AdaRec: Adaptive Recommendation with LLMs via Narrative Profiling and Dual-Channel Reasoning

Meiyun Wang [1] **Charin Polpanumas** [2]

Abstract

We propose **AdaRec**, a few-shot in-context learning framework that leverages Large Language Models (LLMs) for an adaptive personalized recommendation. AdaRec introduces **narrative profiling**, transforming user-item interactions into natural language representations to enable unified task handling and enhance human readability. Centered on a **bivariate reasoning paradigm**, AdaRec employs a **dual-channel architecture** that integrates **horizontal behavioral alignment**—discovering peer-driven patterns—with **vertical causal attribution**—highlighting decisive factors behind user preferences. Unlike existing LLM-based approaches, AdaRec **eliminates manual feature engineering** through semantic representations and supports **rapid cross-task adaptation** with minimal supervision. Experiments on real e-commerce datasets demonstrate that AdaRec outperforms both machine learning models and LLM-based baselines by up to **8%** in few-shot settings. In zero-shot scenarios, it achieves up to a **19%** improvement over expert-crafted profiling, showing effectiveness for long-tail personalization with minimal interaction data. Moreover, lightweight fine-tuning on synthetic data generated by AdaRec matches the performance of fully fine-tuned models, highlighting its efficiency and generalization across diverse tasks.[1].

1. Introduction

Recommender systems are critical to e-commerce, social media, and digital services, where delivering personalized content demands continuous adaptation to dynamic user preferences and evolving contexts (Hansen et al., 2020). Traditional approaches, including collaborative filtering (He et al., 2017) and feature-based machine learning models (Weng & Liu, 2004; Wang et al., 2015), have achieved early success but struggle with manual feature engineering, limited generalization, and inflexibility in responding to changing user behaviors (Hu et al., 2020).

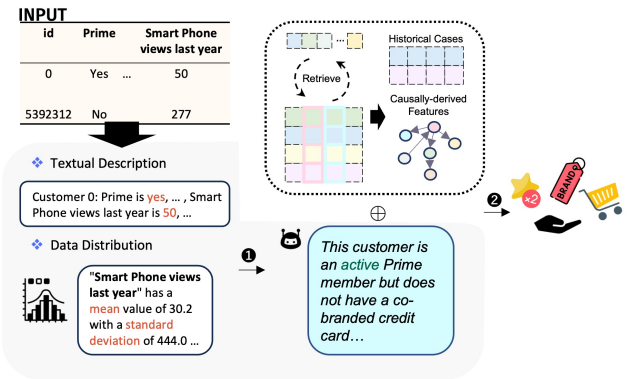

Figure 1: AdaRec: Few-shot adaptive recommendation with LLMs via narrative profiling and dual-channel reasoning.

The rise of Large Language Models (LLMs) introduces new opportunities for recommender systems by leveraging semantic understanding and in-context learning capabilities (Wu et al., 2024). Current LLM-driven recommendation research follows four main directions: narrative-based methods generate static textual profiles but lack dynamic adaptability to evolving user preferences (Mysore et al., 2023); similarity-based approaches retrieve historical cases without causal reasoning, limiting robustness to data shifts (Xie et al., 2024); agent-based frameworks improve personalization via interaction but incur high operational costs (Wang et al., 2023); and methods such as fine-tuning-based techniques, multi-modal personalization systems, and feedback-driven exploration strategies (Lin et al., 2024; Zhang et al., 2024) offer strong task-specific performance by tightly integrating LLMs into recommendation workflows. However, these methods rely on substantial computational resources and complex training pipelines. This highlights the need for a **lightweight**, **adaptable**, and **explainable** solution that

[1]The University of Tokyo [2]Amazon.com, Inc.. Correspondence to: Meiyun Wang <omiun20@g.ecc.u-tokyo.ac.jp>.

Proceedings of the 1st ICML Workshop on Foundation Models for Structured Data, Vancouver, Canada. 2025. Copyright 2025 by the author(s).

[1]Codes are available: https://anonymous.4open.science/r/AdaRec-CE5C

minimizes task-specific engineering overhead while maintaining high recommendation quality.

To address these challenges, we propose **AdaRec**, an LLM-based few-shot recommendation framework for adaptive personalization in tabular data. AdaRec introduces **narrative profiling**, converting user-item interactions into natural language to enhance flexibility and human-readable reasoning. It further employs a **dual-channel reasoning** architecture, integrating **horizontal behavioral alignment** with **vertical causal attribution** to combine peer pattern discovery and dynamic feature focus for robust, explainable recommendations.

The key contributions of this work are: (1) **Efficient zero-shot and few-shot adaptation**, eliminating costly fine-tuning and manual feature engineering; (2) **Enhanced robustness and explainability** through integrated causal reasoning and narrative profiling; (3) **Seamless cross-task generalization**, enabling rapid deployment across diverse recommendation scenarios.

Experiments on real-world e-commerce datasets show that AdaRec consistently outperforms traditional machine learning models and recent LLM-based baselines, achieving up to **8%** improvement on benchmark datasets and up to **19%** gains in zero-shot settings over expert-crafted profiling. These results demonstrate that AdaRec is a practical and adaptable solution for recommendation challenges, balancing high performance with reduced manual effort and improved task flexibility.

2. Related Work

Before the rise of Large Language Models (LLMs), recommendation systems for tabular data focused on explicit user behavior modeling and feature interaction design. Representative methods include DIN (Zhou et al., 2018) for adaptive interest extraction, DIEN (Zhou et al., 2019) for temporal preference modeling, MIMN (Pi et al., 2019) and SIM (Pi et al., 2020) for long-term behavior handling, and STAR (Sheng et al., 2021) for multi-domain adaptation. While effective, these approaches rely heavily on manual feature engineering and task-specific architectures, limiting flexibility and explainability.

LLMs introduce new opportunities by addressing data sparsity, enhancing interaction via natural language, and offering strong generalization without complex feature processing. Fine-tuning methods like InstructRec (Zhang et al., 2023) and TALLRec (Bao et al., 2023) improve task-specific performance but demand high computational costs and retraining. Lightweight solutions such as RecMind (Wang et al., 2023), narrative-driven frameworks (Mysore et al., 2023), and similarity-based recommenders (Xie et al., 2024) leverage in-context learning (ICL) and external knowledge but lack explainability. AdaRec addresses these gaps by combining LLM-driven flexibility and explainability with dual-channel reasoning, enabling lightweight, adaptive recommendations for tabular data without extensive fine-tuning or manual intervention.

3. AdaRec

AdaRec is a lightweight in-context learning framework that transforms tabular user data into natural language profiles via narrative profiling, enhancing recommendations through structured reasoning. By combining similarity retrieval and causal inference, AdaRec delivers personalized and explainable recommendations.

Narrative Profiling. For each user feature vector $\mathbf{x}_\theta \in \mathbb{R}^d$, AdaRec leverages LLMs to generate context-aware natural language descriptions. Instead of predefined rules, AdaRec provides global statistical distributions and user-specific data via carefully designed prompts. The LLM interprets each feature f_i with value $x_{\theta,i}$ based on its distribution $\mathcal{D}_i$, generating context-aware qualitative descriptions: $\Psi(f_i, x_{\theta,i}, \mathcal{D}_i) \rightarrow \mathcal{T}$, where $\mathcal{T}$ denotes the natural language description space, and Ψ represents the LLM-driven mapping function that contextualizes feature values within their statistical distributions. Table 4 shows an example of the narrative profiling prompt, where the input includes raw feature values with exact numbers. Figure 3 illustrates the generated narrative profiles, where numerical features are described using relative trends and contextualized language.

Historical Case Discovery. For a target user feature vector $\mathbf{x}_\theta$, we compute cosine similarity with each user feature vector $\mathbf{x}_j$ in the training set as $\mathcal{S}(\mathbf{x}_\theta, \mathbf{x}_j) = \frac{\mathbf{x}_\theta \cdot \mathbf{x}_j}{\|\mathbf{x}_\theta\| \|\mathbf{x}_j\|}$.

We first select the top-η_1 most similar users to compute mutual information (MI) for feature importance weighting. Then, we select the top-η_2 users ($\eta_2 < \eta_1$) as the reference set for causal structure learning using FCI.[2] From this set, we select representative cases $\mathcal{H}$ for few-shot reasoning, forming the foundation of **Pattern Analysis**.

Causal Structure Learning. On the reference set (top-η_2 users), we apply the Fast Causal Inference (FCI) algorithm (Spirtes et al., 1995) to discover causal relationships between features $\mathbb{F}$ and the target variable $y \in \{0, 1\}$. To quantify the importance of each causal feature $f_i \in \mathbb{F}$ connected to y, we leverage the previously computed mutual information scores. We select the top-p causal features as $\mathcal{F}_c$, ensuring downstream reasoning focuses on key behavioral drivers.

These causal features provide focused guidance for the subsequent reasoning process and are integrated into the struc-

[2]Top-η_2 users are selected to balance FCI speed and performance.

tured reasoning prompt as the basis for **Factor Analysis**.

Structured Reasoning Framework. The structured reasoning prompt in AdaRec combines four components: the task description T, **Factor Analysis** (from causal features $\mathcal{F}_c$), **Pattern Analysis** (from representative cases $\mathcal{H}$), and the narrative profile $\Phi(\mathbf{x}_\theta)$. These elements guide the LLM to generate personalized and explainable recommendations: $R = \text{LLM}\Big(\xi\big(T, \mathcal{F}_c, \mathcal{H}, \Phi(\mathbf{x}_\theta)\big)\Big)$, where $\xi(\cdot)$ constructs a coherent reasoning input, and R denotes the generated recommendation output. Table 5 shows an example of the structured reasoning prompt.

4. Experiments

4.1. Tasks

We evaluate AdaRec on two tasks: Customer Response Prediction (binary classification) and Brand Recommendation (top-3 selection from 17 brands). For Customer Response Prediction, the dataset contains 5,000 training samples, 600 validation samples, and 600 test samples, each with 119 features (115 numeric, 4 categorical). For Brand Recommendation, we use 4,692 training samples, 587 validation samples, and 587 test samples, each with 111 features (103 numeric, 8 categorical).

4.2. Evaluation Metrics

We report Precision, Recall, and F1 score for customer response prediction. For brand recommendation, we adopt Expected CTR (Li et al., 2011), which measures the overlap between predicted and ground truth brand sets, weighted by click behaviour.

4.3. Models

Baselines. - **LightGBM** (Ke et al., 2017): Trained on 1.14M historical samples with 119 features for customer response prediction.[3] - **Hierarchical RNN** (Du et al., 2015): For brand recommendation, trained on 2B+ user-brand historical interactions with 111 features. - **MINT** (Mysore et al., 2023): Leverages LLMs to generate synthetic narrative queries from user-item interaction, enabling retrieval-based narrative-driven recommendations. - **NBCRS** (Xie et al., 2024): Recommends by reusing crowd-written answers linked to similar conversational contexts, combining neighbourhood-based retrieval for an efficient conversational recommendation. - **RecMind** (Wang et al., 2024): An autonomous LLM-powered agent that utilizes external knowledge and a self-inspiring planning algorithm to provide personalized recommendations. **Expert Profiling**: Manually designed customer profiles by marketing experts based on 26 selected key features, providing static and quantitative summaries.

Figure 3 highlights the difference between expert profiling and AdaRec's narrative profiling.

Backbone Models. AdaRec is evaluated using Claude-3.5-Sonnet, Llama-3.1-70B, and Qwen-2.5-32B with greedy decoding (temperature = 0) and a maximum of 1000 output tokens.

Parameters. We set $k = 5$ for the number of representative cases used in few-shot reasoning. We set $\eta_1 = 2000$, $\eta_2 = 1000$, use a significance level of $\alpha = 0.1$, and select up to $p = 15$ causal features in the causal analysis phase. To evaluate AdaRec's adaptability across different tasks, we apply supervised fine-tuning (SFT) for the Customer Response Prediction task (2 epochs, learning rate = 1×10^{-4}) and Kahneman-Tversky Optimization (KTO) for the Brand Recommendation task (2 epochs, learning rate = 5×10^{-5}). All experiments are conducted on 4 NVIDIA V100 GPUs.

4.4. Results

RQ1: How effective is our approach compared to baselines? Table 1 shows that AdaRec consistently outperforms all baselines across both tasks and profiling strategies. On the Customer Response Prediction task, AdaRec with Claude3.5 achieves an F1 score of 94.33%, exceeding the strongest ML baseline (LightGBM at 86.67%) by **8%**. For the Brand Recommendation task, AdaRec with Qwen2.5 reaches a CTR of 10.3%, outperforming all baselines.

In zero-shot settings, our narrative profiling shows significant advantages. For example, using Qwen2.5, the narrative profiling approach achieves an F1 score of 74.13%, clearly surpassing expert profiling (55.58%) with a **19%** relative improvement.

Meanwhile, AdaRec performs well with both expert-designed and LLM-generated narrative profiles. In the simpler Customer Response Prediction task, Narrative + AdaRec achieves performance close to Expert + AdaRec (e.g., 91% vs. 94% F1 with Claude3.5). In contrast, for the more challenging Brand Recommendation task, Narrative + AdaRec outperforms Expert + AdaRec, with Qwen2.5 achieving a CTR of 10.3% compared to 9.5%. These results indicate that narrative profiling not only reduces manual feature engineering but also enhances performance in complex recommendation scenarios.

RQ2: Can AdaRec generalize across tasks without retraining? To assess AdaRec's ability to capture transferable customer representations, we conduct a cross-task evaluation: fine-tuning the model on one task (e.g., Customer Response Prediction), then directly applying it to another task (e.g., Brand Recommendation).

[3] ML baselines use >1M samples; LLM-based models use fewer due to inference cost.

Model	strategy	n-shot	CustomerResponse			BrandRec
			P	R	F1	CTR
	ML Baseline		86.67	86.67	86.67	8.57
	MINT		74.83	74.84	74.83	8.4
	NBCRS		84.01	84.01	84	8.9
	RecMind		79.25	78.62	78.54	9
Llama3.1	expert	0	70.30	55.29	44.98	8.1
	AdaRec(expert)	5	90.01	88.57	88.40	8.3
	narrative	0	75.40	68.18	65.62	8.1
	AdaRec(narrative)	5	90.38	90.03	89.98	8.5
Qwen2.5	expert	0	74.53	61.58	55.58	9.2
	AdaRec(expert)	5	92.72	92.66	92.66	9.5
	narrative	0	74.36	74.20	74.13	9.8
	AdaRec(narrative)	5	91.65	90.83	90.91	**10.3**
Claude3.5	expert	0	75.57	63.40	58.30	8.3
	AdaRec(expert)	5	**94.42**	**94.35**	**94.33**	8.3
	narrative	0	77.56	74.12	73.19	8.4
	AdaRec(narrative)	5	91.25	91.03	90.99	8.7

Table 1: Performance Comparison of Different Models and Strategies. All metrics are shown in percentages (%). The best results are in bold.

As shown in Table 2, AdaRec maintains strong performance across tasks. In the zero-shot setting, the F1 score on Customer Response Prediction drops only slightly (from 75.7% to 73.5%), and Brand Recommendation CTR remains stable (9.0% vs. 9.1%). In the 5-shot setting, cross-task results closely match the in-task upper bound: 87.4% vs. 87.6% F1 and 9.7% vs. 9.7% CTR.

These results show that AdaRec learns generalizable representations of customer behavior that transfer well across tasks. The narrative profile supports robust reasoning across settings, demonstrating its potential to eliminate the need for repeated training and feature redesign in multi-objective customer modeling.

Setting	Model	n-shot	CustomerResponse (F1)			BrandRec (CTR)
			P	R	F1	CTR
Same-task evaluation (fine-tuned and tested on the same task)						
In-task	Qwen2.5	0	75.76	75.69	75.65	9.0
In-task	Qwen2.5	5	88.44	87.62	87.59	9.7
Cross-task evaluation (fine-tuned on one task, tested on another)						
Cross-task	Qwen2.5	0	73.72	73.53	73.45	9.1
Cross-task	Qwen2.5	5	88.37	87.44	87.40	9.7

Table 2: Cross-task generalization: AdaRec is fine-tuned on one task and directly evaluated on another. Results show minimal performance degradation.

RQ3: How robust is AdaRec to variations in narrative profiles? We evaluate AdaRec's robustness by using narrative profiles generated by different LLMs (Claude3.5, Llama3.1, Qwen2.5). As shown in Table 3, AdaRec consistently delivers strong performance regardless of profiling source.

In the zero-shot setting, Customer Response Prediction F1 scores remain stable (71.9% to 74.9%), with Qwen2.5-generated profiles achieving the highest CTR (9.8%) for Brand Recommendation.

In the 5-shot setting, AdaRec achieves over 90% F1 across all profiling sources, with Claude3.5 profiles slightly outperforming others in Customer Response Prediction. CTR values remain competitive (8.9% to 9.8%), confirming AdaRec's ability to effectively leverage diverse narrative profiles.

These results demonstrate that AdaRec is robust to variations in narrative profile generation, ensuring reliable performance across different LLM providers and adaptable to various business environments.

Profiling Source	CustomerResponse (F1)	BrandRec (CTR)	n-shot
Claude3.5	74.9 / 94.4	8.5 / 9.8	0 / 5
Llama3.1	71.9 / 93.9	8.1 / 8.9	0 / 5
Qwen2.5	74.1 / 90.9	9.8 / 9.8	0 / 5

Table 3: AdaRec performance using narrative profiles from different LLMs. Metrics are shown as 0-shot / 5-shot.

Figure 2: The ablation study of AdaRec.

RQ4: How do different components contribute to AdaRec's performance? An ablation study (Figure 2) shows that narrative profiling provides a solid baseline (F1: 74.1%, CTR: 9.8%), causal feature weights offer moderate gains (F1: 76.0%, CTR: 9.8%), and historical behavior patterns deliver the largest improvement (F1: 90.9%, CTR: 10.3%). These results highlight the key role of historical patterns and the synergy of all components in achieving AdaRec's high performance.

5. Conclusion

This paper presents AdaRec, an adaptive recommendation framework that leverages LLMs to generate narrative profiles from user-item interactions for flexible and explainable personalization. Its dual-channel architecture, combining causal reasoning and similarity retrieval, delivers strong zero-shot and few-shot performance while reducing manual feature engineering and training costs. Experiments on real e-commerce datasets demonstrate notable improvements over ML models and LLM-based baselines. AdaRec supports real-time preference adaptation and explainable recommendations, making it practical for dynamic environments. Future work will focus on online A/B testing and extending to multi-modal tasks to enhance performance in real business applications.

References

Bao, K., Zhang, J., Zhang, Y., Wang, W., Feng, F., and He, X. Tallrec: An effective and efficient tuning framework to align large language model with recommendation. In *Proceedings of the 17th ACM Conference on Recommender Systems*, pp. 1007–1014, 2023.

Du, Y., Wang, W., and Wang, L. Hierarchical recurrent neural network for skeleton based action recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1110–1118, 2015.

Hansen, C., Hansen, C., Maystre, L., Mehrotra, R., Brost, B., Tomasi, F., and Lalmas, M. Contextual and sequential user embeddings for large-scale music recommendation. In *Proceedings of the 14th ACM Conference on Recommender Systems*, pp. 53–62, 2020.

He, X., Liao, L., Zhang, H., Nie, L., Hu, X., and Chua, T.-S. Neural collaborative filtering. In *Proceedings of the 26th international conference on world wide web*, pp. 173–182, 2017.

Hu, L., Xu, S., Li, C., Yang, C., Shi, C., Duan, N., Xie, X., and Zhou, M. Graph neural news recommendation with unsupervised preference disentanglement. In Jurafsky, D., Chai, J., Schluter, N., and Tetreault, J. (eds.), *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 4255–4264, Online, July 2020. Association for Computational Linguistics. doi: 10.18653/v1/2020.acl-main.392. URL https://aclanthology.org/2020.acl-main.392.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30, 2017.

Li, L., Chu, W., Langford, J., and Wang, X. Unbiased offline evaluation of contextual-bandit-based news article recommendation algorithms. In *Proceedings of the fourth ACM international conference on Web search and data mining*, pp. 297–306, 2011.

Lin, X., Wang, W., Li, Y., Yang, S., Feng, F., Wei, Y., and Chua, T.-S. Data-efficient fine-tuning for llm-based recommendation. In *Proceedings of the 47th international ACM SIGIR conference on research and development in information retrieval*, pp. 365–374, 2024.

Mysore, S., McCallum, A., and Zamani, H. Large language model augmented narrative driven recommendations. In *Proceedings of the 17th ACM Conference on Recommender Systems*, pp. 777–783, 2023.

Pi, Q., Bian, W., Zhou, G., Zhu, X., and Gai, K. Practice on long sequential user behavior modeling for click-through rate prediction. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pp. 2671–2679, 2019.

Pi, Q., Zhou, G., Zhang, Y., Wang, Z., Ren, L., Fan, Y., Zhu, X., and Gai, K. Search-based user interest modeling with lifelong sequential behavior data for click-through rate prediction. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, pp. 2685–2692, 2020.

Sheng, X.-R., Zhao, L., Zhou, G., Ding, X., Dai, B., Luo, Q., Yang, S., Lv, J., Zhang, C., Deng, H., et al. One model to serve all: Star topology adaptive recommender for multi-domain ctr prediction. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, pp. 4104–4113, 2021.

Spirtes, P., Meek, C., and Richardson, T. Causal inference in the presence of latent variables and selection bias. In *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence*, pp. 499–506, 1995.

Wang, H., Wang, N., and Yeung, D.-Y. Collaborative deep learning for recommender systems. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 1235–1244, 2015.

Wang, Y., Jiang, Z., Chen, Z., Yang, F., Zhou, Y., Cho, E., Fan, X., Huang, X., Lu, Y., and Yang, Y. Recmind: Large language model powered agent for recommendation. *arXiv preprint arXiv:2308.14296*, 2023.

Wang, Y., Jiang, Z., Chen, Z., Yang, F., Zhou, Y., Cho, E., Fan, X., Lu, Y., Huang, X., and Yang, Y. Recmind: Large language model powered agent for recommendation. In *Findings of the Association for Computational Linguistics: NAACL 2024*, pp. 4351–4364, 2024.

Weng, S.-S. and Liu, M.-J. Feature-based recommendations for one-to-one marketing. *Expert Systems with Applications*, 26(4):493–508, 2004.

Wu, L., Zheng, Z., Qiu, Z., Wang, H., Gu, H., Shen, T., Qin, C., Zhu, C., Zhu, H., Liu, Q., et al. A survey on large language models for recommendation. *World Wide Web*, 27(5):60, 2024.

Xie, Z., Wu, J., Jeon, H., He, Z., Steck, H., Jha, R., Liang, D., Kallus, N., and McAuley, J. Neighborhood-based collaborative filtering for conversational recommendation. In *Proceedings of the 18th ACM Conference on Recommender Systems*, pp. 1045–1050, 2024.

Zhang, C., Wu, S., Zhang, H., Xu, T., Gao, Y., Hu, Y., and Chen, E. Notellm: A retrievable large language model for note recommendation. In *Companion Proceedings of the ACM on Web Conference 2024*, pp. 170–179, 2024.

Zhang, J., Xie, R., Hou, Y., Zhao, X., Lin, L., and Wen, J.-R. Recommendation as instruction following: A large language model empowered recommendation approach. *ACM Transactions on Information Systems*, 2023.

Zhou, G., Zhu, X., Song, C., Fan, Y., Zhu, H., Ma, X., Yan, Y., Jin, J., Li, H., and Gai, K. Deep interest network for click-through rate prediction. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, pp. 1059–1068, 2018.

Zhou, G., Mou, N., Fan, Y., Pi, Q., Bian, W., Zhou, C., Zhu, X., and Gai, K. Deep interest evolution network for click-through rate prediction. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pp. 5941–5948, 2019.

A. Prompt Template for AdaRec

System Prompt
You are a customer profile generator. Below is the data distribution for each feature:
`[Textual Distribution]`[4]
'Number of category purchased in the last 360 days.' has a mean value of 11.2 with a standard deviation of 6.6. The minimum observed value is 0.0, while the maximum is 34.0. Approximately 25% of values are below 6.0, the median (50th percentile) is 12.0, and 75% fall below 16.0 . . . .
Using this information, generate a clear and cohesive profile for the customer. For non-numerical features, emphasize specific values. For numerical features, describe relative trends without exact numbers. Present as a single fluid paragraph without extra formatting. Customer Profile:
Customer Profile
Number of category purchased in the last 360 days is 6. Number of category viewed in the last 30 days is 3. Number of benefits used in the last 360 days is 4 . . .

Table 4: Example of narrative profiling prompt.

System Prompt
As the Senior Marketing Manager at [company], your task is to recommend three brands for the promotional carousel. The promotion has two components: 1. Condition: Customers must purchase X units. 2. Reward: Customers receive 10% in Points.
`[brand description]`
Available brands:
cocacola: Grocery beverages . . .

Based on a customer profile, please recommend three brand names for the customer.

Factor Analysis
Reference: Factors and their importance ranking that affect brand recommendations based on historical data.
`[Factor Analysis Guidelines]`[5]

Pattern Analysis
Reference: Below are preferences from similar customer profiles.
`[Pattern Analysis Guidelines]`

Reference Cases:
Customer Profile: Number of days visited in the last 360 days is 315, Number of category viewed in the last 30 days is 5, Total mobile app visits in the last 360 days is 522 . . .
Based on the information above, please recommend three brand names in the following JSON format: `{'brand': 'brand1, brand2, brand3', 'confidence': confidence, 'reason': reason}`
Narrative Profile
This customer is an active Prime member ... focusing primarily on wireless products and sports items...

Table 5: Example of the structured reasoning prompt.

B. Comparison of Expert Profiling and Narrative Profiling

Figure 3 compares expert profiling and narrative profiling.

C. Case Study

Figure 4 compares expert reasoning with AdaRec. While experts rely on descriptive statistics, AdaRec captures nuanced behavioral patterns and underlying customer preferences, offering both accurate predictions and explainable insights.

[4]See the code for full details of [Textual Distribution].
[5]Full versions of [Factor Analysis Guidelines] and [Pattern Analysis Guidelines] are available in the codes.

❖ Expert Profiling

The customer currently has been with [COMPANY] for 6,407 days, has 116 Points, does not own any credit card, In the last 360 days, the customer visited [COMPANY] 35 days, last visited [COMPANY] 1 day ago, visited 0 time using mobile app, visited 62 times using desktop, viewed products from 3 different categories...

❖ Narrative Profiling

This customer is not a current Prime member and does not have a co-branded credit card, preferring to use other benefits such as free shipping on orders within a certain timeframe. They tend to engage moderately with promotional content, with a slightly above average percentage of paid sessions in the past month. In terms of purchasing habits, they have bought items from a smaller number of categories compared to the average, focusing their purchases more narrowly. Their recent activity shows a lower frequency of category browsing, with fewer categories viewed in the last month than the typical customer. Despite this, they actively utilize a moderate number of benefits over the year, suggesting a selective but consistent engagement with available offers...

Figure 3: Expert Profiling vs. Narrative Profiling.

❖ Expert

✖ The customer has a long purchasing history (6,407 days), consistent engagement (11 active days), and significant spending during previous deals. With past participation in two deal events, they might find the current offer attractive.

❖ AdaRec

✔ The customer tends to avoid promotions and discounts, as evidenced by their minimal engagement with promotional content. Their sporadic recent activity and preference for straightforward purchasing suggest they are less likely to participate in deals requiring specific conditions.

Figure 4: AdaRec vs. expert analysis: capturing behavioral patterns for explainable recommendations.