When Will Customers Buy? A Deep Learning Approach Incorporating Adaptive Irregularity for Next Purchase Prediction

Jessica Qiuhua Sheng[1]
Department of Systems and Operations Management
California State University-Northridge
18111 Nordhoff Street, Northridge, CA 91330-8309, USA
qiuhua.sheng@csun.edu

Da Xu
Department of Marketing, Analytics, and Professional Sales
University of Mississippi
University, Mississippi, 38677, USA
dxu@olemiss.edu

Pouyan Eslami
Department of Systems and Operations Management
California State University-Northridge
18111 Nordhoff Street, Northridge, CA 91330-8309, USA
pouyan.eslami@csun.edu

Daeeun Daniel Choi
Department of Systems and Operations Management
California State University-Northridge
18111 Nordhoff Street, Northridge, CA 91330-8309, USA
dchoi@csun.edu

ABSTRACT

The ability to accurately predict the timing of the next purchase is critical for business decision-making yet challenging. Shopping regularity is often disrupted by negligible transaction costs of e-commerce and the ease of responding to promotions, fostering increasingly irregular behavior and thereby complicating prediction efforts. In addition, existing predictive models often overlook how buying in one product category affects future purchases in others. To address these issues, this study proposes a deep learning framework that integrates the purchase irregularity, captures category-specific purchase patterns, and learns cross-category interactions for effective next purchase time prediction at product category level. Specifically, we model purchase irregularity as a latent state that adaptively captures whether a purchase in each product category tends to follow a routine pattern or not. Then we utilize the LSTM networks to capture recurring purchase patterns based on past inter-purchase intervals. Finally, a self-attention mechanism is applied to capture interactions of shopping behaviors among distinct product categories, learning how the timing of purchases in one category can affect purchasing behavior in others. Experimental evaluations on a large-scale retail dataset demonstrate the effectiveness of our approach. The proposed model improves purchasing time prediction and enables businesses to better anticipate demand fluctuations and optimize resource allocation in online marketplaces.

Keywords: Next Purchase Prediction; E-commerce; Deep Learning; Self-attention; Purchase Irregularity

---

Cite: Sheng, J. Q., Xu, D., Eslami, P. & Choi, D. D. (2026). When Will Customers Buy? A Deep Learning Approach Incorporating Adaptive Irregularity for Next Purchase Prediction. *Journal of Electronic Commerce Research, 27*(4), 335-350.
[1] Corresponding author

1. Introduction

Predicting consumer purchase behavior on e-commerce platforms is essential for developing effective business strategies, as it enables businesses to optimize marketing efforts, enhance customer engagement, improve inventory management, and deliver personalized shopping experiences (Qiu et al., 2015; Yang & Pan, 2004). Prior studies on consumer behavior prediction have identified three fundamental aspects that predict consumers' purchasing behavior: "what" a consumer will purchase, "if" they will make a purchase, and "when" the purchase occurs. These dimensions form a comprehensive framework that supports tailored promotional strategies in e-commerce. Studies addressing the "what" aspect focus on predicting the items customers are likely to purchase, typically using recommender systems that utilize user-level data, such as explicit preferences (e.g., ratings or likes) and implicit signals (e.g., browsing behavior and past purchases), to generate personalized suggestions (Bodapati, 2008; Huang et al., 2019; Zhang & Pennacchiotti, 2013). These methods have been instrumental in improving customer satisfaction and driving sales by delivering relevant recommendations. The "if" aspect of purchase behavior research examines the likelihood that a customer will purchase within a given time frame, often employing classification or probabilistic models to inform customer retention efforts and demand forecasting (Chaudhuri et al., 2021; Martínez et al., 2020).

These two dimensions, what customers buy ("what") and if they will make a purchase within a specific time period ("if"), play a critical role in dictating the timing of purchases ("when"), as these products may be needed at specific moments or occasions. For example, seasonal items like decorations are purchased closer to the year's holidays, whereas school supplies might be bought before the semester begins. The timing of consumer purchases represents a critical dimension of behavioral analysis that can drive strategic decision-making across business functions (Anıl & Akcayol, 2020; Grigoraș & Leon, 2023; Manzoor & Akoglu, 2017). Accurately predicting when a customer is likely to make a purchase (of specific products) allows firms to optimize the timing of personalized promotions, emails, and advertisements, increasing the likelihood that a customer will respond to these efforts by completing a transaction (Kim et al., 2024; Lismont et al., 2018). From an operational standpoint, purchase timing forecasts also support proactive inventory planning and resource allocation (Ha et al., 2002; Yang & Pan, 2004).

To accurately predict the timing of consumer purchases, businesses increasingly rely on advanced quantitative models, including machine learning and deep learning techniques, which are capable of uncovering complex patterns from extensive consumer data. Traditional machine learning methods, such as XGBoost (Chen & Guestrin, 2016; Wang et al., 2023), feed-forward neural networks (Droomer & Bekker, 2020), and association rule-based methods (Qiu et al., 2015), primarily rely on handcrafted features (e.g., average inter-purchase time, past frequency, and recency) to predict the timing of a purchase. These methods rely on engineered features extracted from historical events and cannot inherently maintain "memory" about past transactions. Deep learning architectures that have emerged focus on sequence-aware models with memory components, such as recurrent neural networks (RNNs) (Song, 2017) and long short-term memory (LSTM) (Sakar et al., 2019). These architectures are particularly effective at capturing sequential and temporal patterns in transaction data.

Despite the effectiveness of sequence-aware models, predicting the timing of consumer purchases is a complex task fraught with challenges. First, most prior studies treat each customer purchase as a single event without considering the distinct demands among product categories; therefore, methods proposed in prior studies are not optimal in predicting the timing of purchases for certain product categories (Anıl & Akcayol, 2020; Saito et al., 2021). In fact, predicting the timing of the next purchasing for certain product category is more essential because different product categories have different consumption cycles (e.g., perishable foods vs. electronics)[2]. Understanding consumption timeline at product category level could help businesses optimize inventory and promotions (Rekasiute et al., 2022). Second, accurately predicting timing of next purchase faces the challenge of "collapse of regularity" (Kahn & Schmittlein, 1989; Kim & Park, 1997). Traditionally, shopping habits may follow relatively predictable schedules, with consumers planning purchases at regular intervals (Kim & Park, 1997). However, the rise of e-commerce has disrupted shopping regularity, making it easier for consumers to act on impulsive decisions, promotional incentives, event-driven needs, or lifestyle changes (Kahn & Schmittlein, 1989; Kim & Park, 1997; Koschate-Fischer et al., 2018), thereby making it more difficult to predict the timing of the next purchase. Third, relationships among distinct product categories further complicate predicting customer buying behavior: purchases in one product category often influence the timing of another, with complementary goods reinforcing simultaneous purchases and substitutes delaying subsequent buys (Chintagunta & Haldar, 1998; Seetharaman et al., 1999). Such

[2] This study focuses on predicting when the next purchase is at product category level. We argue that predicting purchase timing at product category level is more practical and actionable than at specific product/items. Specific product/item purchases can be sporadic. In contrast, product categories demonstrate more stable consumption patterns that are suitable for demand forecasting and behavior mining.

interactions highlight that consumer purchase behavior is not isolated; rather, it is embedded within a broader, dynamically evolving decision context.

Motivated by these challenges, this study aims to predict the timing of consumer purchases at the product-category level through a deep learning approach. This task facilitates demand forecasting and customer behavior analysis, providing valuable, actionable insights for business decision-making. The novelty of this method lies in two aspects. First, we introduce the purchase irregularity, a latent state that represents whether a purchase in each product category is likely to follow a routine pattern or not. Consumer shopping behaviors are characterized by mixtures of regularity and irregularity, where regularity refers to consistent, routinized purchase intervals and irregularity reflects disruptions of regularity caused by impulsive decisions, promotions, or life changes (Azizah et al., 2022; Koschate-Fischer et al., 2018; Mela et al., 1998). Consequently, the observed time intervals between purchases reflect this dual nature. The purchase irregularity state is designed to adaptively capture such dual nature and aims to enhance predictions of the next purchase time. Second, we combine category-specific LSTMs with the self-attention mechanism to capture cross-category interactions for next purchase time prediction. Specifically, we compute similarity scores between the upcoming purchase and a user's historical purchases, then integrate these scores into category-specific LSTM networks. The similarity scores are used to weight historical consumption events, thereby emphasizing past purchases that exhibit purchase irregularities similar to those of the forthcoming purchase. A subsequent self-attention is further applied upon the outputs of the category-specific LSTMs to account for the cross-category interactions. In this way, the proposed method allows parallel execution of category-specific LSTMs and enables the model to leverage complementary temporal data across categories, producing a synergistic representation for effective predictions. Finally, the proposed method combines these components into one end-to-end deep learning model to predict when each product category will be purchased next for a customer.

2. Literature Review

Understanding online purchase behavior is a primary driver of e-commerce success. According to the cognitive–habit model (Wood & Rünger, 2016), consumer behavior and decision-making are jointly reflected by deliberate cognitive evaluations and automatic, habitual responses. This perspective applies directly to the e-commerce context, where online purchase behaviors range from routinized to irregular.

Within the cognitive–habit framework, routinized purchases are primarily governed by habitual processes that develop through repeated, satisfactory interactions in stable contexts. Prior research on social trust and satisfaction (Gefen, 2000; Pavlou, 2003) can be positioned as critical antecedents to habit formation. Specifically, trust in an online platform reduces perceived uncertainty and cognitive effort, while satisfaction reinforces positive behavioral outcomes, jointly facilitating the internalization of routine purchases (Bhattacherjee, 2001; Eid, 2011; Gefen, 2000). As habit strength increases, purchasing behavior becomes increasingly automatic and less reliant on deliberate intention or evaluative judgment. Consequently, this habit-dominant pathway explains why consumers frequently repurchase online with minimal cognitive engagement.

In contrast, irregular purchases arise when habitual cues are weak, disrupted, or absent, thereby activating deliberate cognitive processes (Shah et al., 2014). Within the framework of the cognitive–habit model, the theory of planned behavior provides a subsidiary mechanism that operates when habitual guidance is weak and explains how consumers form purchase decisions once deliberation is required by emphasizing the roles of attitudes, subjective norms, and perceived behavioral control (Ajzen, 1991). For example, attitudes toward online shopping may be shaped by beliefs about convenience, price advantages, and variety (Kim & Eastin, 2011). Consumers who hold favorable attitudes toward the price advantages of online shopping may respond opportunistically to temporary discounts or flash sales, leading to irregular purchasing behavior (Dholakia, 2000). Subjective norms reflect individuals' perceptions of social expectations regarding purchase behavior (Ajzen, 1991). Such perceived social pressures (e.g., positive online reviews) can exert strong normative influence, increasing the likelihood of impulsive purchases even when consumers initially have no intention to buy a product (Liu et al., 2022; Park & Lee, 2009; Wang & Chou, 2014). Moreover, perceived behavioral control captures individuals' assessments of how easy or difficult it is to complete a purchase (Ajzen, 1991). E-commerce platforms enhance perceived behavioral control by simplifying the purchasing process through features such as one-click checkout, stored payment information, and AI-driven auto-fill forms (Unal & Park, 2023; Venkatesh & Davis, 2000). As a result, these platforms reduce cognitive and operational barriers to purchase execution, thereby increasing the likelihood of non-routine purchases. Beyond deliberation, event-driven needs and lifestyle changes can alter goals, which disrupt established routines and introduce variability in purchase timing (Kahn & Schmittlein, 1989; Kim & Park, 1997; Koschate-Fischer et al., 2018). Building on prior work on purchase regularity (Platzer & Reutterer, 2016; Reutterer et al., 2021; Vakratsas & Bass, 2002), we define purchase irregularity as the extent to which a consumer's purchase timing deviates from a stable, routinized rhythm.

Empirically, this deviation is reflected in greater variability in inter-purchase intervals relative to the consumer's baseline behavior.

Beyond theoretical frameworks that explain online purchasing behaviors, accurately predicting when the next purchase occurs requires robust quantitative methods. Prior data-driven approaches for predicting the timing of the next purchase typically focus on modeling the following aspects: occasion-driven purchases, seasonal and promotional incentives, consumer online behaviors, and sequential behavior patterns. Methods focusing on occasion-driven predictions mainly capture shifts in intrinsic user behavior triggered by specific occasions (e.g., birthdays, anniversaries, celebrations). For example, analyzing long-term behavior patterns and incorporating shifts in intrinsic user behavior from occasions enhance prediction of the next purchasing item within a short time frame (Wang et al., 2020). Methods based on seasonal and promotional predictions leverage product seasonality and promotional cycles, such as end-of-season sales, to identify future purchases. For example, the seasonality of products and end-of-season sales are integrated into recommendation systems to suggest the right products at the right time proactively (Keerthika & Saravanan, 2020). Demographic data (e.g., age, location, income level) combined with consumer online behavioral data (e.g., user interactions or browsing patterns on an e-commerce platform) are also commonly used to predict future purchase behavior (Liu & Ma, 2019). Finally, sequence-based methods leverage deep learning and sequential pattern mining to analyze detailed consumer-item interaction histories, thereby improving predictions of future purchases (Verma, 2020).

As customer purchase decisions are often influenced by complementary, substitutive, and co-purchased items, unveiling relationships among them is crucial for predicting consumer purchase behavior in e-commerce. For example, item relationship graph neural networks (IRGNN) model complementary/substitute relationships among items, while a co-purchase graph connects items if they are frequently purchased together within an order (Linden et al., 2003; Liu et al., 2021). GraphSAGE, an inductive GNN, captures relationships between products for better prediction of complementary or substitute product links (Ahmed et al., 2021). The directed graph-based GNN learns asymmetric product relationships, effectively predicting relevant related items while mitigating selection bias (Virinchi et al., 2022). Deep learning methods encode product assortments into dense latent embeddings, capturing implicit relationships among products based on their co-purchase patterns (Gabel & Timoshenko, 2022).

Existing studies reveal several research gaps. First, although prior research has focused on recurring purchasing behavior, limited attention has been given to the irregular patterns emerging in online shopping. The rapid expansion of e-commerce introduces significant variability and unpredictability in purchase timing. Therefore, there is a need for a method that can measure purchase irregularity and adaptively predict when the next purchase occurs. Second, existing methods often treat a customer's purchase as a single event, making predictions at the transaction level, rather than at the product category (or product) level; however, product category level predictions are more meaningful and valuable for stakeholders on e-commerce platforms. Third, conventional graphical models typically represent user-product or co-purchase relationships using simplistic directed edges, inadequately capturing the complexity of cross-category purchase interactions over multiple shopping events. These methods often fail to address how buying decisions in one product category dynamically influence subsequent purchasing behaviors in other product categories over time. Addressing these limitations requires novel analytical approaches capable of modeling both the stochastic nature of purchasing behaviors and cross-category purchase interactions to enhance future purchase timing prediction.

3. Method Design and Theory Foundation

3.1. Problem Definition

Let $U = [u_1, u_2, \ldots, u_N]$ be a set of customers, and N represents the total number of users. We use c to denote a specific product category, where $c \in [1, 2, \ldots, C]$, and C is the total number of categories. The timing at which customer u_i purchased product category c is denoted as $\boldsymbol{T}_i^c = \left[t_{i,1}^c, t_{i,2}^c, \ldots, t_{i,s}^c\right]$, where s is the total number of purchases of u_i for category c. The time intervals of two consecutive purchases of category c are represented as $\boldsymbol{I}_i^c = \left[I_{i,1}^c, I_{i,2}^c, \ldots, I_{i,s-1}^c\right]$, where $I_{i,j}^c = \Delta(t_{i,j+1}^c - t_{i,j}^c)$ for $j = 1, 2, \ldots, s-1$. We formally define the purchase timing prediction problem as follows: Given u_i's entire inter-purchase interval sequences $\{\boldsymbol{I}_i^c\}_{c=1}^C$, the objective is to effectively predict u_i's next purchasing time y_i^c for a specific product category c.

Figure 1 shows the framework architecture of the proposed method, which consists of three key components: purchase irregularity modeling, category-level purchase pattern mining, and cross-category relationship learning. Together, these components are connected in a pipeline, and data flows sequentially through these components. The purchase irregularity modeling component adaptively estimates the likelihood of the purchase being routine or irregular at each time step. The category-level purchase pattern mining component applies category-specific LSTMs to learn the unique consumption patterns over time associated with each product category; that is, each LSTM is trained separately on the sequential record of a specific product category to capture its consumption pattern. Based on the outputs from LSTMs, the cross-category relationship learning component applies a self-attention mechanism to capture the interactions of consumption patterns across different product categories. It learns how much past time intervals across product category contribute to predicting next purchase time. We next provide the theoretical rationale and logic underlying our method design, along with technical details of each component.

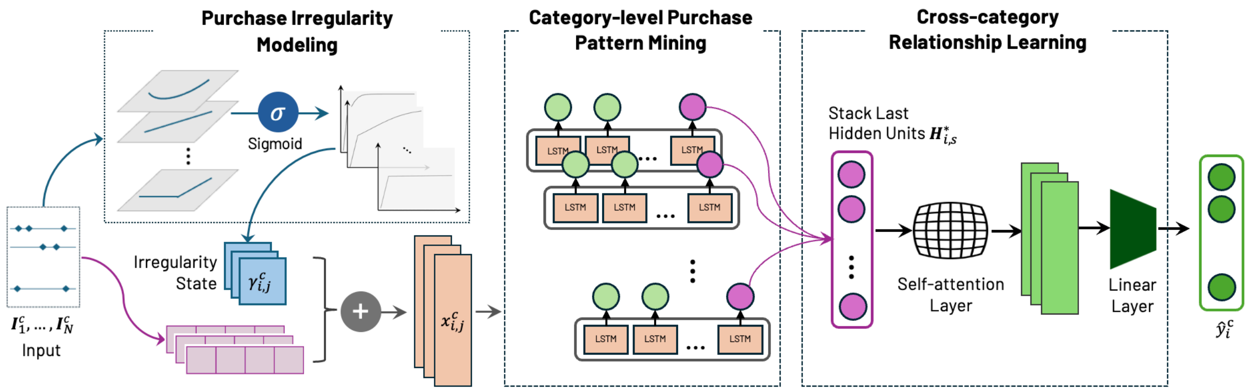

Figure 1. Design Framework

3.2. Purchase Irregularity Modeling

The observed time intervals between purchases are jointly shaped by both regular and irregular purchasing behaviors (Bawa & Ghosh, 1991; Gupta, 1991). On one hand, purchase behavior can occur at consistent, routinized intervals. On the other hand, these routines can often be disrupted by factors such as impulsive decisions, event-driven purchasing, or life changes (Azizah et al., 2022; Koschate-Fischer et al., 2018; Mela et al., 1998), introducing irregularity into consumers' shopping patterns. To capture this dual nature, we introduce purchase irregularity, a latent state that indicates whether the purchase (in each product category) is likely to follow a routine pattern or not. The computation of purchase irregularity is grounded in state dependence effect, a concept originating in dynamic choice theory and habit formation, indicating past experiences shape the underlying state that governs what happens next (Heckman, 1981). In economics, state dependence is evident in phenomenon such as unemployment dynamics, where being unemployed today increases the probability of being unemployed in the future. Extended in consumer behavior research, it manifests as persistence in shopping and brand choice, as past consumption experiences foster loyalty, guide search, and facilitate learning (Dubé et al., 2010). Building on this foundation, our study quantifies purchase irregularity within a state dependent framework, in which past shopping behaviors (characterized by varying levels of regularity or irregularity) shape future behavioral tendencies toward regularity or irregularity.

We define the purchase irregularity as a latent state $\gamma_{i,j}^{c}$ representing whether customer u_i's purchasing behavior for product category c leans toward more routine or irregular at time $t_{i,j}^{c}$. The value of $\gamma_{i,j}^{c}$ ranges from [0, 1], determined by the customer's purchasing behavior before $t_{i,j}^{c}$, where 0 corresponds to entirely routine purchase behavior, and 1 corresponds to irregular purchase behavior. Specifically, the category-specific irregularity $\gamma_{i,j}^{c}$ is computed as a linear combination of three key components derived from the customer's past purchasing behavior: a weighted average of all previous irregularity state with temporal decay rate λ, a tail-sensitive standard deviation $\beta_{i,j}^{c}$, and a category-specific bias term $b_{i,j}^{c}$. The rationale of this design is to capture both the regularity and volatility of purchasing behavior at the product category level. The weighted average of previous irregularity states $\frac{\sum_{\tau=1}^{j-1} \lambda^{(j-1)-\tau} * \gamma_{i,\tau}^{c}}{\sum_{\tau=1}^{j-1} \lambda^{(j-1)-\tau}}$ preserves behavioral continuity by emphasizing recent purchasing tendencies, the tail-sensitive standard deviation $\beta_{i,j}^{c}$ introduces sensitivity to abrupt deviations in purchase timing that signal irregular shifts, and the category-specific bias term $b_{i,j}^{c}$ provides contextual adjustment, recognizing that baseline regularity differs inherently across product categories. $\gamma_{i,j}^{c}$ is calculated as follows:

$$\gamma_{i,j}^{c} = \sigma\left[w_1^c \cdot \left(\frac{\sum_{\tau=1}^{j-1} \lambda^{(j-1)-\tau} * \gamma_{i,\tau}^{c}}{\sum_{\tau=1}^{j-1} \lambda^{(j-1)-\tau}}\right) + w_2^c \cdot \beta_{i,j}^{c} + b_{i,j}^{c}\right] \quad (1)$$

$$\beta_{i,j}^{c} = \sqrt{\frac{1}{j-1}\sum_{\tau=1}^{j-1}(I_{i,\tau}^{c} - \bar{I}_{i}^{c})^{2} + \delta_{i,j}^{c}} \tag{2}$$

$$\delta_{i,j}^{c} = \frac{1}{j-1}\sum_{\tau=1}^{j-1}\left[\max\left(I_{i,\tau}^{c} - \rho_{i}^{c}, 0\right)\right]^{2} \tag{3}$$

where w_1^c, w_2^c, and $b_{i,j}^c$ are learnable parameters to control the balance between the weighted average of past irregularity states and temporal irregularity term for each category, and $\sigma[\cdot]$ is a sigmoid function that maps values to the range [0, 1]. The weighted average of past irregularity states is first scaled by a temporal decay factor $\lambda^{(j-1)-\tau}$: More recent purchase intervals are assigned higher weights, while older intervals are progressively downweighed by a decay factor. This ensures that the current irregularity state reflects not only the cumulative history of purchase behavior but also emphasizes the most recent deviations in purchasing patterns. $\beta_{i,j}^c$ is computed as the tail-sensitive standard deviation of past inter-purchase intervals up to time $t_{i,j-1}^c$, where $I_{i,\tau}^c$ denotes the inter-purchase interval for category c at τ and $\bar{I}_i^c$ is the mean of previous inter-purchase intervals. $\delta_{i,j}^c$ is used to emphasize the contribution of those rare and long inter-purchase time intervals. It ensures that a few long and infrequent inter-purchase intervals (e.g., when a customer buys a certain product category only occasionally) are not diluted by numerous regular and short intervals, thereby preventing the model from overfitting to high-frequency noise from short, erratic purchases and enabling it to better capture meaningful deviations. To decide what counts as "unusually long," we set a cutoff (e.g., 90th percentile) based on a chosen percentile ρ_i^c of the customer's past shopping intervals. Therefore, $\beta_{i,j}^c$ indicates subsequent irregularity based on the volatility of the customer's previous purchasing behavior. The learned irregularity state $\gamma_{i,j}^c$ is further combined with inter-purchase interval sequence $\boldsymbol{I}_i^c$ to serve as input for the category-level purchase pattern mining component.

3.3. Category-Level Purchase Pattern Mining

The category-level purchase pattern mining component builds multiple LSTMs to capture customers' purchase patterns of each product category; the category-specific LSTMs independently learn sequential patterns from past purchases of its corresponding product category. Each LSTM takes each consumer's inter-purchase interval sequence $\boldsymbol{I}_i^c$ and irregularity state $\gamma_{i,j}^c$ as input, allowing our method to learn not only how often purchases occur for a particular product category, but also how irregular they are. Specifically, the model processes all product categories in parallel, each with its own LSTM, denoted as $LSTM^{(c)}$. Let s denote the maximum sequence length, to predict the purchasing time at the next time step $s+1$, the $LSTM^{(c)}$ compares prior irregularity state with the irregularity state at time step $s+1$, given higher attention to inter-purchase intervals exhibiting similar irregularity state. The similarities between irregularity states are calculated using a Gaussian kernel:

$$\eta_{i,j}^{c} = \exp\left(\frac{-\left(\gamma_{i,s+1}^{c} - \gamma_{i,j}^{c}\right)^{2}}{2\phi^{2}}\right), j \in [1, 2, \dots s] \tag{4}$$

where ϕ is a hyperparameter that controls the smoothness of the kernel function, and the similarity score $\eta_{i,j}^c$ ranges between 0 and 1, with 1 indicating similarity and 0 denoting dissimilarity. The input of $LSTM^{(c)}$ is further weighted by the similarity score as:

$$x_{i,j}^{c} = I_{i,j}^{c} * \eta_{i,j}^{c}, j \in [1, 2, \dots s] \tag{5}$$

where $x_{i,j}^c$ is the input of $LSTM^{(c)}$ at step j. $LSTM^{(c)}$ then processes input $x_{i,j}^c$ as follows:

$$\left(h_{i,j}^{c}, o_{i,j}^{c}\right) = LSTM^{(c)}\left(x_{i,j}^{c}, h_{i,j-1}^{c}, o_{i,j-1}^{c}\right) \tag{6}$$

where $h_{i,j}^c$ is the hidden state and $o_{i,j}^c$ is the cell state. After processing s steps, $LSTM^{(c)}$ returns a sequence of hidden states $[h_{i,1}^c, h_{i,2}^c, \dots, h_{i,s}^c]$. The final hidden states of LSTM output will first be normalized using an $L2$ normalization function $f(\cdot)$ and then stacked as $\boldsymbol{H}_{i,s}^*$, which will be used as the input of the cross-category relationship learning component.

$$\boldsymbol{H}_{i,s}^{*} = f_{stack}\left(f(h_{i,s}^{1}), f(h_{i,s}^{2}), \dots, f(h_{i,s}^{C})\right) \tag{7}$$

3.4. Cross-category Relationship Learning

The cross-category relationship learning component builds on the economic theory of complementary and substitute goods (Deaton & Muellbauer, 1980). A complementary relationship occurs when the consumption of one product increases the utility of another (e.g., milk and cereal) (Yalcin et al., 2013). A substitute relationship arises when products fulfill similar needs and can replace one another (Milgrom & Strulovici, 2009; Walters, 1991). In consumer shopping behavior, these effects manifest dynamically over time: complements frequently co-occur within the same basket, whereas substitutes shape purchases across time (Mani et al., 2022). Building upon this, our method leverages the self-attention mechanism to model the associations among product categories. This approach aligns with this theory as attention weights naturally reflect the varying strengths of relationships between product categories, magnifying interactions when products are consumed jointly.

We apply a self-attention mechanism to $\boldsymbol{H}_{i,s}^{*}$ to capture the cross-category relationship. In this way, the model can learn how one purchase decision might inform another regarding shopping time intervals. Through the self-attention mechanism, $\boldsymbol{H}_{i,s}^{*}$ will be projected into three learnable spaces: queries $\boldsymbol{Q}_i$, keys $\boldsymbol{K}_i$, and values $\boldsymbol{V}_i$.

$$\boldsymbol{Q}_i = \boldsymbol{H}_{i,s}^{*} \cdot \boldsymbol{W}_Q \tag{8}$$

$$\boldsymbol{K}_i = \boldsymbol{H}_{i,s}^{*} \cdot \boldsymbol{W}_K \tag{9}$$

$$\boldsymbol{V}_i = \boldsymbol{H}_{i,s}^{*} \cdot \boldsymbol{W}_V \tag{10}$$

$$\boldsymbol{\alpha}_i^c = \mathrm{softmax}\left(\frac{\boldsymbol{q}_i^c \boldsymbol{K}_i^{T}}{\sqrt{H}}\right) \tag{11}$$

where $\boldsymbol{W}_Q, \boldsymbol{W}_K, \boldsymbol{W}_V$ are weight matrices for query, key, and value, respectively. The attention weight $\boldsymbol{\alpha}_i^c$ is further used to compute the output for each category as a weighted sum of the values $\boldsymbol{v}_i^c$.

$$\boldsymbol{z}_i^c = \sum_{c=1}^{C} \boldsymbol{\alpha}_i^c \boldsymbol{v}_i^c \tag{12}$$

$$\hat{y}_i^c = f'(\boldsymbol{z}_i^c) \tag{13}$$

The predicted purchase time $\hat{y}_i^c$ is further calculated based on $\boldsymbol{z}_i^c$ (equation 12), where $f'(*)$ is a linear transformation, projecting the attention layer's output $\boldsymbol{z}_i^c$ to a scalar value $\hat{y}_i^c$ to predict next purchase time for product category c. Our method is then trained based on a mean squared error (MSE) loss:

$$\mathcal{L}(y_i^c, \hat{y}_i^c) = \frac{1}{N}\sum_{i=1}^{N}(y_i^c - \hat{y}_i^c)^2 \tag{14}$$

4. Data

4.1. Data Collection

We evaluated the proposed model using a publicly available Kaggle dataset[3]. This dataset was collected from Instacart; it records detailed transactions of each consumer's purchase, with the time interval between two consecutive purchases available; each transaction also includes information on the product categories purchased by the consumer. There are 21 distinct product categories included in the dataset. Similar to (Guidotti et al., 2017), we exclude customers with purchase histories of fewer than 10 or more than 40 transactions to reduce sparsity and extreme cases. After applying this criterion, 1,586,576 transactions from 77,678 customers remain for evaluation.

4.2. Preliminary Data Analysis

We perform a preliminary analysis on this data with a focus on purchase count distribution and inter-purchase time intervals for each product category (i.e., the time interval between consecutive purchases for each product category). This analysis evaluates the degree of randomness inherent in the purchase behavior, thereby validating the rationale for incorporating category-level purchasing irregularity into our method design.

[3] https://www.kaggle.com/datasets/yasserh/instacart-online-grocery-basket-analysis-dataset

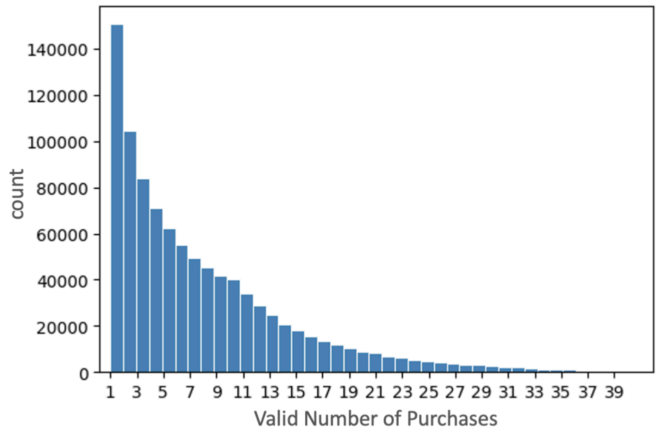

Figure 2. Valid Purchase Sequence Lengths for All User-Product Category Pairs

In our dataset, the purchase records are relatively sparse, and a substantial number of user-product category pairs have zero purchases because users do not purchase all product categories during the observation window. Figure 2 illustrates the distribution of observation counts for user–category pairs with at least one recorded purchase. From the data, we observe most sequences are relatively short, while longer sequences are concentrated in the most frequently purchased product categories (Figure 3). Figure 3 further illustrates this heterogeneity by presenting boxplots of valid sequence lengths for the five most and least frequently purchased product categories, summarizing their distributions and central tendencies (mean and median)[4]. Clear differences in customer engagement exist across product categories. With respect to purchase frequency, produce and dairy eggs exhibit larger median and mean values than other product categories (both above 10), shown in Figure 3(a), suggesting that these items are purchased more frequently over time and generate a greater number of purchase records. In contrast, categories such as alcohol, pets, and bulk display much lower central tendencies and result in fewer purchase records, shown in Figure 3(b).

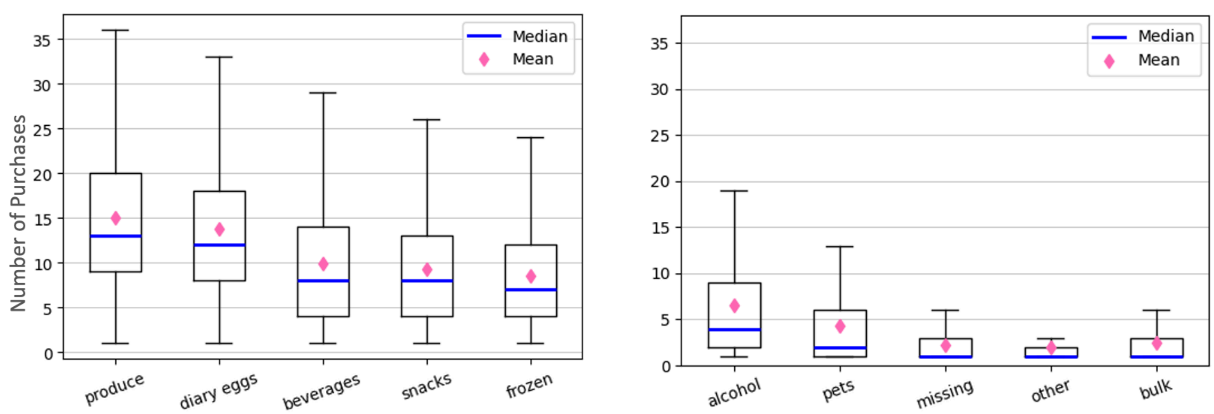

Figure 3(a). Purchase Frequency Distributions for the Five Most Purchased Product Categories

Figure 3(b). Purchase Frequency Distributions for the Five Least Purchased Product Categories

Figure 3. Purchase Frequency Distributions for the Most and Least Purchased Product Categories

Figure 4(a) shows the average number of days between consecutive purchases within each product category. Categories such as other, bulk, pets, international, and personal care have longer average intervals, indicating less frequent purchases. Conversely, categories like produce, dairy eggs, and snacks exhibit shorter intervals, suggesting more frequent purchases. Figure 4(b) presents the normalized entropy of time intervals per product category, which was calculated to quantify the irregularity of inter-purchase intervals across product categories. Interestingly, it

[4] In the original dataset, the product category field includes the labels “missing” and “other”. The category “missing” indicate no product category information was available. The category “other” aggregates products that do not fall into any of the predefined product categories due to the ambiguous classification.

displays an opposite pattern compared to Figure 4(a): categories with longer average intervals tend to have lower entropy (less diversity), suggesting relatively regular, though infrequent, purchasing behavior. In contrast, categories with shorter average intervals have higher entropy, indicating more variability or irregularity in purchasing timing. Frequently purchased categories (shorter average intervals such as produce, dairy eggs, and snacks etc.) tend to have highly variable purchasing behavior. These frequently purchased items are consumed regularly and likely influenced by fluctuations in daily consumption needs, promotions, and short-term events (e.g., hosting guests or holidays), which can lead to more irregular and volatile purchasing patterns. However, less frequently purchased categories (longer average intervals such as bulk, personal care, and pets) tend to have relatively more regular purchasing intervals. One possible explanation is that such products are usually purchased in larger quantities or for planned use, indicating more predictable purchasing cycles (e.g., monthly restocking of pet food or occasional alcohol purchases). As the need for replenishment is easier to plan, the variability in timing is lower.

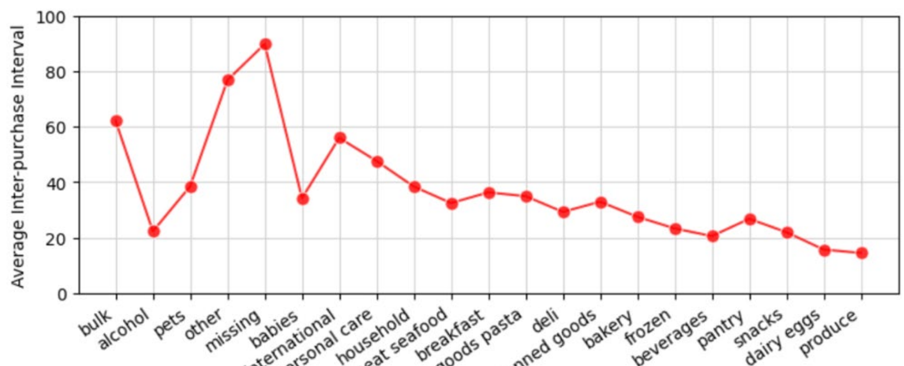

Figure 4(a). Average Shopping Time Intervals across Different Product Categories

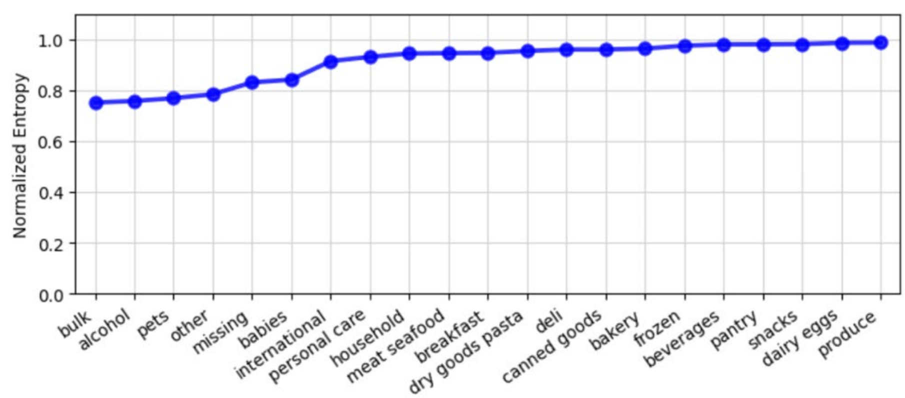

Figure 4(b). Normalized Entropy of Shopping Time Intervals across Different Product Categories

Figure 4. Average and Normalized Entropy of Shopping Time Intervals across Different Product Categories

5. Evaluation

5.1. Evaluation Design

To evaluate the prediction performance of the proposed method, we extracted each customer's last purchase history (i.e., transaction) as test data and utilized the remaining transactions for model training, mimicking the "next purchase time" prediction. This method ensures that the method is trained on historical data while being evaluated on unseen future purchase behavior.

Several prevalent benchmarks were also considered and included in the evaluation for comparison, including Graph_N2V, Graph_GAT, LSTMs_CNN, LSTMs_Linear, LSTMs_MLP, and LSTMs_AttnSC. These benchmark methods share a core LSTM (Hochreiter & Schmidhuber, 1997) component, which processes historical purchase data to capture unique recurring trends. Specifically, Graph_N2V is the benchmark using deep learning graphic method to model the cross-category interactions (Grover & Leskovec, 2016; Linden et al., 2003). It first constructs a co-purchase graph where two products are connected if they are frequently purchased together in the same order. Next, Node2Vec (Grover & Leskovec, 2016) is used to generate node embeddings for these product categories, and these embeddings further serve as inputs to an LSTM for next purchase time prediction on each product category. Graph_GAT builds the same co-purchase graph and then a GAT (graph attention network) learns node embeddings (Velickovic et al., 2017) to predict the next purchase time. LSTMs_CNN uses a convolutional neural network (CNN) after the LSTMs to extract and integrate the cross-category interactions through convolutional layers, capturing localized interaction patterns across different products and time intervals (Krizhevsky et al., 2012). This model treats each product category independently by assigning a separate LSTM to each. For LSTMs_Linear, a single linear layer is applied after the LSTMs to summarize the cross-category effects for purchase time prediction. LSTMs_MLP, on the other hand, replaces the linear layer with a multilayer perceptron (MLP), aiming to model these interactions more flexibly and nonlinearly. Last, LSTMs_AttnSC employs an attention layer to summarize cross-category interactions while incorporating stochastic random noise to reflect disruptions in shopping regularity (Vaswani et al., 2017). Specifically, it parameterizes a latent variable for each product category using learnable mean and log-variance parameters, enabling the model to adapt to irregular purchase timing variations explicitly.

All methods were evaluated by comparing their predicted next purchase times with the actual purchase times on product categories that were included in the next purchase. The proposed and benchmark methods were evaluated using user-averaged mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE), shown in equations 15-17.

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\sum_{c=1}^{C}|\hat{y}_i^c - y_i^c| \cdot m_i^c}{\sum_{c=1}^{C} m_i^c}\right) \tag{15}$$

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\sum_{c=1}^{C}(\hat{y}_i^c - y_i^c)^2 \cdot m_i^c}{\sum_{c=1}^{C} m_i^c}\right) \tag{16}$$

$$\mathrm{RMSE} = \frac{1}{N}\sum_{i=1}^{N}\sqrt{\frac{\sum_{c=1}^{C}(\hat{y}_i^c - y_i^c)^2 \cdot m_i^c}{\sum_{c=1}^{C} m_i^c}} \tag{17}$$

where y_i^c represents the actual next purchase time for user u_i and product category c, $\hat{y}_i^c$ denotes the predicted next purchase time, and m_i^c is the mask for category c, with $m_i^c = 1$ indicating product category c is included in user u_i's next purchase and 0 otherwise.

5.2. Prediction Performances of Each Method

Table 1. Prediction Results of the Proposed and Benchmark Methods

	MAE		**MSE**		**RMSE**	
	$mean$	std	$mean$	std	$mean$	std
Graph_N2V	17.398	0.120	1068.791	11.224	23.165	0.156
Graph_GAT	16.569	0.054	982.530	12.397	22.093	0.088
LSTMs_CNN	13.279	0.067	474.673	4.129	16.287	0.077
LSTMs_Linear	9.810	0.148	172.461	7.070	10.698	0.209
LSTMs_MLP	9.647	0.204	167.016	2.756	10.476	0.105
LSTMs_AttnSC	9.094	0.101	146.383	4.604	9.985	0.124
Proposed Method	**8.860**	0.047	**141.178**	3.863	**9.661**	0.054

Table 1 summarizes prediction performances of the proposed and benchmark methods. We repeat the evaluation 10 times with random parameter initializations and present the mean and standard deviation of each method. The results show that our method achieves the best performance over benchmark methods across all error metrics (MAE, MSE, and RMSE). We observe that methods based on a single LSTM have higher errors compared to those relying on multiple LSTMs. This is likely due to the unique purchasing patterns of different product categories that cannot be effectively captured by a single unified LSTM. Whereas dedicated LSTM models for each product category are capable of learning category-specific temporal relationships, thereby enhancing prediction accuracy (i.e., MAE ranges from 8.860 to 13.279 compared to 16.569–17.398). The proposed method has a mean MAE (8.860) which is substantially lower than benchmarks using multiple LSTMs: LSTM-CNN (13.279), LSTM-Linear (9.810), LSTMs_AttnSC (9.094), and LSTM-MLP (9.647). The results indicate that predicted times generated by the proposed method are closer to the actual purchase times on average, demonstrating better accuracy. The proposed method also demonstrates consistent improvements across all evaluation metrics compared to the best-performing baseline model (LSTMs_AttnSC). Specifically, the MAE decreases from 9.094 to 8.860, representing an improvement of approximately 2.6%. A similar trend holds for both MSE and RMSE, where the proposed method achieves the lowest mean (141.178 for MSE and 9.661 for RMSE). Compared to the best-performing baseline LSTMs_AttnSC, the MSE is reduced from 146.383 to 141.178, achieving about a 3.6% improvement, while the RMSE decreases from 9.985 to 9.661, corresponding to a 3.2% improvement. The results suggest fewer large errors in next purchase time predictions, highlighting improved reliability. In addition, our method maintains relatively small standard deviations across these metrics, suggesting more consistent performance.

5.3. Ablation Analysis

We apply a stacked ablation design so that we can isolate the contribution of each component (category-specific purchase irregularity and self-attention) while keeping the LSTM as the common backbone, shown in Figure 5.

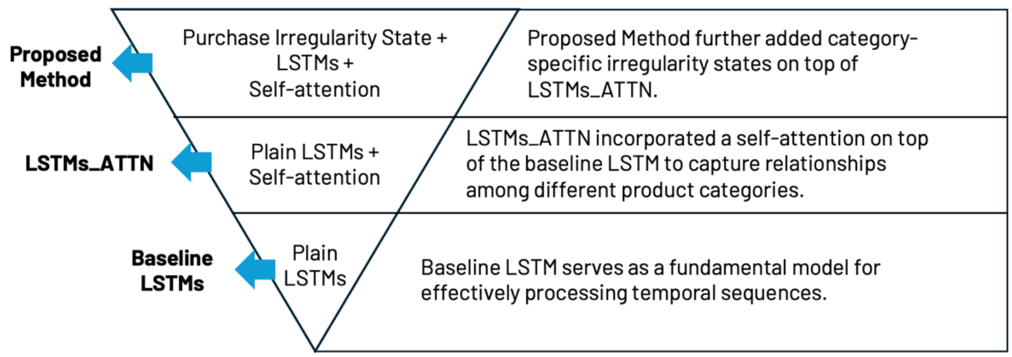

Figure 5. Ablations Analysis Framework.

The baseline LSTMs serve as a fundamental model that effectively processes temporal sequences, where both category-specific purchase irregularity and the self-attention mechanism are removed. LSTMs_ATTN extends the baseline LSTM by incorporating a self-attention mechanism to capture relationships among different product categories but still removes category-specific purchase irregularity. The Proposed Method builds on LSTMs_ATTN by further introducing category-specific purchase irregularity states, thereby having both self-attention and category-specific purchase irregularity incorporated.

Table 2. Prediction Outcomes based on Ablation Analysis

	MAE		MSE		RMSE	
	$mean$	std	$mean$	std	$mean$	std
Baseline LSTMs	9.852	0.154	176.947	5.150	10.841	0.219
LSTMs_ATTN	9.485	0.102	165.478	2.878	10.475	0.147
Proposed Method	8.860	0.047	141.178	3.863	9.661	0.054

The results in Table 2 demonstrate a clear progression in performance improvement as additional modeling components are introduced. Specifically, the baseline LSTMs model shows a relatively large MAE of 9.852 (with a standard deviation $std = 0.154$) and an MSE of 176.947 ($std = 5.150$). By considering the cross-category relationships (LSTMs_ATTN), the MAE decreases substantially to 9.485 ($std = 0.102$), and the MSE drops to 165.478 ($std = 2.878$). Last, the proposed method, which combines LSTM, self-attention, and purchase irregularity, further lowers the MAE to 8.860 ($std = 0.047$) and the MSE to 141.178 ($std = 3.863$). A similar trend is observed in the RMSE, decreasing from 10.841 (for LSTMs) to 10.475 (for LSTMs_ATTN) and 9.661 for the proposed approach. These results suggest that learning cross-category interaction and distinct purchasing patterns (routine versus irregular) are useful for predicting customers' next purchase time.

5.4. Attention Weight Interpretation

The heat map in Figure 6 illustrates the matrix of the pairwise attention weight learned by the self-attention mechanism in our model. It represents how different product categories interact when predicting the next purchase time. In this matrix, rows and columns denote queries $\boldsymbol{Q}_i$ and keys $\boldsymbol{K}_i$ corresponding to each category, respectively.

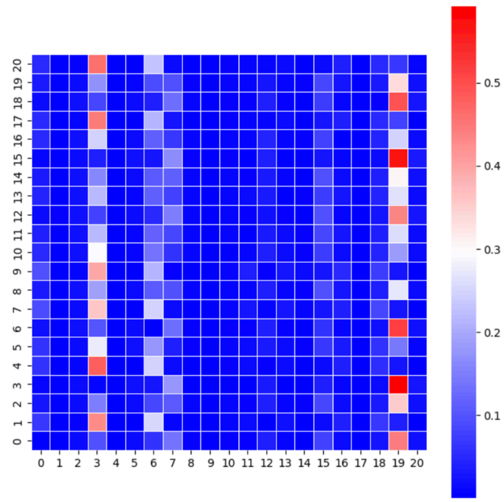

Figure 6. Attention Weights Matrix Explaining the Temporal Interactions among Product Categories

We observe several distinct vertical red bands appear around category indices 3 (produce), 6 (beverages), 7 (pets), and 19 (deli), shown in Figure 6. These product categories consistently receive strong attention from many others. This pattern demonstrates that these categories may act as temporal anchors, meaning their purchase timing provides useful contextual cues for predicting when other products are likely to be purchased. The attention mechanism enables the model to uncover hidden relationships among product categories from actual purchasing data. These relationships can be complex and context-dependent; for example, certain products may be purchased together under specific circumstances but not consistently across all shopping occasions. By relying on data, the model reveals how product categories dynamically influence one another in real-world purchasing behavior.

6. Contributions

Practical Implications: This study offers actionable insights with practical implications for businesses. With the ability to predict the timing of a consumer's next purchase by product category, the proposed method enables businesses to optimize inventory, enhance personalized marketing, and improve customer retention. Online grocery stores, such as Instacart, can utilize such predictions to manage perishable inventory more efficiently, ensuring that fresh produce and dairy products are stocked to meet customer demand while minimizing spoilage. With accurate purchase time prediction, large online vendors and platforms can improve inventory management and reduce shipping times by pre-positioning product categories in warehouses closer to customers who are likely to buy. Businesses can also use the predicted timing to send targeted promotions or reminders for replenishment at the right time. In this way, it increases the likelihood of conversion and improves customer satisfaction. In addition, our method allows businesses to identify irregularities in consumer behavior, such as delays in expected purchases. This could facilitate proactive outreach to retain customers.

Research Implications: In this study, we follow the framework of the cognitive-habit model and introduce purchase regularity and irregularity as methodological artifacts that represent a design innovation. The way we compute the purchase irregularity is theoretically grounded by the cognitive-habit framework, which distinguishes between routinized, habit-driven purchases and irregular purchases that arise when habitual cues are disrupted. Our design operationalizes this distinction by tracking how an individual customer's purchasing behavior shifts from regularity toward irregularity at the product-category level. The evaluation of the proposed method demonstrates that translating a well-established behavioral framework into a computational representation using large-scale behavioral data yields improved predictive performance in forecasting the timing of subsequent purchases. Future studies could explore integrating richer external contextual factors, such as promotional events and customer demographic information, into purchase irregularity modeling. This could capture granular shopping behavior and lead to more contextually relevant predictions. In addition, our study includes cross-category relationship modeling by using a self-attention mechanism. Future methods could advance the self-attention mechanism by incorporating hierarchical frameworks. Such extensions would capture deeper interactions among business entities (e.g., products, users, service providers, etc.) for enhanced purchase behavior prediction. Finally, as deep-learning models increase in complexity, our method embraces interpretability. Developing deep learning methods that explain model predictions in terms of customer-level insights, such as reasons behind purchasing irregularity or cross-category interactions, can enhance usability by decision-makers.

7. Conclusion

In this work, we propose a novel deep learning-based method to predict the timing of consumers' subsequent purchases across different product categories in online retail environments. This approach enhances the accuracy and interpretability of purchase time prediction by addressing critical challenges such as disrupted shopping regularity and overlooked cross-category influences. We first introduce purchase irregularity, a latent state that represents whether a purchase in each product category is likely to follow a routine pattern or not. The following integration of category-specific LSTMs enables the model to capture unique purchasing rhythms across product types while incorporating stochastic noise to reflect the inherent randomness in consumer behavior. Finally, the self-attention mechanism effectively models inter-category interactions, revealing how purchase behavior in one category can inform purchase timing predictions in others. The evaluation results on large-scale e-commerce datasets demonstrate that the proposed model outperforms existing baselines, offering actionable insights for demand forecasting, personalized marketing, and inventory planning. By anticipating when consumers are likely to make their next purchase, businesses can improve inventory management, make more informed decisions, and deliver more timely promotions.

This study has several limitations. First, the dataset used in this study primarily contains transactional data; therefore, the proposed method might not adequately account for consumer-specific factors beyond historical purchase patterns, such as demographics, lifestyle, or external events influencing purchasing decisions. Future research could collect more comprehensive data, including consumer-specific attributes and external behavior indicators, to improve predictive accuracy. Second, another limitation is that the dataset has been anonymized with respect to calendar dates and thus doesn't contain detailed transaction timestamps. Only interval-based information is available. As a result, we are unable to fully capture temporal cycles such as weekly or seasonal shopping behaviors, or to explicitly model the impact of holiday promotions. Future work could address this if richer temporal information is accessible. Third, the performance of the proposed method still largely relies on sufficient historical transaction data, limiting its effectiveness for new or infrequent customers. Future work may focus on investigating techniques such as transfer learning that could address limitations related to insufficient historical data for new customers.

REFERENCES

Ahmed, F., Cui, Y., Fu, Y., & Chen, W. (2021). A graph neural network approach for product relationship prediction. International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, V03AT03A036.

Ajzen, I. (1991). The Theory of planned behavior. *Organizational Behavior and Human Decision Processes*.

Anıl, U., & Akcayol, M. A. (2020). Deep learning based prediction model for the next purchase. *Advances in Electrical and Computer Engineering*, *20*(2).

Azizah, F. D., Nur, A. N., & Putra, A. H. P. K. (2022). Impulsive buying behavior: Implementation of IT on technology acceptance model on E-Commerce purchase decisions. *Golden Ratio of Marketing and Applied Psychology of Business*, *2*(1), 58-72.

Bawa, K., & Ghosh, A. (1991). The covariates of regularity in purchase timing. *Marketing letters*, *2*(2), 147-157.

Bhattacherjee, A. (2001). An empirical analysis of the antecedents of electronic commerce service continuance. *Decision Support Systems*, *32*(2), 201-214.

Bodapati, A. V. (2008). Recommendation systems with purchase data. *Journal of marketing research*, *45*(1), 77-93.

Chaudhuri, N., Gupta, G., Vamsi, V., & Bose, I. (2021). On the platform but will they buy? Predicting customers' purchase behavior using deep learning. *Decision Support Systems*, *149*, 113622.

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 785-794.

Chintagunta, P. K., & Haldar, S. (1998). Investigating purchase timing behavior in two related product categories. *Journal of marketing research*, *35*(1), 43-53.

Deaton, A., & Muellbauer, J. (1980). *Economics and consumer behavior*. Cambridge university press.

Dholakia, U. M. (2000). Temptation and resistance: An integrated model of consumption impulse formation and enactment. *Psychology & Marketing*, *17*(11), 955-982.

Droomer, M., & Bekker, J. (2020). Using machine learning to predict the next purchase date for an individual retail customer. *South African Journal of Industrial Engineering*, *31*(3), 69-82.

Dubé, J. P., Hitsch, G. J., & Rossi, P. E. (2010). State dependence and alternative explanations for consumer inertia. *The RAND Journal of Economics*, *41*(3), 417-445.

Eid, M. I. (2011). DETERMINANTS OF E-COMMERCE CUSTOMER SATISFACTION, TRUST, AND LOYALTY IN SAUDI ARABIA. *Journal of Electronic Commerce Research*, *12*(1).

Gabel, S., & Timoshenko, A. (2022). Product choice with large assortments: A scalable deep-learning model. *Management science*, *68*(3), 1808-1827.

Gefen, D. (2000). E-commerce: the role of familiarity and trust. *Omega*, *28*(6), 725-737.

Grigoraş, A., & Leon, F. (2023). Transformer-based model for predicting customers' next purchase day in e-commerce. *Computation*, *11*(11), 210.

Grover, A., & Leskovec, J. (2016). node2vec: Scalable feature learning for networks. Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining, 855-864.

Gupta, S. (1991). Stochastic models of interpurchase time with time-dependent covariates. *Journal of marketing research*, *28*(1), 1-15.

Ha, S. H., Bae, S. M., & Park, S. C. (2002). Customer's time-variant purchase behavior and corresponding marketing strategies: an online retailer's case. *Computers & Industrial Engineering*, *43*(4), 801-820.

Heckman, J. J. (1981). Heterogeneity and state dependence. In *Studies in labor markets* (pp. 91-140). University of Chicago Press.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735-1780.

Huang, C., Wu, X., Zhang, X., Zhang, C., Zhao, J., Yin, D., & Chawla, N. V. (2019). Online purchase prediction via multi-scale modeling of behavior dynamics. Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining, 2613-2622.

Kahn, B. E., & Schmittlein, D. C. (1989). Shopping trip behavior: An empirical investigation. *Marketing letters*, *1*, 55-69.

Keerthika, K., & Saravanan, T. (2020). Enhanced product recommendations based on seasonality and demography in ecommerce. 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), 721-723.

Kim, B.-D., & Park, K. (1997). Studying patterns of consumer's grocery shopping trip. *Journal of retailing*, *73*(4), 501-517.

Kim, S., & Eastin, M. S. (2011). Hedonic tendencies and the online consumer: An investigation of the online shopping process. *Journal of Internet Commerce*, *10*(1), 68-90.

Kim, S., Shin, W., & Kim, H.-W. (2024). Predicting online customer purchase: The integration of customer characteristics and browsing patterns. *Decision Support Systems*, *177*, 114105. https://doi.org/https://doi.org/10.1016/j.dss.2023.114105

Koschate-Fischer, N., Hoyer, W. D., Stokburger-Sauer, N. E., & Engling, J. (2018). Do life events always lead to change in purchase? The mediating role of change in consumer innovativeness, the variety seeking tendency, and price consciousness. *Journal of the academy of marketing science*, *46*(3), 516-536.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, *25*.

Linden, G., Smith, B., & York, J. (2003). Amazon. com recommendations: Item-to-item collaborative filtering. *IEEE Internet computing*, *7*(1), 76-80.

Lismont, J., Ram, S., Vanthienen, J., Lemahieu, W., & Baesens, B. (2018). Predicting interpurchase time in a retail environment using customer-product networks: An empirical study and evaluation. *Expert Systems with Applications*, *104*, 22-32. https://doi.org/https://doi.org/10.1016/j.eswa.2018.03.016

Liu, W., Zhang, Y., Wang, J., He, Y., Caverlee, J., Chan, P. P., Yeung, D. S., & Heng, P.-A. (2021). Item relationship graph neural networks for e-commerce. *IEEE Transactions on Neural Networks and Learning Systems*, *33*(9), 4785-4799.

Liu, X., Wu, X., Shi, W., Tong, W., & Ye, Q. (2022). The impacts of electronic word-of-mouth on high-involvement product sales: moderating effects of price, brand origin, and number of customers. *Journal of Electronic Commerce Research*, *23*(3), 177-189.

Liu, Z., & Ma, X. (2019). Predictive Analysis of User Purchase Behavior based on Machine Learning. *International Journal of Smart Business and Technology*, *7*(1), 45-56.

Mani, V., Thomas, D. J., & Bansal, S. (2022). Estimating substitution and basket effects in retail stores: Implications for assortment planning. *Management science*, *68*(7), 5002-5024.

Manzoor, E., & Akoglu, L. (2017). Rush! targeted time-limited coupons via purchase forecasts. Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1923-1931.

Martínez, A., Schmuck, C., Pereverzyev Jr, S., Pirker, C., & Haltmeier, M. (2020). A machine learning framework for customer purchase prediction in the non-contractual setting. *European Journal of Operational Research*, *281*(3), 588-596.

Mela, C. F., Jedidi, K., & Bowman, D. (1998). The long-term impact of promotions on consumer stockpiling behavior. *Journal of marketing research*, *35*(2), 250-262.

Milgrom, P., & Strulovici, B. (2009). Substitute goods, auctions, and equilibrium. *Journal of Economic theory*, *144*(1), 212-247.

Park, C., & Lee, T. M. (2009). Antecedents of online reviews' usage and purchase influence: an empirical comparison of US and Korean consumers. *Journal of interactive Marketing*, *23*(4), 332-340.

Pavlou, P. A. (2003). Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model. *International Journal of Electronic Commerce*, *7*(3), 101-134.

Platzer, M., & Reutterer, T. (2016). Ticking away the moments: Timing regularity helps to better predict customer activity. *Marketing science*, *35*(5), 779-799.

Qiu, J., Lin, Z., & Li, Y. (2015). Predicting customer purchase behavior in the e-commerce context. *Electronic commerce research*, *15*, 427-452.

Rekasiute, L., Palenzuela, A. J. J., Salnaite, N., Cuenca, R. C., & Frasincar, F. (2022). Predicting the category of customers' next product to buy in web shops. Proceedings of the 37th ACM/SIGAPP Symposium on Applied Computing, 1862-1871.

Reutterer, T., Platzer, M., & Schröder, N. (2021). Leveraging purchase regularity for predicting customer behavior the easy way. *International Journal of Research in Marketing*, *38*(1), 194-215. https://doi.org/https://doi.org/10.1016/j.ijresmar.2020.09.002

Saito, R., Otake, K., & Namatame, T. (2021). Customer Visit Prediction Using Purchase Behavior and Tendency. International Conference on Human-Computer Interaction, 120-129.

Sakar, C. O., Polat, S. O., Katircioglu, M., & Kastro, Y. (2019). Real-time prediction of online shoppers' purchasing intention using multilayer perceptron and LSTM recurrent neural networks. *Neural Computing and Applications*, *31*(10), 6893-6908.

Seetharaman, P., Ainslie, A., & Chintagunta, P. K. (1999). Investigating household state dependence effects across categories. *Journal of marketing research*, *36*(4), 488-500.

Shah, D., Kumar, V., & Kim, K. H. (2014). Managing customer profits: The power of habits. *Journal of Marketing Research*, *51*(6), 726-741.

Song, H. S. (2017). Comparison of performance between MLP and RNN model to predict purchase timing for repurchase product. *Journal of Information Technology Applications and Management*, *24*(1), 111-128.

Guidotti, R., Rossetti, G., Pappalardo, L., Giannotti, F., & Pedreschi, D. (2017). Next basket prediction using recurring sequential patterns. *arXiv preprint arXiv*:1702.07158.

Unal, M., & Park, Y.-H. (2023). Fewer clicks, more purchases. *Management science*, *69*(12), 7317-7334.

Vakratsas, D., & Bass, F. M. (2002). The relationship between purchase regularity and propensity to accelerate. *Journal of retailing*, *78*(2), 119-129.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*.

Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., & Bengio, Y. (2017). Graph attention networks. *stat*, *1050*(20), 10-48550.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management science*, *46*(2), 186-204.

Verma, A. (2020). Consumer behaviour in retail: next logical purchase using deep neural network. *arXiv preprint arXiv:2010.06952*.

Virinchi, S., Saladi, A., & Mondal, A. (2022). Recommending related products using graph neural networks in directed graphs. Joint European Conference on Machine Learning and Knowledge Discovery in Databases, 541-557.

Walters, R. G. (1991). Assessing the impact of retail price promotions on product substitution, complementary purchase, and interstore sales displacement. *Journal of marketing*, *55*(2), 17-28.

Wang, E. S.-T., & Chou, N. P.-Y. (2014). Consumer characteristics, social influence, and system factors on online group-buying repurchasing intention. *Journal of Electronic Commerce Research*, *15*(2), 119-132.

Wang, J., Louca, R., Hu, D., Cellier, C., Caverlee, J., & Hong, L. (2020). Time to Shop for Valentine's Day: Shopping Occasions and Sequential Recommendation in E-commerce. Proceedings of the 13th International Conference on Web Search and Data Mining, 645-653.

Wang, W., Xiong, W., Wang, J., Tao, L., Li, S., Yi, Y., Zou, X., & Li, C. (2023). A user purchase behavior prediction method based on xgboost. *Electronics*, *12*(9), 2047.

Wood, W., & Rünger, D. (2016). Psychology of habit. *Annual review of psychology*, *67*(1), 289-314.

Yalcin, T., Ofek, E., Koenigsberg, O., & Biyalogorsky, E. (2013). Complementary goods: Creating, capturing, and competing for value. *Marketing science*, *32*(4), 554-569.

Yang, J.-S., & Pan, J. C.-H. (2004). Just-in-time purchasing: an integrated inventory model involving deterministic variable lead time and quality improvement investment. *International Journal of Production Research*, *42*(5), 853-863.

Zhang, Y., & Pennacchiotti, M. (2013). Predicting purchase behaviors from social media. Proceedings of the 22nd international conference on World Wide Web, 1521-1532.