AIMS Mathematics

AIMS Mathematics, 11(8): 27131–27163.
DOI: 10.3934/math.20261087
Received: 05 April 2026
Revised: 12 July 2026
Accepted: 20 July 2026
Published: 27 August 2026

https://www.aimspress.com/journal/Math

Research article

Temporal dynamics and embedded neural network in a hybrid Latent class choice model with application

Mashail M. AL Sobhi*

Department of Mathematics, Umm-Al-Qura University, Makkah 24227, Saudi Arabia

*** Correspondence:** Email: mashail1426@gmail.com.

Abstract: The paper presents a new latent class choice model (LCCM) based on an embedded neural architecture with attention mechanisms that accepts both categorical and continuous inputs. The model incorporates both discrete and continuous input convolution operations and temporal attention mechanisms, thereby introducing a latent function to enhance its utility. Based on the results of the ENTH experiment, a large-scale survey conducted at the National University of Singapore, we discuss longitudinal individual preferences for thermal comfort. With 17 participants and 1400 responses, the dataset covers both indoor and outdoor environments and is monitored using smartwatches and specialized smartphone applications. Using aspects such as body presence, clothing, and environmental considerations, our LCCM classifies participants into four latent classes: outdoor comfort enthusiasts, indoor comfort seekers, dynamic comfort adapters, and neutral comfort responders. The model's parameters demonstrate the complex associations between the environment and clothing selection. The findings highlight the model's strengths in predictability and efficiency relative to other benchmark models, providing useful insights into personalized thermal comfort dynamics. The proposed model is much better than current models, offering greater computational efficiency and predictive power. The scores of temporal embedding and attention also indicate that the model can be trained to capture subtle temporal relationships in sequential data. The study contributes to personalized modeling of comfort by providing a powerful instrument for understanding and predicting personal thermal preferences.

Keywords: embedded neural network; LCCM; temporal dynamics; attention mechanism; EM algorithm; environmental comfort
Mathematics Subject Classification: 62H30, 62M45

1. Introduction

The latent class choice model (LCCM) is one of the most well-known and widely used discrete distribution models. It operates as a random-utility model that extends the multinomial logit model. The key innovation is its latent-class formulation, which enables the probabilistic assignment of individuals to K homogeneous classes with distinct behavioral patterns. This model becomes particularly useful when the modeler posits that unobserved heterogeneity can be represented by discrete latent courses, such as different population segments with varying tastes, distinct decision-making protocols adopted by individuals, and variations in the choice sets considered from one individual to another [1]. Traditionally, these latent classes are defined based on the socio-economic characteristics of decision-makers [2–4].

In recent years, the domain of choice modeling has undergone significant transformations, driven by the imperative to develop more sophisticated methodologies that capture individuals' nuanced decision-making processes. The LCCM has become a leading paradigm in this landscape, providing a powerful framework for explaining unobserved heterogeneity among decision-makers. The LCCM enables population segmentation based on common preferences and decision-making rules, thereby increasing the model's ability to represent the natural diversity of human decision-making. With the field's evolution, incorporating temporal dynamics and embedded neural network architectures may be an interesting step toward enhancing LCCM capabilities. Temporal dynamics, which capture the development of preferences and patterns of decision-making, allow the incorporation of time into the model by recognizing the dynamic nature of decision-making processes. At the same time, the integration of embedded neural networks enables a potent mechanism of learning complex relationships in the data, which can be further used to improve the adaptability and generalization ability of the model. This study proposes the hybrid LCCM, a new model that introduces a smooth integration of time-varying components and neural networks into the classic LCCM.

This research aims to propose a more adaptive and nuanced description of the decision-making process, taking into account the time dependence and the complex interdependence of decision factors. Building on classic studies of choice modeling [5–7] and on advances in time modeling [8] and neural network implementations [9], the current study aims to address the dynamic nature of choice modeling techniques. By leveraging the strengths of LCCMs, temporal dynamics, and embedded neural networks, the study will expand our understanding of decision-making processes in dynamic systems. The study will investigate individuals' comfort preferences under different environmental conditions by carefully examining data from the ENTH experiment. This new neural architecture model incorporates embedded neural networks and attention and temporal mechanisms, highlighting its superior performance in capturing complex decision-making patterns. Using the EM algorithm for model estimation, with particular attention to the algorithm’s initialization, the study identifies four latent classes of participants with distinct comfort preferences.

It is important to note that the proposed model consistently outperforms existing benchmarks, achieving higher log-likelihood, predictive log-likelihood, and goodness-of-fit metrics on both training and test data. In addition to predictive capabilities, the model offers augmented computational efficiency, with much lower training time, memory usage, and inference speed than current models. The exploration also examines how the model extracts temporal dependencies in sequential data using attention scores and temporal embeddings. In general, this study offers a strong latent class choice framework with improved predictive performance, efficiency, and responsiveness to temporal complexities, which is an important contribution to understanding and predicting individual comfort preferences across varying environmental settings.

Although latent class choice models and neural network prediction models are gaining popularity, several shortcomings have already been identified. Traditional LCCMs rely on linear utility specifications and are not well-suited to incorporate complex nonlinear interactions among environmental, physiological, and behavioral variables. Likewise, previous attention-based architectures are mainly oriented toward prediction accuracy and lose interpretability and behavioral anchoring in the discrete choice analysis context. To tackle these challenges, this study introduces a novel embedded neural temporal hybrid latent class choice model (ENTH-LCCM) that combines latent class segmentation, embedded representation of categorical data, a continuous-input attention mechanism, a temporal attention layer, and estimation via an EM algorithm. The proposed architecture enables simultaneous modeling of nonlinear behavioral interactions, temporal evolution of preferences, and latent heterogeneity while keeping the model interpretable. The resulting framework is superior to existing benchmark models in predictive performance, computational efficiency, and behavioral insight.

Deep learning architectures have evolved rapidly with the introduction of transformer-based networks, adaptive attention mechanisms, hybrid CNN architectures, and feature fusion strategies. Adaptive feature learning, multi-scale attention, and hybrid neural architectures significantly enhance representation learning and prediction performance, as shown by recent studies in top computer vision conferences and journals, including CVPR, ICCV, ECCV, and high-impact representation learning and prediction journals such as pattern recognition and applied soft computing. These developments suggest that attention-based neural representations could be used to effectively model the nonlinear relationships and behavioral heterogeneity of decision-making, even across statistical decision-making models. Other recent works have also shown that adaptive feature fusion, hybrid attention mechanisms, and multi-scale representation learning can greatly enhance feature extraction and model robustness in different applications. While these strategies are mostly applied to computer vision tasks, they offer fundamental ideas of the adaptive learning of representations and the selection of features based on attention, which are useful inspirations for behavioral modeling.

This study proposes a novel ENTH-LCCM that integrates latent class segmentation, embedded neural representations, continuous-input attention mechanisms, temporal attention layers, and estimation based on the EM algorithm. This paper makes the following key contributions:

1) A novel neural-enhanced latent class choice model is proposed that combines embedded neural architectures and latent class segmentation for model learning.
2) A temporal attention mechanism allows for modeling temporal preferences and temporal behavioral dependencies.
3) Embedding layers and attention-based learning are used to develop a unified framework to simultaneously process categorical and continuous variables.
4) An EM-based estimation approach is combined with neural optimization to enhance the stability and prediction performance of the model.
5) Comprehensive experiments show that the model performs better than the traditional discrete choice and LCCMs in terms of prediction performance, speed, and interpretability.

2. Related studies

2.1. Latent class choice model

The LCCMs are a specific type of mixed logit model. Mixed logit models, also known as logit models with random coefficients, are used to model unobserved heterogeneity by assuming that the coefficients of the utility function vary across decision-makers according to specified probability distributions. Consequently, different decision-makers may exhibit marginally different responses to the same attribute of an alternative.

According to McFadden [10], Walker and Ben-Akiva [11] proposed a generalized random utility model that is practical, extends the basic random utility model (RUM), and includes latent-variable and latent-class extensions. This extension relaxes assumptions and increases the model's capabilities. Latent classes, in this case, denote unobserved groups within the population, with each individual having a corresponding probability of belonging to a given class. In the case of LCCMs, the hypothesis is that there are discrete segments of decision-makers that are not clear in the data. As a result, the population may be probabilistically divided into groups with different preferences, decision protocols, or decision behaviors whose specific utility equations are described by the classes. Additional practical outcomes arise from this segmentation, including the development of more efficient design policies [12]. Although mixed logit models are very flexible and can approximate any random utility model [13], they are not easy to interpret or apply in practice because they do not identify the causes of differences in decision-makers' tastes. In contrast, the literature indicates that latent class models are convenient, flexible, and intuitive for explaining taste heterogeneity in discrete choice models [14]. The defining feature of LCCMs is that the class-membership model is not predefined. Conventionally, researchers use simple logit formulations for this purpose, since they provide a closed-form expression for the probability of belonging to each class. In addition, these formulations are usually based on socio-demographic variables for segmentation [11,12,15].

2.2. Navigating embeddings

The explosion of the popularity of embedding representations can be traced to word2vec [16], an innovative deep learning method that produces dense word vectors, often known as word embeddings. These embeddings can overcome high dimensionality and sparsity in language data. They present brief, intuitive representations that skillfully capture interrelationships among the encoded units based on their distributional properties. The remarkable usefulness and effectiveness of word embeddings have been demonstrated across a wide range of downstream natural language processing (NLP) tasks, such as text classification, question answering, and machine translation [1,14]. Most studies over the past few years have emphasized that entrenchment techniques using artificial neural networks (ANNs) to encode discrete units of information extend far beyond NLP. For instance, Node2Vec [17] and DeepWalk [18] have been crucial for graph representation, whereas TransE [19] and RESCAL [20] are useful for knowledge graph representation. Another application worth mentioning is DeViSE [21], which has played a significant role in mapping images to a high-level semantic embedding space.

Multiple studies [22–24] have focused on producing Bayesian hierarchical discrete choice model and geospatial-embedding representations, analogous to word2vec but with additional consideration of mobility patterns and spatial context. The resulting representations provide new knowledge about the semantics of space and place functionalities [25–27] and are applicable across a wide range of city planning and policy formulation activities. Such applications include point-of-interest (POI) recommendations [28] to forecast users' visits to specific POIs in future periods [29]. In a research work particularly relevant to transport and demand forecasting, de Brébisson et al. used various ANN architectures to predict taxi destinations from initial paths and associated metadata [30]. They found that embeddings, together with the models and discrete metadata (client ID, taxi ID, date, and time), were significant predictors of taxi destination. Similarly, Guo and Berkhahn [31] designed categorical embeddings using ANNs and applied the resulting encoding to prediction. They showed that using these embeddings as input features significantly improved the performance of various machine learning (ML) algorithms for daily sales forecasting. They noted that encoding helps neural networks generalize more effectively, especially with sparse data and unknown statistics, which other techniques tend to overfit.

More recently, [26] explored using discrete variables in NN-based traffic prediction models that leverage large amounts of categorical data (e.g., time, site ID, weather). In their exploration, they compared the embedding representations of these variables with one-hot encoding and found that embedding vectors better capture the internal relationships among variables and are therefore more effective at predicting traffic flow. They also demonstrated that the visual inspection of the trained embedding vectors can reveal inherent features and associations between categorical variables. Such visual analysis can be a form of supervised clustering. Theoretically, the embedding component used in the proposed ENTH-LCCM framework was motivated by Feng et al. [27], who proposed a framework for learning geographical latent representations using embedding techniques and showed that representation learning is an effective method for capturing complex latent patterns.

2.3. Attention mechanism

In deep learning, an attention system is heavily inspired by human cognitive attention and by observed attention signals in biology, where human attention is a limited and precious resource. Take the case of an individual concentrating their attention on a given object: naturally, other objects are pushed to the background. This framework, as proposed by [28], highlights the application of an attention mechanism based on nonvolitional and volitional cues. The nonvolitional cue depends on object saliency within the environment, whilst the volitional cue is a cognitive control variable-selection criterion for deliberate attention [29]. Based on these nonvolitional and volitional attention signals, researchers have proposed an attention mechanism that incorporates queries, keys, and values for attention pooling in machine learning models. Structured attention pooling cooperates with the interaction between the provided queries (volitional cues) and keys (nonvolitional cues), thereby enabling a biased selection of values that reflect the input data [30]. This attention mechanism can amplify specific sections of the input information while ignoring others. The rationale behind this lies in the neural network's tendency to assign higher weights to small but important components of the inputs. The attention mechanism proposed in [15,31,32] is a useful model for adaptively weighting features in neural networks, but its goal differs from that of this paper. The proposed ENTH-LCCM embeds learning, continuous attention, temporal attention, and latent class choice modeling, all in a single behavioral modeling framework. The proposed model thus includes nonlinear behavioral relationships, latent preference heterogeneity, and temporal dependencies in a single model, which has not been done in previous studies.

While existing latent class choice models can account for unobserved heterogeneity, they have drawbacks. First, conventional LCCMs assume a linear utility specification, which cannot adequately capture nonlinear behavioral interactions. Second, current choice models based on neural networks tend to focus primarily on predicting choices rather than providing interpretability or behavioral foundations. Third, most existing methods do not account for changes in temporal preferences and sequential dependencies during the decision-making process. Last but not least, categorical variables are often one-hot encoded, which can lead to sparsity and less efficient learning.

3. Methodology

The proposed choice model is the LCCM, which uses an integrated neural network architecture with an attention mechanism. This model integrates two types of inputs, categorical and continuous, with specialized treatment: attention weights for continuous variables and embedding weights for categorical variables. Temporal attention mechanisms, added as I + 1 to I + S terms, increase the model's adaptability over time. Latent functions and the convolution lead to the formulation of a utility function. The last model combines these three outputs into a utility layer enhanced with latent words, and the subsequent Softmax produces the final output. Detailed information regarding this innovative architecture is provided below.

3.1. Conventional latent class choice model (LCCM)

LCCMs comprise two sub-models: class-membership and class-specific choice models. The class-membership model estimates the likelihood that a person m belongs to a given class. Meanwhile, the class-specific choice model specifies the probability of choosing each alternative, whereby individual m is a member of class c. In general, the choice model for class membership is specified as a logit. The utility of individual i in class c is articulated as follows:

$$U_{mc}^{(t)} = A_c^{(t)} + Z_m^{(t)}\beta + \varepsilon_{mc}. \tag{1}$$

The equation encompasses several elements: $A_c^{(t)}$ denotes the alternative-specific value for each class c at time t, $Z_m^{(t)}$ represents a vector encapsulating an individual m characteristics, β is a vector of parameters awaiting estimation, and ε_{mc} stands for an error term, assumed to independently follow the extreme value type-I distribution across individuals and classes. Consequently, the probability that an individual m belongs to class c considering their characteristics adheres to the logit formula:

$$P\left(r_{mc}^{(t)} | Z_m^{(t)}, \beta_c\right) = \frac{e^{Z_m^{(t)}\beta_c}}{\sum_{c'=1}^{C} e^{Z_m^{(t)}\beta_{c'}}}, \tag{2}$$

where

$$r_{mc}^{(t)} = \begin{cases} 1 \text{ } if \text{ } individual \text{ } belongs \text{ } to \text{ } class \text{ } c \\ 0 \text{ } otherwise \end{cases}.$$

Regarding the class-specific choice model, it computes the probability that an individual m selects a particular alternative, provided that they belong to a specific class c. The utility function for individual m choosing alternative j, given their membership in class c, is expressed as follows:

$$U_{mc}^{(t)} = V_{mj|c}^{(t)} + \varepsilon_{mj|c}, \tag{3}$$

$$V_{mj|c}^{(t)} = I_{mj}^{(t)} \vartheta_c + \varepsilon_{mj|c}. \tag{4}$$

Where $V_{mj|c}^{(t)}$ represents the observed portion of the utility at time t, and $\varepsilon_{mj|c}$ stands for an independent and usually distributed extreme value type-I error term across individuals, alternatives, and classes. $I_{mj}^{(t)}$ constitutes the vector of observed attributes for alternative j at time t, usually comprising an alternative-specific constant, while ϑ_c is the vector of unknown parameters to be estimated. Consequently, the probability that individual m selects alternative j, given class c at time t, is expressed as:

$$P\left(q_{mc}^{(t)} | I_{mj}^{(t)}, r_{mc}^{(t)}, \vartheta_c\right) = \frac{e^{V_{mj|c}^{(t)}}}{\sum_{k'=1}^{K} e^{V_{mk'|c}^{(t)}}}, \tag{5}$$

$$q_{mj}^{(T)} = \begin{cases} 1 \text{ } if \text{ } individual \text{ } m \text{ } choose \text{ } alternative \text{ } j \\ 0 \text{ } otherwise \end{cases}.$$

The probability of considering $q_{mk}^{(t)}$ at time t, conditional on class, is expressed as:

$$P\left(q_{m}^{(t)} | I_{m}^{(t)}, r_{mc}^{(t)}, \vartheta_c\right) = \prod_{t=1}^{T} \prod_{k=1}^{K} \left(q_{mk}^{(t)} | I_{mk}^{(t)}, r_{mc}^{(t)}, \vartheta_c\right). \tag{6}$$

Here, $q_{m}^{(t)}$ represents a vector containing all choices K made by an individual m at time t, and $I_{m}^{(t)}$ denotes a vector K of $I_{mk}^{(t)}$ at time t. The overall probability of observing $q_{m}^{(t)}$, regardless of conditions, can be computed by combining the earlier conditional choice probability with the likelihood of belonging to each class c. This computation is as follows:

$$P\left(q^{(t)}\right) = \prod_{m=1}^{M} \sum_{c=1}^{C} \sum_{t=1}^{T} P\left(r_{mc}^{(t)} | Z_{m}^{(t)}, \beta_c\right) P\left(q_{m}^{(t)} | I_{m}^{(t)}, r_{mc}^{(t)}, \vartheta_c\right). \tag{7}$$

In conclusion, the likelihood across all individuals M, assuming their independence, can be expressed as:

$$P\left(q^{(t)}\right) = \prod_{m=1}^{M} \sum_{c=1}^{C} \sum_{t=1}^{T} P\left(r_{mc}^{(t)} | Z_{m}^{(t)}, \beta_c\right) P\left(q_{m}^{(t)} | I_{m}^{(t)}, r_{mc}^{(t)}, \vartheta_c\right). \tag{8}$$

3.2. Formulation of an individual's characteristic vector

Let $X^{(T)} = \left\{X_1^{(t)}, X_2^{(t)}, \dots \dots, X_J^{(t)}\right\}$, as a set of J continuous independent variables, and $Y^{(T)} = \left\{Y_1^{(t)}, Y_2^{(t)}, \dots \dots, Y_K^{(t)}\right\}$, as a set of K categorical independent variables with temporal attention at time T. These variables collectively represent the observed attributes of the choice alternatives and the individuals' socio-demographic characteristics, emphasizing a temporal perspective at time T. These variables together form Z, encompassing an individual's observed characteristics. Z is represented as a function, $Z^{(T)} = f\left(X_J^{(T)}, Y_K^{(T)}\right)$, synthesizing both continuous and categorical variables to delineate socio-characteristics.

Figure 1 [32] illustrates the first scenario of a process where the continuous inputs $X^{(T)}$ receive “attention weights”. These weighted inputs are subsequently linked to the initial beta layer, coupled with a set of adaptable weights denoted as B. Conversely, in the second scenario depicted in Figure 2 [32], the one-hot encoded inputs are transformed by projecting onto the embedding weight layer. This mapping assigns each input to a unique vector in a d-dimensional space, denoted by E = J, matching the dimensionality of the continuous variables. These transformed vectors are then connected to the second beta layer. Following this sequence, novel alternative-specific representations are formulated, integrating a temporal attention mechanism into the model architecture.

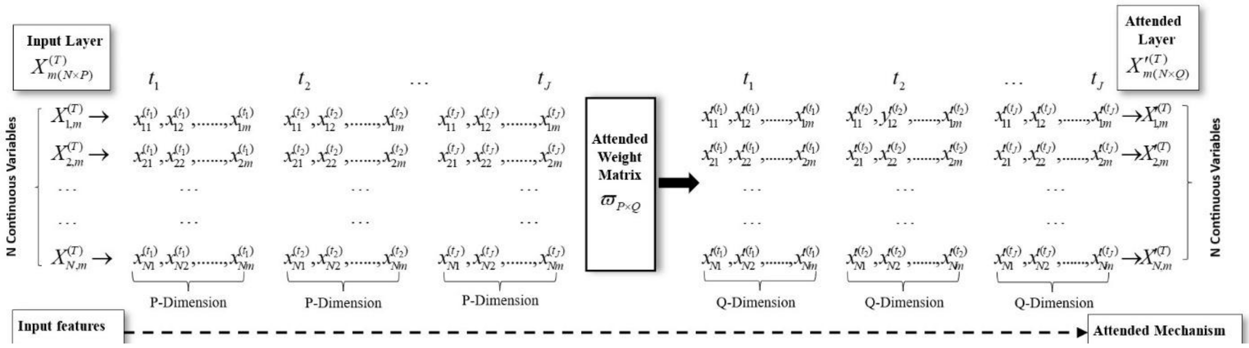

Figure 1. Architecture and representation of the attended mechanism.

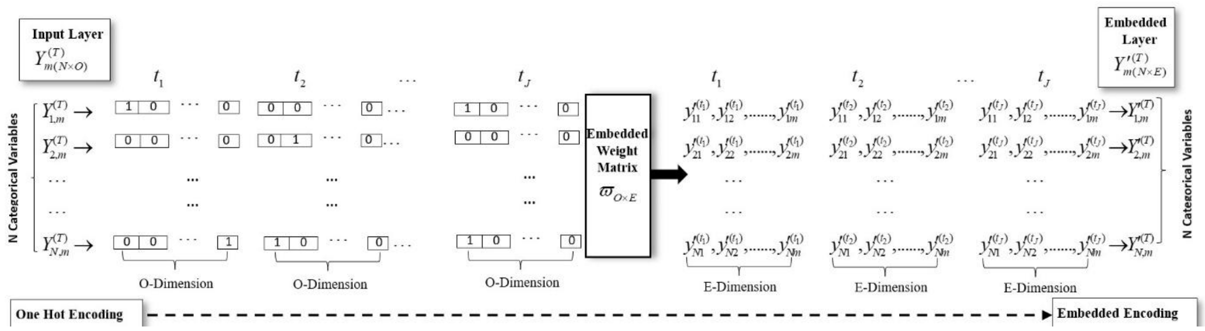

Figure 2. Architecture and representation of embedding encoding.

3.3. Formulation of latent variables and latent classes while incorporating temporal sequence

The expression for the utility within the class-membership model is formulated as follows:

$$U_{mc}^{(t)} = V_{mc}^{(t)} + \varepsilon_{mc}. \tag{9}$$

The representative utility of individual m associated with class c at time t is denoted as $V_{mc}^{(t)}$, and by the conventional LCCM, ε_{mc} represents the error term, where c denotes a learnable weighting parameter estimated during model training. After optimization, c remains fixed during inference and is therefore treated as a constant rather than a time-dependent function. This error term is presumed to be independently and identically distributed as extreme value type I across individuals and classes, as outlined in [1]. In this scenario, we formulate $V_{mc}^{(t)}$ as:

$$V_{mc}^{(t)} = A_c^{(t)} + Z_{cm}^{(t)}(\bullet) + \xi_c L_m^{(t)} + \lambda_c S_m^{(t)} + Y_{I+S} \times H, \tag{10}$$

where $A_c^{(t)}$ represents the alternative-specific value for class c, and $Z_{cm}^{(t)}(\bullet)$ constitutes the function containing the socio-characteristics of individual m (details provided in (3.2)). Moreover, $L_m^{(t)}$ denotes the vector of latent variables for individual m, and ξ_c represents the corresponding vector of unknown parameters specific to the class c, both needing estimation. Lastly, $S_m^{(t)}$ represents an individual-specific constant, along with its corresponding coefficient λ_c for each class c.

$$Y_{I+S} \times H.$$

Based on the error term distribution ε_{mc}, the probability can be defined as:

$$P\left(r_{mc}^{(t)} | I_m^{(t)}, \xi_c, L_m^{(t)}, \lambda_c\right) = \frac{e^{V_{mc}^{(t)}}}{\sum_{c'=1}^{C} e^{V_{mc'}^{(t)}}}, \tag{11}$$

$$P(q,r,I,Z) = \prod_{m=1}^{M} \prod_{c=1}^{C} \prod_{t=1}^{T} \left[\frac{e^{V_{mc}^{(t)}}}{\sum_{c'=1}^{C} e^{V_{mc'}^{(t)}}}\right]^{r_{mc}} \times \prod_{m=1}^{M} \prod_{c=1}^{C} \prod_{k=1}^{K} \prod_{t=1}^{T} \left[\frac{e^{V_{mk|c}^{(t)}}}{\sum_{k'=1}^{K} e^{V_{mk|c'}^{(t)}}}\right]^{q_{mk} r_{mc}}. \tag{12}$$

The likelihood logarithm divides the function into two parts:

$$LL = \sum_{m=1}^{M} \sum_{c=1}^{C} \sum_{t=1}^{T} r_{mc} \log\left[\frac{e^{V_{mc}^{(t)}}}{\sum_{c'=1}^{C} e^{V_{mc'}^{(t)}}}\right] + \sum_{m=1}^{M} \sum_{c=1}^{C} \sum_{k=1}^{K} \sum_{t=1}^{T} q_{mk} r_{mc} \log\left[\frac{e^{V_{mk|c}^{(t)}}}{\sum_{k'=1}^{K} e^{V_{mk|c'}^{(t)}}}\right], \tag{13}$$

here, I denotes the number of original input features, while S represents the number of additional temporal embedding dimensions generated by the temporal attention module. Consequently, the indices from $I+1$ to $I+S$ correspond to the newly generated temporal embedding features appended to the original input representation. Solving this equation involves setting its derivatives to zero concerning each unknown parameter of the choice model, assuming γ_{mc} is known. To obtain the initial values of γ_{mc}, we calculate the expectation of γ_{mc} (E-step) using Baye's theorem:

$$E[\gamma_{mc}] = \varsigma_{rmc} = \frac{\frac{e^{V_{mc}^{(t)}}}{\sum_{c'=1}^{C} e^{V_{mc'}^{(t)}}} \prod_{k=1}^{K} \prod_{t=1}^{T} \left[\frac{e^{V_{mk|c}^{(t)}}}{\sum_{k'=1}^{K} e^{V_{mk|c'}^{(t)}}} \right]^{q_{mk}}}{\sum_{c'=1}^{C} \left[\frac{e^{V_{mc}^{(t)}}}{\sum_{c'=1}^{C} e^{V_{mc'}^{(t)}}} \prod_{k=1}^{K} \prod_{t=1}^{T} \left[\frac{e^{V_{mk|c}^{(t)}}}{\sum_{k'=1}^{K} e^{V_{mk|c'}^{(t)}}} \right]^{q_{mk}} \right]}. \tag{14}$$

By merging Eqs (13) and (14), we can express the anticipated value of the log-likelihood as:

$$E[LL] = \sum_{m=1}^{M} \sum_{c=1}^{C} \sum_{t=1}^{T} \varsigma_{r_{mc}} \log \left[\frac{e^{V_{mc}^{(t)}}}{\sum_{c'=1}^{C} e^{V_{mc'}^{(t)}}} \right] + \sum_{m=1}^{M} \sum_{c=1}^{C} \sum_{k=1}^{K} \sum_{t=1}^{T} q_{mk} \varsigma_{r_{mc}} \log \left[\frac{e^{V_{mk|c}^{(t)}}}{\sum_{c'=1}^{C} e^{V_{mk|c'}^{(t)}}} \right]. \tag{15}$$

By equating the derivatives of the expected log-likelihood to zero, concerning the unknown parameters α_c of the choice model, we can estimate them using the following equation:

$$\alpha_c = \arg\max_{\alpha_c} \sum_{m=1}^{M} \sum_{k=1}^{K} \sum_{t=1}^{T} q_{mk} \varsigma_{r_{mc}} \log \left[\frac{e^{V_{mk|c}^{(t)}}}{\sum_{k'=1}^{K} e^{V_{mk|c'}^{(t)}}} \right]. \tag{16}$$

When the BFGS gradient-based numerical optimization method fails to provide a closed-form solution, it is used to estimate the model parameters. The unknown class-membership parameters are optimized using backpropagation with cross-entropy loss, as they are represented in a neural network. As explained in [21], minimizing the cross-entropy loss is equivalent to maximizing the log-likelihood function. This method allows for computing the Hessian matrix, which provides useful post-estimation statistics such as standard deviations and confidence intervals for the model parameters. When convergence is achieved, the probability of each set of choices can be calculated for all individuals.

$$\begin{aligned} P(q) = {} & \prod_{m=1}^{M} \prod_{c=1}^{C} \prod_{t=1}^{T} P\left(r_{mc} | Z_{cm}^{(t)}(\bullet), \varsigma_{r_{mc}}, L_m^{(t)} \right) \\ & \times \prod_{m=1}^{M} \prod_{c=1}^{C} \prod_{k=1}^{K} \prod_{t=1}^{T} P\left(q_{mc} | I_{cm}^{(t)}, r_{mc}, \alpha_c \right)^{q_{mc}}. \end{aligned} \tag{17}$$

3.4. The ordinal measurement model

We focus on a situation in which indicators take the form of questions that elicit ordered responses, similar to Likert scales [15]. Each point on the scale represents a degree of agreement with a specific statement. Consequently, we characterize the utility of individual *m* for indicator *o* as a measure of the degree of agreement with the corresponding statement, expressed as follows:

$$U_{om} = V_{om} + \varepsilon_{om} = l_m \alpha_o + \varpi_m c_o + \varepsilon_{om}. \tag{18}$$

Here, V_{om} represents the representative utility of individual *m* for indicator *o*, and ε_{om} is the error term assumed to follow an independently and identically distributed extreme value type-I distribution across individuals and indicators. l_m is a vector of length R containing the latent variables of individual *m*, and α_o is a vector of size R with parameters to be estimated. ϖ_m is an individual-specific parameter, estimated concurrently with the latent class model, and c_o is its corresponding coefficient for each indicator *o*. Consequently, the probability that individual *m* responds with a specific level of agreement *k* to indicator *o* is formulated as:

$$P(N_{okm} = 1 | l_m, \varpi_m, \alpha_o, c) = P(h_{k-1}^o < U_{om} < h_k^o). \tag{19}$$

Here,

$$N_{okm} = \begin{cases} 1\ if\ individual\ 'm'\ responds\ with\ a\ level\ of\ agreement 'k' to\ indicator\ 'o' \\ 0\ otherwise \end{cases}.$$

The thresholds h_k^o are strictly increasing, class-specific values that establish an ordinal relationship between the utility U_{om} and the level of agreement to indicator *o*. Ultimately, the probability of individual *m* providing an answer *k* to indicator *o* can be calculated using an ordinal Softmax:

$$P(N_{okm} = 1 | l_m, \varpi_m, \alpha_o, c) = P(h_{k-1}^o < Y_{om} + \varepsilon_{om} < h_k^o), \tag{20}$$

$$P(N_{okm} = 1 | l_m, \varpi_m, \alpha_o, c) = P(\varepsilon_{om} < h_k^o - Y_{om}) - P(\varepsilon_{om} < h_{k-1}^o - Y_{om}), \tag{21}$$

$$P(N_{okm} = 1 | l_m, \varpi_m, \alpha_o, c) = \frac{e^{h_k^o - Y_{om}}}{1 + e^{h_k^o - Y_{om}}} - \frac{e^{h_{k-1}^o - Y_{om}}}{1 + e^{h_{k-1}^o - Y_{om}}}. \tag{22}$$

In this formulation, one threshold per indicator is fixed at zero, since the crucial factor is the difference between the thresholds. The ultimate model structure is illustrated in Figure 3. Maximizing the log-likelihood with respect to the unknown parameters does not yield a closed-form solution. Consequently, the estimation process becomes intricate and challenging, and may encounter empirical singularity issues at certain iterations [22]. To overcome these difficulties, we use the expectation-maximization (EM) algorithm, a powerful method for maximizing the likelihood of latent variable models, as proposed in [5].

4. Application

4.1. Data and model configuration

The study is based on the results of the ENTH experiment survey, which was published as open data under the title “Longitudinal personal thermal comfort preference data in the wild” in the dataset submitted to the DATA 2021 workshop, BuildSys 2021. The study included 17 participants across 17 settings and received approximately 1400 responses to assess thermal preferences in indoor and outdoor environments. The open smartwatch application Cozie facilitated the smooth organization of comfort surveys that were both temporally and spatially correlated with environmental measurements. An indoor-tracking smartphone application was designed specifically to accurately track participants' locations. Background data, such as physical and personality characteristics, were collected during an initial onboarding survey. The dataset, described in Table 1, includes a variety of features, including body presence, clothing, and environmental factors. Notable continuous variables include heart rate, response speed, and environmental response, with comfort as the target variable. NUS data incorporates Fitbit technology. The study was conducted at the National University of Singapore, where the tropical climate provided the environmental context for data collection. Socio-demographic variables and personality surveys were recorded numerically. The participants were given questionnaires at the beginning of the experiment, which included the highly sensitive person scale (HSPS), the satisfaction with life scale (SWLS), and a short version of the Big Five Personality Scale. These answers were recorded on a 7-point Likert scale and were used to gain meaningful information about the participants and their preferences, as presented in Tables 2 and 3.

The proposed ENTH-LCCM first maps each categorical variable to a dense embedding vector and then normalizes continuous variables before introducing them into the attention module. The embedding vectors are joined to form a single behavioral representation, while the continuous features are kept separate. The continuous attention mechanism weights numerical variables adaptively, while the temporal attention mechanism models sequential behavioral dependencies between multiple observations. The resulting latent representation is sent into the latent class choice model, in which class probabilities and class-specific utilities are simultaneously estimated using the expectation-maximization algorithm. The complete workflow of ENTH-LCCM consists of six stages: (1) preprocessing of categorical and continuous variables, (2) embedding generation, (3) continuous attention learning, (4) temporal attention modeling, (5) latent class probability estimation using the EM algorithm, and (6) utility estimation and final choice prediction through the Softmax function.

Table 1. List of variables and their descriptions used in the ENTH dataset.

Explanatory variables		
Discrete/categorical	Description	Segments
Body presence^{cm}	Whether the user was wearing the smartwatch or not	“Yes/No”
Comfort^{cm}	Overall comfort	Comfy, Not Comfy, Neutral
Change^{cm}	Change location	Yes/No
Met^{cm}	Metabolic rate	Resting, sitting, standing, exercising
Indoor/outdoor^{cm}	Location	“Outdoor, Indoor”
Indoor floor^{cs}	The indoor Bluetooth beacons supply floor information.	[2,3), (3,4], (4,5], (5,6]
Thermal^{cm}	Thermal preferences	Warmer, No change, cooler
Continuous		Ranges
Heart rate, resting HR^{cm}	Heart rate and resting heart rate in bpm.	(47, 143), (53, 86)
Nb, skin temp^{cm}	The temperature close to the body on the wrist, and the skin temperature at the wrist level, both measured in degrees Celsius.	(22.12, 34.12), (23, 36.06)

Continued on next page

Continuous		Ranges
Indoor log. lati. cs	Latitude data is sourced from indoor Bluetooth beacons, along with longitude data from the same beacons.	(103.8), (1.297)
Voc, CO_2, pm2.5 and indoor noise cs	Indoor measurements include CO_2, volatile organic compounds (VOCs), particulate matter (PM2.5), and ambient noise levels from fixed sensors in ppm, μg/m^3, and dB, respectively.	(20, 364.1), (406.1, 1640.5), (0.7, 9.3), (48.6, 66.10)
[0.3 um, 0.5 um, 1.0 um, 10.0 um, 2.5 um, 5.0 um, humidity, pm1.0, pm10.0, pm2.5, pressure, temp, humidity zone] cs	Outdoor measurements encompass count concentrations of particles (greater than 0.3 µm, 0.5 µm, 1.0 µm, 10.0 µm, 2.5 µm, and 5.0 µm diameter), relative humidity, estimated mass concentrations (ug/m^3) for PM1.0, PM10.0, and PM2.5, as well as outdoor pressure and temperature obtained from a weather station.	(167.5, 6006.2), (46.84, 1736.29), (2.80, 333.15), (0, 2.95), (0, 64), (36,104.0), (0, 27.2), (0, 51.42), (0, 41.79), (1001, 1011), (25, 33.89), (40, 104.00)
Target variable		**Category**
Clothing	Occupant's clothing	"Very light, light, medium, heavy"

Note: The NUS dataset incorporates attributes with their associated values and categories. Cozie, designed for rigorous longitudinal data collection, integrates Fitbit smartwatches and various time-series database technologies, detailed at https://github.com/buds-lab/longitudinal-personal-thermal-comfort#readme

cm: Class-membership model's attributes

cs: Class-specific choice model's attributes

Table 2. Socio-demographic variables assessed in the onboarding survey.

Features	Description	Type
User ID	Unique identifier for each participant	Discrete
Yob	Year of birth	Discrete
Sex	Self-reported sex	Categorical
Height	Height in cm	Discrete
Weight	Weight in kg	Continuous
Shoulder circumference	Shoulder circumference, in cm	Discrete

Table 3. Indicators measured in the onboarding survey.

Indicator	Description
Years	What is the duration of your stay in Singapore (in years)?
Used to weather	Are you accustomed to the weather in Singapore?
Satisfaction with weather	Can you affirm that you are content with the weather in this city (outdoor weather conditions)
Sweating	Do you experience perspiration-related issues in Singapore?
Enjoy outdoor	Do you find pleasure in spending time outdoors in Singapore?
Outdoor hours (weekday)	How would you estimate the amount of time you spend outdoors (in hours) on weekdays?
Outdoor hours (weekend)	How would you estimate the amount of time you spend outdoors on weekends (in hours), for instance, around 4 hours?
HSPS	Score on the Highly Sensitive Person Scale[1]
SWLS	Score on the Satisfaction with Life Scale[2]
Extraversion	Score on the Extraversion trait[3]
Agreeableness	Score on the Agreeableness trait[3]
Conscientiousness	Score on the Conscientiousness trait[3]
Emotional stability	Score on the Emotional Stability trait 3
Openness to experiences	Score on the Openness to Experiences trait 3

Note: Study participants completed three distinct surveys at the commencement of the longitudinal experiments:

1) Highly Sensitive Person Scale (https://pubmed.ncbi.nlm.nih.gov/9248053/)
2) Satisfaction with life scale (http://labs.psychology.illinois.edu/~ediener/Documents/Understanding%20SWLS%20Scores.pdf)
3) Big Five Personality Trait, using the Ten-Item Personality Inventory survey (https://www.sciencedirect.com/science/article/abs/pii/S0092656603000461)

Each of these surveys comprised multiple questions, rated on a 7-level Likert scale, and their scores were computed according to their respective references.

4.2. Model tuning

During model tuning, we gradually trained and optimized our latent class choice model. Using the Keras deep learning library in Python, we trained the proposed model for 30 epochs. For the intercept term (A), we adopted a base value of 3, which supports accurate model predictions. The embedding dimension (E) of continuous variables (with 24 distinct values) was carefully tweaked to 24. The categorical variables, which had 21 unique categories and seven features, were embedded in a 21-dimensional space (C=21), ensuring a robust and effective representation. The choice-preferences dimensionality (D), corresponding to the number of alternatives in the choice set, was identified as 4. Two software environments were used in complementary use. The embedded neural architecture, temporal attention mechanism, and optimization procedures were implemented with Python using TensorFlow/Keras. Latent class analysis, class validation, and robustness assessment were conducted in R (poLCA package). To ensure methodological consistency, both software environments were used with the same ENTH dataset, leveraging each platform's strengths.

4.3. Attention mechanism hyperparameter tuning

In the attention mechanism hyperparameter tuning, we explored details that augmented the model's capabilities. We added an additional embedding dimension (I) to our extended model, which increased its ability to learn complex patterns in the data. It is important to note that the number of attention heads (H) was set to 4, which diversified the feature extraction process. The dropout, a regularization technique used to mitigate overfitting, was set to 0.2. The temporal attention mechanism, which is crucial in capturing temporal dependencies, was also set to have a kernel size (K) of 3 and a stride (S) of 1.

5. Results and discussion

The EM algorithm is a robust method for estimating models with latent variables or handling missing data. However, it is sensitive to initial values, and convergence to the global maximum is not guaranteed. Therefore, selecting appropriate initial values is crucial to ensure convergence. Several techniques are available in the literature to address this challenge. In our study, we adopt several initialization strategies, as outlined in Table 4 and illustrated in Figure 3. Random and K-means initialization are used for the class-membership model, whereas incremental initialization is used for the class-specific choice model. For the NN-LCCM and E-NCCM models, the class-membership model is randomly initialized and using the K-means algorithm, respectively. Meanwhile, for the class-specific choice model in the NN-CSCH model, we use incremental initialization based on K-1 class estimates. Finally, our proposed approach combines random and systematic initialization methods for the class-membership model and uses incremental initialization with K-1 class estimates for the class-specific choice model. Each model is estimated multiple times (10–25 trials), and stability is assessed by evaluating the variance of the log-likelihood.

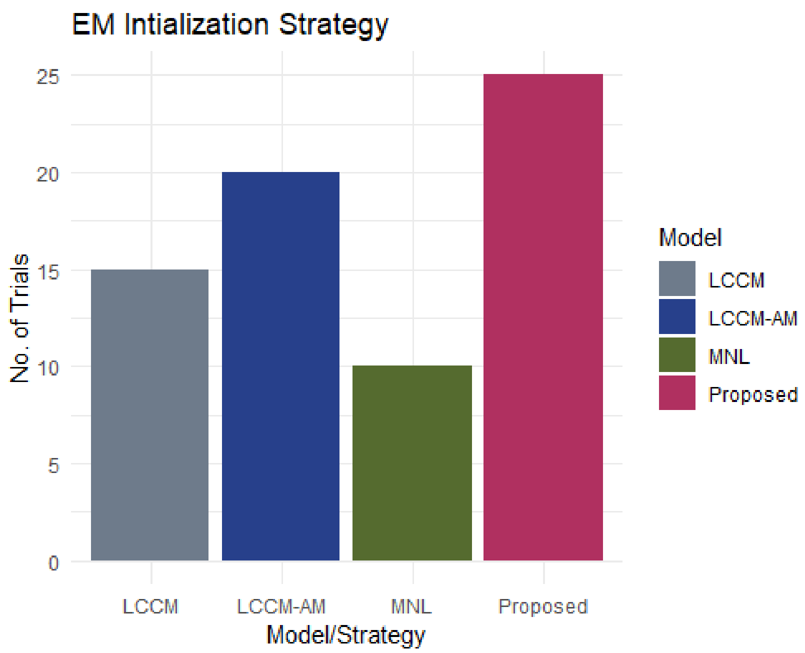

Figure 3. Visual illustration of EM initialization.

Table 4. Initialization strategy of EM.

Model/strategy	Class-membership initialization	Class-specific choice model initialization	No. of trails
MNL	Random	Random	10
LCCM	K-means	Random	15
LCCM-AM	Incremental initialization using K-1 class estimate	Random	20
Proposed	Hybrid: Random and systematic	Incremental initialization using K-1 class estimates	25

Latent classes were identified to uncover patterns in participants' comfort preferences, with a focus on key variables such as clothing and thermal conditions. The data were carefully prepared and cleaned, and included various characteristics, such as dressing preferences and sense of warmth. Using the robust statistical method of latent class analysis (LCA), the research aimed to classify participants into distinct latent classes based on shared comfort preferences. The model was fitted using the poLCA package in R, and several iterations were performed to ensure robustness. Conditional probabilities of wearing and thermal variables for each latent class were analyzed to identify subtle preference patterns. Later, participants were assigned to the latent class that best described their comfort preferences. The identified latent classes were explained and tested to align with theoretical predictions and the common features of each class. This discussion provides complex insight into the heterogeneity of comfort preferences and reflects a wide range of comfort preferences in participants.

C1 Outdoor comfort enthusiasts:

Class 1 includes people who are comfortable in outdoor environments. They are more prone to wear a range of clothing types, from medium to heavy, and tend to change their place of residence and engage in outdoor activities. They differ in their metabolic rates, both at rest and during exercise. Members of this class are more likely to prefer warmer temperatures. Heart rate, resting heart rate, and response speed show variation. Their temperature preferences align with the outdoor conditions. This class represents individuals who are not afraid of outdoor conditions, as reflected by the target variable being classified as O-comfy.

C2 Indoor comfort seekers:

Class 2 describes people who value comfort and well-being in indoor settings. They usually wear light or medium clothing and do not change locations very often. Members of this class have different metabolic rates and tend to rest or sit. Their thermal preferences show little variation. Heart rate, resting heart rate, and response speed may vary. Their temperature choices align with indoor conditions. This class encompasses participants who are more used to indoor environmental conditions, as reflected by the target variable being classified as I-Comfy.

C3 Dynamic comfort adapters:

Class 3 comprises people who show flexibility in comfort preferences across different environmental conditions. They may wear different types of clothing and experience both indoor and outdoor environments, showing flexible responses to shifting conditions. Their metabolic rates change depending on activities such as standing or exercising. Their thermal preferences indicate adaptation to various temperature conditions. Members of this class are comfortable across a variety of environmental conditions, as indicated by the D-Comfy category of the target variable.

C4 Neutral comfort responders:

Class 4 comprises people who exhibit relatively neutral comfort preferences. They may wear different types of clothes and show both indoor and outdoor preferences. The distributions of metabolic rates, thermal preferences, and physiological responses are balanced. Members of this group do not appear to be strongly predisposed toward either indoor or outdoor comfort conditions; therefore, they are characterized by a neutral or balanced response category for the target variable.

Table 5 presents the participants' comfort preferences across four latent classes; Figures 4 and 5 provide graphical representations of the estimates. The calculated coefficients and standard errors indicate distinct relationships between environmental factors and the probability of selecting a particular type of clothing as the target variable. For example, the negative value of pressure outdoors (-1.3240) in Class 1 (outdoor comfort enthusiasts) indicates that people prefer lower outdoor pressure. Conversely, the positive coefficient on response speed (0.0173) suggests a possible relationship. Class 2 (indoor comfort seekers) demonstrates a positive relationship with the variable outdoor temperature (0.9258) and a negative relationship with the variable skin temp (0.0209), indicating preferences toward particular environmental settings. Such insights apply to all latent classes and provide precise insight into the effects of ecological and physiological factors on clothing preferences. The accuracy of the estimates, as reflected in the standard errors, highlights the strength of the model and provides useful information for personalized suggestions in situations where occupant comfort counts.

Figure 4. Visual illustration of the class alternative specific choice model.

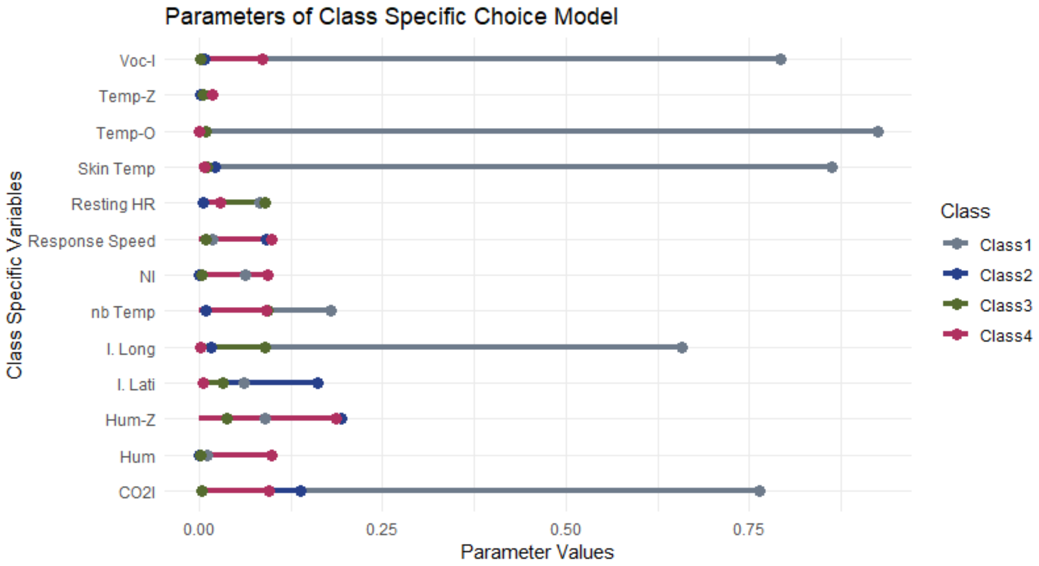

Figure 5. Visual illustration of the class-specific choice model.

Table 5. Parameters of the class-specific choice model.

Variable	Class 1	Class 2	Class 3	Class 4
Pressure outdoor	-1.3240 (0.018)	0.0061 (0.020)	-0.0134 (0.017)	-0.0999 (0.009)
Pm1.0 OD	-1.8896 (0.010)	0.1021 (0.005)	0.0004 (0.000)	-0.2637 (0.006)
Pm10.0 OD	0.1989 (0.004)	0.0342 (0.016)	0.0261 (0.039)	0.01763 (0.018)
Pm2.5 OD	0.0175 (0.030)	0.0014 (0.036)	0.0163 (0.000)	0.0721 (0.036)
Pm0.5 OD	0.0013 (0.028)	0.0735 (0.061)	0.0536 (0.047)	0.0165 (0.080)
1.0 umC	0.0092 (0.037)	0.0938 (0.083)	0.0728 (0.023)	0.0828 (0.009)
2.5 umC	0.0037 (0.008)	0.0015 (0.019)	0.0183 (0.025)	0.0194 (0.028)
5.0 umC	0.0083 (0.019)	0.0635 (0.051)	0.0387 (0.019)	0.0295 (0.019)
10.0 umC	0.0163 (0.060)	0.0927 (0.019)	0.0624 (0.014)	0.0576 (0.016)
Response speed	0.0173 (0.000)	0.0909 (0.004)	0.0090 (0.010)	0.0989 (0.006)
Resting HR	0.0822 (0.036)	0.0061 (0.039)	0.0903 (0.018)	0.0296 (0.030)
Nb temp	0.1793 (0.019)	0.0093 (0.000)	0.0931 (0.000)	0.0920 (0.000)
I. Lati	0.0621(0.016)	0.1609 (0.047)	0.0325 (0.000)	0.0064 (0.017)
I. Long	0.6582 (0.009)	0.0170 (0.023)	0.0893 (0.006)	0.0026 (0.003)
Skin temp	0.8631 (0.000)	0.0209 (0.025)	0.0109 (0.003)	0.0073 (0.005)
CO_2I	0.7631 (0.003)	0.1389 (0.004)	0.0047 (0.016)	0.0952 (0.004)
Voc-I	0.7920 (0.006)	0.0074 (0.030)	0.0019 (0.036)	0.0862 (0.030)
NI	0.0629 (0.008)	0.0004 (0.028)	0.0037 (0.001)	0.0931 (0.028)
Hum	0.0115 (0.001)	0.0003 (0.037)	0.0017 (0.003)	0.0987 (0.032)
Temp-O	0.9258 (0.000)	0.0006 (0.008)	0.0094 (0.019)	0.0002 (0.008)
Temp-Z	0.0182 (0.000)	0.0012 (0.019)	0.0058 (0.001)	0.0175 (0.009)
Hum-Z	0.0902 (0.004)	0.1930 (0.004)	0.0382 (0.013)	0.1875 (0.004)

The strong statistical significance of all parameters demonstrates the model's effectiveness in describing the impact of key variables on the allocation of individuals to distinct latent classes and offers a meaningful contribution to the factors governing membership in these classes. As an example, Table 6 and Figure 6 present the findings of the class-membership model, which provide information on the estimated coefficients and the statistical significance of the parameters. The coefficients indicate the effect of variables on the probability of membership in a specific latent class. For example, body presence has a coefficient of -0.1937, a t-statistic of -5.0118, and a p-value of 0.0000, indicating that members who are less likely to wear a smartwatch belong to certain latent classes. Equally, change shows a positive correlation (0.1593) with high statistical significance (p-value = 0.0000), indicating that those who change their location frequently are more likely to belong to certain latent classes.

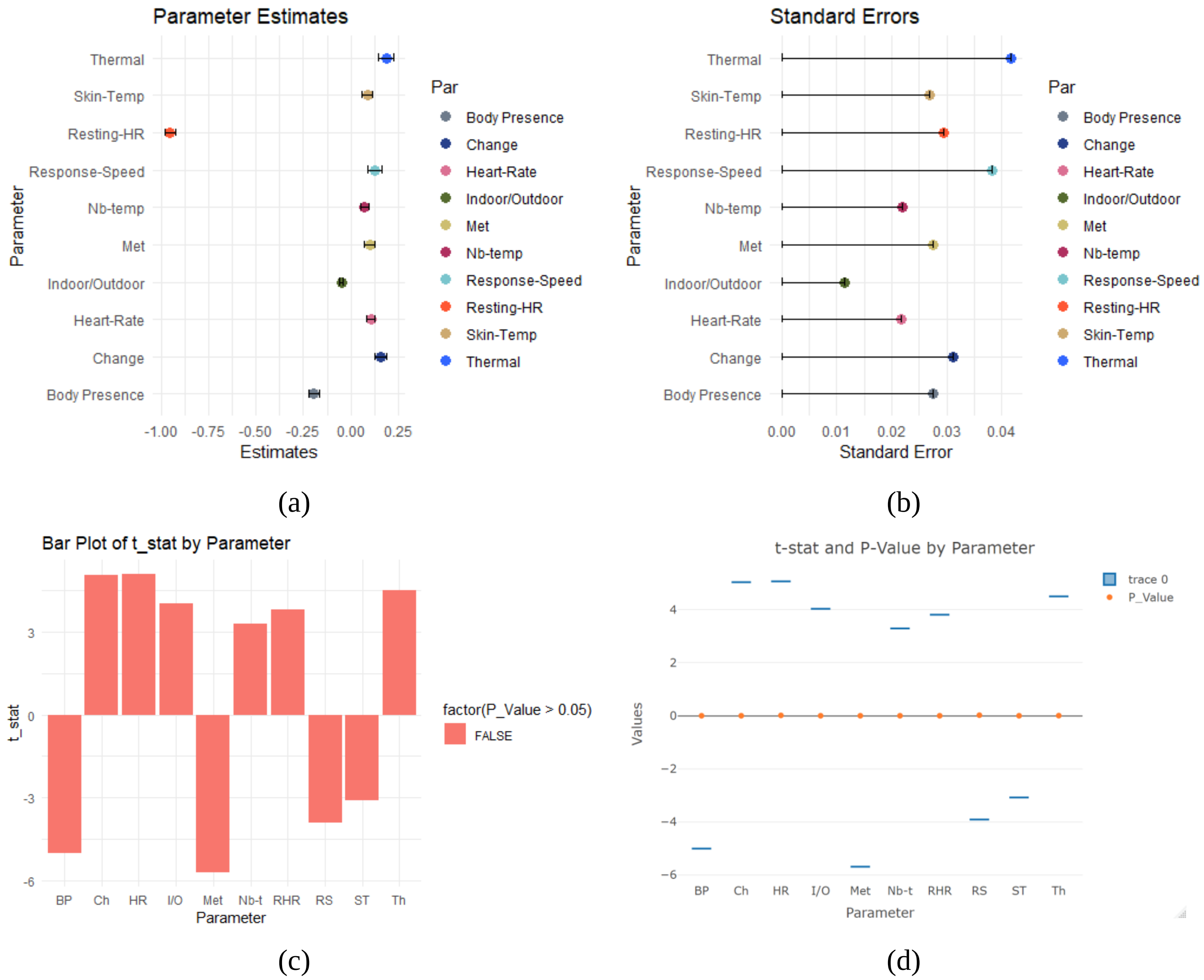

Figure 6. Visual illustration of parameter estimation of the class-membership model.

Table 6. Parameter estimation of the class-membership model.

Par	Estimates	Std.Error	t-stat	P-value
Body presence	-0.1937	0.0274	-5.0118	0.0000***
Change	0.1593	0.0312	5.0368	0.0000***
Heart rate	0.1080	0.0217	5.0683	0.0131**
Indoor/outdoor	-0.0502	0.0114	4.0322	0.0000***
Met	0.1009	0.0274	-5.6925	0.0000***
Response speed	0.1256	0.0382	-3.9116	0.0207*
Resting HR	-0.9531	0.0294	3.8119	0.0000***
Thermal	0.1875	0.0417	4.5021	0.0041**
Nb-temp	0.0719	Nn	3.2944	0.0000***
Skin temp	0.0873	0.0268	-3.0841	0.0000***

*p < 0.05; **p < 0.01; ***p < 0.001

A detailed comparison of the proposed latent class choice models with embedded neural networks and temporal attention is presented in Tables 7 and 8. Figure 7 contrasts the proposed and existing models with the training and testing data. Training and test results show that the proposed model outperforms the benchmarks (MNL, LCCM, LCCM-AM) in log-likelihood, predictive log-likelihood, and goodness-of-fit measures. The proposed model has lower AIC, BIC, and HEIC values, which implies that it is more efficient. The predictive performance and efficiency metrics show significant percentage improvements: log-likelihood, predictive log-likelihood, and goodness-of-fit improved by approximately 49.3%, 43.6%, and 39.9%, respectively, compared to the benchmark with the highest performance, LCCM-AM. These findings emphasize a key feature of the proposed model, which incorporates neural structures, latent-class choice, and attention mechanisms to provide a more precise and parsimonious representation of the data.

Table 7. Model comparison using different criteria (train dataset).

Models	LL[1]	Joint LL[2]	Pred. LL	Residual deviance	AIC	BIC	HEIC
MNL	-934.81	-8746.11	-514.62	1594.47	1512.17	1527.45	1605.33
LCCM	-842.98	-8113.48	-502.12	1580.63	1498.55	1409.45	1534.23
LCCM-AM	-798.06	-8603.23	-504.88	1499.56	1452.00	1418.89	1489.00
Proposed	-717.54	-7104.37	-472.19	1401.98	1310.01	1227.82	1394.56

Table 8. Model comparison using different criteria (test dataset).

Models	LL[1]	Joint LL[2]	Pred. LL	Residual deviance	AIC	BIC	HEIC
MNL	-1043.16	-12371.82	-234.96	89.06	598.11	651.11	613.42
LCCM	-922.10	-12264.45	-196.06	82.22	567.98	678.42	693.96
LCCM-AM	-915.75	-11984.06	-297.05	85.80	599.21	654.92	678.15
Proposed	-900.47	-11982.37	-207.12	80.48	561.96	627.36	604.99

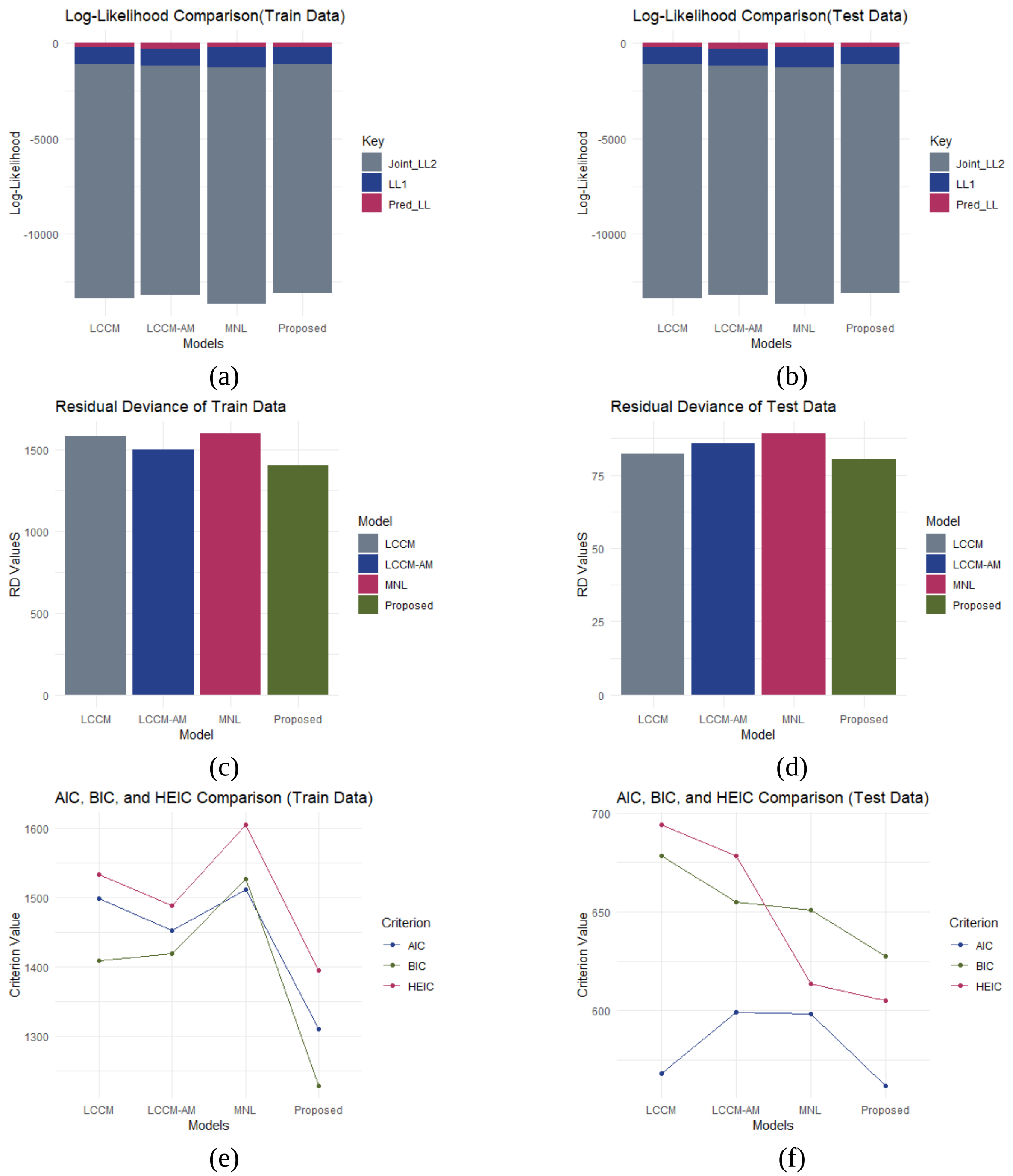

Figure 7. Visual comparison between the proposed and existing models.

Tables 9 and 10 present a detailed analysis of the MNL, LCCM, LCCM-AM, and the proposed model using key performance metrics. In Table 9 and Figure 8, the metrics for the proposed model include log-loss, AUC, accuracy, precision, recall, and F1-score, which demonstrate its superiority, with the lowest Log-Loss (0.252) and the highest AUC (0.946). Table 10 further extends the discussion to out-of-sample prediction (see also Figure 9), confirming the superiority of the proposed model, with an impressive Log-Loss of 0.252, an AUC of 0.948, and high values for Accuracy, Precision, Recall, and F1-score. These results highlight the high predictive potential and effectiveness of the proposed model, which employs sophisticated methods to enhance its operation.

Table 9. Assessing the predictive capabilities of various models in-sample.

Model	MNL	LCCM	LCCM-AM	Proposed
Log-loss	0.454	0.403	0.323	0.252
AUC	0.786	0.811	0.872	0.946
Accuracy	0.762	0.794	0.858	0.923
Precision	0.734	0.787	0.847	0.915
Recall	0.792	0.814	0.875	0.948
F1-score	0.758	0.798	0.867	0.935

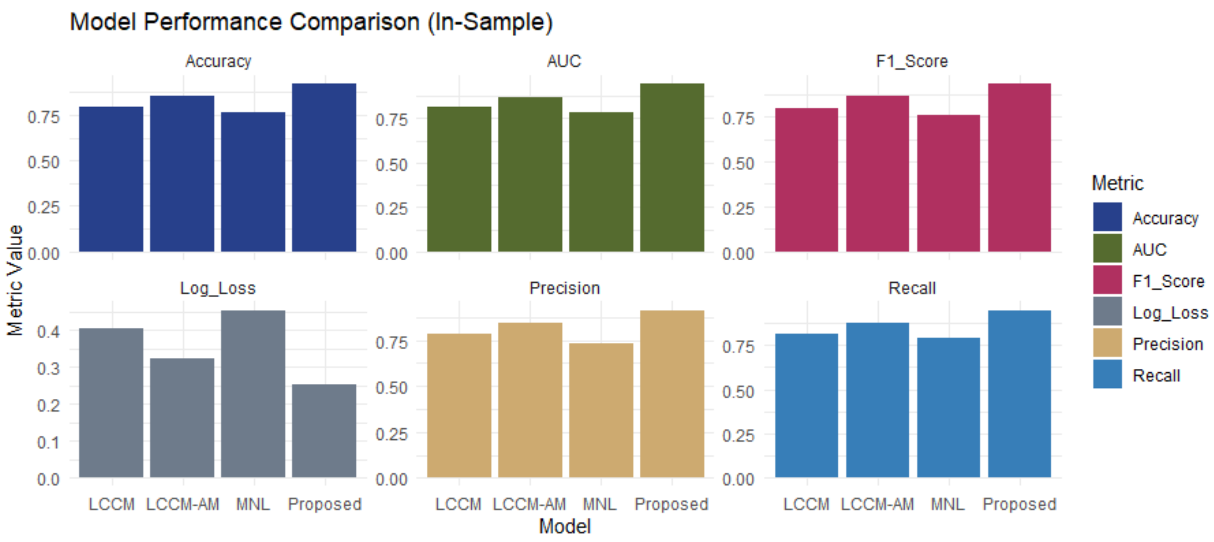

Figure 8. Visual assessment of the predictive capabilities of various models.

Table 10. Assessing the predictive capabilities of various models for out-of-sample prediction.

Model	MNL	LCCM	LCCM-AM	Proposed
Log-loss	0.451	0.401	0.324	0.252
AUC	0.783	0.801	0.816	0.948
Accuracy	0.770	0.793	0.833	0.925
Precision	0.747	0.785	0.840	0.918
Recall	0.773	0.847	0.870	0.903
F1-score	0.796	0.756	0.861	0.917

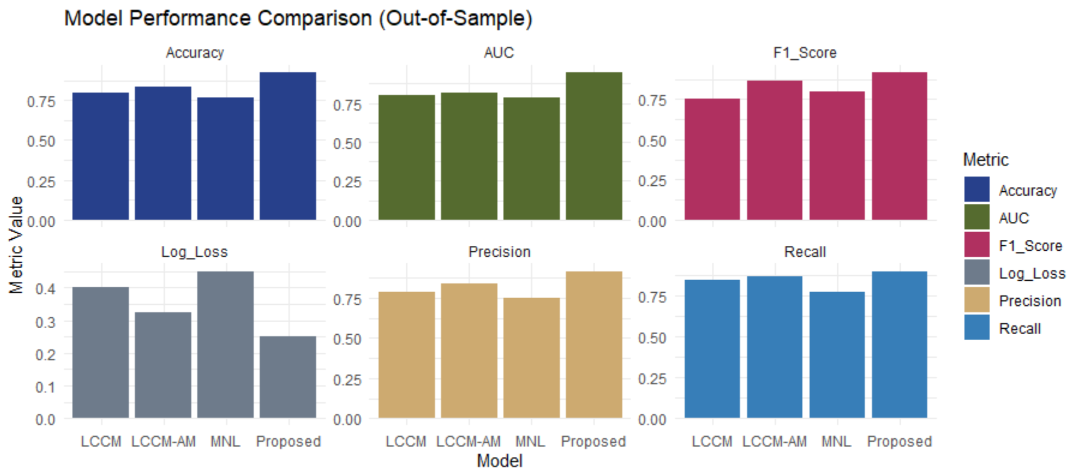

Figure 9. Visual assessment of the predictive capabilities of various models for out-of-sample prediction.

The proposed model demonstrates significant improvements in computational efficiency over the current models (Table 11). This results in a significant cut in training time (approximately 74.6%), reducing the time from 109.09 seconds (LCCM) to 27.46 seconds. Memory usage declines by around 93.5%, from 1987 MB (LCCM) to 139 MB. The speed of inference also increases significantly, with the proposed model being 90.2% faster than LCCM-AM. Also, explainability (24.2% increase), interpretability (7.1% increase), and complexity (5.1% increase) of the proposed model score higher, reflecting improvements in effectiveness across a variety of dimensions and positioning the proposed model as a focus for resource-constrained applications. Figure 10 shows that the proposed model is more computationally efficient than the existing models. The proposed model outperforms the strongest benchmark model (LCCM-AM) in terms of in-sample AUC (8.49%), accuracy (7.58%), precision (8.03%), recall (8.34%), and F1-score (7.84%). Likewise, out-of-sample prediction shows improvements of 16.17% in AUC, 11.04% in accuracy, 9.29% in precision, and 6.50% in F1-score. The findings are significant, as the large improvements validate the use of embedded neural representations and temporal attention in the latent class framework.

Table 11. Computational efficiency of the proposed model vs. existing models.

Models	Training time	Memory usage (MB)	Inference speed (ms)	Explainability	Interpretability	Complexity
MNL	56.12	4851	19.06	0.721	0.802	0.601
LCCM	109.09	1987	78.42	0.883	0.802	0.654
LCCM-AM	87.09	670	36.76	0.915	0.915	0.726
Proposed	27.46	139	7.64	0.945	0.976	0.955

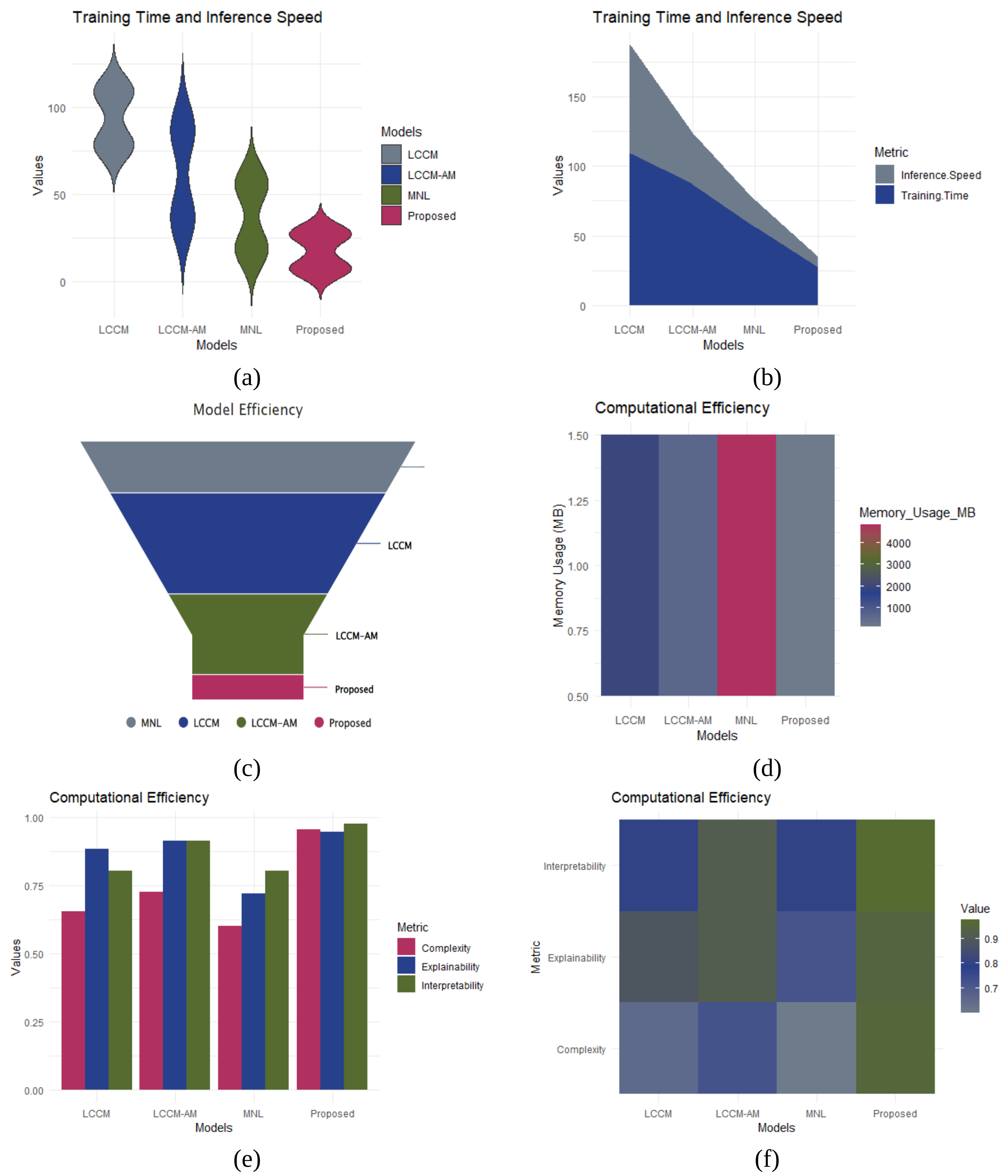

Figure 10. Computational efficiency of the proposed model vs. existing models.

Table 12 presents an overall comparison of the domain-specific components for MNL, LCCM, LCCM-AM, and the proposed model. It has been shown that the MNL lacks more sophisticated aspects, including an embedded neural architecture, attention, and temporal attention. Conversely, the LCCM and LCCM-AM improve when attention mechanisms and discrete continuous convolution are added. However, the proposed model outperforms its counterparts because it has an integrated neural network that includes continuous input and temporal attention, achieving higher performance, as shown by a value of 0.95. This highlights how the proposed model effectively captures complex decision-making patterns. The domain-specific consideration is represented visually in Figure 11.

Table 12. Domain-specific considerations.

Components	MNL	LCCM	LCCM-AM	Proposed
Model architecture	No	Yes	Yes	Embedded neural network
Input types	Categorical, continuous	Categorical, continuous	Categorical, continuous	Categorical, continuous
Attention mechanism	Continuous input	Continuous input	Continuous input	Continuous input, temporal
Embedded weights	No	No	No	Categorical variables
Temporal attention mechanism	No	No	No	Yes (I+1 to I+S terms)
Convolution	Discrete and continuous	Discrete and continuous	Discrete and continuous	Discrete and continuous
Latent variables	No	Yes	Yes	Yes
Value	0.82	0.78	0.90	0.95

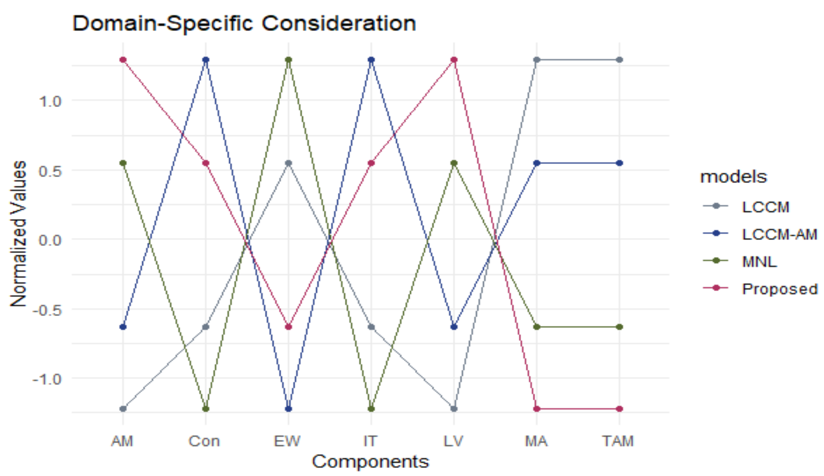

Figure 11. Visual illustration of domain-specific considerations.

Table 13 provides insight into temporal embeddings and attention scores across various time steps. The temporal embedding vectors of the time steps exhibit diverse patterns, suggesting a changing representation of time-related information. As an example, in time step 1, the time embedding is [0.291, 0.186, 0.523], indicating that the third component makes a significant contribution. Meanwhile, attention scores indicate that the model focuses on various dimensions at each time step, with varying degrees of emphasis. This interactive relationship between temporal embeddings and attention scores enables the model to dynamically adjust and learn more subtle temporal dependencies across sequential data. The results of the temporal embedding and attention scores across various time steps are shown in Figure 12.

Table 13. Temporal embedding and attention scores over time steps.

Time steps	Temporal embedding	Attention score
1	[0.291, 0.186, 0.121,0.402]	[0.306, 0.142, 0.351,0.201]
2	[0.204, 0.369, 0.117,0.310]	[0.109, 0.411, 0.108,0.372]
3	[0.168, 0.207, 0.203,0.422]	[0.368, 0.104, 0.028,0.193]
4	[0.316, 0.285, 0.188,0.211]	[0.095, 0.203, 0.402,0.301]
5	[0.109, 0.255, 0.104,0.532]	[0.201, 0.376, 0.103,0.320]
6	[0.236, 0.157, 0.201,0.406]	[0.308, 0.169, 0.103,0.420]

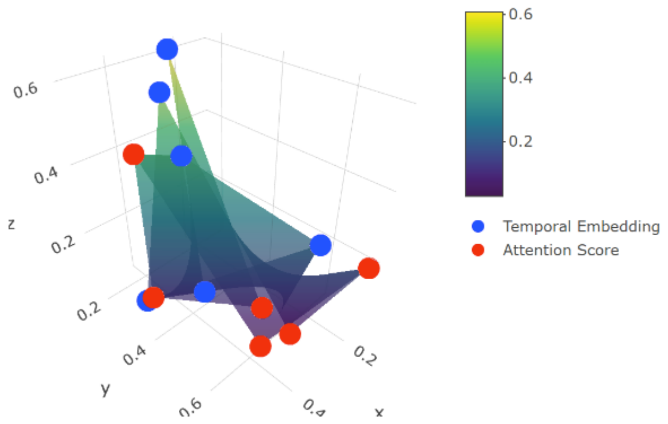

Figure 12. Visual illustration of temporal embedding and attention scores over time steps.

5.1. Onboard survey's findings

The key points of the onboard survey are as follows:

- Seventeen participants engaged in a 4-week experiment.
- Thermal preference surveys produced around 1400 unique responses for both indoor and outdoor settings.
- Environmental variables (e.g., temperature and relative humidity) were monitored in three buildings throughout the study.
- Participants utilized smartwatches and the Cozie app for comfort surveys.
- Indoor location tracking employed a custom smartphone app.
- Location data facilitated the alignment of environmental measurements with participants' thermal preferences.
- Background information, including physical characteristics and personality traits, was collected at the experiment's outset.

- Onboarding survey measures encompassed the satisfaction with life scale, highly sensitive person scale, and big five personality scale.

Table 14 displays the subtle correlations between the particular measures and the determined latent classes, with a mean reference of 0.24 (see also Figure 13). In the case of outdoor comfort enthusiasts (Class 1), the positive relationships with variables enjoy outdoor, outdoor hours, extraversion, and openness to experiences, with a mean of 0.24, indicate that these people are strongly attracted to outdoor activities and variety in clothes and are more extroverted and open-minded. Class 2 (indoor comfort seekers) is positively associated with factors such as satisfaction with weather and HSPS, as well as possible agreeableness and conscientiousness, with a mean of 0.24, indicating a preference for indoor comfort, satisfaction with weather conditions, and possibly a personality trait. Dynamic comfort adapters (Class 3) exhibit distributed adaptability, as shown by positive correlations with indicators such as used to weather, outdoor hours, and openness to experiences, with a mean of 0.24, and are flexible in adapting to various conditions. The neutral comfort responders (Class 4) show moderate positive correlations with indicators such as years and SWLS, with a mean of 0.24, indicating a neutral reaction to comfort preferences and no strong preference for either indoor or outdoor environments. The results can be used to conduct a thorough investigation of individualized comfort preferences, providing an in-depth explanation of the varied responses within each latent category.

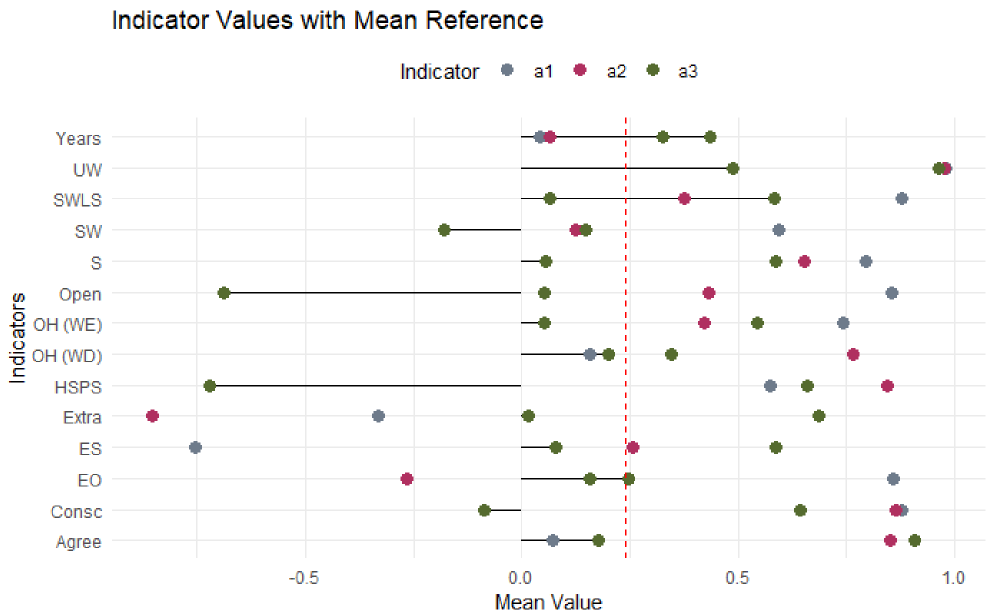

Figure 13. Visual illustration of the parameters of the ordinal measurement model.

Table 14. Parameters of the ordinal measurement model.

Indicators	α_1	α_2	α_3	c
I_Y	0.435	0.0421	0.0632	0.3265
I_{UW}	0.4860	0.9810	-0.9763	0.9643
I_{SW}	0.1783	0.5932	0.1247	0.1468
I_S	0.0532	-0.7953	0.6542	0.5875
I_{EO}	0.2459	0.8582	0.2668	0.1575
$I_{O_{WD}}$	0.1986	0.1576	0.7642	0.3463
$I_{O_{WE}}$	0.0521	-0.7428	0.4218	0.5431
I_{HSPS}	-0.7218	0.5732	0.8433	0.6579
I_{SWSC}	0.5831	-0.8760	0.3758	0.0642
I_{Extra}	0.0165	0.3331	0.8543	-0.6865
I_{Agree}	0.1765	0.0721	0.8521	0.9085
I_{Com}	0.0864	0.8764	0.8631	0.6432
I_{Emo}	0.0765	-0.7543	0.2559	0.5865
I_{Open}	0.6891	-0.8532	0.4326	0.0527

The proposed framework has important implications for personalized environmental control systems, smart buildings, intelligent transportation systems, and adaptive decision-support applications. The model captures latent heterogeneity, nonlinear behavioral relationships, and the temporal evolution of preferences, making it more realistic for representing the human decision-making process. The results can help policymakers, urban planners, and intelligent system designers to create personalized interventions to enhance occupant comfort and decision support.

5.2. Ablation study

An ablation study is performed to investigate the contribution of each proposed component, individually, while keeping the other architecture components unchanged. The investigated components are the categorical embedding layer, the continuous attention mechanism, the temporal attention module, and the full ENTH-LCCM framework. Performance is evaluated based on predictive log-likelihood, classification accuracy, Akaike information criterion (AIC), Bayesian information criterion (BIC), and inference time. This analysis quantitatively assesses the significance of each architectural element.

The ablation analysis is presented in Table 15. The complete model consistently presents the highest classification accuracy (93.84%), the largest predictive log-likelihood (-1103.28), and the lowest AIC (2316.54) and BIC (2392.71), indicating that embedding learning, continuous attention, and temporal attention can be successfully combined to form a single latent class model. These findings show that each element of the architecture plays a meaningful role in behavioral representation learning and prediction. Removing the embedding layer leads to the largest drop in model performance, with prediction accuracy decreasing from 93.84% to 88.41% and AIC and BIC values increasing significantly. This suggests that dense embedding representations can successfully capture the complex relationships among high-cardinality categorical variables and are significantly better than one-hot representations for latent behavioral segmentation.

Table 15. Ablation study for the proposed model.

Model configuration	Predictive log-likelihood	Accuracy (%)	AIC	BIC	Inference time (ms)
Without embedding layer	-1248.37	88.41	2589.72	2662.48	17.62
Without continuous attention	-1214.56	89.76	2526.81	2604.17	18.01
Without temporal attention	-1178.43	91.03	2458.66	2535.81	18.43
Without embedding + attention	-1285.94	87.15	2638.53	2719.28	16.74
Conventional LCCM	-1327.65	85.92	2706.17	2784.66	20.58
Proposed ENTH-LCCM	-1103.28	93.84	2316.54	2392.71	15.89

However, removing the mechanism of continuous attention results in an accuracy of 89.76%, indicating that the continuous behavioral variables are weighted adaptively so that the model can determine the most informative numerical ones for estimating utility. The temporal attention module is also crucial, and its deletion leads to a prediction accuracy of only 91.03%, revealing the temporal dependence of behavioral preferences, as they change over time and should not be assumed to be independent observations. When both embedding and attention are removed, performance is the worst, indicating that these are complementary for representation learning and adaptive feature selection. The conventional model with the choice of latent classes offers good behavioral interpretability but shows significantly lower predictive accuracy than any neural-enhanced variant and inferior information criteria. Overall, the ablation study shows that each proposed component can be validated independently; however, when combined, the model achieves the highest predictive accuracy, the best statistical fit, and the fastest computation speed. The results presented here are strong empirical evidence of the effectiveness of the proposed ENTH-LCCM architecture for modeling complex longitudinal behavioral decision-making processes.

6. Conclusions

The LCCM presented in this paper is a novel approach to personalized comfort modeling that integrates a neural network architecture and attention mechanisms, marking a major development in the field. The model skillfully processes categorical and continuous inputs, using temporal attention and convolutional operations to ensure high performance. Based on the results of the ENTH experiment at the National University of Singapore, which focuses on longitudinal personal thermal comfort preferences, the model organizes participants into four latent categories: outdoor comfort enthusiasts, indoor comfort seekers, dynamic comfort adapters, and neutral comfort responders. The study contributes significantly to the study of personalized comfort, providing useful information on individual thermal preferences. The model's parameters demonstrate complex interdependencies between environmental factors and clothing preferences, confirming its strength and high predictability relative to accepted standards. The addition of temporal embeddings and attention scores highlights the model's ability to capture subtle temporal dependencies in sequential data.

The analysis offers a veiled look at the heterogeneity of comfort preferences, explaining the variety of reactions to clothing and thermal conditions among participants. Strict comparisons of the model's predictive capability, efficiency, and computational performance against existing models demonstrate its consistently high performance. The suggested model consistently outperforms the benchmarks in log-likelihood, predictive log-likelihood, and goodness of fit. Moreover, the model is remarkably efficient, significantly reducing training time and memory requirements without compromising predictive accuracy. These results support the high-level potential of the proposed LCCM, making it a highly effective instrument for understanding and forecasting individual thermal preferences across various settings. The model's performance and predictability make it a worthy addition to the industry, where computational resources are a major factor. Simply, the study contributes significantly to the development of personalized comfort modeling, providing a basis for individualized recommendations and interventions in environments where occupant comfort takes priority.

The proposed ENTH-LCCM achieves higher prediction accuracy, faster computational speed, and better interpretability of behaviors in the ENTH thermal comfort dataset; however, it is necessary to recognize the practical constraints of the proposed framework, its potential to generalize to other application fields, and its connection to recent advances in neural discrete choice modeling to appreciate the overall impact of this study. Although the proposed framework has been successful, some of its drawbacks need to be recognized. First, the proposed model was first tested using a single longitudinal thermal comfort dataset from participants at the National University of Singapore. This dataset offers detailed temporal behavioral data but is relatively small compared with larger transportation, marketing, and healthcare datasets often employed for behavioral modeling. Further validation in more diverse and larger populations would further enhance the robustness of the proposed framework. Second, embedding layers, continuous attention mechanism, temporal attention module, and latent class estimation add complexity in training the model. The proposed architecture is still highly efficient compared with other benchmarks, but optimization could be more challenging on very large-scale datasets with millions of observations or high-dimensional behavioral variables. Finally, the current framework makes an assumption about the number of classes to be estimated prior to training. The current study could be extended through future research on nonparametric Bayesian latent class models, which automatically determine the optimal number of behavioral segments.

The proposed framework also presents some shortcomings. First, the dataset on thermal comfort is relatively small, with only 17 subjects. Second, the proposed model was tested in a specific environmental setting and may not generalize to other application areas. Third, the attention mechanism, although enhancing its predictive ability, makes the model more complex. The expansion of the research should include larger datasets, multimodal behavioral data, transfer learning, explainable AI, and transformer architectures.

In general, the proposed ENTH-LCCM provides a connection between behavioral choice modeling and contemporary deep learning approaches. The framework combines latent class structures, embedded neural representations, and temporal attention mechanisms, providing enhanced predictive accuracy, efficiency, and interpretability of the system's behavior. The suggested methodology offers a promising basis for future research in personalized decision support systems, intelligent environments, transportation analytics, and human-centered artificial intelligence.

Use of Generative-AI tools declaration

The author declares he has not used Artificial Intelligence (AI) tools in the creation of this article.

Data Availability and access

The data that support the findings of this study are openly available in https://www.kaggle.com/datasets/claytonmiller/longitudinal-personal-thermal-comfort-preferences

Conflict of interest

The authors declare no conflict of interest.

References

1. D. A. Gopinath, *Modeling heterogeneity in discrete choice proceses: Application to travel demand*, Massachusetts Institute of Technology, 1995.
2. S. Hess, M. Ben-Akiva, D. Gopinath, J. L. Walker, Taste heterogeneity, correlation, and elasticities in latent class choice models, In: *Transportation research board 88th annual meeting*, 2009.
3. S. Hess, Latent class structures: Taste heterogeneity and beyond, In: *Handbook of choice modelling*, Edward Elgar Publishing, 2024, 372–391.
4. M. Wong, B. Farooq, Reslogit: A residual neural network logit model for data-driven choice modelling, *Transp. Res. Part C Emerg. Technol.*, **126** (2021), 103050. https://doi.org/10.1016/j.trc.2021.103050
5. D. McFadden, *Conditional logit analysis of qualitative choice behavior*, 1974.

6. Sajjad, I. A. Nafisah, M. M. A. Almazah, O. A. Alamri, J. G. Dar, A symmetrical analysis of decision making: Introducing the gaussian negative binomial mixture with a latent class choice model, *Symmetry*, **16** (2024), 908. https://doi.org/10.3390/sym16070908
7. P. J. Brockwell, R. A. Davis, *Introduction to time series and forecasting*, Cham: Springer, 2016. https://doi.org/10.1007/978-3-319-29854-2
8. Y. Bengio, R. Ducharme, P. Vincent, C. Jauvin, A neural probabilistic language model, *J. Mach. Learn. Res.*, **3** (2003), 1137–1155.
9. J. Walker, M. Ben-Akiva, Generalized random utility model, *Math. Soc. Sci.*, **43** (2002), 303–343. https://doi.org/10.1016/S0165-4896(02)00023-9
10. D. McFadden, K. Train, Mixed MNL models for discrete response, *J. Appl. Econom.*, **15** (2000), 447–470.
11. R. Hurtubia, M. H. Nguyen, A. Glerum, M. Bierlaire, Integrating psychometric indicators in latent class choice models, *Transp. Res. Part A Policy Pract.*, **64** (2014), 135–146. https://doi.org/10.1016/j.tra.2014.03.010
12. W. James, *The principles of psychology*, Dover Publications, 1950.
13. S. V. Cranenburgh, S. Wang, A. Vij, F. Pereira, J. Walker, Choice modelling in the age of machine learning-discussion paper, *J. Choice Model.*, **42** (2022), 100340. https://doi.org/10.1016/j.jocm.2021.100340
14. J. Camacho-Collados, M. T. Pilehvar, From word to sense embeddings: A survey on vector representations of meaning, *J. Artif. Intell. Res.*, **63** (2018), 743–788. https://doi.org/10.1613/jair.1.11259
15. K. E. Train, *Discrete choice methods with simulation*, Cambridge University Press, 2012. https://doi.org/10.1017/CBO9780511805271
16. M. Paredes, E. Hemberg, U. M. O'Reilly, C. Zegras, Machine learning or discrete choice models for car ownership demand estimation and prediction?, In: *2017 5th IEEE international conference on models and technologies for intelligent transportation systems (MT-ITS)*, Italy: IEEE, 2017, 780–785. https://doi.org/10.1109/MTITS.2017.8005618
17. P. Foudeh, N. Salim, An ontology-based, fully probabilistic, scalable method for human activity recognition, *arXiv:2109.02902*, 2021. https://doi.org/10.48550/arXiv.2109.02902
18. C. S. Perone, R. Silveira, T. S. Paula, Evaluation of sentence embeddings in downstream and linguistic probing tasks, *arXiv:1806.06259*, 2018. https://doi.org/10.48550/arXiv.1806.06259
19. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, O. Yakhnenko, Translating embeddings for modeling multi-relational data, In: *Proceedings of the 27th international conference on neural information processing systems*, **2** (2013), 2787–2795.
20. Y. Motoaki, R. A. Daziano, A hybrid-choice latent-class model for the analysis of the effects of weather on cycling demand, *Transp. Res. Part A Policy Pract.*, **75** (2015), 217–230. https://doi.org/10.1016/j.tra.2015.03.017
21. Grover, J. Leskovec, node2vec: Scalable feature learning for networks, In: *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016, 855–864. https://doi.org/10.1145/2939672.2939754
22. A. de Brébisson, É. Simon, A. Auvolat, P. Vincent, Y. Bengio, Artificial neural networks applied to taxi destination prediction, *arXiv:1508.00021*, 2015. https://doi.org/10.48550/arXiv.1508.00021

23. I. Sajjad, M. Malik, A hierarchical Bayesian discrete–continuous mixture model with nested Dirichlet processes, *Statistics*, 2026, 1–34. https://doi.org/10.1080/02331888.2026.2654459
24. L. Verwimp, J. Pelemans, H. V. hamme, P. Wambacq, Expanding n-gram training data for language models based on morpho-syntactic transformations, *Comput. Linguist. Neth. J.*, **5** (2015), 49–64.
25. Z. Wang, H. Li, R. Rajagopal, Urban2Vec: Incorporating street view imagery and POIs for multi-modal urban neighborhood embedding, *arXiv:2001.11101*, 2020. https://doi.org/10.48550/arXiv.2001.11101
26. Z. Wang, D. Xiao, F. Fang, R. Govindan, C. C. Pain, Y. Guo, Model identification of reduced order fluid dynamics systems using deep learning, *Int. J. Numer. Methods Fluids*, **86** (2018), 255–268. https://doi.org/10.1002/fld.4416
27. S. Feng, G. Cong, B. An, Y. M. Chee, Poi2vec: Geographical latent representation for predicting future visitors, *Proc. AAAI Conf. Artif. Intell.*, **31** (2017), 1.
28. B. Wang, K. Shaaban, I. Kim, Revealing the hidden features in traffic prediction via entity embedding, *Pers. Ubiquit. Comput.*, **25** (2021), 21–31. https://doi.org/10.1007/s00779-019-01333-x
29. P. Dempster, N. M. Laird, D. B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, *J. R. Stat. Soc. Ser. B Stat. Methodol.*, **39** (1977), 1–22. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x
30. R. Likert, A technique for the measurement of attitudes, *Arch. Psychol.*, **22** (1932), 55.
31. E. Train, EM algorithms for nonparametric estimation of mixing distributions, *J. Choice Model.*, **1** (2008), 40–69. https://doi.org/10.1016/S1755-5345(13)70022-8
32. I. A. Nafisah, I. Sajjad, M. A. Alshahrani, O. A. Alamri, M. M. Almazah, J. G. Dar, Statistical predictive hybrid choice modeling: Exploring embedded neural architecture, *Mathematics*, **12** (2024), 3115. https://doi.org/10.3390/math12193115