International Journal of Innovative Research and Scientific Studies, 8(5) 2025, pages: 2370-2380

IJIRSS
International Journal of Innovative Research & Scientific Studies

ISSN: 2617-6548

URL: www.ijirss.com

Real-time data fusion for thermal comfort prediction using transformer models

Bibars Amangeldy[1,2], Nurdaulet Tasmurzayev[1,2*], Baglan Imanbek[2], Serik Aibagarov[1,2], Zukhra Abdiakhmetova[2]

[1]LLP «DigitAlem», Almaty 050042, Kazakhstan.
[2]Al-Farabi Kazakh National University, Almaty 050040, Kazakhstan.

Corresponding author: Nurdaulet Tasmurzayev (*Email: tasmurzayev.n@gmail.com*)

Abstract

Maintaining optimal thermal comfort in buildings is critical for occupant well-being and energy efficiency. This study introduces a permutation-importance-based feature selection method for multiclass thermal comfort classification using ensemble algorithms. Data comprising environmental (air temperature, humidity, CO_2 concentration) and physiological (SpO_2, blood pressure, BMI, HRV metrics) variables were collected from 1536 samples and labeled on the ASHRAE 7-point scale. Decision Tree, AdaBoost, and CatBoost models were trained on the full feature set, then re-evaluated using only the ten most informative predictors identified via permutation importance. Results show that reduced-feature models match or slightly outperform full-feature counterparts: CatBoost accuracy improved from 0.961 to 0.971, AdaBoost from 0.841 to 0.857, and Decision Tree from 0.688 to 0.825, while dramatically lowering sensor and computational requirements. Paired t-tests confirmed no significant performance loss. This streamlined approach enables cost-effective, real-time thermal comfort monitoring and supports the deployment of intelligent HVAC systems with minimal hardware.

Keywords: CatBoost, Feature selection, Permutation importance, Sensor reduction, Thermal comfort.

DOI: 10.53894/ijirss.v8i5.9474
Funding: This work is supported by the Committee of Science of the Republic of Kazakhstan (Grant number AP23488794).
History: Received: 26 June 2025 / **Revised:** 30 July 2025 / **Accepted:** 1 August 2025 / **Published:** 26 August 2025

Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.
Transparency: The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.
Publisher: Innovative Research Publishing

1. Introduction

In modern conditions of rising energy consumption and intensifying climate challenges, maintaining optimal thermal comfort in indoor environments has become critically important for the design, operation, and management of buildings. Thermal comfort is regarded as a key parameter of indoor environmental quality (IEQ), directly affecting occupants' well-being, health, and cognitive productivity [1, 2]. Insufficiently balanced temperature and humidity conditions can lead to reduced concentration, increased fatigue, and even the development of chronic illnesses. The relevance of this research is justified by the need to develop adaptive and energy-efficient microclimate control strategies capable of ensuring a high level of comfort with minimal energy expenditure.

Perception of thermal comfort is determined by a combination of physical and physiological factors. Key environmental parameters include air temperature (Ta), relative humidity (RH), mean radiant temperature (Tmrt), air velocity (va), and indoor pollutant concentration (such as CO_2) [3-5]. Individual user characteristics, such as age, gender, clothing insulation (Clo), metabolic rate, and body mass index (BMI), also play a significant role [2]. Recent studies have expanded the set of thermal-state markers to include skin temperature (SKT) at various body sites, heart rate variability (HRV), blood pressure, and respiratory rate, allowing for more accurate assessments of comfort and physiological adaptation [6].

Building construction features and ventilation strategies strongly influence the indoor microclimate. Natural and mechanical ventilation, the architectural layout of windows and doors, and the performance of HVAC systems determine the distribution of temperature, air velocity, and pollutant levels within spaces [7]. The diversity of climate zones, from Mediterranean to cold, and from hot-humid to heat-wave-prone megacities, imposes specific requirements on system selection and configuration [8-10]. Usage patterns further affect microclimate characteristics: university lecture halls often suffer from CO_2 buildup during classes, dormitories face high occupancy densities in cold seasons [8, 11] and childcare facilities demand consideration of spatial variations at children's heights [12].

One of the main operational challenges is balancing comfort with energy expenditure. Intensive ventilation to reduce CO_2 levels can significantly increase energy consumption, especially in regions with extreme temperatures [11]. Approaches to address this include occupancy-based ventilation optimization, heat-recovery exchangers (for example, total heat exchangers), and advanced HVAC control systems that adapt to changing conditions and loads [13-15]. However, many existing solutions remain costly and energy-intensive due to the need for numerous sensors and complex control algorithms [2].

Traditional building control systems rely on static rules and do not account for dynamic external conditions or individual user preferences. They poorly adapt to fluctuations in occupancy and activity levels and neglect physiological states, reducing the precision of comfort maintenance [16, 17]. Modern demands call for intelligent systems that can rapidly respond to changing conditions, forecast energy needs, and manage the indoor climate efficiently without undue energy waste.

The rapid development of machine learning (ML) and artificial intelligence (AI) technologies opens new opportunities for predicting and optimizing thermal comfort and building energy efficiency [18, 19]. Algorithms such as XGBoost, Random Forest, LightGBM, CatBoost, hybrid neural networks (LSTM, CNN), and reinforcement learning (RL) methods are applied to forecast energy consumption [20] optimize green building design [13] estimate unmeasured indoor parameters (temperature, humidity, CO_2) with limited sensors [15] develop advanced HVAC control strategies [21] and assess comfort based on combined environmental and physiological data, including computer vision techniques [1, 22, 23].

Digital Twin technologies, integrating data from IoT devices and sensors, enable the creation of virtual replicas of indoor spaces for real-time microclimate monitoring and optimization [20]. AIoT systems enable intelligent control of actuators such as ceiling fans based on sensor data and machine learning algorithms to achieve energy savings and enhanced comfort [24]. ML-based occupancy detection methods optimize energy use according to actual space utilization [16] while predictive maintenance systems for HVAC increase reliability and microclimate stability [14].

Despite significant progress, limitations persist related to data quality and volume, model generalizability, and computational resources. Furthermore, the unique requirements of different building types and user categories necessitate adaptive, customized solutions. In this work, we propose a permutation-importance-based feature selection method for multiclass thermal comfort classification. We demonstrate that using a small set of the most informative features achieves accuracy comparable to a full-feature model, while significantly reducing the number of required sensors and computational overhead.

2. Methodology

This section outlines the methodological pipeline implemented in the study. It includes a description of the dataset, the applied machine learning algorithms, the cross-validation and hyperparameter tuning procedures, the permutation-based approach for determining feature importance, and the evaluation metrics used to assess model performance.

Each subsection provides detailed explanations of the corresponding components within the predictive modeling workflow.

2.1. Dataset

The dataset comprises 1,536 samples with 24 features, encompassing a rich array of physiological, environmental, and anthropometric parameters. Environmental variables include indoor air temperature, relative humidity, carbon dioxide (CO_2) concentration, and outdoor temperature. Physiological signals were collected through wearable and medical-grade devices, capturing heart rate (HR), heart rate variability (HRV) metrics (such as SDNN, rMSSD, LF/HF ratio), blood oxygen saturation (SpO_2), and systolic and diastolic blood pressure. In addition, the dataset includes anthropometric measurements such as body mass index (BMI), fat and muscle percentage, and absolute body weight and muscle mass. Each data point is labeled with a thermal comfort score based on the ASHRAE 7-point thermal sensation scale, ranging from 3 (cold) to +3 (hot). Linear correlation of the features shown in Figure 1.

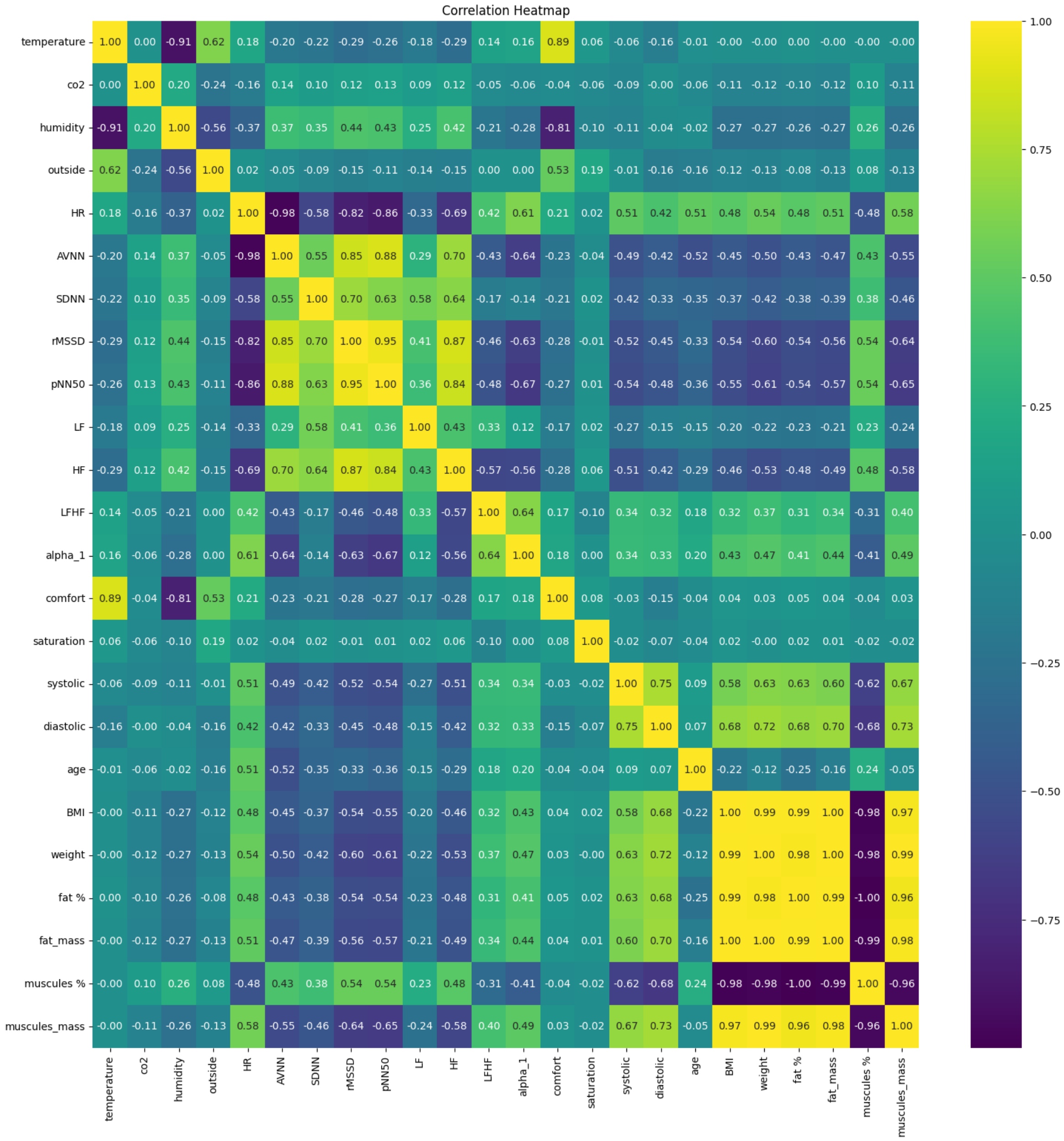


Figure 1.
Correlation heatmap of dataset.

2.2. Machine Learning Methods

The predictive framework incorporates three classification algorithms of varying complexity: Decision Tree (DT), Adaptive Boosting (AdaBoost), and Categorical Boosting (CatBoost). DT serves as an interpretable baseline model, while AdaBoost and CatBoost offer more robust performance through ensemble and gradient boosting techniques. The next section provides a detailed overview of each machine learning model, explaining their main features and how they contribute to the predictive system.

The DT algorithm is a non-parametric method that splits data into branch-like segments, forming an inverted tree structure. It handles large, complex datasets efficiently and can be trained on one dataset while another is used for validation to determine the optimal tree size [25]. AdaBoost is a well-known ensemble learning method that enhances the performance of weak classifiers by combining them through weighted voting. It operates by selecting the optimal threshold along a feature to split the data into two classes, typically labeled -1 and 1 [26]. CatBoost is a machine learning algorithm based on an improved version of Gradient Boosting Decision Trees (GBDT), designed to handle categorical features effectively. It reduces information loss and prevents target leakage through ordered boosting. By using random permutations, it also minimizes overfitting, making it suitable for small datasets and a wide range of applications [27].

2.3. Grid Search and Cross Validation

Hyperparameter tuning was performed through an extensive grid search approach, in which pre-specified sets of parameter values were attempted systematically. Performance in each setting was calculated through the same ten-fold cross-validation procedure described above. A hyperparameter grid was tailored to each model according to its structure, and the best setting was selected as that with the greatest mean validation performance. This ensured refinement of model performance and assured that follow-up comparisons among classifiers reflected their optimal performance under consistent validation settings.

For ensuring stable and unbiased performance estimation, a ten-fold cross-validation approach was employed. The data set was split into ten equal-sized subsets. For every fold, the model was trained on nine subsets and tested on the remaining one. The experiment was repeated ten times, and each subset was employed once as a validation set. The performance metrics thus computed were then averaged across all the folds, providing a global measure of the generalizing capability of the model and minimizing the variance due to the individual train-test split.

2.4. Permutation Importance

Permutation importance is a model-agnostic technique for estimating the contribution of individual features to a model's predictive performance. The method operates by measuring the decrease in model accuracy (or another performance metric) after randomly shuffling the values of a single feature, thereby disrupting its relationship with the target variable. If the feature is important, this permutation will lead to a significant drop in performance. Conversely, shuffling a feature with little or no predictive value will have a minimal effect on the output.

Formally, let f be the trained model and X the feature matrix. For each feature x_j, a permuted version $X_{\pi(j)}$ is created by randomly shuffling the values of x_j across all samples. The model is then evaluated on the modified dataset $X_{\pi(j)}$, and the change in performance (accuracy, F1-score, or AUC) is recorded as the importance score for that feature.

Permutation importance offers several advantages: it is easy to implement, applicable to any model, and directly tied to the model's performance on a validation set. Unlike methods such as SHAP or LIME, it does not require internal access to the model and thus can be used with black-box systems. However, it assumes that the features are not highly correlated, as permutation of one feature may implicitly affect others if multicollinearity is present.

In this study, permutation importance was computed using the trained model on a held-out test set. The top-ranked features, as determined by the largest performance drop upon permutation, were then used to construct reduced-feature models. This approach allowed for the evaluation of whether similar classification performance could be maintained using a smaller, more interpretable subset of input variables.

2.5. Evaluation

To assess the performance of the classification models, multiple evaluation methods were employed, including the confusion matrix, standard classification metrics (accuracy, precision, recall, and F1-score), and statistical significance testing using the paired *t*-test.

2.5.1. Confusion Matrix

The confusion matrix provides a detailed overview of the model's performance by comparing predicted labels to true class labels. For multiclass classification, the matrix contains counts of true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN) for each class. This tool enables the identification of class-specific errors and imbalances in prediction.

2.5.2. Classification Metrics

Standard classification metrics were calculated for each class as well as in macro-averaged and weighted-averaged forms:

1. Accuracy:

$$Accuracy = \frac{\sum_{i=1}^{K} TP_i}{\sum_{i=1}^{K} (TP_i + TN_i + FP_i + FN_i)} \tag{1}$$

2. Precision:

$$Precision_i = \frac{TP_i}{TP_i + FP_i} \tag{2}$$

3. Recall:

$$Recall_i = \frac{TP_i}{TP_i + FN_i} \tag{3}$$

4. F1-score:

$$F1_i = 2 * \frac{Precision_i * Recall_i}{Precision_i + Recall_i} \tag{4}$$

Macro-averaged values treat all classes equally, while weighted averages take class support (i.e., number of instances per class) into account.

2.5.3. Paired t-Test

To statistically compare the performance of models trained on the full feature set and those trained on only the top 10 most important features, a paired t-test was conducted. This test evaluates whether the mean performance difference between two related sets of observations is statistically significant.

Let $D_i = x_i - y_i$ denote the difference in a performance metric (accuracy or F1-score) for the model i, where x_i is the score from the full model and y_i is from the reduced model. The *t-statistic* is computed as:

$$t = \frac{\bar{D}}{s_D / \sqrt{n}} \tag{5}$$

where:

- $\bar{D}$ is the mean of the differences,
- s_D is the standard deviation of the differences,
- n is the number of paired observations (across cross-validation folds).

A corresponding p-value is reported to determine statistical significance. A threshold of α=0.05 is used to evaluate whether the difference is statistically meaningful.

3. Results

This section presents the results of training and evaluating various machine learning models for thermal comfort classification. The analysis includes model performance across key evaluation metrics, comparisons between full-feature and reduced-feature configurations, interpretation of feature importance, and a discussion of the practical implications and limitations of the proposed approach.

3.1. Model Performance on Full Feature Set

The complete 23-feature dataset (1,228 training and 308 test instances, 7 classes) was used to benchmark models. Summary metrics appear in Table 1 while Figure 2 shows the confusion matrices.

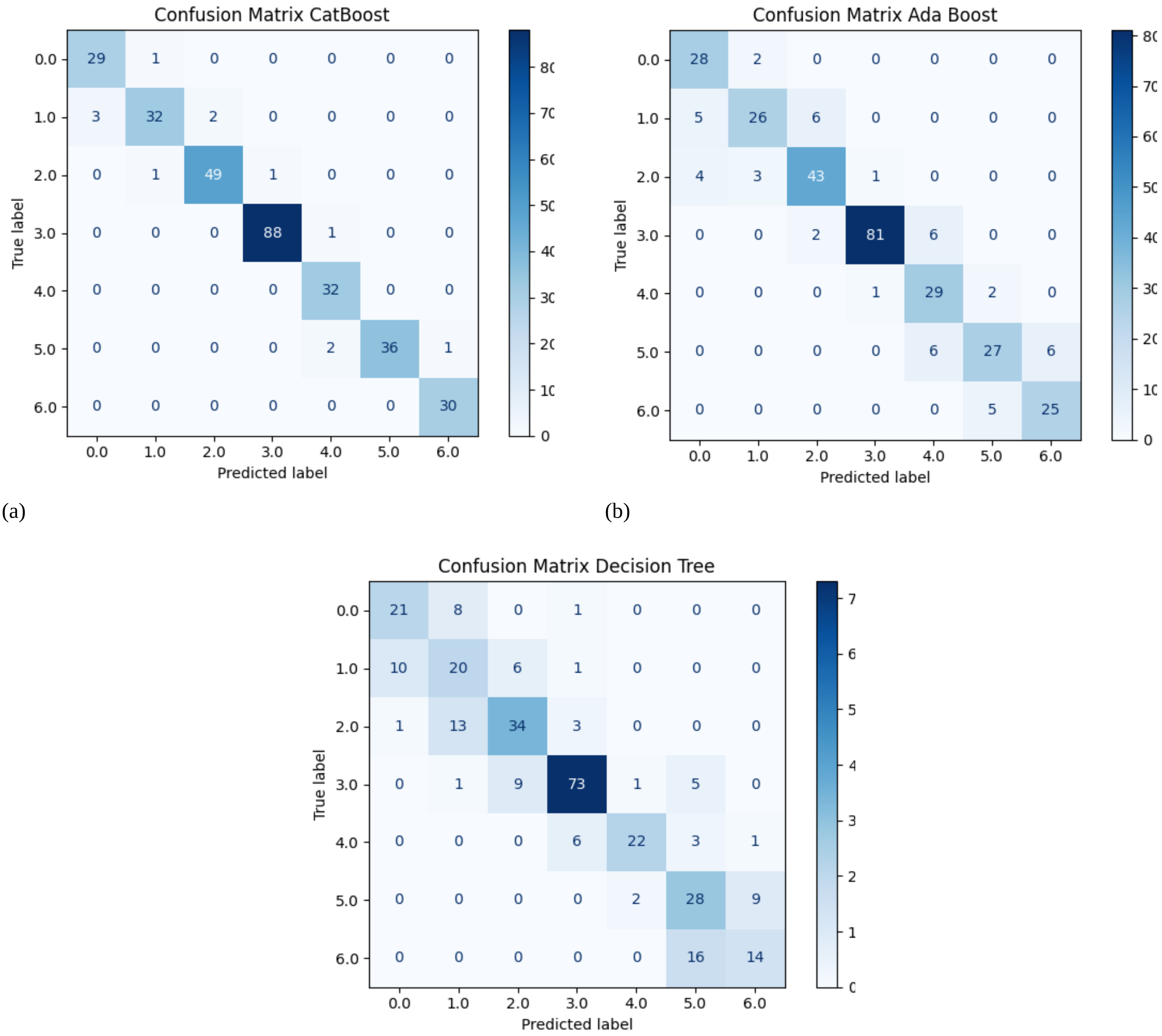
Table 1.
Performance of models on the whole dataset.

Evaluation metrics	AdaBoost	CatBoost	Decision Tree
Accuracy	0.841	0.961	0.689
Cross-validation accuracy	0.856	0.949	0.744
F1-Score	0.82	0.96	0.66
Precision	0.82	0.95	0.67
Recall	0.83	0.96	0.66

CatBoost emerges as the leading model. It achieves a test accuracy of 0.961, closely matching its 10-fold cross-validation mean of 0.949, thereby demonstrating strong generalization. In the classification report (Table 1), CatBoost's class-level F1-scores cluster tightly between 0.90 and 0.99, with precision and recall remaining above 0.90 for every class, and recall reaching 1.00 for classes 4 and 6.

AdaBoost ranks second with a test accuracy of 0.841 (cross-validation = 0.856). It excels on class 3 (F1 = 0.94) and maintains balanced precision and recall on classes 0 and 2 (≈ 0.84 each). Nevertheless, precision drops to 0.71 for class 4, and recall decreases to 0.69 for class 1, indicating room for improvement in specific categories.

The decision tree trails with a test accuracy of 0.689 (cross-validation = 0.744). Its performance is uneven: while classes 3 and 4 achieve respectable F1-scores of 0.84 and 0.77, respectively, F1-scores drop below 0.55 for classes 1 and 6. The macro-averaged recall and F1-score are 0.66 (Table 1).



(a) (b) (c)

Figure 2.
Confusion matrix of models.

In summary, CatBoost not only delivers the highest overall accuracy but also maintains uniformly strong precision, recall, and F1-scores across all seven classes. AdaBoost serves as a competitive, lightweight alternative, though its per-class metrics vary more widely. By contrast, a single Decision Tree, despite its interpretability, does not capture the complex decision boundaries required by this multi-class task, leading to markedly lower average metrics.

3.2. Feature Importance Analysis

After training the models on the full feature set, we quantified each predictor's contribution using permutation importance. Figure 3 presents the ranked importance for the Decision Tree, AdaBoost, and CatBoost models. Several consistent patterns emerge.

3.2.1. Dominance of Ambient Temperature

For every algorithm, temperature is the single most influential feature. It accounts for 27% of total importance in the Decision Tree, 42% in CatBoost, and more than half (≈ 55%) in AdaBoost. This aligns with domain intuition: perceived environmental comfort and, by extension, the target class depend heavily on air-temperature fluctuations.

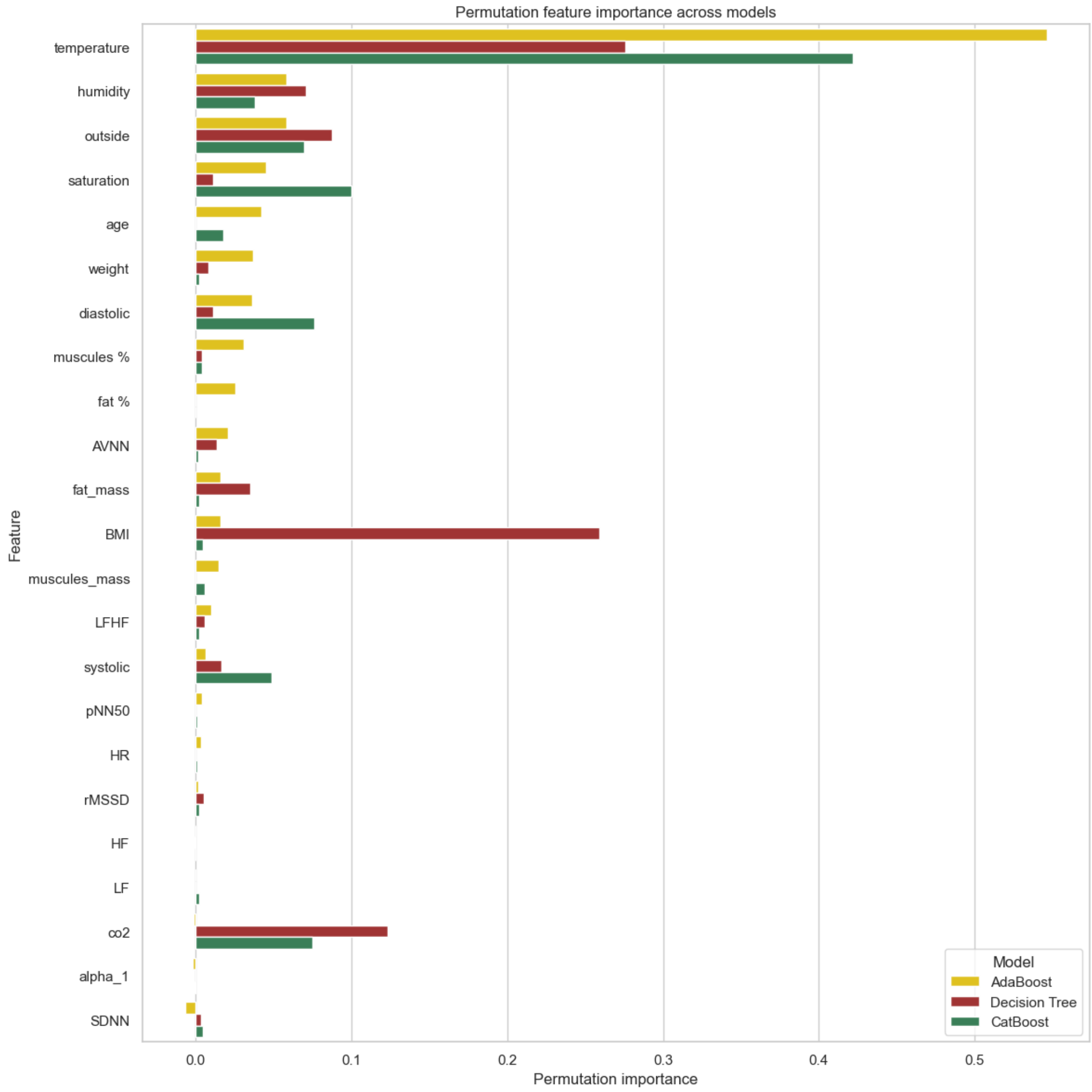


Figure 3.
Permutation feature importance across models.

3.2.2. Environmental Variables vs. Physiological Metrics

Beyond temperature, all three models highlight a core set of environmental parameters: humidity, outdoor temperature (outside), and CO_2 concentration, though their ranks differ. The percentages represent each feature's relative contribution to the model's predictive power.

CatBoost places CO_2 fourth (7.5 %), directly after SpO_2 saturation (9.9 %) and diastolic blood pressure (7.6 %), indicating an interaction between air quality and physiological response.

AdaBoost assigns nearly equal weight (approximately 6%) to humidity and outdoor temperature, while assigning a slightly negative importance to CO_2. This suggests that its ensemble of weak learners relies less on CO_2 for class discrimination.

In the Decision Tree, BMI (25.9%) rivals temperature, whereas humidity and outdoor temperature jointly contribute only 15.8%. This illustrates the tree's tendency to split early on a small subset of strongly discriminative features.

3.2.3. Cardiovascular and Body-Composition Indicators

Physiological features such as BMI, systolic/diastolic pressure, SpO_2 saturation, body-fat mass, and muscle mass collectively account for:

~30 % of total importance in CatBoost,

~25 % in AdaBoost, and

~35 % in the Decision Tree.

Notably, SpO_2 saturation is the second most important variable for CatBoost and the fourth for AdaBoost, underscoring its predictive value when combined with high-capacity models.

3.2.4. Heart-Rate-Variability (HRV) Measures

HRV-derived metrics (AVNN, SDNN, rMSSD, pNN50, LF, HF, LF/HF) show a modest but non-zero influence in CatBoost and AdaBoost, whereas they contribute minimally in the Decision Tree. This is expected: tree-based ensembles can exploit subtle nonlinear interactions that single trees miss.

3.2.5. Near-zero or Negative Scores

A few predictors (HF in AdaBoost and CatBoost, alpha-1 in all three models) receive near-zero or slightly negative importance, indicating either redundancy given other variables or noise concerning the classification objective.

These insights can guide future feature-engineering efforts: most of the predictive power resides in a compact subset of ambient and vital-sign variables, implying that sensor configurations prioritizing temperature, basic air-quality measures, and key physiological indicators would suffice for reliable real-time deployment.

3.3. Performance on Reduced Feature Set

After retraining each model on its ten most influential features (as determined by permutation importance), overall performance was preserved or improved across the board. Table 2 compares accuracy on the full feature set with accuracy on the reduced set and lists the metrics and 10-fold cross-validation (CV) means for the new models. The confusion matrix of each model shown in Figure 4.

Table 2.
Performance of models on top 10 features.

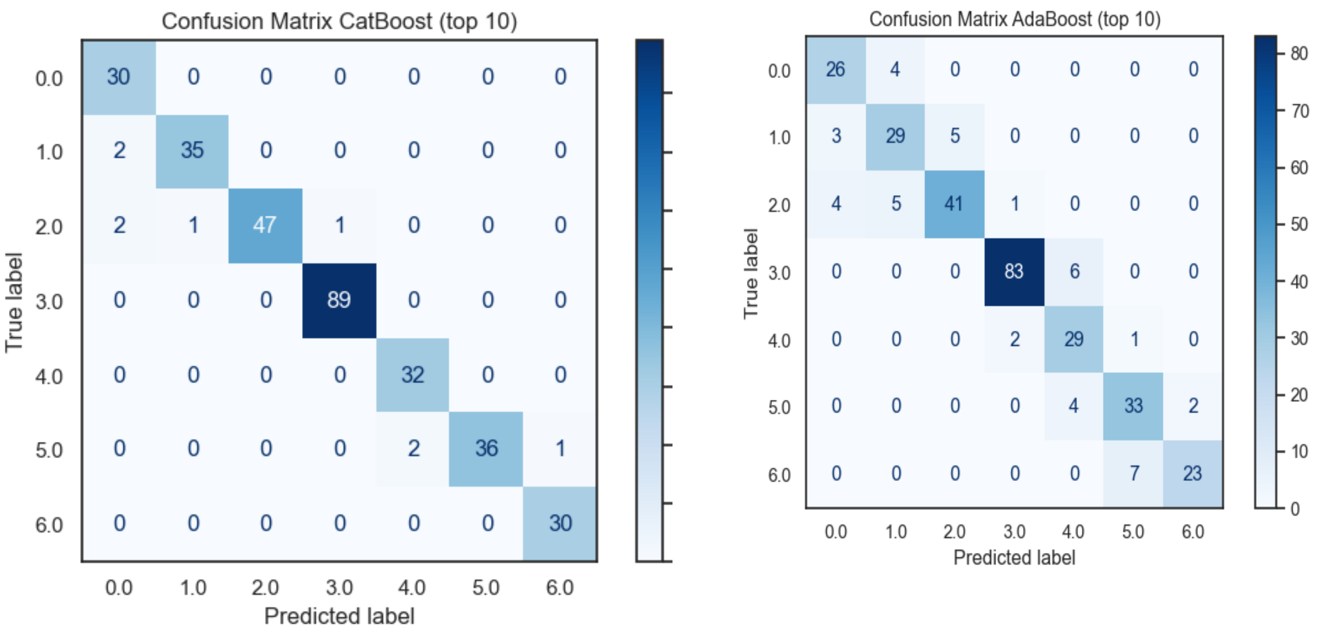
Evaluation metrics	AdaBoost	CatBoost	Decision Tree
Accuracy (full)	0.841	0.961	0.688
Accuracy (top 10)	0.857	0.971	0.825
Δ Accuracy	+0.016	+0.010	+0.137
Cross-validation accuracy (top 10)	0.856	0.972	0.824
F1-Score	0.84	0.96	0.70
Precision	0.85	0.96	0.69
Recall	0.84	0.96	0.69

With only ten inputs, the tree becomes markedly simpler and far less prone to overfitting. The 13.7-point jump in accuracy is accompanied by sizeable gains in precision for classes 0, 2, 5, and 6, and a recall increase of 0.27 for class 1. The macro F1 score of 0.80 now approaches that of AdaBoost, showing that pruning noisy features is more beneficial than granting the tree additional split options.

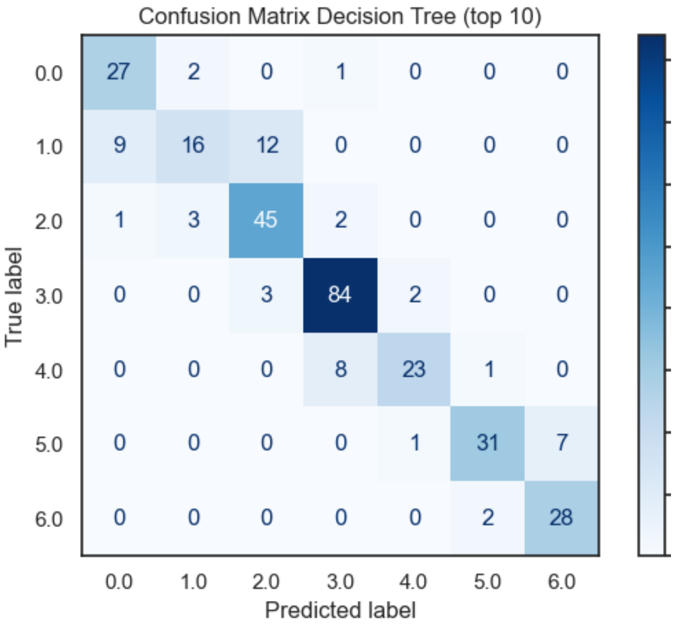
AdaBoost already uses ensembles of shallow trees to dilute noise; nevertheless, trimming the feature space yields a consistent accuracy increase to 0.857 and raises the worst per-class F1 score from 0.76 to 0.80. Cross-validation remains virtually unchanged, confirming that the extra features were largely redundant rather than harmful.

Even the strongest model picks up a small improvement, pushing accuracy to 0.971 and aligning the CV estimate at 0.972. The narrower feature set removes occasional misclassifications in minority classes without degrading any other metric, implying that CatBoost extracts almost all of its predictive power from the leading ten variables.

Because each algorithm indeed retains, it slightly improves generalization with only ten predictors, the monitoring system can dispense with several low-value sensors (such as alpha-1, HF, fat %, muscle %). For real-time deployments, this reduces hardware costs, data collection overhead, and latency while maintaining or enhancing decision quality.



(a) (b)



(c)

Figure 4.
Confusion matrix of models (top 10 features).

In short, the reduced feature set delivers a tangible performance dividend for the Decision Tree, a modest uplift for AdaBoost, and a fine-tuning effect for CatBoost. These results reinforce the conclusion that a compact collection of temperature, core environmental readings, and key physiological markers is sufficient to drive highly accurate multi-class predictions.

3.4. Statistical Comparison

To assess whether reducing the feature set impacted predictive performance, we conducted paired t-tests using ten-fold cross-validation on the entire dataset, comparing the accuracy of pretrained models. Each pair consisted of a model trained on the full set of 23 features and a counterpart trained on the reduced set containing only the top 10 features selected via permutation importance. Importantly, both models were evaluated on the same cross-validation splits but were trained in advance on their respective feature sets.

For the Decision Tree classifier, the difference in accuracy was not statistically significant ($t = -0.357$, $p = 0.729$), with a negligible effect size (Cohen's $d = -0.11$). For AdaBoost and CatBoost, the pretrained models achieved identical accuracy scores across all folds, resulting in zero variance in the differences. Consequently, the t-statistic and effect size were both zero.

These results demonstrate that reducing the feature set to the top 10 variables did not degrade performance. The pretrained reduced-feature models retained the same level of accuracy as the full-feature models, while offering advantages in terms of simplicity, faster inference, and reduced data requirements.

4. Conclusion and Future Work

This study demonstrates that multi-class thermal comfort classification can be performed accurately using a significantly smaller set of input variables than initially employed. Among the three algorithms tested, CatBoost achieved the best results, with cross-validated accuracy ranging from approximately 0.95 to 0.97 and macro-averaged F1 scores between 0.96 and 0.97. AdaBoost was a close second, offering a lighter-weight alternative, while a single Decision Tree provided an interpretable baseline. Permutation-based importance analysis revealed that most predictive power is concentrated in ten key variables, primarily indoor temperature, basic environmental readings such as humidity, outdoor temperature, CO_2 levels, and a few physiological indicators, including SpO_2, blood pressure values, and BMI. Models retrained on only these ten features matched, and in some cases slightly outperformed, their full-feature counterparts, while also reducing training time and simplifying interpretation. Paired t-tests confirmed that the accuracy differences between full and reduced models were statistically insignificant. These findings suggest that a lean sensor package is sufficient for reliable real-time comfort prediction, which can lower both hardware costs and computational overhead in smart-building deployments.

Future research should consider several extensions of this work. Temporal architectures such as LSTMs or transformer-based models may be explored to capture short-term trends in environmental and physiological signals, potentially yielding more stable predictions. Incorporating online learning and drift detection mechanisms would allow models to adapt automatically to seasonal changes and evolving occupant behavior. Model-agnostic interpretability techniques (for example, SHAP, LIME) could be applied at the occupant level to generate personalized explanations that foster user trust. Finally, deploying the lean, top-10 feature model on embedded edge hardware would provide a benchmark for latency, power consumption, and on-device inference feasibility, informing large-scale adoption in intelligent-building systems.

References

[1] T. Rus, R.-P. Moldovan, M. I. Pop, and A.-M. Moldovan, "Assessing the interplay of Indoor environmental quality, energy use, and environmental impacts in educational buildings," *Applied Sciences*, vol. 15, no. 7, p. 3591, 2025. https://doi.org/10.3390/app15073591

[2] S. Dimitroulopoulou *et al.*, "Indoor air quality guidelines from across the world: An appraisal considering energy saving, health, productivity, and comfort," *Environment International*, vol. 178, p. 108127, 2023. https://doi.org/10.1016/j.envint.2023.108127

[3] L.-R. Jia, Q.-Y. Li, X. Chen, C.-C. Lee, and J. Han, "Indoor thermal and ventilation indicator on university students' overall comfort," *Buildings*, vol. 12, no. 11, p. 1921, 2022. https://doi.org/10.3390/buildings12111921

[4] H. Xi, B. Wang, and W. Hou, "Machine learning-based prediction of indoor thermal comfort in traditional Chinese dwellings: A case study of hankou lifen," *Case Studies in Thermal Engineering*, vol. 61, p. 105048, 2024. https://doi.org/10.1016/j.csite.2024.105048

[5] J. Park *et al.*, "Field test of machine-learning based mean radiant temperature estimation methods for thermal comfort-integrated air-conditioning control improvement and energy savings," *Energy Reports*, vol. 11, pp. 5682-5702, 2024. https://doi.org/10.1016/j.egyr.2024.05.040

[6] M. Luo, F. Guo, H. Tang, R. Ming, L. Huang, and H. Zhao, "A systematic review of the influence of physiological factors on outdoor thermal comfort," *Human Settlements and Sustainability*, vol. 1, no. 1, pp. 27–40, 2025. https://doi.org/10.1016/j.hssust.2025.02.001

[7] R. Nagy, E. Krídlová Burdová, K. Harčárová, and S. Vilčeková, "Influence of the heating system on the indoor environmental quality—case study," *Buildings*, vol. 12, no. 8, p. 1088, 2022. https://doi.org/10.3390/buildings12081088

[8] I. Nasir, H. Haider, M. Shafiquzzaman, M. Alinizzi, G. Hu, and A. R. Ghumman, "Investigating the effects of occupancy and natural ventilation on the indoor air quality of dormitories in cold regions," *Buildings*, vol. 15, no. 6, p. 896, 2025. https://doi.org/10.3390/buildings15060896

[9] P. Aparicio-Ruiz, E. Barbadilla-Martín, J. Guadix, and J. Nevado, "Analysis of variables affecting indoor thermal comfort in mediterranean climates using machine learning," *Buildings*, vol. 13, no. 9, p. 2215, 2023. https://doi.org/10.3390/buildings13092215

[10] J. Gamero-Salinas, D. López-Hernández, P. González-Martínez, A. Arriazu-Ramos, A. Monge-Barrio, and A. Sánchez-Ostiz, "Exploring indoor thermal comfort and its causes and consequences amid heatwaves in a Southern European city—An unsupervised learning approach," *Building and Environment*, vol. 265, p. 111986, 2024. https://doi.org/10.1016/j.buildenv.2024.111986

[11] P. Romero, V. Valero-Amaro, J. I. Arranz, F. J. Sepúlveda, and M. T. Miranda, "Indoor air quality and thermal comfort in university classrooms in southwestern spain: A longitudinal analysis from pandemic to post-pandemic," *Buildings*, vol. 15, no. 5, p. 829, 2025. https://doi.org/10.3390/buildings15050829

[12] J. M. M. Luque, J. L. S. Jiménez, and M. Ruiz de Adana, "Spatial and temporal distribution of CO2 and thermal comfort conditions in a day care center," *Atmosphere*, vol. 15, no. 12, p. 1500, 2024. https://doi.org/10.3390/atmos15121500

[13] S. Mahmood, H. Sun, E.-S. M. El-Kenawy, A. Iqbal, A. H. Alharbi, and D. S. Khafaga, "Integrating machine and deep learning technologies in green buildings for enhanced energy efficiency and environmental sustainability," *Scientific Reports*, vol. 14, no. 1, p. 20331, 2024. https://doi.org/10.1038/s41598-024-70519-y

[14] A. Kaligambe, G. Fujita, and T. Keisuke, "Estimation of unmeasured room temperature, relative humidity, and CO2 concentrations for a smart building using machine learning and exploratory data analysis," *Energies*, vol. 15, no. 12, p. 4213, 2022. https://doi.org/10.3390/en15124213

[15] L. Chang, I. Permana, F. Wang, and B. Prasetyo, "Improving indoor air quality and thermal comfort using a total heat exchanger ventilation system for an office building," *Thermal Science,*, vol. 28, no. 6, pp. 4531–4544, 2024. https://doi.org/10.2298/TSCI240110118C

[16] A. N. Sayed, F. Bensaali, Y. Himeur, G. Dimitrakopoulos, and I. Varlamis, "Enhancing building sustainability: A Digital Twin approach to energy efficiency and occupancy monitoring," *Energy and Buildings*, vol. 328, p. 115151, 2025. https://doi.org/10.1016/j.enbuild.2024.115151

[17] M. He, H. Liu, Z. Fang, B. He, and B. Li, "High-temperature and thermal radiation affecting human thermal comfort and physiological responses: An experimental study " *Journal of Building Engineering*, vol. 86, p. 108815, 2024. https://doi.org/10.1016/j.jobe.2024.108815

[18] X. Zhou, H. Du, S. Xue, and Z. Ma, "Recent advances in data mining and machine learning for enhanced building energy management," *Energy*, vol. 307, p. 132636, 2024. https://doi.org/10.1016/j.energy.2024.132636

[19] R. Hiroki *et al.*, "Long-term changes of universal thermal climate index (UTCI) estimated from weather stations and gradient-boosted decision trees throughout Japan," *International Journal of Climatology*, vol. 45, no. 8, p. e8843, 2025. https://doi.org/10.1002/joc.8843

[20] F. Han, F. Du, S. Jiao, and K. Zou, "Predictive analysis of a building's power consumption based on digital twin platforms," *Energies*, vol. 17, no. 15, p. 3692, 2024. https://doi.org/10.3390/en17153692

[21] R. Levinson *et al.*, "Hot, cold, or just right? An infrared biometric sensor to improve occupant comfort and reduce overcooling in buildings via closed-loop control," *Energy and Buildings*, vol. 312, p. 114063, 2024. https://doi.org/10.1016/j.enbuild.2024.114063

[22] X. Chen, H. Zhao, B. Wang, and B. Xia, "Study of factors influencing thermal comfort at tram stations in Guangzhou based on machine learning," *Buildings*, vol. 15, no. 6, p. 865, 2025. https://doi.org/10.3390/buildings15060865

[23] Silvestri A. et al, "Real building implementation of a deep reinforcement learning controller to enhance energy efficiency and Indoor temperature control," *Applied Energy*, vol. 368, p. 123447, 2024. https://doi.org/10.1016/j.apenergy.2024.123447

[24] Khan H. R. et al, "Design and experimental results of an AIoT-enabled, energy-efficient ceiling fan system," *Sustainability*, vol. 16, no. 12, p. 5047, 2024. https://doi.org/10.3390/su16125047

[25] Y. Song and Y. Lu, "Decision tree methods: Applications for classification and prediction," *General Psychiatry*, vol. 27, no. 2, 2015. https://doi.org/10.11919/j.issn.1002-0829.215044

[26] Y. Ding, H. Zhu, R. Chen, and R. Li, "An efficient AdaBoost algorithm with the multiple thresholds classification," *Applied Sciences* vol. 12, no. 12, p. 5872, 2022. https://doi.org/10.3390/app12125872

[27] Y. Zhao and H. Zhao, "A hybrid machine learning framework by incorporating categorical boosting and manifold learning for financial analysis," *Intelligent Systems with Applications,* vol. 25, p. 200473, 2025. https://doi.org/10.1016/j.iswa.2024.200473