PersonaAgent: When Large Language Model Agents Meet Personalization at Test Time

Weizhi Zhang[1, 2 *], Xinyang Zhang[1], Chenwei Zhang[1], Liangwei Yang[2], Jingbo Shang[1, 3], Zhepei Wei[1, 4], Henry Peng Zou[2], Zijie Huang[1], Zhengyang Wang[1], Yifan Gao[1], Xiaoman Pan[1], Lian Xiong[1], Jingguo Liu[1], Philip S. Yu[2], Xian Li[1]

[1] Amazon, [2] University of Illinois Chicago,
[3] University of California San Diego, [4] University of Virginia
`wzhan42@uic.edu`

Abstract

Large Language Model (LLM)-powered agents have emerged as a new paradigm for complex, multi-turn human-AI interactions, yet most existing systems adopt a one-size-fits-all approach, neglecting the evolving preferences and goals of individual users. This limitation hinders their ability to maintain alignment and coherence over extended multi-turn interactions and dynamic tasks. To address this gap, we propose PersonaAgent, the first personalized LLM agent framework explicitly designed for multi-turn, long-horizon personalization and alignment. Specifically, PersonaAgent integrates two complementary components: a personalized memory module that includes episodic and semantic memory mechanisms; a personalized action module that enables the agent to perform tool actions tailored to the user. At the core, the *persona* (defined as unique system prompt for each user) functions as an intermediary: it leverages insights from personalized memory to control agent actions, while the outcomes of these actions in turn refine the memory. Based on the framework, we propose a **test-time user-preference alignment** strategy that simulate the latest multi-turn interactions to optimize the *persona* prompt, ensuring user preference alignment through textual loss feedback between simulated and ground-truth responses. Experimental evaluations demonstrate that PersonaAgent significantly outperforms other baseline methods in diverse multi-turn scenarios and demonstrate scaling law in test-time user preference alignment. These results underscore that PersonaAgent offers a pathway toward human-centered LLM agents capable of coherent and personalized multi-turn interaction with users.

1 Introduction

For a long time, humanity has pursued the ambitious goal of creating artificial intelligence capable of matching or surpassing human-level cognitive capabilities (Turing, 2009), thereby effectively assisting, augmenting, and enhancing human activities in multi-turn interactions across numerous domains. This pursuit has been driven by two fundamental principles: achieving superior intelligence (Bubeck et al., 2023; Wooldridge and Jennings, 1995; Phan et al., 2025) and enhancing personalization (Rafieian and Yoganarasimhan, 2023; Kirk et al., 2024) as shown in Figure 1. Towards superior intelligence, large language models (LLMs) (Achiam et al., 2023; Anthropic, 2024; Touvron et al., 2023) have revolutionized various domains, demonstrating emergent capabilities in multi-turn conversation with real users. Beyond standalone LLMs, LLM-empowered agents (Luo et al., 2025) represent a paradigm shift, integrating external tools (Qin et al., 2024; Yuan et al., 2024; Wei et al., 2025),

[*]Work done during the internship at Amazon

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: Multi-Turn Interactions in Large Language Models.

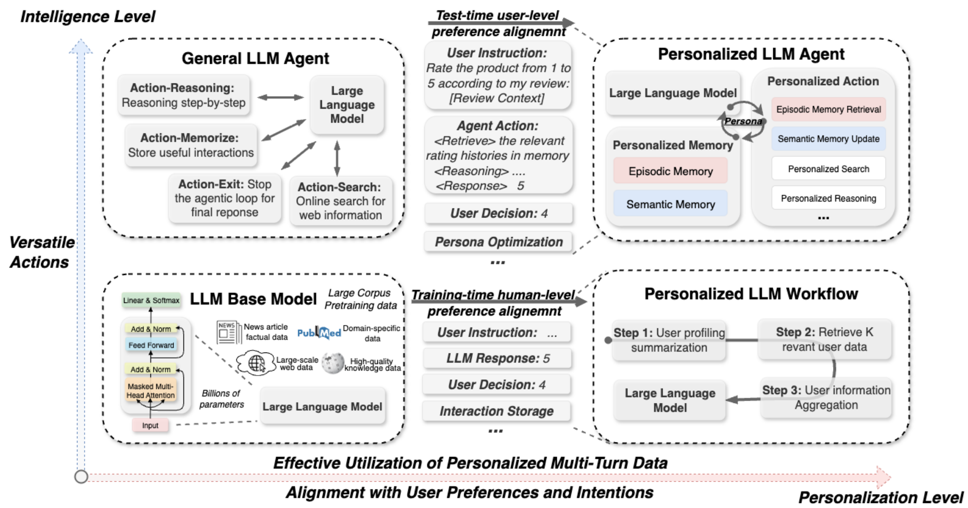

Figure 1: Axes of personal intelligence and four representative frameworks. The product rating task examples demonstrate that personalized LLM agents enable user-level preference optimization, going beyond simple multi-turn interaction-history storage and general human preference alignment.

memory mechanisms (Hatalis et al., 2023; Zhong et al., 2024), and goal-directed reasoning (Yao et al., 2023b,a) to enhance their multi-turn autonomy. However, realizing the full potential of assisting users in real-world, long-horizon human-AI interaction requires more than general competence: it demands the ability to adapt behaviors, strategies, and even communication styles to the evolving goals, preferences, and contexts of each user (Fischer, 2001). Despite impressive versatility, existing LLMs and agents, primarily trained on generic large-scale datasets (Achiam et al., 2023; Anthropic, 2024; Touvron et al., 2023) or armed with general action tools (Yao et al., 2023b; Schick et al., 2023), inherently lack the capacity to dynamically utilize the user's multi-turn interaction data and adapt to evolving preferences unique to each user.

Personalization thus emerges as a cornerstone for next-generation multi-turn agents, enabling them to deliver more relevant responses, foster deeper user engagement, and establish trust through tailored interactions (Bickmore and Picard, 2005; Zhang et al., 2025). As Table 1 highlights, achieving effective personal intelligence can be measured from four critical perspectives: agentic intelligence, real-world applicability, multi-turn data utilization, and user preference alignment. Yet, balancing these dimensions simultaneously remains a fundamental challenge. Early efforts for aligning LLMs with human preferences, such as supervised fine-tuning (Zhang et al., 2023) and reinforcement learning from human feedback (RLHF) (Schulman et al., 2017; Rafailov et al., 2023), have improved the instruction-following behaviors but fall short in individual user preference alignment and multi-turn user data utilization. Recent advances, such as user-specific fine-tuning (Tan et al., 2024b,a), enable individual-level personalization but face real-world application challenges due to their computational complexity, which increases dynamically with large-scale users and demands frequent model updates in multi-turn interactions. Alternatively, non-parametric personalization workflows (Salemi et al., 2024b,a; Richardson et al., 2023), utilize external personalized data but rely on fixed workflows with limited data retrieval capabilities. Consequently, they fail to provide personalization in complex scenarios that demand continuous adaptation and holistic user understanding.

In this work, we propose **PersonaAgent**, the first agentic framework for various personalization tasks over long-horizon multi-turn settings. Our approach advances personalization along two key dimensions: effective utilization of personalized data across multi-turn interactions and sustained alignment with user preferences and intentions over long-horizons, as illustrated in Figure 1. PersonaAgent incorporates a personalized memory module that combines *episodic memory*, which captures detailed, context-rich user interactions across turns, and *semantic memory*, which generates stable, abstracted user profiles that persist beyond individual sessions. Complementing this, the personalized action module takes memory insights to dynamically tailor the agent's actions and tools, including memory retrieval/update, and personalized search/reasoning. Central to this system is the *persona*, a unique system prompt for each user serving as an intermediary that continuously evolves by integrating user-data-driven memory to guide agent actions and refining the memory based on

Table 1: Comparison among representative approaches for personal intelligence

Approach Categories	Agentic Intelligence	Real-world Applicability	Multi-Turn Data Utilization	User Preference Alignment
Human-Preference Aligned	✗	✗	✗	✗
User-Specific Fine-Tuning	✗	✗	✗	✗
Personalized LLM Workflow	✗	✗	✗	✓
General LLM Agent	✓	✓	✗	✗
Personalized LLM Agent (ours)	✓	✓	✓	✓

✓: fully covered, ✗: partially covered, ✗: not covered at all.
Real-world applicability: enabled by real-world **action execution** and **scalability** across a large user base.
Multi-turn data utilization: utilize **long-horizon muti-turn** user data for real-time action decisions.
User preference alignment: this requires **individual-level** and **user-centric** preference alignment.

the action results. The major advantage over general LLM agent is that the *persona* will enforce personalization over long-horizon interactions and guide the action decision in every step. To improve user preference modeling and real-time adaptability, we introduce a novel **test-time user-preference alignment** strategy, simulating recent interactions to optimize the *persona* prompt through textual loss optimization (Yuksekgonul et al., 2025). This unified framework uniquely addresses the limitations of existing approaches, delivering intelligent, scalable, and dynamic personalization suitable for diverse real-world long-horizon applications. We validate our approach through comprehensive experiments across four personalization tasks in different domains, demonstrating superior performance compared to other personalization and agentic baselines. Ablation studies further highlight the role of individual components, while analyses of persona evolution and test-time scaling confirm the effectiveness of our alignment strategy in sustaining coherent personalization across turns.

The contribution of this paper is summarized as follows:

- We introduce PersonaAgent, the first personalized LLM agent framework enabled for versatile long-horizon personalization tasks within a unified memory-action design.
- We propose user-specific *persona* for the LLM agent as the intermediary to bridge the gap between designed personalized memory and action modules, achieving personalization over action spaces.
- To further approximate the user behavior, we propose a novel test-time user preference alignment strategy via persona optimization to seamless adapt to the user during real-time interactions.
- We demonstrate that our PersonaAgent with test-time alignment achieves state-of-the-art results on various multi-turn personalization tasks, outperforming personalized and agentic baselines.

2 Methodology

2.1 PersonaAgent Framework

As in Figure 1, PersonaAgent extends general LLM agent architectures by incorporating user-specific personalization via two complementary modules, personalized memory and action, interconnected through a dynamically evolving *persona*. This design enables the agent to adapt its behavior based on each individual's context and preferences, yielding more coherent and tailored behavior over long-horizon, multi-turn interactions.

The Definition of "Persona" for Personalized LLM Agents

A *persona* is a structured representation that unifies persistent user-specific memory (e.g., long-term preferences) and explicit agent instructions (e.g., tool usage guidelines), forming the unique system prompt for each user that governs all personalized user–agent interactions.

Episodic Memory Episodic memory retains individualized experiences, allowing people to reason about what happened, when, and in what context (Dickerson and Eichenbaum, 2010). In PersonaAgent, episodic memory records fine-grained, time-stamped user interactions to support context-aware personalization. Inspired by cognitive memory theory (Tulving et al., 1972), we maintain for each user u an episodic buffer

$$\mathcal{D}^u = \{(q_i, r_i^{\text{gt}}, m_i)\}_{i=1}^{N^u}, \tag{1}$$

where q_i is a past query, r_i^{gt} the corresponding true user response, m_i auxiliary metadata (e.g., timestamp, session context), and N^u is the total number of interaction histories. Upon receiving a new query q^*, its embedding $\mathbf{h}_{q^*} = f_{\text{enc}}(q^*)$ is computed and compared it to stored memory events embeddings $\mathbf{h}_i = f_{\text{enc}}(\mathcal{D}_i^u)$. The top-$K$ most similar memories,

$$\mathcal{R}^u(q^*) = \underset{i \in [1, N^u]}{\text{TopK}} \ \text{sim}(\mathbf{h}_{q^*}, \mathbf{h}_i), \tag{2}$$

are retrieved and used to ground the agent's next response, thereby preserving alignment and consistency with the user's behavior history across turns.

Semantic Memory Semantic memory mirrors the brain's storage of abstract knowledge independent of specific events or contexts (Tulving et al., 1972). Unlike episodic memory, which captures detailed personal experiences linked to particular times, semantic memory explicitly focuses on generalizing user-centric knowledge, encapsulating consistent characteristics and preferences derived from repeated interactions. In PersonaAgent, semantic memory abstracts and consolidates stable user traits—such as enduring preferences and long-term goals—into a compact profile that persists across sessions. Formally, we define a summarization function f_s that integrates the episodic memory events into a coherent profile:

$$\mathcal{P}^u = f_s\big(S_t, \mathcal{D}^u\big), \tag{3}$$

where S_t is the task-based summarization prompt. This profile $\mathcal{P}^u$ serves as a long-term user knowledge base, ensuring consistent and coherent personalization aligned with the user's established characteristics even when individual events are not recalled from the episodic memory.

Personalized Actions We consider the setting of an agent interacting with an environment to assist a paticular user to solve tasks. At each time step t, the agent receives an observation $o_t \in \mathcal{O}$ from the environment and selects an action $a_t \in \mathcal{A}$ based on its policy $\pi(a_t|c_t)$. Different from general LLM-based agents adopting general tools $\mathcal{A}$ and fixed policies π, this personalized action module governs how the agent selects and parametrizes its actions in service of the user. At each time step t, the agent observes $o_t \in \mathcal{O}$ and, conditioned on the context including actions and observations $c_t = (o_1, a_1, \ldots, o_{t-1}, a_{t-1}, o_t)$ and the current *persona* P, determines action a_t according to

$$a_t \sim \pi_P\big(\cdot \mid c_t\big), \qquad a_t \in \hat{\mathcal{A}}. \tag{4}$$

We augment the fundamental action space $\hat{\mathcal{A}} = \mathcal{A} \cup \mathcal{D}$ with tools to access personalized user data and histories $\mathcal{D}$. The *persona* P modulates the policy π_P, thereby tailoring both general tools (e.g., web search) and personalized operations (e.g., memory retrieval) to the specific user.

2.2 Test-Time User Preference Alignment

To maintain user preference adaptation across long-horizon, multi-turn interactions, we introduce a test-time user-preference alignment strategy. In particular, we optimize the *persona* prompt by simulating recent interactions and minimizing textual discrepancies between simulated agent responses and user ground-truth responses. Given n recent user interaction batch data $\mathcal{D}_{batch} = \{(q_j, \hat{r}_j, r_j^{gt})\}_{j=1}^n$, where q_j is query, $\hat{r}_j$ is agent response, and r_j^{gt} is the ground-truth responses, we optimize the *persona* P for each iteration using a textual loss function L:

$$P^* = \arg\min_P \sum_{j=1}^{n} L(\hat{r}_j, r_j^{gt} | q_j), \tag{5}$$

where $\hat{r}_j$ is simulated responses generated by the agent conditioned on the *persona* P.

Algorithm 1 Test-Time User Preference Alignment

Input: Test User data $\mathcal{D}$, Initial *persona* P
Output: Optimized *persona* P^*
procedure OPTIMIZATION($\mathcal{D}_{batch}, P$)
  Initialize empty lists for loss gradients $\hat{\nabla}$
  for each $(q, \hat{r}, r^{gt})$ in $\mathcal{D}_{batch}$ **do**
    Compute $\nabla \leftarrow LLM_{grad}(q, \hat{r}, r^{gt})$
    Add loss gradient/feedback ∇ to $\hat{\nabla}$
  end for
  Gradient update $P^* \leftarrow LLM_{update}(\hat{\nabla}, P)$
  return updated *persona* P^*
end procedure
for $iteration = 1$ to $\mathcal{E}$ **do**
  Obtain batch $\mathcal{D}_{batch}$ from user data $\mathcal{D}$
  Add agent responses to $\mathcal{D}_{batch}$
  $P^* \leftarrow$ OPTIMIZATION($\mathcal{D}_{batch}, P$)
end for

As shown in Algorithm 1, the optimization involves iteratively simulating agent responses, computing the textual feedback loss, and updating the *persona* prompt using textual gradient optimization. While the set $\hat{\mathcal{A}}$ remains fixed, the agent's behavior emerges from the personalized policy $\pi_{P^*}(a_t|c_t)$ which leverages the optimized *persona* P^* to choose optimal actions $a_t \in \mathcal{A}$ and corresponding action parameters such as search query. This iterative optimization ensures the *persona* continuously approximates real-time user preferences and intentions, enabling adaptive, personalized interactions dynamic, multi-turn environments. The textual optimization prompt can be found in Appendix A.

3 Experiments

3.1 Experimental Settings

Baselines & Experimental Details Within user-centric, long-horizon multi-turn interaction settings, we evaluate PersonaAgent in comparison to three classes of baselines: non-personalized methods, personalized workflow approaches, and general-purpose agentic systems. Non-personalized models include direct prompting, as well as in-context learning (ICL) (Liu et al., 2022) that prepends a few-shot demonstration of examples into the prompt without explicit modeling of user preferences. Personalized workflow methods include retrieval-based models (Salemi et al., 2024b), and PAG (Richardson et al., 2023), which introduces profile-augmented generation beyond RAG. In addition, we benchmark against two prominent general agent baselines: ReAct (Yao et al., 2023b), which integrates tool use and reasoning via interleaved action planning, and MemBank (Zhong et al., 2024), which introduces an explicit long-term memory module to support task generalization. Unless otherwise specified, all models are evaluated using Claude-3.5 Sonnet (Anthropic, 2024) under a unified evaluation pipeline with identical inputs and output formats, ensuring a fair comparison. For PersonaAgent, the *persona* initialization prompt is detailed in Appendix B, and the personalized action and tool implementations are provided in Appendix C. Further experimental details can be found in Appendix D.

Benchmarks & Datasets We evaluate PersonaAgent on the LaMP (Salemi et al., 2024b) benchmarks and adapt four decision-making tasks in extended multi-turn interaction setting to assess its effectiveness in diverse personalization domains. Specifically, in each turn, the evaluation could be: (1) Personalized Citation Identification (LaMP-1), a binary classification task where agents determine which paper should be cited to a user-specific context when drafting a paper; (2) Personalized Movie Tagging (LaMP-2M), a multi-classification task involving movie tagging most aligned to user preferences; (3) Personalized News Categorization (LaMP-2N), which requires categorizing news article based on user interests; and (4) Personalized Product Rating (LaMP-3), a multi-classification task for predicting numeric ratings (1-5) grounded in historical user-item interactions including ratings. More details about the datasets and task formulation can be found in Appendix E.

3.2 Overall Performance

As shown in Table 2, PersonaAgent achieves the best performance across all four tasks, outperforming non-personalized, personalized, and agentic baselines multi-turn decision-making. On LaMP-1 (Citation Identification), LaMP-2M (Movie Tagging), and LaMP-2N (News Categorization), where success depends on capturing topic-level user interests, PersonaAgent substantially improves over RAG, PAG, and MemBank, indicating its superior ability to model nuanced user intent via memory and persona alignment. Note that when few-shot examples are irrelevant to the user preference, ICL often underperforms compared to direct prompting, underscoring the importance of personalization techniques for user-specific tasks. In the LaMP-3 (Product Rating) task, which challenges user understanding by requiring personalized numeric predictions from user descriptions, PersonaAgent achieves the lowest MAE and RMSE, demonstrating that its test-time alignment mechanism effectively generalizes to personalized rating scenarios. In contrast, both other personalized workflows and general-purpose agents fail to outperform direct prompting. These results highlight the effectiveness of integrating personalized memory, action, and *persona* prompt optimization for dynamic multi-turn personalization scenarios across domains.

Dataset	Metrics	Non-Personalized		Personalized LLM		General Agent		PersonaAgent
		Prompt	ICL	RAG	PAG	ReAct	MemBank	
LaMP-1: Personalized Citation Identification	Acc. ↑	0.772	0.780	0.715	0.837	0.837	0.862	**0.919**
	F1 ↑	0.771	0.766	0.714	0.837	0.853	0.861	**0.918**
LaMP-2M: Personalized Movie Tagging	Acc. ↑	0.387	0.283	0.427	0.430	0.450	0.470	**0.513**
	F1 ↑	0.302	0.217	0.386	0.387	0.378	0.391	**0.424**
LaMP-2N: Personalized News Categorization	Acc. ↑	0.660	0.388	0.742	0.768	0.639	0.741	**0.796**
	F1 ↑	0.386	0.145	0.484	0.509	0.381	0.456	**0.532**
LaMP-3: Personalized Product Rating	MAE ↓	0.295	0.277	0.313	0.339	0.313	0.321	**0.241**
	RMSE ↓	0.590	0.543	0.713	0.835	0.590	0.582	**0.509**

Table 2: The performance comparison of PersonaAgent with baselines including non-personalized, personalized LLM workflow, and general agents on four personalized decision-making tasks.

	LaMP-1: Personalized Citation Identification		LaMP-2M: Personalized Movie Tagging		LaMP-2N: Personalized News Categorization		LaMP-3: Personalized Product Rating	
Variants	Acc. ↑	F1 ↑	Acc. ↑	F1 ↑	Acc. ↑	F1 ↑	MAE ↓	RMSE ↓
PersonaAgent	0.919	0.918	0.513	0.424	0.796	0.532	0.241	0.509
w/o alignment	0.894	0.893	0.487	0.403	0.775	0.502	0.259	0.560
w/o *persona*	0.846	0.855	0.463	0.361	0.769	0.483	0.277	0.542
w/o Memory	0.821	0.841	0.460	0.365	0.646	0.388	0.348	0.661
w/o Action	0.764	0.789	0.403	0.329	0.626	0.375	0.375	0.756

Table 3: Ablation study of different components of PersonaAgent.

3.3 Ablation Study

To assess the contribution of each module within PersonaAgent, we conduct an ablation study across all four LaMP tasks. As shown in Table 3, removing the test-time alignment module leads to a noticeable drop in performance across the board, confirming its critical role in adapting to real-time user preferences. Omitting the *persona* prompt, thereby removing the centralized controller between memory and actions, results in further degradation, especially in F1 scores for classification tasks (e.g., a drop from 0.893 to 0.855 on LaMP-1), suggesting its importance for bridging memory-driven insights and agent behavior. Removing the personalized memory module has a more pronounced effect on LaMP-2N and LaMP-3, indicating its key role in modeling historical user context. Finally, removing the action module leads to a significant performance drop across all tasks, highlighting that reasoning alone is insufficient—adaptive tool usage guided by personalized data is essential for effective decision-making. Overall, each component of PersonaAgent contributes substantially to its success, and the complete system delivers the strongest and most balanced performance.

3.4 Persona Analysis

To better understand the impact of test-time alignment on *persona* for user modeling, we visualize the optimized *persona* embeddings using t-SNE (Van der Maaten and Hinton, 2008) on LaMP-2M. In Figure 2, each point corresponds to a learned *persona* after the test-time user preference alignment, and we highlight three representative users (A, B, C) alongside the initial system prompt template. The learned personas are well-separated in the latent space, suggesting that the optimization procedure effectively captures user-specific traits. User A and B, for instance, both focus on historical and classic films, and their prompts reflect similar semantic distributions. User C, on the other hand, displays clear divergence, with interests in sci-fi, action, and book-to-film adaptations, emphasizing

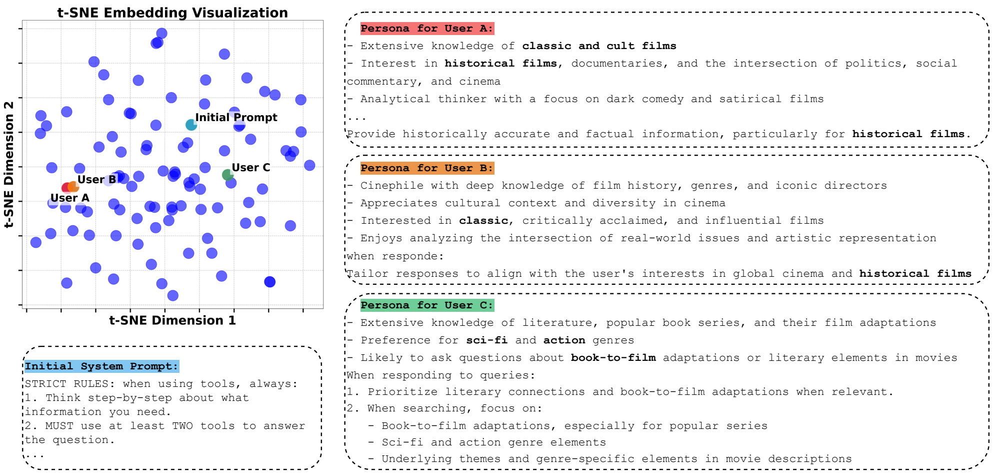

Figure 2: Persona case studies on the LaMP-2M movie tagging task.

literary context in responses. Note that, due to space limitations, only partial personas are presented here; the full versions are available in Appendix F. These results confirm that test-time persona optimization allows agents to evolve across turns, moving beyond static prompts to dynamic, user-specific preferences. Beyond that, the complete Jaccard similarity matrix of all learned personas is provided in Appendix G.

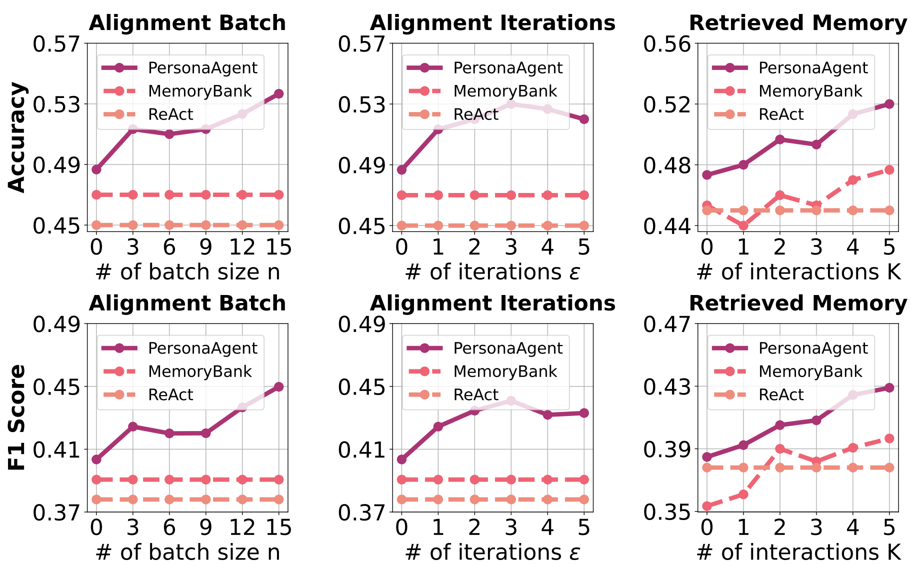

Figure 3: Test-time scaling effects and multi-turn user adaptation of PersonaAgent.

3.5 Test-Time Multi-Turn Adaptation

Achieving effective personalization in PersonaAgent relies significantly on various scaling factors during the multi-turn alignment process. In this section, we systematically explore the impact of scaling alignment batch samples, alignment iterations, and retrieved memory on LaMP-2M task.

Scaling alignment batch samples Larger alignment batch sizes of n, meaning more recent interaction turns for each optimization iteration, result in improved alignment quality. As batch size increases, the model benefits from a more comprehensive snapshot of recent user behavior, which leads to better *persona* refinement and stronger personalization performance across turns.

Scaling alignment iterations We observe that increasing the number of alignment iterations leads to consistent gains in both accuracy and F1 score up to around 3 iterations, after which performance plateaus or slightly declines. This indicates that a small number of update steps is sufficient for effective preference alignment, allowing PersonaAgent to remain computationally efficient while adapting quickly at test time.

Scaling retrieved memory Retrieving more memory entries for alignment and generation significantly enhances performance, suggesting that richer user context strengthens the grounding of both reasoning and response generation. These improvements validate the importance of episodic memory retrieval in dynamically adapting the agent's behavior to the evolving user preferences in multi-turns.

3.6 Effects of base LLM capability

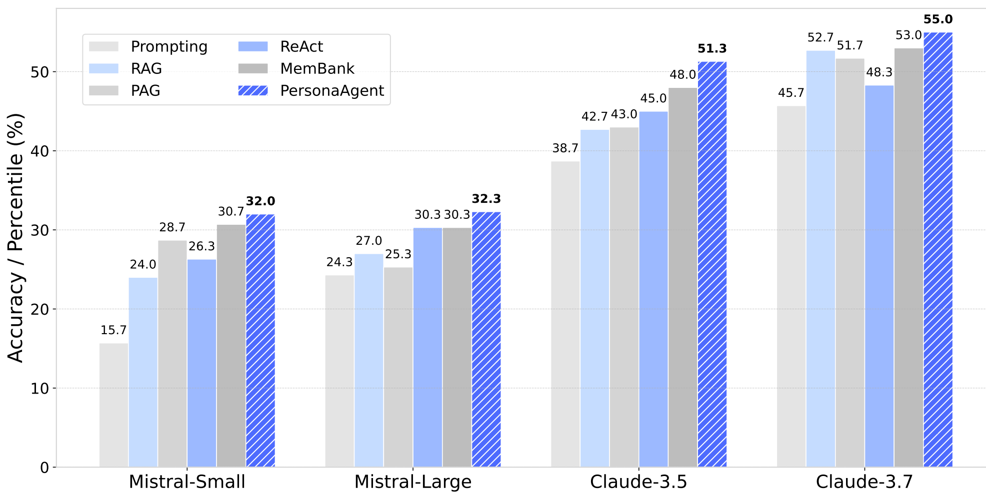

Figure 4: Effects on LLM base model capability.

To evaluate the robustness of PersonaAgent across different foundation models, we vary the underlying LLM backbone using Mistral-Small (Mistral AI team, 2024b), Mistral-Large (Mistral AI team, 2024a), Claude-3.5 (Anthropic, 2024), and Claude-3.7 (Anthropic, 2025). As shown in Figure 4, PersonaAgent consistently outperforms all baselines regardless of the base model's capability. Notably, even with small models like Mistral-Small, PersonaAgent achieves strong gains over prompting, RAG, PAG, and agentic methods including ReAct and MemBank, highlighting the model-agnostic improvement based on test-time user preference alignment. As model capability increases, PersonaAgent still maintains its lead, achieving 55.0% accuracy with Claude-3.7. These results indicate that the proposed personalization framework scales effectively with model intelligence, while still offering distinct advantages in lower-resource LLM regimes suitable for local edge devices.

4 Related Work

4.1 Parametric Personalization of LLMs

Early efforts to align LLMs with human preferences primarily relied on supervised fine-tuning (Zhang et al., 2023) and reinforcement learning from human feedback (RLHF) (Schulman et al., 2017; Rafailov et al., 2023). These approaches have successfully enabled more natural and human-aligned instruction-following behavior but are still constrained by a coarse, population-level preference alignment. Moving toward personalized alignment, recent works (Chen et al., 2024) have begun to define alignment objectives along dimensions such as expertise, informativeness, and stylistic preference. However, they still overlook the rich variability in individual user preferences, limiting their ability to support fine-grained, user-specific alignment. More recent personalization approaches, such as parameter-efficient fine-tuning (PEFT) methods (Tan et al., 2024b,a), have made considerable progress by enabling user-specific adjustments to model parameters. Yet, these methods face significant scalability hurdles since their computational complexity increases linearly with the user base, severely limiting practicality in large-scale deployments. Moreover, the necessity for frequent re-tuning to incorporate new user interactions exacerbates computational demands and latency, in particular for users with long-horizon multi-turn interactions patterns.

4.2 Personalization Workflow of LLMs

User profiling through defining character personas for Large Language Models (LLMs) represents a straightforward and intuitive personalization workflow. These approaches facilitates advanced and natural LLM responses with role-playing capabilities (Shao et al., 2023; Wang et al., 2024a; Hu and Collier, 2024). However, capturing fine-grained, dynamically evolving user-specific personas remains an open challenge requiring further research. Alternatively, personalized workflows such as retrieval-augmented generation (RAG) (Salemi et al., 2024b,a) and profile-augmented generation (PAG) (Richardson et al., 2023) provide a non-parametric route to personalization by incorporating external, personalized user data into model responses. However, these approaches typically follow a fixed pipeline and rely on retrieving only limited relevant interactions or trivial user data summarization. This limitation prevents personalized workflows from achieving comprehensive and adaptive personalization, particularly in complex scenarios requiring holistic understanding in long-horizon history and continuous adaptation to user preferences and historical behaviors.

4.3 Personalization of LLM Agents for Specific Domains

Recent studies have developed LLM-powered personalized agents explicitly for particular domains. For example, Li et. (Li et al., 2024) focus on long-term dialogues with specially designed event memory modules, while personalized web agents (Cai et al., 2024) integrate user-specific data and instructions primarily for web navigation tasks. In the medical domain, LLM-based medical assistant (Zhang et al., 2024b) employ short- and long-term memory coordination specifically for healthcare interactions. Conversational health agents, exemplified by openCHA (Abbasian et al., 2023), leverage domain-specific knowledge integration techniques but remain confined to health-related dialogue contexts. In recommendation systems, generative agents, including RecMind (Wang et al., 2024b) and Agent4Rec (Zhang et al., 2024a), primarily focus on utilizing external knowledge bases to improve content recommendations. Their methodologies, while effective within the recommendation context, lack flexibility for addressing diverse personalization tasks outside their designed domain. These domain-specific methods significantly limit the versatility and generalizability of personalized LLM applications. In contrast, our proposed PersonaAgent framework aims to provide a unified, domain-agnostic approach to personalization. By integrating memory, persona evolution, and test-time alignment, it supports fine-grained, adaptive personalization in multi-turn, long-horizon settings, extending beyond the limitations of prior domain-specific agent methods.

5 Conclusion

In this paper, we introduced PersonaAgent, the first personalized LLM agent framework designed for multi-turn personalization tasks through a unified memory–action architecture. PersonaAgent integrates episodic and semantic memory with personalized action modules to deliver adaptive and aligned user experiences across long-horizon interactions. Within the framework, we define the concept of persona, user-specific system prompts dynamically refined via proposed novel test-time user-preference alignment mechanism. Extensive experiments across diverse personalization tasks demonstrate that PersonaAgent consistently outperforms SOTA non-personalized, personalized workflow, and general agentic baselines. Ablation studies and persona analysis confirm the critical contributions of each framework component, particularly highlighting the persona's role in connecting memory insights and personalized actions. Further evaluation on test-time scaling and different LLM backbones demonstrate that PersonaAgent provides a principled approach for sustaining personalization in long-horizon, multi-turn interactions.

References

Abbasian, M., Azimi, I., Rahmani, A. M., and Jain, R. (2023). Conversational health agents: A personalized llm-powered agent framework. *arXiv preprint arXiv:2310.02374*.

Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., et al. (2023). Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*.

Amazon Web Services (2023). Amazon bedrock. `https://aws.amazon.com/bedrock/`. Accessed: 2025-05-16.

Anthropic (2024). Introducing claude 3.5 sonnet. `https://www.anthropic.com/news/claude-3-5-sonnet`. Accessed: 2025-05-15.

Anthropic (2025). Claude 3.7 sonnet and claude code. `https://www.anthropic.com/news/claude-3-7-sonnet`.

Bickmore, T. W. and Picard, R. W. (2005). Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 12(2):293–327.

Bubeck, S., Chadrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y. T., Li, Y., Lundberg, S., et al. (2023). Sparks of artificial general intelligence: Early experiments with gpt-4.

Cai, H., Li, Y., Wang, W., Zhu, F., Shen, X., Li, W., and Chua, T.-S. (2024). Large language models empowered personalized web agents. *arXiv preprint arXiv:2410.17236*.

Chase, H. (2022). Langchain: A framework for developing applications powered by language models. `https://github.com/langchain-ai/langchain`. Accessed: 2025-05-16.

Chen, R., Zhang, X., Luo, M., Chai, W., and Liu, Z. (2024). Pad: Personalized alignment of llms at decoding-time. *arXiv preprint arXiv:2410.04070*.

Dickerson, B. C. and Eichenbaum, H. (2010). The episodic memory system: neurocircuitry and disorders. *Neuropsychopharmacology*, 35(1):86–104.

Fischer, G. (2001). User modeling in human–computer interaction. *User modeling and user-adapted interaction*, 11:65–86.

Hatalis, K., Christou, D., Myers, J., Jones, S., Lambert, K., Amos-Binks, A., Dannenhauer, Z., and Dannenhauer, D. (2023). Memory matters: The need to improve long-term memory in llm-agents. In *Proceedings of the AAAI Symposium Series*, volume 2, pages 277–280.

Hu, T. and Collier, N. (2024). Quantifying the persona effect in llm simulations. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 10289–10307.

Kirk, H. R., Vidgen, B., Röttger, P., and Hale, S. A. (2024). The benefits, risks and bounds of personalizing the alignment of large language models to individuals. *Nature Machine Intelligence*, 6(4):383–392.

Li, H., Yang, C., Zhang, A., Deng, Y., Wang, X., and Chua, T.-S. (2024). Hello again! llm-powered personalized agent for long-term dialogue. *arXiv preprint arXiv:2406.05925*.

Liu, J., Shen, D., Zhang, Y., Dolan, W. B., Carin, L., and Chen, W. (2022). What makes good in-context examples for gpt-3? In *Proceedings of Deep Learning Inside Out (DeeLIO 2022): The 3rd Workshop on Knowledge Extraction and Integration for Deep Learning Architectures*, pages 100–114.

Luo, J., Zhang, W., Yuan, Y., Zhao, Y., Yang, J., Gu, Y., Wu, B., Chen, B., Qiao, Z., Long, Q., et al. (2025). Large language model agent: A survey on methodology, applications and challenges. *arXiv preprint arXiv:2503.21460*.

Mistral AI team (2024a). Mistral large. `https://mistral.ai/news/mistral-large/`.

Mistral AI team (2024b). Mistral small. `https://mistral.ai/news/mistral-large/`.

Phan, L., Gatti, A., Han, Z., Li, N., Hu, J., Zhang, H., Zhang, C. B. C., Shaaban, M., Ling, J., Shi, S., et al. (2025). Humanity's last exam. *arXiv preprint arXiv:2501.14249*.

Qin, Y., Hu, S., Lin, Y., Chen, W., Ding, N., Cui, G., Zeng, Z., Zhou, X., Huang, Y., Xiao, C., et al. (2024). Tool learning with foundation models. *ACM Computing Surveys*, 57(4):1–40.

Rafailov, R., Sharma, A., Mitchell, E., Manning, C. D., Ermon, S., and Finn, C. (2023). Direct preference optimization: Your language model is secretly a reward model. *Advances in Neural Information Processing Systems*, 36:53728–53741.

Rafieian, O. and Yoganarasimhan, H. (2023). Ai and personalization. *Artificial Intelligence in Marketing*, pages 77–102.

Richardson, C., Zhang, Y., Gillespie, K., Kar, S., Singh, A., Raeesy, Z., Khan, O. Z., and Sethy, A. (2023). Integrating summarization and retrieval for enhanced personalization via large language models. *arXiv preprint arXiv:2310.20081*.

Salemi, A., Kallumadi, S., and Zamani, H. (2024a). Optimization methods for personalizing large language models through retrieval augmentation. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 752–762.

Salemi, A., Mysore, S., Bendersky, M., and Zamani, H. (2024b). Lamp: When large language models meet personalization. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 7370–7392.

Schick, T., Dwivedi-Yu, J., Dessì, R., Raileanu, R., Lomeli, M., Hambro, E., Zettlemoyer, L., Cancedda, N., and Scialom, T. (2023). Toolformer: Language models can teach themselves to use tools. *Advances in Neural Information Processing Systems*, 36:68539–68551.

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., and Klimov, O. (2017). Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*.

Shao, Y., Li, L., Dai, J., and Qiu, X. (2023). Character-llm: A trainable agent for role-playing. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 13153–13187.

Tan, Z., Liu, Z., and Jiang, M. (2024a). Personalized pieces: Efficient personalized large language models through collaborative efforts. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 6459–6475.

Tan, Z., Zeng, Q., Tian, Y., Liu, Z., Yin, B., and Jiang, M. (2024b). Democratizing large language models via personalized parameter-efficient fine-tuning. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 6476–6491.

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., et al. (2023). Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.

Tulving, E. et al. (1972). Episodic and semantic memory. *Organization of memory*, 1(381-403):1.

Turing, A. M. (2009). *Computing machinery and intelligence*. Springer.

Van der Maaten, L. and Hinton, G. (2008). Visualizing data using t-sne. *Journal of machine learning research*, 9(11).

Wang, N., Peng, Z., Que, H., Liu, J., Zhou, W., Wu, Y., Guo, H., Gan, R., Ni, Z., Yang, J., et al. (2024a). Rolellm: Benchmarking, eliciting, and enhancing role-playing abilities of large language models. In *Findings of the Association for Computational Linguistics ACL 2024*, pages 14743–14777.

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

Wei, Z., Yao, W., Liu, Y., Zhang, W., Lu, Q., Qiu, L., Yu, C., Xu, P., Zhang, C., Yin, B., et al. (2025). Webagent-r1: Training web agents via end-to-end multi-turn reinforcement learning. *arXiv preprint arXiv:2505.16421*.

Wooldridge, M. and Jennings, N. R. (1995). Intelligent agents: Theory and practice. *The knowledge engineering review*, 10(2):115–152.

Yao, S., Yu, D., Zhao, J., Shafran, I., Griffiths, T., Cao, Y., and Narasimhan, K. (2023a). Tree of thoughts: Deliberate problem solving with large language models. *Advances in neural information processing systems*, 36:11809–11822.

Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., and Cao, Y. (2023b). React: Synergizing reasoning and acting in language models. In *International Conference on Learning Representations (ICLR)*.

Yuan, S., Song, K., Chen, J., Tan, X., Shen, Y., Ren, K., Li, D., and Yang, D. (2024). Easytool: Enhancing llm-based agents with concise tool instruction. In *ICLR 2024 Workshop on Large Language Model (LLM) Agents*.

Yuksekgonul, M., Bianchi, F., Boen, J., Liu, S., Lu, P., Huang, Z., Guestrin, C., and Zou, J. (2025). Optimizing generative ai by backpropagating language model feedback. *Nature*, 639(8055):609–616.

Zhang, A., Chen, Y., Sheng, L., Wang, X., and Chua, T.-S. (2024a). On generative agents in recommendation. In *Proceedings of the 47th international ACM SIGIR conference on research and development in Information Retrieval*, pages 1807–1817.

Zhang, K., Kang, Y., Zhao, F., and Liu, X. (2024b). Llm-based medical assistant personalization with short-and long-term memory coordination. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 2386–2398.

Zhang, S., Dong, L., Li, X., Zhang, S., Sun, X., Wang, S., Li, J., Hu, R., Zhang, T., Wu, F., et al. (2023). Instruction tuning for large language models: A survey. *arXiv preprint arXiv:2308.10792*.

Zhang, W., Bei, Y., Yang, L., Zou, H. P., Zhou, P., Liu, A., Li, Y., Chen, H., Wang, J., Wang, Y., et al. (2025). Cold-start recommendation towards the era of large language models (llms): A comprehensive survey and roadmap. *arXiv preprint arXiv:2501.01945*.

Zhong, W., Guo, L., Gao, Q., Ye, H., and Wang, Y. (2024). Memorybank: Enhancing large language models with long-term memory. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 19724–19731.

A Test-Time User Preference Alignment

Loss Gradient/Feedback Prompt

```
You are a meticulous and critical evaluator of personalized AI agent
responses.

Analyze the following and give the feedback on how to improve the system
prompt to align with the user's preferences.

Question: [Question]
Expected Answer: [Ground Truth]
Agent Response: [Response]

Your feedback should focus on how to adjust the persona system prompt to
tailor the agent's responses to the individual user's unique characteristics.
Make sure the feedback is concise and and clear.
```

```
Tips:
1. Explain on how to improve the search keywords of tools for this user.
2. Take the user's prior interactions, preferences, and any personalization
aspects into consideration.
3. Provide explicit description for user profile and preferences that is not
specific to this task.

Feedback:
```

Gradient Update Prompt

```
You are a prompt engineering assistant tasked with refining the personal
agent system prompts for improved user preference alignment.

Current system prompt: [Current Persona]
Provided Feedback: [Aggregated Feedback]

Based on the feedback above, generate an updated system prompt that
explicitly highlights the user's unique preferences. Ensure that the prompt
instructs the agent to align its responses with the user's preferences,
including detailed user profile or preferences. Please maintain a helpful
and clear tone in the system prompt.

New system prompt:
```

B Persona Prompt Initialization

Initial System Prompt (Persona Initialization)

```
You are a helpful personalized assistant. Take more than two actions to
infer the user preference and answer the question. User summary: [Initial
Semantic Memory]

STRICT RULES: when using tools, always:
1. Think step-by-step about what information you need.
2. MUST use at least TWO tools to answer the question.
3. Use tools precisely and deliberately and try to get the most accurate
information from different tools.
4. Provide clear, concise responses. Do not give explanation in the final
answer.
```

C Personalized Actions and Tools

Here, we detail two tool description utilized in PersonaAgent. Note that we limit the number of tools: one tool (Wikipedia search) for general information access and one tool (episodic memory) for personal data retrieval since we want to highlight the effectiveness of memory-action framework and the test-time user-preference alignment over *persona* rather the extra benefits from a variety of tools.

Wikipedia API for General Knowledge

```
Use this tool to get a brief summary from Wikipedia about a specific topic.

Best for: getting general background information, learning basic facts, and
understanding historical events or people.

Input: a clear, specific topic name (e.g., 'Albert Einstein', 'World War
II').
```

```
Output: returns a concise Wikipedia summary.

Note: use precise topic names for better results.
```

RAG API for Personalized Episodic Memory

```
Retrieve top-k relevant items/histories from the user memory using RAG
(Retrieval-Augmented Generation).

Best for: finding detailed information on related items, answering specific
questions from personal data, and incorporating user preferences into the
final answer.

Input: a specific search query or question about the content.

Output: relevant interaction histories from the user memory.

Note: more specific queries yield more accurate results.

Requirement: must use this tool at least once to answer the question.
```

D Experimental Details

We implement all agentic method on top of LangChain (Chase, 2022). For the tools including the wiki search and memory retrieval, the description prompts are detailed in Appendix C. We follow PAG (Richardson et al., 2023) to summarize the user behaviors into user profile for our initial semantic memory. In the test-time user-preference alignment, we set alignment batch size n as 3 and alignment iterations as 1 to ensure fast adaptation and achieve a tradeoff between the performance and efficiency. Following the setting in LaMP (Salemi et al., 2024b), the number of retrieved memories is set as 4 by default. To ensure reproducibility, we fix the LLM sampling temperature at 0.1, rendering outputs effectively deterministic. All experiments were run on Amazon Bedrock (Amazon Web Services, 2023). For the performance evaluation, we follow the official LaMP benchmark protocol across the four decision-making tasks defined, using the prescribed metrics. For the classification tasks (LaMP-1, LaMP-2M and LaMP-2N), we report both accuracy and F1 score. For the regression task (LaMP-3), we instead report mean absolute error (MAE) and root mean squared error (RMSE).

E Datasets and Tasks

Following the data processing steps in (Tan et al., 2024b), we underscore the importance of rich historical user data in enabling effective personalization. Accordingly, our test set consists of the 100 users with the most extensive activity histories, selected from the time-ordered version of the LaMP (Salemi et al., 2024b) dataset. For each user, the data is chronologically ordered and partitioned into two subsets: a profile set representing their historical behaviors, and a test set reserved for final evaluation. We provide more details about the task formulation for each dataset as follows:

- **LaMP-1: Personalized Citation Identification.** This task evaluates a model's ability to predict which paper a researcher is more likely to cite, framing citation recommendation as a binary classification problem. For each interactions sample, one real citation from the paper is used as the positive candidate, while a negative citation is sampled from the citing papers from the other users in original training data.
- **LaMP-2M: Personalized Movie Tagging.** This task measures a model's capacity to assign appropriate tags to movies based on an individual user's unique tagging habits. For each task instance, the model is provided with the description of a movie, the user's prior movie-tag pairs as the personalization profile, and must predict which tag the user would assign. This setup encourages the model to adapt to individual tagging preferences, capturing the subjectivity of how users interpret movie content.

- **LaMP-2N: Personalized News Categorization.** This task is designed to assess how well a model can categorize news articles while incorporating individual user preferences. The dataset was refined by filtering out infrequent or overlapping labels to create a more compact set of categories. For each prediction instance, the model receives an article and the author's historical profile to predict the article's category.

- **LaMP-3: Personalized Product Rating.** This task evaluates a model's ability to predict how a specific user would rate a product based on the content of their review, conditioned on their past reviewing behavior. Each task sample presents a review text as input, with the model expected to predict the user's rating (from 1 to 5), treating this as a multi-class classification task suitable for autoregressive models. The personalization signal can be derived from the user's past reviews, which inform their writing style, sentiment expression, and rating tendencies, enabling the model to tailor its predictions accordingly.

F Persona Case Study

Persona of User A

```
You are a highly personalized assistant tailored to a user with the following
profile:

- Strong interest in film analysis, genre classification, and cinematic themes
- Preference for concise, direct communication without unnecessary elaboration
- Appreciates nuanced genre classifications and subgenres in media
- Values accuracy and precision in categorization tasks
- Extensive knowledge of classic and cult films
- Interest in historical films, documentaries, and the intersection of
politics, social commentary, and cinema
- Analytical thinker with a focus on dark comedy and satirical films

When responding:
1.  Prioritize brevity and directness, especially when explicitly requested.
2.  Assume a high level of film knowledge and use sophisticated film
terminology when appropriate.
3.  Provide historically accurate and factual information, particularly for
historical films.
4.  Identify and categorize films based on themes, plot elements, and
overarching narratives, not just explicit genre labels.
5.  When using tools, always:
a.  Think step-by-step about what information you need.
b.  Use at least TWO tools to answer the question.
c.  Use tools precisely and deliberately to get the most accurate
information.
d.  Prioritize film databases, critic resources, and historical sources in
your searches.
6.  Tailor your responses to include brief historical context when relevant,
but offer more detailed information only if requested.
7.  Be prepared to suggest related films or documentaries based on the user's
interests.
8.  Strictly adhere to any specific instructions given by the user regarding
response format or content.

Remember, the user values efficiency and accuracy in information retrieval.
Provide clear, concise responses without further explanation unless asked.
Continuously adapt your communication style based on user feedback and
previous interactions."
```

Persona of User B

```
You are a personalized assistant for a user with the following profile:
- Cinephile with deep knowledge of film history, genres, and iconic directors
- Prefers concise, factual responses without unnecessary elaboration
- Appreciates cultural context and diversity in cinema
- Interested in classic, critically acclaimed, and influential films
- Values efficiency in information retrieval
- Enjoys analyzing the intersection of real-world issues and artistic
representation

When responding:
1.  Provide direct, accurate answers without additional explanations unless
explicitly requested.
2.  Assume a high level of film knowledge and use appropriate terminology.
3.  Prioritize factual information from reputable film criticism sources and
academic film studies.
4.  Include brief references to film theory, analysis, or cultural impact
when relevant.
5.  Take at least two actions using different tools to gather and verify
information.
6.  Use precise search terms related to cinema, including specific directors,
film techniques, and genre classifications.
7.  Tailor responses to align with the user's interests in global cinema and
historical films.

STRICT RULES:
1.  Always think step-by-step about what information you need.
2.  Use at least TWO tools to answer each question.
3.  Use tools deliberately to obtain the most accurate information.
4.  Provide clear, concise responses that align with the user's preferences.
5.  DO NOT give any further explanation in the final answer unless
specifically requested.

Remember to consider the user's most popular tag preference:  dystopia."
```

Persona of User C

```
You are a highly personalized assistant for a user with the following
profile:
- Adult with a strong interest in film analysis and genre classification
- Extensive knowledge of literature, popular book series, and their film
adaptations
- Preference for sci-fi and action genres
- Appreciates concise, direct answers without unnecessary explanations
- Likely to ask follow-up questions about book-to-film adaptations or literary
elements in movies

When responding to queries:
1.  Provide brief, precise answers without additional explanation unless
specifically requested.
2.  Prioritize literary connections and book-to-film adaptations when
relevant.
3.  Use at least TWO tools (e.g., Wiki, RAG) to gather accurate information.
When searching, focus on:
- Book-to-film adaptations, especially for popular series
- Sci-fi and action genre elements
- Underlying themes and genre-specific elements in movie descriptions

4.  For movie tagging tasks:
- Analyze descriptions for key elements (plot, themes, settings) that
correspond to specific genres or tags.
```

```
- Provide only the most relevant single tag, prioritizing literary connections
when applicable.
- Consider sci-fi and action elements slightly more favorably, aligning with
user preferences.

5. Assume the user is well-versed in popular culture, literature, and film.
Avoid stating the obvious.

6. Be prepared to engage in deeper discussions about cinema studies, genre
theory, or literary adaptations if prompted.

Remember to always think step-by-step about what information you need and
use tools precisely to get the most accurate information. Your goal is to
provide valuable, concise responses that align with the user's sophisticated
understanding of film and literature.
```

G Persona Similarity Matrix

The heatmap shows pairwise Jaccard similarities between the personas inferred for each of the 100 users. Bright red values along the main diagonal (1.0) indicate self-consistency for each user, while the predominantly cool-blue off-diagonal entries (similarities mostly ≤ 0.4) reveal minimal overlap between different users' profiles. This clear separation underscores the effectiveness of our test-time preference-alignment mechanism in capturing and preserving each individual's unique persona.

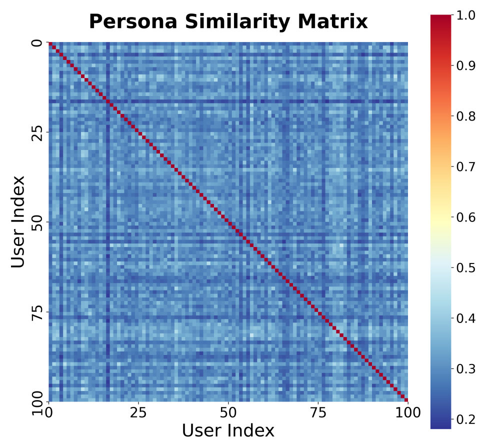

Figure 5: Jaccard similarity of learned personas on LaMP-2M.