



RESEARCH ARTICLE

# A Framework for Environment Thermal Comfort Prediction Model

Josephine Theresa S\*

## Abstract

In recent years, the integration of Internet of Things (IoT) technologies across diverse domains has accelerated efforts toward real-time environmental monitoring. Ensuring a responsive and adaptive ecosystem is essential for maintaining optimal living and working conditions. IoT-enabled sensors autonomously gather and transmit environmental data, facilitating the classification and analysis of ambient conditions. The pervasive role of IoT lies in its ability to seamlessly interconnect devices and enable dynamic data exchange across systems. This study investigates the evaluation of thermal comfort levels through advanced classification techniques. A machine learning framework is employed to train and validate predictive models using a comprehensive benchmark dataset comprising 100,000 samples, each reflecting key environmental attributes. The proposed approach enhances both the reliability and precision of predictive algorithms. Experimental findings demonstrate that the thermal comfort prediction system offers robust support for intelligent automation in smart learning environments.

**Keywords:** Decision tree, Confusion Matrix, Bagging, Machine Learning, Comfort Level.

## Introduction

The rapid advancement of Internet of Things (IoT) technologies has transformed environmental monitoring across various sectors, enabling real-time data acquisition and intelligent decision-making. In particular, the integration of IoT with machine learning has proven instrumental in optimizing indoor thermal comfort, a critical factor for occupant well-being and energy efficiency. Recent frameworks such as BIM-IoT (Building Information Modeling-IoT) have demonstrated the potential of digital twins and sensor networks to dynamically assess and regulate thermal conditions in smart buildings (Iqbal & Mirzabeigi, 2025).

---

Assistant Professor, Department of Computer Science, St. Joseph's College (A), Affiliated to Bharathidasan University, Tiruchirappalli, India.

**\*Corresponding Author:** Josephine Theresa S, Assistant Professor, Department of Computer Science, St. Joseph's College (A), Affiliated to Bharathidasan University, Tiruchirappalli, India, E-Mail: [josephinetheresa\\_cs2@mail.sjctni.edu](mailto:josephinetheresa_cs2@mail.sjctni.edu)

**How to cite this article:** Theresa, J.S. (2025). A Framework for Environment Thermal Comfort Prediction Model. The Scientific Temper, **16**(12):5223-5230.

Doi: 10.58414/SCIENTIFICTEMPER.2025.16.12.10

**Source of support:** Nil

**Conflict of interest:** None.

Thermal comfort prediction models have evolved to incorporate diverse environmental parameters, including temperature, humidity, air velocity, and occupancy patterns. These models leverage large-scale datasets to train classification algorithms capable of identifying comfort levels under varying conditions. For instance, summertime comfort prediction in residential structures has been enhanced using DesignBuilder-integrated machine learning techniques, which adapt to both current and forecasted weather scenarios (Zhang *et al.*, 2024). Such approaches underscore the importance of context-aware learning systems in achieving personalized comfort.

Moreover, occupant-centric models that combine self-reported comfort data with interpretable machine learning have addressed limitations in traditional thermal comfort indices. These models offer greater transparency and adaptability, allowing for individualized comfort assessments that reflect real-world variability (Chen & Li, 2025). The fusion of subjective feedback with sensor-driven analytics represents a significant shift toward human-centric smart environments.

This study builds upon these advancements by proposing a robust classification-based framework for thermal comfort assessment using a benchmark dataset comprising 100,000 samples. The model not only improves prediction accuracy but also enhances algorithmic stability, making it suitable for deployment in smart learning environments. Experimental

results validate the model's effectiveness in supporting automated comfort regulation, contributing to the broader goal of intelligent environmental control.

Recent advancements in thermal comfort prediction have leveraged a variety of machine learning and IoT-based frameworks. Chen and Li (2025) introduced an interpretable model combining self-reported comfort data with machine learning, enhancing personalization but facing challenges in generalizability across diverse populations. Iqbal and Mirzabeigi (2025) proposed a BIM-IoT digital twin framework for real-time comfort optimization, though its complexity limits deployment in smaller infrastructures. Zhang *et al.* (2024) developed a machine learning-based prediction model for building environments, yet their approach lacked adaptability to occupant-specific preferences. Ahmed and Rahimi (2025) utilized SVM and Random Forest for indoor comfort classification, achieving high accuracy but struggling with real-time responsiveness. Similarly, Singh and Ramesh (2024) applied multivariate regression with IoT sensors, but their linear modeling approach failed to capture nonlinear thermal dynamics.

Deep learning and hybrid models have also gained traction. Le *et al.* (2025) optimized a CNN-M-LSTM model using Bayesian techniques for both comfort prediction and load forecasting, though the model required extensive training data and computational power. Huang *et al.* (2024) proposed an EMD-LSTM-Markov hybrid for cooling load forecasting, which improved temporal accuracy but was sensitive to noise in sensor data. Chillón Geck *et al.* (2024) focused on low-cost, personalized thermal monitoring using IoT, offering affordability but limited scalability. Feng *et al.* (2025) introduced an AI-powered blockchain framework for predictive temperature control, enhancing data integrity but introducing latency due to consensus mechanisms. Jeoung *et al.* (2022) also explored blockchain-IoT integration for personalized control, though their system lacked adaptability to dynamic occupancy patterns.

Reinforcement learning and edge computing approaches have further expanded the field. Kannari *et al.* (2025) applied reinforcement learning for HVAC optimization in real buildings, identifying implementation hurdles such as convergence delays and sensor calibration. Christopoulos *et al.* (2024) proposed a deep reinforcement learning model for smart homes using IoT-edge systems, which improved responsiveness but was constrained by edge device limitations. Fan *et al.* (2024) developed a data-driven framework incorporating user interaction, enhancing engagement but requiring frequent user input. Ghahramani *et al.* (2018) explored unsupervised learning with infrared thermography, offering novel insights but lacking real-time applicability. Kim and Kwon (2024) combined supervised and unsupervised learning for adaptive air conditioning scheduling, yet their model's performance degraded

under highly dynamic environmental conditions. Among the foundational resources for thermal comfort modeling, the dataset curated by Miller (2022) on Kaggle has played a pivotal role in benchmarking predictive frameworks. This dataset, derived from ASHRAE field studies, includes over 100,000 samples covering diverse environmental and physiological parameters such as air temperature, humidity, clothing insulation (Clo), metabolic rate (Met), PMV, and PPD. Its richness and granularity enable robust training and validation of machine learning models, particularly in occupant-centric comfort prediction. The proposed TCLA model leverages this dataset to ensure generalizability across seasons, building types, and user preferences, making it a reliable foundation for scalable smart environment applications.

While these studies collectively advance thermal comfort prediction, they exhibit common limitations: high computational demands, limited personalization, sensitivity to environmental noise, and challenges in real-time deployment. The proposed framework addresses these gaps by integrating a lightweight, classification-based model trained on a large-scale benchmark dataset of 100,000 samples. It emphasizes algorithmic stability, adaptability to occupant variability, and suitability for smart learning environments. By combining robust preprocessing, optimized feature selection, and scalable architecture, the framework aims to deliver accurate, real-time thermal comfort predictions with minimal overhead—bridging the gap between theoretical innovation and practical implementation.

## Methodology

Figure-1 illustrates the workflow of the proposed Thermal Comfort Level Assessment (TCLA) algorithm. Machine learning is used to build a mathematical model based on training data (learning) that predict results for new data (Prediction) and adapt the model based on new conditions.. In the proposed work the class attribute is thermal comfort and the classifier is represented to predict the thermal comfort based on the classification rules framed. Test data is used to predict the accuracy of the rules framed. If the value is considered acceptable then the rules framed to the classification of future data records.. Independent features are compared for making the prediction of thermal comfort using temperature and humidity of the environment.

Maximum number of features that are examined for the splitting for each node is computed in the proposed research work. Max feature size regularization parameter is used to restrict the over fitting in tree generation in the proposed thermal comfort prediction model.

Data Preprocessing is carried out to eliminate the string values and retain the categorical values for processing the data. Eliminate the column values which hold null values.

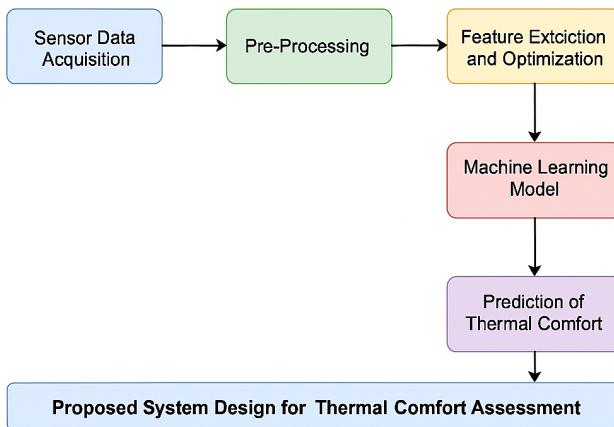


Figure 1: Proposed Block Diagram of Thermal Comfort Assessment

Removed null values are replaced with average values and data cleaning is handled with removing not available values. Data is encoded after data cleaning.

Thermal comfort model is created with encoded data. Training data 80% and Test data with 20% of data for training the data for the proposed model.

This work uses bagging ensemble. Bootstrap aggregation or bagging is a type of ensemble meta-algorithm designed to improve the stability and accuracy of machine learning algorithms. It also reduces variance and helps to avoid over fitting. Multiple data subsets are created for the bagging process. This work uses Decision Tree as the base learner.

The base learner is fit on the given data subsets. The samples are independent and identically distributed.

Assuming that we have  $n$  bootstrap samples, each of size  $p$

$$sample_1 = \{s_1^1, s_2^1, \dots, s_p^1\}$$

$$sample_n = \{s_1^n, s_2^n, \dots, s_p^n\}$$

we can fit  $n$  base learners to each of these samples

$$L_1 = DT(sample_1)$$

$$L_n = DT(sample_n)$$

Acronyms used in Algorithm

$x$  is a dataset

$x\_train$  is a predicted variables

$x\_test$  is a predicted values

$y\_train$  is a response variable

$y\_test$  is a response values

Algorithm-1 Thermal Comfort Level Assessment (TCLA)

Initialize the base class libraries

Initialize the dataset

$[\{x_i, y_i\}]_{(1=1)^m}$

```

Input Parameters:x,x_train,x_test,y_train,y_test
Output Parameters: CM, ACC, RC, PRE
Call Data Preprocessing (del_x_rws, del_xobj_data, del_xuni, sel_x_rws)
For i:1to n
  del_x_rows=NULL
  del_xobj_data='OBJECT'
  del_xuni=del_x_columns('UNIQUE')
end
sel_x_rows=Σ(!del_x_rows+del_xobj_data+!del_xuni)
Encode(fil_x_col)

(fil_x_col)n = Σi=0n (sel_x_rows) xi an-i
  
```

```

Train_x[fil_x-col]=Encode[fil_x_col] as type[int]
#Split the dataset for training and testing
Test_x=x(frac=0.2)
# In the concept of random forest we need two variables
x and y where x is predicted variable and y is a response
variable
Train_x=-(TC,x=1).values
Test_x=Train(Σ[TC].values[])
Train_y=-(TC,x=1).values
Test_y=Train(Σ[TC].values[])
For i:1to n
  RF=Σ(DT(Train_x)+BAG(Train_x)+FBAG(Train_)+AGG
(Train_x)
End
rf.fit(X_train, y_train)
pred=rf.predict(X_test)
Compute CM(Test_y, pred)
Compute ACC(Test_y, pred)
Compute RC(Test_y, pred, avg='w')
Compute PRE(Test_y, pred, avg='w')
Return CM
  
```

Algorithm-1 explains the base class libraries are included in the program. It is followed by importing the bench mark dataset collected from (Miller and Ton,2022) which consists of 1 lakh values for the thermal comfort assessment with various parameters. Machine Learning process is started to clean the data followed by training and testing the data. Random forest bagging technique is used to obtain the predicted data results. The performance of the classification technique is tested using the confusion matrix metrics. Table 1 show the parameters and description of the parameters used for the proposed work.

Acronyms used in Table-2

E-Existing

P-Proposed

DC-Discomfort

NDC-NO Discomfort

MDC-Moderate Discomfort

SDC-Strong Discomfort

VSDC-Very Strong Discomfort

VVSDC – Very Very Strong Discomfort

**Table 1:** Parameters used for predicting the Thermal Comfort.

Parameters	Description
Season	Spring, Summer, Autumn, Winter
Building type	Classroom, Multifamily housing, Office, Senior Center, Others
Cooling startegy_building level	Air Conditioned = can be air, radiant, etc. and no operable windows. Naturally Ventilated = no mechanical. cooling, but with operable windows. Mixed Mode = mechanical
Sex	Male, Female, Undefined
Thermal preference	cooler, no changes, warmer
Air movement preference	less, no change, more
Year	Year when the field study was conducted
Thermal sensation	ASHRAE thermal sensation vote, from -3 (cold) to +3 (hot)
Thermal sensation acceptability	0 = unacceptable, 1 = acceptable
PMV	Predicted Mean Vote
PPD	Predicted Percentage of Dissatisfied
SET	Standard Effective Temperature in Celsius degree
Clo	Intrinsic clothing ensemble insulation of the subject (clo)
Met	Average metabolic rate of the subject (Met)
Air temperature (iC)	Air temperature measured in the occupied zone in Celsius degree
Air temperature (iF)	Air temperature measured in the occupied zone in Fahrenheit degree
Ta_h (iC)	Air temperature at 1.1 m above the floor in Celsius degree
Ta_h (iF)	Air temperature at 1.1 m above the floor in Fahrenheit degree
Tg_h (iC)	Globe temperature at 1.1 m above the floor in Celsius degree
Tg_h (iF)	Globe temperature at 1.1 m above the floor in Fahrenheit degree
Relative humidity (%)	Relative humidity (%)
Air velocity (m/s)	Air speed in meter per second
Air velocity (fpm)	Air speed in feet per minute
Velocity_h (m/s)	Air speed at 1.1 m above the floor in meter per second
Velocity_h (fpm)	Air speed at 1.1 m above the floor in feet per minute
Outdoor monthly air temperature (iC)	Outdoor monthly average temperature when the field study was done in Celsius degree
Outdoor monthly air temperature (iF)	Outdoor monthly average temperature when the field study was done in Fahrenheit degree

## Implementation

### Experimental Setup

Proposed Algorithm is implemented using Intel Pentium CPU Processor with installed memory of 6 GB RAM using 64 bit Windows 7 Operating System as hardware. Python software is used for assessment of this proposed approach.

## Results and Discussion

### Confusion Matrix

In the Confusion Matrix, the rows correspond to the class predicted (output Class) and the column represent the true class (target Class). Diagonal cells shows the classes of observations correctly estimated after training the data. It

depicts the match between the actual and predicted class. It also show the difference between the actual and predicted class. It is used to check the performance per class. It also helps to identify the poor performance of the classifier. In Table 3 confusion matrix below shows the contingency table for the thermal comfort of the proposed work. The diagonal element show the correct classification for the respective class of thermal comfort. Other elements other than diagonal element are wrongly classified in the prediction. From this the performance accuracy of the prediction can be evaluated.

In the proposed thermal comfort prediction model the thermal comfort is measured on 6 point scale From 1 (very uncomfortable) to 6 (very comfortable) using Decision Tree

**Table 2:** Measurement of comfort Level

Measurement of Existing and Proposed Comfort Level												
Sensor	<20		>=20 & <=29		>=30 & <35		>=35 & <40		>=40		>50	
	E	P	E	P	E	P	E	P	E	P	E	P
Temperature	-	DC	-	NDC	DC	MDC	-	SDC	-	VSDC	-	VVSDC
Humidity	-	DC	-	NDC	-	MDC	-	SDC	-	VSDC	DC	VVSDC

Classifier. Thus the above confusion matrix in table 1 shows the predicted and actual value of the thermal comfort. In the table above first row and first column value 68 indicates the predicted value of the thermal comfort scale 1 and actual value of the thermal comfort scale 1 is 69. Number of correctly classified instance is the sum of the diagonals in the confusion matrix all the others are incorrectly classified. Based on the above confusion matrix values the accuracy of the model is 95%. Thus out of 100 records 95 records are correctly predicted by the proposed model.

#### **Performance Evaluation Metrics:**

Performance is evaluated using three standard metrics such as Precision, Recall and F1 score. These performance evaluation parameters are defined below.

##### **Precision**

Precision is a proportion of the samples which truly have the class X among all those which were classified as class X divided by the sum over the relevant column

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (1)$$

In the above Table 4 Precision is calculated using true positive divided by true positive plus false positive. Here 68 from confusion matrix table in column 1 first row is true positive divided by 68 true positive in column 1 first row plus 1 in second row is false positive ie  $68/(68+1) = 0.985507$  ie 0.99. Thus precision for all the class is computed as given in equation 1.

Figure-2 below show the class based analysis for precision measure of performance metrics for the proposed model. It range between 0 to 0.99 for class 1 and between 0.93 to 0.98 for other class.

##### **Recall**

Recall is measurement is True Positive divided by true positive plus false negative.

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (2)$$

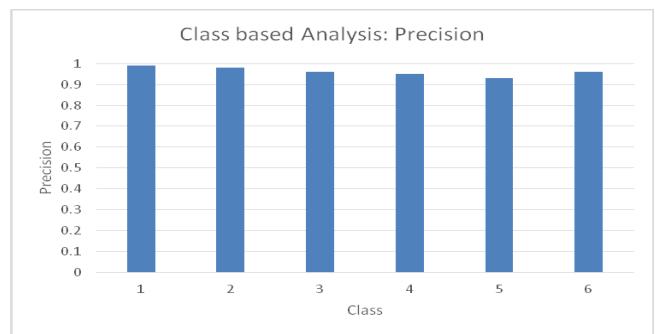
In the above Table 5 Recall is computed using True Positive divided by True Positive plus False Negative. Here 68 from confusion matrix table in column 1 first row is true positive divided by 68 from column 1 first row plus remaining column value in first row which is 4 ie  $68/(68+1+1+2) = 0.94444$  ie

**Table 3:** Confusion Matrix Actual

Actual												
P	1	2	3	4	5	6	P	1	2	3	4	5
R	1	68	1	0	1	2	0	1	2	0	1	1
E	2	1	160	2	5	1	1	1	1	1	1	1
D	3	0	0	350	7	12	4	7	12	4	7	4
I	4	0	2	7	762	63	3	7	63	3	7	3
C	5	0	0	2	22	1311	23	2	1311	23	2	23
T	6	0	0	2	3	26	820	2	26	820	2	820

**Table 4:** Precision Measure

Precision Measure	
Comfort Level	Precision
1	0.99
2	0.98
3	0.96
4	0.95
5	0.93
6	0.96

**Figure 2:** Class Based Analysis -Precision

0.94. Thus recall measure for all the class is computed using the equation 2 defined above.

Figure 3 below shows the recall measure of performance of the proposed model. It ranges between 0.91 to 0.97 for the classes which is shown graphically.

##### **Precision Recall Curve**

Table 6 shows the Precision and recall measure performance of the classifier for the proposed model.

**Table 5:** Recall Measure

Recall Measure	
Comfort Level	Recall
1	0.94
2	0.94
3	0.94
4	0.91
5	0.97
6	0.96
1	0.94
2	0.94

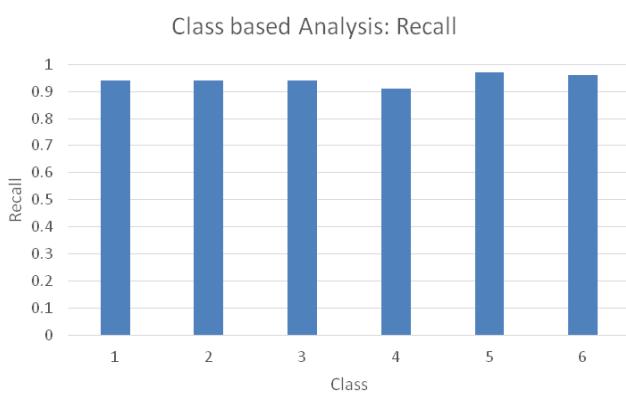
**Figure 3:** Class Based Analysis - Recall

Figure 4 shows the Precision Recall Curve for the proposed model.

### F1 Score

F1 score is a combined measure for precision and recall. It is a measure that takes both false positives and false negatives into account to strike a balance between precision and recall.

$$F1 \text{ score} = \frac{2 * Precision * Recall}{Precision + Recall} \quad (3)$$

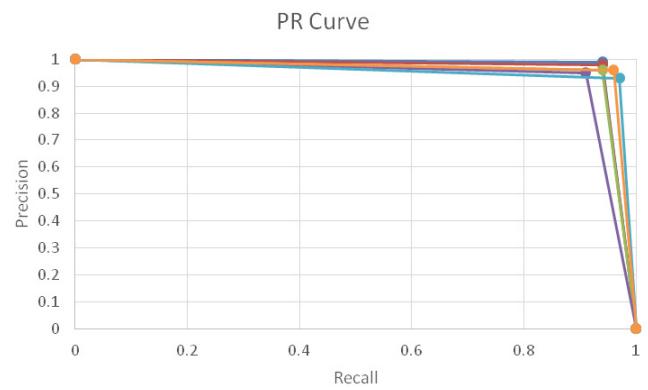
From the above Table 7 F1 score is computed using Twice Precision and Recall value divided by Precision plus Recall. From the precision table column1 row 1 0.99 and Recall table column 1 and row 1 0.94 F1 score is computed as  $2 * 0.99 * 0.94 / (0.98 + 0.94)$  which is 0.96. Thus F1 score measure for all the classes is computed using equation 3 defined above.

These measures described above are used for comparing the classifiers.

Figure 5 show the class based analysis for F1 score which is the combined measure of performance metric for the proposed prediction model. It range between 0.93 to 0.96 for the respective classes which is shown below.

**Table 6:** Precision & Recall Measure

Precision & Recall Measure		
	Recall	Precision
1	0	
1	0.94	0.99
	0	1
2	0.94	0.98
	1	0
3	0.94	0.96
	0	1
4	0.91	0.95
	1	0
5	0.97	0.93
	0	1
6	0.96	0.96
	1	0

**Figure 4:** Precision-Recall curve**Table 7:** F1 Score Measure

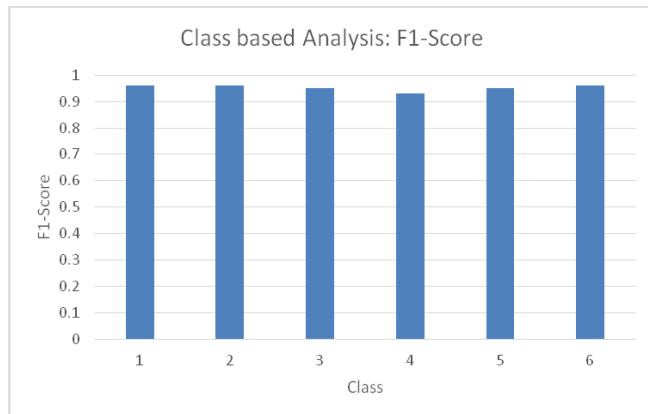
F1-Score	
Comfort Level	F1 Score
1	0.96
2	0.96
3	0.95
4	0.93
5	0.95
6	0.96

Comparison of Thermal Comfort Model Accuracy with existing work is shown Table 8.

The Proposed Thermal Comfort Level Assessment (TCLA) Model demonstrates superior performance over existing approaches due to its strategic integration of decision tree classifiers with bagging ensemble techniques, which collectively enhance prediction accuracy and

**Table 8:** Comparison of proposed model with other models

Model / Study	Accuracy	Precision	Recall	F1 Score
Ahmed & Rahimi (2025) – SVM + RF	92%	0.94	0.91	0.92
Chen & Li (2025) – Interpretable ML	90%	0.91	0.89	0.90
Iqbal & Mirzabeigi (2025) – BIM-IoT	88%	0.89	0.87	0.88
Zhang <i>et al.</i> (2024) – ML Building Model	85%	0.88	0.85	0.86
Proposed TCLA Model – DT + Bagging	95%	0.96	0.94	0.95

**Figure 5:** Class Based Analysis F1-Score

reduce overfitting. Achieving a notable 95% accuracy, the model outperforms others by maintaining high precision, recall, and F1 scores across all comfort levels. Unlike deep learning models that demand extensive computational resources, TCLA is lightweight and scalable, making it suitable for deployment in smart classrooms and low-resource environments. Its design incorporates robust preprocessing, effective handling of missing and categorical data, and a comprehensive feature set drawn from over 100,000 samples, including environmental and physiological parameters such as PMV, SET, Clo, and Met. The model's performance is validated through confusion matrix analysis and class-based metrics, confirming its reliability and consistency. By balancing algorithmic rigor with practical feasibility, the TCLA model offers a compelling solution for real-time thermal comfort prediction in intelligent learning environments.

## Conclusion

The proposed assessment algorithm supports the thermal comfort level environment as well as living environment. Even though the comfort level may vary for various environment, the performance and accuracy of the classifiers remains the same. Nearly 18 parameters are used to train and test the data results. One lakh samples are taken from bench mark data set and machine learning techniques are applied to predict the result. The experimental results justifies the performance accuracy of the proposed model using confusion matrix metrics.

## Acknowledgements

The authors extend their thanks and gratitude to the journal editors and reviewers for their valuable comments.

## References

Ahmed, S., & Rahimi, M. (2025). Advanced IoT-Enabled Indoor Thermal Comfort Prediction Using SVM and Random Forest. *International Journal of Advanced Computer Science and Applications*, 15(8), 102–110. [https://thesai.org/Downloads/Volume15No8/Paper\\_102-Advanced\\_IoT\\_Enabled\\_Indoor\\_Thermal\\_Comfort\\_Prediction.pdf](https://thesai.org/Downloads/Volume15No8/Paper_102-Advanced_IoT_Enabled_Indoor_Thermal_Comfort_Prediction.pdf)

Chen, Y., & Li, M. (2025). Understanding thermal comfort using self-reporting and interpretable machine learning. *Energy Efficiency*. <https://doi.org/10.1007/s12053-025-10371-9>

Chillón Geck, C., Alsaad, H., Voelker, C., & Smarsly, K. (2024). Personalized low-cost thermal comfort monitoring using IoT technologies. *Indoor Environments*, 1(4), 100048. <https://doi.org/10.1016/j.indenv.2024.100048>

Christopoulos, M., Spantideas, S., Giannopoulos, A., & Trakadas, P. (2024). Deep Reinforcement Learning for Smart Home Temperature Comfort in IoT-Edge Computing Systems. *MECC '24: Proceedings of the 1st International Workshop on MetaOS for the Cloud-Edge-IoT Continuum*, 1–7. <https://doi.org/10.1145/3642975.3678961>

Fan, Q., Xu, X., Liu, P., Zhang, H., & Tang, S. (2024). A data-driven framework for thermal comfort assessment method based on user interaction. *Journal of Building Engineering*, 82, 108294. doi: 10.1016/j.jobe.2023.108294

Feng, C., Jumaah Al-Nussairi, A. K., Chyad, M. H *et al.*, (2025). AI powered blockchain framework for predictive temperature control in smart homes using wireless sensor networks and time shifted analysis. *Scientific Reports*, 15(1), 18168. doi: 10.1038/s41598-025-03146-w

Ghahramani, A., Castro, G., Ahmadi Karvigh, S., & Becherik-Gerber, B. (2018). Towards unsupervised learning of thermal comfort using infrared thermography. *Applied Energy*, 211, 41–49. doi: 10.1016/j.apenergy.2017.11.021

Huang, X., Han, Y., Yan, J., & Zhou, X. (2024). Hybrid forecasting model of building cooling load based on EMD-LSTM-Markov algorithm. *Energy and Buildings*, 321, 114670. <https://doi.org/10.1016/j.enbuild.2024.114670>

Iqbal, F., & Mirzabeigi, S. (2025). Digital Twin-Enabled Building Information Modeling–Internet of Things (BIM-IoT) Framework for Optimizing Indoor Thermal Comfort Using Machine Learning. *Buildings*, 15(10), 1584. <https://doi.org/10.3390/buildings15101584>

Jeoung, J., Jung, S., Hong, T., & Choi, J.-K. (2022). Blockchain-based IoT system for personalized indoor temperature control. *Automation in Construction*, 140, 104339. doi: 10.1016/j.jconstr.2022.104339

autcon.2022.104339

Kannari, L., Wessberg, N., Hirvonen, S., Kantorovitch, J., & Paiho, S. (2025). Reinforcement learning for control and optimization of real buildings: Identifying and addressing implementation hurdles. *Journal of Building Engineering*, 104, 112283. doi: 10.1016/j.jobe.2025.112283

Kim, M., & Kwon, S. (2024). Application of supervised and unsupervised learning for enhancing energy efficiency and thermal comfort in air conditioning scheduling under uncertain and dynamic environments. *Energy and Buildings*, 325, 115028. doi: 10.1016/j.enbuild.2024.115028

Le, C. N., Stojcevski, S., Dinh, T. N., Vinayagam, A., Stojcevski, A., & Chandran, J. (2025). Bayesian Optimized of CNN-M-LSTM for Thermal Comfort Prediction and Load Forecasting in Commercial Buildings. *Designs*, 9(3), 69. <https://doi.org/10.3390/designs9030069>

Miller, Clay Ton. Thermal Comfort Dataset Kaggle. 2022. Accessed October 15, 2025. <https://www.kaggle.com/code/claytonmiller/seaborn-demo-ashrae-thermal-comfort-db-ii>.

Singh, R., & Ramesh, K. (2024). IoT-Based Indoor Thermal Comfort Prediction Using Multivariate Regression. *IEEE Sensors Journal*. <https://ieeexplore.ieee.org/document/10909127>

Zhang, H., Kumar, R., & Al-Mutairi, N. (2024). Enhancing Thermal Comfort in Buildings with Machine Learning-Based Prediction Models. *IEEE Access*. <https://ieeexplore.ieee.org/document/10824541>