IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 20 November 2023, accepted 12 January 2024, date of publication 14 February 2024, date of current version 29 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3366453

RESEARCH ARTICLE

Improved Thermal Comfort Model Leveraging Conditional Tabular GAN Focusing on Feature Selection

MD. SHAJALAL[1,2], MILAD BOHLOULI[2,3], HARI PRASANNA DAS[4], ALEXANDER BODEN[1,5], AND GUNNAR STEVENS[2,5]

[1]Fraunhofer-Institute for Applied Information Technology (FIT), 53757 Sankt Augustin, Germany
[2]IT-Security and Consumer Computing, University of Siegen, 57072 Siegen, Germany
[3]Petanux GmbH, 53129 Bonn, Germany
[4]Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, CA 94720, USA
[5]Institute for Consumer Informatics, Bonn-Rhein-Sieg University of Applied Sciences, 53757 Sankt Augustin, Germany

Corresponding author: Md. Shajalal (md.shajalal@fit.fraunhofer.de)

This work was supported by the European Union (EU) Horizon 2020 Marie Sklodowska-Curie International Training Network GECKO under Grant 955422.

ABSTRACT The indoor thermal comfort in both homes and workplaces significantly influences the health and productivity of inhabitants. The heating system, controlled by Artificial Intelligence (AI), can automatically calibrate the indoor thermal condition by analyzing various physiological and environmental variables. To ensure a comfortable indoor environment, smart home systems can adjust parameters related to thermal comfort based on accurate predictions of inhabitants' preferences. Modeling personal thermal comfort preferences poses two significant challenges: the inadequacy of data and its high dimensionality. An adequate amount of data is a prerequisite for training efficient machine learning (ML) models. Additionally, high-dimensional data tends to contain multiple irrelevant and noisy features, which might hinder ML models' performance. To address these challenges, we propose a framework for predicting personal thermal comfort preferences, combining the conditional tabular generative adversarial network (CTGAN) with multiple feature selection techniques. We first address the data inadequacy challenge by applying CTGAN to generate synthetic data samples, incorporating challenges associated with multimodal distributions and categorical features. Then, multiple feature selection techniques are employed to identify the best possible sets of features. Experimental results based on a wide range of settings on a standard dataset demonstrated state-of-the-art performance in predicting personal thermal comfort preferences. The results also indicated that ML models trained on synthetic data achieved significantly better performance than models trained on real data. Overall, our method, combining CTGAN and feature selection techniques, outperformed existing known related work in thermal comfort prediction in terms of multiple evaluation metrics, including area under the curve (AUC), Cohen's Kappa, and accuracy. Additionally, we presented a global, model-agnostic explanation of the thermal preference prediction system, providing an avenue for thermal comfort experiment designers to consciously select the data to be collected.

INDEX TERMS Personal thermal comfort, generative adversarial network, feature selection, machine learning, data inadequacy.

I. INTRODUCTION

Occupants' well-being, health, and productivity significantly depend on thermal comfort both at home and in the workplace [1], [2], [3], [4], [5], [6]. A notable portion of the total energy consumption is attributed to the HVAC (heating, ventilation, and air conditioning) system, accounting for nearly half of the overall energy use in corporate and residential buildings [5]. Additionally, these buildings contribute to almost 40% of CO_2 gas emissions [1], [5]. The advancements in sensor technology over the last two decades have played a crucial role in shaping the concept of smart home systems

The associate editor coordinating the review of this manuscript and approving it for publication was Gina Tourassi.

to reality, empowering inhabitants to control and monitor the indoor environment within their homes and workplaces [2], [3], [4], [5], [6], [7]. Environmental parameters related to thermal comfort, such as temperature and humidity, can be adjusted using multiple machine learning-based systems with human-in-the-loop interaction [1].

In general, artificial intelligence (AI)-based techniques can be applied to have energy-efficient and comfortable indoor environment inside buildings [8], [9]. It is also evident that researchers often leveraged AI-enabled techniques for energy aware and comfortable built environment. However, the primary objective is to save energy and decrease the carbon-di-oxide footprints. The notable smart home energy-aware applications that generally applied deep learning (DL) and ML models can be energy demand forecasting, adjusting indoor environment by predicting thermal comfort preferences, etc [1]. However, in this work we focus on personal thermal comfort preference prediction. Recently, there has been a considerable attention in applying ML models for thermal comfort preference prediction tasks [1], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20].

Generally, the task of personal thermal comfort preference prediction can be classified into two different categories, global and personal. In global thermal comfort (GTC) preference prediction task, the model tries to predict the overall thermal comfort preference in the rooms/zones. On the other hand, since the thermal comfort of different person varied widely, personal thermal comfort (PTC) preference prediction refers to identifying an occupant's individual thermal comfort [1]. Based on the preference prediction system's output, smart home systems can control and adjust the environment to provide pleasant and comfortable living space.

In most of the existing studies [2], [10], [15], [18], authors applied their predictive models to high dimensional features to capture relation between the data and occupants' thermal preference. There are two major challenges associated with modeling the personal thermal comfort preference prediction: one is the high dimensionality of the data including environmental and physiological features, and other one is the lake of adequate amount of data samples to train efficient predictive ML model. This is expected that the data to predict thermal comfort preference will be high dimensional, since it considers every possible attributes that are related to the occupants indoor HVAC comfort. On the other hand, the data collection from real subjects with right annotation procedure is very time consuming and costly.

Generally, ML models needs adequate data to train and this is a prime requirement in any predictive models. To mitigate the data availability problem, an effective synthetic data generation technique addressing associated challenges can be a game changer. The high-dimensionality might be the curse in modeling indoor thermal comfort preference. Because there might have some features that are not relevant and can even downgrade the performance of the predictive model. Hence, identifying the possible relevant set of features is a prerequisite of the system with high-dimensional data.

In this research, we propose a new indoor thermal comfort preference prediction system by addressing the above-mentioned challenges by incorporating CTGAN and multiple feature selection techniques. First, we address the data inadequacy challenge by employing one of the most successful synthetic data generation techniques that incorporate the multi-modal distribution in the numeric features with mode-specific normalization technique. In addition with the data adequacy problem, datasets related to the PTC preference are generally imbalanced, which might make the performance biased towards the majority class samples. By incorporating CTGAN, we address also the data imbalance problem by synthetically generating the data for minority class samples.

The best set of relevant features generally provides high performance in predictive modeling in case of high-dimensional datasets. Before applying feature selection techniques, we conducted experiments to determine whether highly correlated features exist in the PTC dataset. Our hypothesis for this experiment was that if we found more correlated features related to PTC preference prediction, we could then make use of feature selection techniques to filter out irrelevant, noisy, and redundant features. To achieve this, we carried out experiments on a PTC preference prediction dataset [2]. The correlation among the 82 features, based on Pearson's correlation coefficient analysis, is illustrated in the heatmap representation in Fig. 1.

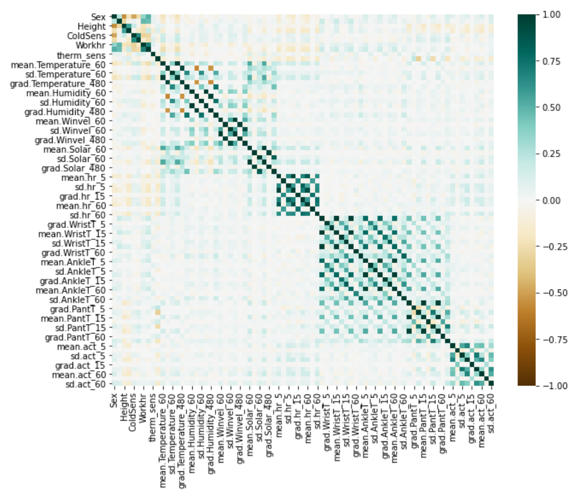

FIGURE 1. Heatmap depicting correlation coefficient among different features.

We observed that more than 20 features out of the 82 different physiological, environmental, and weather features are highly correlated ($\rho \geq 0.80$) [1]. With these findings, we developed the idea to apply various feature selection techniques to filter out irrelevant features. Our previous paper, based on the preliminary findings on the effect of applying feature selections, is published in the proceedings of the ACM International Conference on Systems for

Energy-Efficient Buildings, Cities, and Transportation (ACM BuildSys 2022) [1]. Inspired by the impressive preliminary findings, this extended research applied multiple feature selection techniques and introduced CTGAN to generate synthetic data samples for effectively training ML models.

We conducted experiments by training six different ML models on a standard PTC prediction dataset by generated synthetic data samples employing CTGAN for thermal comfort prediction. We leveraged the best-selected feature set, applying multiple feature selection techniques, for training the ML models. Since the dataset was imbalanced, we carefully utilized multiple evaluation metrics, including *Cohen's Kappa* and *Area Under the Curve (AUC)*, along with *accuracy*, that can measure the prediction performance a ML model for imbalance data distribution.

The experimental results, encompassing a wide range of settings, demonstrated the superiority of the proposed method with synthetically generated data focusing on feature selection techniques. The performance of predictive models trained on synthetic data significantly outperformed the baseline, as well as models trained on real data with feature selections. Compared to known related works, our methods also achieved much higher performance in terms of all evaluation metrics (AUC, Kappa, and Accuracy). The contributions of this research are summarized as follows:

- We introduced CTGAN, a synthetic data generation technique, to address the problem of data inadequacy by generating new personal thermal comfort data for individuals.
- We employed multiple feature selection techniques to identify the best possible set of relevant features for effectively modeling personal thermal comfort preference prediction.
- The experimental results demonstrated the superiority of our framework in modeling thermal comfort preference prediction. We achieved significantly higher performance after applying feature selection techniques and CTGAN, a synthetic data generation technique. The combination of both techniques showed a significant improvement in performance compared to known related methods.

In the remainder of the paper, we present state-of-the-art on personal thermal comfort preference prediction in section II. We then present our proposed thermal comfort modeling framework combining on CTGAN and feature selection techniques in section III. In section IV, we discuss about the dataset, evaluation metrics, experimental design and the findings by demonstrating results for wide range of experiments. Finally, we conclude our findings with future research direction in section V.

II. LITERATURE REVIEW

The prior works on modeling PTC preferences are associated with the experiments on the data collected from living labs [21], [22], [23], [24], [25], [26], [27]. Generally, a large number of features are included, often the amount of data samples available after cleaning the data of any missing value in the features decreases [21], [24], [27], especially for experiments including physiological metrics for an occupant. Consequently, classical ML models often have more predictive power compared to deep learning based models [28], [29], [30], [31]. Therefore, unlike the latter models, which automate the feature engineering by learning from the data, with the classical models, feature engineering plays a key role in the predictive power.

However, most of the datasets for PTC preference prediction task are small in sample size and the sample distributions among different classes are quite imbalanced. Since ML models need adequate data for learning the pattern from the samples, it is challenging to train models with small dataset. In addition, collecting big dataset from the participants in a living lab setting is quite time consuming and costly. Therefore, synthetically generating data samples on the available data has got considerable attention in thermal comfort modeling research in recent time [32], [33], [34], [35], [36].

To address the data inadequacy challenges in thermal comfort preference prediction task, several methods has been proposed that generate synthetic data [32], [33], [34]. Synthetic data generation techniques are also often applied for balancing the data [37], [38]. Quintana et al. [32] employed conditional generative adversarial networks to generate synthetic data samples for minority class to address class imbalance problem. Similarly, conditional Wasserstein GAN has been applied by Yoshikawa et el. [33] for balancing the thermal comfort preference prediction dataset. Das et al. [34] also applied basic GAN architecture for the same purpose.

Evidently, the inclusion of redundant features degrades the performance of the ML model [39]. Feature selection has also been of interest of similar fields, e.g. in occupancy prediction where the objective is to predict the occupancy count of the rooms in a building, adaptive lasso feature selection has been used to select the most relevant features [40]. Similarly in [41] authors use genetic algorithms for feature selection.

Feature selection techniques based on manual observation that evaluate the best combination by the prediction performance of the trained model. For instance in [21], authors tried out the various combinations of the input features and concluded that skin temperature and heating settings are the best predictors for thermal comfort. In a similar study [27], authors examined the skin temperature at 6 points on the body and evaluated the various combinations, besides also proposing a new feature representative of the body's average temperature based on Ramanathan's formula as a combined feature [27]. In another study [42], authors defined 3 feature sets and evaluated their predictive power via precision and recall. The above-mentioned approaches, despite being the most prevalent approaches adopted in the literature, require significant background knowledge regarding the features and significant manual labor. Also, the manual approach might

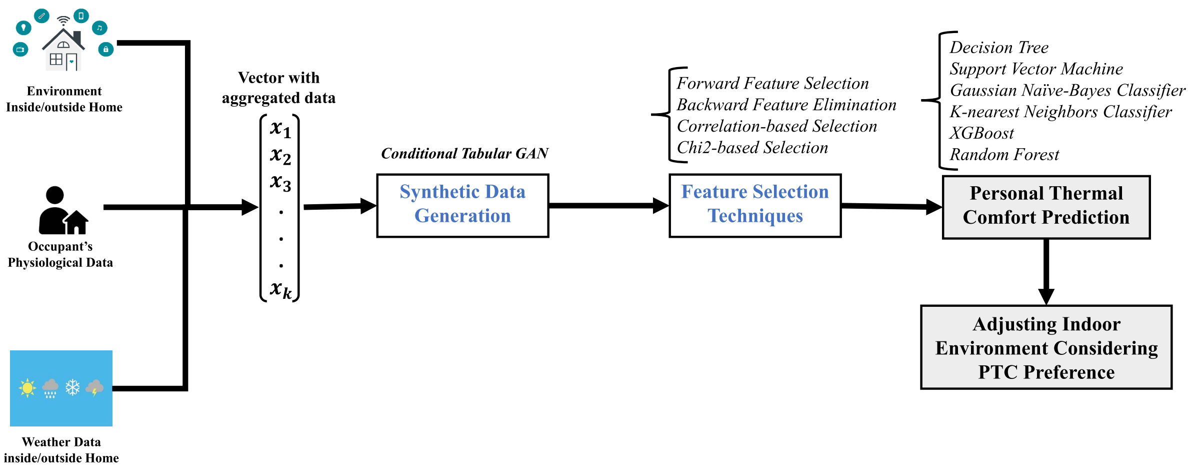

FIGURE 2. A high-level building block of the proposed PTC preference prediction with synthetic data using CTGAN focusing on feature selection.

not necessarily result in the best combination, especially for datasets with high dimensions.

Another selection method is related to the prior knowledge from thermal comfort domain and literature. For example, in [43], authors have defined new features based on the ASHRAE standard [44] which would be representative for model structure and heat balance of the body. In another study [45], authors derived new features based on the polynomial basis function to capture the relation between the environmental features and thermal perception. Although these groups of studies may introduce new crucial features that may promote the predictive power, they still have to employ a selection process to omit the less productive features. Similar to the first group of the studies, this approach also requires prior knowledge while introducing new features.

One of the most classical ways is to apply feature selection techniques to find the relation between the input variables and the target variable. For example, in the study conducted in [46], authors tried to find the relation between the thermal sensation and the air quality via multi-linear regression and hypotheses testing. However, in their study they do not differentiate between heating and cooling. Similarly, in [19], authors used Lasso feature selection to select among the features that collected from an experimental study. Additionally, in the study done in [2], authors employed Pearson correlation coefficient among the features and the target to measure the importance of the introduced features.

In this research, we applied conditional tabular GAN [35], [36] inspired by the success achieving new benchmark. CTGAN can solve the multimodal distributions in the numeric data points by incorporating mode-specific normalization technique and address challenge associated with categorical features applying variational Gaussian-Mixture model [35], [36]. Adequate training data, complemented by synthetic data generated through CTGAN, ensures the effective training of ML models. Furthermore, this can address the issue of data imbalance in the training set, ensuring that the model does not exhibit bias towards any majority class. On the other hand, we incorporated multiple feature selection techniques to identify the best set of features that help the ML models to achieve higher performance in thermal comfort preference prediction.

III. METHODOLOGY

The overview of our proposed PTC preference prediction framework is illustrated in Fig. 2. We first generate synthetic data applying CTGAN and then apply four different feature selection techniques to identify the best possible sets of features. With the selected features, we employed six different ML classification models to predict occupants' thermal comfort preference.

A. SYNTHETIC DATA GENERATION WITH CTGAN

A significant challenge in developing predictive models for PTC preferences arises from the scarcity of adequate data. The dataset [2] utilized in our experimental work, as described in greater detail in Section IV-A was sourced from a group of 14 individuals. These subjects participated in data collection activities that involved the annotation of their thermal comfort preferences while residing in living laboratories located in Berkeley and San Francisco.

It is worth noting that acquiring large datasets specially for training of state-of-the-art ML models for PTC preference prediction is expected to be very expensive and time consuming. Consequently, there is potential value in generating synthetic data through the application of robust data generation techniques. This approach is inspired by the remarkable achievements of generative adversarial networks (GAN) when applied to tabular data, as evidenced by prior works [35], [36]. Specifically, we explore the application of conditional tabular GAN (CTGAN), which offers particular advantages in addressing challenges related to mixed data types and multi-modal distributions when generating synthetic tabular data.

Unlike other GAN-based methods including WGAN [47] and WGAN-GP [48], CTGAN can capture the heterogeneity

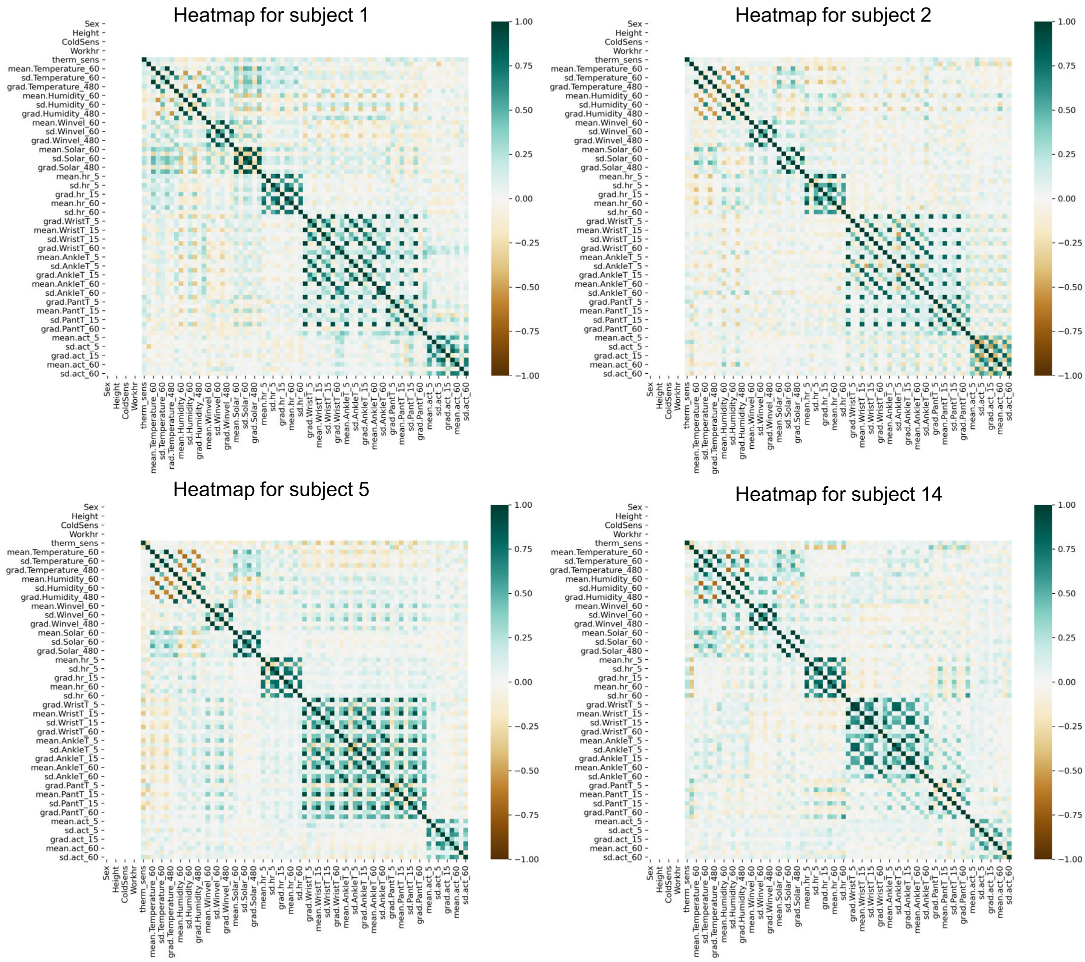

FIGURE 3. Heatmaps for four different subjects highlighting the correlation among different features.

of the real-world data [35], [36]. To handle mixed data in creating synthetic data, CTGAN developed a full workflow from data preprocessing to modifying GAN architecture. The major challenge that CTGAN solved is non-Gaussian multimodal distribution by introducing a mode-specific normalization technique. It handles this problem by following multimodal distributions. By applying a variational Gaussian mixture model (VGM), it can represent each continuous real-valued feature in a one-hot vector that indicates the sampled mode and the normalized value [35], [36]. To tackle challenges posed by categorical features, CTGAN introduced the sparsity of one-hot-encoded vectors in real-valued data with probability distributions [35], [36]. Further, it introduced a conditional data generator that gets ride of the challenges posed by multimodal and imbalanced data distributions. The detail description of CTGAN can be found in [35] and [36].

B. FEATURE SELECTION TECHNIQUES

Classifiers are often misled by redundant, correlated and noisy features. In case of a high dimensional dataset, selecting best features' set before applying classifier would be a better approach for modeling thermal comfort. In this section, we will outline our visual exploration of feature redundancy. This exploration indicates the usefulness of feature selection techniques to filter out less relevant features. Next we introduce four feature selection techniques in our study to filter out redundant and correlated features to improve the performance of PTC prediction model.

1) VISUAL EXPLORATION

We explore the correlation among different features across occupants numerically as well as visually to detect patterns in them. Heatmaps in Fig. 3, show the feature correlation for four different occupants. We can see that there are a substantial number of features that are correlated to each other. However, we can also see that the correlation coefficients among different features are quite different among different occupants. These observations and findings illustrate why PTC prediction is a challenging task. The detailed view also shows that there are some common patterns and correlations among some features across the occupants. With this preliminary analysis, we hypothesize that the elimination of these correlated and redundant features might improve the PTC prediction model.

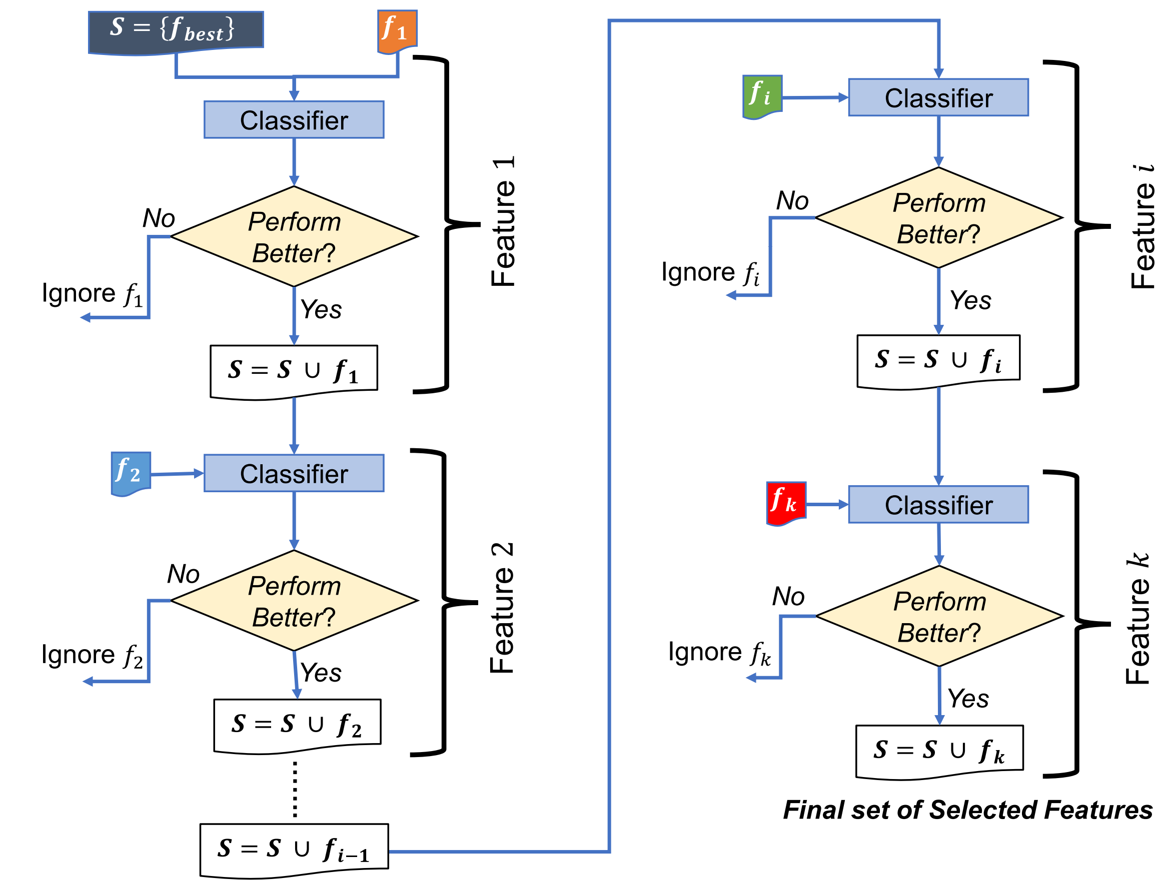

FIGURE 4. **The workflow of the forward feature selection (FFS) technique (Figure created based on [1]).**

2) CORRELATION-BASED FEATURE SELECTION

Since one of our primary objectives is to filter out irrelevant features before applying classifiers, we conducted a correlation analysis across all features. In correlation-based feature selection, we consider a feature as redundant if it has a high correlation coefficient ($\rho \geq 0.80$) and remove it from the list [1]. We compute the Pearson correlation coefficient between two features as follows:

$$\rho(f_a, f_b) = \frac{\sum_{i=1}^{n} (f_{a_i} - \bar{f_a})(f_{b_i} - \bar{f_b})}{\sqrt{\sum_{i=1}^{n} (f_{a_i} - \bar{f_a})^2 \cdot \sum_{i=1}^{n} (f_{b_i} - \bar{f_b})^2}} \quad (1)$$

where f_a and f_b are two features from the list of features $F = \{f_1, f_2, f_3, \ldots, f_k\}$. The average feature values for two different features f_a and f_b are denoted by $\bar{f_a}$ and $\bar{f_b}$, respectively [1].

3) CHI-SQUARE TEST-BASED FEATURE SELECTION

The target of this technique is to select those features, which have higher dependency with the response. In statistics, Chi-Square test is a prominent technique applied to test the independence of two different events. In this research, however, we employed Chi-Square test as a tool to select the best set of features. Given two different variables, Chi-square test computes how the expected count E deviates from the observed count O for those two variables. The computation is done as $\mathcal{X}^2 = \sum \frac{(O_i - E_i)^2}{E_i}$. We employ this test to determine the relationship between specific features and the labeled response. Chi-Square will return a smaller value when two features are independent. In other words, the observed count is close to the expected count. Hence, the higher the Chi-Square value the feature is more dependent on the response and that feature should be selected for training the model. We applied iterative approach to have Chi-Square test for each feature and selected the best features' set. To select an optimal value for the number of selected features, we make use of a grid search. The details on parameter tuning is presented in section IV-D1.

4) SUPERVISED FORWARD FEATURE SELECTION

We applied supervised forward feature selection (FFS), which enables the selection of more insightful features through a greedy iterative selection procedure. We present the FFS procedure in Fig. 4 [1]. This approach first utilizes every feature individually and applies a baseline classifier to predict occupants' personal thermal comfort. By comparing the performance of all individual features, it selects the best-performing one as f_{best}. The first selected feature f_{best} is then added to the selected feature set S. Subsequently, FFS then combines the remaining features f_i one at a time to the selected feature set S and applies the classifier separately [1].

Considering the performance of the classifier, FFS selects the feature f_i if the combination achieved better performance than the previous classifier with the already selected feature set S. This greedy approach continues for the rest of the features. Applying this FFS approach allows us to select the best set of features that are effective in predicting occupants' PTC preferences efficiently [1].

5) SUPERVISED BACKWARD FEATURE ELIMINATION

The working principle of backward feature elimination (BFE) is the opposite of forward feature selection. Unlike the forward feature selection approach, it first applies all the features feeding to a classifier to model PTC and then computes the classification performance. After computing the performance, it iteratively discards one feature at a time and checks whether the performance of the model increases or decreases without that feature. If the performance decreases, then it hypothesizes that the feature has an important role in modeling PTC. Consequently, it includes that feature in the list of important features. In the opposite case, it ignores the feature and discards it as less relevant.

IV. EXPERIMENTS

This section presents the wide range of experimental setups and performance evaluations of our proposed methods that validate the efficiency in modeling personal thermal comfort in terms of multiple evaluation metrics on a PTC dataset.

A. DATASET

We conducted experiments on a PTC preference prediction dataset collected by Liu et al. [2]. The dataset was collected and annotated by 14 different subjects living in the areas of Berkeley and San Francisco. During the study, the authors measured the skin temperature from different parts of the subjects' bodies and the surrounding room temperature where the subjects were present. Additionally, they also measured the activity and heart rate of the subjects using accelerometers and polar sensors, respectively. The experiments spanned 14 days, with each subject expected to provide their thermal comfort preference 12 times a day, categorized as "Cooler," "Warmer," or "No Change." Out of the collected 3848 samples, the distribution across different classes is quite imbalanced. Fig. 5 illustrates the percentage of samples across different classes, showing that 68.5% of samples are for "No Change," while 16.5% and 15% are for "Cooler" and "Warmer," respectively.

B. DATA PRE-PROCESSING

The values of the features in the dataset vary widely in terms of their units and ranges. In addition, there are some missing values which we tackled by applying median values. After that we applied min-max normalization [49] to map each variables' values to a certain range [0,1]. We also grouped the dataset based on individual occupant and analyzed the features for individual occupants.

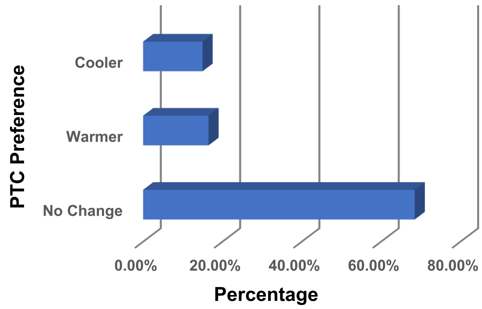

FIGURE 5. Distribution of samples over PTC preferences.

C. EVALUATION METRICS

In the assessment of any ML model, it is crucial to consider the evaluation metrics with respect to the characteristics of the dataset. According to the literature on PTC models [4], Accuracy, AUC (Area Under Curve), and Cohen's kappa are the three widely used evaluation metrics.

The validity and usefulness of evaluation metrics also depends on the specific domain and characteristics of the datasets. In our case, it is essential that the evaluation metric are sensitive to the class-imbalance. For instance, using accuracy alone will be problematic, since this metric will not reflect the class-wise prediction performance. Considering the imbalance distribution in the personal thermal comfort datasets, a model classifying all the samples as "no change" with an 80% share in the original dataset, would result in an 80% accuracy, which does not necessarily suggest the strength of the classification.

In the following formulations, True Positive (TP) is an outcome where the model correctly predicts the positive class, True Negative (TN) is an outcome where the model correctly predicts the negative class, False Positive (FP) is an outcome where the model incorrectly predicts the positive class, False Negative (FN) is an outcome where the model incorrectly predicts the negative class.

1) ACCURACY

As shown in equation 2, *Accuracy* only requires the class labels for evaluation and does not examine the separability strength of the model. Nonetheless, it has also been reported to be compared with the previous studies.

$$accuracy_{class1} = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

2) AREA UNDER THE CURVE (AUC)

In contrast to Accuracy, AUC [50] also considers how well the predicted classes are distinguished, by taking the prediction probabilities of each class into account. Technically, AUC is the area under the curve of the ROC (Receiver Operative Characteristics) which is the representation of TPR (True Positive Rate) with respect to the FPR (False Positive Rate), defined in equations 3 and 4, respectively when the decision boundary is moved through the data points.

$$TPR = \frac{TP}{TP + FP} \tag{3}$$

$$FPR = \frac{FP}{TN + FP} \tag{4}$$

Fundamentally, this metric is proposed for binary classification problem, however, in order to apply it to the multiclass cases, the One vs Rest approach has been used.

3) COHEN'S KAPPA

Cohen's Kappa [50] is often an under-utilized, but quite useful metric, which also considers the prediction probabilities of each class. It can be defined as follows:

$$kappa = \frac{P_0 - P_c}{1 - P_c} \tag{5}$$

$$P_0 = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

$$P_c = P(\text{“}Positive\ Classified''') + P(\text{“}Negative\ Classified'') \tag{7}$$

$$P(\text{“}Positive\ Classified'') = \frac{TP + FP}{TP + TN + FP + FN} \tag{8}$$

$$P(\text{“}Negative\ Classified'') = \frac{TN + FN}{TP + TN + FP + FN} \tag{9}$$

Cohen's kappa considers the quantity of the classes. More precisely, it consider the probability of classes being changed, defined as P_C (the observed agreement) and P_0 (the expected agreement). It varies from 0 to 1, with 0 being a random classification.

D. FEATURE SELECTION

We applied the four feature selection techniques described in section III. The number of selected features differ among selection techniques.

1) CHI-SQUARE TECHNIQUE'S PARAMETER TUNING

The Chi-Square-based feature selection technique requires to tune the parameter k, the number of selected features. We applied a grid search to identify the optimal number of selected features and evaluated the performance. The experimental results in terms of AUC are illustrated in Fig. 6. The figure concludes that the optimal number of features (highest AUC) should be $k = 17$ and the performance with those selected features are on the y-axis. Therefore, we select $k = 17$ and apply this parameter value in our chi-square-based feature selection.

2) RESULT OF FEATURE SELECTION TECHNIQUES

None of the introduced feature selection techniques need parameter tuning except the Chi-Square test. For correlation-based selection, we make use of $\rho >= 0.80$ (Eq. 1) as highly correlated features. In turn, we could apply other techniques straight forward. Table 1 presents the detailed results of our feature selection techniques with the number of selected features. The four feature selection techniques chi-square based, correlation based, forward feature selection and backward feature elimination are denoted as Chi-Square, Correlation, FFS and BFE, respectively. As outlined above, in the case of Chi-Square-based feature selection technique, the number of selected features was a result of our parameter tuning. We can see in Table 1, some features are common in all selected feature sets. From 82 different features, we enlisted here the most notable 19 features. Of these 19 features, some features have different varieties i.e., the temperature has different varieties such as mean, gradient, and standard deviation of different time slots.

Similar to temperature, skin temperature on the wrist and ankle, wind speed, and wrist acceleration also have some varieties. Here in this list, we checked if any one of the varieties is selected by the feature selection technique. However, we can see that *age, height, thermal sensitivity, temperature, working hour, heart rate, weight, and subject location* are the common features that all the selection techniques selected as relevant. Other than that, the skin temperature of the wrist, ankle, and body proximity temperature are important features considered by the three selection

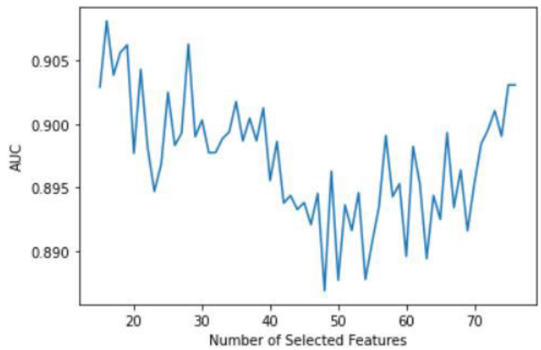

FIGURE 6. Tuning the parameter k, number of selected features in Chi-Square feature selection using grid search.

TABLE 1. Results of applying feature selection techniques and notable selected features.

Features	Chi-Square	Correlation	FFS	BFE
# of features	17	59	27	32
Age	x	x	x	x
Height	x	x	x	x
Therm_sens	x	x	x	x
Temperature	x	x	x	x
Workhr	x	x	x	x
Heart rate	x	x	x	x
Sex	-	x	-	x
ColdSens	-	x	x	x
ColdExp	-	x	-	x
Coffeeintake	-	x	-	x
Location	x	x	x	x
Humidity	-	x	-	x
WristT	x	x	-	x
Winvel	-	x	x	-
AnkleT	-	x	x	x
PantT	-	x	x	x
Solar	x	x	x	x
Act	x	-	x	x

TABLE 2. Performance of applying feature selection techniques in different ML models trained on real data in global thermal comfort prediction. The best results for each feature selection techniques are in bold. The blue-colored values indicate the best performance among all experimental settings.

Feature Selection	Model	Kappa	Accuracy	AUC
Chi2-based Selection	DT	0.6331	0.8374	0.9229
	SVM	0.5977	0.8270	0.9029
	KNN	0.4966	0.7906	0.8574
	BNB	0.4644	0.7464	0.8199
	XGB	**0.6585**	0.8491	**0.9401**
	RF	0.6567	**0.8517**	0.9378
Correlation-based Selection [1]	DT	0.6551	0.8348	0.8964
	SVM	0.5054	0.8023	0.9101
	KNN	0.2830	0.7308	0.7802
	GNB	0.2825	0.5825	0.7267
	XGB	**0.6830**	**0.8595**	**0.9370**
	RF	0.5417	0.8127	0.9167
Forward Feature Selection [1]	DT	0.6471	0.8426	0.8852
	SVM	0.5441	0.8127	0.9167
	KNN	0.3939	0.7646	0.8184
	GNB	0.4627	0.7256	0.7923 5
	XGB	**0.6782**	**0.8569**	0.9362
	RF	0.6109	0.8374	**0.9380**
Backward Feature Elimination	DT	0.6709	0.8426	0.9007
	SVM	0.5474	0.8140	0.9227
	KNN	0.3733	0.7542	0.8029
	GNB	0.4291	0.7113	0.7740
	XGB	**0.6717**	**0.8556**	**0.9438**
	RF	0.5895	0.8296	0.9375

techniques. Similarly thermal sensitivity, cold sensitivity and cold extremity experience also came out to be important in modeling PTC preference.

E. EXPERIMENTAL SETTING

We first applied feature selection techniques and trained the ML models on original data samples collected from 14 different subjects. For all of these feature sets selected based on selection criteria, we carried out a range of experiments to validate the performance of our introduced feature selection techniques. At first, we applied our methods on the whole dataset combining samples from all 14 subjects. The primary intuition was to observe how our feature selection methods perform on overall thermal comfort dataset. In other words, evaluating our methods on the global thermal comfort (GTC) preference in buildings. Then we applied the similar experimental setting to train ML models on synthetically generated data by CTGAN with the selected relevant features sets.

Finally, with the selected features leveraging four different feature selection techniques, we applied six different classical classifiers to model the thermal comfort of the occupants. Then, we applied two best performing models to observe the performance on PTC preference prediction for individual subject. By doing so, we applied our method on 14 different subjects' collected samples separately to model PTC. The remainder of the section presents the experimental results for modeling global and personal thermal comfort and performance comparison with prior works. We repeat all experiments for GTC and PTC with training the models with our generated synthetic data by conditional tabular generative adversarial networks for modeling thermal comfort preference prediction.

F. GLOBAL THERMAL COMFORT PREDICTION PERFORMANCE

We designed the experiments in a way to visualize the performance of the feature selection techniques in modeling thermal comfort preference with the trained models on real data. As we noted earlier that the data might not be adequate to train ML models, we employed CTGAN to generate quality synthetic data and combine it with real data for training the model with feature selection. Therefore, we first illustrate the performance of different feature selection techniques on real data and then we present the results on applying CTGAN for synthetic data generation.

1) PERFORMANCE ON REAL DATA

The performance of six different classifiers trained on real data samples of the global thermal comfort dataset with our introduced feature selection techniques is summarized in Table 2. With the selected features for each feature selection technique, we applied six different classifiers: decision tree (DT), support vector machine (SVM), K-nearest neighbor (KNN), Gaussian Naive Bayes (GNB), XGBoost (XGB), and random forest (RF). This results in a total of 24 experimental setups.

Table 2 shows that XGBoost with the correlation-based feature selection technique achieved better performance among all experimental settings (highlighted in blue) in terms of Cohen's Kappa and Accuracy. However, in terms of AUC, the XGBoost model with the backward feature elimination technique obtained the best performance. Among the six different ML models, the XGBoost model is consistently better across all feature selection techniques, except for Forward feature selection and Chi-square-based selection

TABLE 3. Performance of applying feature selection techniques in different ML models trained on synthetic data with CTGAN in global thermal comfort prediction. The best results for each feature selection techniques are in bold. The blue-colored values indicate the best performance among all experimental settings.

Feature Selection	Model	Kappa	Accuracy	AUC
Chi2-based Selection	DT	0.6400	0.7629	0.8857
	SVM	0.6944	0.7983	0.9245
	KNN	0.5410	0.7020	0.8716
	GNB	0.5746	0.7206	0.8779
	XGB	**0.8265**	**0.8854**	**0.9734**
	RF	0.8054	0.8713	0.9638
Correlation-based Selection	DT	0.5986	0.7366	0.8760
	SVM	0.7449	0.8320	0.9404
	KNN	0.5469	0.7067	0.8702
	GNB	0.5827	0.7261	0.8671
	XGB	**0.8483**	**0.8998**	**0.9786**
	RF	0.8008	0.8686	0.9602
Forward Feature Selection	DT	0.6537	0.7687	0.8836
	SVM	0.7885	0.8603	0.9592
	KNN	0.5265	0.6957	0.8859
	GNB	0.5676	0.7170	0.8715
	XGB	**0.8872**	**0.9253**	**0.9890**
	RF	0.8282	0.8860	0.9765
Backward Feature Elimination	DT	0.6673	0.6673	0.8973
	SVM	0.7983	0.7983	0.9606
	KNN	0.4954	0.6769	0.8730
	GNB	0.6164	0.7485	0.8954
	XGB	**0.8802**	**0.9208**	**0.9885**
	RF	0.8455	0.8979	0.9788

techniques. In both cases, random forest (RF) acquired the highest accuracy for the Chi-square-based selection procedure and the best AUC for the forward feature selection technique. However, it is also observed that the performance difference among correlation-based, forward feature selection, and backward feature elimination is not significant. We can broadly say that among all six different classifiers, XGBoost and random forest are the two best-performing models across all feature selection techniques.

2) PERFORMANCE ON SYNTHETIC DATA.

The performance of introduced feature selection techniques with the synthetically generated data employing the prominent conditional tabular GAN is presented in Table 3. The result shows clearly that the performance has been improved significantly for all ML models with the trained models using synthetic data. The best performing model among all experimental settings is XGBoost that used the selected features based on forward features selection techniques, numerically the performance 20%, 7% and 5% higher than the best performing model trained on original data (Table 2) in terms of Kappa, Accuracy and AUC, respectively. Based on AUC score, XGBoost model on real data with backward feature selection technique achieved highest performance (numerically, 0.9438 (Table 2)) and on contrary the XGBoost model trained on synthetic data generated by CTGAN is way more higher than the the performance (numerically, 0.9885) on same model trained with original data.

The superiority in modeling thermal comfort preference with synthetically generated data is visualized in Fig. 7 and 8 using bar and line chart. In both figures, we highlight the difference in achieving the higher performance in terms of most significant evaluation metrics AUC between model trained with original data and with synthetic data for all 24 experimental settings. In Fig. 7, the black colored bar denotes the performance for trained models on synthetically generated data by CTGAN and the gray colored bars represent the performance for models on original data. Similarly, the orange and blue colored line in Fig. 8 represent the similar performance of models with and without synthetic data.

From both the figures we can see that, the models trained with synthetic data demonstrated higher performance compared to all experimental settings except two models, decision tree for Ch2-based and correlation-based feature selection technique. Fig. 8 clearly illustrates the performance improvements after integrating CTGAN based synthetic data generation techniques. However, based on the facts and findings discussed above we can conclude that feature selection on synthetic data can achieve higher performance in predicting thermal comfort preference and that can be used to calibrate the indoor environment. Hence, this might provide more occupant-friendly environment that can be both healthy and energy efficient in smart home.

G. PERFORMANCE ON PTC PREFERENCE PREDICTION

To predict occupants' PTC preferences, we trained two ML models, including XGBoost and Random Forest, which are the two best-performing models on real and synthetically generated data. We considered the selected features by applying backward feature elimination techniques and trained both models for each subject separately on synthetically generated data leveraging CTGAN. The experimental results in predicting PTC preferences for all 14 different subjects

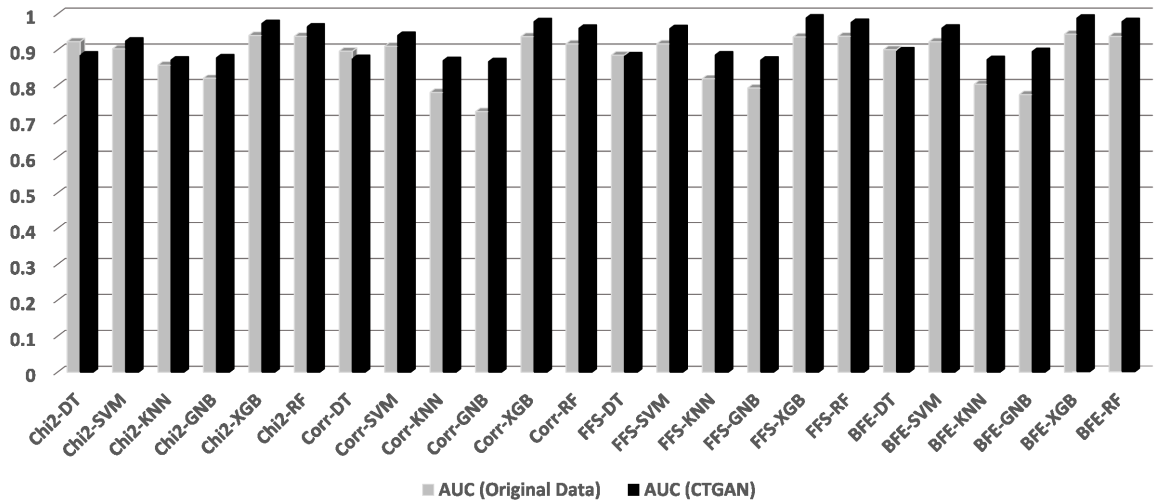

FIGURE 7. The performance comparison of all experimental settings between the models trained on original data and synthetically generated data by CTGAN, respectively.

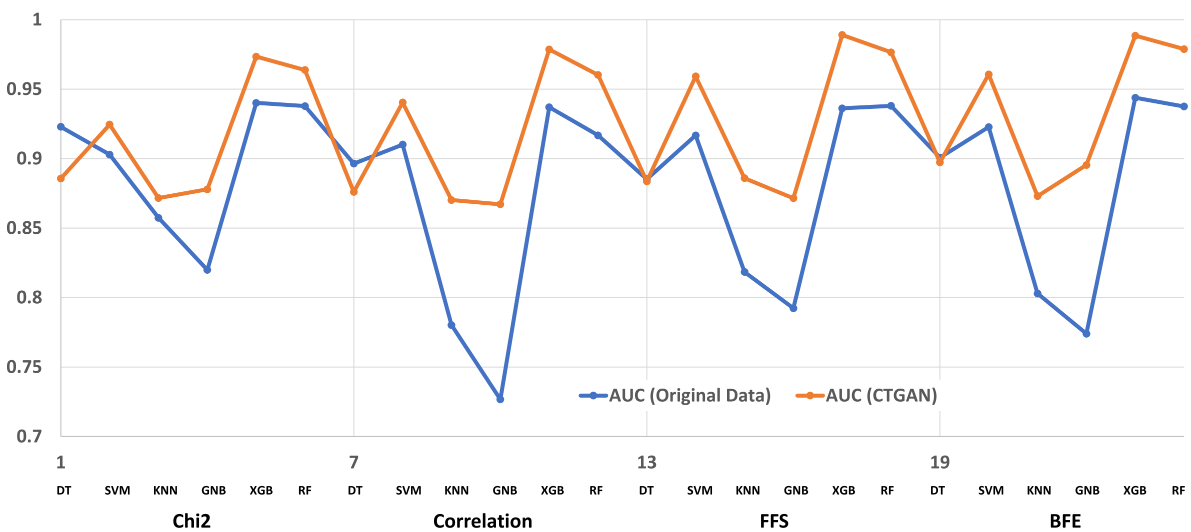

FIGURE 8. The performance comparison of all experimental settings between the models trained on original data and synthetically generated data by CTGAN, respectively.

are presented in Table 4. We also compared the results with existing work [2], where they trained classical ML models on the same dataset. The best PTC preference prediction performance among multiple ML models for every subject [2] is reported in the right-hand side of Table 4.

Our trained models with the selected features on synthetic data by CTGAN outperformed the models in related work in terms of Cohen's Kappa, except for one subject (Subject 2). In terms of the most important metric, AUC, our method, combining synthetic data generation and feature selection techniques, significantly outperformed existing works for all 14 different subjects. AUC is the metric that can better measure a classifier's performance for imbalanced data distribution.

Liu et al. [2] applied multiple classical ML models on the annotated data, considering all features. As we mentioned and analyzed earlier, some features are correlated, redundant, and irrelevant. We observed that our feature selection techniques identify the best sets of features related to modeling PTC preference. Compared to prior research, their model might suffer from irrelevant features that might mislead the classifiers.

One of the major problems in training the PTC preference model is the inadequacy of sufficient data samples. We tackled this problem by introducing CTGAN to generate high-quality synthetic data, and utilizing those data, we trained the models efficiently. Hence, the model has more data points about particular subjects and can learn better to predict thermal comfort preference more accurately. The dataset was imbalanced, CTGAN also solved that problem and mitigated the possible bias-related issues in predicting thermal comfort. Based on the evaluation metrics Kappa and AUC, both

TABLE 4. Performance in modeling PTC preference compared to with baseline.

Sub. ID	Model	[CTGAN +FS]			Liu et al. [2]		
		Kappa	Accuracy	AUC	Kappa	Accuracy	AUC
1	XGB	**0.2791**	0.5294	**0.8315**	0.17	**0.56**	0.68
	RF	0.2718	0.5490	0.7994			
2	XGB	0.3049	0.6666	**0.8202**	**0.51**	**0.74**	0.75
	RF	0.1898	0.5833	0.8065			
3	XGB	**0.7816**	**0.8785**	**0.9823**	0.50	0.77	0.86
	RF	0.7521	0.8598	0.9778			
4	XGB	0.1684	0.5393	**0.8059**	0.07	**0.86**	0.73
	RF	**0.2014**	0.5505	0.7904			
5	XGB	**0.5901**	**0.7528**	0.9311	0.30	0.69	0.79
	RF	0.5191	0.6853	**0.9327**			
6	XGB	**0.3447**	**0.5500**	**0.7973**	0.17	0.53	0.63
	RF	0.2595	0.4875	0.7970			
7	XGB	**0.5399**	**0.7230**	**0.9289**	0.37	0.69	0.77
	RF	0.5334	0.7153	0.9245			
8	XGB	**0.6666**	**0.8974**	**0.9816**	0.33	0.88	0.84
	RF	0.3960	0.7350	0.9610			
9	XGB	**0.3768**	0.6547	**0.9153**	0.21	**0.79**	0.81
	RF	0.3741	0.7261	0.9020			
10	XGB	**0.5430**	**0.7294**	**0.8947**	0.18	0.63	0.67
	RF	0.4430	0.6588	0.8555			
11	XGB	0.43557	0.6590	**0.9149**	0.37	**0.79**	0.79
	RF	**0.4590**	0.7045	0.9109			
12	XGB	**0.4197**	**0.6545**	**0.8756**	0.18	0.63	0.76
	RF	0.3243	0.6363	0.8408			
13	XGB	**0.9000**	**0.9545**	0.9477	0.41	0.75	0.83
	RF	0.7928	0.9090	**0.9643**			
14	XGB	0.1793	0.4018	**0.9416**	0.04	**0.80**	0.75
	RF	**0.1917**	**0.4392**	0.9030			

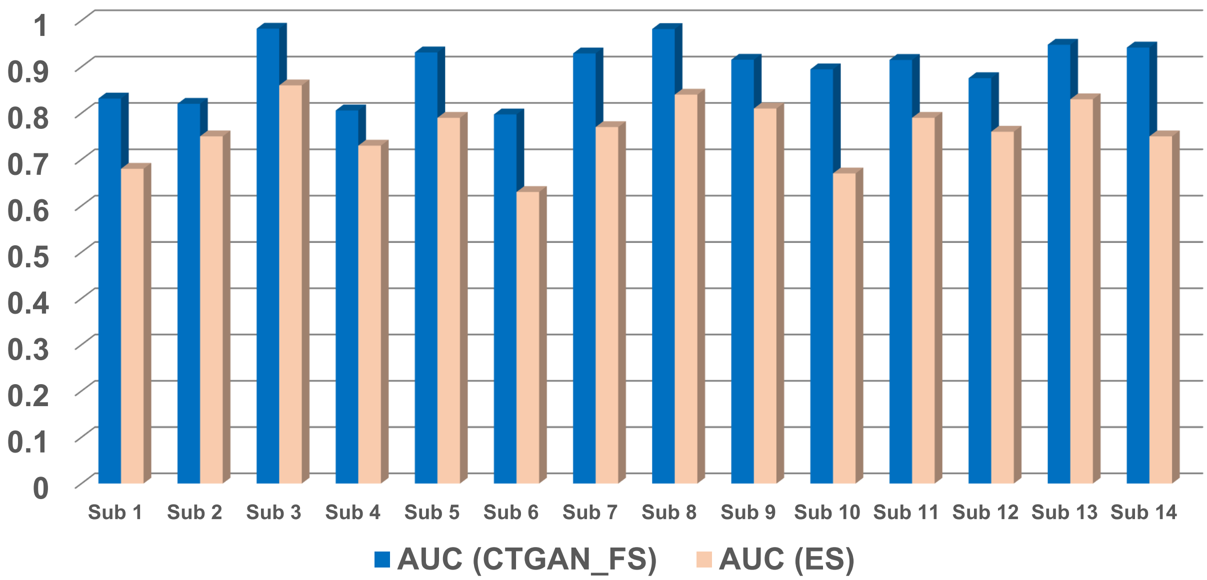

FIGURE 9. Performance comparison with existing study in terms of AUC.

recommended due to the data imbalance issue, our introduced models demonstrated significant improvements. Therefore, we can say that applying CTGAN for data generation for personal thermal comfort with feature selection is an effective approach that can be an effective combination to achieve high performance in PTC prediction.

To point out the performance differences compared to previous work, we present a comparison with related work as a bar chart in Fig. 9 in terms of performance based on AUC. AUC (CTGAN+FS) and AUC (ES) denote the performance of our proposed method combining CTGAN and feature selection (FS) and existing study (ES), respectively, in terms of AUC. We can also observe from the figure that our method outperforms prior work for all 14 subjects' thermal comfort preference prediction. To the best of our knowledge, this research is one of the few studies that conduct an extensive

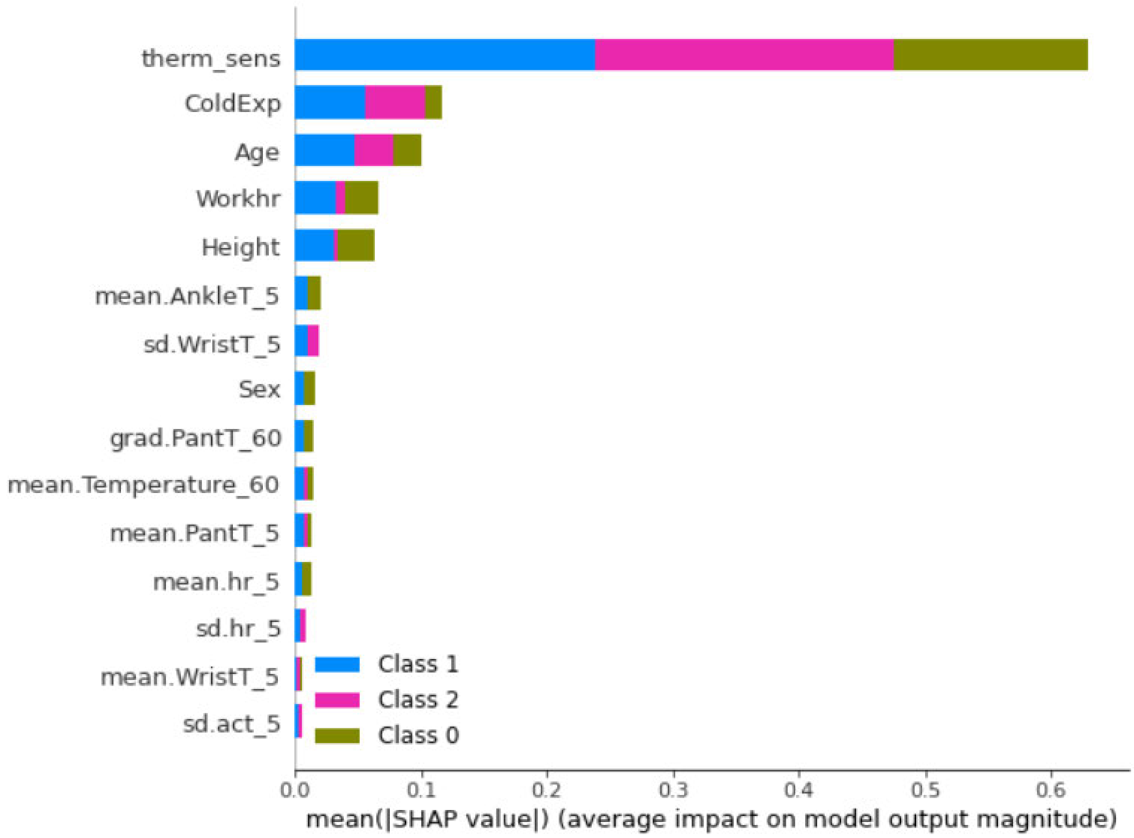

FIGURE 10. Interpretation of PTC model with feature selection using SHAP values.

study on the impact of synthetic data generation by CTGAN and effective feature selection techniques in PTC preference prediction.

H. MODEL INTERPRETABILITY

To understand the priorities in decision-making of PTC model, we applied one of the successful method, shapely additive explanation (SHAP) [51]. The feature interpretation using SHAP is presented in Fig. 10. Note that, we conducted this interpretability experiments with the selected features by applying correlation-based feature selection technique. The figure shows the 15 most important features that contributed most to make the decision of our proposed PTC model.

The top two most important features that the model takes into account are thermal sensation and cold extremity experience. It makes sense that thermal comfort preference should be dependent on these two features most. However, the next three features are age, total working hours per day and height of the subjects. Next, the skin temperature at ankle and wrist contributed the PTC model in decision making. Interestingly, subjects' sex is also an important feature that is considerable in predicting the PTC. In turn, body proximity temperature, outdoor temperature, and heart rate is also considerably important features in modeling PTC preference. This is in line with the feature selection results that we presented in Table 1.

V. CONCLUSION AND FUTURE WORK

This paper proposes a thermal comfort preference prediction method that combines a two-step process involving synthetic data generation using CTGAN and the selection of the best set of features by filtering out irrelevant and noisy features using multiple feature selection approaches. The results on a wide range of experimental settings demonstrated state-of-the-art performance and significantly outperformed existing known related work.

We observed that the ML models trained on synthetic data generated by CTGAN can predict better PTC preference than on original data samples. In addition, the introduction of a series of feature selection techniques helps filter out irrelevant features in modeling PTC preference prediction tasks. The interpretability of the model with SHAP demonstrated that the important features also overlap with the selected features. Since the PTC preference prediction task needs a substantial amount of data samples per subject to train the model efficiently, the incorporation of an effective data generation technique can save both data collection costs and associated time. The findings with feature selection indicate not to collect unnecessary data from the subject and environment and hence it might also save potential cost in sensor-based data collection cost.

In the future, we plan to introduce explainable artificial intelligence (XAI) on a large scale to provide a human-centered explanation so that occupants can understand the reason behind specific indoor parameter changes related to thermal comfort in smart homes. Since the PTC preference prediction model will be used in the smart home to control the indoor environment, an exciting extension of this work would be to conduct a user study to design a new collaborative interface for general users in human-computer interaction (HCI) perspective so that they can also be included in the loop of the smart heating system.

REFERENCES

[1] M. Shajalal, M. Bohlouli, H. P. Das, A. Boden, and G. Stevens, "Focus on what matters: Improved feature selection techniques for personal thermal comfort modelling," in *Proc. 9th ACM Int. Conf. Syst. Energy-Efficient Buildings, Cities, Transp.*, Nov. 2022, pp. 496–499.

[2] S. Liu, S. Schiavon, H. P. Das, M. Jin, and C. J. Spanos, "Personal thermal comfort models with wearable sensors," *Building Environ.*, vol. 162, Sep. 2019, Art. no. 106281.

[3] G. Brager, H. Zhang, and E. Arens, "Evolving opportunities for providing thermal comfort," *Building Res. Inf.*, vol. 43, no. 3, pp. 274–287, May 2015.

[4] K. Liu, T. Nie, W. Liu, Y. Liu, and D. Lai, "A machine learning approach to predict outdoor thermal comfort using local skin temperatures," *Sustain. Cities Soc.*, vol. 59, Aug. 2020, Art. no. 102216.

[5] H. P. Das, S. Schiavon, and C. J. Spanos, "Unsupervised personal thermal comfort prediction via adversarial domain adaptation," in *Proc. 8th ACM Int. Conf. Syst. Energy-Efficient Buildings, Cities, Transp.*, Nov. 2021, pp. 230–231.

[6] M. H. Fakir and J. K. Kim, "Prediction of individual thermal sensation from exhaled breath temperature using a smart face mask," *Building Environ.*, vol. 207, Jan. 2022, Art. no. 108507.

[7] M. Shajalal, A. Boden, and G. Stevens, "Towards user-centered explainable energy demand forecasting systems," in *Proc. 13th ACM Int. Conf. Future Energy Syst.*, Jun. 2022, pp. 446–447.

[8] R. May, X. Zhang, J. Wu, and M. Han, "Reinforcement learning control for indoor comfort: A survey," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 609, Sep. 2019, Art. no. 062011.

[9] J. Ngarambe, G. Y. Yun, and M. Santamouris, "The use of artificial intelligence (AI) methods in the prediction of thermal comfort in buildings: Energy implications of AI-based thermal comfort controls," *Energy Buildings*, vol. 211, Mar. 2020, Art. no. 109807.

[10] M. M. Abdelrahman, A. Chong, and C. Miller, "Personal thermal comfort models using digital twins: Preference prediction with BIM-extracted spatial–temporal proximity data from Build2 Vec," *Building Environ.*, vol. 207, Jan. 2022, Art. no. 108532.

[11] J. Xie, H. Li, C. Li, J. Zhang, and M. Luo, "Review on occupant-centric thermal comfort sensing, predicting, and controlling," *Energy Buildings*, vol. 226, Nov. 2020, Art. no. 110392.

[12] T. Chaudhuri, D. Zhai, Y. C. Soh, H. Li, and L. Xie, "Random forest based thermal comfort prediction from gender-specific physiological parameters using wearable sensing technology," *Energy Buildings*, vol. 166, pp. 391–406, May 2018.

[13] T. Chaudhuri, Y. C. Soh, H. Li, and L. Xie, "A feedforward neural network based indoor-climate control framework for thermal comfort and energy saving in buildings," *Appl. Energy*, vol. 248, pp. 44–53, Aug. 2019.

[14] A. Chennapragada, D. Periyakoil, H. P. Das, and C. J. Spanos, "Time series-based deep learning model for personal thermal comfort prediction," in *Proc. 13th ACM Int. Conf. Future Energy Syst.*, Jun. 2022, pp. 552–555.

[15] N. Somu, A. Sriram, A. Kowli, and K. Ramamritham, "A hybrid deep transfer learning strategy for thermal comfort prediction in buildings," *Building Environ.*, vol. 204, Oct. 2021, Art. no. 108133.

[16] N. Gao, W. Shao, M. S. Rahaman, J. Zhai, K. David, and F. D. Salim, "Transfer learning for thermal comfort prediction in multiple cities," *Building Environ.*, vol. 195, May 2021, Art. no. 107725.

[17] N. Eslamirad, S. M. Kolbadinejad, M. Mahdavinejad, and M. Mehranrad, "Thermal comfort prediction by applying supervised machine learning in green sidewalks of Tehran," *Smart Sustain. Built Environ.*, vol. 9, no. 4, pp. 361–374, Apr. 2020.

[18] R. Escandón, F. Ascione, N. Bianco, G. M. Mauro, R. Suárez, and J. J. Sendra, "Thermal comfort prediction in a building category: Artificial neural network generation from calibrated models for a social housing stock in southern Europe," *Appl. Thermal Eng.*, vol. 150, pp. 492–505, Mar. 2019.

[19] J. Guenther and O. Sawodny, "Feature selection and Gaussian process regression for personalized thermal comfort prediction," *Building Environ.*, vol. 148, pp. 448–458, Jan. 2019.

[20] T. Chaudhuri, D. Zhai, Y. C. Soh, H. Li, and L. Xie, "Thermal comfort prediction using normalized skin temperature in a uniform built environment," *Energy Buildings*, vol. 159, pp. 426–440, Jan. 2018.

[21] K. Katić, R. Li, and W. Zeiler, "Machine learning algorithms applied to a prediction of personal overall thermal comfort using skin temperatures and occupants' heating behavior," *Appl. Ergonom.*, vol. 85, May 2020, Art. no. 103078.

[22] A. Aryal, B. Becerik-Gerber, G. M. Lucas, and S. C. Roll, "Intelligent agents to improve thermal satisfaction by controlling personal comfort systems under different levels of automation," *IEEE Internet Things J.*, vol. 8, no. 8, pp. 7089–7100, Apr. 2021.

[23] P. Jayathissa, M. Quintana, M. Abdelrahman, and C. Miller, "Humans-as-a-sensor for buildings—Intensive longitudinal indoor comfort models," *Buildings*, vol. 10, no. 10, p. 174, Oct. 2020.

[24] J. Lee and Y. Ham, "Physiological sensing-driven personal thermal comfort modelling in consideration of human activity variations," *Building Res. Inf.*, vol. 49, no. 5, pp. 512–524, Jul. 2021.

[25] S. Lee, P. Karava, A. Tzempelikos, and I. Bilionis, "A smart and less intrusive feedback request algorithm towards human-centered HVAC operation," *Building Environ.*, vol. 184, Oct. 2020, Art. no. 107190.

[26] D. Li, C. C. Menassa, V. R. Kamat, and E. Byon, "HEAT—Human embodied autonomous thermostat," *Building Environ.*, vol. 178, Jul. 2020, Art. no. 106879.

[27] C. Shan, J. Hu, J. Wu, A. Zhang, G. Ding, and L. X. Xu, "Towards non-intrusive and high accuracy prediction of personal thermal comfort using a few sensitive physiological parameters," *Energy Buildings*, vol. 207, Jan. 2020, Art. no. 109594.

[28] Y. Feng, S. Liu, J. Wang, J. Yang, Y.-L. Jao, and N. Wang, "Data-driven personal thermal comfort prediction: A literature review," *Renew. Sustain. Energy Rev.*, vol. 161, Jun. 2022, Art. no. 112357.

[29] Y. Wu and B. Cao, "Recognition and prediction of individual thermal comfort requirement based on local skin temperature," *J. Building Eng.*, vol. 49, May 2022, Art. no. 104025.

[30] L. Arakawa Martins, V. Soebarto, and T. Williamson, "A systematic review of personal thermal comfort models," *Building Environ.*, vol. 207, Jan. 2022, Art. no. 108502.

[31] Z. Qavidel Fard, Z. S. Zomorodian, and S. S. Korsavi, "Application of machine learning in thermal comfort studies: A review of methods, performance and challenges," *Energy Buildings*, vol. 256, Feb. 2022, Art. no. 111771.

[32] M. Quintana, S. Schiavon, K. W. Tham, and C. Miller, "Balancing thermal comfort datasets: We GAN, but should we?" in *Proc. 7th ACM Int. Conf. Syst. Energy-Efficient Buildings, Cities, Transp.*, Nov. 2020, pp. 120–129.

[33] H. Yoshikawa, A. Uchiyama, and T. Higashino, "Data balancing for thermal comfort datasets using conditional Wasserstein GAN with a weighted loss function," in *Proc. 8th ACM Int. Conf. Syst. Energy-Efficient Buildings, Cities, Transp.*, Nov. 2021, pp. 264–267.

[34] H. P. Das and C. J. Spanos, "Synthetic personal thermal comfort data generation," in *Proc. 9th ACM Int. Conf. Syst. Energy-Efficient Buildings, Cities, Transp.*, Nov. 2022, pp. 280–281.

[35] L. Xu, M. Skoularidou, A. Cuesta-Infante, and K. Veeramachaneni, "Modeling tabular data using conditional GAN," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, pp. 1–11.

[36] L. Xu, "Synthesizing tabular data using conditional GAN," M.S. thesis, Dept. Elect. Eng. Comput. Sci., Massachusetts Inst. Technol., Cambridge, MA, USA, 2020.

[37] M. Shajalal, A. Boden, and G. Stevens, "Explainable product backorder prediction exploiting CNN: Introducing explainable models in businesses," *Electron. Markets*, vol. 32, no. 4, pp. 2107–2122, Dec. 2022.

[38] M. Shajalal, P. Hajek, and M. Z. Abedin, "Product backorder prediction using deep neural network on imbalanced data," *Int. J. Prod. Res.*, vol. 61, no. 1, pp. 302–319, Jan. 2023.

[39] D. W. Aha, D. Kibler, and M. K. Albert, "Instance-based learning algorithms," *Mach. Learn.*, vol. 6, no. 1, pp. 37–66, Jan. 1991.

[40] W. Wang, T. Hong, N. Xu, X. Xu, J. Chen, and X. Shan, "Cross-source sensing data fusion for building occupancy prediction with adaptive lasso feature filtering," *Building Environ.*, vol. 162, Sep. 2019, Art. no. 106280.

[41] M. Esrafilian-Najafabadi and F. Haghighat, "Impact of predictor variables on the performance of future occupancy prediction: Feature selection using genetic algorithms and machine learning," *Building Environ.*, vol. 219, Jul. 2022, Art. no. 109152.

[42] S. Lu, W. Wang, S. Wang, and E. Cochran Hameen, "Thermal comfort-based personalized models with non-intrusive sensing technique in office buildings," *Appl. Sci.*, vol. 9, no. 9, p. 1768, Apr. 2019.

[43] Q. Zhao, Y. Zhao, F. Wang, J. Wang, Y. Jiang, and F. Zhang, "A data-driven method to describe the personalized dynamic thermal comfort in ordinary office environment: From model to application," *Building Environ.*, vol. 72, pp. 309–318, Feb. 2014.

[44] Z. Wang, H. Zhang, Y. He, M. Luo, Z. Li, T. Hong, and B. Lin, "Revisiting individual and group differences in thermal comfort based on ASHRAE database," *Energy Buildings*, vol. 219, Jul. 2020, Art. no. 110017.

[45] J. Guenther and O. Sawodny, "Feature selection for thermal comfort modeling based on constrained LASSO regression," *IFAC-PapersOnLine*, vol. 52, no. 15, pp. 400–405, 2019.

[46] T. C. T. Cheung, S. Schiavon, E. T. Gall, M. Jin, and W. W. Nazaroff, "Longitudinal assessment of thermal and perceived air quality acceptability in relation to temperature, humidity, and CO_2 exposure in Singapore," *Building Environ.*, vol. 115, pp. 80–90, Apr. 2017.

[47] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 214–223.

[48] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of Wasserstein GANs," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 1–11.

[49] S. Jain, S. Shukla, and R. Wadhvani, "Dynamic selection of normalization techniques using data complexity measures," *Expert Syst. Appl.*, vol. 106, pp. 252–262, Sep. 2018.

[50] N. Japkowicz and M. Shah, *Evaluating Learning Algorithms: A Classification Perspective*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[51] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 1–10.

MD. SHAJALAL received the B.Sc. degree in computer science from the University of Chittagong (CU), Bangladesh, and the M.Eng. degree in computer science from Toyohashi University of Technology (TUT), Japan. He is currently pursuing the Ph.D. degree in explainable AI with the University of Siegen, Germany. He is also a Marie Skodowska-Curie Research Fellow with the Fraunhofer-Institute for Applied Information Technology (FIT), Germany. He was a Visiting Research Fellow with the University of Strathclyde, Glasgow, U.K., and completed an industrial secondment with AGENTS Inc., Berlin, Germany. He is also an Academic Faculty Member of Computer Science (currently on leave) with Hajee Mohammad Danesh Science and Technology University, Bangladesh. He was awarded prestigious Skodowska-Curie Fellowship and MEXT Scholarship for his master's and Ph.D. studies, funded by the European Union (EU) and the Ministry of Education, Culture, Sports, Science, and Technology, Japan, respectively. His research interests include human-centered explainable AI (XAI), natural language processing (NLP), information retrieval (IR), and applied machine/deep learning.

MILAD BOHLOULI received the bachelor's degree in computer science from Iran University of Science and Technology and the master's degree in artificial intelligence from the Amirkabir University of Technology. He is currently a Machine Learning Engineer with Petanux GmbH, Bonn, Germany. He was also a Researcher with the University of Siegen. His research interests include the filed of machine learning, computer vision, and explainable AI.

HARI PRASANNA DAS received the Ph.D. degree from the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley (UC Berkeley), Berkeley. His research interests include the intersection of generative AI, deep learning, computer vision, data-efficient machine learning, and smart buildings. He has been a recipient of the Best Poster Award at ACM BuildSys 2021, the C.V. & Daulat Ramamoorthy Distinguished Research Award by the Department of Electrical Engineering and Computer Sciences, UC Berkeley-based on outstanding contributions to a new research area in computer science and engineering, and the Lotfi A. Zadeh Prize by the Department of Electrical Engineering and Computer Sciences, UC Berkeley, for outstanding contributions to soft computing and its applications.

ALEXANDER BODEN received the Ph.D. degree in business informatics from the University of Siegen, Germany, in an interdisciplinary topic between software engineering and social sciences, and the Habilitation degree in business informatics from the University of Siegen, in 2019. He is currently a Professor in software engineering with Bonn-Rhein-Sieg University of Applied Sciences (HBRS), Bonn, Germany, where he is also the Director of the Institute for Consumer Informatics. He is also affiliated with the Department of User-Centered Computing, Fraunhofer Institute for Applied Information Technology (FIT), where he leads a research team in the field of human-centered computing. His research interests include human-centered AI, human–computer interaction, and business informatics.

GUNNAR STEVENS is currently a Professor in information systems and new media with the University of Siegen, Germany. He is also the Divisional Director of IT Security and Consumer Informatics. He is leading a multidisciplinary research team that is broadly concerned with user-centered computing and IT security, human–computer interaction, and applied machine/deep learning. He is also serving as the Head of the Consumer Informatics/Competence Center SME 4.0 Usability, Bonn-Rhein-Sieg University of Applied Sciences, Bonn, Germany.