one. It is done by calculating the difference between each observation and the mean of the feature and dividing it with the standard deviation as shown in Equation (1) below:

$$z = \frac{x - \mu}{\sigma} \dots \dots (1)$$

Where z is the transformed value of the feature, x is the original value of each descriptor, μ is the mean, and σ is the standard deviation of the feature in the dataset.

3.4. Feature importance

Feature importance is a method of estimating the contribution of a single input variable on the prediction of a machine learning model. It is used to determine which features produce the most bearing on the output and those that produce least bearing on the output. It ranks features giving us an insight into the decision of the model. This enhances interpretability and feature selection and model optimization.

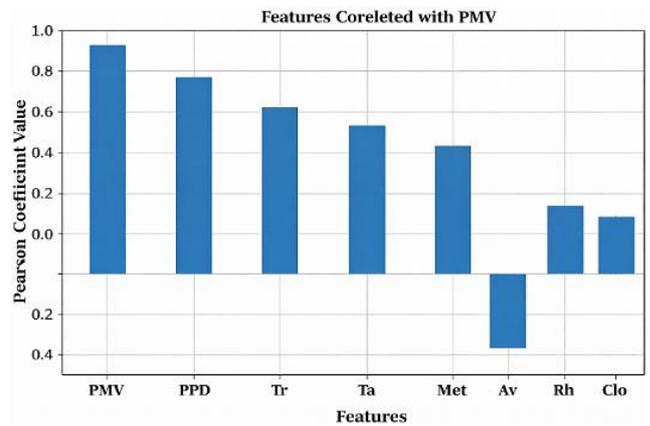


Fig 4: Feature Importance

Figure 4 indicates the Pearson correlation of features with PMV, PMV, PPD, Tr, and Ta have strong positive correlation, Met have moderate positive correlation, Rh and Clo weak positive correlation and Av negative correlation.

3.5. Data Splitting

An 80% and 20% split are used to obtain the training and testing sets respectively. This method maintains the same distribution of classes in the two subsets.

3.6. Proposed Random Forest Model

This work proposes a supervised machine learning-based model, Random Forest, for Air Cooling Control in Thermal HVAC Systems. Random Forest is an ensemble machine learning algorithm that builds and combines multiple decision trees to improve classification or regression performance. Each tree is trained on a random subset of the training data (bagging) and uses a random subset of features at each split, which increases diversity among the trees. The final prediction is made by majority voting (for classification) or averaging (for regression) across all trees. The data is recursively split into partitions. At a particular node, the split is done by asking a question on an attribute. The choice for the splitting criterion is based on some impurity measures such as Shannon Entropy or Gini impurity. Gini impurity is used as the function to measure the quality of split in each node. Gini impurity at node N is given by equation (2).

$$g(N) = \sum_{i \neq j} P(w_i)P(w_j) \dots \dots (2)$$

Where $P(w_i)$ is the proportion of the population with class label i . Another function which can be used to judge the quality of split is Shannon Entropy. It measures the disorder in the information content. In Decision trees, Shannon entropy is used to measure the unpredictability in the information contained in a particular node of a tree. The entropy in a node N can be calculated as follows in equation (3):

$$H(N) = \sum_{i=1}^{i=d} P(w_i) \log_2(Pw_i) \dots \dots \dots (3)$$

Where d is the number of classes considered and $P(w_i)$ is the proportion of the population labeled as i . Entropy is the highest when all the classes are contained in equal proportion in the node. It is the lowest when there is only one class present in a node (when the node is pure). The obvious heuristic approach to choose the best splitting decision at a node is the one that reduces the impurity as much as possible. In other words, the best split is characterized by the highest gain in information or the highest reduction in impurity. The information gain due to a split can be calculated as follows in equation (4):

$$\Delta I(N) = I(N) - P_L * I(N_L) - P_R * I(N_R) \dots \dots (4)$$

Where $I(N)$ is the impurity measure (Gini or Shannon Entropy) of node N , P_L is the proportion of the population in node N that goes to the left child of N after the split and similarly, P_R is the proportion of the population in node N that goes to the right child after the split. N_L and N_R are the left and right child of N , respectively.

3.7. Evaluation metrics

In machine learning, model evaluation is required to attain a measure of the level of agreement between predictions and real results. Four measures, R^2 , MAE, RMSE, and MSE, were used to measure regression performance in this research.

3.7.1. R-Squared

R^2 is a measure of how well the regression model fits the data. The value of R^2 ranges from 0 to 1, with higher values indicating a better fit of the model to the data. An R^2 value of 0 means that the model explains none of the variability of the response data around its mean, while an R^2 value of 1 indicates that the model perfectly predicts the response data. The formula to calculate R^2 is mentioned in equation (5):

$$R^2 = \frac{\sum_{i=1}^n (y_i - y_i^p)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \dots \dots (5)$$

3.7.2. MAE (Mean Absolute Error)

MAE is a commonly used metric for evaluating the accuracy of a predictive model. It measures the average magnitude of errors in a set of predictions, without considering their direction. A lower MAE value suggests better performance. The formula to calculate MAE is given in equation (6):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - y_i^p| \dots (6)$$

Where,

Y is the actual value,

\hat{Y} is the predicted value, and n is the number of observations.

3.7.3. MSE (Mean Squared Error)

The MSE measures the average squared difference between the predicted values and the actual values. In the context of Real Estate Prices prediction, if y_i represents the actual price at the time i and y_i^p represents the predicted price at time, then the MSE is calculated in following equation (7):

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - y_i^p)^2 \dots \dots (7)$$

Where:

n is the total number of observations.

y_i is the actual Real Estate Prices at time i .

y_i^p is the predicted Real Estate Prices at time i .

3.7.4. RMSE (Root Mean Squared Error)

This metric represents the square root of MSE. RMSE quantifies how far the model's predictions deviate from the actual values. Lower RMSE values indicate better model performance. The formula to calculate RMSE is shown in equation (8):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y_i^p)^2} \dots \dots (8)$$

4. Results and Discussion

This part shows the experimental setup along with performance determination of proposed Random Forest (RF) model. Tests on a Linux, Python and scikit-learn used Energy Efficiency data set. RF model had high R^2 of 96.4 percent and extremely low MAE (0.01), MSE (0.01) and RMSE (0.076) indicating high predictive power and efficiency of the model in air cooling control in HVAC.

Table 2: Classification Results of the Proposed RF Model, Air Cooling Control in Thermal HVAC Systems Using Energy Efficiency Dataset

Matrix	RF
R^2	96.4
MAE	0.01
MSE	0.01
RMSE	0.076

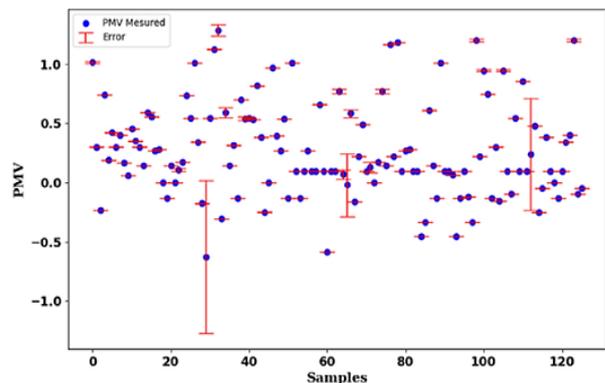


Fig 5: Scatter Plot of Predicted and Actual Values with the Error for the Proposed Model

In Figure 5, the sample values of PMV are mainly within the neutral range of comfort which means that the thermal conditions are stable. The standard deviations are typically low, and there are few cases when samples have greater uncertainty in prediction.

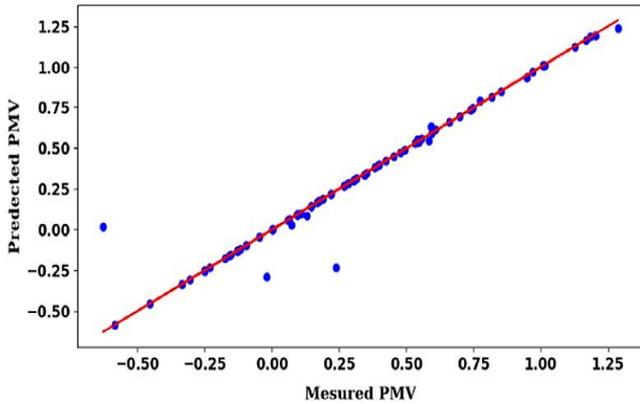


Fig 6: Regression Analysis of the RF Model

The correlation between measured and predicted PMV values illustrated in Figure 6 is linear and most of the data points are near the reference line implying that the accuracy of the prediction is high with a slight deviation.

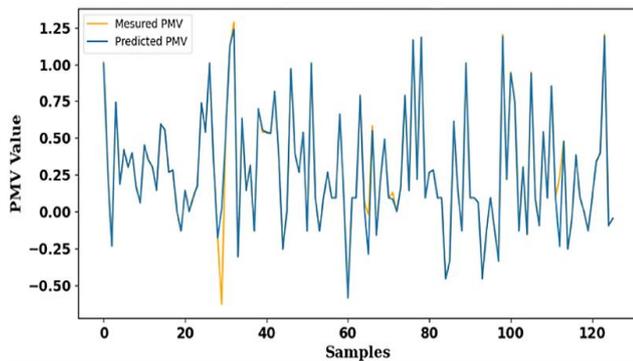


Fig 7: Real and predicted values of the RF model

Figure 7 is the line graph of measured and predicted PMV, which depicts a close resemblance in the manner of trends in all the samples. Minor differences refer to minor errors in prediction and overall consistency assures that the model was reliable in modeling changes in thermal comfort.

4.1. Comparative analysis

Table III presents the comparison of the predictive performance of different models on HVAC air cooling control in terms of Energy Efficiency dataset. The performance of MLP is moderate ($R^2 = 77.7$) and NN is a bit more accurate ($R^2 = 79.5$). LSSVR achieves good results in small error values and high R^2 of 95.7. All other RF models have lower predictive power because the proposed RF model has the lowest MAE (0.01), MSE, and RMSE (0.076), and the highest R^2 (96.4%).

Table 3: Comparison of Different Machine learning and deep learning Models for Air Cooling Control in Thermal HVAC Systems

Model	MAE	MSE	RMSE	R^2
MLP[48]	21.83	3.475	4.672	77.7
NN[49]	-	0.4731	-	79.5
LSSVR[50]	1.583	-	2.008	95.7
RF	0.01	0.01	0.076	96.4

The random forest (RF) model proposed exhibits a strong strength in the high value of R^2 , which is 96.4, which implies excellent prediction and actual values. The high level of determinism is justified by the fact that this model has a great ability to reflect complex nonlinear links in air cooling control of thermal HVAC systems. Also, the ensemble feature of RF enhances robustness and minimizes overfitting, which results in reliable and accurate predictions. All in all, the high-quality R^2 performance shows that the proposed RF approach is effective in comparison with other models.

5. Conclusion and future study

The current trend in building energy consumption around the globe is increasing, specifically in heating, ventilation, and air conditioning (HVAC), and thus it is of interest to optimize HVAC energy consumption. HP systems are very popular in HVAC as they are a major thermal system capable of reusing waste heat and increasing energy efficiency. Nevertheless, the majority of existing studies are concerned with the local optimization of equipment within the computer room, and regarding the optimization techniques, deep reinforcement learning may be used to resolve the issue that deep reinforcement learning cannot converge fast. Thus, based on the final environment, the comparative analysis of the various machine learning models proves that the ensemble-based and kernel-based models are better than the traditional neural network in controlling air cooling in thermal HVAC systems. Whereas the MLP and the NN models were moderate in their predictive power with an R^2 of 77.7 and 79.5, respectively, the LSSVR model displayed a significant enhancement in the R^2 of 95.7, indicating that the model has a great ability to predict nonlinear relationships. Random Forest model was the most accurate with R^2 of 96.4, which is the best in robustness, generalization capacity, and feature complex interaction. These findings verify that the Random Forest is the most appropriate model of those tested to be able to apply it to energy efficiency prediction and air-cooling control problems. The proposed research will be followed by future studies that will verify the proposed model on real-time HVAC operational data of different building settings. Also, the model will be combined with adaptive control or smart building systems based on IoT and the prospects of their implementation in the real world will be discussed.

References

[1] P. B. Patel and S. Kumar Singh, "Efficient Energy Management Through the Integration of HVAC and Power Distribution Systems," in *2025 4th International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, IEEE, Sep. 2025, pp. 85–92. doi:

- 10.1109/ICIMIA67127.2025.11200946.
- [2] P. B. Patel and S. Deshmukh, "Development of Liquid Cooling Solutions for Thermal Management Systems," in *2025 3rd International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, IEEE, Aug. 2025, pp. 2043–2049. doi: 10.1109/ICSCDS65426.2025.11166733.
- [3] P. B. Patel, "Modeling and Analysis of Power Distribution Networks with HVAC Load Management," *TIJER - Int. Res. Journals*, vol. 9, no. 10, pp. 27–33, 2022.
- [4] P. B. Patel and S. Darshan, "Analysis and Modelling of Heat Pump Performance using a Rolling-Piston Compressor," in *2025 3rd International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, IEEE, Aug. 2025, pp. 1–7. doi: 10.1109/ICSCDS65426.2025.11167589.
- [5] M. R. Anand, "Transforming Energy-Intensive Smart Factories with AI: TCN-based Forecasting and DQN-Driven Operational Optimization for Healthcare Manufacturing," *Int. Conf. Intell. Comput. Inf. Control Syst.*, pp. 508–515, 2025.
- [6] C. Tayal, "Secure Data Integration Frameworks for Omni-channel Healthcare Marketing Systems," *Int. J. Eng. Res. Technol.*, vol. 14, no. 10, 2025.
- [7] Pritesh B Patel, "Energy Consumption Forecasting and Optimization in Smart HVAC Systems Using Deep Learning," *Int. J. Adv. Res. Sci. Commun. Technol.*, pp. 780–788, Jun. 2024, doi: 10.48175/IJARST-18991.
- [8] Rutvik Patel, "Sustainability and Energy Management: Trends and Technologies for a Greener Industrial Future," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 4, no. 1, pp. 886–898, Jul. 2024, doi: 10.48175/IJARST-19200E.
- [9] R. Patel, "Machine Learning (ML)-Via Power Usage Effectiveness (PUE) Estimation for Energy-Efficient Cloud Data Centers," in *2025 3rd World Conference on Communication & Computing (WCONF)*, IEEE, Jul. 2025, pp. 1–6. doi: 10.1109/WCONF64849.2025.11233352.
- [10] W. W. Che *et al.*, "Energy consumption, indoor thermal comfort and air quality in a commercial office with retrofitted heat, ventilation and air conditioning (HVAC) system," *Energy Build.*, 2019, doi: 10.1016/j.enbuild.2019.06.029.
- [11] C. Kanthila, A. Boodi, K. Beddiar, Y. Amirat, and M. Benbouzid, "Building Occupancy Behavior and Prediction Methods: A Critical Review and Challenging Locks," 2021. doi: 10.1109/ACCESS.2021.3083534.
- [12] X. Cao, X. Dai, and J. Liu, "Building energy-consumption status worldwide and the state-of-the-art technologies for zero-energy buildings during the past decade," *Energy Build.*, 2016, doi: 10.1016/j.enbuild.2016.06.089.
- [13] Xiaodong Cao, Xilei Dai, and Junjie Liu, "Building energy-consumption status worldwide and the state-of-the-art technologies for zero-energy buildings during the past decade," *Energy Build.*, 2016.
- [14] P. B. Patel, "Comparative Study of Liquid Cooling vs. Air Cooling in Thermal Management," *Int. J. Res. Anal. Rev.*, vol. 8, no. 3, pp. 1–9, 2021.
- [15] R. Patel, "Unlocking the Probable of Waste Heat Retrieval in Data Centers: Industry Insights and Adoption Trends," *Int. J. Eng. Sci. Math.*, vol. 14, no. 05, pp. 130–136, 2024.
- [16] R. Patel and R. Tandon, "Advancements in Data Center Engineering: Optimizing Thermal Management, HVAC Systems, and Structural Reliability," *Int. J. Res. Anal. Rev.*, 2021.
- [17] P. B. Patel, "Predictive Maintenance in HVAC Systems Using Machine Learning Algorithms: A Comparative Study," *Int. J. Eng. Sci. Math.*, vol. 9, no. 11, pp. 118–125, 2019.
- [18] A. N. Naveen Kolli, John Wesley Sajja, "Building Secure AI Agents for Autonomous Data Access in Compliance/Regulatory-Critical Environments," *Comput. Fraud Secur.*, vol. 2024, no. 9, pp. 363–373, 2024, doi: 10.52710/cfs.746.
- [19] C. C. Cheng and D. Lee, "Enabling smart air conditioning by sensor development: A review," 2016. doi: 10.3390/s16122028.
- [20] J. C. P. Cheng *et al.*, "applied sciences Artificial Intelligence for Predictive Maintenance Applications ;," *J. Build. Eng.*, 2024.
- [21] S. Wang, J. Xing, Z. Jiang, and Y. Dai, "A novel sensors fault detection and self-correction method for HVAC systems using decentralized swarm intelligence algorithm," *Int. J. Refrig.*, 2019, doi: 10.1016/j.ijrefrig.2019.06.007.
- [22] S. Garg, "Next-Gen Smart City Operations with AIOps & IoT: A Comprehensive look at Optimizing Urban Infrastructure," *J. Adv. Dev. Res.*, vol. 12, no. 1, 2021.
- [23] S. Garg, "Next-Gen Smart City Operations with AIOps & IoT: A Comprehensive Look at Optimizing Urban Infrastructure," *SSRN Electron. J.*, vol. 12, no. 1, 2025, doi: 10.2139/ssrn.5271046.
- [24] V. Shah, "Analyzing Traffic Behavior in IoT-Cloud Systems: A Review of Analytical Frameworks," *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, vol. 9, no. 3, pp. 877–885, 2023.
- [25] V. Shah, "Networking Challenges in IoT to Deployment of TCP / IP to 6LoWPAN for Next Gen Network System," *KOS J. Sci. Eng.*, vol. 1, no. 1, pp. 1–8, 2024.
- [26] UU Republik Indonesia *et al.*, "PENENTUAN ALTERNATIF LOKASI TEMPAT PEMBUANGAN AKHIR (TPA) SAMPAH DI KABUPATEN SIDOARJO," *Energies*, 2022.
- [27] N. A. Sulaiman, M. P. Abdullah, H. Abdullah, M. N. S. Zainudin, and A. M. Yusop, "Fault detection for air conditioning system using machine learning," *IAES Int. J. Artif. Intell.*, 2020, doi: 10.11591/ijai.v9.i1.pp109-116.
- [28] H. H. Hosamo, P. R. Svennevig, K. Svidt, D. Han, and H. K. Nielsen, "A Digital Twin predictive maintenance framework of air handling units based on automatic fault detection and diagnostics," *Energy Build.*, 2022, doi: 10.1016/j.enbuild.2022.111988.
- [29] R. Patel, "Optimizing Oil and Gas Pipeline Monitoring with AI Machine Learning Based Predictive Framework," in *2025 International Conference on*

- Smart & Sustainable Technology (INCSST), IEEE, Jul. 2025, pp. 1–6. doi: 10.1109/INCSST64791.2025.11210432.
- [30] R. Patel, “Automated Threat Detection and Risk Mitigation for ICS (Industrial Control Systems) Employing Deep Learning in Cybersecurity Defence,” *Int. J. Curr. Eng. Technol.*, vol. 13, no. 06, Dec. 2023, doi: 10.14741/ijcet/v.13.6.11.
- [31] R. Patel, “Predictive Analytics for Assessing the Efficiency of Ground-Source Heat Pumps under Thermal Interference Based on Machine Learning Models,” in *2025 International Conference on Smart & Sustainable Technology (INCSST)*, IEEE, Jul. 2025, pp. 1–6. doi: 10.1109/INCSST64791.2025.11210249.
- [32] V. Thakran, “Utilization of Machine Learning Algorithms in Optimizing Finite Element Modeling and Analysis,” in *2025 International Conference on Smart & Sustainable Technology (INCSST)*, IEEE, Jul. 2025, pp. 1–6. doi: 10.1109/INCSST64791.2025.11210352.
- [33] R. Patel and P. Patel, “Machine Learning-Driven Predictive Maintenance for Early Fault Prediction and Detection in Smart Manufacturing Systems,” *ESP J. Eng. Technol. Adv.*, vol. 4, no. 1, pp. 141–149, 2024, doi: 10.56472/25832646/JETA-V4I1P120.
- [34] R. Patel, “Artificial Intelligence-Powered Optimization of Industrial IoT Networks Using Python-Based Machine Learning,” *ESP J. Eng. Technol. Adv.*, vol. 3, no. 4, pp. 138–148, 2023, doi: 10.56472/25832646/JETA-V3I8P116.
- [35] V. Thakran, “An Analysis of Machine Learning Solutions for Precise Forecasting of Oil and Gas Pipeline,” in *2025 International Conference on Intelligent Computing and Knowledge Extraction (ICICKE)*, IEEE, Jun. 2025, pp. 1–6. doi: 10.1109/ICICKE65317.2025.11136639.
- [36] P. B. Patel, “Thermal Efficiency and Design Considerations in Liquid Cooling Systems,” *Int. J. Eng. Sci. Math.*, vol. 9, no. 11, pp. 118–125, 2019.
- [37] B. Amangeldy, N. Tasmurzayev, T. Imankulov, B. Imanbek, W. Wójcik, and Y. Nurakhov, “A Hybrid Machine Learning Approach for High-Accuracy Energy Consumption Prediction Using Indoor Environmental Quality Sensors,” *Energies*, 2025, doi: 10.3390/en18154164.
- [38] A. M. Zaki *et al.*, “Environmental and energy performance analyses of HVAC systems in office buildings using boosted ensembled regression trees: Machine learning strategy for energy saving of air conditioning and lighting facilities,” *Process Saf. Environ. Prot.*, 2025, doi: 10.1016/j.psep.2025.107214.
- [39] R. Agrawal, T. Imieliński, and A. Swami, “Mining Association Rules Between Sets of Items in Large Databases,” *ACM SIGMOD Rec.*, 1993, doi: 10.1145/170036.170072.
- [40] R. Agrawal, T. Imieliński, and A. Swami, “Mining association rules between sets of items in large databases,” 1993. doi: 10.1145/170035.170072.
- [41] A. Boodi and K. Beddiar, “A Hybrid Approach to Building Thermal Modeling Using Physics-Based Machine Learning,” 2025. doi: 10.1109/uemcon67449.2025.11267566.
- [42] J. Li, Y. Ma, S. Yu, J. Gao, and H. Chen, “Multifaceted Velocity Prediction-Based Bipartite Integration Optimization Strategy of Cabin and Battery Thermal Management for EVs,” *IEEE Trans. Intell. Transp. Syst.*, 2025, doi: 10.1109/TITS.2025.3604838.
- [43] G. Prema Arokia Mary, M. S. Hema, S. Bama, and N. Guptha, “Comparing Deep Reinforcement Learning Methods for HVAC System Optimization with Forecast Parameters Over Various Time Periods,” in *2025 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation, IATMSI 2025*, 2025. doi: 10.1109/IATMSI64286.2025.10984746.
- [44] G. Singh, R. Das, and A. Singh, “Application of Artificial Neural Network in Renewable Energy-Based Building Cooling Systems,” in *Proceedings of the 3rd IEEE International Conference on Power Electronics, Intelligent Control and Energy Systems, ICPEICES 2024*, 2024. doi: 10.1109/ICPEICES62430.2024.10719162.
- [45] N. J. Wange, R. V. S. Praveen, S. K. F. Babavali, P. Rao, P. Gamini, and D. Ramasamy, “Optimizing Thermal Comfort and Energy Efficiency in Smart Buildings using GANs and Reinforcement Learning Models,” in *4th IEEE International Conference on Mobile Networks and Wireless Communications, ICMNWC 2024*, 2024. doi: 10.1109/ICMNWC63764.2024.10872151.
- [46] A. Domingues, A. Barbosa, and P. C. Bartolomeu, “Heat Pump Energy Consumption Optimization Using Machine Learning,” in *IECON 2023- 49th Annual Conference of the IEEE Industrial Electronics Society*, IEEE, Oct. 2023, pp. 1–7. doi: 10.1109/IECON51785.2023.10311953.
- [47] H. Y. Liu, B. Balaji, S. Gao, R. Gupta, and D. Hong, “Safe HVAC Control via Batch Reinforcement Learning,” in *Proceedings - 13th ACM/IEEE International Conference on Cyber-Physical Systems, ICCPS 2022*, 2022. doi: 10.1109/ICCPS54341.2022.00023.
- [48] R. Chaganti *et al.*, “Building Heating and Cooling Load Prediction Using Ensemble Machine Learning Model,” *Sensors*, 2022, doi: 10.3390/s22197692.
- [49] J. Mao, R. Grammenos, and K. Karagiannis, “Data analysis and interpretable machine learning for HVAC predictive control: A case-study based implementation,” *Sci. Technol. Built Environ.*, 2023, doi: 10.1080/23744731.2023.2239081.
- [50] F. Li, L. Li, and F. You, “Enhancing Energy Efficiency In Cooling Systems Through Advanced Machine Learning And Meta-Heuristic Algorithms For Precise Cooling Load Prediction,” *J. Appl. Sci. Eng.*, 2025, doi: 10.6180/jase.202506_28(6).0010.