Modeling Dynamic Consumer Preferences from Few-shot Data: A Meta-Learning Approach

Mingzhang Yin* Khaled Boughanmi† Anirban Mukherjee‡

Abstract

The ability to quickly capture and adapt to customer preferences is central for firms seeking to offer personalized products and improve retention. This objective becomes challenging when individual-level data on customer interactions are limited, as is often the case for new customers or short consumption sessions. To this end, we propose meta-temporal processes (MetaTP), a meta-learning framework that enables scalable personalization from a small number of individual observations. MetaTP is trained across a large collection of session-based tasks, allowing it to improve data efficiency and transfer shared structure across customers. To model customer interactions over time, MetaTP integrates a Transformer-based architecture that captures sequential consumption patterns within sessions. This design uncovers dynamic preference heterogeneity and enables accurate future predictions. We illustrate MetaTP through an application on customer sequential consumption of digital products, focusing on the lukewarm stage of the customer journey, a transition period characterized by limited individual observations. Empirically, MetaTP outperforms a comprehensive set of benchmark methods in few-shot prediction and reveals meaningful patterns of preference evolution through its interpretable parameters. Managerially, we demonstrate how firms can leverage MetaTP to optimize personalized recommendations with limited individual data, including product sequencing decisions and both open-loop and closed-loop session completion strategies.

Keywords: Meta-learning, Few-shot Data, Sequential Model, Personalization, Customer Heterogeneity

*Mingzhang Yin is an Assistant Professor of Marketing at the Warrington College of Business, University of Florida. Email: mingzhang.yin@warrington.ufl.edu.

†Khaled Boughanmi is an Assistant Professor of Marketing at the SC Johnson Graduate School of Management at Cornell University. Email: kb746@cornell.edu.

‡Anirban Mukherjee, Avyayam Holdings. Email: anirban@avyayamholdings.com.

INTRODUCTION

Quickly adapting to customer preferences with minimal individual-level data is crucial to a firm's success. This ability is important in two key scenarios. First, new customers are more likely to churn if they fail to quickly discover desired products or are dissatisfied with their early interactions. Firms thus have a strong incentive to personalize offerings at an early stage to cement customer loyalty. Second, customer preferences are often context-dependent and can vary over time. Thus, firms need to quickly assess what customers prefer in the current consumption context to be able to personalize their offerings. In both scenarios, firms need to infer customer tastes from a small number of individual-level observations, a challenge further compounded by the dynamic and heterogeneous nature of customer preferences. In this paper, we propose a solution that enables data-efficient, interpretable, and large-scale preference inference and personalization using limited individual data.

The need for data-efficient personalization is driven by both technological advancements and evolving social norms. The increasing accessibility of portable devices has shifted user interactions with online products toward short sessions. For example, the average Spotify listening session consists of only 10–20 tracks, yet users exhibit substantial shifts in music tastes across sessions (Brost, Mehrotra, and Jehan 2019). This highlights the importance of customizing content streaming based on a few individual-level observations. Additionally, recent privacy regulations such as the General Data Protection Regulation (GDPR) restrict the collection and storage of sensitive user data (Voigt and dem Bussche 2017). These constraints on the external data collection further motivate firms to tailor their offerings based on early-stage interaction data, making data-efficient personalization both necessary and strategic.

Previous research has developed various methods to analyze individual data and infer heterogeneous preferences. A widely used approach is the hierarchical Bayes (HB) method, as it yields individual-level parameters suitable for personalization (Allenby, Rossi, and McCulloch

2005). However, traditional HB methods often lack sufficient flexibility due to restrictive functional forms, and they often do not scale well to large datasets because of the computationally intensive Markov chain Monte Carlo (MCMC) inference. In contrast, machine learning (ML) approaches capture the relationship between products and customers through scalable optimization. Yet, these approaches typically require a large amount of individual data to fit an accurate predictive model for each customer, which may not be available for short sessions. Furthermore, these predictive models rely on the assumption that the training and testing data are independent and identically distributed (i.i.d.), an assumption that could be violated when the targets are new customers or sessions that exhibit different underlying preferences.

Motivated by these challenges, we develop a new methodology to model dynamic individual preferences from large-scale customer interaction data with sparse individual-level observations. Building on the meta-learning literature in artificial intelligence (AI) (Schmidhuber 1987; Finn, Abbeel, and Levine 2017), we propose a novel framework called the meta-temporal process (MetaTP), which infers customer preferences and predicts future behaviors by efficiently leveraging a limited number of customer interactions with products. MetaTP is a deep-learning-based probabilistic model that mimics the human learning process: learning from past experiences in solving similar problems enables it to solve a new problem quickly with minimal data. To the best of our knowledge, this work is the first in marketing that leverages meta-learning for customer journeys and preference measurement.

To achieve high efficiency with small-scale individual data, we train MetaTP over multiple *tasks* following the meta-learning framework (Finn, Abbeel, and Levine 2017; Gordon et al. 2018; Beck et al. 2023). Each task is constructed by splitting a session's data into a *context set* and a *target set*, where the context set contains only a small number of initial interactions from the session. During training, for each task, the model is provided with the small context set and is optimized to predict the customer responses to the items in the target set. Therefore, each task operates as a supervised learning problem by its own, with the context set playing the role of training data and the target set serving as validation data (see Fig. 1). Due to the

small size of the context set by design, each task is referred to as a *few-shot* learning problem (Brown et al. 2020; Gharoun et al. 2024). By training to solve a large number of few-shot prediction tasks, the model captures the commonalities across these tasks and learns how to effectively leverage the context data. This fast-learning ability enables the model to infer and predict a new session based on a few initial observations as the new context.

In designing MetaTP, we adopt a novel encoder-decoder structure and fit it over the session-based meta-training tasks. The encoder is designed to summarize the context set and capture the relationship between the context and target sets of each task, while the decoder serves as a task-specific (or session-specific) predictive model. Existing meta-learning methods typically assume the data within a task are exchangeable and non-sequential (e.g., for few-shot image classifications (Gharoun et al. 2024)), however, customer sessions are inherently sequential. To capture this temporal structure, we develop a novel integration of the Transformer architecture within the MetaTP encoder. The Transformer-based encoder captures rich information in the context set, including item features, customer responses, and their orderings, and links this information to customer preferences at a specific future position in a session. Thereby, our use of Transformers to estimate individual response functions differs from their conventional application in next-token prediction (Vaswani et al. 2017; Lu and Kannan 2025). By explicitly modeling sequential patterns, our approach extends meta-learning to uncover dynamic preference heterogeneity from limited data and substantially improves the prediction of customers' future responses.

For the MetaTP decoder, we model it as a structured parametric function that yields parameters with clear economic interpretations while retaining substantial flexibility. The structural decoder produces interpretable, individual-level parameters and, through its interaction with the flexible encoder, preserves predictive accuracy with minimal loss of model expressiveness. Finally, we develop a stochastic variational inference approach to calibrate MetaTP efficiently with large-scale data.

Our main application context is the customer consumption sessions of digital products

on online platforms. Using a large-scale music listening session dataset, we demonstrate that our method can quickly model an individual customer's dynamic song-skipping behavior by learning from their sparse observations. We validate the model's predictive advantage by comparison studies with a rich set of baseline methods, including hierarchical Bayes, supervised learning, and deep sequential models. Given only a few context data points from a new session, MetaTP is able to infer the underlying preference and make more accurate predictions of near-term and long-term customer responses than the baseline methods. We demonstrate that MetaTP can continuously and quickly improve its predictive accuracy by adapting to incoming data of a new session. Using controlled synthetic sessions, we further validate that the Transformer-based encoder accurately captures user preferences from the context set in the learned embeddings, under various music features and response patterns in the context sets.

Next, we demonstrate that the parameters of MetaTP are interpretable and can reveal individual-level preference dynamics. It uncovers the evolution in music tastes and customer fatigue over the course of a listening session, facilitating explainable content offerings. Managerially, we illustrate how the platform can utilize MetaTP to optimize personalized product recommendations with high data efficiency. Specifically, given a few initial interactions, MetaTP can assist digital platforms in one-shot completion of sessions with a minimal expected skip rate. Compared to the music recommended by baseline methods, the products suggested by MetaTP align more closely with customer tastes in essential acoustic features. In addition, we demonstrate how MetaTP can sequentially customize a session from scratch for a new customer and dynamically adjust product offerings based on real-time responses. This capability highlights MetaTP's potential as a data-efficient solution for few-shot customer targeting and personalization.

We provide an additional application in Web Appendix J, focusing on the shopping sessions on consumer packaged goods (CPGs). Across eight independent product categories, we find that MetaTP most accurately predicts the future choices of new households given their initial brand selections compared to benchmark models. It provides interpretable parameters

that reflect the dynamics of own-price elasticity, which are well-aligned with major external events like the Great Recession and closely match the price sensitivity patterns identified in the literature (Dew, Ansari, and Li 2020). We demonstrate how the firms can increase the expected revenue with few-shot data by leveraging MetaTP for personalized dynamic pricing.

In summary, we make three methodological contributions. First, we introduce meta-learning to marketing applications as a novel multi-task framework for handling limited individual-level customer data. We convert customer sessions into meta-learning tasks through data splitting and train the model using a multi-task meta-learning framework. This approach fundamentally departs from standard predictive model training based on i.i.d. data and substantially improves prediction and estimation accuracy when individual data are scarce. Through comprehensive comparisons and ablation studies, we demonstrate that framing customer sessions as few-shot learning tasks is crucial for effectively leveraging minimal context data of new sessions. Second, we propose a novel meta-learning model tailored to sequential customer interactions. By integrating a Transformer architecture into an encoder design, the model adapts meta-learning from its traditional focus on non-sequential data to customer sessions that are inherently sequential. This approach uncovers dynamic preference heterogeneity and improves the prediction of customers' future responses within a session. Third, our model achieves a favorable balance between accuracy and interpretability. A structured decoder delivers economically meaningful individual-level parameters, while a flexible encoder sustains strong predictive performance. As a result, the inferred parameters reveal meaningful patterns of preference evolution and support explainable recommendations.

Our method takes a first step toward studying an underexplored lukewarm stage of the customer journey, characterized by observing a short initial period of customer–platform interactions. This transitional phase lies between the cold-start and warm stages, and may be of particular interest to firms, as customer preferences have begun to be revealed through direct product interactions, but loyalty has yet to be established. We demonstrate that these initial interaction data, though small in scale, contain valuable signals that can be effectively

leveraged to enhance customer engagement and retention.

The rest of the paper is organized as follows. The *Related Literature* section positions our paper in the related marketing research areas. We develop the MetaTP model and describe the inference approach in the *Modeling Approach* section. Next, we apply the model to consumption sessions of digital products on online platforms. The section *Model Comparison* compares it with benchmark methods in few-shot predictions. In *Results* section, we demonstrate the interpretable results from the estimated parameters. The *Managerial Applications* section demonstrates how firms can leverage the model for data-efficient product offerings, followed by a conclusion.

RELATED LITERATURE

Our paper relates to several streams of literature. First, our meta-learning approach connects to and extends the hierarchical Bayes (HB) framework for modeling individual heterogeneity. A classic approach in this literature is the hierarchical logit model (Allenby, Rossi, and McCulloch 2005), which places a shared prior distribution to capture dependencies across individuals. Recent methodological developments enhanced the flexibility of HB models by introducing more flexible hierarchical structures. For example, additive regression trees have been used to model individual-level parameters as dependent variables (Wiemann 2025), and Gaussian Process (GP) priors have been employed to capture time-varying heterogeneity across an individual's actions (Dew, Ansari, and Li 2020). Our proposed meta-learning model generalizes the hierarchical Bayes framework by learning a data-dependent prior over predictive functions, parameterized using flexible neural networks, rather than relying on a prespecified prior distribution (Gordon et al. 2018; Grant et al. 2018). Moreover, our approach improves the scalability relative to MCMC-based hierarchical models (Rossi, McCulloch, and Allenby 1995; Dew, Ansari, and Li 2020). By training across multiple meta-learning tasks, our flexible and scalable framework captures both individual and dynamic heterogeneity,

while being specifically designed for settings with limited individual-level data.

Second, our paper is related to the literature on meta-learning and deep learning (Schmidhuber 1987; Liu 2023; Dew et al. 2024). Existing meta-learning methods are typically applied to non-sequential problems, such as classification and regression from a few examples (Koch et al. 2015; Finn, Abbeel, and Levine 2017; Yin et al. 2020; Gharoun et al. 2024). We adapt meta-learning to customer sessions with a novel encoder-decoder model. It incorporates a Transformer structure (Vaswani et al. 2017) to capture the sequential nature of customer sessions, thereby uncovering preference evolution and enhancing sequential prediction. Recent work has adopted the Transformer model in analyzing customers' text, such as inferring brand interest from online comments (Hartmann et al. 2021) and estimating attribute-specific valence in product reviews (Puranam, Kadiyali, and Narayan 2021). Beyond text data, Lu and Kannan (2025) models customers' sequential actions by treating them as word tokens and applying the Transformer to predict the next actions (token). We extend the use of the Transformer by leveraging its architecture to design a context-set encoder within a meta-learning framework, where it generates session-specific predictive functions rather than next tokens.

Finally, our paper relates to the literature on customer relationship management and customer journeys (Følstad and Kvale 2018). When rich individual-level data are available, Li and Ma (2020) models search phrases along the path to purchase to improve marketing efficiency; Padilla, Ascarza, and Netzer (2024) factorize customer-journey parameters into customer-specific and journey-type components, modeling the latter using nonparametric Bayesian priors. In contrast, when individual data are unavailable for new customers, the cold-start problem has received substantial attention, with existing approaches leveraging external information such as demographics and acquisition data for personalization (Padilla and Ascarza 2021). We complement this line of research by focusing on an underexplored lukewarm stage of the customer journey, where a short initial period of customer responses is observed. We demonstrate that this transitional phase between cold-start and warm stages,

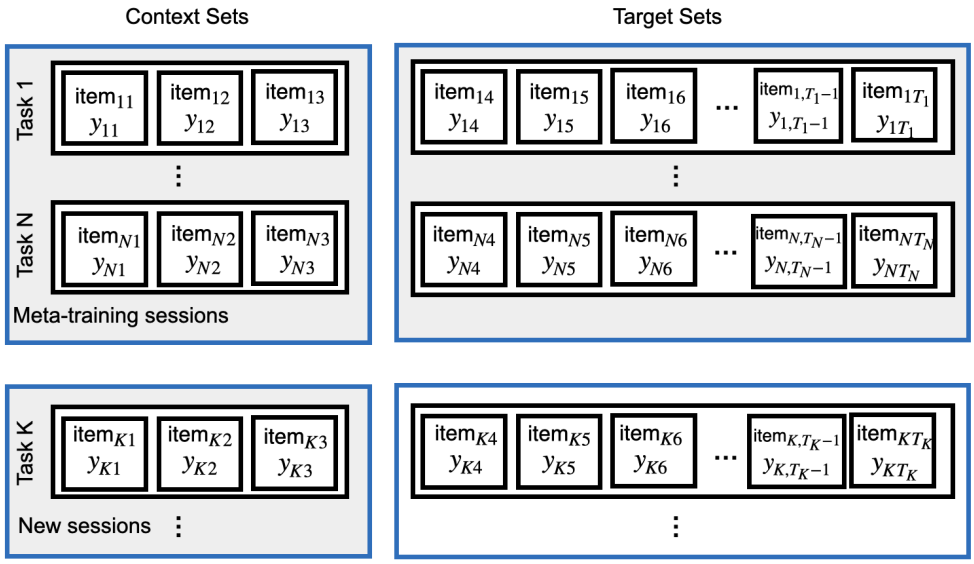


Figure 1: Illustrations of the Meta-learning Framework for Customer Sessions.

despite involving only a small number of observations, provides valuable signals that can be efficiently leveraged to enhance customer engagement and retention.

MODELING APPROACH

We first introduce how to convert consumer sessions into meta-training tasks for few-shot prediction. We then present the probabilistic model of MetaTP and the inference procedure, followed by extending the Transformer structure in meta-learning to capture sequential patterns and individual dynamic heterogeneity.

Data-efficient Learning with Meta-training Tasks

Traditional ML methods often face data efficiency issues. For example, experienced sales experts may infer a customer's needs from just a few observations, whereas conventional machine learning models typically require thousands of data points to achieve good predictive performance. Intuitively, sales experts achieve this data efficiency by repeatedly observing customer behavior and forming predictions based on the experiences of many past customers. Over time, they transfer this fast-learning ability to new customers. Meta-learning formalizes

this intuition by training models across many related learning problems, referred to as meta-training *tasks* that share common structure. Through this multi-task training, the model learns how to learn efficiently from limited data when encountering a new, similar task.

Consider a firm with a potentially large pool of data from N existing customer sessions. Each session i consists of a sequence of data $\mathcal{D}_i = (\boldsymbol{w}_{i1}, \boldsymbol{w}_{i2}, \cdots, \boldsymbol{w}_{iT_i})$, where $i \in \{1, 2, \cdots, N\}$, T_i is the session length, and each $\boldsymbol{w}_{it} = (\boldsymbol{x}_{it}, y_{it})$ is a pair of observed item features $\boldsymbol{x}_{it} \in \mathbb{R}^P$ and a customer response $y_{it} \in \mathbb{N}$. The data are longitudinal and follow a sequential pattern indexed by t, which denotes the position in the sequence of interactions. For example, in a listening session to music tracks, $\boldsymbol{x}_{it}$ could be a vector of acoustic features of a song and y_{it} is the listener response to the t-th song in session i; In a shopping session, $\boldsymbol{x}_{it}$ and y_{it} could be marketing mix variables of competing products and customers' brand choices.

For a new session k ($k \geq N + 1$), a small set of data $\mathcal{D}_k = \boldsymbol{w}_{k,1:t_k}, \boldsymbol{w}_{kt} = (\boldsymbol{x}_{kt}, y_{kt})$ is observed as the context. The context data represents a number of t_k initial interactions from this new session. Different from standard i.i.d. predictions, the data distribution of a new session $p_k(\boldsymbol{w}_{k,1:T_k})$ may differ from those distributions $p_i(\boldsymbol{w}_{i,1:T_i})$ of existing sessions. For example, in listening sessions, the distributions p_k and p_i are determined by the musical tastes expressed in sessions k and i, which may differ because of individual heterogeneity if the sessions come from different listeners, or because of contextual factors, such as time of day or concurrent activities, when the sessions come from the same listener. In this challenging distribution shift scenario, quickly inferring and predicting behaviors in a new session hinges on effective adaptation to the limited context data.

To achieve this data efficiency in a new session, we adopt the meta-learning scheme (Finn, Abbeel, and Levine 2017; Gordon et al. 2018; Beck et al. 2023) and reformulate each session's data into a task for meta-training. We split data $\mathcal{D}_i = \{(\boldsymbol{x}_{it}, y_{it})\}_{t=1}^{T_i}$ of an existing session i into a *context* set $\mathcal{D}_i^c = \{(\boldsymbol{x}_{it}, y_{it})\}_{t=1}^{t_i}$ and a *target* set $\mathcal{D}_i^q = \{(\boldsymbol{x}_{it}, y_{it})\}_{t=t_i+1}^{T_i}$. The context set size is kept small with $t_i \ll T_i$. The goal of a task is to predict the response y_{it} to the product $\boldsymbol{x}_{it}$ in the target set ($\mathcal{D}_i^q$) by learning the response patterns from the context set

($\mathcal{D}_i^c$). Therefore, each task acts as a supervised learning problem by its own, while the context set and target set play the roles of the training data and the validation data, respectively. By learning to solve a large number of tasks during meta-training, the model captures the commonalities across these tasks, which enables it to infer and predict a new session quickly with a small context set.

Fig. 1 illustrates this episodic training framework over multiple meta-training tasks. Each task is represented as a row in the figure, where the context set comprises the initial interactions, and the target set contains the remaining interactions. A meta model first adapts to the context set and produces a task-specific model, which then predicts on the target set. A training task i is "solved" if the model achieves a high likelihood $p_i(y_{it} \mid \boldsymbol{x}_{it}, \mathcal{D}_i^c, t)$ for the data $(\boldsymbol{x}_{it}, y_{it})$ in the target set given the context set $\mathcal{D}_i^c$ and the position t. By meta-training over N tasks, each represents a few-shot prediction problem, the model learns to capitalize on the limited context data to optimize its predictive performance. Unlike typical supervised learning, the outcome of meta-training is not a specific predictive function, instead, it produces a function that can adapt to a new session's context set $\mathcal{D}_k^c$ to generate a new predictive distribution $p_k(y_{kt} \mid \boldsymbol{x}_{kt}, \mathcal{D}_k^c, t)$ (i.e., learning to learn (Finn, Abbeel, and Levine 2017)). As we will see in the application, training over multiple tasks that consist of context and target sets is essential for achieving accurate predictions in new sessions with limited observations.

Having outlined the meta-training schema for customer sessions, we next describe our probabilistic model that is compatible with this schema and how we adjust meta-learning for customer sessions by incorporating Transformer structures.

Meta Temporal Processes

We propose Meta-Temporal Processes (MetaTP), a novel probabilistic model that leverages session-based meta-training tasks to learn efficiently from limited individual-level data. The model incorporates two novel design elements to address key marketing desiderata. First,

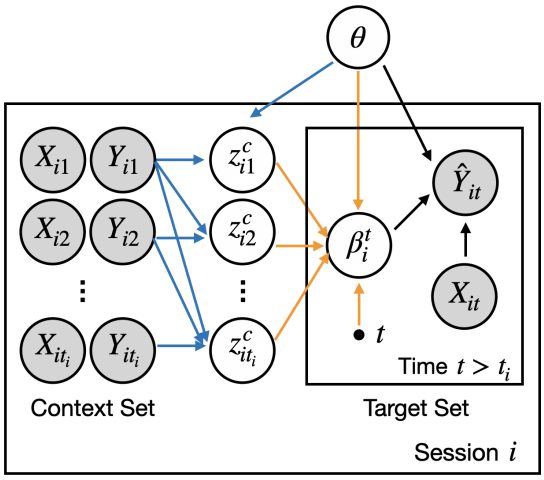


Figure 2: Graphical Model of MetaTP.

Notes. The blue and orange arrows are self- and cross-attention of the encoder, and the black arrows are of the decoder.

MetaTP adopts an encoder–decoder architecture in which the encoder is implemented using flexible neural networks, while the decoder is specified as a structural parametric model. This design yields interpretable parameters that capture individual-level dynamic heterogeneity, while achieving high predictive accuracy, as we later demonstrate. Second, MetaTP integrates a Transformer-based structure to capture sequential patterns within customer sessions, thereby extending meta-learning methods beyond their typical application to exchangeable data.

As an overview, Fig. 2 presents the graphical model of our proposed approach. The model includes a set of global parameters, denoted by $\boldsymbol{\theta}$, which capture statistical structure shared across tasks. As we describe in the next section, $\boldsymbol{\theta}$ consists of high-dimensional parameters associated with Transformer-based neural networks. This high dimensionality provides the flexibility needed for the model to automatically learn inductive biases for few-shot prediction, without requiring them to be manually specified through prior distributions. At the session level, we introduce task-specific latent variables $\boldsymbol{\beta}_i^t \in \mathbb{R}^P$, which are derived from each session's context set ($\mathcal{D}_i^c$) through its learned embeddings ($\boldsymbol{z}_{it}^c$). We next describe the encoder and decoder components of the model.

Flexible Encoder. We use a probabilistic encoder and decoder structure to specify the model

(Kingma and Welling 2013; Burnap, Hauser, and Timoshenko 2023; Dew, Ansari, and Toubia 2022). MetaTP is trained on multiple meta-training tasks, each of which involves predicting outcomes in a target set using information from a given context set. From a Bayesian view, this objective is captured by the posterior predictive distribution $p(y_{it} \mid \boldsymbol{x}_{it}, \mathcal{D}_i^c, t)$ for $(\boldsymbol{x}_{it}, y_{it})$ in the target set $\mathcal{D}_i^q$ of session i. The posterior predictive distribution decomposes as

$$p(y_{it} \mid \boldsymbol{x}_{it}, \mathcal{D}_i^c, t) = \int p(y_{it} \mid \boldsymbol{x}_{it}, \boldsymbol{\beta}_i^t) p(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t) d\boldsymbol{\beta}_i^t, \quad (\boldsymbol{x}_{it}, y_{it}) \in \mathcal{D}_i^q, \tag{1}$$

where we use the latent variable $\boldsymbol{\beta}_i^t \in \mathbb{R}^P$ to encode the context set information that is relevant to predicting target responses. The distribution $p(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t)$ is a posterior distribution of $\boldsymbol{\beta}_i^t$ given the context set $\mathcal{D}_i^c$, while the posterior predictive distribution $p(y_{it} \mid \boldsymbol{x}_{it}, \mathcal{D}_i^c, t)$ marginalizes out the latent variable $\boldsymbol{\beta}_i^t$.

However, Eq. (1) cannot be computed analytically because the posterior is unknown and the integration is intractable for high-dimensional data. This challenge motivates the adoption of variational inference (VI), which jointly estimates the posterior distribution and maximizes the predictive likelihood with high scalability. We introduce a variational encoder distribution $q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t)$ to approximate the posterior of $\boldsymbol{\beta}_i^t$ with parameters $\boldsymbol{\theta}$; its structure is detailed in the following section.

Traditional variational inference maximizes the evidence lower bound (ELBO) of the marginal likelihood (i.e., the evidence $\log p(y_{it})$). In contrast, our objective is to maximize the predictive log-likelihood $\log p(y_{it} \mid \boldsymbol{x}_{it}, \mathcal{D}_i^c, t)$ conditional on data $(\boldsymbol{x}_{it}, \mathcal{D}_i^c, t)$. Following the derivation of variational bounds (Blei, Kucukelbir, and McAuliffe 2017) and the approximation techniques of the neural processes (Garnelo et al. 2018b), we obtain a tractable ELBO of the predictive log-likelihood in Eq. (1) as

$$\mathcal{L}(\boldsymbol{\theta}) = \mathbb{E}_{\beta_i^t \sim q_{\boldsymbol{\theta}}(\beta_i^t \mid \mathcal{D}_i^c, t)}[\log p(y_{it} \mid \boldsymbol{x}_{it}, \boldsymbol{\beta}_i^t)] - \mathrm{KL}(q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \mid \boldsymbol{x}_{it}, y_{it}, \mathcal{D}_i^c, t) || q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t)), \tag{2}$$

with Kullback–Leibler divergence KL$(\cdot||\cdot)$ (see Web Appendix A for a detailed derivation).[1]

For a session i, the variational encoder $q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t)$ infers the customer preference at the position t, given their initial interactions in the context set $\mathcal{D}_i^c = \{(\boldsymbol{x}_{it}, y_{it})\}_{t=1}^{t_i}$. However, the context set contains complicated information about customer preference, as a mix of the high-dimensional product features $\boldsymbol{x}_{it}$, the response y_{it} to the feature $\boldsymbol{x}_{it}$, and the temporal order of $(\boldsymbol{x}_{it}, y_{it})$ pairs. To capture this rich information, we leverage the Transformer structure (Vaswani et al. 2017) to encode the context set and relate it to the target set. Transformer is a state-of-the-art architecture for sequential modeling, known for its success in processing sequences such as natural languages (Devlin et al. 2018).

Our integration of the Transformer differs from traditional applications of the Transformer that autoregressively predict the next token given the previous tokens (Brown et al. 2020; Lu and Kannan 2025). We will illustrate how we construct a Transformer-based encoder and make it compatible with meta-learning in the next section.

Structural Decoder. The decoder $p(y_{it} \mid \boldsymbol{x}_{it}, \boldsymbol{\beta}_i^t)$ represents a local model adapted to a specific session-based task i. It predicts the response y_{it} in a target set given the features $\boldsymbol{x}_{it}$ and the local parameters $\boldsymbol{\beta}_i^t$, where parameters $\boldsymbol{\beta}_i^t \sim q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t)$ summarize the context set. We set the decoder as,

$$p(y_{it} \mid \boldsymbol{x}_{it}, \boldsymbol{\beta}_i^t) = p(y_{it} \mid g^{-1}(\boldsymbol{x}_{it}^\top \boldsymbol{\beta}_i^t)). \tag{3}$$

This is a generalized linear model (GLM) (Nelder and Wedderburn 1972) that incorporates the features through a linear interaction $\boldsymbol{x}_{it}^\top \boldsymbol{\beta}_i^t$ and a nonlinear link function g, where the expected outcome $\mathbb{E}[y_{it} \mid \boldsymbol{x}_{it}, \boldsymbol{\beta}_i^t] = g^{-1}(\boldsymbol{x}_{it}^\top \boldsymbol{\beta}_i^t)$. By selecting an appropriate link function g (e.g., logistic or softmax functions), GLMs can model different types of outcome variables y_{it} (e.g., binary or categorical), and the product $\boldsymbol{x}_{it}^\top \boldsymbol{\beta}_i^t$ represents the underlying systematic utility.[2]

[1] Technically, the posterior $p(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t)$ should be used in the KL term, but as it is not available, we use the variational approximation similar to Garnelo et al. (2018a,b).

[2] MetaTP is a modular encoder–decoder framework that accommodates either linear or nonlinear decoder models for y_{it}. For predictive purposes, model selection can be based on the few-shot

Our use of a GLM decoder enables a clear and direct interpretation of parameters $\boldsymbol{\beta}_i^t$, and is related to the notion of parameter functions in economic modeling (Farrell, Liang, and Misra 2020). As demonstrated in our applications, $\boldsymbol{\beta}_i^t$ may represent the dynamic music tastes of the acoustic features $\boldsymbol{x}_{it}$ in listening sessions, and can uncover price sensitivity and elasticity in shopping sessions. The encoder-decoder structure of MetaTP balances model flexibility and interpretability by employing advanced neural networks for the encoder to ensure flexibility, while using a relatively simple decoder to preserve marketing structure and interpretability. As we will show, this comes with minimal compromise in the model's predictive capacity. MetaTP therefore combines the strengths of traditional parametric structural models and purely predictive ML approaches.

Sequential Modeling with Transformer-based Encoder

We leverage the Transformer structure (Vaswani et al. 2017) to design the MetaTP encoder $q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \,|\, \mathcal{D}_i^c, t)$ that encodes the context set into embeddings, and connects the context and target sets. The core of the Transformer architecture is the attention mechanism, which we use to capture relationships among customer interactions. Technically, attention is a mapping that takes the key, value, and query matrices $\mathbf{K} \in \mathbb{R}^{t \times m_k}$, $\mathbf{V} \in \mathbb{R}^{t \times m_v}$, and $\mathbf{Q} \in \mathbb{R}^{t' \times m_k}$ as the inputs and outputs an embedding for each vector in the query matrix. Here, t is the number of keys and values, and t' is the number of queries; each element in the key and query matrices is dimension m_k, and each element in the value matrix is dimension m_v. In our application, $\mathbf{K}$ and $\mathbf{V}$ are constructed from the session's context set, and $\mathbf{Q}$ is constructed from either the context set or the target set, depending on how we use the attention mechanism. The attention produces the output $\mathbf{S} \in \mathbb{R}^{t' \times m_v}$ by

$$\mathbf{S} = \text{softmax}(\frac{\mathbf{Q}\mathbf{K}^\top}{\sqrt{m_k}})\mathbf{V}, \tag{4}$$

predictive accuracy of new sessions under different context sizes.

where each row of $\mathbf{S}$ is a m_v-dimensional embedding for a query element. The architecture of the MetaTP encoder is visualized in Fig. 3.

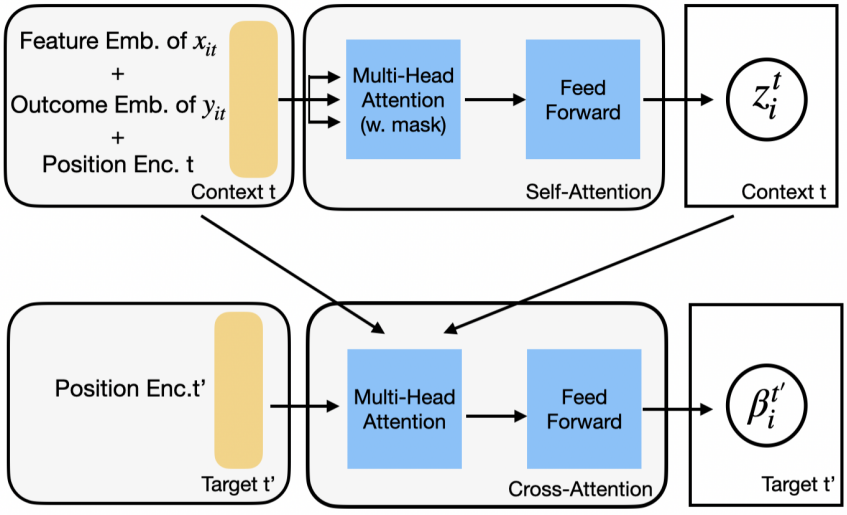


Figure 3: The Architecture of Encoder.

Summarizing context set by self-attention. The first step of the MetaTP encoder is to construct the matrices $\mathbf{K}, \mathbf{V}, \mathbf{Q}$ from the context set $\mathcal{D}_i^c$, similar to the self-attention process in Vaswani et al. (2017). For the context set of session i, it transforms the item features $X_i = [\boldsymbol{x}_{i1}, \cdots, \boldsymbol{x}_{it_i}]^\top$, the responses $Y_i = [y_{i1}, \cdots, y_{it_i}]^\top$, and the positions $1, 2, \cdots, t_i$ to the embeddings as,

$$
\begin{aligned}
&E_i^x = \sigma(X_i W_1 + \boldsymbol{b}_1)W_2 + \boldsymbol{b}_2, \quad E_i^y = Y_i \mathbf{U}, \\
&(E_i^p)_{t,j} = \cos\big(t/10000^{(j-1)/m}\big) \text{ if } j \text{ is even, and } \sin\big(t/10000^{j/m}\big) \text{ if } j \text{ is odd,}
\end{aligned} \tag{5}
$$

where the feature embedding E_i^x is the output of a two-layer transformation with ReLU activation σ, E_i^y is the output of a linear transformation, and $(W_1, W_2, \mathbf{U}, \boldsymbol{b}_1, \boldsymbol{b}_2)$ are the embedding parameters. The sequential information is captured by the positional encoding E_i^p, which follows the sinusoidal encoding from Vaswani et al. (2017). To aggregate the three components, the context set is encoded as $E_i^c = E_i^x + E_i^y + E_i^p \in \mathbb{R}^{t_i \times m}$, which is then mapped to be the $\mathbf{K}, \mathbf{V}, \mathbf{Q}$ matrices for attention mechanism in Eq. (4). In a nutshell, the

output of self-attention is a context set embedding $\boldsymbol{z}_i^c = [\boldsymbol{z}_{i1}^c, \boldsymbol{z}_{i2}^c, \cdots, \boldsymbol{z}_{it_i}^c]$, where each element $\boldsymbol{z}_{it}^c \in \mathbb{R}^{m_v}$ is a vector that summarizes the context data $(\boldsymbol{x}_{it'}, y_{it'})_{t' \leq t}$ before position t.

Connecting to target set by cross-attention. Having obtained the context embedding $\boldsymbol{z}_i^c$, next, we use the attention mechanism to model the relationship between the context set and the target set, similar to the cross-attention mechanism of Transformer.

The cross-attention uses the embedding functions (Eq. (5)) to transform $(\boldsymbol{x}_{it}, y_{it}, t)$ of the context set and $(\boldsymbol{x}_{it'}, y_{it'}, t')$ of the query set into an element in the key and query matrices, and linearly transform the learned context embedding $\boldsymbol{z}_i^c$ into the value matrix. These matrices are mapped to $\tilde{\mathbf{S}}_i \in \mathbb{R}^{t_i' \times m_v}$ by the attention mechanism in Eq. (4) where t_i' is the target set size of session i. Importantly, each row $\tilde{\mathbf{S}}_{i,t}$ of $\tilde{\mathbf{S}}_i$ depends on all the data of the context set $\mathcal{D}_i^c$ and its position t in the session, which captures ordering and dynamic heterogeneity. Finally, we map the embedding $\tilde{\mathbf{S}}_{i,t}$ to the distributional parameters $(\boldsymbol{\mu}_{it}, \boldsymbol{\sigma}_{it})$ by a feed-forward neural network, and set the variational distribution of the task-specific parameter $\boldsymbol{\beta}_i^t$ as $q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t) = \mathcal{N}(\boldsymbol{\mu}_{it}(\tilde{\mathbf{S}}_{i,t}(\mathcal{D}_i^c, t)), \text{diag}(\boldsymbol{\sigma}_{it}^2(\tilde{\mathbf{S}}_{i,t}(\mathcal{D}_i^c, t))))$ where $\boldsymbol{\sigma}_{it} \in \mathbb{R}^P$ and $\text{diag}(\boldsymbol{\sigma}_{it}^2(\tilde{\mathbf{S}}_{i,t}(\mathcal{D}_i^c, t)))$ is a diagonal matrix with $\boldsymbol{\sigma}_{it}^2$ on the diagonal.

We denote all the encoder parameters as $\boldsymbol{\theta}$, including those of the attention and feed-forward neural networks, which are shared across the meta-training tasks. In practice, we utilize multi-head attention to capture the multifaceted relations between the context and target sets, and apply causal masking to prevent future information leakage (Vaswani et al. 2017). The Web Appendix C provides complete details of the encoder structure.

Our encoder design provides two benefits over conventional meta-learning and Transformer models. First, meta-learning is typically applied to exchangeable data (Gharoun et al. 2024). Leveraging the Transformer enables modeling the sequential pattern of customer interactions, which is essential for estimating dynamic heterogeneity and predicting near-term responses (see the *Model Comparison* section). Second, standard Transformers are usually applied to predict the next token $\boldsymbol{x}_{t+1}$ from an existing token sequence $\boldsymbol{x}_{1:t}$ (Vaswani et al. 2017; Lu and Kannan 2025). Unlike next-token prediction, our use of the Transformer generates a

predictive function for each session, which predicts how a customer in session i would respond (y_{it}) to an item $(\boldsymbol{x}_{it})$ at time t.

So far, we have considered sessions separately. When multiple sessions are known to belong to the same customer, our model can leverage these sessions to predict future interactions in the target set. A straightforward way is to concatenate the past sessions with the new session's context set to be an updated context set. This allows cross-attention to relate the target set to both the new session's initial interactions and historical interactions from the same customer, analogous to in-context learning in language models (Brown et al. 2020). Another approach is to leverage the concept of hierarchical attention (Yang et al. 2016), commonly used in modeling documents, sentences, and words in texts, to capture the relationships between customers, sessions, and interactions; we present the details of this approach as a model extension and an interesting future direction in Web Appendix A.

Inference and Few-shot Prediction

The MetaTP model can be inferred by maximizing the ELBO in Eq. (2) with the stochastic gradient ascent. The inference procedures and the full algorithm of MetaTP are described in Web Appendix A.

At the test time, the context data $\mathcal{D}_k$ for a new session k, $k \geq N+1$, consists of the initial interactions with the platform. Note that the size of $\mathcal{D}_k$ can be *arbitrary* since the MetaTP encoder can handle varied sequence length with the attention mechanism[3]. As demonstrated in the empirical study, the predictive accuracy improves quickly when more context data becomes available in a new session.

Using the optimized parameter $\hat{\boldsymbol{\theta}}$, the prediction of future events of a new session k given its context data $\mathcal{D}_k$ can be computed by the Monte Carlo estimate of the predictive distribution in Eq. (1). The prediction of the future behavior $\hat{y}_{kt}$, $t > t_k$ is obtained by taking samples $l = 1, \cdots, L$ from the predictive distribution as $\boldsymbol{\beta}_{kl}^t \sim q_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\beta}_k^t \,|\, \mathcal{D}_k, t)$ and

[3]This is similar to the ability of Transformer in handling different text length.

$\hat{y}_{ktl} \sim p(y_{kt} \mid \boldsymbol{x}_{kt}, \boldsymbol{\beta}_{kl}^t)$. For continuous outcomes, we take the predicted outcome as the conditional expectation $\mathbb{E}_{y_{kt} \mid \boldsymbol{x}_{kt}, \mathcal{D}_k^c}[y_{kt}]$ approximated by $\sum_{l=1}^{L} \hat{y}_{ktl}/L$. For discrete outcomes, we take the prediction as the predictive mode, approximated by the most frequent outcome in the samples $\{\hat{y}_{ktl}\}_{l=1}^{L}$. The adaptation to the new sessions does not require refitting the model on the original large data set $\{\mathcal{D}_i\}_{i=1}^{N}$. Hence, the computation for the new sessions is very efficient, allowing for individual adaptation on the fly. Furthermore, without constraints on the size of the context set $\mathcal{D}_k$, the model can progressively refine its performance as new data arrives throughout the session. This facilitates a seamless and fast transition from the initial cold start stage to a more informed warm stage.

Our proposed model makes several methodology contributions. By reformulating customer session data into training tasks, it introduces a data-efficient approach to learn from limited individual data using meta-learning. Moreover, the model captures the sequential patterns of customer interactions through a novel Transformer integration, which enhances sequential prediction and provides a scalable approach to uncovering customers' dynamic heterogeneity. In addition, the encoder–decoder design yields interpretable individual-level parameters and addresses the tension between model interpretability and flexibility.

Having introduced our modeling framework, we demonstrate its versatility using an empirical application to digital product consumption (e.g., music) on streaming platforms, with an additional application to customers' shopping sessions presented in Web Appendix J. We will illustrate how the model improves few-shot prediction, generates interpretable marketing insights, and assists platforms with product offerings in the following sections.

EMPIRICAL SETTINGS OF THE APPLICATION

We illustrate the value of our approach for managing customer interactions on digital platforms such as Spotify, TikTok, and Amazon. Users' interaction with the digital products on these platforms is often formed as sessions that are taken within a period of time (Wikipedia 2025). New customers are frequently added to these platforms with unknown preferences, and

even for existing users, preferences can differ across sessions due to unrecorded factors like emotion and social environment. Thus, it is beneficial for the platforms to rapidly personalize content based on minimal initial interactions of a new session. We demonstrate the usefulness of our approach in an application involving music listening sessions.

Music consumption is an economically important activity, valued at $29 billion in 2023 (Statista 2024). Reducing the song-skipping rate in listening sessions, as a direct measure of engagement, is economically crucial for streaming services, artists, and consumer welfare. On average, listeners skipped roughly 15 songs per hour in 2018, with skip rates reaching 50% in some segments (Owsinski 2018). Streaming payouts typically range from 0.003 to 0.005 per play (Unchained Music 2024). A back-of-the-envelope calculation suggests that a 1% reduction in skipping could generate an additional 8.9–14.5 million in monthly revenue for platforms and artists. Therefore, quickly understanding consumer preferences is key for platforms to deliver relevant content and support creators in distributing their work.

We use a publicly available listening sessions data set (Brost, Mehrotra, and Jehan 2019) that covers a time span from July 15th, 2018, to September 18th, 2018. Each session contains 10 to 20 songs/tracks that a customer has listened to in sequence. We focus on the customer skipping behavior of each track, which is defined by the platform as whether the track was played below a time threshold. We retain sessions played in shuffled mode to address concerns about song sequencing endogeneity and to enable interpretable preference parameter estimation. We randomly sample $M = 10,000$ sessions, corresponding to 50,704 tracks and 167,880 interactions. This scale allows us to estimate the benchmark methods, especially those with MCMC inference, in a reasonable amount of time for comparative studies. When assessing the scalability of MetaTP, we will analyze it on a larger collection of $M = 100,000$ sessions. The product features $\boldsymbol{x}_{it}$ in this application are the music acoustic features, and the outcome is coded as $y_{it} = 1$ when a track is skipped and $y_{it} = 0$ otherwise. Fig. 4 summarizes the average skipping rate at each track position. We note that the skipping rate in the initial few tracks increases sharply, which highlights the importance of capturing customer

preferences with few-shot data.

For meta-learning, we convert each listening session into a training task. We set the context set size as five based on the observation in Fig. 4, and take the rest of the tracks in the session as the target set. The context size is a model hyperparameter – a small training context size makes the training tasks challenging, hence a model meta-trained on these tasks is more likely to perform well when a new session has a similar scale context set. We will analyze the choice of context set size in the ablation studies. The context set size during the testing can be arbitrary, regardless of what context set size is set in the model training.

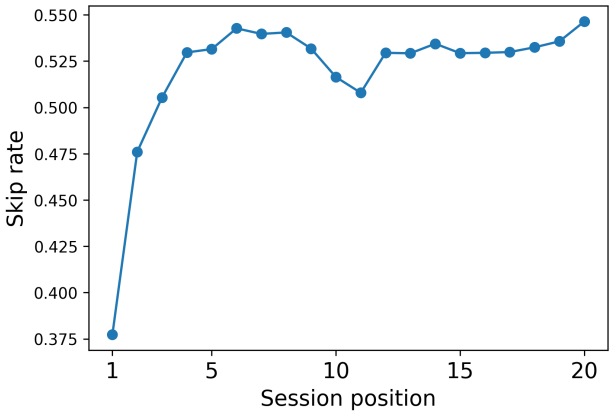


Figure 4: Skipping Rate at Each Session Position.

The data set does not include customer-specific information like demographics; thus, the personalized content streaming relies on capturing the pattern of the interactions in the context set. The acoustic features include Acousticness, Beat strength, Bounciness, Danceability, Energy, Flatness, Key, Mode, Mechanism, Liveness, Loudness, Instrumentalness, Organism, Speechiness, Tempo, and Valence, describing the technical aspects of the music tracks and the felt experience of listeners (Boughanmi and Ansari 2021). Web Appendix B contains the definition, summary statistics, and preprocessing details of these features.

MODEL COMPARISON

We validate the performance of MetaTP against the benchmarks on the few-shot prediction and use the comparison to reveal the effectiveness of MetaTP components. To this end, we randomly select 80% of the sessions indexed from 1 to N for model training, 5% hold-out sessions for model selection, and the remaining sessions indexed from $N+1$ to M for model validation. The number of attention heads and the learning rate were selected by hold-out data, and we found that the model is not particularly sensitive to these hyperparameters. After meta-training over existing listening sessions, we compute the posterior predictive distribution of the future responses for a new session, as described in the *Modeling Approach* section. We predict the binary skip by taking $\hat{y}_{kt} = \mathbf{1}[p(y_{kt} \mid \boldsymbol{x}_{kt}, \mathcal{D}_k) > 0.5]$ with the indicator function $\mathbf{1}[\cdot]$. The implementation details of MetaTP are presented in the Web Appendix E.

Benchmark Models

We systematically compare MetaTP with a range of benchmarks, categorized as Bayesian approaches, sequential models, and supervised learning methods. The comparisons highlight the key properties of MetaTP, such as its few-shot data efficiency, scalability, and ability to generalize to new data distributions.

***Benchmark Family I*: Hierarchical Bayesian Methods.** The Hierarchical Bayes Logit model (HB) is a probabilistic method that offers individual heterogeneity and interpretability. We consider two variants of the HB model as benchmarks. In the traditional HB setting (Allenby, Rossi, and McCulloch 2005), each session i has a local parameter $\boldsymbol{\beta}_i$ with a normal distribution as the shared prior distribution. However, traditional HB model only infers $\boldsymbol{\beta}_i$ for the existing sessions $i = 1, \cdots, N$. To deal with a new session as in our application context, we derive Static HB and Adaptive HB as two HB variants.

- *Adaptive HB* is derived in the Web Appendix D. Essentially, denoting all the observed data of existing sessions as $\mathcal{M} = \{(\boldsymbol{x}_{it}, y_{it})\}_{i=1:N}^{t=1:T_i}$ and parameters shared by sessions

as $\boldsymbol{\theta}$, we derive the posterior predictive distribution based on the d-separation of its graphical model as

$$p(y_{kt} \mid \boldsymbol{x}_{kt}, \mathcal{D}_k, \mathcal{M}) \propto \iint p(y_{kt} \mid \boldsymbol{x}_{kt}, \boldsymbol{\beta}_k) \big[\prod_{(\boldsymbol{x}'_k, y'_k) \in \mathcal{D}_k} p(y'_k \mid \boldsymbol{x}'_k, \boldsymbol{\beta}_k) \big] p(\boldsymbol{\beta}_k \mid \boldsymbol{\theta}) p(\boldsymbol{\theta} \mid \mathcal{M}) d\boldsymbol{\theta} d\boldsymbol{\beta}_k, \quad (6)$$

for a new session k ($k > N$) with context data $\mathcal{D}_k$. Intuitively, the samples of $\boldsymbol{\beta}_k \sim p(\boldsymbol{\beta}_k \mid \boldsymbol{\theta})$ are adapted to a new session k by re-weighting the samples according to their predictive accuracy on the context data $\mathcal{D}_k$. Adaptive HB uses the same set of data as MetaTP, comparing with which highlights MetaTP's scalability and advantage in flexible, sequential modeling.

- *Static HB* is a standard HB model that does not fit new data after model calibration. Its posterior predictive distribution is

$$p(y_{kt} \mid \boldsymbol{x}_{kt}, \mathcal{M}) \propto \iint p(y_{kt} \mid \boldsymbol{x}_{kt}, \boldsymbol{\beta}_k) p(\boldsymbol{\beta}_k \mid \boldsymbol{\theta}) p(\boldsymbol{\theta} \mid \mathcal{M}) d\boldsymbol{\theta} d\boldsymbol{\beta}_k. \quad (7)$$

The details of the prior distributions, the derivation of the posterior predictive distribution are presented in the Web Appendix D. We use Hamiltonian Monte Carlo (HMC), implemented in Stan, to estimate the HB benchmark models. The HB models can also be estimated by variational Bayes (VB) methods for better computational efficiency, but we find the accuracy with VB is inferior to that of HMC; hence, we defer the results from VB estimation to Web Appendix D.

***Benchmark Family II*: Sequential Models.** These are the neural network-based predictive models that capture sequential patterns but are trained over the whole sequence without data splitting, rather than meta-learning tasks.

- *Transformer* is a benchmark that leverages the self-attention introduced in the *Modeling Approach* section. It first takes the preceding sequence $(\boldsymbol{x}_{i,1:t-1}, y_{i,1:t-1})$ of session i and outputs its embedding $\boldsymbol{z}_{it}$, which is then combined with $\boldsymbol{x}_{it}$ and passed through a

Table 1: A Qualitative Comparison of Benchmark Methods.

	Interpretability	Varying Context Size	High Scalability	Individual Heterogeneity	Dynamic Heterogeneity
Static HB	x			x	
Adaptive HB	x	x		x	
Transformer		x	x		x
RNN		x	x		x
SLWC			x		x
MetaTP	x	x	x	x	x

Notes. The list of full benchmarks is in Web Appendix F.

feed-forward neural net to predict y_{it}. The comparison to Transformer highlights the importance of meta-learning across multiple tasks, as the Transformer is not trained on few-shot learning tasks with context and target set splitting.

- *Recurrent Neural Networks (RNNs)* such as Long Short-Term Memory (LSTM) are popular sequential models. RNNs summarize historical track features $\boldsymbol{x}_{i,1:t}$ into a hidden state based on gating functions and then combine with features $\boldsymbol{x}_{i,t+1}$ for the next event y_{t+1} prediction.

***Benchmark Family III*: Supervised Learning Methods.** We consider the supervised learning benchmarks that take the concatenated context set as input to predict skipping.

We design a benchmark named *supervised learning with context (SLWC)* that treats each customer session as one example, and concatenates all the information available to MetaTP as the inputs to a predictive model $f_\theta(\cdot)$. Specifically, the input is $[\boldsymbol{x}_{i1}, \cdots, \boldsymbol{x}_{it_i}, y_{i1}, \cdots, y_{it_i}, \boldsymbol{x}_{it}, \mathbf{1}(1), \cdots, \mathbf{1}(t_i), \mathbf{1}(t)]$ that include the track features and responses in the context set, the feature $\boldsymbol{x}_{it}$ of the current track, and an indicator $\mathbf{1}(t)$ of the positions. A predictive model generates the outcome as $\hat{y}_{it} = f_{\boldsymbol{\theta}}([\boldsymbol{x}_{i1}, \cdots, \boldsymbol{x}_{it_i}, y_{i1}, \cdots, y_{it_i}, \boldsymbol{x}_{it}, \mathbf{1}(1), \cdots, \mathbf{1}(t_i), \mathbf{1}(t)])$. We set the predictive model $f_{\boldsymbol{\theta}}$ as a linear function (*SLWC-L*) or a nonlinear Xgboost function (Chen and Guestrin 2016) (*SLWC-NL*). Though SLWC is a strong predictive benchmark that adopts the data splitting idea of meta-learning, it learns a single mapping $f(\mathcal{D}_i^c, \boldsymbol{x}_{it}, t) \to y_{it}$

for different sessions i. In contrast, MetaTP learns different mappings $f_i(\boldsymbol{x}_{it}) \rightarrow y_{it}$ for each session i with the parameters adapted from the context set. A limitation of SLWC is its fixed context set at the test time and its lack of interpretability, which limits its applicability, as illustrated in our application.

Comparisons with SLWC confirm MetaTP's strong few-shot predictive accuracy. More importantly, it highlights the effectiveness of MetaTP in adapting to the varying sizes of context data, a capability that is essential for both handling new sessions with different starting points and managing a session at different stages.

Table 1 provides a qualitative comparison between these benchmarks and MetaTP in terms of their applicable domains. We will refer to this table in the subsequent sections to select proper benchmarks based on their applicability.

In addition to the comparable methods above, we also include (i) two simple baselines that predict $\hat{y}_{kt}$ by sampling from a Bernoulli distribution with probability as the average skipping rate for the position t in the training sessions (*Simple Global* model), or by sampling from a Bernoulli distribution with probability as the average skipping rate in the context set (*Simple Local* model); (ii) *Sequential Models*: the recurrent neural networks (RNNs) with Gated recurrent units (GRUs); and (iii) *Supervised Learning Methods:* Logistic Regression (LR) and Random Forest (RF) as two more conventional non-sequential predictive models, and their fine-tuning variants. The detailed description, implementation, and evaluation of the benchmark models are presented in the Web Appendix F.

Few-shot Predictions

We evaluate the future skipping prediction on new sessions not seen during model training, given a context set consisting of the initial five interactions. We measure the accuracy by predicting the first song skipping after the context set, and predicting all future skippings in a session, which reflects both the short and long-term predictive ability (Granroth-Wilding and Clark 2016). We use predictive accuracy, recall, precision, and AUC as evaluation metrics

(definitions are provided in Web Appendix E).

Table 2: Overall and First Event Predictive Results.

	First Event				Overall			
	Accuracy	Precision	Recall	AUC	Accuracy	Precision	Recall	AUC
Hierarchical Bayes								
Static HB	54.60%	0.568	0.892	0.527	52.80%	0.546	0.725	0.51
Adaptive HB	66.30%	0.721	0.653	0.725	59.30%	0.631	0.588	0.617
Sequential Models								
RNN-LSTM	54.50%	0.570	0.746	0.538	54.18%	0.560	0.700	0.545
Transformer	63.70%	0.566	0.894	0.593	57.32%	0.620	0.901	0.690
Supervised Models								
SLWC-L	66.55%	0.656	0.841	0.710	59.95%	0.603	0.751	0.620
SLWC-NL	71.95%	0.740	0.765	0.755	60.32%	0.616	0.698	0.628
MetaTP	73.20%	0.756	0.766	0.776	62.86%	0.638	0.708	0.664

Notes. Comparison of *few-shot* predictions between MetaTP and Bayesian methods, sequential methods, and supervised learning approaches. All predictions are for new sessions, given the initial five interactions. See Table F2 in the Web Appendix for a full set of benchmark results.

Table 2 reports the predictive results for the main benchmarks, with additional benchmarks provided in Web Appendix F. MetaTP demonstrates a consistent performance advantage over all alternatives. Fig. 5 further plots predictive accuracy across target positions. As the target position increases, the prediction moves further away from the observed context, and the prediction becomes more challenging. MetaTP maintains higher accuracy than other benchmarks at every target position, highlighting its strength in long-term prediction.

Compared to the HB methods, the predictive lift of MetaTP stems from its episodic training across multiple tasks and its nonlinear, sequential modeling capabilities. These properties of MetaTP lead to a high model flexibility and data efficiency, which is validated through the higher predictive accuracy in new sessions with a small context set. The accuracy gap is larger for the later positions than for the first position, indicating that MetaTP better captures long-term dependency. In the Web Appendix G, we conduct a detailed ablation study on the

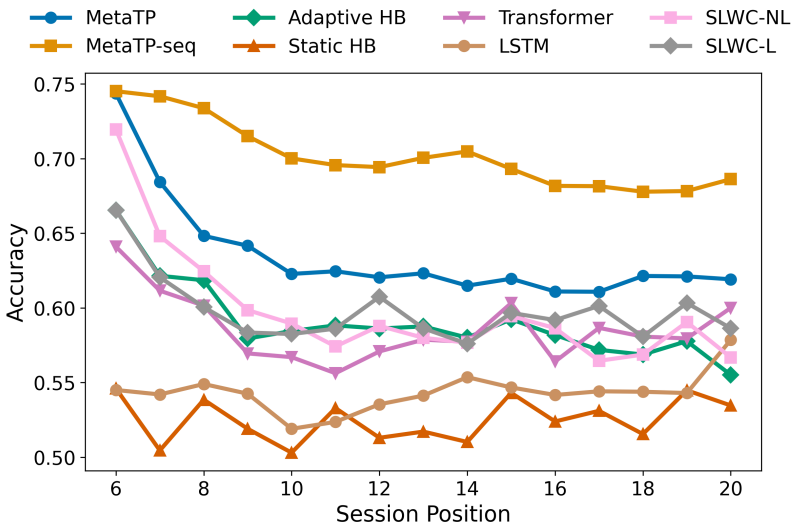


Figure 5: Predictive Accuracy at Each Target Position.

contribution of each model component of MetaTP to the performance gains over Adaptive HB. We find that the meta-learning training scheme yields the largest benefits, as training over multiple tasks with small context sets encourages the efficient use of few-shot data.

An important advantage of MetaTP over HB methods is the scalability. Specifically, generating 1,000 samples using HMC in Stan with 4 parallel chains (250 samples per chain) takes around 50 times longer than MetaTP inference. We find that, although more efficient inference of the HB model with variational Bayes is possible, the accuracy drops significantly, potentially due to the simplified variational distribution and posterior approximation errors. We present the results of the HB model with variational Bayes in Web Appendix F.

To further analyze scalability, with a larger dataset comprising ten times more sessions, we report the detailed computation time of MetaTP for training set sizes ranging from 10,000 to 100,000 sessions in Web Appendix H, where MetaTP's wall-clock time grows mildly with the sample size. This scalability translates directly into improved prediction. Increasing the number of training sessions tenfold raises both the overall and first-event accuracy by about 1%, whereas such scaling would be computationally intensive for the HB baselines.

Compared with sequential models, MetaTP achieves substantially higher accuracy in

predicting new sessions from only a few initial interactions. Since the sequential models are flexible neural networks, this improvement underscores the strength of MetaTP's meta-learning approach. By learning session-specific predictive functions, MetaTP more effectively captures both individual-level and dynamic preference heterogeneity.

Finally, our approach achieves a higher accuracy than the designed SLWC, which is a strong baseline in terms of prediction, as it adopts the data splitting of meta-learning and is oriented towards few-shot prediction by training on small context sets. SLWC, similar to other supervised learning methods, fits a single predictive function. MetaTP, however, learns a function that generates a unique predictive function $p_{\boldsymbol{\beta}_i^t}(y_{it} \mid \boldsymbol{x}_{it})$ for each session i, which better captures the heterogeneous product preference underlying each session. The theoretical analysis in Farrell, Liang, and Misra (2020) suggests that this generating function of functions has more efficient convergence properties with finite data, which may lead to the higher accuracy. Nevertheless, SLWC requires a fixed context-set size, which is often impractical when the context set size varies with a user's starting point or with the session evolution.

To capitalize on the high first-event accuracy, we can also generate a prediction $\hat{y}_{kt}$ by MetaTP using all the preceding new session data, and we label this approach as MetaTP-seq. That is, when predicting the skipping at the position t, we let the context set consist of all the observations from 1 to $t-1$ as $\mathcal{D}_k^c = \{\boldsymbol{x}_{kt'}, y_{kt'}\}_{t'=1:t-1}$. The context set is dynamically updated while the model does not need to be re-estimated. As expected, MetaTP-seq has the highest accuracy at all positions due to the increased context size; we will demonstrate how it can assist firm decisions in the *Managerial Applications* subsection.

Fast Adaptation with Streaming-in Data

Our meta-learning model can quickly adapt to the initial interaction data of a new session and enhance its model fitting as data streams in. In Appendix F, Adaptive HB allows for varying context size and has the highest accuracy; thus, we compare the adaptation ability

on a new context set between MetaTP and Adaptive HB. Although MetaTP is calibrated on training tasks with five context points, the Transformer encoder enables it to handle arbitrary context sizes at test time without requiring a refit of the model.

As shown in Fig. 6, MetaTP and Adaptive HB both benefit from an increased number of context points. However, the improvement of MetaTP is significantly faster than that of Adaptive HB, and the performance gap increases with more context points. It indicates that MetaTP can leverage limited data more efficiently and can learn faster with new data streaming in. This fast adaptation ability is crucial for improving customer satisfaction and reducing the churn rate at the early stage of a customer relationship.

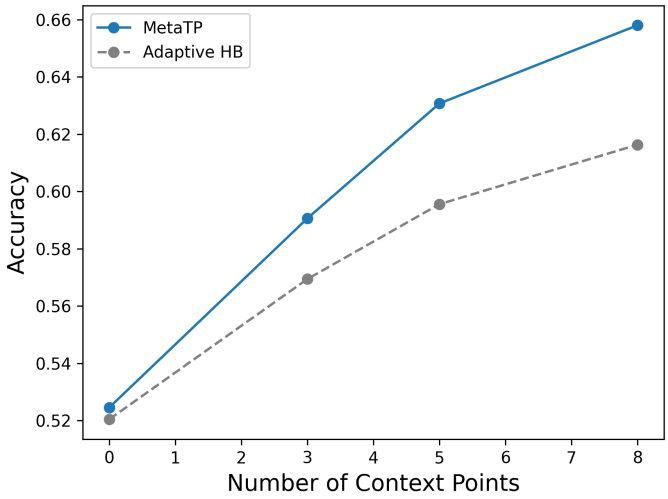


Figure 6: Predictive Accuracy for Different Numbers of Context Points of New Sessions.

Notes. MetaTP quickly improves accuracy with more available new data, seamlessly and efficiently transitioning from cold stage to warm stage.

Ablation Studies and Boundary Conditions

MetaTP models customer sessions by meta-learning across multiple supervised tasks and by using a Transformer to capture sequential patterns. To isolate the contribution of these two components, we construct two ablation variants. The first is the Transformer baseline introduced in the previous section, which captures sequential structure but is not trained

through meta-learning. The second one modifies MetaTP by removing the position encoding, thereby not modeling sequential patterns while retaining the meta-learning framework. The results are shown in Appendix Table G1. We find that meta-learning with tasks provides the greatest gains for few-shot prediction in new sessions. Sequential modeling contributes more to first-event prediction than to overall performance, which is essential for applications such as ordering and closed-loop session completion, as we later demonstrate.

We further explore the source of performance gain over Adaptive HB, such as nonlinearity, meta-learning task construction, and sequential modeling (see Web Appendix G for details). Aligned with the previous study, we find that these components have a greater impact on early position predictions than on later ones, with meta-learning task construction providing the largest benefit for few-shot prediction.

Next, we explore the boundary conditions of MetaTP. Although MetaTP can generalize to new tasks by adapting to context data, the new task cannot be arbitrary and must be related to existing tasks. As the data distribution of a new session p_k becomes increasingly different from the distributions p_i of existing sessions, it becomes more challenging to make accurate predictions. We explore how the accuracy changes over the distribution shifts. The shift of a new session k can be captured by the metric $\min_{1\leq i\leq N} ||\boldsymbol{z}_k^c - \boldsymbol{z}_i^c||_2$, where the $\boldsymbol{z}_k^c$ and $\boldsymbol{z}_{1:N}^c$ are context embeddings by self-attention for the new session and existing sessions, respectively. It measures the similarity between the new session and the most similar existing session. We find that the accuracy of MetaTP and SLWC decreases as the distribution shift increases. However, MetaTP shows the greatest advantage over SLWC in moderately dissimilar cases, those that are challenging but still possible to predict, which validates the improved generalizability of MetaTP to new sessions. The Web Appendix G presents the results and visualizations of the boundary condition study.

Finally, we assess the trade-off between model interpretability and flexibility by replacing the MetaTP linear decoder with a nonlinear feedforward neural network. The resulting improvement in predictive accuracy is modest, with only a 0.2% increase in overall and

Table 3: Synthetic Context Sets for Sessions that Prefer Vibrant (V) and Mellow (M) music.

Context Set 1				Context Set 2			
Track	Artist	Type	Skip	Track	Artist	Type	Skip
DJ Hi-Tek Rulez	Die Antwoord	V	0	Like Home	Trent Reznor and Atticus Ross	M	0
Like Home	Trent Reznor and Atticus Ross	M	1	DJ Hi-Tek Rulez	Die Antwoord	V	1
El Borrach	Los Tucanes De Tijuana	V	0	Miracle	London Philharmonic Orchestra	M	0
Miracle	London Philharmonic Orchestra	M	1	El Borrach	Los Tucanes De Tijuana	V	1
Haters Still Mad	Lil' Flip	V	0	Did You Know It Was Me?	Barbra Streisand	M	0
Did You Know It Was Me?	Barbra Streisand	M	1	Haters Still Mad	Lil' Flip	V	1
Reppin' Uptown	McGruff	V	0	Many Mothers	Junkie XL	M	0
Many Mothers	Junkie XL	M	1	Reppin' Uptown	McGruff	V	1
Act Fore...The End?	The Roots	V	0	Chelsea Bridge	Henry Mancini	M	0
Chelsea Bridge	Henry Mancini	M	1	Act Fore...The End?	The Roots	V	1
Context Set 3				**Context Set 4**			
Track	Artist	Type	Skip	Track	Artist	Type	Skip
DJ Hi-Tek Rulez	Die Antwoord	V	0	Like Home	Trent Reznor and Atticus Ross	M	0
El Borrach	Los Tucanes De Tijuana	V	0	Miracle	London Philharmonic Orchestra	M	0
Haters Still Mad	Lil' Flip	V	0	Did You Know It Was Me?	Barbra Streisand	M	0
Reppin' Uptown	McGruff	V	0	Many Mothers	Junkie XL	M	0
Act Fore...The End?	The Roots	V	0	Chelsea Bridge	Henry Mancini	M	0
Like Home	Trent Reznor and Atticus Ross	M	1	DJ Hi-Tek Rulez	Die Antwoord	V	1
Miracle	London Philharmonic Orchestra	M	1	El Borrach	Los Tucanes De Tijuana	V	1
Did You Know It Was Me?	Barbra Streisand	M	1	Haters Still Mad	Lil' Flip	V	1
Many Mothers	Junkie XL	M	1	Reppin' Uptown	McGruff	V	1
Chelsea Bridge	Henry Mancini	M	1	Act Fore...The End?	The Roots	V	1

Notes. Synthetic context sets with two music preferences and two skipping patterns. Context Sets 1 and 3 represent sessions preferring vibrant music, while Context Sets 2 and 4 represent sessions preferring mellow music.

first-event accuracy. This suggests that MetaTP's combination of a flexible Transformer encoder and an interpretable structural decoder strikes a favorable balance between accuracy and interpretability. While this balance may vary across application domains, the nonlinear decoder remains a viable alternative when predictive performance is the primary objective.

RESULTS

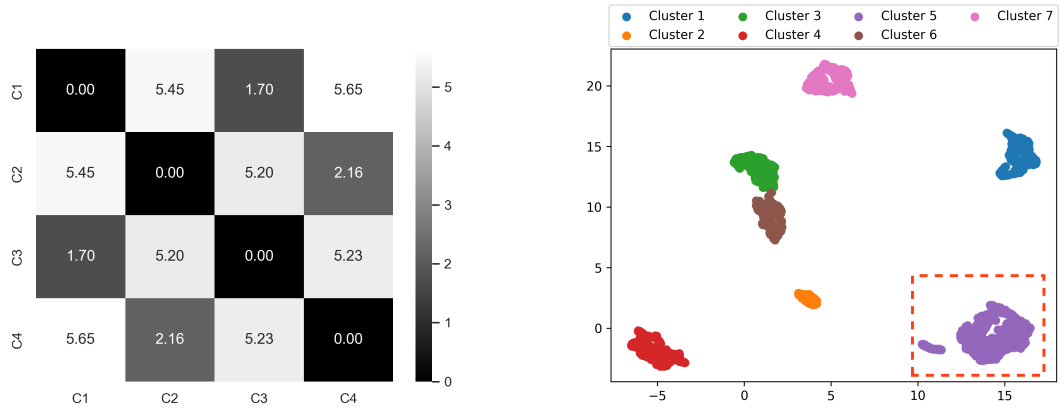
Having validated the performance improvements of MetaTP on the new sessions, we now leverage MetaTP to generate substantive understandings of dynamic preferences using the few-shot individual data. Specifically, we demonstrate that MetaTP, after fitting to the training tasks, can produce meaningful context embeddings and uncover interpretable parameters.

Context Set Embeddings

We demonstrate that the MetaTP encoder can accurately capture customer preferences from the context set and show that customer segments with high future engagement can be identified from the embeddings.

The preference pattern in the context set is reflected as a combination of the high-dimensional track features, the associated skipping behaviors, and the sequence order. MetaTP encodes this pattern as the embedding $\boldsymbol{z}^c_{it_i} \in \mathbb{R}^m$ by self-attention, which depends on the whole context set up to position t_i of the listening session. We create four synthetic context sets that provide the ground truth of music tastes. These context sets are shown in Table 3 with different music preferences and skipping patterns. The set consists of vibrant and mellow music, where vibrant tracks are randomly sampled from songs with low acousticness, high speechiness, and high energy, which are mainly hip-hop and rap music. The mellow tracks are conversely sampled from songs with high acousticness, low speechiness, and low energy, which are mainly orchestral, emotional and cinematic musics. The listeners in Context Sets 1 or 3 skip all the mellow music but do not skip vibrant music, who prefer vibrant and energetic music, while the skipping is reversed for Context Sets 2 and 4 where the mellow and soft music is preferred. Context Sets 1 and 2 (or 3 and 4) have the same skipping sequence but different feature-response pairing; Context Sets 1 and 3 (or 2 and 4) have the same pairing but differ in the sequential order.

Fig. 7a shows the pairwise distances between the embeddings of these context sets. We find that the sessions with similar music preferences are close to each other, i.e., Context Sets 1 and 3 have the smallest distance, followed by the embedding distance between Context Sets 2 and 4. In addition, the embedding distance is influenced by the order of sequences, where the embeddings of the Context Sets 1 and 3 (or 2 and 4) have small but non-zero distance. These findings suggest that the context embedding successfully encodes the information on customer preferences from complex skipping patterns.



(a) Pairwise distance of synthetic context sets **(b)** Clustering

Figure 7: Embedding of the Context Set.

Table 4: Cluster Statistics.

Cluster	1	2	3	4	5	6	7
Size	281	82	249	289	621	212	266
Context set listening rate	1.07%	81.00%	40.08%	100.00%	62.61%	60.43%	20.00%
Target set listening rate	38.90%	50.17%	35.72%	60.65%	59.40%	41.50%	37.53%

Notes. The context set embedding identifies heterogeneous segments of customer sessions. Some segment consists of sessions with high engagement.

This meaningful embedding space can be leveraged to identify the subgroup of sessions with high future engagement. We use Uniform Manifold Approximation and Projection (UMAP) (McInnes et al. 2018; Dew 2023) to reduce the embedding $\boldsymbol{z}^c_{it_i}$ of the new sessions to two dimensions and visualize it in Fig. 7b. The embeddings segment the sessions into 7 well-separated clusters. We characterize these segments using the listening rate of each cluster in the context and target set, defined as the percentage of tracks not being skipped. The values are shown in Table 4. Among the segments, we find Cluster 5 of particular interest, which is the largest cluster and contains 29.4% of total sessions. The average listening rate of this cluster is 62.6% in the context set and is as high as 59.4% in the target set (the second highest among all clusters). This suggests that Cluster 5 represents the Enthusiasts who persistently engage throughout the listening session. In comparison, Top Listeners (Cluster 4) with 100% listening rate in the context set have a future listening rate of 60.7%, only slightly

higher than that of the Enthusiasts. The Top Listener cluster may contain passive listeners whose high listening rate can not reflect their engagement. Moderate Listeners (Cluster 6) have a context-set listening rate similar to the Enthusiasts but have a much lower future listening rate of 41.5%. Comparing Enthusiasts with Top Listeners and Moderate Listeners indicates that the context embedding of MetaTP contains essential predictive information of subsequent behaviors beyond the basic context set listening rate. The platform can use this embedding to identify consumers with high engagement for future tracks.

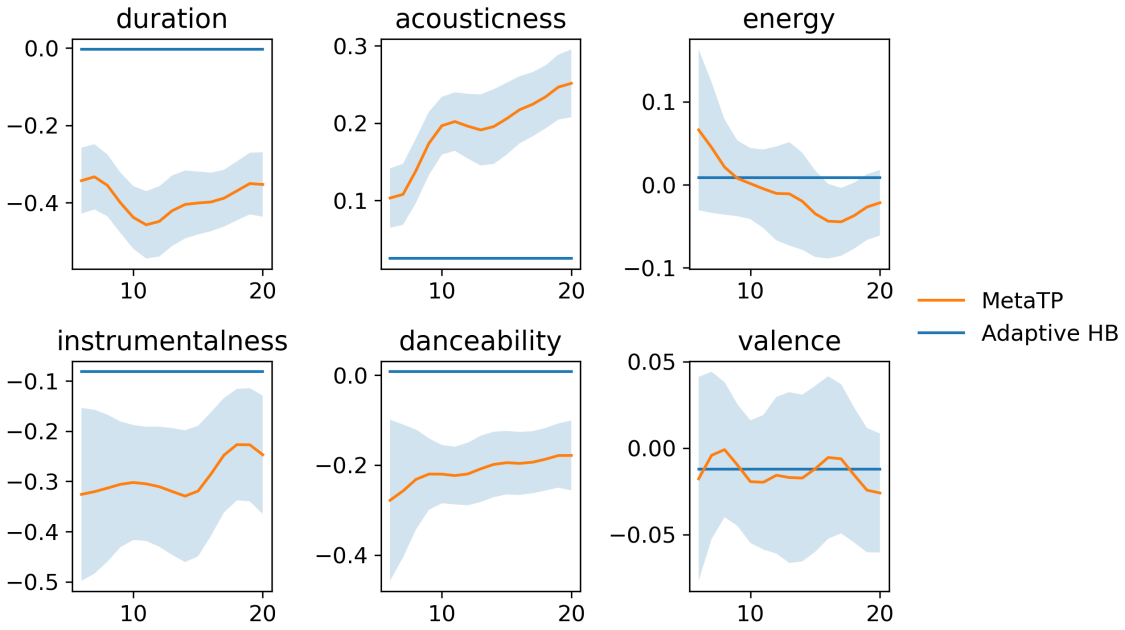


Figure 8: Preference Dynamics Averaged over Sessions.

Notes. The solid line and the shaded region are the mean and standard deviation across sessions; the x-axis is the session position.

Dynamic Preference Heterogeneity

The structured decoder of MetaTP enables the session-level parameter $\boldsymbol{\beta}_i^t$ to be interpreted as a preference measurement. Specifically, for the binary skipping, we use a logit model as the MetaTP decoder with $\mathbb{E}[y_{it} \,|\, \boldsymbol{x}_{it}] = 1/(1+\exp(\boldsymbol{x}_{it} \cdot \boldsymbol{\beta}_i^t))$. Accordingly, $\boldsymbol{\beta}_i^t \in \mathbb{R}^P$ represents the session-specific preference across the music features at the t of session i. An element $\boldsymbol{\beta}_{ip}^t$ quantifies the change in the log-odds of track skipping given a one-unit change in feature x_{itp}.

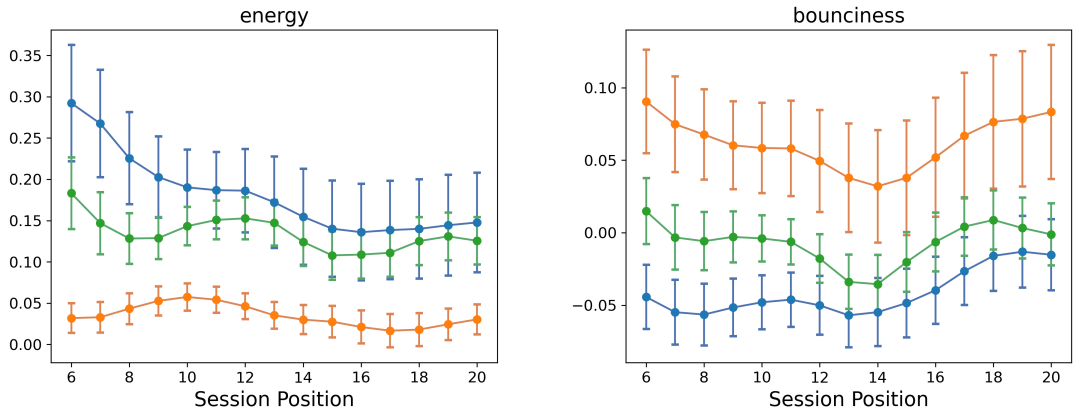


Figure 9: The Session-level Preference Parameters.

Notes. The estimated parameters uncovers rich preference heterogeneity across time and sessions.

Fig. 8 shows how music preferences evolve over time for future tracks in the target set at the population level. The results for all the features are in the Web Appendix I. Since all the processed features are positive and $y_{it} = 1$ is being skipped, the sign of β_{ip}^t shows whether the acoustic feature is positively or negatively linked to non-skipping. A larger β_{ip}^t indicates a higher probability of non-skipping and a stronger preference for the p-th acoustic feature. Among the set of interpretable baseline models (cf. Table 1), we present the estimated parameters of Adaptive HB as a reference, which has the highest predictive accuracy among the interpretable benchmark models.

From Fig. 8, the acousticness of a track is positively related to non-skipping, while the track duration and instrumentalness exhibit a negative correlation. MetaTP parameters reveal interesting dynamic patterns that cannot be captured by the Adaptive HB baseline. Across time, on average, the preference for acousticness (e.g., less electric instruments) and instrumentalness (e.g., less vocal) increases, whereas the preference for energy (e.g., loudness and intensity) decreases. These patterns suggest that continuing a listening session increases fatigue, shifting choices toward mellow, less energetic music, which echoes the customer fatigue patterns identified in the searching sessions (Ursu, Zhang, and Honka 2023). However, the fatigue may not be reflected in the track length, as the tendency to skip long tracks

increases during the early period and then decreases. This suggests that the platform offers lower-energy tracks in the later stage of a session as breaks to reduce the cognitive cost of continuation and increase engagement.[4]

Beyond the aggregated patterns, the estimated $\boldsymbol{\beta}_i^t$ uncovers time-varying preferences for each individual session. Fig. 9 plots the variables $\beta_{ip}^t \sim q_{\boldsymbol{\theta}}(\beta_{ip} \mid \mathcal{D}_i^c, t)$ sampled for three random sessions. On the left panel, the parameter for the energy feature suggests that customers exhibit different fatigue patterns, which can be leveraged by the firm to provide personalized music at specific session positions.

Interpretability enables the platform not only to decide which products to recommend, but also to explain why they are recommended. Such explanations have been shown to foster positive consumer responses (Chen, Tian, and Jiang 2024) and are widely adopted by digital platforms such as Netflix. In Web Appendix I, we illustrate these implications through a case study in which the platform transparently communicates why specific music tracks are recommended, based on the inferred dynamics of a customer's music preferences.

Having shown the substantive insights generated by MetaTP, we now move to showcase its use in managerial tasks.

MANAGERIAL APPLICATIONS

We demonstrate how MetaTP can inform the managerial actions of platforms for customizing and extending new sessions based on few-shot interactions. This scenario is particularly useful for the firms to quickly personalize and maximize their user consumption, which often closely relates to the user loyalty and freemium user conversions (Zhang et al. 2022).

As the tracks are anonymized in our data, we collected an additional 192,000 tracks using the Spotify API, with track names, artist information, and acoustic features, to facilitate an understanding of the targeted product offerings.

[4]In the second application to shopping sessions, Web Appendix J shows that the inferred dynamics of price sensitivity are consistent with external events and are aligned with prior findings in the literature (Dew, Ansari, and Li 2020).

Optimal Product Sequencing

We start with a scenario where customers have a listening list, and the platform aims to provide a listening order for the list based on the initial interactions. MetaTP can optimize the sequential ordering of a list to improve customer engagement without changing the list content. The cross-attention mechanism of the encoder can take the position encodings for each permutation $\pi(\cdot)$ of the future tracks positions $(\pi(t_i + 1), \cdots, (\pi(T_i))$ and compute the skipping probabilities. We select the permutation with the lowest average skipping probability for the tracks in the target set.

An example of the re-ordered tracks for a session from our data set is shown in Table 5. Interestingly, we see that the optimal order tends to place the non-skipped tracks at the early positions. This indicates that placing the preferred products at early positions may generate a positive carry-over effect that influences subsequent customer decisions.

We further examine the effect of personalization on the observed sessions. We randomly select 50 sessions from the data and take the first five interactions of each session as the context set. The observed skipping rate is 58.2% of the future target set. By re-ordering the tracks to their optimal sequence, the predicted skipping rate reduces by 4.9% to 53.3%.

Table 5: Example of the Optimal Ordering of the Target Set.

Original target set	Artist	Skip	Reordered target set
Silver Tiles	Matt and Kim	1	Don't Let Him Steal Your Heart
Dear Avery	The Decemberists	1	October Road
Don't Let Him Steal Your Heart	Phil Collins	0	Ordinary Just Won't Do
Ordinary Just Won't Do	Commissioned	0	Lay It All On Me
October Road	James Taylor	0	Dear Avery
Lay It All On Me	Rudimental	1	Silver Tiles

One-shot Open-loop Session Completion

MetaTP can generate an optimal continuation of a listening session with the tracks from a candidate pool. Specifically, for each session position in the target set, we can estimate the skipping probability for a candidate track using its acoustic features conditional on the context data, and then complete the session using the track with the highest non-skipping probability at that position. This application is important for settings where platforms expect quick customization using a minimal number of customer interactions.

Table 6: Session Completion for Heterogeneous Context Sets.

Session Completion for a Context Set Preferring Vibrant Music				
Track	Artist	Acousticness	Energy	Speechiness
Sucka Ass Niggas	Spice 1	0.006	0.754	0.122
Tits On The Radio	Scissor Sisters	0.012	0.694	0.042
The Cut Song	Slimm Calhoun	0.003	0.673	0.118
Babies Makin' Babies (Rap)	Beats International	0.047	0.871	0.077
Body II Body	Samantha Mumba	0.036	0.912	0.041
Gata Gangster	Daddy Yankee	0.004	0.829	0.131
Smooth (Short)	Wreckx-N-Effect	0.014	0.885	0.183
Act a Donkey	5th Ward Boyz	0.009	0.827	0.267
High On The Boogie	Stargard	0.050	0.785	0.070
This Bitch	Insane Clown Posse	0.015	0.879	0.094
Session Completion for a Context Set Preferring Mellow Music				
Track	Artist	Acousticness	Energy	Speechiness
Naked Moon - Live Version	Pat Metheny Group	0.966	0.228	0.032
Con Te Partiro	Chris Botti	0.965	0.170	0.040
Letter from Home - Live Version	Pat Metheny Group	0.923	0.211	0.056
In The Kingdom of Peace	Jean-Luc Ponty	0.963	0.228	0.035
Calliope	Tom Waits	0.989	0.227	0.034
My Man's Gone Now	Buddy Rich	0.986	0.229	0.050
When It's Sleepy Time Down South	Al Hirt	0.982	0.042	0.037
O Sacred Head Now Wounded	Jim Brickman	0.989	0.152	0.075
Hark! The Herald Trumpets Sing	Mannheim Steamroller	0.957	0.184	0.033
On The Battlefield For My Lord	Brian McKnight	0.988	0.324	0.054

Notes. The next 10 recommended tracks for two sessions. MetaTP accurately adapts to heterogeneous preferences from a context set with only 10 tracks. The top session has Context Set 3 (vibrant music lover), and the bottom session has Context Set 2 (mellow music lover) from Table 3.

Session Completion with Heterogeneous Preferences. We first evaluate how the model completes the individual listening sessions and adapt to heterogeneous needs. Specifically, we take the synthetic Context Sets 2 and 3 from Table 3, which exhibit different preferences for the mellow music and vibrant music.

Table 6 shows the completed session for the two context sets by MetaTP. For ease of interpretation, we show the Acousticness, Energy, and Speechiness for each track. Vibrant music typically exhibits low acousticness, high speechiness, and energy, whereas mellow music shows the opposite pattern. Given only ten interactions as the context set, the recommended tracks by MetaTP clearly match the taste of customers – for example, the tracks for the vibrant music lover are by artists such as Scissor Sisters and Daddy Yankee, who are known for energetic dance-pop and reggaeton music. The tracks for the mellow music lover are by pianists and jazz musicians such as Jim Brickman and Chris Botti, and even the proposed track by Buddy Rich, best known for high-energy drumming and virtuosic solos, is a slow and soothing ballad. In these completions, MetaTP effectively adapts to session heterogeneity, efficiently delivering accurate personalization from the small context sets. Such a fast adaptation mechanism is not hard-coded but emerges from meta-learning across thousands of training tasks designed to infer preferences from few-shot data.

Next, we evaluate the quality of session completion by comparing MetaTP with Adaptive HB. Because future-track recommendation is often required at different stages of a session, we benchmark against Adaptive HB, which has the best performance among those that can adapt to varying context sizes (cf. Table 1). Table H2 in Appendix I shows the tracks recommended by Adaptive HB.

Although MetaTP and Adaptive HB both manage to adjust music energy based on early interactions, MetaTP's completed sessions much more faithfully reflect customer tastes. From the context set (set 3 in Table 3), the customer favors rap and hip-hop music with high speechiness and rhythmic emphasis (e.g., Die Antwoord and Lil' Flip), which matches the music suggested by MetaTP in Table 6 (e.g., Spice 1 and Slimm Calhoun). In contrast,

Adaptive HB offers a heterogeneous mix of rock (e.g., John Fogerty), country (e.g., Tim McGraw), and Caribbean soca (e.g., Kevin Lyttle) music, many of which have low speechiness and emphasize melody over rhythm. A similar pattern appears for the mellow-music listener (set 2 of Table 3), where the listener shows preference to orchestral and cinematic tracks in the first 10 interacted tracks. This subtle taste is accurately captured by MetaTP, reflected in its consistent completion in music style, while Adaptive HB suggests comedy-talk, pop, and folk music that deviate substantially from the user's tastes.

Overall, these comparisons highlight that MetaTP's advantage extends beyond predictive accuracy to the quality of managerial actions it supports. By meta-learning from few-shot data, MetaTP can effectively infer the underlying product preference of a session at the early stage of a customer journey.

In addition, we present the one-shot completion for two context sets from the observed hold-out sessions, one with a low skipping rate and the other with a high skipping rate for low danceability tracks, in Web Appendix I. Similar to the synthetic context sets, we find that MetaTP can complete the sessions by offering personalized tracks with appropriately adapted danceability levels.

Table 7: Joint Session Completion for a Customer Group.

Track	Genre	Valence	Danceability	Acousticness	Prob 1	Prob 2	Prob 3
Jimmy Collins' Wake	Folk	0.96	0.55	0.05	0.71	0.77	0.80
Counting Stars	Pop	0.89	0.63	0.01	0.71	0.76	0.80
Tear Jerky	Indie	0.96	0.64	0.01	0.71	0.76	0.79
Good Times, Cheap Wine	Country	0.94	0.79	0.05	0.71	0.76	0.79
I Wanna Be Your Boyfriend	Indie	0.90	0.45	0.00	0.71	0.76	0.79
So So Long	Folk	0.96	0.63	0.06	0.71	0.76	0.79
Los Vaquetones	Latin	0.96	0.76	0.34	0.71	0.76	0.79
Old Time Rock And Roll	Rock	0.94	0.54	0.10	0.70	0.76	0.81
God Save Rock n Roll	Rock	0.90	0.71	0.01	0.69	0.77	0.81
Hot Summer Night	Pop	0.96	0.68	0.04	0.70	0.77	0.80

Notes. Based on limited interaction data, MetaTP accurately completes the session with music products satisfying the preference of each customer in the group with high non-skipping probability.

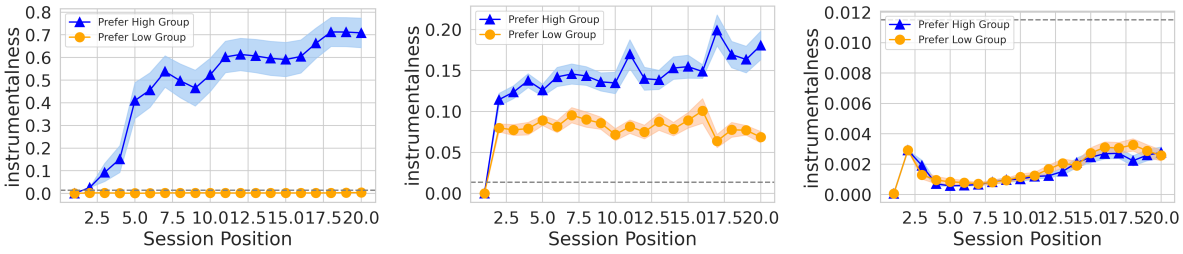
Session Completion for Customer Groups. MetaTP can complete a music session jointly

for a group of listeners. Creating group playlists is popular on streaming platforms, allowing shared music lists for social events. For example, Spotify launched the Collaborative Playlist feature in 2008 and introduced the Jam service to provide personalized and real-time listening sessions for a group in 2023 (Spotify Newsroom 2023). MetaTP can leverage the context sets of group members and complete the session using a sequence of tracks that appeal to all the group members. Each track is selected to maximize the probability that no group member will skip it. Table 7 shows a list of recommended tracks by MetaTP for a group of three individuals from our data with varied context sets: one has few skips for high valence tracks, one has a high skipping rate for low danceability tracks, and another has a low skipping rate for low acousticness tracks. Table 7 contains key acoustic features of the recommended tracks and the non-skipping probability by each group member. The recommended tracks cater to the group's personalized preferences, meaning the tracks are generally happy, danceable, and less acoustic. The non-skipping probability is uniformly high for each group member.

Adaptive Session Completion

Beyond the open-loop completion, we can leverage the MetaTP-seq discussed in the *Model Comparison* section to enable closed-loop session completion by incorporating real-time customer feedback for future product offerings (Chung, Rust, and Wedel 2009). As our data set does not have dynamic feedback for the recommended tracks, we simulate virtual sessions that are dynamically constructed with a set of pre-specified customer response rules. The MetaTP model suggests the next track to be included in the session based on the simulated customer interactions with previous tracks.

Each virtual new customer k is first given one randomly selected track. The customer response y_{k1} to this track and its features $\boldsymbol{x}_{k1}$ is used as the initial context set, which is then leveraged to suggest the next track. At any position t in the sequence, MetaTP takes the data $\{\mathcal{D}_k^c(t) = \{\boldsymbol{x}_{kt'}, y_{kt'}\}_{t'=1}^{t}\}$ as the context set and completes the session at position $t+1$ by a track $\boldsymbol{x}_{k,t+1}$ with the lowest skipping probability. After observing the skipping behavior



(a) MetaTP **(b)** Adaptive HB **(c)** LSTM

Figure 10: Adaptive Session Completion for Two Heterogeneous Groups.

Notes. The dashed horizontal line is the threshold level of instrumentalness that separates the Prefer High and Prefer Low groups. MetaTP achieves accurate customization for new customers from these two groups after adapting to a few interactions.

$y_{k,t+1}$, the context set is updated as $\mathcal{D}_k^c(t+1) = \mathcal{D}_k^c(t) \cup \{(\boldsymbol{x}_{k,t+1}, y_{k,t+1})\}$.

Specifically, we study whether MetaTP can quickly uncover customer preferences when data streams in for a new session. We use preferences pertaining to specific acoustic features to define two groups of synthetic customers. We define a Prefer-High group that prefers a high value of that acoustic feature. Customers in this group are assumed to skip a track with a probability of 0.01 if the acoustic feature of this track is above the third quartile, and otherwise will skip with a probability of 0.99. Similarly, a Prefer-Low group will skip with a probability of 0.99 (or 0.01) if the acoustic feature is below (or above) the third quartile. These hypothesized customers and their preference patterns are likely to differ from those in any sessions that appear in the MetaTP training data; therefore, the model needs to generalize to new sessions with underlying product preference shifts, rather than simply matching the preferences in the training sessions.

In Fig. 10, we display the average acoustic feature level of the recommended tracks for the two customer groups that prefer high and low instrumentalness (see Web Appendix I for a similar study on the acousticness feature). The magnitude of the instrumentalness is averaged over the top 20 tracks of the recommendation list and 50 customers from the two groups to reduce sampling variance. We compare the adaptation ability of MetaTP with Adaptive HB and LSTM, which are comparable methods that can be applied to an increasing

size of the context set.

The figure demonstrates MetaTP's fast customization ability. The two distinctive groups begin with similar initial tracks, but after a small number of interactions, the model automatically figures out the appropriate products to offer. The instrumentalness of the tracks recommended by MetaTP for the Prefer-High group is substantially higher than that for the Prefer-Low group, and in both cases the recommended tracks fall well within each group's preferred instrumentalness range. In contrast, Adaptive HB, while able to differentiate the two groups, still assigns tracks with undesirably high instrumentalness to the Prefer-Low group. LSTM performs less selectively, offering tracks with similarly low instrumentalness to both groups, which fails to meet the preferences of the Prefer-High group.

Table 8: Average Skip Rate with Growing Context Sizes.

	After 1st	After 5th	After 10th	After 15th
Adaptive HB	0.449	0.437	0.428	0.425
LSTM	0.500	0.499	0.502	0.512
MetaTP	0.226	0.190	0.174	0.154

Table 8 shows the average skip rate of future tracks after the customer has interacted with the initial 1, 5, 10, and 15 tracks as the context set. Both MetaTP and Adaptive HB reduce the skip rate as the context set size increases based on their adaptation ability. However, MetaTP can adapt to heterogeneous needs underlying the few-shot interactions more accurately and achieves an over 20% skip rate reduction compared to Adaptive HB.

As an example, Fig. 11 visualizes a sampled session for a customer in the Prefer-High group and a customer in the Prefer-Low group for the acousticness feature. Both customers start from the track The Twelve Days of Christmas, and the subsequent session for the customers from the Prefer-High group mainly consists of tracks like Space Kay, Kitchen Girl, I Ain't Gonna Be The First To Cry with acousticness over 0.8, whereas the session for the Prefer-low customer mainly consists of songs like Crazy Rap and No Hate with heavy

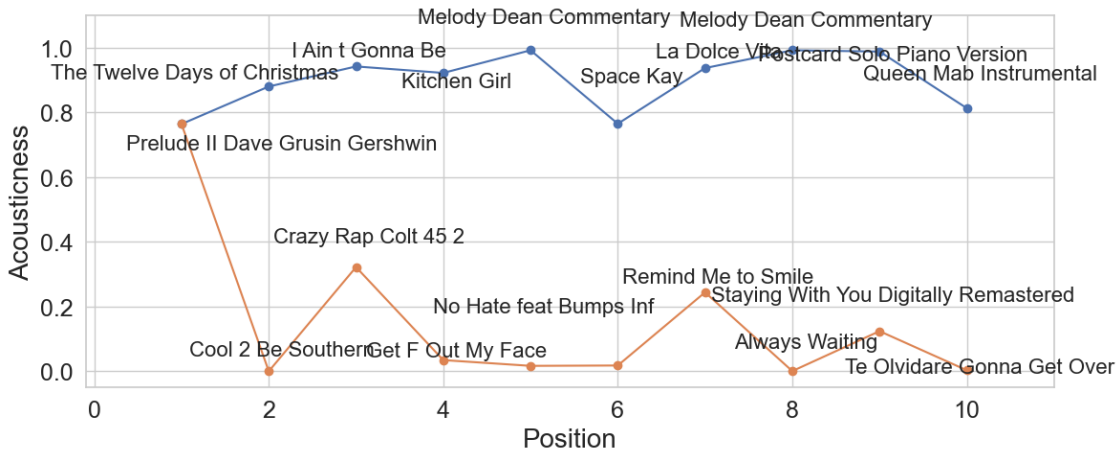


Figure 11: Sampled Sessions for the Customers with High and Low Preferences for the Acousticness Feature.

Notes. The tracks corresponding to the blue and orange curves represent a customer from the Prefer-high group and a customer from the Prefer-low group, respectively.

electronic sounds and acousticness below 0.4.

In summary, this study illustrates that MetaTP can quickly estimate customer preferences from a few interactions, and can continuously and efficiently improve as it transitions from the cold-start stage to the warm stage.

CONCLUSION

In this paper, we developed a novel meta-learning framework for quickly inferring product preferences and predicting future behaviors for new customer sessions. Using a small number of initial interactions, the framework generates accurate future predictions and provides personalized product offerings. The high data efficiency for new sessions is largely obtained from a new meta-training framework over multiple tasks constructed from customer sessions. The key benefits of our encoder-decoder model are its accuracy with limited individual-level data, scalability and interpretability. The Transformer-based encoder captures rich information in the context set and the sequential pattern, while the structural decoder enables interpretation of the inferred individual parameters. Together, the model achieves high few-shot predictive accuracy, uncovers individual dynamic heterogeneity, and stikes a

favorable balance of interpretability and flexibility.

Our proposed approach is versatile and well-suited for a diverse range of marketing applications. We apply it to the consumption of digital music products in online listening sessions as our main application, and also to the choice of CPG products in offline purchasing contexts as our second application in the Web Appendix. We validate the data efficiency and generalization ability of MetaTP via its highly accurate few-shot predictions within new, unseen sessions. Using a large-scale listening session data, we demonstrate that MetaTP can quickly learn from the streaming-in new data on the fly and efficiently transition from a cold-start phase to a warm phase. The interpretable parameters reveal the evolution of customer preferences to acoustic features in the listening session and the customer fatigue patterns. We demonstrate that digital platforms can leverage MetaTP for personalized content offerings. The context embedding accurately captures the underlying music tastes, and the model can use the small context set to suggest future tracks of a new session that are highly aligned with the customer's needs. The model can adaptively complete a session from scratch by iteratively incorporating customer responses on the fly. Beyond the consumption sessions, our proposed meta-learning framework can potentially be extended to new customer acquisitions, where customer preferences need to be quickly inferred from limited initial data, such as from brief onboarding surveys.

This paper makes several direct contributions. Methodologically, we introduce a multi-task meta-learning framework that reformulates customer sessions as few-shot prediction tasks, enabling accurate prediction and estimation when individual-level data are limited. By integrating Transformer architectures into an encoder–decoder design, our approach adapts meta-learning from typical non-sequential settings to sequential customer interactions. Thereby, it uncovers dynamic preference heterogeneity and improves the prediction of future responses. At the same time, the model strikes a favorable balance between flexibility and interpretability, which yields interpretable individual-level parameters and supports explainable recommendations. In terms of application, our work complements the customer

journey literature by examining an underexplored lukewarm stage that bridges cold-start and warm phases, and we demonstrate that effectively leveraging limited early interaction data can significantly enhance customer engagement and retention.

Although our model offers numerous advantages over competing frameworks, it's crucial to acknowledge some limitations of the present approach. First, in the model specification, we adopt a structural decoder model for interpretability. When the predictive accuracy or product recommendation is the primary objective, utilizing more sophisticated decoders could potentially enhance predictive performance. Second, though the Transformer structure is flexible to adapt to a wide range of data, our current applications only deal with structured product features. We leave it as a future direction to analyze customer few-shot interactions with unstructured data, such as images on Instagram and short videos on TikTok. Finally, for model estimation, we currently employ a variational inference approach for the task-level parameter and a point estimate for the global parameters. This inference is efficient and scales to hundreds of thousands of sessions. Providing accurate uncertainty estimation for high-dimensional neural network models is challenging and may substantially increase estimation time. To improve uncertainty estimation, Bayesian Neural Networks, which model the weights of a neural network as random variables, might be a feasible approach. We look forward to exploring these directions as future research.

This work lays the foundation for several promising future research directions. First, current tasks are constructed by splitting customer data at the session level. A promising approach is to construct tasks automatically by detecting distribution shifts and preference changes. Continual learning and change point detection have been applied to meta-learning (Harrison et al. 2020) and can be explored with customer sessions. Our approach provides interpretability through the probabilistic framework. An extension would be to further explore which context points drive the individual's future behavior. This can potentially be enabled by analyzing the attention mechanism and the attention weights. At a more granular level, the influence of product features on customer decisions can be quantified using methods

such as saliency maps (Yang, Zhang, and Zhang 2021). Furthermore, researchers can impose economic structures in the MetaTP decoder to answer substantial policy questions (Donnelly et al. 2021) and to study dynamic sequences, such as price and demand (Gordon 2009; Naik et al. 2015; Misra, Schwartz, and Abernethy 2019).

We provide a flexible and data-efficient approach for modeling new customer sessions. It offers new inference and prediction tools for marketers in two practical yet underexplored scenarios: when the pooled data is rich but individual-level data is scarce, and when customer data distributions at deployment differ from those during model calibration. By establishing connections between meta-learning methods and existing marketing practices, such as hierarchical Bayes and supervised learning, we hope that this research opens up new directions for modeling general customer dynamic patterns from hierarchical, sequential, and few-shot observations.

Disclosure: All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The authors have no funding to report. We use the large language model Claude to verify the grammar accuracy.

REFERENCES

Allenby, Greg, Peter Rossi, and Robert McCulloch (2005), "Hierarchical Bayes models: A practitioners guide," *SSRN*.

Beck, Jacob, Risto Vuorio, Evan Zheran Liu, Zheng Xiong, Luisa Zintgraf, Chelsea Finn, and Shimon Whiteson (2023), "A survey of meta-reinforcement learning," *arXiv preprint arXiv:2301.08028*.

Bernanke, Ben S (2010), "Causes of the recent financial and economic crisis," *Statement before the Financial Crisis Inquiry Commission, Washington, September*, 2.

Blei, David M, Alp Kucukelbir, and Jon D McAuliffe (2017), "Variational inference: A review for statisticians," *Journal of the American statistical Association*, 112 (518), 859–877.

Boughanmi, Khaled and Asim Ansari (2021), "Dynamics of musical success: A machine learning approach for multimedia data fusion," *Journal of Marketing Research*, 58 (6), 1034–1057.

Bronnenberg, Bart J, Michael W Kruger, and Carl F Mela (2008), "Database paper—The IRI marketing data set," *Marketing science*, 27 (4), 745–748.

Brost, Brian, Rishabh Mehrotra, and Tristan Jehan "The music streaming sessions dataset," "WWW," (2019).

Brown, Tom, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell et al. (2020), "Language models are few-shot learners," *NeurIPS*.

Burnap, Alex, John R Hauser, and Artem Timoshenko (2023), "Product aesthetic design: A machine learning augmentation," *Marketing Science*.

Chen, Changdong, Allen Ding Tian, and Ruochen Jiang (2024), "When post hoc explanation knocks: Consumer responses to explainable AI recommendations," *Journal of Interactive Marketing*, 59 (3), 234–250.

Chen, Tianqi and Carlos Guestrin "Xgboost: A scalable tree boosting system," "SIGKDD," (2016).

Chung, Tuck Siong, Roland T Rust, and Michel Wedel (2009), "My mobile music: An adaptive personalization system for digital audio players," *Marketing science*, 28 (1), 52–68.

Cleveland, William S (1979), "Robust locally weighted regression and smoothing scatterplots," *Journal of the American statistical association*, 74 (368), 829–836.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova (2018), "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*.

Dew, R, A Ansari, and Y Li (2020), "Modeling dynamic heterogeneity using Gaussian processes," *JMR, Journal of marketing research*.

Dew, Ryan (2023), "Adaptive Preference Measurement with Unstructured Data," *SSRN*.

Dew, Ryan, Asim Ansari, and Olivier Toubia (2022), "Letting logos speak: Leveraging multiview representation learning for data-driven branding and logo design," *Marketing Science*, 41 (2), 401–425.

Dew, Ryan, Nicolas Padilla, Lan E Luo, Shin Oblander, Asim Ansari, Khaled Boughanmi, Michael Braun, Fred M Feinberg, Jia Liu, Thomas Otter et al. (2024), "Probabilistic Machine Learning: New Frontiers for Modeling Consumers and their Choices," *Available at SSRN 4790799*.

Donnelly, Robert, Francisco JR Ruiz, David Blei, and Susan Athey (2021), "Counterfactual inference for consumer choice across many product categories," *Quantitative Marketing and Economics*, pages 1–39.

Farrell, Max H, Tengyuan Liang, and Sanjog Misra (2020), "Deep learning for individual heterogeneity: An automatic inference framework," *arXiv preprint arXiv:2010.14694*.

Finn, Chelsea, Pieter Abbeel, and Sergey Levine "Model-agnostic meta-learning for fast adaptation of deep networks," "ICML," (2017).

Følstad, Asbjørn and Knut Kvale (2018), "Customer journeys: a systematic literature review," *Journal of service theory and practice*, 28 (2), 196–227.

Garnelo, Marta, Dan Rosenbaum, Christopher Maddison, Tiago Ramalho, David Saxton, Murray Shanahan, Yee Whye Teh, Danilo Rezende, and S M Ali Eslami "Conditional Neural Processes," Jennifer Dy and Andreas Krause, editors, "ICML," (2018a).

Garnelo, Marta, Jonathan Schwarz, Dan Rosenbaum, Fabio Viola, Danilo J Rezende, SM Eslami, and Yee Whye Teh (2018b), "Neural processes," *arXiv preprint arXiv:1807.01622*.

Gharoun, Hassan, Fereshteh Momenifar, Fang Chen, and Amir H Gandomi (2024), "Meta-learning approaches for few-shot learning: A survey of recent advances," *ACM Computing Surveys*, 56 (12), 1–41.

Gordon, Brett R (2009), "A dynamic model of consumer replacement cycles in the PC processor industry," *Marketing Science*, 28 (5), 846–867.

Gordon, Jonathan, John Bronskill, Matthias Bauer, Sebastian Nowozin, and Richard Turner "Meta-Learning Probabilistic Inference for Prediction," "International Conference on Learning Representations," (2018).

Granroth-Wilding, Mark and Stephen Clark "What happens next? event prediction using a compositional neural network model," "Proceedings of the AAAI Conference on Artificial Intelligence," (2016).

Grant, Erin, Chelsea Finn, Sergey Levine, Trevor Darrell, and Thomas Griffiths "Recasting Gradient-Based Meta-Learning as Hierarchical Bayes," "International Conference on Learning Representations," (2018).

Harrison, James, Apoorva Sharma, Chelsea Finn, and Marco Pavone (2020), "Continuous meta-learning without tasks," *Advances in neural information processing systems*, 33, 17571–17581.

Hartmann, Jochen, Mark Heitmann, Christina Schamp, and Oded Netzer (2021), "The power of brand selfies," *Journal of Marketing Research*, 58 (6), 1159–1177.

Kingma, Diederik P and Max Welling (2013), "Auto-encoding variational bayes," *arXiv preprint arXiv:1312.6114*.

Koch, Gregory, Richard Zemel, Ruslan Salakhutdinov et al. "Siamese neural networks for one-shot image recognition," "ICML deep learning workshop," Vol. 2., pages 1–30, Lille (2015).

Li, Hongshuang and Liye Ma (2020), "Charting the path to purchase using topic models," *Journal of Marketing Research*, 57 (6), 1019–1036.

Liu, Xiao (2023), "Deep learning in marketing: A review and research agenda," *Artificial Intelligence in Marketing*, 20, 239–271.

Lu, Zipei and P. K. Kannan (2025), "AI for Customer Journeys: A Transformer Approach," *Journal of Marketing Research*.

McInnes, Leland, John Healy, Nathaniel Saul, and Lukas Großberger (2018), "UMAP: Uniform Manifold Approximation and Projection," *Journal of Open Source Software*, 3 (29), 861.

Misra, Kanishka, Eric M Schwartz, and Jacob Abernethy (2019), "Dynamic online pricing with incomplete information using multiarmed bandit experiments," *Marketing Science*, 38 (2), 226–252.

Naik, Prasad A et al. (2015), "Marketing dynamics: A primer on estimation and control," *Foundations and Trends® in Marketing*, 9 (3), 175–266.

Nelder, John Ashworth and Robert WM Wedderburn (1972), "Generalized linear models," *Journal of the Royal Statistical Society Series A: Statistics in Society*, 135 (3), 370–384.

Nussberger, Anne-Marie, Lan Luo, L Elisa Celis, and Molly J Crockett (2022), "Public attitudes value interpretability but prioritize accuracy in Artificial Intelligence," *Nature communications*, 13 (1), 5821.

Owsinski, Bobby "Song Skip Rates Are Changing The Way Hits Are Made," (2018) https://forbes.com/sites/bobbyowsinski/2018/11/17/song-skip-rates/.

Padilla, Nicolas and Eva Ascarza (2021), "Overcoming the Cold Start Problem of Customer Relationship Management Using a Probabilistic Machine Learning Approach," *Journal of marketing research*.

Padilla, Nicolas, Eva Ascarza, and Oded Netzer (2024), "The customer journey as a source of information," *Quantitative Marketing and Economics*, pages 1–40.

Puranam, Dinesh, Vrinda Kadiyali, and Vishal Narayan (2021), "The impact of increase in minimum wages on consumer perceptions of service: A transformer model of online restaurant reviews," *Marketing Science*.

Rossi, Peter E, Robert E McCulloch, and Greg M Allenby "Hierarchical modelling of consumer heterogeneity: an application to target marketing," "Case Studies in Bayesian Statistics, Volume II," pages 323–349, Springer (1995).

Rossi, Peter E, Robert E McCulloch, and Greg M Allenby (1996), "The value of purchase history data in target marketing," *Marketing Science*, 15 (4), 321–340.

Schmidhuber, Jürgen *Evolutionary principles in self-referential learning, or on learning how to learn: the meta-meta-... hook* Ph.D. thesis, Technische Universität München (1987).

Smith, Adam N, Stephan Seiler, and Ishant Aggarwal (2023), "Optimal price targeting," *Marketing Science*, 42 (3), 476–499.

Spotify Newsroom "Spotify Jam: Personalized, Real-Time Listening Sessions for Free and Premium Users," (2023) https://newsroom.spotify.com/2023-09-26/spotify-jam-personalized-collaborative-listening-session-free-premium-users/.

Statista "Global Revenue of the Music Industry from 2005 to 2023," https://www.statista.com/statistics/272305/global-revenue-of-the-music-industry/ (2024) Accessed: 17 Nov. 2024.

Unchained Music "Spotify Royalty Calculator: How Much Does Spotify Pay per Stream?," (2024) https://www.unchainedmusic.io/blog-posts/spotify-royalty-calculator.

Ursu, Raluca M, Qianyun Zhang, and Elisabeth Honka (2023), "Search gaps and consumer fatigue," *Marketing Science*.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin (2017), "Attention is all you need," *Advances in neural information processing systems*, 30.

Voigt, Paul and Axel Von dem Bussche (2017), *The EU General Data Protection Regulation (GDPR): A Practical Guide* Cham: Springer International Publishing, 1st edition.

Wiemann, Thomas "Personalization with HART," (2025) Working paper.

Wikipedia "Session (web analytics) — Wikipedia," (2025) https://en.wikipedia.org/wiki/Session_(web_analytics).

Yang, Jeremy, Juanjuan Zhang, and Yuhan Zhang (2021), "Engagement that Sells: Influencer Video Advertising on TikTok," *SSRN*.

Yang, Zichao, Diyi Yang, Chris Dyer, Xiaodong He, Alex Smola, and Eduard Hovy "Hierarchical attention networks for document classification," "NAACL," pages 1480–1489 (2016).

Yin, Mingzhang, George Tucker, Mingyuan Zhou, Sergey Levine, and Chelsea Finn "Meta-learning without memorization," "International Conference on Learning Representations," (2020).

Zhang, Dennis J, Ming Hu, Xiaofei Liu, Yuxiang Wu, and Yong Li (2022), "NetEase cloud music data," *Manufacturing & Service Operations Management*, 24 (1), 275–284.

Web Appendix

These materials have been supplied by the authors to aid in the understanding of their paper. The AMA is sharing these materials at the request of the authors.

CONTENTS

WEB APPENDIX A: MODEL INFERENCE, ALGORITHM, AND EXTENSION

Derivation of the ELBO. As the first step for the inference algorithm, we derive a lower bound for the intractable posterior predictive distribution $p(y_{it} \mid \boldsymbol{x}_{it}, \mathcal{D}_i^c)$ as follows:

$$
\begin{aligned}
&\log p(y_{it} \mid \boldsymbol{x}_{it}, \mathcal{D}_i^c) \\
=&\log \int p(y_{it} \mid \boldsymbol{x}_{it}, \boldsymbol{\beta}_i^t) p(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t) d\boldsymbol{\beta}_i^t \\
=&\log \int p(y_{it} \mid \boldsymbol{x}_{it}, \boldsymbol{\beta}_i^t) \frac{p(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t)}{q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t)} q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t) d\boldsymbol{\beta}_i^t \\
=&\log \mathbb{E}_{\boldsymbol{\beta}_i^t \sim q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t)} \left[p(y_{it} \mid \boldsymbol{x}_{it}, \boldsymbol{\beta}_i^t) \frac{p(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t)}{q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t)} \right] \\
\geq&\mathbb{E}_{\boldsymbol{\beta}_i^t \sim q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t)} [\log p(y_{it} \mid \boldsymbol{x}_{it}, \boldsymbol{\beta}_i^t)] - \mathrm{KL}(q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \mid \boldsymbol{x}_{it}, y_{it}, \mathcal{D}_i^c, t) || p(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t)) \\
\approx&\mathbb{E}_{\boldsymbol{\beta}_i^t \sim q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t)} [\log p(y_{it} \mid \boldsymbol{x}_{it}, \boldsymbol{\beta}_i^t)] - \mathrm{KL}(q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \mid \boldsymbol{x}_{it}, y_{it}, \mathcal{D}_i^c, t) || q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t)) \\
:=&\mathcal{L}(\boldsymbol{\theta}).
\end{aligned}
\tag{8}
$$

The inequality is by Jensen's Inequality. Since the density of the true posterior $p(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t)$ is unknown, we follow Garnelo et al. (2018b) to approximate it with the variational posterior $q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t)$ in the last equation. The KL divergence term measures the difference between two posterior distributions that only differ in one data point $(\boldsymbol{x}_{it}, y_{it})$, so the KL divergence is often small and often dropped during the model training (Garnelo et al. 2018a; Singh et al. 2019).

Inference Objective. The inference with empirical data takes the following steps. We obtain the global model parameters $\boldsymbol{\theta}$ by maximizing $\mathcal{L}(\boldsymbol{\theta})$. The predictive likelihood of data, averaged over all existing sessions, becomes

$$
\mathbb{E}\left[\log p_{\boldsymbol{\theta}}(y_{it} \mid \boldsymbol{x}_{it}, \mathcal{D}_i^c)\right] = \mathbb{E}[\log \int p_{\boldsymbol{\theta}}(y_{it} \mid \boldsymbol{x}_{it}, \boldsymbol{\beta}_i^t) p(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t) d\boldsymbol{\beta}_i^t] \gtrsim \mathbb{E}[\mathcal{L}(\boldsymbol{\theta})] \tag{9}
$$

where the expectation is over uniformly distributed session $i \in \{1, \cdots, N\}$ and target position

$t \in \{t_i + 1, \cdots, T_i\}$. Optimizing the ELBO can be operationalized as repeating the following steps until convergence:

- Randomly sample a subset of training sessions/tasks $\mathcal{A} \subset \{1, 2, \cdots, N\}$;
- For each $i \in \mathcal{A}$ and target point $t \in \{t_i + 1, \cdots, T_i\}$, for $l = 1, \cdots, L$, sample $\boldsymbol{\beta}_{il}^t \sim q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \,|\, \mathcal{D}_i^c, t)$;
- Maximize the predictive likelihood using stochastic gradient ascent with stepsize η

$$\boldsymbol{\theta}^{\text{new}} \leftarrow \boldsymbol{\theta} + \eta \nabla_{\boldsymbol{\theta}} \frac{1}{|\mathcal{A}|} \sum_{i \in \mathcal{A}} \Big\{ \frac{1}{T_i - t_i} \sum_{t=t_i+1}^{T_i} \log \frac{1}{L} \sum_{l=1}^{L} p(y_{it} \,|\, \boldsymbol{x}_{it}, \boldsymbol{\beta}_{il}^t(\boldsymbol{\theta})) - \tag{10}$$

$$\text{KL}(q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \,|\, \boldsymbol{x}_{it}, y_{it}, \mathcal{D}_i^c, t) || q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \,|\, \mathcal{D}_i^c, t)) \Big\}. \tag{11}$$

We randomly sample a mini-batch of sessions $\mathcal{A} \subset \{1, 2, \cdots, N\}$ and compute the stochastic gradient to achieve high computational efficiency (Ansari, Li, and Zhang 2018). The stepsize η can be scheduled by methods such as Adam (Kingma and Ba 2014) and the gradient can be computed by auto-differentiation tools such as PyTorch.

Full Algorithm. Alg. 1 describes the steps of the proposed method MetaTP.

Algorithm 1: Meta-Temporal Processes

Input : Data $\{\mathcal{D}_i^c, \mathcal{D}_i^q\}_{i=1}^N$, encoder $q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \,|\, \mathcal{D}_i^c, t)$, decoder $p(y_{it} \,|\, \boldsymbol{x}_{it}, \boldsymbol{\beta}_i^t)$, batchsize B

Output : Model with parameter $\boldsymbol{\theta}$

Initialize parameter $\boldsymbol{\theta}$ randomly ;

while *not converged* **do**

  Sample random set of customers $\mathcal{A} \subset \{1, 2, \cdots, N\}$ with $|\mathcal{A}| = B$;

  for $i \in \mathcal{A}$ **do**

    Compute context embedding $\boldsymbol{z}_i^c$ by self-attention with $\mathcal{D}_i^c$;

    Compute $\boldsymbol{\mu}_{it}$, $\boldsymbol{\sigma}_{it}^2$ by cross-attention with $\boldsymbol{z}_i^c, \mathcal{D}_i^c, t$;

    Sample local variable $\boldsymbol{\beta}_{il}^t \sim q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \,|\, \boldsymbol{\mu}_{it}(\mathcal{D}_i^c, t), \text{diag}(\boldsymbol{\sigma}_{it}^2(\mathcal{D}_i^c, t)))$, for $l = 1 : L$, $t = t_i + 1 : T_i$;

  end

  Update $\boldsymbol{\theta}$ acoording to Eq. (11) with step size η.

end

Extension to Multiple Sessions Per Customer. When multiple sessions belong to the same customer, we use the index set $\mathcal{N}(i) \subset \{1, 2, \cdots, N\}$ to denote the sessions related to the same customer as the session i.

We can leverage the hierarchical structure to borrow information across these sessions. Specifically, we extend the concept of hierarchical attention (Yang et al. 2016), commonly used in modeling documents, sentences, and words in texts, to capture the relationships between customers, sessions, and interactions. The idea is to aggregate multiple sessions of a customer according to their context set similarity to the focal session i. By self-attention, the whole context set of each session $j \in \mathcal{N}(i)$ is encoded in the embedding $\boldsymbol{z}^c_{jt_j}$, where t_j is the last element in the context set. We then compute the context sets similarity by session-level attention as

$$\alpha_{ij} = \frac{\exp((\boldsymbol{z}^c_{jt_j})^\top \boldsymbol{z}^c_{it_i})}{\sum_{j \in \mathcal{N}(i)} \exp((\boldsymbol{z}^c_{jt_j})^\top \boldsymbol{z}^c_{it_i})}, \tag{12}$$

where the inner product measures the similarity of session j and i, both belonging to the same customer, and $\sum_{j \in \mathcal{N}(i)} \alpha_{ij} = 1$. The final context embedding $\boldsymbol{z}^c_i$ for session i is computed as $\tilde{\boldsymbol{z}}^c_i = \sum_{j \in \mathcal{N}(i)} \alpha_{ij} \boldsymbol{z}^c_j$, which aggregates the embeddings $\boldsymbol{z}^c_j$ according to the similarity weights α_{ij}. The updated $\tilde{\boldsymbol{z}}^c_i$ then replaces $\boldsymbol{z}^c_i$ in Alg. 1.

The aggregation allows session i to share information with other sessions in $\mathcal{N}(i)$ from the same customer. When a customer is new to the platform with a new session i, the embedding $\tilde{\boldsymbol{z}}^c_i$ will depend only on its own context set since $\mathcal{N}(i) = \{i\}$ and $\alpha_{ii} = 1$.

WEB APPENDIX B: DIGITAL PLATFORM DATA SUMMARY

The music session data consists of 16 acoustic fingerprints of the tracks; among them, Mode and Key are coded as one-hot categorical features, and the rest are min-max standardized as numerical features with the summary statistics shown in Table B1. The definition of these acoustic fingerprints is presented in Table B2. The other features include an eight-dimensional acoustic embedding vector provided by Spotify, as well as duration, release year, and US popularity estimate for a track. We preprocess the data by one-hot encoding the categorical features, such as Mode and Key, using dummy variables for the categories. The continuous features, such as Valence, Danceability, and Acousticness, are min-max normalized to be within the range from 0 to 1 to put these on the same scale.

Table B1: Summary statistics of the numerical acoustic fingerprints.

	Acousticness	Beat strength	Bounciness	Danceability	Energy	Flatness	Mechanism
Min	0	0	0	0	0	0	0
Q1	0.03	0.44	0.46	0.57	0.51	0.9	0.45
Mean	0.22	0.55	0.6	0.68	0.63	0.91	0.6
Median	0.12	0.56	0.61	0.7	0.63	0.92	0.64
Q3	0.34	0.67	0.74	0.8	0.76	0.94	0.76
Max	1	1	1	1	1	1	1
	Liveness	Loudness	Instrumentalness	Organism	Speechiness	Tempo	Valence
Min	0	0	0	0	0	0	0
Q1	0.1	0.84	0	0.21	0.05	0.44	0.28
Mean	0.19	0.86	0.03	0.36	0.15	0.56	0.46
Median	0.13	0.87	0	0.32	0.09	0.57	0.44
Q3	0.24	0.89	0	0.49	0.21	0.66	0.63
Max	1	1	1	1	1	1	1

Table B2: Definitions of Acoustic Fingerprints

Acoustic Fingerprints	Definition
Duration	The track duration in minutes.
Popularity	How popular a track is relative to all other tracks.
Acousticness	Degree to which a track is acoustic (e.g., sounds by acoustic guitar, piano, strings).
Beat Strength	Strength of beats in a track.
Bounciness	Rhythmic lift of a track.
Danceability	Track's suitability for dancing.
Energy	Track's intensity and dynamism.
Flatness	Signal's lack of fluctuation.
Mechanism	Mechanical sound presence in music.
Liveness	Likelihood track was recorded in front of a live audience.
Loudness	Track's average decibel level.
Instrumentalness	Track's lack of vocal content.
Organism	Organic sound presence in music.
Speechiness	Spoken word presence in a track.
Tempo	Beats per minute of a track.
Valence	The degree to which a song conveys positive emotions.
Key	Pitch center of a track using standard pitch class notation.
Mode	Scale type (major=1, minor=0) of a track.

WEB APPENDIX C: DETAILS OF THE TRANSFORMER MODEL

We design the encoder $q_{\boldsymbol{\theta}}(\boldsymbol{\beta}_i^t \mid \mathcal{D}_i^c, t)$ using the Transformer model and the attention mechanism therein (Vaswani et al. 2017). Here, we first introduce the general techniques of the attention mechanism and then illustrate how to leverage attention to infer the time-varying individual preference from music sessions. The attention mechanisms are illustrated in Fig. C1, which are components of the Transformer model depicted in Fig. 3.

Attention Mechanism. Attention is a mapping that takes the key, value, and query matrices as the inputs and outputs an embedding for each vector in the query matrix. The key matrix $\mathbf{K} \in \mathbb{R}^{t \times m_k}$ contains m_k-dimensional keys, the value matrix $\mathbf{V} \in \mathbb{R}^{t \times m_v}$ contains m_v-dimensional values, and the query matrix $\mathbf{Q} \in \mathbb{R}^{t' \times m_k}$ contains the m_k-dimensional queries. The number of keys and values is t and the number of queries is t'. The output embeddings S are computed according to Eq. (4) as

$$\mathbf{S} = \text{softmax}(\frac{\mathbf{Q}\mathbf{K}^\top}{\sqrt{m_k}})\mathbf{V},$$

where $\mathbf{S} \in \mathbb{R}^{t' \times m_v}$ contains m_v-dimensional embeddings for each query element. Intuitively, the embedding is a weighted combination of the values in $\mathbf{V}$, and the attention weights $\mathbf{W}^a = \text{softmax}(\frac{\mathbf{Q}\mathbf{K}^\top}{\sqrt{m_k}})$ measure the alignment between the query elements and the keys.

In practice, we use multi-head attention to capture heterogeneous dependency structures between the queries and keys. Consider $H > 1$ heads. Each head $h \in \{1, \cdots, H\}$ has its own query $\mathbf{Q}_h$, key $\mathbf{K}_h$ and value $\mathbf{V}_h$ matrices and outputs $\mathbf{S}_h$ according to Eq. (4). The multi-head attention output combines the outputs of all the attention heads by $\mathbf{S} = [\mathbf{S}_1, \cdots, \mathbf{S}_H]\mathbf{W}^o \in \mathbb{R}^{t' \times m}$ where $\mathbf{W}^o \in \mathbb{R}^{Hm_v \times m}$ is an aggregation matrix. We use the attention mechanism to first generate embeddings for the context set and then for the target set.

Self-attention for context embedding. The context set contains the initial session's track features and skipping behaviors. The context set size of session i is t_i. For each

context data $(\boldsymbol{x}_{it}, y_{it})$, we transform the track features $\boldsymbol{x}_{it}$ into an embedding $E_{it}^x \in \mathbb{R}^m$ with dimension m by a feed-forward network, and transform the skipping y_{it} (one-hot encoded) by an embedding matrix $\mathbf{U}$. Written in a matrix form, the transformations are

$$\begin{aligned} E_i^x &= \sigma(X_i W_1 + \boldsymbol{b}_1) W_2 + \boldsymbol{b}_2, \quad X_i = [\boldsymbol{x}_{i1}, \cdots, \boldsymbol{x}_{it_i}]^\top, \\ E_i^y &= Y_i \mathbf{U}, \quad Y_i = [y_{i1}, \cdots, y_{it_i}]^\top. \end{aligned} \tag{13}$$

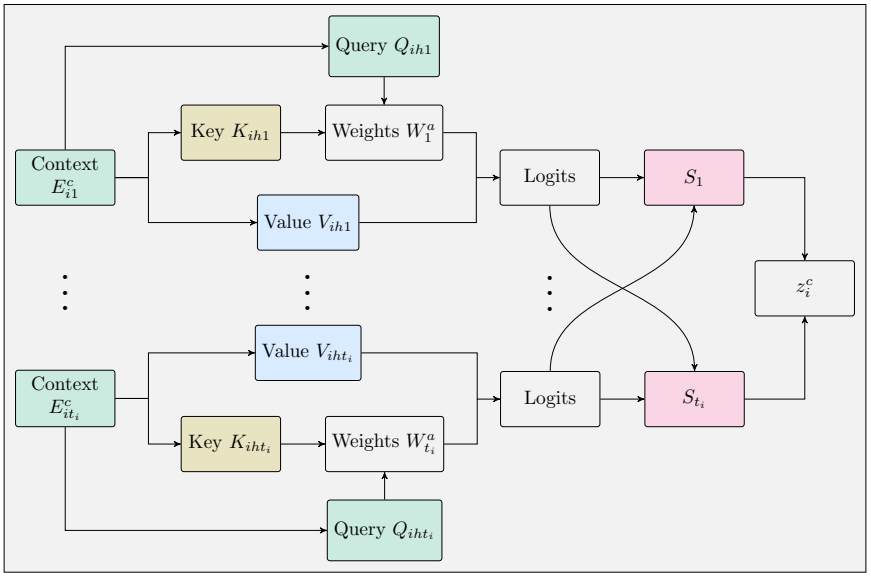
We apply the position encoding in the Transformer models (Vaswani et al. 2017) to define an m-dimensional temporal encoding for each track position in the context set,

$$(E_i^p)_{j,t} = \begin{cases} \cos\left(t/10000^{\frac{j-1}{m}}\right), & \text{if } j \text{ is even} \\ \sin\left(t/10000^{\frac{j}{m}}\right), & \text{if } j \text{ is odd} \end{cases} \tag{14}$$

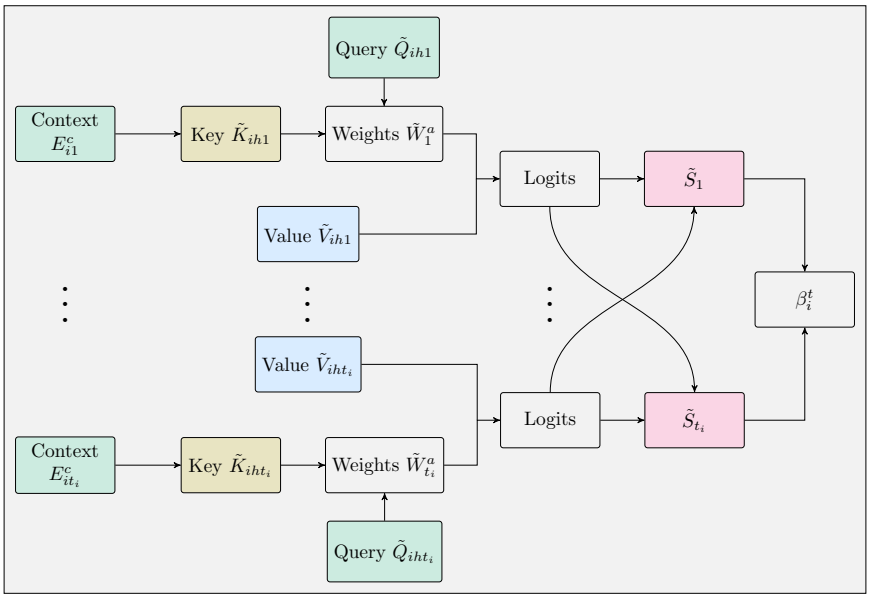
for element $j = 1, \cdots, m$ and the position encoding matrix $E_i^p = [E_i^{p,1}, \cdots, E_i^{p,t_i}]^\top$. The context set of consumer i is encoded as $E_i^c = E_i^x + E_i^y + E_i^p \in \mathbb{R}^{t_i \times m}$. The query, key, and value matrices for each attention head h are computed as $\mathbf{Q}_{ih} = E_i^c \mathbf{W}_h^{\mathbf{Q}} \in \mathbb{R}^{t_i \times m_k}$, $\mathbf{K}_{ih} = E_i^c \mathbf{W}_h^K \in \mathbb{R}^{t_i \times m_k}$, and $\mathbf{V}_{ih} = E_i^c \mathbf{W}_h^{\mathbf{V}} \in \mathbb{R}^{t_i \times m_v}$ with linear transformation weights $\mathbf{W}_h^Q, \mathbf{W}_h^K \in \mathbb{R}^{m \times m_k}$ and $\mathbf{W}_h^V \in \mathbb{R}^{m \times m_v}$. We then apply the attention mechanism, as described in the previous section, to map the inputs $\{\mathbf{Q}_{ih}, \mathbf{K}_{ih}, \mathbf{V}_{ih}\}_{h=1}^H$ to the output matrix $\mathbf{S} \in \mathbb{R}^{t_i \times m}$ where H is the number of attention heads. Finally, this output matrix $\mathbf{S}$ is fed through a position-wise feed-forward neural network $f_{\boldsymbol{\eta}_0}(\cdot)$ with neural network parameters $\boldsymbol{\eta}_0$, generating the embeddings of the context set $\boldsymbol{z}_i^c = f_{\boldsymbol{\eta}_0}(\mathbf{S})$ with $\boldsymbol{z}_i^c = [\boldsymbol{z}_{i1}^c, \cdots, \boldsymbol{z}_{it_i}^c] \in \mathbb{R}^{t_i \times m}$. We apply the causal masking technique by setting the attention weights for future data as zero so that the embedding $\boldsymbol{z}_{it}^c$ at the session position t is independent of future tracks $\{\boldsymbol{x}_{it'}, y_{it'}\}_{t'>t}$[5].

Cross-attention for query embedding. The self-attention captures the dependency

[5]Causal masking in Transformers is a technique used to ensure that during training, a model only has access to present and past data, but not future data. In the context of a sequence, this means each position can only attend to positions before it (or itself), maintaining the causal or sequential order of the data.



(a) Self-attention



(b) Cross-attention

Figure C1: Visualization of Attention Mechanism.

within the context set and encodes it in the embedding $\boldsymbol{z}_i^c$. Next, we apply the cross-attention, the attention mechanism between the context set and the target set, to learn the dynamic pattern of future sessions given the context. The target set size of session i is t_i' The position encoding E_i^t is similarly computed for the target data with $t > t_i$. The query, key, and value matrices are computed as $\tilde{\mathbf{Q}}_{ih} = E_i^t \tilde{\mathbf{W}}_h^{\mathbf{Q}} \in \mathbb{R}^{t_i' \times m_k}$, $\tilde{\mathbf{K}}_{ih} = E_i^c \tilde{\mathbf{W}}_h^K \in \mathbb{R}^{t_i \times m_k}$, and $\tilde{\mathbf{V}}_{ih} = \boldsymbol{z}_i^c \tilde{\mathbf{W}}_h^{\mathbf{V}} \in \mathbb{R}^{t_i \times m_v}$ with weights $\tilde{\mathbf{W}}_h^Q, \tilde{\mathbf{W}}_h^K \in \mathbb{R}^{m \times m_k}$ and $\tilde{\mathbf{W}}_h^V \in \mathbb{R}^{m \times m_v}$. Eventually, the output $\tilde{\mathbf{S}}_i \in \mathbb{R}^{t_i' \times m_v}$ by the attention mechanism in Eq. (4) with inputs $\{\tilde{\mathbf{Q}}_{ih}, \tilde{\mathbf{K}}_{ih}, \tilde{\mathbf{V}}_{ih}\}_{h=1}^H$ is mapped to the mean and variance of the approximate posterior distribution as $\boldsymbol{\beta}_i^t \sim \mathcal{N}(\boldsymbol{\mu}_{\eta_1}(\tilde{\mathbf{S}}_{i,t}), \sigma_{\eta_1}^2(\tilde{\mathbf{S}}_{i,t})\mathbf{I})$ where $\boldsymbol{\eta}_1$ are the parameters for a feed-forward network, where $\tilde{\mathbf{S}}_{i,t}$ is the row of $\tilde{\mathbf{S}}_i$ corresponding to the position t (i.e., the $(t - t_i)$-th row). We use $\boldsymbol{\theta}$ to denote all the parameters shared across the tasks as $\boldsymbol{\theta} = \{\mathbf{W}_1, \mathbf{W}_2, \boldsymbol{b}_1, \boldsymbol{b}_2, U, \{\mathbf{W}_h^Q, \mathbf{W}_h^K, \mathbf{W}_h^V\}_{h=1}^H, \boldsymbol{\eta}_0, \{\tilde{\mathbf{W}}_h^Q, \tilde{\mathbf{W}}_h^K, \tilde{\mathbf{W}}_h^V\}_{h=1}^H, \boldsymbol{\eta}_1\}$.

WEB APPENDIX D: DERIVATION OF ADAPTIVE MIXED LOGIT MODEL

Following the typical hierarchical Bayes (HB) setting, each session i is associated with a set of parameters $\boldsymbol{\beta}_i$ which come from a multivariate normal population distribution $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. We use normal priors for the population mean and a separation-strategy prior (Barnard, McCulloch, and Meng 2000) for the population covariance matrix by decomposing it as $\boldsymbol{\Sigma} = \text{Diag}(\boldsymbol{\tau})\,\mathbf{L}\,\text{Diag}(\boldsymbol{\tau})$, where $\boldsymbol{\tau}$ is a vector of standard deviations, $\text{Diag}(\boldsymbol{\tau})$ is a diagonal matrix with diagonal vector $\boldsymbol{\tau}$, and $\mathbf{L}$ is a correlation matrix. In particular, we use independent Half-Cauchy priors on the standard deviations in $\boldsymbol{\tau}$ and an LKJ prior (Lewandowski, Kurowicka, and Joe 2009) on $\mathbf{L}$ to obtain a flexible prior. The full probabilistic model is

$$\begin{aligned} &y_{it}|\boldsymbol{x}_{it}, \boldsymbol{\beta}_i \sim \text{Cat}(\sigma(\boldsymbol{x}_{it}^\top \boldsymbol{\beta}_i)), \quad \boldsymbol{\beta}_i \sim N(\boldsymbol{\mu}, D(\boldsymbol{\tau})\mathbf{L}D(\boldsymbol{\tau})) \\ &\tau_p \sim \text{HalfCauchy}(0, a), \quad \mathbf{L} \sim \text{LKJ}(b), \quad \mu_p \sim N(0, c^2) \end{aligned} \tag{15}$$

for $i = 1, \cdots, N$ and $t = 1 \cdots, T_i$, where $\sigma(x) = 1/(1 + \exp(-x))$ is the sigmoid function, $\mathbf{I}_p$ is an identity matrix, and a, b, c are hyperparameters. With a slight abuse of notation, we denote $\boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{\tau}, \mathbf{L})$ as the global variables for the HB methods shared by the sessions and $\mathcal{M} = \{(\boldsymbol{x}_{it}, y_{it})\}_{i=1:N}^{t=1:T_i}$ as all the observed data of existing sessions. The global and local latent variables are inferred by the posterior distribution $p(\{\boldsymbol{\beta}_i\}_{i=1:N}, \boldsymbol{\theta} \,|\, \mathcal{M})$, which can be approximated by sampling methods such as Hamiltonian Monte Carlo (HMC).

We use HMC to draw V posterior samples for each latent variable

$$\boldsymbol{\beta}_i^v, \boldsymbol{\tau}^v, \boldsymbol{\mu}^v, \mathbf{L}^v \sim p(\boldsymbol{\beta}_i, \boldsymbol{\tau}, \boldsymbol{\mu}, \mathbf{L} \,|\, \mathcal{D}_1, \cdots, \mathcal{D}_N), \quad v = 1, \cdots, V. \tag{16}$$

Suppose for a new customer $k > N$, we observe some additional new context data $\mathcal{D}_k$. Let all the training data be $\mathcal{M} = \{\mathcal{D}_1, \cdots, \mathcal{D}_N\}$. Then the posterior predictive distribution

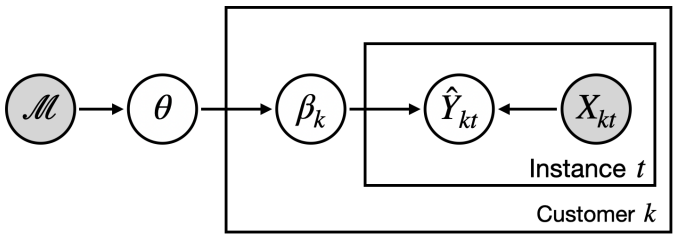


Figure D1: Posterior Prediction of the Adaptive HB Logit Model.

for the data in the target set of the new customer is

$$\begin{aligned} p(y_{kt} \,|\, \boldsymbol{x}_{kt}, \mathcal{D}_k, \mathcal{M}) &= \int p(y_{kt} \,|\, \boldsymbol{x}_{kt}, \boldsymbol{\beta}_k, \mathcal{D}_k, \mathcal{M}) p(\boldsymbol{\beta}_k \,|\, \mathcal{D}_k, \mathcal{M}) d\boldsymbol{\beta}_k \\ &= \int p(y_{kt} \,|\, \boldsymbol{x}_{kt}, \boldsymbol{\beta}_k) p(\mathcal{D}_k \,|\, \boldsymbol{\beta}_k, \mathcal{M}) p(\boldsymbol{\beta}_k \,|\, \mathcal{M}) d\boldsymbol{\beta}_k \cdot \text{Const.} \\ &= \int p(y_{kt} \,|\, \boldsymbol{x}_{kt}, \boldsymbol{\beta}_k) p(\boldsymbol{y}_k^c \,|\, \boldsymbol{x}_k^c, \boldsymbol{\beta}_k, \mathcal{M}) p(\boldsymbol{x}_k^c \,|\, \boldsymbol{\beta}_k, \mathcal{M}) p(\boldsymbol{\beta}_k \,|\, \mathcal{M}) d\boldsymbol{\beta}_k \cdot \text{Const.} \\ &= \iint p(y_{kt} \,|\, \boldsymbol{x}_{kt}, \boldsymbol{\beta}_k) p(\boldsymbol{y}_k^c \,|\, \boldsymbol{x}_k^c, \boldsymbol{\beta}_k) p(\boldsymbol{\beta}_k \,|\, \boldsymbol{\theta}) p(\boldsymbol{\theta} \,|\, \mathcal{M}) d\boldsymbol{\theta} d\boldsymbol{\beta}_k \cdot \text{Const.} \end{aligned} \tag{17}$$

where $\boldsymbol{\theta} = (\boldsymbol{\tau}, \boldsymbol{\mu}, \mathbf{L})$. The derivation is based on the graphical model of inference and prediction for the HB Logit model in Fig. D1 and the corresponding d-separation rules. The first equality is because $\boldsymbol{x}_{kt} \perp\!\!\!\perp \boldsymbol{\beta}_k$ not conditional on y_{kt}; the second equality is by the Bayes rule $p(\boldsymbol{\beta}_k \,|\, \mathcal{D}_k, \mathcal{M}) = p(\mathcal{D}_k \,|\, \mathcal{M}, \boldsymbol{\beta}_k) p(\boldsymbol{\beta}_k \,|\, \mathcal{M}) / p(\mathcal{D}_k \,|\, \mathcal{M})$, and $y_{kt} \perp\!\!\!\perp \mathcal{D}_k, \mathcal{M} \,|\, \boldsymbol{x}_{kt}, \boldsymbol{\beta}_k$, where the multiplicative constant is $1/p(\mathcal{D}_k \,|\, \mathcal{M})$; the third equality is by $\mathcal{D}_k = (\boldsymbol{x}_k^c, \boldsymbol{y}_k^c)$; the last equality is by $\boldsymbol{y}_k^c \perp\!\!\!\perp \mathcal{M} \,|\, \boldsymbol{x}_k^c, \boldsymbol{\beta}_k$ and $\boldsymbol{x}_k^c \perp\!\!\!\perp \boldsymbol{\beta}_k, \mathcal{M}$, where the multiplicative constant is $p(\boldsymbol{x}_k^c)/p(\mathcal{D}_k \,|\, \mathcal{M})$.

We draw V samples $\boldsymbol{\theta}^v \sim p(\boldsymbol{\theta} \,|\, \mathcal{M})$, $\boldsymbol{\beta}_k^v \sim p(\boldsymbol{\beta}_k \,|\, \boldsymbol{\theta}^v)$ and compute $p(y_{kt} \,|\, \boldsymbol{x}_{kt}, \boldsymbol{\beta}_k^v)$, $p(\boldsymbol{y}_k^c \,|\, \boldsymbol{x}_k^c, \boldsymbol{\beta}_k^v)$. Samples from the marginal posterior $p(\boldsymbol{\theta} \,|\, \mathcal{M})$ are distributionally the same as the samples obtained from the full posterior $p(\{\boldsymbol{\beta}_i\}_{i=1:N}, \boldsymbol{\theta} \,|\, \mathcal{M})$ using HMC, so the samples of $\boldsymbol{\theta}$ by Eq. (16) can be directly used. Finally, the Monte-Carlo estimate of Eq. (17) is by

$$\hat{p}(y_{kt} \,|\, \boldsymbol{x}_{kt}, \mathcal{D}_k, \mathcal{M}) = \sum_{v=1}^{V} w_v p(y_{kt} \,|\, \boldsymbol{x}_{kt}, \boldsymbol{\beta}_k^v), \quad w_v = \frac{p(\boldsymbol{y}_k^c \,|\, \boldsymbol{x}_k^c, \boldsymbol{\beta}_k^v)}{\sum_{v'=1}^{V} p(\boldsymbol{y}_k^c \,|\, \boldsymbol{x}_k^c, \boldsymbol{\beta}_k^{v'})}. \tag{18}$$

We take the predicted outcome as the one that has the highest predictive likelihood. Intuitively,

the weight w_v scales according to how accurate each posterior sample $\boldsymbol{\beta}_k^v$ predicts the new individual's context data $\mathcal{D}_k^c$.

The HB methods are implemented with CmdStanPy, a Python interface for Stan (Carpenter et al. 2017). We used four parallel chains to generate posterior samples from the No-U-Turn sampler (Hoffman, Gelman et al. 2014) and checked for convergence using the trace plots of the unknowns, calculated the effective sample size, and monitored the mixing rate.

WEB APPENDIX E: IMPLEMENTATION DETAILS

For the proposed method, we set the dimension of context points initial encoding E_i^c as 64. The dimensions m_k, m_v for the key, query, and value matrices of attention are set as 16, and we use 4 attention heads. The feed-forward neural network has two layers where the latent layer dimension is 128, and the activation function is $\text{ReLU}(x) = \max\{0, x\}$. We use a dropout with a rate of 0.1 for all the neural network parameters during the model training. We use the Adam optimizer (Kingma and Ba 2014) implemented in PyTorch with a learning rate of $1e-4$. The learning rate is decayed by a factor of 0.9 every 20 epochs. As shown in Fig. E1, the algorithm converges in around 2,000 iterations, where each iteration processes 200 sessions. The number of attention heads and the learning rate are selected on the 5% hold-out data; we find the model is not sensitive to these hyperparameters.

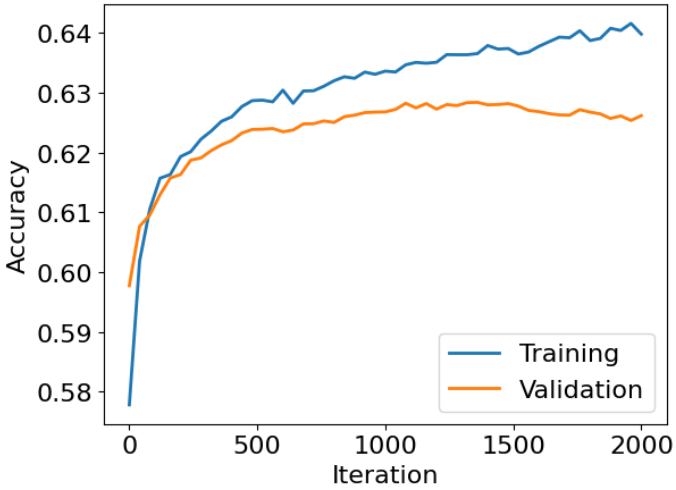


Figure E1: Trace Plot of Model Fitting.

The logistic regression and random forest are implemented using the *LogisticRegression* and *RandomForestClassifier* modules of the Scikit-learn library in Python (Pedregosa et al. 2011). For the FT-LR method, the logistic regression model is first calibrated on all the training sessions, then the pre-trained model is fine-tuned on the new session context data by performing stochastic gradient descent. The method is implemented by the *SGDClassifier*

module and the *partial_fit* function therein.[6] For the FT-RF, a random forest model is pre-trained on the training sessions with the warm_start parameter set to True, which allows fitting additional weak-learners, i.e., the classification trees, to an already fitted model.[7] We set the number of trees for the pre-trained model as 20, and use an additional 2 trees to fit the context data of a new customer. The number of trees is chosen to maximize the predictive accuracy on a hold-out set.

[6] https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html
[7] https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html

WEB APPENDIX F: ADDITIONAL MODEL COMPARISONS

We provide details of the evaluation metrics, additional benchmarks, the ablation study of MetaTP, and the analysis of boundary conditions in this section.

Evaluation Metrics

For the prediction of all future events, we evaluate the methods with metrics defined as

$$\text{Accuracy} = \frac{1}{\sum_{k=N+1}^{M}(T_k - t_k)} \sum_{k=N+1}^{M} \sum_{t=t_k+1}^{T_k} \mathbf{1}[\hat{y}_{kt} = y_{kt}],$$

$$\text{Recall} = \frac{\sum_{k=N+1}^{M} \sum_{t=t_k+1}^{T_k} \mathbf{1}[\hat{y}_{kt} = 1 \text{ and } y_{kt} = 1]}{\sum_{k=N+1}^{M} \sum_{t=t_k+1}^{T_k} \mathbf{1}[y_{kt} = 1]},$$

$$\text{Precision} = \frac{\sum_{k=N+1}^{M} \sum_{t=t_k+1}^{T_k} \mathbf{1}[\hat{y}_{kt} = 1 \text{ and } y_{kt} = 1]}{\sum_{k=N+1}^{M} \sum_{t=t_k+1}^{T_k} \mathbf{1}[\hat{y}_{kt} = 1]},$$

and AUC. AUC is the area under the Receiver Operating Curve (ROC), a widely used metric for evaluating the overall predictive performance (Bishop and Nasrabadi 2006). Similar definitions apply to the first event prediction by fixing $t = t_k + 1$.

Additional Benchmarks

In addition to the benchmarks in Table 2, we further compare MetaTP with Gated recurrent units (GRUs) as a type of recurrent neural network (RNN) for sequential modeling. In a session i, the RNNs compute a hidden state $h_{i,t-1}$ based on the past track features $\boldsymbol{x}_{i,1:t-1}$ up to time $t-1$, and predict the outcome y_{it} based on $h_{i,t-1}$ and $\boldsymbol{x}_{it}$.

We also make a comparison with Logistic Regression (LR) and Random Forest (RF), a linear and a nonlinear model widely used with tabulated data. We further compare with their fine-tuning variants. Fine-tuning is widely used to adapt a model pre-trained on large datasets to specific tasks (Ziegler et al. 2019). We evaluate the Fine-tuned Logistic Regression (FT-LR)

and Fine-tuned Random Forest (FT-RF) that adapt to new sessions. Specifically, FT-LR first calibrates a logistic regression model on all the training sessions, computes a stochastic gradient using the new session context data, and then updates the model parameters by a one-step gradient descent (Pedregosa et al. 2011). FT-RF first fits a random forest model for the training sessions and keeps the fitted model fixed. Then it introduces additional simple tree models to fit the new customer data and combines the fitted forest model and the tree model as an ensemble.

Finally, we notice that a group of sequential models, such as the session-based recommendation systems (Hidasi et al. 2015) and the GPT language models (Brown et al. 2020), focus on predicting the next item (e.g., the clicks and word tokens) given the past history. These models decompose a sequence distribution $p(\boldsymbol{x}_{i1}, \cdots, \boldsymbol{x}_{iT}) = p(\boldsymbol{x}_{i1}) \prod_{t=2}^{T} p(\boldsymbol{x}_{it} \,|\, \boldsymbol{x}_{i,1:t-1})$ and focus on fitting $p(\boldsymbol{x}_{it} \,|\, \boldsymbol{x}_{i,1:t-1})$. Since our setting is to learn a meta model that generates function mappings from $\boldsymbol{x}_{it}$ to y_{it} rather than from $\boldsymbol{x}_{i,1:t-1}$ to $\boldsymbol{x}_{it}$, we do not compare with this line of methods.

Table F1 and Table F2 present a qualitative and quantitative comparison of all the benchmarks and MetaTP, respectively. The first-event and overall accuracy for the simple global baseline and local baseline described in the *Application to Digital Streaming Platforms* section of the main paper are $(56.1\%, 60.9\%)$ and $(50.4\%, 50.2\%)$, respectively. LR and RF have poor performance because they do not adapt to new sessions. Among the adaptive methods, FT-LR and FT-RF are based on fine-tuning heuristics, which underperform compared to Adaptive HB, SLWC, and MetaTP.

Table F1: A Qualitative Comparison of All Benchmark Methods.

	Interpretability	Varying Context Size	High Scalability	Dynamic Heterogeneity
Static HB	×			
Adaptive HB	×	×		
SLWC			×	×
Transformer		×	×	×
RNN		×	×	×
MetaTP	×	×	×	×
Static SL	×		×	
Finetuned SL	×	×	×	

Notes. HB = Hierarchical Bayes and SL = supervised learning. The rows above the dashed line are in the main paper.

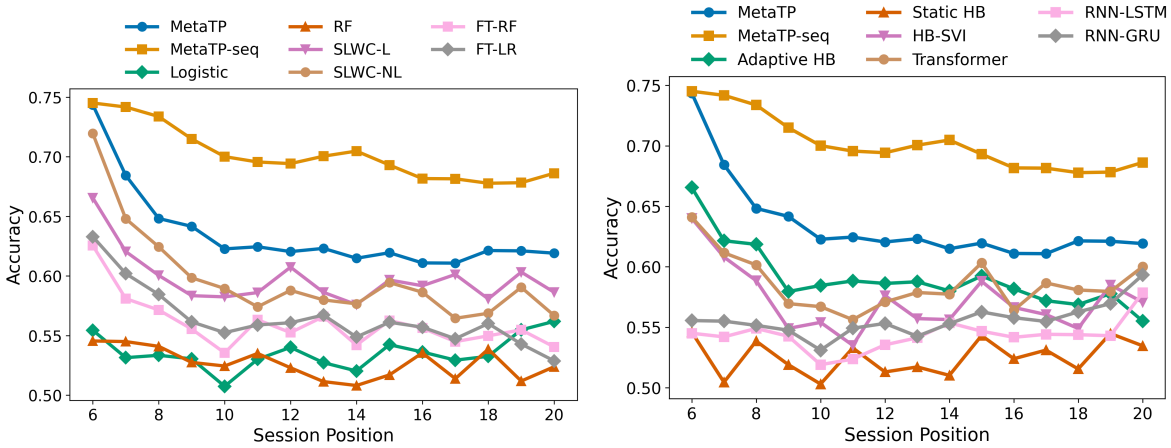


Figure F1: Predictive Accuracy at Each Target Position for All Benchmark Models.

Table F2: Overall and First Event Predictive Results for All Benchmarks.

	First Event				Overall			
	Accuracy	Precision	Recall	AUC	Accuracy	Precision	Recall	AUC
Hierarchical Bayes								
Static HB	54.60%	0.568	0.892	0.527	52.80%	0.546	0.725	0.51
Adaptive HB	66.30%	0.721	0.653	0.725	59.30%	0.631	0.588	0.617
Adapt. HB-VI	62.55%	0.576	0.811	0.601	57.61%	0.716	0.543	0.708
Sequential Models								
RNN-LSTM	54.50%	0.570	0.746	0.538	54.18%	0.560	0.700	0.545
Transformer	63.70%	0.566	0.894	0.593	57.32%	0.620	0.901	0.690
RNN-GRU	55.55%	0.563	0.908	0.529	55.38%	0.557	0.843	0.549
Supervised Models								
SLWC-L	66.55%	0.656	0.841	0.710	59.95%	0.603	0.751	0.620
SLWC-NL	71.95%	0.740	0.765	0.755	60.32%	0.616	0.698	0.628
LR	55.70%	0.577	0.762	0.541	53.60%	0.549	0.755	0.528
RF	54.40%	0.598	0.570	0.560	52.50%	0.56	0.565	0.529
FT-LR	63.30%	0.721	0.545	0.717	56.80%	0.627	0.489	0.605
FT-RF	62.60%	0.684	0.617	0.685	56.20%	0.598	0.571	0.514
MetaTP	73.20%	0.756	0.766	0.776	62.90%	0.638	0.708	0.664

Notes. The blue-marked methods are the additional baselines.

WEB APPENDIX G: ABLATION STUDIES AND BOUNDARY CONDITIONS

Ablation Studies

Table G1 shows the ablation results with Transformer (a sequential model but not trained with meta-learning tasks) and MetaTP-NS, which modifies MetaTP by removing the position encoding, thereby not modeling sequential patterns while retaining the meta-learning framework.

Table G1: Comparison between MetaTP and Transformer (not meta-trained), and MetaTP-NS (MetaTP without sequential modeling).

	First Event				Overall			
	Accuracy	Precision	Recall	AUC	Accuracy	Precision	Recall	AUC
Transformer	63.70%	0.566	0.894	0.593	57.32%	0.620	0.901	0.690
MetaTP-NS	70.50%	0.720	0.770	0.744	62.73%	0.640	0.705	0.660
MetaTP	73.20%	0.756	0.766	0.776	62.90%	0.638	0.708	0.664

Table G2: The Percentage of Accuracy Gain That Remains After Removing Various Components of Our Model.

	Remaining percentage of gain in accuracy			
	No attention	Single head	Large context set in training	No sequential
Overall Accuracy	84.14%	86.36%	73.28%	78.06%
First Accuracy	67.39%	89.86%	32.61%	62.32%

Next, we explore the source of performance gain over Adaptive HB due to modeling differences. We conduct the following leave-one-out ablation studies to evaluate the accuracy gains of MetaTP over Adaptive HB that remains. These studies quantify the importance of the nonlinearity, meta-learning task construction, and sequential modeling on MetaTP's performance. The results are in Table G2.

• *The contribution of nonlinear Transformer.* Our model incorporates the nonlinear Transformer in designing the encoder model, while the HB is a linear logit model. We remove the attention mechanism from the full model by manually setting the attention weights in Eq. (4) uniformly as $1/t_i$ for each context point. We also remove the multiple attention heads from the full model by setting the head number to 1, while keeping everything else the same. The results in Table G2 show the remaining proportion of the accuracy gain after these changes, respectively. We find an evident accuracy contribution by the multi-head attention, especially on the first event accuracy.

• *The importance of meta-training tasks.* MetaTP uses an episodic learning framework with multiple tasks, while the HB model models the entire data within a task. We examine how the performance changes without the few-shot training tasks. Specifically, we change the context size to be the maximal 9 since the shortest session length is 10. This makes the training task easier to solve because each task has richer information from the context set. Results in Table G2 show that the model's performance in accuracy is significantly impacted by these weaker meta-training tasks. This is because when the model is trained on easy tasks, it does not effectively learn how to solve few-shot problems. Thereby, this ablation study underscores the importance of meta-training the model on the few-shot learning tasks.

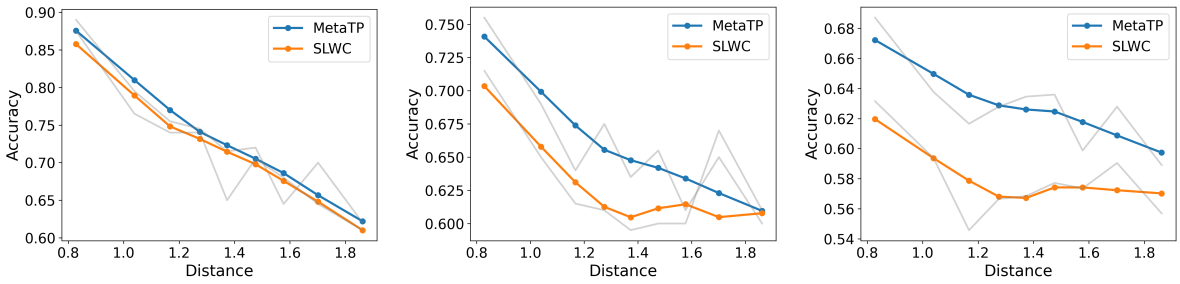
• *The impact of sequential model.* MetaTP models the sequential pattern of the data while Adaptive HB assumes individual data are exchangeable. We remove the position encoding in MetaTP to evaluate the contribution of sequential modeling. The model accuracy, as shown in Table G2, again drops significantly, which confirms the contribution of the sequential model, especially in the first event accuracy.

These leave-one-out ablation studies provide the quantification of importance for the nonlinearity, task construction, and sequential modeling on model performance. We find these components generally have a higher impact on the first positions than the later positions. For this empirical study, a proper construction of meta-learning tasks and sequential modeling stand out as crucial elements in achieving high performance.

Boundary Conditions

MetaTP can generalize to an unseen distribution $p_k(x_{kt}, y_{kt})$ by adapting to the context set. However, such generalization is not arbitrary as the distribution p_k has to be related to the distribution $p_1, \cdots, p_N$ of the training tasks. We explore the boundary conditions on this generalization ability to the distribution shift. Specifically, we take the context set embedding $\boldsymbol{z}_i^c$, then compute the distance between $\boldsymbol{z}_k^c$ of new session k to $\{\boldsymbol{z}_1^c, \boldsymbol{z}_2^c, \cdots, \boldsymbol{z}_N^c\}$ of the training sessions as $\min_{1 \leq i \leq N} ||\boldsymbol{z}_k^c - \boldsymbol{z}_i^c||_2$ with Euclidean norm $|| \cdot ||_2$. This distance between a point and a set represents the dissimilarity of a new session to the existing sessions.

In Fig. G1, we plot the predictive accuracy against the dissimilarity for the first, second, and overall positions in the target set, applying locally estimated scatter plot smoothing (LOESS) to aid interpretability. We also provide the accuracy of SLWC as a reference. There are two key observations: first, the accuracy drops as the dissimilarity increases, which validates the embedding distance as a proper measure that the platforms can use to decide the generalization boundaries. Second, the advantage of MetaTP over SLWC is most evident for the positions beyond the first and for the sessions with a moderate level of dissimilarity. This observation reveals that the advantage of MetaTP is most significant on the challenging but not impossible cases.



(a) First position **(b)** Second position **(c)** Overall

Figure G1: Accuracy Changes with New Session Similarity to Existing Sessions.

Notes. The predictive accuracy for a new session drops when its dissimilarity to the existing sessions increases, smoothed using LOESS.

WEB APPENDIX H: COMPUTATION TIME

Fig. H1 shows that MetaTP's wall-clock time grows approximately linearly with the sample size; the reported time is measured from initialization to convergence. Fig. H2 shows that MetaTP's predictive accuracy improves as the sample size increases.

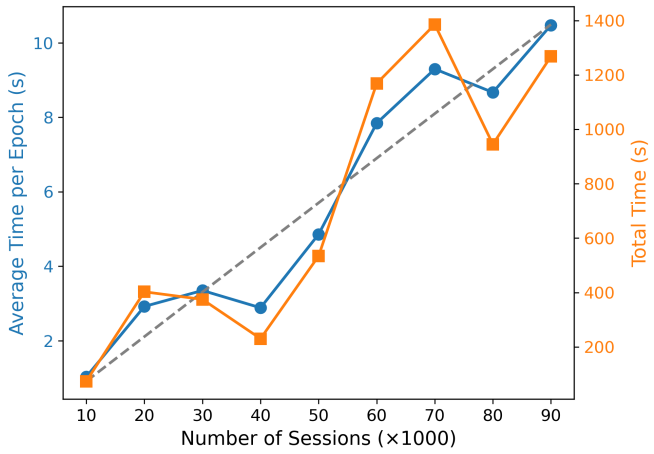


Figure H1: The Computational Time with the Sample Sizes.

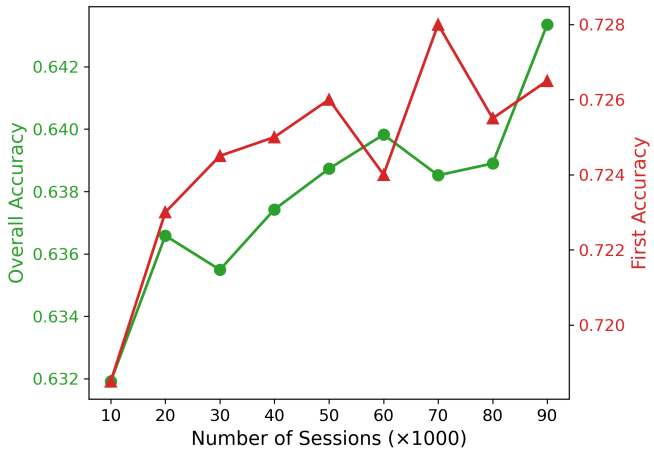


Figure H2: The Predictive Accuracy with the Sample Sizes.

WEB APPENDIX I: ADDITIONAL APPLICATION RESULTS

One-shot session completion

Table H1: One-shot Session Completion.

Context set			Session Completion	
Track	Skip	Danceability	Track	Danceability
Beautiful Zombies	1	0.15	Ms. Sweetwater (Skit)	0.00
One and Only	0	0.45	My Man's Gone Now	0.36
Waltz for the Lonely	0	0.51	The Map Room: Dawn	0.09
Last Days of My Bitter Heart	0	0.28	Portrait of Tracy - Live	0.46
Corona Radiata	0	0.09	Naked Moon - Live Version	0.34
Pavane	0	0.17	Calliope	0.26
Est Secretum	0	0.15	Con Te Partiro	0.08
Morning In The Bush	0	0.34	Warszawa	0.26
Spanish Love Song	0	0.32	Entr'acte - Live in Concert	0.17
Innocent Soul	0	0.09	Whatever (Strings)	0.56
Average	0.1	0.26	-	0.26

(a) Session completion for a customer with a *low* skip rate for low danceability tracks.

Context set			Session Completion	
Track	Skip	Danceability	Track	Danceability
My Kind Of Perfect	0	0.53	Cold Steel Canyons	0.73
Bagatelle in A Minor"	1	0.33	Tierra Del Fuego	0.52
Hemp	1	0.25	Sally Goodin'	0.37
Talking Bout My Baby	1	0.25	Move To The City - Live	0.38
Mr. Longbottom Flies	1	0.07	Speed of Life - Live	0.45
Valtari	1	0.17	Spanish Lady - From Slane Castle	0.66
Final Fantasy: Main Theme	1	0.13	Outkast - New"	0.36
Cloudy This Morning	1	0.27	Eres Divina	0.73
Dolphin Dance	1	0.23	Jump Down	0.69
Real World Applications	1	0.34	Juanita -Flor De Walamo	0.70
Average	0.9	0.23	-	0.51

(b) Session completion for a customer with a *high* skip rate for low danceability tracks.

Table H1 contains the one-shot session completion for two context sets sampled from our observed data set. One context set shows a high skip rate for low danceability tracks, while the other context set shows a low skip rate for low danceability tracks.

Table H2 shows the session completion by Adaptive HB, where one context set shows preference to vibrant, vocal music, and the other shows preference to mellow and cinematic music.

Table H2: Session Completion for Heterogeneous Context Sets by Adaptive HB.

Session Completion for a Context Set Preferring Vibrant Music				
Track	Artist	Acousticness	Energy	Speechiness
Pinch Me - Live	Barenaked Ladies	0.006	0.932	0.037
I Like It, I Love It	Tim McGraw	0.020	0.961	0.032
Flux And Flow	Lights	0.010	0.910	0.042
Don't Get Me Wrong - Live	Pretenders	0.023	0.918	0.031
Down On The Corner - Live 1997	John Fogerty	0.019	0.847	0.046
Drunk Tweets	The Lawrence Arms	0.001	0.950	0.053
When Jesus Left Birmingham	John Mellencamp	0.020	0.920	0.037
Hollywood Nights - Live	Bob Seger	0.003	0.992	0.090
I Got It - Spragga down 1 db - Final	Kevin Lyttle	0.024	0.938	0.041
Halfway To The Moon	Phish	0.006	0.897	0.029
Session Completion for a Context Set Preferring Mellow Music				
Track	Artist	Acousticness	Energy	Speechiness
The First Queer Santy Claus	Larry The Cable Guy	0.984	0.026	0.000
Meet James Ensor	They Might Be Giants	0.918	0.100	0.063
Glean	They Might Be Giants	0.971	0.074	0.067
Check Writer - Skit	Method Man	0.988	0.116	0.000
Lochinvar	Rod Stewart	0.857	0.110	0.044
True Blue - Live	Dolly Parton	0.968	0.187	0.046
Don't	Elvis Presley	0.966	0.072	0.029
Frankie In The Rain	Jennifer Warnes	0.972	0.156	0.028
The Man Who Sends Me Home	Laura Nyro	0.987	0.028	0.038
Early in the Morning - Remastered	Harry Nilsson	0.988	0.038	0.067

Notes. The next 10 recommended tracks for two sessions by Adaptive HB. The top session has Context Set 3 (vibrant music lover), and the bottom session has Context Set 2 (mellow music lover) from Table 3.

Dynamic preferences

Fig. H3 shows the preference dynamics for all the numerical acoustic fingerprints over the population.

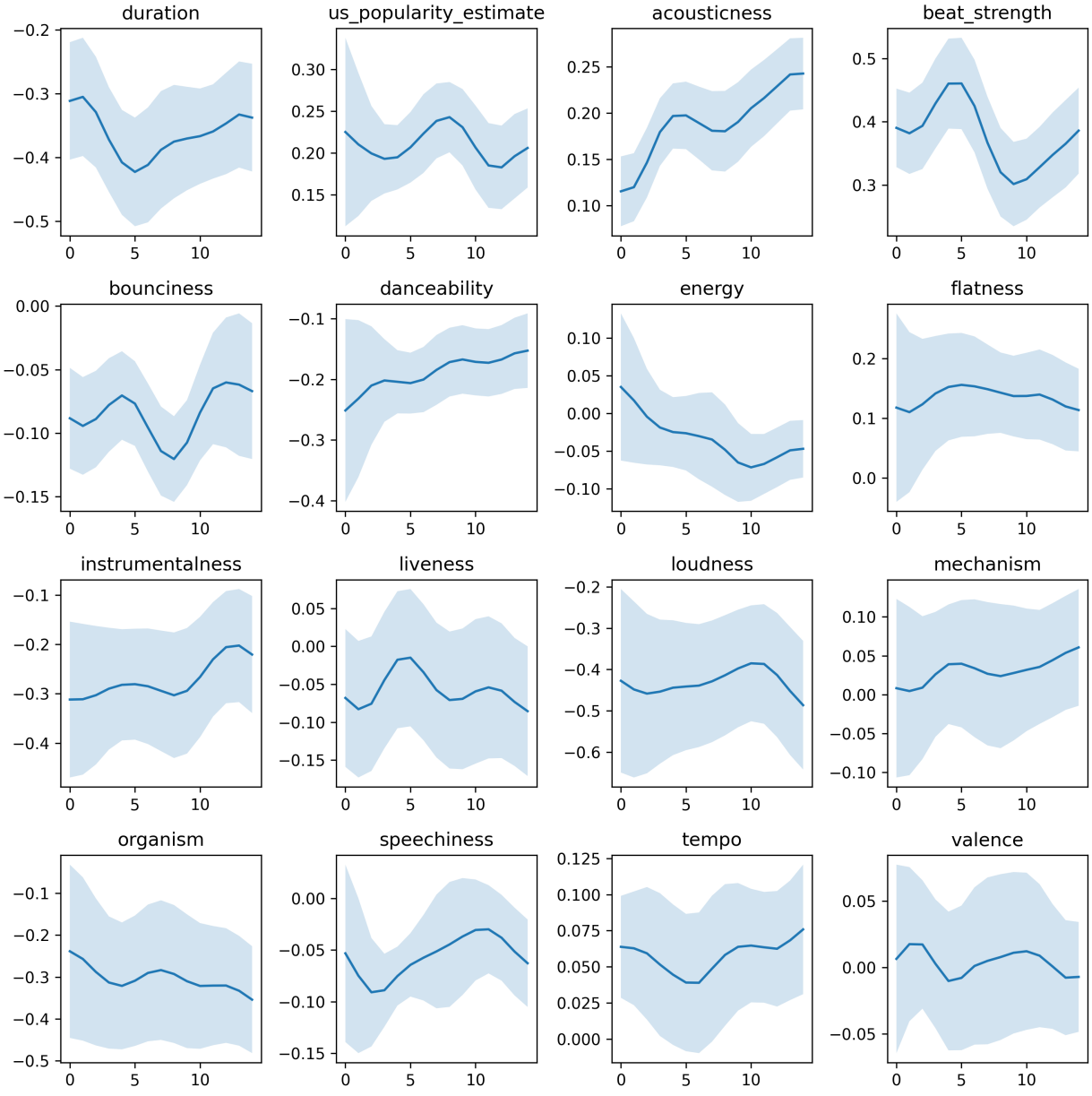


Figure H3: Preference Dynamics Distributed over the Population.

Fig. H4 plots the individual trace of β_{ip}^{t} around the mean curve. We notice that the individual preference for acousticness centers around the mean curve, while the preference

for energy exhibits two types of patterns. One subgroup prefers high-energy tracks at the beginning and gradually becomes neutral, while the other subgroup prefers soft tracks.

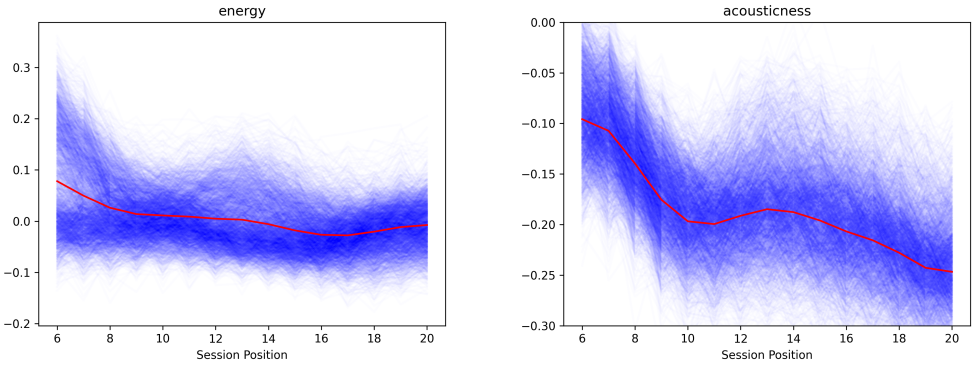


Figure H4: Dynamics of Preference Variables for Multiple Individuals.

To explore the driving factors behind the dynamic patterns, we focus on Listeners and Skippers as two subgroups of the new customer sessions. The Listeners are defined as the sessions with the top 10% listening rate, and the Skippers are defined as the sessions with the bottom 10% listening rate in the context set. The preference dynamics of these two subgroups are shown in Fig. H5. We find that the level of historical engagement in the context set significantly influences the evolution of preference dynamics. From the duration and valence panels, the Listeners are more tolerant of long tracks and prefer negative emotional songs more than Skippers. We notice that the reduced interest in the energy feature is mainly among the Listeners, and the increased interest in the instrumentalness is more evident for the Skippers. The subgroup analysis demonstrates the heterogeneous patterns of customer fatigue and identifies the listening rate of the context sets as a driving factor.

Adaptive session completion for acousticness groups

In Fig. H6, we simulate another two Prefer-High and Prefer-Low groups based on the acousticness feature. Both MetaTP and Adaptive HB recommend proper tracks to each group in the end, but LSTM cannot. MetaTP has more advantage over Adaptive HB for the groups segmented by instrumentalness than acousticness. The reason might be the instrumentalness

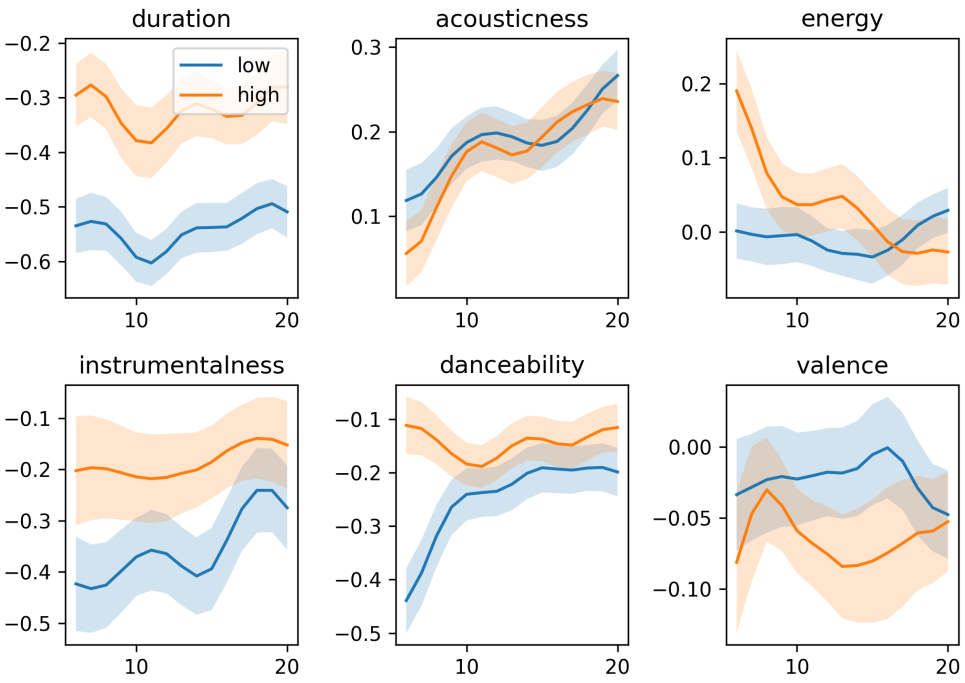
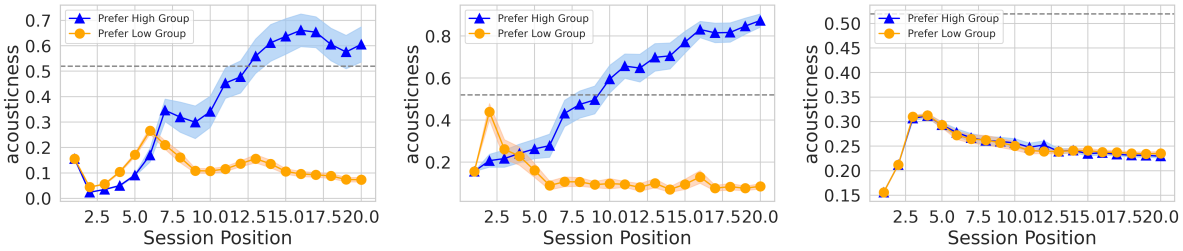


Figure H5: (Color online) Heterogeneous Preference Dynamics for Subgroups.

Notes. The orange curve corresponds to the high listening rate group, and the blue curve corresponds to the low listening rate group. The dynamics are over the target set.

distribution has a higher skewness of 2.60 (most tracks are vocal) than the skewness 0.84 of acousticness, suggesting that MetaTP model can handle more challenging cases with its flexibility.



(a) MetaTP **(b)** Adaptive HB **(c)** LSTM

Figure H6: Adaptive Session Completion for Two Heterogeneous Groups.

Notes. The dashed horizontal line is the threshold level of acousticness that separates the Prefer High and Prefer Low groups. MetaTP achieves accurate customization for new customers from these two groups after adapting to a few interactions.

Explainable Recommendations

We show how digital platforms can leverage our approach to provide structured feedback to customers regarding the recommendations they receive.

Figure H7: Netflix Recommendation Explanation based on Past Streaming Behaviors.

Digital platforms have increasingly invested in explanation systems that help users understand why particular products, content, or advertisements are recommended to them.

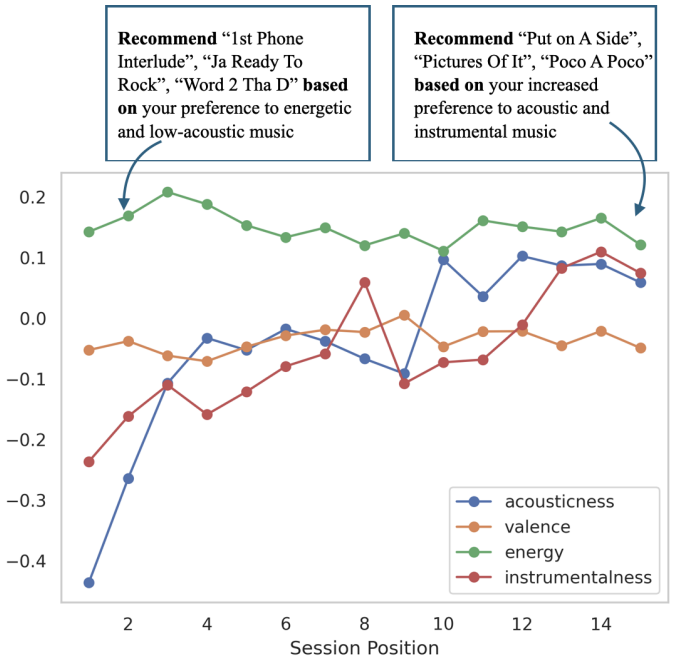


Figure H8: Dynamic Preferences for Music Features for an Example Customer, and the Explainable Recommendations at Two Positions of This Session.

Offering explanations for recommendations has been shown to foster positive consumer responses (Chen, Tian, and Jiang 2024) and to build trust in artificial intelligence systems (Nussberger et al. 2022), as it makes the underlying rationale of algorithmic recommendations more transparent. Figure H7 illustrates a prominent example of explanation strategies used by Netflix. The platform explains its recommendations primarily through users' past interaction behavior across similar content.

Our approach extends these explanation mechanisms by allowing platforms to generate explanations grounded in users' estimated latent preferences rather than surface-level correlations alone. Specifically, the model decomposes recommendations into interpretable preference dimensions, enabling platforms to explain not only what is recommended, but which underlying preference components drive the recommendation.

We illustrate this capability in the context of music streaming sessions, where few-shot interactions are observed. We consider a representative example from the application. Fig. H8 visualized the preference dynamics of an example customer over multiple music features. For

this customer, the model estimates that the preferences for valence and energy remain stable across sessions, while the preferences for acousticness and instrumentalness increase with session progression. Fig. H8 shows the recommendation at an early and a late position in the session by accounting for the evolving preferences. Importantly, beyond recommending a session, the platform can communicate the rationale underlying the recommended listening session through a transparent, user-facing explanation such as "From your prior interactions on the platform, we recommend "Put on A Side", "Pictures Of It", "Poco A Poco" based on your increased preference to acoustic and instrumental music." This form of explanation provides users with more meaningful and diagnostic feedback, while also offering platforms a principled way to audit, refine, and personalize their recommendation strategies.

WEB APPENDIX J: APPLICATION TO CONSUMER SHOPPING SESSIONS

In this application, we apply the proposed MetaTP to study the few-shot learning problems in consumer purchase sessions. Specifically, we study consumer brand choices in the IRI Consumer Packaged Goods (CPG) data (Bronnenberg, Kruger, and Mela 2008). We consider the consumer choices at the week level from January 1, 2006, to December 31, 2011, which covers 311 weeks. A goal is to examine whether a consumer's brand preferences can be uncovered from a few choices, which provides companies with crucial information for timely product targeting and pricing (Smith, Seiler, and Aggarwal 2023). We also explore whether the model can offer interpretable and managerially relevant information from its estimated parameters, such as the sensitivity to marketing variables.

Table H3: Summary Statistics of the Purchase Sessions

Category	Brands	Sessions	Interactions	Avg. Length	Std. Length
Cereal	5	387	7130	18.4	8.9
Chips	4	1158	36201	31.3	23
Coffee	5	551	13148	23.9	17.9
Peanut Butter	5	617	13634	22.1	13
Soda	7	2037	120293	59.1	61.8
Soup	4	1411	46869	33.2	26.4
Tissues	4	525	12874	24.5	19.8
Tuna	4	1091	28954	23.1	12

Task Construction and Few-shot Prediction

We consider the shopping session of a household as one task in meta-learning because the brand choice distribution given the marketing variables differs across households. We analyze the sessions with length over 10, take the context set $\mathcal{D}_i^c$ as the first 10 choices, and the target set $\mathcal{D}_i^q$ as the remaining choices for each household i. The product categories we analyzed are cereal, chips, coffee, peanut butter, soda, soup, tissues, and tuna. We analyze each category independently. The observed covariates for each brand are the brand indicator, price, coupon

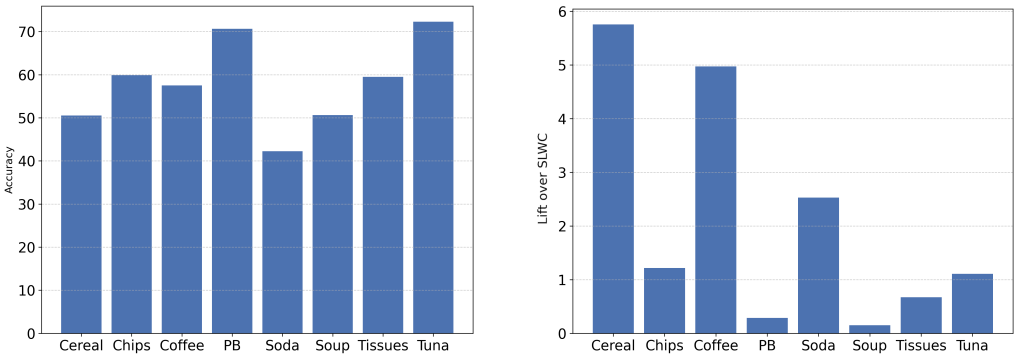


Figure H9: Accuracy (%) of MetaTP (Left) and it Minus the Accuracy of SLWC-NL (Right).

rate, and the percentage of observations in which a brand is featured or displayed (Ftdsp). The summary statistics of each product category are in Table H3. For the MetaTP, the decoder is a logit model as described in the *Modeling Approach* section.

We demonstrate MetaTP's effective few-shot prediction and compare it with SLWC-NL, the benchmark model with the best predictive performance in our first application. We take 80% of sessions for model training, and make predictions on the target set of the remaining new sessions given their context set with ten observed choices. The results in Fig. H9 show the overall accuracy of MetaTP on the left panel and its difference from SLWC-NL on the right panel. The predictive accuracy of MetaTP is higher than that of SLWC-NL across all the product categories. Note that SLWC-NL is a blackbox model, while MetaTP balances predictive performance with interpretability yet achieves a higher accuracy. Next, we demonstrate several applications of MetaTP that go beyond mere prediction of customer product choices.

Interpretable Sensitivity Dynamics

The estimated parameters of MetaTP reveal dynamic individual preferences for a variety of marketing mix variables. Fig. H10 visualizes the evolution of the aggregated $\sum_i \boldsymbol{\beta}_i^t / N$ for different product categories where $\boldsymbol{\beta}_i^t$ is estimated by its posterior mean. The price is always

negatively related to the brand choice, but the magnitude varies across product categories. Customers are more price sensitive with tuna and peanut butter and are less price sensitive with soda. The price sensitivity of tuna tends to decrease over time. The coupons are generally positively related to brand choice, though we find the coupons are particularly effective for soup and peanut butter, but not as effective with soda. The positive effects of coupons also increase for peanut butter while decreasing for soda. In terms of featured display, it is positively associated with brand choice, though the effects on chips, soup, and tuna tend to wind down over time, which might be a result of promotion saturation or fatigue. This trend offers signals for a company to revise its display promotion strategies.

In Fig. H11, we show the individual-level dynamics by plotting its deviation from the population mean as $\boldsymbol{\beta}_i^t - \sum_i \boldsymbol{\beta}_i^t / N$ with 5 randomly sampled individuals per product category. Most households remain above or below the mean consistently over time, but some are converging to or diverging from the mean. The individual dynamics provide a timely opportunity for the firms to conduct intervention, for example, by offering targeted coupons when the coupon effect becomes more evident for the focal household.

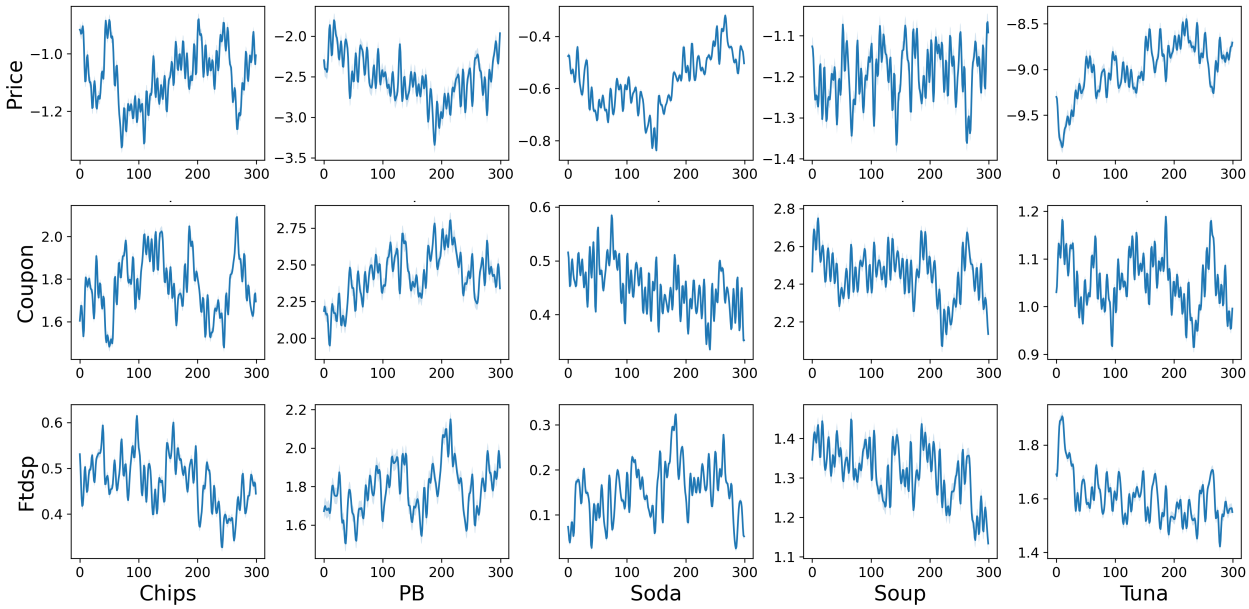


Figure H10: Dynamic of Sensitivity for Price, Coupon and Display over Weeks (the X-axis).

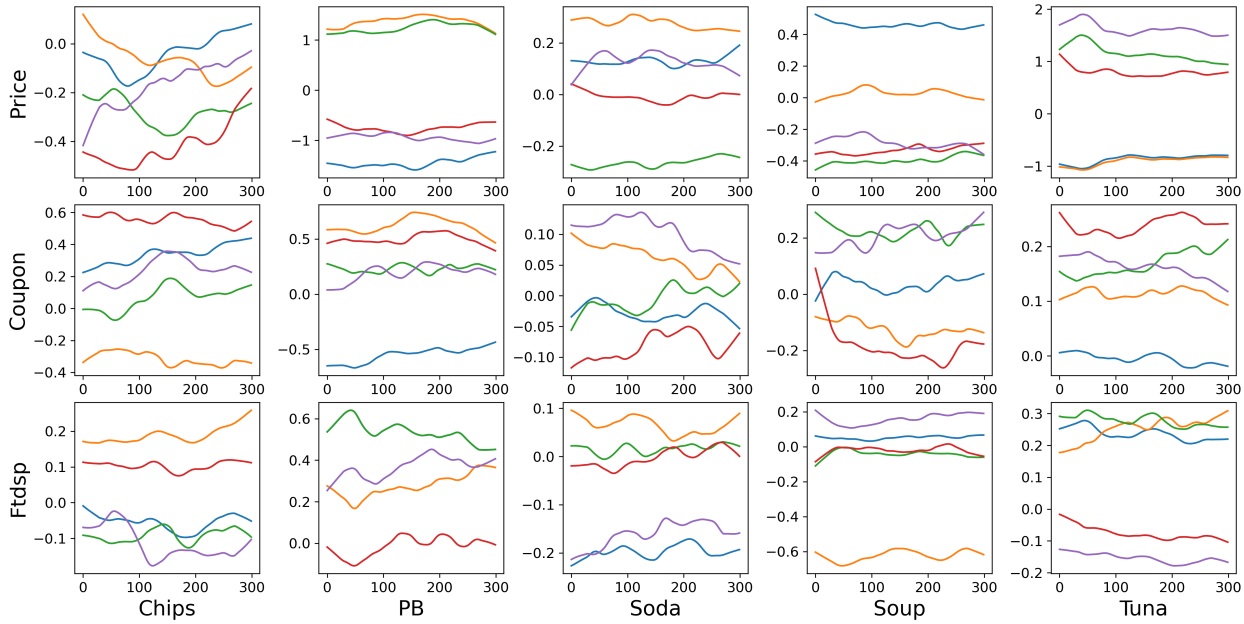


Figure H11: Dynamic Individual Deviation from the Mean (X-axis is the week).

Next, we use the outputs of MetaTP to compute the price elasticity, which reflects how the choice of a brand reacts to changes in price. Specifically, we use the own-price elasticity defined as $\eta_{ib}^t = \boldsymbol{\beta}_{ip}^t \times \text{Price}_b^t \times (1 - p_{ib}^t)$, where $\boldsymbol{\beta}_{ip}^t$ is the element of $\boldsymbol{\beta}_i^t$ for the price, Price_b^t is the price of brand b at week t, and p_{ib}^t is the probability of household i choosing brand b at week t which is estimated from the logit model decoder. We select the soda category for a case study because the household purchases of soda are dense and consistent over time, as shown in Table H3, with the longest session lengths and the largest number of choices. Fig. H12 shows the average price elasticity over the households for each brand and at each week. We use two dashed lines to mark December 1, 2007, and June 1, 2009, placing the period of the Great Recession in between (Bernanke 2010).

Fig. H12 left panel shows the average elasticity over the households, which generates several interesting insights. First, the elasticity exhibits significant heterogeneity across brands, indicating different levels of brand loyalty among customers. Second, the recession caused a significant increase in the magnitude of own-price elasticity. The elasticity increased more for the brands that customers were already highly price-sensitive to before the recession.

Brands with low own-price elasticity tend to maintain their low levels during the recession. We also notice that the price elasticity has an increasing trend for several brands before the recession period, which might be due to rising costs and economic uncertainty. This suggests that the change in price elasticity identified by MetaTP for the fast-moving consumer goods might serve as a valuable forecasting signal for major market shifts.

Compared to the results of (Dew, Ansari, and Li 2020) over the same dataset (Appendix Figures 7-12), we note that our estimated parameter dynamics exhibit a very similar pattern. Both approaches find that the brands experienced a clear increase in average price elasticity during the Great Recession. This similarity corroborates the accuracy of our estimation. However, our analysis is conducted at the weekly level, while Dew, Ansari, and Li (2020) is at the monthly level. The scalability of our approach, therefore, enables more fine-grained demand management for firms.

Beyond the average analysis, the right panel shows the estimated price elasticity for a randomly selected household. The individual elasticity can be significantly different from the average in magnitude and dynamic trends. For example, the household is more sensitive to the price of brand 7 and relatively less sensitive to the price of brand 1. This household-level diversity uncovered by MetaTP can be utilized for personalization. Next, we show managerial implications of MetaTP for retailers with an application of targeted pricing.

Few-shot Targeted Pricing for New Households

MetaTP enables a firm to quickly set a targeted price for a new household, such as through coupon delivery, after observing a few brand choices from this household. We follow the setting of Rossi, McCulloch, and Allenby (1996) and suppose the product manufacturer margin is M for a brand b, the estimated net revenue R_{ibt} for the household i to choose this brand at time t is

$$R_{ibt} = p(y_i^t = b \,|\, \hat{\boldsymbol{x}}_{it}^{\Delta}, \boldsymbol{\beta}_i^t)(M + \Delta). \tag{19}$$

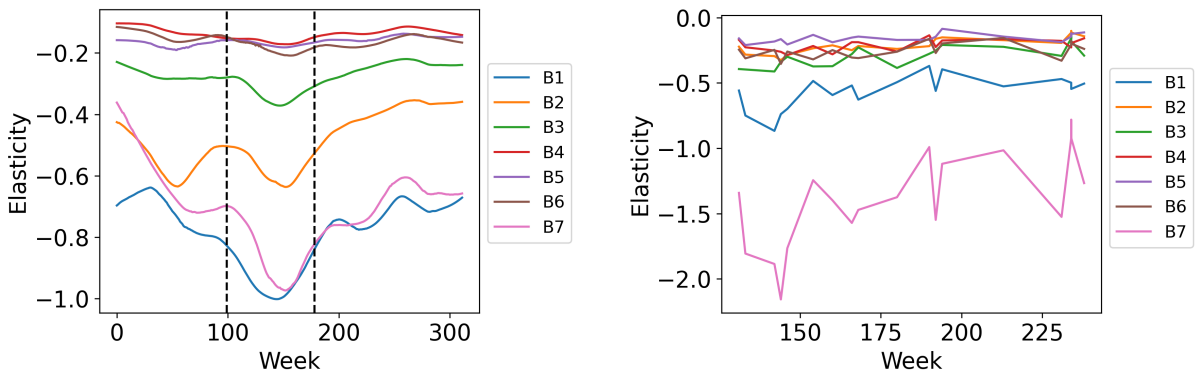


Figure H12: Price Elasticity Dynamics.

Notes. The left panel is the average price elasticity (smoothed by LOESS (Cleveland 1979) for interpretability); the right panel is the price elasticity for a randomly selected household. B1 to B7 are seven soda brands.

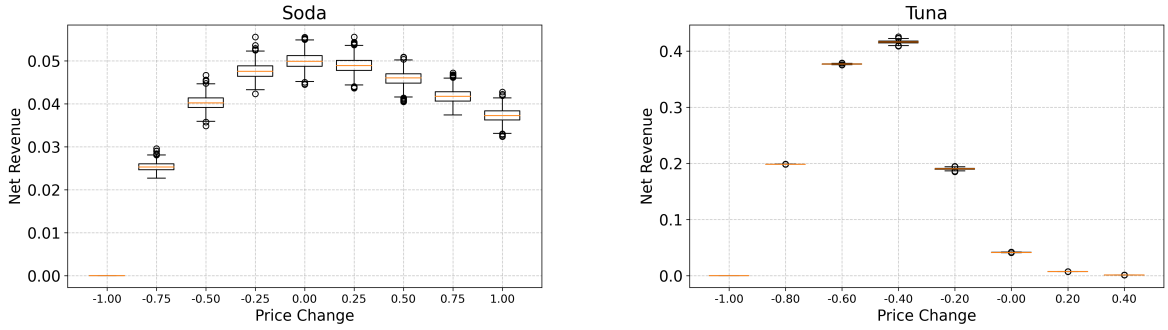


Figure H13: Household-level Posterior Distribution of the Estimated Net Revenue across Prices.

Here, $\Delta \in \mathbb{R}$ is the adjusted price for this household. For example, a $\Delta < 0$ can be implemented by personalized coupons with face value Δ. $\hat{\boldsymbol{x}}_{it}^{\Delta}$ are the covariates at time t except the price of the brand b being replaced by its original price plus Δ. $p(y_i^t = b \,|\, \hat{\boldsymbol{x}}_{it}^{\Delta}, \boldsymbol{\beta}_i^t)$ is the probability of household i choosing brand b given its adjusted price and the parameters $\boldsymbol{\beta}_i^t$ sampled from MetaTP encoder.

Rossi, McCulloch, and Allenby (1996) show that the uncertainty of the estimated choice probability plays an important role in avoiding the overestimation of net revenue. To capture the revenue uncertainty, MetaTP can sample $\boldsymbol{\beta}_i^t \sim q(\boldsymbol{\beta}_i^t \,|\, \boldsymbol{x}_{it}, y_{it}, \mathcal{D}_i^c, t)$ from the encoder and compute the posterior distribution of R_{ibt} using these parameter samples. In this analysis, we randomly sample a new, unseen household from the hold-out sessions of each product

category and select one brand as the focal brand. We use this household's first ten choices as the context set and consider the targeted pricing of the focal brand for their next choice. The margin M is assumed to be \$1.

Fig. H13 shows the posterior distribution of the net revenue by the boxplots across various price change Δ. The optimal price can be determined by the posterior mean, which represents the Bayesian optimal decision strategy to maximize the expected revenue. We find that the optimal price of the soda brand for the target household is at its current price. However, the optimal price of the tuna brand would be reached with a \$0.4 discount, which significantly increases the expected net revenue from this household.

WEB APPENDIX REFERENCES

Ansari, Asim, Yang Li, and Jonathan Z Zhang (2018), "Probabilistic topic model for hybrid recommender systems: A stochastic variational Bayesian approach," *Marketing Science*, 37 (6), 987–1008.

Barnard, John, Robert McCulloch, and Xiao-Li Meng (2000), "Modeling covariance matrices in terms of standard deviations and correlations, with application to shrinkage," *Statistica Sinica*, pages 1281–1311.

Bishop, Christopher M and Nasser M Nasrabadi (2006), *Pattern recognition and machine learning* Springer.

Brown, Tom, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell et al. (2020), "Language models are few-shot learners," *NeurIPS*.

Carpenter, Bob, Andrew Gelman, Matthew D Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus A Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell (2017), "Stan: A probabilistic programming language," *Journal of statistical software*, 76.

Garnelo, Marta, Dan Rosenbaum, Christopher Maddison, Tiago Ramalho, David Saxton, Murray Shanahan, Yee Whye Teh, Danilo Rezende, and S M Ali Eslami "Conditional Neural Processes," Jennifer Dy and Andreas Krause, editors, "ICML," (2018a).

Garnelo, Marta, Jonathan Schwarz, Dan Rosenbaum, Fabio Viola, Danilo J Rezende, SM Eslami, and Yee Whye Teh (2018b), "Neural processes," *arXiv preprint arXiv:1807.01622*.

Hidasi, Balázs, Alexandros Karatzoglou, Linas Baltrunas, and Domonkos Tikk (2015), "Session-based recommendations with recurrent neural networks," *arXiv preprint arXiv:1511.06939*.

Hoffman, Matthew D, Andrew Gelman et al. (2014), "The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo.," *J. Mach. Learn. Res.*, 15 (1), 1593–1623.

Kingma, Diederik P and Jimmy Ba (2014), "Adam: A method for stochastic optimization," *arXiv preprint 1412.6980*.

Lewandowski, Daniel, Dorota Kurowicka, and Harry Joe (2009), "Generating random correlation matrices based on vines and extended onion method," *Journal of multivariate analysis*, 100 (9), 1989–2001.

Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg et al. (2011), "Scikit-learn: Machine learning in Python," *the Journal of machine Learning research*, 12, 2825–2830.

Singh, Gautam, Jaesik Yoon, Youngsung Son, and Sungjin Ahn (2019), "Sequential neural processes," *Advances in Neural Information Processing Systems*, 32.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin (2017), "Attention is all you need," *Advances in neural information processing systems*, 30.

Ziegler, Daniel M, Nisan Stiennon, Jeffrey Wu, Tom B Brown, Alec Radford, Dario Amodei, Paul Christiano, and Geoffrey Irving (2019), "Fine-tuning language models from human preferences," *arXiv*.