

# 000 DESCRIPTIVE HISTORY REPRESENTATIONS: LEARNING 001 002 REPRESENTATIONS BY ANSWERING QUESTIONS 003 004

005 **Anonymous authors**

006 Paper under double-blind review

## 007 008 ABSTRACT

009 010  
011 Effective decision making in partially observable environments requires compressing long  
012 interaction histories into informative representations. We introduce *descriptive history*  
013 *representations (DHRs)*: sufficient statistics characterized by their capacity to answer  
014 *relevant questions* about past interactions and potential future outcomes. DHRs focus on  
015 capturing the information necessary to address task-relevant queries, providing a structured  
016 way to summarize a history for optimal control. We propose a learning framework, in-  
017 volving representation, decision, and question-asking components, optimized using a joint  
018 objective that balances reward maximization with the representation’s ability to answer  
019 informative questions. This yields representations that capture the salient historical details  
020 and predictive structures needed for effective decision making. We validate our approach on  
021 recommendation tasks with public movie and shopping datasets, generating interpretable  
022 textual *user profiles* which serve as sufficient statistics for predicting preference-driven  
023 behavior of users.

## 024 1 INTRODUCTION

025 Reinforcement learning (RL) in partially observable environments is challenging: agents must make decisions  
026 based on long interaction histories. This is particularly acute in domains like recommender systems, where  
027 an agent must infer a user’s latent preferences (the unobserved state) from their observed behavior to make  
028 personalized decisions (Åström, 1965; Sondik, 1971; Kaelbling et al., 1998). Representing this history in a  
029 compact, informative way is crucial. Traditional approaches in partially observable Markov decision processes  
030 (POMDPs) use *belief states* (Åström, 1965; Sondik, 1971; Kaelbling et al., 1998), but these often require a  
031 pre-specified Markovian state space, which is impractical for complex user modeling. Conversely, directly  
032 using the entire history becomes computationally intractable as it grows. *Predictive state representations*  
033 (*PSRs*) (Littman and Sutton, 2001; Boots et al., 2011) address this by representing the state as a set of  
034 predictions about future observations, conditioned on specific action sequences.

035 In this work, we introduce *descriptive history representations (DHRs)*, a framework that learns representations  
036 that focus on *answering questions*. Instead of predicting low-level observations, we construct history  
037 representations that answer broad classes of predictive questions. These questions, ideally formulated flexibly  
038 (e.g., in natural language), aim to capture essential information for effective decision-making. DHRs shift the  
039 representation burden away from predicting specific, low-level observations to a higher level of abstraction,  
040 ensuring they answer *relevant queries*. For example, consider a recommender agent that interacts with users.  
041 A *user profile*, representing a user’s past interactions with the agent, should suffice to answer queries such as:  
042 “Does the user prefer Item 1 to Item 2?”; “How likely is the user to abandon the session?”; or open queries  
043 like “Write a review this user might provide for the following product: <description>.” Such queries can  
044 be far more natural, task-relevant, and easier to reason about, than the precise sequence of low-level future  
045 observations (e.g., specific interactions) that simply correlate with user satisfaction.

047 To *learn* DHRs, we employ a multi-agent cooperative training paradigm. A *representation (DHR) encoder*  
 048 learns to construct a summary of the history, while an *answer agent* learns to generate answers to queries  
 049 about the history / future using *only* the generated summary. Finally, a *decision agent* determines the next  
 050 action to take given the current DHR summary. The agents' objectives are aligned with expected reward  
 051 maximization in the underlying environment.

052 Our key contributions are as follows. (1) We introduce DHRs, defined by their ability to answer relevant  
 053 questions about past interactions and future outcomes. We formally establish them as *sufficient statistics*  
 054 (formally defined in Section 2) for effective decision-making. (2) We propose a multi-agent framework for  
 055 learning DHRs, using representation, decision, and question-asking components, jointly optimized to balance  
 056 reward maximization with the representation's ability to answer informative questions. (3) We demonstrate  
 057 the efficacy of the learned DHRs on recommendation domains, using public movie and shopping datasets,  
 058 showcasing their ability to generate predictive textual user profiles and improve recommendation quality.

## 059 2 PROBLEM SETTING

060 We consider a *partially observable environment (POE)* defined by the tuple  $(\mathcal{O}, \mathcal{A}, T)$ , where  $\mathcal{O}$  is the  
 061 observation space,  $\mathcal{A}$  is the action space, and  $T : (\mathcal{O} \times \mathcal{A})^* \mapsto \Delta_{\mathcal{O}}$  is the transition function, mapping  
 062 (action-observation) histories to a distribution over observations. We assume a *reward function* over histories  
 063  $R : (\mathcal{O} \times \mathcal{A})^* \mapsto \mathbb{R}$ . At each time  $t$ , the environment is in a history state  $h_t = (o_1, a_1, \dots, o_t)$ . An  
 064 agent takes an action  $a_t \in \mathcal{A}$ , receives reward  $r_t = R(h_t, a_t)$ , and the environment transitions to a new  
 065 state via  $o_{t+1} \sim T(\cdot|h_t, a_t)$ . A *policy*  $\pi : (\mathcal{O} \times \mathcal{A})^* \times \mathcal{O} \mapsto \Delta_{\mathcal{A}}$  maps histories to distributions over  
 066 actions. Let  $\Pi$  denote the set of policies. The *value* of policy  $\pi \in \Pi$  is its expected cumulative return:  
 067  $V(\pi) = \mathbb{E} \left[ \sum_{t=1}^H r(h_t, a_t) | a_t \sim \pi(h_t) \right]$ , where  $H$  is the horizon. Our objective is to find a policy  $\pi^*$  with  
 068 maximum value, i.e.,  $\pi^* \in \arg \max_{\pi} V(\pi)$ .

069 We use the notion of an *f-sufficient statistic* below. Let  $f : \mathcal{H} \times \Pi \mapsto \mathbb{R}$  be a specific function of interest. We  
 070 say a mapping  $E : \mathcal{H} \mapsto \mathcal{Z}$  is an *f*-sufficient statistic of  $\mathcal{H}$  if, for any  $h, \pi$ , there is some  $\tilde{f} : \mathcal{Z} \times \Pi \mapsto \mathbb{R}$   
 071 such that  $f(h; \pi) = \tilde{f}(E(h); \pi)$  (Rémond, 1984). Examples include value-sufficient statistics for POMDPs  
 072 such as belief state representations (Åström, 1965), and PSRs (Littman and Sutton, 2001). The main task of  
 073 our work is to learn such a history-to-state mapping that both effectively summarizes historical information  
 074 and is amenable to sequential decision making.

075 Given a probability function  $y \sim q(y|x)$ , we often write  $y \sim q_x$ . Let  $\mathcal{H}_t$  be the set of histories of length  $t$ ,  
 076 and  $\mathcal{H} = \bigcup_{t=1}^H \mathcal{H}_t$ . Similarly,  $\Omega_t$  is the set of future realizations beginning at time  $t$ , with  $\Omega = \bigcup_{t=1}^H \Omega_t$ . Thus  
 077  $(h_t, \omega_t) \in \mathcal{H} \times \Omega$  such that  $h_t = (o_1, a_1, \dots, o_t)$  and  $\omega_t = (a_t, o_{t+1}, \dots, o_H)$ . For any  $\pi \in \Pi$ ,  $d^\pi(\omega|h)$  is  
 078 the probability of a future  $\omega$  given  $\pi$  starting at history state  $h$ .

## 079 3 DESCRIPTIVE HISTORY REPRESENTATIONS (DHRs)

080 This section formally defines DHRs. We first introduce the *Question-Answer-space (QA-space)*, which  
 081 structures historical information through questions and answers, and then detail how DHRs serve as compact,  
 082 actionable summaries derived from these spaces.

### 083 3.1 QA-SPACES

084 Informally, a *QA-space* comprises questions and answers over histories. For example, in a conversational  
 085 recommender, useful questions might query latent user preferences (e.g., “which brand does the user prefer?”),  
 086 with answers being (distributions over) preference attributes (e.g., “the user prefers brand A over B with  
 087 probability 0.8”). Alternatively, a question could probe user behavior (e.g., “Write a review this user might

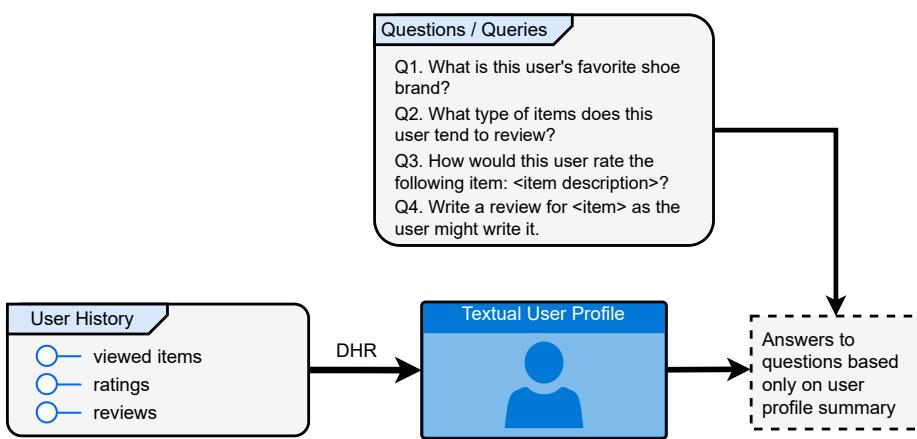


Figure 1: An illustrative example of a descriptive history representation (DHR) which maps a user's history to a compact user profile, which is sufficient for answering questions about the user's preferences.

provide for item A"), answered by a distribution over possible reviews. Formally, a QA-space defines questions and answers which adhere to a particular functional form, where answers are outcomes of questions.

**Definition 1** (QA-space). A QA-space is a tuple  $(\mathcal{Q}, \mathcal{Y}, \mathcal{X}, \nu)$ , where  $\mathcal{Q}$  is a question space,  $\mathcal{Y}$  is an answer space,  $\mathcal{X}$  is a context space, and  $\nu : \mathcal{X} \times \mathcal{Q} \mapsto \Delta_{\mathcal{Y}}$  is an answer function.<sup>1</sup>

It is important to note that the set of sufficient questions  $Q_x^*$  is determined by the context  $x$  (e.g., the history) and the underlying environment dynamics; it exists independently of any specific representation we aim to learn. While QA-spaces are general, in this work we focus on QA-spaces that ask *semantically meaningful* and *interpretable* questions about histories. Thus, the context space  $\mathcal{X}$  is the space of histories  $\mathcal{H}$ , and the answer function  $\nu$  poses informative questions about those histories (e.g., summarization of past interactions, or predictions of future events). Notice that the answer function  $\nu$  must generally return a *distribution* over the answer space given the partially observable nature of the environment.

We are primarily interested in QA-spaces where the collective answers to questions provide all necessary information for a given purpose (like decision-making or prediction), thereby acting as sufficient statistics. In the recommender setting, a QA-space with user preference questions can serve as a sufficient statistic for future behaviors (e.g., item acceptance, session abandonment) by encoding inferred latent preferences. This motivates the definition of a *sufficient QA-space*, which asks questions that induce a sufficient statistic.

**Definition 2** (*f*-Sufficient QA-space). A QA-space  $(\mathcal{Q}, \mathcal{Y}, \mathcal{X}, \nu)$  is *f*-sufficient, if for any  $x \in \mathcal{X}$ , there is a subset  $Q_x^* \subseteq \mathcal{Q}$  such that  $\{(q, \nu(x, q))\}_{q \in Q_x^*}$  is an *f*-sufficient statistic. We call  $Q_x^*$  the set of (*f*-) sufficient questions for context  $x$ , and  $Q^* = \bigcup_{x \in \mathcal{X}} Q_x^*$  the set of sufficient questions.

### 3.2 EXAMPLES OF QA-SPACES

**Belief States (Kaelbling et al. (1998)).** For POMDPs with state space  $\mathcal{S}$ , belief states  $b(s|h)$  are distributions over states which constitute a sufficient statistic for value prediction. We define a QA-space for belief states as follows. Let  $\mathcal{Q} = \mathcal{S}$ ,  $\mathcal{Y} = [0, 1]$ , and answer function  $\nu(h, s) = b(s|h)$ . For any history  $h \in \mathcal{H}$ , the set of question-answer pairs  $\{(s, b(s|h))\}_{s \in \mathcal{S}}$  constitutes a sufficient QA-space.

<sup>1</sup>When  $\mathcal{Y}$  is continuous,  $\nu$  maps questions and contexts to probability measures on the Borel sets of  $\mathcal{Y}$ .

**Predictive State Representations (PSRs, Littman and Sutton (2001)).** A PSR is defined by a vector of probabilities over a set of core tests  $(P(\omega^1|h), \dots, P(\omega^k|h))$ , for any history  $h \in \mathcal{H}$ . Like belief states, PSRs are sufficient statistics. To construct a QA-space for PSRs, let  $\mathcal{Q}$  be the set of tests,  $\mathcal{Y} = [0, 1]$  as before, and define an answer function as  $\nu(h, \omega) = P(\omega|h)$ .

**Item Recommendation (informal).** Consider a user interacting with a recommender, where actions are recommended item slates, and observations are user interactions (e.g., choices, ratings, etc.). The recommender exploits a *user profile* representing user preferences learned from interactions (see Fig. 1). Given a user’s history and candidate items, useful questions assess preferences over pairs of items, or the likelihood of accepting recommendations. A sufficient statistic includes answers which are predictive of preferences, behavior, and suffice for optimal recommendations.

### 3.3 DESCRIPTIVE HISTORY REPRESENTATIONS OF QA-SPACES

While QA-spaces offer a powerful mechanism for identifying vital information within a history  $h$  by finding answers  $\nu(h, q)$  to a set of sufficient questions, they do not inherently define how this distilled information should be structured into an effective representation. We next leverage the insights obtained from QA-spaces to learn a condensed representation, which retains the information required to answer *sufficient* questions.

**Definition 3** (Descriptive History Representation). *Let  $(\mathcal{Q}, \mathcal{Y}, \mathcal{H}, \nu)$  be an  $f$ -sufficient QA-space. An embedding  $E : \mathcal{H} \mapsto \mathcal{Z}$  is called a Descriptive History Representation (DHR) if there exists  $\nu_A : \mathcal{Z} \times \mathcal{Q} \mapsto \Delta_{\mathcal{Y}}$  such that  $\nu_A(z, q) = \nu(h, q)$ , for any  $h \in \mathcal{H}$ , where  $z = E(h)$ , and any sufficient question  $q \in Q_h^*$ .*

A DHR encapsulates the essence of the history for question answering by learning a representation  $E : \mathcal{H} \mapsto \mathcal{Z}$  that maps a history  $h \in \mathcal{H}$  to a compact representation  $z \in \mathcal{Z}$ . In general,  $\mathcal{Z}$  can be arbitrary; e.g., if  $\mathcal{Z} \subseteq \mathbb{R}^d$ , then  $E$  is a classic embedding. In our recommendation example,  $z \in \mathcal{Z}$  could be a *textual user profile*, describing a user’s preferences in natural language (see Fig. 1 for an illustration). The central idea of a DHR is that the representation acts as a proxy for the history, preserving precisely the information required to answer the relevant questions defined by the QA-space. Thus, instead of computing  $\nu(h, q)$  from the history  $h$ , a *compressed* answer function  $\nu_A : \mathcal{Z} \times \mathcal{Q} \rightarrow \Delta_{\mathcal{Y}}$  produces the same answer distribution using only  $z = E(h)$ . A DHR, denoted by  $z = E(h)$ , is a powerful concept because it effectively compresses a potentially complex history  $h$  into a more compact form  $z$ . By preserving the means to answer sufficient questions, the DHR encapsulates all task-relevant information defined by the QA-space, suggesting that it is a sufficient statistics of a given task. Indeed, the following theorem establishes this connection (see Appendix E for proof).

**Theorem 1.** *Let  $E : \mathcal{H} \mapsto \mathcal{Z}$  be a DHR (Definition 3). Then it is also an  $f$ -sufficient statistic.*

## 4 LEARNING DESCRIPTIVE HISTORY REPRESENTATIONS

Our goal is to learn a DHR that is effective for downstream decision making. Specifically, we learn a *DHR policy*  $\pi = (E, \pi_D)$  consisting of two components: a *DHR embedding*  $E : \mathcal{H} \mapsto \Delta_{\mathcal{Z}}$  (i.e., a history *encoder*); and a *decision policy*  $\pi_D : \mathcal{Z} \mapsto \Delta_{\mathcal{A}}$  over summarized histories. The DHR policy induces a policy over histories in the obvious way:  $\pi(h) = \mathbb{E}_{z \sim E(h)}[\pi_D(z)]$ . To ensure  $E$  is indeed a DHR (i.e., a sufficient statistic), we train an *answer function*  $\nu_A : \mathcal{Z} \times \mathcal{Q} \mapsto \mathcal{Y}$  for any sufficient question  $q \in Q_h^*$  (see Definition 3).

**Generating Sufficient Questions and Answers.** To train  $\nu_A$ , we need to define the set of sufficient questions  $Q^*$ . We introduce an oracle *QA-generator*—only required during training—which provides ground-truth question-answer pairs  $(q, y)$  that define the content used to train the DHR. **This follows the paradigm of Learning Using Privileged Information (LUPI) (Vapnik et al., 2015), where extra information is available during training but not at test time.** To obtain a QA-generator, we use full trajectory realizations; i.e., a trajectory  $\tau := (h, \omega)$ , where  $h$  is the history, and  $\omega$  is its future realization. The QA generator leverages both

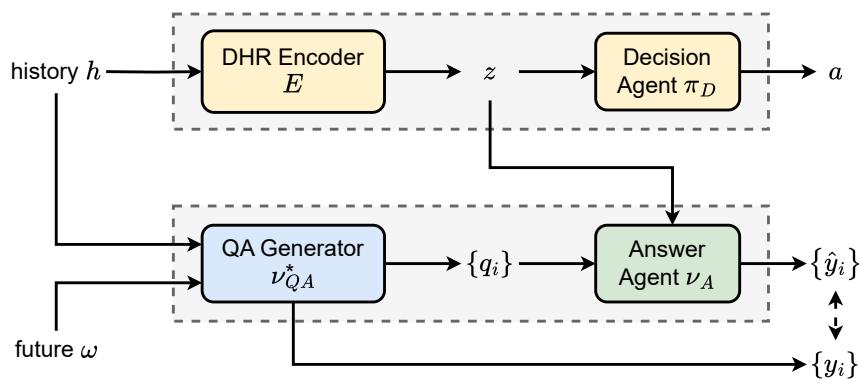


Figure 2: An illustration of our learning framework. A policy  $\pi$  is composed of a DHR encoder  $E$  and a decision policy  $\pi_D$ . The DHR embedding takes a history  $h$  and creates a representation  $z$ . The decision policy takes actions in the environment given the representation  $z$ . In order to learn a DHR ( $E$ ), we use a QA generator, responsible for generating history ( $h$ ) and future ( $\omega$ )-dependent question-answer pairs. Finally, an answer agent is responsible for answering said questions with access to only the representation  $z$ . The training procedure is described in Sec. 4 and Algorithm 1.

$h$  and  $\omega$  (the privileged information) to construct questions and their corresponding ground-truth answers. Designing this oracle can be challenging. In Sec. 5, we detail how we design a QA-generator using LLMs and future user interactions for recommendation tasks. We refer to Appendix B for a detailed discussion on the construction and training of a QA-generator. Formally, a QA generator is a mapping  $\nu_{QA}^* : \mathcal{H} \times \Omega \mapsto \Delta_{\mathcal{Q} \times \mathcal{Y}}$ , which we decompose as  $\nu_{QA}^*(h, \omega) = \nu_A^*(y|q, h, \omega)\nu_Q^*(q|h, \omega)$ .

**Training Objective.** We define the joint objective of the DHR encoder, answer agent, and decision agent to maximize value and match the distribution of ground-truth answers (i.e., match  $\nu_A^*$ ):

$$\max_{E, \nu_A, \pi_D} (1 - \lambda)V(\pi) - \lambda D_f \left( d^{\nu_A^*} \parallel d^{\nu_A} \right), \quad (\text{OPT 1})$$

where  $\lambda \in [0, 1]$  balances between RL and DHR learning, and the distributions are defined as  $d^{\nu_A}(y, h, q) = P(y, h, q \mid q \sim \nu_Q^*(h, \omega), y \sim \nu_A(z, q), z \sim E(h), \omega \sim d^\pi(\omega \mid h))$ , and  $d^{\nu_A^*}(y, h, q) = P(y, h, q \mid q, y \sim \nu_{QA}^*(h, \omega), \omega \sim d^\pi(\omega \mid h))$ . Here,  $D_f$  denotes a specific  $f$ -divergence. Directly optimizing Eq. (OPT 1) is challenging, so we solve its variational form. Let  $f^*$  be the convex conjugate of  $f$ , then the dual formulation of Eq. (OPT 1) is:

$$\max_{\pi_D, E, \nu_A} \min_{g: \mathcal{H} \times \mathcal{Q} \times \mathcal{Y} \mapsto \mathbb{R}} \mathbb{E} \left[ (1 - \lambda)r(h, a) + \lambda \mathbb{E}_{y \sim \nu_A(q, h, \omega)} [f^*(g(h, q, y))] - \lambda \mathbb{E}_{y \sim \nu_A^*(z, q)} [g(h, q, y)] \right]. \quad (\text{OPT 2})$$

This max-min objective can be solved by iteratively training a discriminator  $g$  and the agents  $(\pi_D, E, \nu_A)$ . Note that when the QA-generator  $\nu_{QA}^*$  is fixed (our primary setting), the target distribution  $d^{\nu_A^*}$  is stationary. The optimization in Eq. (OPT 2) thus becomes a stable density ratio estimation task, avoiding the instability common in adversarial methods (e.g., GANs) where the target distribution is moving.

**DHR Learning (DHRL).** We now detail the DHRL algorithm (Algorithm 1). The framework is depicted in Fig. 2. DHRL trains the DHR encoder, answer agent, decision agent, and discriminator, with parameters  $\theta_E, \theta_A, \theta_D, \theta_g$ . It first samples trajectories using  $E$  and  $\pi_D$  (line 3). These are used in hindsight: each

235 **Algorithm 1** Descriptive History Representation Learning (DHRL)

---

```

236
237 1: Initialize  $\theta_E, \theta_A, \theta_D, \theta_g$ .
238 2: for each training iteration do
239 3:   Sample trajectory  $\tau = (o_1, a_1, \dots, o_H)$  and rewards  $r_1, \dots, r_H$ .  $\triangleright$  Using  $E, \pi_D$ 
240 4:   for  $t = 1, \dots, H$  do
241 5:     Split trajectory to history  $h_t$  and future  $\omega_t$ .
242 6:     Sample QA pairs  $\{(q_{kt}, y_{kt})\}_{k=1}^K \sim \nu_{QA}^*(h_t, \omega_t)$ .
243 7:     Sample representation  $z_t \sim E(h_t)$ .
244 8:     Predict answers  $\hat{y}_{kt} = \nu_A(z_t, q_{kt})$  for  $k = 1 \dots K$ .
245 9:   end for
246 10:  Update  $\theta_g$  according to  $\frac{1}{HK} \sum_{t=1}^H \sum_{k=1}^K \nabla_{\theta_g} [f^*(g_{\theta_g}(h_t, q_{kt}, \hat{y}_{kt})) - g_{\theta_g}(h_t, q_{kt}, y_{kt})]$ .
247 11:  Update  $\theta_D, \theta_A, \theta_E$  via RL using rewards  $r^D, r^A, r^E$  (See Sec. 4).
248 12: end for
249

```

---

250 trajectory is split into histories  $h_t$  and futures  $\omega_t$  (line 5). The QA-generator  $\nu_{QA}^*$  creates ground-truth pairs  
 251  $(q_k, y_k)$  (line 6). A representation  $z_t \sim E(h_t)$  is sampled (line 7), and the answer agent predicts answers  
 252  $\hat{y}_k = \nu_A(z_t, q_k)$  (line 8). The collected data is used to update the discriminator  $g$  (line 10). The agents are  
 253 updated using RL (line 11), with reward signals derived from Eq. (OPT 2). Specifically, the decision agent  
 254  $\pi_D$  uses reward  $r^D = r(h, a)$ , the answer agent uses reward  $r^A = f^*(g(h, q, \hat{y}))$ , and the DHR embedding  
 255 uses reward  $r^E = (1 - \lambda)r^D + \lambda r^A$ . The DHRL framework is flexible and supports both *online* and *offline*  
 256 training paradigms. In an online setting, trajectories are actively sampled from the environment using the  
 257 current policy. In an offline setting, they are simply drawn from a fixed, pre-collected dataset. **Crucially, the**  
 258 **QA-generator, answer agent, and discriminator are used only during training.** During deployment the agent  
 259 *does not have access* to the future. It simply uses the DHR encoder on the current history  $h_t$  to produce the  
 260 representation  $z_t$  for the decision agent. The future trajectory is only required for training, not execution,  
 261 **making the inference process efficient.**

262 **5 EXPERIMENTS**

263 Our experiments are designed to validate DHRL, focusing on its ability to generate predictive representations  
 264 that answer relevant questions. We demonstrate that DHRL can produce high-quality DHRs in recommendation  
 265 domains, where we use DHRs to generate textual user profiles which (1) accurately answer questions  
 266 about user preferences and future behavior; (2) induce informative, coherent summaries of user history; and  
 267 (3) effectively support a downstream recommendation task.

268 **5.1 EXPERIMENTAL SETUP**

269 **Datasets and Task.** We use two datasets in our experiments: MovieLens 25M (Harper and Konstan, 2015),  
 270 and Amazon Reviews (Ni et al., 2019) (specifically, the *Clothing, Shoes and Jewelry* category). User histories  
 271  $h_t$  consist of sequences of observations (item titles, ratings, and, for Amazon, descriptions, prices and reviews)  
 272 and actions (recommended items). This history serves as input to the encoder  $E$ , which creates a textual user  
 273 profile. Future user interactions  $\omega_t$  are only used in hindsight to construct the QA pairs for training.

274 **Question Generation.** For both MovieLens and Amazon, questions  $q_k$  compare pairs of held-out movies  
 275 (e.g., “Rank the following movies based on the user’s preferences: [<movie\_1>, <movie\_2>]”), with the  
 276 ground-truth answer  $y_k$  taken from future user ratings in  $\omega_t$ . For Amazon Reviews, we augment the question  
 277 set with review generation queries  $q_{\text{review}}$  which ask “Write a review for item <item description> as <user>  
 278 might write it.” Here, the ground-truth answer  $y_{\text{review}}$  is the user’s textual review of that item from  $\omega_t$ .

282 Table 1: Performance of DHRL with default settings (256-token profile, 5 questions, 10 interactions, TV-divergence)  
 283 after 1,000 iterations for Gemma V3 4B, 12B, and Gemini 1.5 Flash models. **Bold** indicates statistically significant  
 284 improvement over the corresponding baseline ( $p < 0.05$ ). See Appendix D for confidence intervals.

286 <b>Model &amp; Method</b>	287 <b>Prediction Accuracy (w.r.t. GT)</b>	288 <b>Rec. Reward (<math>r^D</math>)</b>	289 <b>Profile-History Consistency (AI/Human)</b>	290 <b>Prediction Fidelity (AI/Human)</b>	291 <b>Review Quality (AI/Human)</b>
292 Task: recommendation using Amazon products user profiles					
Gemma V3 4B	0.34	0.54	3.46 / 3.28	0.44 / 0.47	0.15 / 0.27
Gemma V3 4B+DHRL	<b>0.71</b>	<b>0.83</b>	<b>3.41 / 3.85</b>	<b>0.56 / 0.53</b>	<b>0.85 / 0.73</b>
Gemma V3 12B	0.67	0.78	4.32 / 4.18	0.34 / 0.39	0.17 / 0.35
Gemma V3 12B+DHRL	<b>0.75</b>	<b>0.84</b>	<b>4.39 / 4.37</b>	<b>0.66 / 0.61</b>	<b>0.83 / 0.65</b>
Gemini 1.5 Flash	0.69	0.80	4.41 / 4.30	0.38 / 0.41	0.17 / 0.38
Gemini 1.5 Flash+DHRL	<b>0.74</b>	<b>0.86</b>	<b>4.46 / 4.34</b>	<b>0.62 / 0.59</b>	<b>0.83 / 0.62</b>
300 Task: recommendation using MovieLens user profiles					
Gemma V3 4B	0.37	0.61	3.72 / 3.96	0.32 / 0.45	-
Gemma V3 4B+DHRL	<b>0.74</b>	<b>0.79</b>	<b>3.84 / 4.24</b>	<b>0.68 / 0.55</b>	-
Gemma V3 12B	0.58	0.64	4.66 / 4.58	0.35 / 0.42	-
Gemma V3 12B+DHRL	<b>0.84</b>	<b>0.93</b>	<b>4.62 / 4.71</b>	<b>0.76 / 0.66</b>	-
Gemini 1.5 Flash	0.62	0.68	4.69 / 4.53	0.38 / 0.39	-
Gemini 1.5 Flash+DHRL	<b>0.82</b>	<b>0.92</b>	<b>4.74 / 4.70</b>	<b>0.77 / 0.62</b>	-

305 **Training Methodology.** We train  $E$  and  $\nu_A$  via standard policy gradient (PG). We use Gemini 1.5 Flash  
 306 (Team et al., 2024) and Gemma V3 (4B, 12B) (Team et al., 2025) as our language models. All agents use the  
 307 same base model and share parameters. Our decision agent  $\pi_D$  is fine-tuned to make item recommendations  
 308 to the user given their generated profile (DHR), receiving a reward based on the user’s rating for the item.  
 309 Unless otherwise stated, we use a Gemma V3 4B model, a 256-token profile limit,  $K = 5$  item comparison  
 310 questions, one review question (Amazon only), a history length of  $H = 10$  interactions, and TV-distance.

311 **Evaluation Metrics.** The generated DHRs (user profiles  $z$ ) and the accuracy of the answer agent  $\nu_A$  are  
 312 evaluated using three main criteria: predictive accuracy, recommendation reward, and human evaluation.

313 For ranking questions, which compare pairs of items, the *predictive accuracy* is the fraction of pairwise  
 314 comparisons the agent predicts correctly with respect to ground-truth future user ratings. We also define a  
 315 *recommendation reward* to assess the utility of generated DHRs for our downstream task. Specifically, given  
 316 a DHR and a held-out set of unseen items,  $\pi_D$  recommends an item. The reward is defined by the user’s true  
 317 rating for this item (normalized to  $[0, 1]$ ).

318 Finally, we use Gemini 2.5 Pro and 24 human raters<sup>2</sup> to assess three qualitative aspects: (1) *Profile-History*  
 319 *Consistency*: Raters are given a user interaction history  $h_t$  and its generated textual profile  $z_t$ . They rate (1-5  
 320 scale) how accurately and coherently  $z_t$  summarizes the information in  $h_t$ . (2) *Prediction Fidelity*: Raters  
 321 are given a generated user profile  $z_t$  and asked to predict the user’s preferences over held-out item pairs. We  
 322 report the win-rate between the rater accuracy for a baseline-generated profile vs. its DHRL counterpart. Note  
 323 that human performance on this task may be inherently limited. (3) *Review Quality* (Amazon only): Raters  
 324 are shown the profile  $z_t$  and two reviews for a held-out item, one from a baseline model and one from the  
 325 DHRL model. We report the win-rate for the DHRL-generated review, indicating its perceived authenticity.

326  
 327 <sup>2</sup>Raters were paid contractors. They received their standard contracted wage, which is above the living wage in their  
 328 country of employment.

329  
330 Table 2: Comparison with state-of-the-art recommender system baselines. Our DHRL-enhanced model outperforms  
331 both the base LLM and specialized RecSys models. **Results significantly outperform baselines (p<0.05).**

332	333	334	Dataset	Model & Method	335	Prediction	Rec.
336	337	338	Amazon	Gemma V3 12B (Baseline)	339	Accuracy	Reward
340	341	342	MovieLens	DCLMDB (Huang et al., 2024)	343	(w.r.t. GT)	( $r^D$ )
343	344	345	Amazon	DCRec(Yang et al., 2023)	346	0.71	0.81
344	345	346	Amazon	<b>Gemma V3 12B + DHRL (ours)</b>	347	<b>0.75</b>	<b>0.84</b>
345	346	347	MovieLens	Gemma V3 12B (Baseline)	348	0.58	0.64
346	347	348	MovieLens	DCLMDB (Huang et al., 2024)	349	0.80	0.88
347	348	349	MovieLens	DCRec(Yang et al., 2023)	350	0.78	0.85
348	349	350	MovieLens	<b>Gemma V3 12B + DHRL (ours)</b>	351	<b>0.84</b>	<b>0.93</b>

## 344 5.2 RESULTS

345 Table 1 summarizes the performance of DHRL under our default configuration. Incorporating DHRL leads to  
346 substantial improvements in predictive accuracy for ranking tasks. Furthermore, AI and human evaluators  
347 rated the generated profiles favorably for consistency with user history and preference predictive ability,  
348 suggesting that textual DHRs encapsulate relevant user information effectively. We find human and AI  
349 evaluation results to be highly correlated. The notably higher recommendation rewards attained by DHRL-  
350 enhanced models confirm that DHR representations are beneficial for downstream decision-making. For  
351 Amazon Reviews, review quality metrics also indicate a strong preference for reviews conditioned on DHR  
352 profiles over baseline-generated reviews.

353 We compare our DHRL framework against state-of-the-art methods for recommendation, including:  
354 DCLMDB (Huang et al., 2024) and DCRec (Yang et al., 2023). **The “Gemma V3 (Baseline)” uses the**  
355 **same LLM backbone trained on the same data via standard predictive objectives.** The results in Table 2 show  
356 that while these methods improve upon the base LLM, our DHRL framework consistently achieves superior  
357 performance in both prediction accuracy and recommendation reward. This demonstrates that DHRL is not  
358 only a general representation learning framework but also a highly competitive method in the recommendation  
359 domain.

360 The impact of history length on reward is shown in Fig. 3 (left). DHRL consistently outperforms the baseline.  
361 Notably, optimal performance for DHRL is observed with shorter histories (5-10 interactions). This suggests  
362 that relatively concise histories suffice for learning effective DHRs in our benchmark tasks, and excessively  
363 long histories may introduce noise. Fig. 3 (middle) shows the impact of DHR profile length. Performance  
364 generally improves as the token limit increases from shorter lengths, allowing more detailed information.  
365 However, this trend does not continue indefinitely; beyond an optimal length (around 256 tokens), gains  
366 diminish. This might stem from challenges in exploiting an overly large representation space. Fig. 3 (right)  
367 shows that recommendation reward generally increases with  $K$ , the number of guiding questions. This  
368 emphasizes the importance of using a sufficiently diverse set of questions to guide the representation learning  
369 process towards a more robust representation of user preferences.

370 We further validate two key aspects of the DHRL framework via ablation studies (see Appendix D.1). First,  
371 we confirm that the DHR closely approximates a sufficient statistic: providing the decision agent with access  
372 to the raw history in addition to the DHR ( $\pi_D(z_t, h_t)$ ) yielded negligible improvement (+0.5% reward) over  
373 using the DHR alone ( $\pi_D(z_t)$ ). Second, we demonstrate the necessity of the joint optimization (OPT 1).  
374 Training the representation using only the QA objective ( $\lambda = 1$ ) significantly underperformed the full DHRL  
375 method, confirming that balancing the QA framework with reward maximization is crucial.

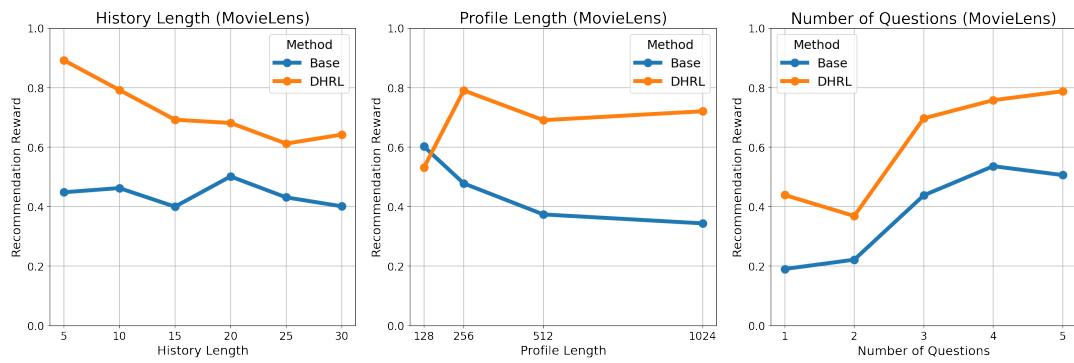


Figure 3: Ablation study on the history length (left), profile length (middle) and number of questions (right) on the MovieLens dataset, showing recommendation reward. DHRL (orange) consistently outperforms the baseline (blue). Short histories (5-10 interactions) are often sufficient. Performance peaks at an intermediate profile length (e.g., 256 tokens). More questions generally improve recommendation quality.

We also experimented with **learning the QA-generator** (see Appendix B). While this led to a marginal improvement in reward (2-3%), the fixed QA-generator bootstrapped with a powerful, pre-trained LLM was already remarkably effective. This highlights the practical applicability of our approach, leveraging off-the-shelf LLMs to create a rich QA-space without complex adversarial training.

Finally, the choice of  $f$ -divergence significantly impacts performance. Among the tested divergences for the Amazon dataset, TV-distance yielded the best results with a Prediction Accuracy of 0.71 and a Recommendation Reward of 0.83. This substantially outperformed both  $\chi^2$ -divergence (0.42 accuracy, 0.61 reward) and KL-divergence (0.40 accuracy, 0.58 reward), highlighting that the specific mechanism used to align the answer agent’s output distribution with the target distribution is a critical factor in DHRL.

## 6 RELATED WORK

Learning history representations is crucial for POMDPs (Kaelbling et al., 1998). PSRs (Littman and Sutton, 2001) use predictions of future action-observation sequences (“tests”), though as discussed above, rely on manually engineered, often low-level tests. In contrast, DHRs learn representations based on answering high-level, semantically meaningful questions about the future, shifting the focus to task-relevant understanding.

Predictive principles have been adapted to deep learning methodologies. Predictive-state decoders (Venkatraman et al., 2017) and PSRNNs (Downey et al., 2017) predict future observations, features, or latent states. While leveraging prediction, they typically focus on low-level features, in contrast to DHRs, which use a broader, more abstract set of questions (the QA-space) to shape representations, ensuring they answer informative queries beyond standard next-observation prediction. **This QA-space can be viewed as learning a set of General Value Functions (GVFs) (Sutton et al., 2011), where questions define the cumulant and policy.**

Recent work has emphasized self-prediction in latent space (Ni et al., 2024; Schwarzer et al., 2020), learning representations by predicting future latent states or values. These methods yield latent representations implicitly defined by the prediction objective. DHRs, via the QA framework, explicitly define representation content through their questions, allowing for greater interpretability and control. Other approaches approximate belief states, like the Wasserstein Belief Updater (Avalos et al., 2023), or use frozen pre-trained LLMs for history compression (Paischer et al., 2023). DHRs use a multi-agent framework to learn the representation encoder jointly with answer and decision modules, guided by reward maximization and question-answering accuracy.

423 Finally, LLMs have been used to generate textual user profiles (Hou et al., 2024; Zhang et al., 2023) to enhance  
 424 interpretability. DHRs differ as they are directly optimized to serve as sufficient statistics, with information  
 425 encoded not just by generative summarization, but by their ability to answer task-relevant predictive questions.  
 426

## 427 7 CONCLUSION

430 This paper introduced descriptive history representations (DHRs), a framework for learning history summaries  
 431 by focusing on their ability to answer task-relevant questions. We demonstrated its effectiveness on challenging  
 432 recommendation domains, where our learning approach generates interpretable textual user profiles that act  
 433 as sufficient statistics. These profiles lead to strong predictive accuracy and outperform powerful baselines in  
 434 downstream recommendation tasks.

435 Our work yields two key insights. First, DHRs offer a compelling method for representation learning that  
 436 enhances interpretability and aligns representations with high-level goals via an explicit QA-space. This  
 437 provides a structured way to inject task-relevant priors into the learning process, moving beyond simple  
 438 predictive objectives. Second, we find that modern LLMs are effective at generating semantically rich  
 439 questions, serving as powerful, off-the-shelf QA-generators, which makes the DHR framework highly  
 440 practical for a wide range of applications. Future work can explore learning the QA-generator for domains  
 441 where strong priors are unavailable, and applying the DHR framework to other **partially observable domains**  
 442 such as robotics (where questions might resemble GVF), human-computer interaction, and dialogue systems.

## 443 REFERENCES

445 Karl Johan Åström. Optimal control of markov processes with incomplete state information i. *Journal of*  
 446 *mathematical analysis and applications*, 10:174–205, 1965.

447 Raphael Avalos, Florent Delgrange, Ann Nowé, Guillermo A Pérez, and Diederik M Roijers. The wasserstein  
 448 believer: Learning belief updates for partially observable environments through reliable latent space models.  
 449 *arXiv preprint arXiv:2303.03284*, 2023.

450 Byron Boots, Sajid M Siddiqi, and Geoffrey J Gordon. Closing the learning-planning loop with predictive  
 451 state representations. *International Journal of Robotics Research*, 30(7):954–966, 2011.

452 Carlton Downey, Ahmed Hefny, Byron Boots, Geoffrey J Gordon, and Boyue Li. Predictive state recurrent  
 453 neural networks. *Advances in Neural Information Processing Systems*, 30, 2017.

454 F Maxwell Harper and Joseph A Konstan. The movielens datasets: History and context. *Acm transactions on*  
 455 *interactive intelligent systems (tiis)*, 5(4):1–19, 2015.

456 Yupeng Hou, Junjie Zhang, Zihan Lin, Hongyu Lu, Ruobing Xie, Julian McAuley, and Wayne Xin Zhao.  
 457 Large language models are zero-shot rankers for recommender systems. In *European Conference on*  
 458 *Information Retrieval*, pages 364–381. Springer, 2024.

459 Zhirong Huang, Shichao Zhang, Debo Cheng, Jiuyong Li, Lin Liu, and Guixian Zhang. Debiased contrastive  
 460 representation learning for mitigating dual biases in recommender systems. *CoRR*, 2024.

461 Leslie Pack Kaelbling, Michael L Littman, and Anthony R Cassandra. Planning and acting in partially  
 462 observable stochastic domains. *Artificial intelligence*, 101(1-2):99–134, 1998.

463 Michael Littman and Richard S Sutton. Predictive representations of state. *Advances in neural information*  
 464 *processing systems*, 14, 2001.

470 Jianmo Ni, Jiacheng Li, and Julian McAuley. Justifying recommendations using distantly-labeled reviews and  
 471 fine-grained aspects. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language  
 472 Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP  
 473 2019*, pages 188–197, 2019.

474

475 Tianwei Ni, Benjamin Eysenbach, Erfan SeyedSalehi, Michel Ma, Clement Gehring, Aditya Mahajan, and  
 476 Pierre-Luc Bacon. Bridging state and history representations: Understanding self-predictive rl. *arXiv  
 477 preprint arXiv:2401.08898*, 2024.

478

479 Fabian Paischer, Thomas Adler, Markus Hofmarcher, and Sepp Hochreiter. Semantic helm: A human-readable  
 480 memory for reinforcement learning. *Advances in Neural Information Processing Systems*, 36:9837–9865,  
 481 2023.

482

483 Marcel Rémon. On a concept of partial sufficiency: L-sufficiency. *International Statistical Review*, pages  
 127–135, 1984.

484

485 Max Schwarzer, Ankesh Anand, Rishab Goel, R Devon Hjelm, Aaron Courville, and Philip Bachman. Data-  
 486 efficient reinforcement learning with self-predictive representations. *arXiv preprint arXiv:2007.05929*,  
 487 2020.

488

489 Edward Jay Sondik. *The optimal control of partially observable Markov processes*. Stanford University,  
 1971.

490

491 Richard S Sutton, Joseph Modayil, Michael Delp, Thomas Degris, Patrick M Pilarski, Adam White, and Doina  
 492 Precup. Horde: A scalable real-time architecture for learning knowledge from unsupervised sensorimotor  
 493 interaction. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume  
 494 2*, pages 761–768, 2011.

495

496 Gemini Team, Petko Georgiev, Ving Ian Lei, Ryan Burnell, Libin Bai, Anmol Gulati, Garrett Tanzer, Damien  
 497 Vincent, Zhufeng Pan, Shibo Wang, et al. Gemini 1.5: Unlocking multimodal understanding across millions  
 498 of tokens of context. *arXiv preprint arXiv:2403.05530*, 2024.

499

500 Gemma Team, Aishwarya Kamath, Johan Ferret, Shreya Pathak, Nino Vieillard, Ramona Merhej, Sarah  
 501 Perrin, Tatiana Matejovicova, Alexandre Ramé, Morgane Rivière, et al. Gemma 3 technical report. *arXiv  
 502 preprint arXiv:2503.19786*, 2025.

503

504 Vladimir Vapnik, Rauf Izmailov, et al. Learning using privileged information: Similarity control and  
 505 knowledge transfer. *J. Mach. Learn. Res.*, 16(1):2023–2049, 2015.

506

507 Arun Venkatraman, Nicholas Rhinehart, Wen Sun, Lerrel Pinto, Martial Hebert, Byron Boots, Kris Kitani,  
 508 and J Bagnell. Predictive-state decoders: Encoding the future into recurrent networks. *Advances in Neural  
 509 Information Processing Systems*, 30, 2017.

510

511 Yuhao Yang, Chao Huang, Lianghao Xia, Chunzhen Huang, Da Luo, and Kangyi Lin. Debiased contrastive  
 512 learning for sequential recommendation. In *Proceedings of the ACM web conference 2023*, pages 1063–  
 513 1073, 2023.

514

515 Junjie Zhang, Ruobing Xie, Yupeng Hou, Xin Zhao, Leyu Lin, and Ji-Rong Wen. Recommendation as  
 516 instruction following: A large language model empowered recommendation approach. *ACM Transactions  
 on Information Systems*, 2023.

---

**A SOCIETAL IMPACT STATEMENT**

517 Descriptive history representations (DHRs), particularly when viewed as interpretable textual user profiles,  
518 offer specific societal benefits but also pose certain risks. We believe the benefits outweigh the risks if  
519 specific risk-mitigation strategies are employed in the use of DHRs. On the positive side, this approach can  
520 significantly enhance the transparency of user models in partially observable environments like recommender  
521 systems. By generating textual summaries that explain the basis for decisions, users might gain a clearer  
522 understanding of why certain recommendations are made, fostering trust and potentially allowing for more  
523 informed interactions. Moreover, DHRs can support a greater degree of user control and agency. The QA-  
524 space itself, guiding the DHR to focus on task-relevant information, could lead to more nuanced and genuinely  
525 helpful personalization, as the system learns to explicitly answer questions about user preferences and future  
526 behavior rather than just correlations with low-level signals. This explicit focus on question-answering for  
527 representation learning offers a pathway to models whose internal reasoning is more transparent.

528 However, the very interpretability and specificity of DHR-generated textual profiles introduce distinct  
529 challenges. Such concentrated, human-readable summaries of user history, even if aimed at better decision-  
530 making, could inadvertently reveal sensitive information or enable more precise inference of private attributes  
531 if not carefully managed, posing additional privacy risks beyond those of opaque embedding-based profiles.  
532 The content and biases of these textual profiles are also directly influenced by the choice of questions in the QA-  
533 space and the underlying language models used for generation—an improperly designed QA-space or biased  
534 LLM could lead to profiles that amplify societal biases or misrepresent users in an understandable but harmful  
535 way. Furthermore, while textual profiles aim for clarity, they are still summaries, and the nuances captured  
536 or missed are dictated by the QA-space, potentially leading to over-simplification or misinterpretations that  
537 could be exploited if the system is designed to generate persuasive, profile-aligned outputs. Responsible  
538 development necessitates careful design of the QA-space to ensure it elicits beneficial and fair representations,  
539 alongside ongoing scrutiny of the generated textual profiles for unintended consequences.

540  
541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

---

**564 B QA GENERATOR (DESIGN AND ADVERSARIAL TRAINING)**

565  
 566 The descriptive history representation (DHR) framework relies on a question-answer space (QA-space) to  
 567 guide the learning of informative representations. A crucial component is the QA generator ( $\nu_{QA}^*$ ), which  
 568 provides question-answer pairs  $(q, y)$  based on the history  $h$  and future outcomes  $\omega$ . This section elaborates  
 569 on how this was handled in our experiments, including an adversarial approach to training the QA generator,  
 570 and discusses the design of sufficient question sets.  
 571

**572 B.1 EXPERIMENT DESIGN**

573  
 574 Our work explores two approaches for the QA generator. For our main experiments, we used a pre-defined  
 575 QA generator to isolate and evaluate the core components of DHR learning in a controlled setting. This  
 576 is implemented by procedurally generating questions and their ground-truth answers based on observed  
 577 user histories and future interactions from the datasets. Additionally, as detailed in Section B.4, we also  
 578 explored learning the generator adversarially. We found that the practical gains were minor, reinforcing the  
 579 effectiveness and efficiency of using a powerful, fixed LLM to bootstrap the QA-space.

580 Our questions are aimed at capturing salient aspects of user preferences and future behavior, rather than  
 581 predicting raw, low-level observations. In our experiments, ranking questions include comparing pairs of  
 582 items based on predicted user preference, such as “Does the user prefer Item A or Item B?” The ground-truth  
 583 answers for these is derived from the user’s future ratings of these items.

584 For the Amazon dataset, we also include review generation questions, asking to generate a textual review  
 585 for an item as the user might write it; e.g., “Write a review this user might provide the following product:  
 586 <description>.” The ground-truth answer here is the actual review written by the user for that item in their  
 587 future interactions.

588 Our questions rely on natural language templates, whose slots are filled with specific items or contexts  
 589 from the user’s future. This approach ensures that questions are not overly specific to individual data points  
 590 but rather represent generalizable queries about user behavior and preferences, leveraging the richness and  
 591 flexibility of natural language. This methodology moves beyond attempting to predict the exact sequence  
 592 of raw future observations (e.g., raw ratings or clicks). Instead, it focuses the representation learning on  
 593 summarizing history in a way that supports answering these more abstract, semantically meaningful questions,  
 594 which are directly relevant to the decision-making task, like making good recommendations. Careful design  
 595 of these questions is crucial for ensuring the DHR learns to encapsulate information relevant to understanding  
 596 user preferences and predicting their future actions in a structured manner.  
 597

**598 B.2 DESIGNING SUFFICIENT QUESTION SETS**

599  
 600 The choice of questions fundamentally defines the information that the DHR aims to capture. A set of  
 601 questions  $Q^*$  is deemed sufficient if the answers to these questions provide all necessary information for a  
 602 specific purpose, such as maximizing reward in an RL task.

603 Designing a set of sufficient questions is analogous to other critical design choices in RL, such as defining an  
 604 action space (or a set of options), an abstraction of the state space, or crafting a reward function. In many  
 605 RL applications, these components are manually engineered based on domain knowledge and the specific  
 606 goals of the agent, and are often refined iteratively. For instance, in robotics, actions might be joint torques,  
 607 and rewards might be based on task completion. Similarly, a set of questions for DHRs can be designed by  
 608 domain experts to probe the most relevant aspects of the environment’s state or history.

609 A heuristic approach for creating a sufficient question set is to iteratively refine or expand the set of questions  
 610 based on the agent’s performance or by analyzing what information seems to be missing from the DHR. For

611 example, in a dialogue system, initial questions might focus on user intent or sentiment, with later additions  
 612 querying user knowledge or specific entities. In autonomous driving, questions could range from simple  
 613 presence detection like "Is there a pedestrian in the crosswalk?" to more complex predictions like "What is  
 614 the predicted trajectory of the vehicle ahead?"

615 A well-designed set of questions can be highly beneficial. Questions, especially in natural language, make  
 616 the DHR's content more interpretable, as we can understand what information the representation is trying  
 617 to encode. It also explicitly directs the representation learning process towards capturing information  
 618 deemed critical by the designer. Furthermore, compared to learning representations based on predicting  
 619 all raw future observations—as seen in some PSR or belief state approaches—focusing on a curated set of  
 620 high-level questions can be more tractable and generalizable, particularly in complex environments with  
 621 high-dimensional observation spaces.

622 While learning the questions themselves is a desirable long-term goal, designing the question set using the  
 623 strong prior of large language models (LLMs) is a practical and powerful approach, which injects domain  
 624 knowledge and task relevance into the representation learning process. This practicality and power stem  
 625 from the vast world knowledge and semantic understanding embedded within LLMs, acquired during their  
 626 pre-training on diverse and extensive corpora. Consequently, an LLM can be prompted to generate candidate  
 627 questions with domain-specific nuances—for example, recognizing relevant attributes for movies, such as  
 628 genre, director, or thematic elements; or for products, such as brand, material, or user reviews.

629 Beyond domain specifics, LLMs possess a strong grasp of task-relevant concepts. For a recommendation task,  
 630 this includes notions of preference, comparison, a user's potential future behavior, or even the underlying  
 631 reasons for a choice, enabling the generation of questions that directly probe these critical aspects. This  
 632 allows the LLM-guided design process to formulate questions that elicit more abstract and semantically rich  
 633 answers than merely predicting low-level future observations. For instance, rather than focusing solely on  
 634 predicting the next click, questions can probe comparative preferences (e.g., "Would the user prefer item A  
 635 over item B given their history?") or solicit generative summaries of latent preferences (e.g., "Describe the  
 636 user's taste profile based on past interactions.").

637 The process might involve leveraging an LLM to generate a broad suite of potential questions tailored to the  
 638 specific domain and task, from which a human designer can then curate, refine, or further specialize the set,  
 639 significantly augmenting the human designer's capacity to craft a comprehensive and effective question set.

640 By injecting such structured and high-level priors into the question design phase, the DHR learning process  
 641 is more directly guided towards capturing the most salient information for effective decision-making and  
 642 interpretability, aligning the representation with the core objectives of the task.

### 644 645 646 647 648 B.3 COLD-START SCENARIOS

649  
 650  
 651 In cold-start scenarios, where a new user has a sparse interaction history, the DHR framework remains robust.  
 652 The DHR encoder  $E$  is trained on population data, allowing it to learn generalizable patterns and priors  
 653 regarding user preferences. When presented with a short history  $h_{\text{sparse}}$ , the encoder leverages these learned  
 654 priors to generate a profile  $z$ . Just as a standard recommender model infers "likely future items" based on  
 655 limited data, the DHR infers "likely answers to profile questions." The QA-space provides a structured format  
 656 for encoding these priors, enabling the system to generate meaningful initial profiles even with minimal  
 657 user-specific data.

658 B.4 ADVERSARIAL TRAINING OF THE QA GENERATOR  
659660 While our main results use a fixed QA-space, we also conducted experiments where the QA generator itself  
661 was learned. To achieve this, we formulated the task as an adversarial learning problem, where a QA agent, or  
662 generator, learns to pose questions and provide answers that are maximally informative or challenging for the  
663 DHR encoder and answer agent.664 The original optimization problem (Eq. (OPT 1)) is designed to maximize task reward while minimizing the  
665 divergence between the DHR’s predicted answer distribution and a fixed, ground-truth answer distribution.  
666 To learn the QA generator, we introduced an adversarial objective.667 We refine the objective in Eq. (OPT 1) as follows:  
668

669 
$$\max_{E, \nu_A, \pi_D} \min_{\nu_{QA}^*} (1 - \lambda)V(\pi) - \lambda D_f \left( d^{\nu_A^*} \middle\| d^{\nu_A} \right), \quad (\text{OPT 3})$$
  
670

671 where the adversarial QA generator  $\nu_{QA}^*$  (which generates  $(q, y)$  pairs from  $h, \omega$ ) tries to select questions  
672 and answers to maximize the divergence  $D_f$ , making it harder for the DHR  $E$  and answer agent  $\nu_A$  to match  
673 its outputs. The DHR encoder  $E$  and answer agent  $\nu_A$  continue to try to minimize this divergence while  
674 maximizing task reward.675 The dual variational form, analogous to Eq. (OPT 2), becomes  
676

677 
$$\max_{\pi_D, E, \nu_A} \min_{\substack{g: \mathcal{H} \times \mathcal{Q} \times \mathcal{Y} \mapsto \mathbb{R} \\ \nu_{QA}^*: \mathcal{H} \times \Omega \mapsto \Delta_{\mathcal{Q} \times \mathcal{Y}}}} \mathbb{E} \left[ (1 - \lambda)r(h, a) + \lambda \mathbb{E}_{\substack{q \sim \nu_Q^*(h, \omega) \\ y \sim \nu_A(q, h, \omega)}} [f^*(g(h, q, y))] - \lambda \mathbb{E}_{\substack{q \sim \nu_Q^*(h, \omega) \\ y \sim \nu_A^*(z, q)}} [g(h, q, y)] \right]. \quad (\text{OPT 4})$$
  
678  
679

680 In this formulation,  $\nu_{QA}^*$  is the adversarial QA generator that produces  $(q, y)$  pairs given history  $h$  and future  
681  $\omega$ . The agent’s answer network  $\nu_A$  predicts  $\hat{y}$  for a question  $q$  (generated by  $\nu_{QA}^*$ ) using the DHR  $z = E(h)$ .  
682 The discriminator  $g$  tries to distinguish between answers from  $\nu_A$  and  $\nu_{QA}^*$ . The adversarial QA generator  
683  $\nu_{QA}^*$  is trained to generate  $(q, y)$  pairs that are hard for  $\nu_A$  to predict correctly, effectively maximizing the  
684 objective for  $g$ . Eq. (OPT 4) maintains strong duality, allowing for an iterative training algorithm where we  
685 updated  $\pi_D, E, \nu_A, g$ , and  $\nu_{QA}^*$ .  
686687 Our experiments with this adversarial setup (results in Table 3) showed a marginal reward improvement of  
688 2-3%. This demonstrates the potential benefits of learning the QA-space, such as the automated discovery  
689 of more informative questions. However, we also found that ensuring the learned questions remained  
690 semantically meaningful and interpretable was a significant challenge. Without strong regularization (e.g.,  
691 encouraging question diversity or biasing towards human-understandable questions), the generator could  
692 find trivial or uninformative ways to maximize the divergence. Given the substantial effectiveness of the  
693 fixed QA-generator bootstrapped with a powerful LLM and the added complexity of adversarial training, we  
694 conclude that the fixed approach remains highly practical.  
695696 Learning an adversarial QA generator is a complex but promising research direction which may significantly  
697 enhance the adaptability of the DHR framework. Our current work lays the foundation by demonstrating  
698 the effectiveness of DHRs given a well-defined QA-space, and future work can build upon this to explore  
699 dynamic and learned QA-spaces more thoroughly.

705 Table 3: Comparison of DHRL with a fixed QA-generator versus an adversarially trained QA-generator for  
 706 the Gemma V3 12B model.  
 707

708 <b>Dataset</b>	709 <b>Model &amp; Method</b>	710 <b>Prediction Accuracy</b>	711 <b>Rec. Reward (<math>r^D</math>)</b>
710 Amazon	Gemma V3 12B + DHRL (fixed QA)	0.75	0.84
	Gemma V3 12B + DHRL (adversarial QA)	<b>0.77</b>	<b>0.86</b>
712 MovieLens	Gemma V3 12B + DHRL (fixed QA)	0.84	0.93
	Gemma V3 12B + DHRL (adversarial QA)	<b>0.86</b>	<b>0.95</b>

## 715 C IMPLEMENTATION DETAILS

716 This appendix provides additional details regarding the implementation of our DHR learning (DHRL)  
 717 framework.

718 **Models and Parameterization** The DHR encoder ( $E$ ), answer agent ( $\nu_A$ ), and decision agent ( $\pi_D$ ) are  
 719 implemented using LLMs. As mentioned in the main paper (Section 5.1), primary LLM models include  
 720 Gemini 1.5 Flash and Gemma V3 (4B, 12B). The DHR encoder and answer agent share the same base LLM  
 721 architecture. A fixed anchor LLM (typically the original or supervised fine-tuned version of the model),  
 722 is used to provide the reference distribution for the KL-divergence regularization term in the policy loss.  
 723 The value network, responsible for estimating advantages and providing outputs for the DHR discriminator  
 724 ( $g$ ), defaults to a Gemma V3 4B model unless specified otherwise. The discriminator  $g$  outputs are derived  
 725 from the value network. This is achieved by feeding the value network QA pair samples (one reference, one  
 726 target, as required by DHR). The value network’s output vocabulary is conceptually partitioned: one segment  
 727 is used for standard value prediction, while another segment provides the scalar outputs leveraged by the  
 728 discriminator logic.

729 **Training Framework and Algorithm** The core learning algorithm is implemented as an actor-critic method.  
 730 The DHR encoder acts as the policy, and the answer agent components (value function and discriminator  
 731 logic) form parts of the critic and learning signal. Training is run for up to 3000 optimization steps. As noted  
 732 in the main paper, convergence was generally observed within 1000 iterations, which forms the basis for our  
 733 reported results.

734 **The multi-agent training paradigm introduces computational overhead compared to standard supervised  
 735 fine-tuning. In our experiments, the DHRL training procedure resulted in an approximately 3x increase in  
 736 training time compared to the predictive baseline (Gemma V3 Baseline).**

737 **Key Hyperparameters** Several hyperparameters governed the training process. These are summarized in  
 738 Table 4.

739 **Loss Function Details (DHR)** The choice of  $f$ -divergence for the discriminator was crucial. The default  
 740 was TV-distance, as reported in the main paper. KL divergence and Chi-squared divergence were also  
 741 explored. The DHR encoder’s policy was regularized using KL divergence w.r.t. the fixed anchor (SFT)  
 742 model, with the KL weight ( $\alpha$ ) annealing. The total loss driving updates included: a policy gradient loss  
 743 term derived from advantages; a value function loss term; the KL regularization term for the policy; and  
 744 the discriminator loss term. The advantages were computed based on a sum of the environment reward  $r^D$   
 745 (related to downstream task performance) and the DHR answer reward  $r^A$ .

752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
Table 4: Key Hyperparameters for DHRL Experiments.

Hyperparameter	Default Value
DHR Balancing Factor ( $\lambda$ )	0.01
Batch Size	128
LLM Input Token Length	6144
User Profile ( $z$ ) Max Token Length	256 (default)
User History Length ( $H$ )	10 (default)
Number of Ranking Questions ( $K_q$ )	5 (default)
Policy Update Delay	20 steps
Policy Warmup Steps	20 steps

## D ADDITIONAL RESULTS

Tables 6 and 7 provide the complete table of results from Table 1 with 95% confidence intervals. We also provide missing results for ablations for Amazon in Fig. 5, showing similar results to Fig. 3. We also show in Fig. 5 (left) the learning curves for the ablation over profile lengths. We found that the Amazon dataset in particular is very sensitive to longer profile lengths. Future work can further explore the effects of the DHR length and its relation to the choice of DHR templates with their predictive capabilities.

### D.1 ADDITIONAL ABLATION STUDIES

This section presents additional ablation studies investigating the sensitivity to the balancing hyperparameter  $\lambda$ , the necessity of joint optimization, and the sufficiency of the learned DHR.

**Sensitivity to  $\lambda$ .** The hyperparameter  $\lambda$  in (OPT 1) balances the RL objective ( $V(\pi)$ ) and the DHR learning objective (the  $f$ -divergence). We analyzed the sensitivity of the recommendation reward to different values of  $\lambda \in \{0.01, 0.1, 0.5\}$  using the Gemma V3 12B model. As shown in Figure 4, the performance is robust across the tested range, with optimal performance achieved when  $\lambda$  is between 0.01 (our default) and 0.1.

**Necessity of Joint Optimization (FT-QA).** To understand whether the performance gains are attributable to the structured information provided by the QA-generator or the joint optimization framework (OPT 1), we conducted an ablation study (FT-QA). In this ablation, we trained the model using only the QA objective ( $\lambda = 1$ ), effectively performing supervised fine-tuning on the generated QA pairs. We then evaluated the resulting representation on the downstream recommendation task.

The results in Table 5 show that FT-QA significantly underperforms the full DHR method on both datasets. On MovieLens, FT-QA improves over the standard predictive baseline, suggesting the QA structure itself is beneficial. However, on Amazon, FT-QA performs worse than the baseline. Crucially, the joint optimization of DHR consistently outperforms both approaches by a significant margin. This confirms that balancing QA fidelity with reward maximization is essential for learning representations effective for decision-making.

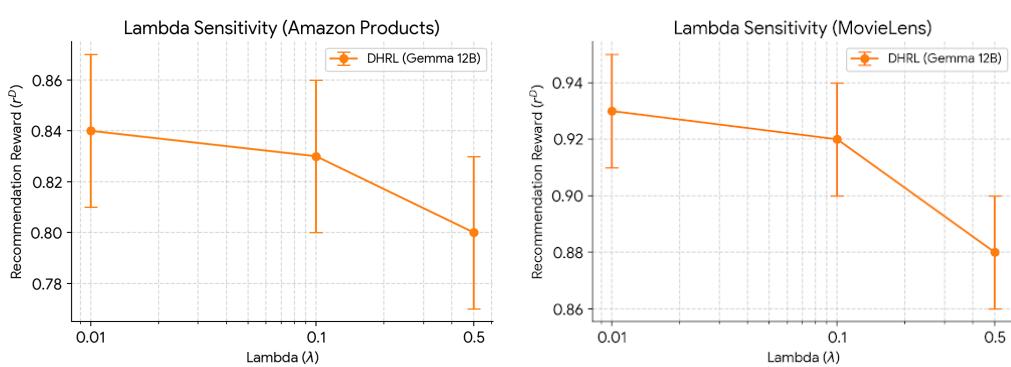


Figure 4: Sensitivity analysis of the recommendation reward to the balancing hyperparameter  $\lambda$  on Amazon (left) and MovieLens (right) datasets.

Table 5: Ablation study on the necessity of joint optimization (Gemma V3 12B). FT-QA ( $\lambda = 1$ ) uses only the QA objective. DHR uses the joint RL+QA objective.

Dataset	Method	Prediction Accuracy	Rec. Reward ( $r^D$ )
Amazon	Baseline (Predictive)	0.67	0.78
	FT-QA ( $\lambda = 1$ )	0.65	0.71
	<b>DHR (Joint, <math>\lambda = 0.01</math>)</b>	<b>0.75</b>	<b>0.84</b>
MovieLens	Baseline (Predictive)	0.58	0.64
	FT-QA ( $\lambda = 1$ )	0.72	0.78
	<b>DHR (Joint, <math>\lambda = 0.01</math>)</b>	<b>0.84</b>	<b>0.93</b>

**Testing DHR Sufficiency.** To empirically test whether the learned DHR  $z_t$  is a sufficient statistic of the history  $h_t$ , we conducted an ablation where the Decision Agent  $\pi_D$  is given access to both the DHR and the raw history:  $\pi_D(z_t, h_t)$ . We compared this against the standard DHR approach  $\pi_D(z_t)$ . Using the Gemma V3 12B model, we observed a negligible improvement in recommendation reward (+0.5% on both datasets). This strongly suggests that the DHR  $z_t$  captures nearly all the necessary information from  $h_t$  for the decision-making task, closely approximating a sufficient statistic.

Table 6: Performance (with 95% CI) of DHRL with default settings (profile with max. 256 tokens, 5 questions, 10 history interactions, TV-divergence) after 1,000 iterations for Gemma V3 4B, 12B, and Gemini 1.5 Flash models.

Model & Method	Prediction Accuracy (w.r.t. GT)	Rec. Reward ( $r^D$ )
Task: recommendation using Amazon products user profiles		
Gemma V3 4B	$0.34 \pm 0.04$	$0.54 \pm 0.04$
Gemma V3 4B+DHRL	$0.71 \pm 0.04$	$0.83 \pm 0.03$
Gemma V3 12B	$0.67 \pm 0.04$	$0.78 \pm 0.03$
Gemma V3 12B+DHRL	$0.75 \pm 0.04$	$0.84 \pm 0.03$
Gemini 1.5 Flash	$0.69 \pm 0.04$	$0.80 \pm 0.03$
Gemini 1.5 Flash+DHRL	$0.74 \pm 0.04$	$0.86 \pm 0.03$
Task: recommendation using MovieLens user profiles		
Gemma V3 4B	$0.37 \pm 0.04$	$0.61 \pm 0.04$
Gemma V3 4B+DHRL	$0.74 \pm 0.04$	$0.79 \pm 0.03$
Gemma V3 12B	$0.58 \pm 0.04$	$0.64 \pm 0.04$
Gemma V3 12B+DHRL	$0.84 \pm 0.03$	$0.93 \pm 0.02$
Gemini 1.5 Flash	$0.62 \pm 0.04$	$0.68 \pm 0.04$
Gemini 1.5 Flash+DHRL	$0.82 \pm 0.03$	$0.92 \pm 0.02$

Table 7: Performance (with 95% CI) of DHRL with default settings (profile with max. 256 tokens, 5 questions, 10 history interactions, TV-divergence) after 1,000 iterations for Gemma V3 4B, 12B, and Gemini 1.5 Flash models.

Model & Method	Profile-History Consistency (AI/Human)	Prediction Fidelity (AI/Human)	Review Quality (AI/Human)
Task: recommendation using Amazon products user profiles			
Gemma V3 4B	$3.46 \pm 0.16 / 3.28 \pm 0.16$	$0.44 \pm 0.04 / 0.47 \pm 0.04$	$0.15 \pm 0.03 / 0.27 \pm 0.04$
Gemma V3 4B+DHRL	$3.41 \pm 0.16 / 3.85 \pm 0.15$	$0.56 \pm 0.04 / 0.53 \pm 0.04$	$0.85 \pm 0.03 / 0.73 \pm 0.04$
Gemma V3 12B	$4.32 \pm 0.12 / 4.18 \pm 0.13$	$0.34 \pm 0.04 / 0.39 \pm 0.04$	$0.17 \pm 0.03 / 0.35 \pm 0.04$
Gemma V3 12B+DHRL	$4.39 \pm 0.12 / 4.37 \pm 0.12$	$0.66 \pm 0.04 / 0.61 \pm 0.04$	$0.83 \pm 0.03 / 0.65 \pm 0.04$
Gemini 1.5 Flash	$4.41 \pm 0.12 / 4.30 \pm 0.13$	$0.38 \pm 0.04 / 0.41 \pm 0.04$	$0.17 \pm 0.03 / 0.38 \pm 0.04$
Gemini 1.5 Flash+DHRL	$4.46 \pm 0.11 / 4.34 \pm 0.12$	$0.62 \pm 0.04 / 0.59 \pm 0.04$	$0.83 \pm 0.03 / 0.62 \pm 0.04$
Task: recommendation using MovieLens user profiles			
Gemma V3 4B	$3.72 \pm 0.15 / 3.96 \pm 0.14$	$0.32 \pm 0.04 / 0.45 \pm 0.04$	-
Gemma V3 4B+DHRL	$3.84 \pm 0.15 / 4.24 \pm 0.13$	$0.68 \pm 0.04 / 0.55 \pm 0.04$	-
Gemma V3 12B	$4.66 \pm 0.09 / 4.58 \pm 0.10$	$0.35 \pm 0.04 / 0.42 \pm 0.04$	-
Gemma V3 12B+DHRL	$4.62 \pm 0.10 / 4.71 \pm 0.09$	$0.76 \pm 0.04 / 0.66 \pm 0.04$	-
Gemini 1.5 Flash	$4.69 \pm 0.09 / 4.53 \pm 0.11$	$0.38 \pm 0.04 / 0.39 \pm 0.04$	-
Gemini 1.5 Flash+DHRL	$4.74 \pm 0.08 / 4.70 \pm 0.09$	$0.77 \pm 0.03 / 0.62 \pm 0.04$	-

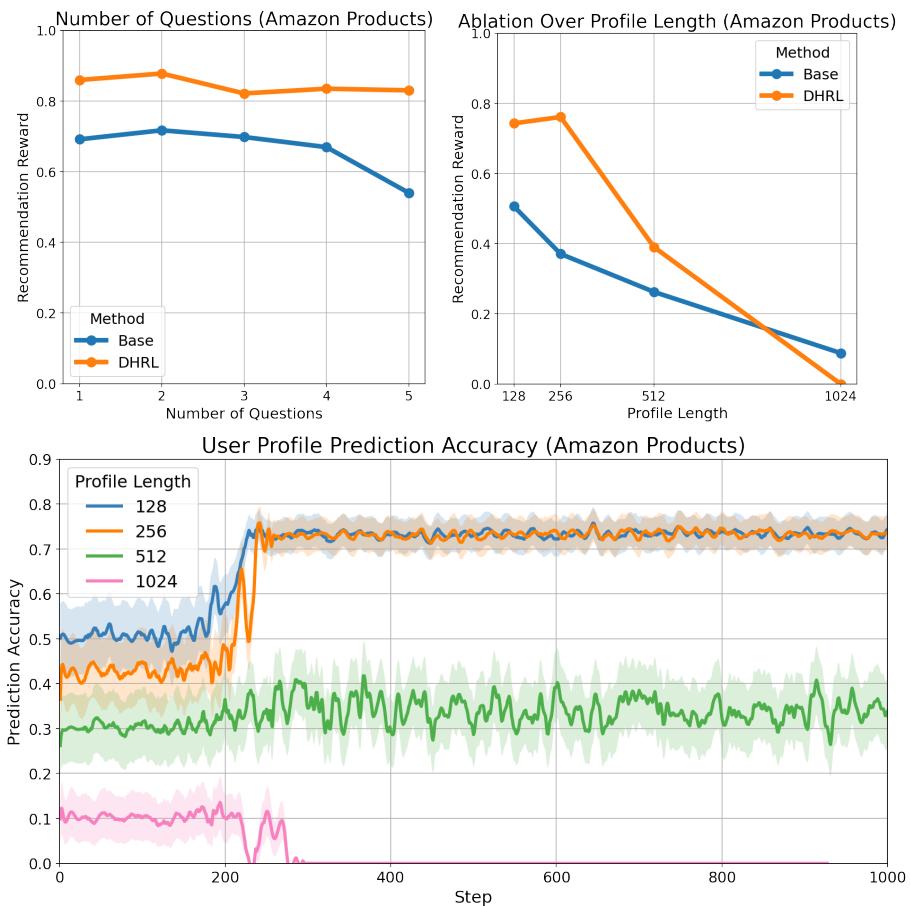


Figure 5: Ablation study on the history length (left) and profile length (right) for the Amazon dataset, showing recommendation reward. Bottom plot shows learning curves for different profile lengths. Lower profile lengths were critical to ensure convergence.

## E PROOF OF THM. 1

**Theorem 1.** *Let  $E : \mathcal{H} \mapsto \mathcal{Z}$  be a DHR (Definition 3). Then it is also an  $f$ -sufficient statistic.*

*Proof.* Let  $E_{DHR} : \mathcal{H} \rightarrow \mathcal{Z}$  be a Descriptive History Representation. We aim to show that  $E_{DHR}$  is an  $f$ -sufficient statistic. According to the definition of  $f$ -sufficiency (Section 2), this means there exists a function  $g^\pi : \mathcal{Z} \mapsto \mathbb{R}$  such that for any history  $h \in \mathcal{H}$  and policy  $\pi \in \Pi$ :

$$f^\pi(h) = g^\pi(E_{DHR}(h)). \quad (1)$$

We are given an  $f$ -sufficient QA-space (Definition 2). This means for any history  $h$ , there is a set of sufficient questions  $Q_h^*$ . The set of sufficient question-answer pairs is defined as:

$$S_{QA}(h) := \{(q, \nu(h, q)) \mid q \in Q_h^*\}.$$

940  
941 Since  $S_{QA}(h)$  is defined as an  $f$ -sufficient statistic, there must exist a function  $F^\pi$  such that:  
942  
943

$$f^\pi(h) = F^\pi(S_{QA}(h)). \quad (2)$$

944  
945  
946 We are also given that  $E_{DHR}$  is a DHR (Definition 3). This implies the existence of an answer function  $\nu_A$   
947 such that for any  $h$  and any  $q \in Q_h^*$ :

$$\nu_A(E_{DHR}(h), q) = \nu(h, q). \quad (3)$$

948  
949  
950 We now construct the required function  $g^\pi(z)$ . We must define  $g^\pi(z)$  such that it depends only on  $z$  (and  $\pi$ ),  
951 not on a specific history  $h$ .

952 Let  $z \in \mathcal{Z}$ . We define the set of constructible QA pairs from  $z$ , denoted  $S_{constr}(z)$ . This set includes all QA  
953 pairs where the question  $q$  is sufficient for *at least one* history  $h'$  that maps to  $z$ , and the answer is generated  
954 by  $\nu_A$  using  $z$ :

$$S_{constr}(z) := \{(q, \nu_A(z, q)) \mid \exists h' \in \mathcal{H} \text{ s.t. } z = E_{DHR}(h') \wedge q \in Q_{h'}^*\}.$$

955  
956  
957 We define  $g^\pi(z)$  using the function  $F^\pi$  from the sufficiency definition (Eq. 2):  
958

$$g^\pi(z) := F^\pi(S_{constr}(z)). \quad (4)$$

959  
960 To prove the theorem (Eq. 1), we must verify that  $g^\pi(E_{DHR}(h)) = f^\pi(h)$ . Let  $z_h = E_{DHR}(h)$ . Using Eq.  
961 2 and Eq. 4, the proof reduces to showing:

$$F^\pi(S_{constr}(z_h)) = F^\pi(S_{QA}(h)).$$

962  
963 We analyze the relationship between the two sets  $S_{QA}(h)$  and  $S_{constr}(z_h)$ .

964  
965 **Step A: Show  $S_{QA}(h) \subseteq S_{constr}(z_h)$ .**

966 Let  $(q_0, y_0) \in S_{QA}(h)$ . By definition,  $q_0 \in Q_h^*$  and  $y_0 = \nu(h, q_0)$ . By the DHR property (Eq. 3),  
967  $y_0 = \nu_A(z_h, q_0)$ .

968 To show  $(q_0, y_0) \in S_{constr}(z_h)$ , we must satisfy the condition in the definition of  $S_{constr}(z_h)$ :  $\exists h'$  s.t.  
969  $z_h = E_{DHR}(h')$  and  $q_0 \in Q_{h'}^*$ . We can simply choose  $h' = h$ . Since  $z_h = E_{DHR}(h)$  and  $q_0 \in Q_h^*$ , the  
970 condition is met. Thus,  $S_{QA}(h) \subseteq S_{constr}(z_h)$ .

971  
972 **Step B: Establish the equality of  $F^\pi$ .**

973 It is possible that  $S_{constr}(z_h)$  contains more QA pairs than  $S_{QA}(h)$ . (This occurs if another history  $h'' \neq h$   
974 maps to  $z_h$ , and  $Q_{h''}^* \not\subseteq Q_h^*$ ).

975 However,  $S_{QA}(h)$  is, by definition, a sufficient statistic for  $f^\pi(h)$ . A fundamental property of sufficient  
976 statistics is that augmenting a sufficient set of information with redundant but consistent information does not  
977 change the resulting value.

978 We must ensure  $S_{constr}(z_h)$  is consistent. By its definition, all answers in  $S_{constr}(z_h)$  are generated by the  
979 single function  $\nu_A(z_h, \cdot)$ . Therefore, for any question  $q$ , the answer is uniquely determined by  $z_h$ , ensuring  
980 consistency across the set.

987 Since  $S_{QA}(h) \subseteq S_{constr}(z_h)$ ,  $S_{constr}(z_h)$  is consistent, and  $S_{QA}(h)$  is sufficient, it follows that:

$$988 \quad 989 \quad 990 \quad 991 \quad 992 \quad 993 \quad 994 \quad 995 \quad 996 \quad 997 \quad 998 \quad 999 \quad 1000 \quad 1001 \quad 1002 \quad 1003 \quad 1004 \quad 1005 \quad 1006 \quad 1007 \quad 1008 \quad 1009 \quad 1010 \quad 1011 \quad 1012 \quad 1013 \quad 1014 \quad 1015 \quad 1016 \quad 1017 \quad 1018 \quad 1019 \quad 1020 \quad 1021 \quad 1022 \quad 1023 \quad 1024 \quad 1025 \quad 1026 \quad 1027 \quad 1028 \quad 1029 \quad 1030 \quad 1031 \quad 1032 \quad 1033$$

$$F^\pi(S_{QA}(h)) = F^\pi(S_{constr}(z_h)).$$

We have shown that  $f^\pi(h) = F^\pi(S_{QA}(h)) = F^\pi(S_{constr}(z_h)) = g^\pi(z_h)$ . Therefore,  $f^\pi(h) = g^\pi(E_{DHR}(h))$ . This completes the proof that  $E_{DHR}$  is an  $f$ -sufficient statistic.  $\square$

1034 F LLM PROMPTS  
1035

1036  
1037 Below we provide the complete prompts used in DHRL. The prompts are provided for the Amazon domain.  
1038 The prompt used for MovieLens is identical except for removal of the review question.

1039 Text in **blue** refers to parts of the prompt that are not changed. Text in **red** refers to placeholders that are  
1040 replaced with e.g., titles, descriptions, user profiles, etc.

1041  
1042 F.1 DHR ENCODER  
1043

1044 Your task is to write a user profile in the shopping domain for Clothing, Shoes, and Jewelry. The user profile  
1045 should describe the user's preferences well.

1046 Your user profile will later be used to predict the user's preferences and reviews over new items.

1047 You will be given a list of products the user has rated (ratings 1 to 5, where 5 is the highest rating).

1049 You will also be given reviews for some of these products.

1050 You will then be asked to write a user profile describing the preferences of the user.

1051  
1052 Here is an example of a user profile given a history:

1053 #BEGIN EXAMPLE  
1054

1055 User History:

1056 Item #1

1057 Item title: Robert Graham Men's Shipwreck Long Sleeve Button Down Shirt

1058 Item description: N/A

1059 Item price: N/A

1060 User rating: 5.0

1061 User review: Fantastic shirt. Great price. I saw the same shirt at Nordstrom for 229.00

1062

1063

1064 Item #2

1065 Item title: Robert Graham Men's Whitehorse Long Sleeve Button Down Shirt

1066 Item description: "The geometric pattern on this dress shirt creates an optical illusion to enhance your  
1067 wardrobe. Crafted from Egyptian cotton for comfort.", "The geometric pattern on this dress shirt creates an  
1068 optical illusion to enhance your wardrobe. Crafted from Egyptian cotton for comfort."

1069 Item price: N/A

1070 User rating: 5.0

1071 User review: Thank you for such a great deal on a great shirt

1072

1073

1074 Item #3

1075 Item title: Robert Graham Men's Shawn-Long Sleeve Woven Shirt

1076 Item description: "When is a polo not just a polo? When it's made from exceptional mercerized pique cotton  
1077 and boasting our signature space dyed tipping at the collar to take it from ordinary to extraordinary."

1078 Item price: N/A

1079 User rating: 5.0

1080 User review: I have never had so many complements on a shirt.

1081 Thank you for the great deal  
1082

---

1083  
1084  
1085 Item #4  
1086 Item title: Robert Graham Men's Redzone Pixel Patterned Button-Front Shirt with Convertible Cuffs  
1087 Item description: N/A  
1088 Item price: N/A  
1089 User rating: 5.0  
1090 User review: love it

---

1091  
1092  
1093 Item #5  
1094 Item title: Robert Graham Men's Shipwreck-Long Sleeve Button Down Shirt  
1095 Item description: "Every stylish guy needs a sophisticated striped shirt in his arsenal. Crafted from premium  
1096 Egyptian cotton with space dyed stripes and a pop of paisley embroidery at the cuffs and inside placket for a  
1097 little extra edge, this shirt has you covered at the office or a weekend away"  
1098 Item price: N/A  
1099 User rating: 5.0  
1100 User review: love it

---

1101  
1102  
1103 Item #6  
1104 Item title: Robert Graham Men's Maitai-Long Sleeve Button Down Shirt  
1105 Item description: "Rain check. No way. Pack a mai tai and check's in for sunny days. Ombre squares of crisp  
1106 bright white, ocean blues and poolside purples along with a dash of sunny blue sky over embroidery are the  
1107 perfect cocktail for any getaway"  
1108 Item price: N/A  
1109 User rating: 5.0  
1110 User review: love it

---

1111  
1112  
1113 Item #7  
1114 Item title: Lucky 21 Women's Plus Size Cuffed Sleeve One Button Boyfriend Blazer  
1115 Item description: "This boyfriend blazer is perfect for office and work wear, but also for a night out with  
1116 friends! It features one button and cuffed sleeves."  
1117 Item Price: \$8.00 - \$34.99  
1118 User rating: 1.0  
1119 User review: sent back

---

1120  
1121  
1122  
1123 Item #8  
1124 Item title: Robert Graham Men's Seattle Long Sleeve Button Down Shirt  
1125 Item description: "Hombre fade window pane print with satin stitch window pane edging and floral  
1126 embroidery woven. Has digital printed neck band, back yoke and under collar."  
1127 Item price: N/A

1128 User rating: 5.0  
1129 User review: Great shirt. Robert graham is the king  
1130  
1131  
1132  
1133 Item #9  
1134 Item title: EleQueen 925 Sterling Silver Full Cubic Zirconia Bridal Flower Stud Earrings 15mm  
1135 Item description: N/A  
1136 Item Price: \$16.99  
1137 User rating: 5.0  
1138 User review: The wife likes  
1139  
1140  
1141 Item #10  
1142 Item title: Leveret Women's Robe Soft Micro Fleece Plush Shawl Collar Bathrobe Robe (Size Small-  
1143 XXLarge)  
1144 Item description: N/A  
1145 Item Price: \$39.99 - \$49.99  
1146 User rating: 5.0  
1147 User review: Wife likes  
1148  
1149  
1150 User Profile:  
1151  
1152 #PROFILE#  
1153  
1154 1. Brand Loyalty:  
1155 • Extremely High Loyalty to Robert Graham: The vast majority (7 out of 10) of rated items are  
1156 from this brand. All Robert Graham items are rated 5/5. This is the single strongest signal in  
1157 the data. The user explicitly calls Robert Graham "king".  
1158 • Open to Other Brands, but with Caution: Purchased items from Lucky 21 (blazer, rated poorly),  
1159 EleQueen (earrings, positive), and Leveret (robe, positive). This suggests a willingness to  
1160 explore, but success outside of Robert Graham is mixed.  
1161 2. Product Category Preferences:  
1162 • Men's Shirts (Dominant): Specifically, long-sleeve button-down shirts are the primary interest.  
1163 Focus is on dress shirts and slightly more casual shirts (like a refined polo).  
1164 • Women's Gifts/Items (Secondary): Purchases include a women's blazer, earrings, and a robe.  
1165 Reviews indicate these are likely gifts ("Wife likes"). The low rating on the blazer suggests  
1166 a different preference set for women's clothing, or simply a bad experience with that specific  
1167 item.  
1168 • No interest shown in shoes, and lower interest in jewelry.  
1169 3. Style Preferences (Men's Shirts):  
1170 • Bold Patterns and Colors: Descriptions of favored items include "geometric pattern," "ombre  
1171 squares," "pixel patterned," "striped," "floral embroidery." This user is not afraid of visually  
1172 interesting shirts.  
1173 • High-Quality Materials: Mentions of "Egyptian cotton," "mercerized pique cotton," "satin  
1174 stitch," and "925 Sterling Silver" (for the earrings) indicate a preference for premium materials.

1175           • Details and Accents: Appreciates features like "paisley embroidery," "space-dyed tipping,"  
1176            "convertible cuffs," and "digital printed neckband." These small details are important.  
1177           • Business Casual, elevated.

1178           4. Style Preferences (Gifts/Women's Items):

1180           • Classic/Elegant: The earrings are described as "Bridal Flower Stud" with Cubic Zirconia,  
1181            suggesting a classic, somewhat formal style. The robe is a "Soft Micro Fleece Plush Shawl  
1182            Collar," indicating a preference for comfort and a traditional design.  
1183           • Less information than men's.

1184           5. Price Sensitivity:

1185           • Value-Oriented, but Willing to Pay for Quality: The user repeatedly mentions "great deal"  
1186            and compares the price favorably to Nordstrom's. This suggests a search for good value, but  
1187            not necessarily the absolute lowest price. They are willing to pay for a Robert Graham shirt,  
1188            implying a higher price bracket acceptance for favored brands.  
1189           • Wide Price Range: Has purchased items ranging from under \$10 (potentially the blazer,  
1190            depending on the specific price within the range) to likely over \$100 (based on the Nordstrom  
1191            comparison for Robert Graham shirts).

1192           6. Review Style:

1193           • Concise and Positive (when satisfied): Uses short phrases like "love it," "Fantastic shirt," "Great  
1194            shirt."  
1195           • Value-Focused: Often mentions price and deals in positive reviews.  
1196           • Direct (when dissatisfied): Simply states "sent back" for the negative review.  
1197           • Expressive: show appreciation and gratefulness.

1198           7. Purchase Channel: The purchase channel is not described.

1199           8. Location/Region: No address, etc. is listed.

1200           9. Purchase Frequency: Cannot be determined from this data alone. More history would be needed.

1201           10. . Inferred Gender and Demographic Most likely Male, purchasing gifts for female.

1202           #END#

1203           #END EXAMPLE

---

1204           The above is just an example, you do not have to follow this template exactly.  
1205           You can try different ways to describe the user's preferences, and you can also be more comprehensive in  
1206           your description, if needed.

---

1207           Now it's your turn.

1208           Below is a list of items the user has rated:

1209           1. Title: <item 1 title>

1210           Description: <item 1 description>

1211           Price: <item 1 price>

1212           Review: <item 1 review>

1213           2. Title: <item 2 title>

1214           Description: <item 2 description>

1215           Price: <item 2 price>

1222 Review: <item 2 review>  
1223  
1224 3. Title: <item 3 title>  
1225 Description: <item 3 description>  
1226 Price: <item 3 price>  
1227 Review: <item 3 review>  
1228  
1229 4. Title: <item 4 title>  
1230 Description: <item 4 description>  
1231 Price: <item 4 price>  
1232 Review: <item 4 review>  
1233  
1234 5. Title: <item 5 title>  
1235 Description: <item 5 description>  
1236 Price: <item 5 price>  
1237 Review: <item 5 review>  
1238  
1239 6. Title: <item 6 title>  
1240 Description: <item 6 description>  
1241 Price: <item 6 price>  
Review: <item 6 review>  
1242  
1243 7. Title: <item 7 title>  
1244 Description: <item 7 description>  
1245 Price: <item 7 price>  
1246 Review: <item 7 review>  
1247  
1248 8. Title: <item 8 title>  
1249 Description: <item 8 description>  
1250 Price: <item 8 price>  
Review: <item 8 review>  
1251  
1252 9. Title: <item 9 title>  
1253 Description: <item 9 description>  
1254 Price: <item 9 price>  
1255 Review: <item 9 review>  
1256  
1257 10. Title: <item 10 title>  
1258 Description: <item 10 description>  
1259 Price: <item 10 price>  
Review: <item 10 review>  
1260  
1261  
1262  
1263 Write a user profile to describe this user.  
1264 Ignore information that seems irrelevant or not informative.  
1265 Limit the profile to a maximum of 166 words.  
1266 Even though the length of the example profile above might be different, there is a strict limit of 166 words to  
1267 your output.  
1268

1269  
1270 Output should be in the following format:  
1271 #PROFILE# <user\_profile> #END#  
1272

1273 F.2 ANSWER AGENT  
1274

1275 Below is a profile of a user's preferences:  
1276

1277 **<user profile (from DHR encoder)>**  
1278

---

1280 Below are 6 items the user hasn't rated yet:  
1281

1282 1. Title: <item 1 title>  
1283 Description: <item 1 description>  
1284 Price: <item 1 price>

1285 2. Title: <item 2 title>  
1286 Description: <item 2 description>  
1287 Price: <item 2 price>

1288 3. Title: <item 3 title>  
1289 Description: <item 3 description>  
1290 Price: <item 3 price>

1291 4. Title: <item 4 title>  
1292 Description: <item 4 description>  
1293 Price: <item 4 price>

1294 5. Title: <item 5 title>  
1295 Description: <item 5 description>  
1296 Price: <item 5 price>

1297 6. Title: <item 6 title>  
1298 Description: <item 6 description>  
1299 Price: <item 6 price>

---

1300  
1301  
1302  
1303  
1304  
1305  
1306 Below are a set of 6 questions about the items above:

1307 (Q1) Rank the items <id 5> and <id 6> based on the user's preferences.  
1308 (Q2) Rank the items <id 6> and <id 2> based on the user's preferences.  
1309 (Q3) Rank the items <id 2> and <id 1> based on the user's preferences.  
1310 (Q4) Rank the items <id 1> and <id 3> based on the user's preferences.  
1311 (Q5) Rank the items <id 3> and <id 4> based on the user's preferences.  
1312 (Q6) Write a review for item <id 1> as the user would write it.

---

1313  
1314 Each of the 6 questions either asks you to rank two items based on the user's preferences, or to  
1315 write a review for an item in the way the user might write it.

1316 For ranking questions, a higher rank means the user would rate the item higher.  
1317 If there's a tie, pick the order randomly.  
1318 For review questions, you should write a raw review for the item as the user would write it.  
1319  
1320 Formatting your answer:  
1321 Your output should be in the following format. For each ranking question you should output a line with its  
1322 prediction.  
1323 For a review question, you should output a review as the user would write it for that item.  
1324  
1325 For question k your output answer should be in the following format:  
1326 (Ak) #PREDICTION# [item\_id, item\_id] #END#  
1327 if it's a ranking question, and  
1328 (Ak) #PREDICTION# user review #END#  
1329 if it's a review question.  
1330 For the ranking question, the list is in order of the user's preferences for those item ids in that  
1331 question.  
1332 For the review question, don't prefix it with anything else (like "User review"). Just write the raw review the  
1333 user would write between the #PREDICTION# and #END#.  
1334  
1335 For example, assume you are asked in question 1 to rank item ids 1 and 2, and you believe the  
1336 user would rate item 1 higher than item 2, and in question 2 you're asked rank item ids 1, 3, where you  
1337 think the user would rate item 3 higher than item 1. And assume in question 3 you are asked to write a review  
1338 for item id 1, and let's say the user liked that item.  
1339 Then your complete output should be:  
1340 (A1) #PREDICTION# [1, 2] #END#  
1341 (A2) #PREDICTION# [3, 1] #END#  
1342 (A3) #PREDICTION# This shirt is great! I loved the color and the material. #END#  
1343  
1344 Output should be exactly in the above format. Do not output anything else.  
Remember to use #PREDICTION# and #END# for every answer.  
1345  
1346  
1347  
1348  
1349  
1350  
1351  
1352  
1353  
1354  
1355  
1356  
1357  
1358  
1359  
1360  
1361  
1362

1363 G RATER STUDY

1364

1365

1366

1367

1368

1369

1370

1371

1372

1373

1374

1375

1376

1377

1378

1379

1380

1381

1382

1383

1384

1385

1386

1387

1388

1389

1390

1391

1392

1393

1394

1395

1396

1397

1398

1399

1400

1401

1402

1403

1404

1405

1406

1407

1408

1409 Below we show examples of rater forms including the contexts and questions:

## 1410 Sample Rater From for Prediction and Review Accuracy using Amazon Reviews

1411 We have the following summary about preference of a user:

1412 This user favors comfortable, stylish women's clothing, particularly sandals and dresses. They appreciate value and are drawn to visually appealing designs, as evidenced by the high rating on the silver chain. A preference for easy-to-wear items is clear, with ratings of 4.0 on the sandals and dress. They seem to like simple, classic styles and are happy to purchase items within a reasonable price range. The 5/5 rating on the flip flops indicates a practical focus on comfort and style.

1413 1. Title: BIADANI Women Button Down Long Sleeve Basic Soft Knit Cardigan Sweater Item Item description: "Casual, Elegant Fitted Knit Cardigan

1414 Sweaters That is Versatile and Made Out of High Quality Material and Luciously Soft."

1415 Price: \$10.01 - \$25.80

1416 2. Title: Moxeay Women Sexy Sleeveless Spaghetti Strap Mini Club Dress

1417 Item Item description: "Material: Polyester (soft fabric make you comfortable to wear) Style: Spaghetti strap mini dress Features: U neck/V neck, 1418 Sleeveless, Spaghetti strap,Backless, High Waist, Back zipper Dress length: Mini length Occasion: Casual, Beach, Club, Prom, Banquet, Party Evening 1419 Attractive style and beautiful design, sexy summer sleeveless dress Ladies Dress ONLY, other accessories photographed not included. U neck style 1420 size: S( US 0 ) Bust:25.19Waist:24.40Length:23.62M(US 2 ) Bust:26.77Waist:25.98Length:24.01L(US 4 ) Bust:28.34Waist:27.55Length:24.40XL(US 6 ) Bust:31.49Waist:29.92Length:24.80Lace V neck style size: S:Bust 32.28;Waist 27.55;Length 30.31M:Bust 33.85;Waist 29.13;Length 30.70L:Bust 35.43;Waist 30.70;Length 31.10XL:Bust 36.22;Waist 31.49;Length 31.49Embroidery V neck style: S:Waist 29.92",Bust 34.64",Length 36.22" M:Waist 31.49",Bust 37.00",Length 36.61" L:Waist 33.07",Bust 39.37",Length 37.00" XL:Waist 33.85",Bust 41.73",Length 37.40" NOTE: 1.Size is Asian sizes,pls allow 1-2 inch size deviation due to manual measurement. 2.Colors may slightly different due to the lighting and monitor. Package Content: 1 x Women dress (Packed in Moxeay designed outpacking!)"

1421 Price: \$14.99 - \$21.99

1422 3. Title: Miusol Women's Casual Flare Floral Contrast Evening Party Mini Dress

1423 Item description: unknown

1424 Price: unknown

1425 4. Title: Damask Embossed Metal Business Card Case

1426 Item Item description: "This business card case features an embossed gold damask pattern on both sides. The case holds up to 10 business cards securely and snaps shut when not in use. Measures 3.75x 2.5x 0.25."

1427 Price: \$3.95

1428 5. Title: Honolulu Jewelry Company Sterling Silver 1mm Box Chain Necklace, 14- 36

1429 Item Item description: "Nickel free sterling silver 1mm box chain. Comes in different sizes with a spring clasp. Rhodium finished to prevent tarnishing. Made in Italy. Gift box included. From Honolulu Jewelry Company, Honolulu, Hawaii."

1430 Price: \$8.99

1431 6. Title: Milumia Women's Button up Split Floral Print Flowy Party Maxi Dress

1432 Item description: unknown

1433 Price: \$18.99 - \$35.99

1434 (Q1) Rank the items id 2 and id 3 based on the user's preferences.

1435 (Q2) Rank the items id 3 and id 6 based on the user's preferences.

1436 (Q3) Rank the items id 6 and id 5 based on the user's preferences.

1437 (Q4) Rank the items id 5 and id 4 based on the user's preferences.

1438 (Q5) Rank the items id 4 and id 1 based on the user's preferences.

1439 (Q6) Pick a review for item id 1 as the user would write it:

1440 REVIEW1: Love,love,love the color of this cardigan. Will be buying more from this seller. I also love the button detail on the sleeve/cuff. Would highly 1441 recommend!

1442 REVIEW2: This cardigan is so soft and comfy! It's a great basic piece that goes with everything. The fit is perfect, and it's really easy to throw on. I 1443 got it for a steal - it's definitely worth the price.

## 1457 Sample Rater Form for Profile Consistency using Amazon Reviews

1458 Given a user with the following purchase history:

1459 1. Title: Zumba Carpet Gliders for Shoes

1460 Item Item description: "Dont get stuck on the carpet floor.Zumba Carpet Glidersallows you to step, shake, swivel and spin on carpet with ease and reduced risk of injury."

1461 Price: unknown

1462 User Rating: 1.0

1463 2. Title: Sterling Silver Leverback Earrings Black Pear Teardrop Made with Swarovski Crystals

1464 Item Item description: "Black faceted crystal teardrops hang from sterling silver rings and leverback earwires. Solid 925 sterling silver, teardrops are approx 5/8 x 3/8 inches. Made with Swarovski Crystals. Gift box or organza bag included, color or style may vary. See Joyful Creations store for matching necklace."

1465 Price: \$17.99

1466 User Rating: 4.0

1467 3. Title: Wipu fashion vintage luxurious pink color crystal brand designer statement women necklace(B379)

1468 Item Item description: "Material:alloy rhinestone ,crystal Size: necklace ribbon chain is 35cm ,pendant is 17cm\*12cm; shipping from China, usually take about 7-15days arrival,if you not accept please don't order, thanks!"

1469 Price: unknown

1470 User Rating: 4.0

1471 4. Title: Kooljewelry Sterling Silver Bead and Diamond-Cut Ball Station Necklace (14, 16, 18, 20, 22, 24, 30, or 36 inch)

1472 Item description: unknown

1473 Price: \$25.99

1474 User Rating: 5.0

1475 5. Title: Dearfoams Women's Sequin Flat Slipper

1476 Item Item description: "Dearfoams is a slipper brand with great awareness. They understand exactly what customers are looking for- from pampering your feet in cozy softness to keeping you one step ahead of the game."

1477 Price: unknown

1478 User Rating: 4.0

1479 6. Title: Women's Chiffon Beachwear Dress Swimwear Bikini Cover-up Made in The USA

1480 Item Item description: "SHORE TRENDZ quality constructed cover-up Chiffon dresses are created with lovely detail. Buy with confidence as they are MADE IN THE USA. These sexy Chiffon dresses have finished edges and lovely pattern detail! We appreciate you visiting and welcome you to check out our store at SHORE TRENDZ for more great items!!!"

1481 Price: \$11.99

1482 User Rating: 1.0

1483 7. Title: Dearfoams Women's Lurex Sweater Knit Ballerina Slipper

1484 Item Item description: "This slipper features a cable knit upper with silver lurex, yarn pom embellishment with silver lurex, and elasticized throat line for secure fit. Brushed terry lining and insole, 10mm high density poly foam insole, with durable, skid resistant, TPR outsole.", "Dearfoams is a slipper brand with great awareness. They understand exactly what customers are looking for- from pampering your feet in cozy softness to keeping you one step ahead of the game."

1485 Price: unknown

1486 User Rating: 3.0

1487 8. Title: Nine West Women's Able Synthetic Platform Pump

1488 Item Item description: "Nine West offers a quick edit of the runways – pinpointing the must have looks of the season, and translating what is fun, hip, and of the moment. It is trend-right footwear that you will reach for in your closet again and again. Nine West is sure to be your trusted resource for everyday chic style."

1489 Price: \$29.99 - \$68.98

1490 User Rating: 5.0

1491 9. Title: Annie Shoes Women's Devine Dress Pump

1492 Item description: unknown

1493 Price: unknown

1494 User Rating: 4.0

1495 10. Title: TinkSky Wedding Tiara Rhinestones Crystal Bridal Headband Pageant Princess Crown

1496 Item description: unknown

1497 Price: \$8.99

1498 User Rating: 3.0

1499 We want to use the following summary to capture user preference from the above purchase history:

1. **Brand Affinity:** Shows a strong preference for Nine West and Annie Shoes, evidenced by the 5-star ratings. A secondary interest in Sterling Silver jewelry.
2. **Style:** Appreciates fashionable slippers and pumps; likely enjoys comfortable yet stylish footwear.
3. **Price Range:** Primarily purchases items in the 10–50 range.
4. **Negative Feedback:** The low rating of the beachwear dress and the tiara headband suggests a critical eye towards embellishments and potentially lower quality.
5. **Purchase type** Most likely female.

1500 Does the above summary faithfully capture user preference from their purchase history?

1501 1 - definite no

1502 2

1503 3

4

5 - definite yes

## Sample Rater Form for Prediction Accuracy using MovieLens

We have the following summary about preference of a user:

This user demonstrates a preference for comedies with a strong comedic impact, leaning towards the more witty and absurdist side. They enjoy action and adventure elements woven into their comedic entertainment, showing an appreciation for thrill and humor combined. While the user appreciates lightheartedness, they are not averse to a more serious tone - the presence of a thriller and a few dramas suggests a willingness to explore different genres, although they seem to lean towards more lighthearted plots. While the user appears to enjoy older comedies, they don't shy away from more contemporary fare, indicating a versatile taste across various time periods. They likely seek out movies with a clear comedic focus rather than those with a heavy dramatic weight. The user isn't afraid to give lower ratings to movies they don't enjoy, suggesting a discerning taste and a desire for quality comedic entertainment.

We will ask questions based on the following list of movies. Please do your own research (using IMDB) if you are not familiar with those movies.

1. Title: Hangover, The (2009)

---

2. Title: Old Boy (2003)

---

3. Title: Sympathy for Mr. Vengeance (Boksuneun naui geot) (2002)

---

4. Title: Let the Right One In (Låt den rätte komma in) (2008)

---

5. Title: Spanking the Monkey (1994)

---

6. Title: Visitor Q (Bizita Q) (2001)

---

(Q1) Rank the movies id 2 and id 6 based on the user's preferences.

(Q2) Rank the movies id 6 and id 5 based on the user's preferences.

(Q3) Rank the movies id 5 and id 1 based on the user's preferences.

(Q4) Rank the movies id 1 and id 3 based on the user's preferences.

(Q5) Rank the movies id 3 and id 4 based on the user's preferences.

### 1551 Sample Rater Form for Profile Consistency using MovieLens

1552 Given a user with the following history:

1553 1. Title: RoboCop (1987)

1554 User Rating: 3.5

1555 2. Title: Chasing Amy (1997)

1556 User Rating: 3.5

1557 3. Title: Grosse Pointe Blank (1997)

1558 User Rating: 3.0

1559 4. Title: Arachnophobia (1990)

1560 User Rating: 3.0

1561 5. Title: Mary Poppins (1964)

1562 User Rating: 1.5

1563 6. Title: Ice Age (2002)

1564 User Rating: 2.0

1565 7. Title: No Country for Old Men (2007)

1566 User Rating: 4.0

1567 8. Title: Wayne's World (1992)

1568 User Rating: 4.5

1569 9. Title: Bad Boys (1995)

1570 User Rating: 2.5

1571 10. Title: Planet of the Apes (1968)

1572 User Rating: 3.5

1573 We want to use the following summary to capture user preference from the above history:

1574 This user enjoys a mix of genres, primarily leaning towards action, comedy, and sci-fi. They demonstrate a preference for films from the 80s and 90s and exhibit a decent tolerance for older movies. Their ratings suggest they favor films with a good balance of action, humor, and engaging plots. The user is not averse to darker or more somber films as long as the storytelling is strong (e.g., No Country for Old Men). Their rating of 'Mary Poppins', however, shows a potential dislike for saccharine, overly-sentimental films. The user shows definite interest in classic sci-fi but is less enthusiastic towards animation and may be reluctant to watch family-oriented content.

1575 Does the above summary faithfully capture user preference from their history?

1576 1 - definite no

1577 2

1578 3

1579 4

1580 5 - definite yes

1581 1551

1582 1552

1583 1553

1584 1554

1585 1555

1586 1556

1587 1557

1588 1558

1589 1559

1590 1560

1591 1561

1592 1562

1593 1563

1594 1564

1595 1565

1596 1566

1597 1567

1598 H QUALITATIVE RESULTS  
15991600 H.1 AMAZON PROFILE EXAMPLE #1  
16011602 **User History:**  
16031604 1. **Title:** New Balance Women's WW665 Walking Shoe  
16051606 **Item description:** "New Balance is dedicated to helping athletes achieve their goals. It's been their  
1607 mission for more than a century to focus on research and development. It's why they don't design  
1608 products to fit an image. They design them to fit. New Balance is driven to make the finest shoes  
1609 for the same reason athletes lace them up: to achieve the very best.", "Get fit and stay chic with the  
1610 WW665 walking shoe from New Balance. Ample mesh allows refreshing breathability while the  
1611 cushy sole delivers comfort stride after stride. A sporty look and soft color palette make this an ideal  
1612 find for your fitness routine."1613 **Price:** \$36.75 - \$62.39  
16141615 **User Rating:** 5.0  
16161617 2. **Title:** Dickies Women's Relaxed Fit Straight Leg Cargo Pant Fade & Wrinkle Resistant  
16181619 **Item description:** "Inseam 32 inches. Fit tip: For accuracy, measure yourself in your undergarments.  
1620 Give all measurements in inches. If your measurements are between sizes, order the larger size."  
16211622 **Price:** \$28.99 - \$86.53  
16231624 **User Rating:** 2.0  
16251626 3. **Title:** Simulated Pink Pearl Rondelle Stretch Bracelet Silver Plated Description:  
16271628 **Item description:** unknown  
16291630 **Price:** \$12.99  
16311632 **User Rating:** 5.0  
16331634 4. **Title:** Eye Catching Women Leather Bracelet Silver Color Beads Cuff Jewelry with Magnetic Clasp  
1635 7.5(Black)  
16361637 **Item description:** "Another eye catching Urban Jewelry bracelet,silver beads color make a stunning  
1638 splash along a luxe multi midnight black leather bracelet with magnetic closure.Provides a touch  
1639 of nature combined style... Guaranteed to add a unique, trendy touch to almost all outfits  
1640 Shipping & Delivery From The USAUrban Jewelry is located in New York City and Offer Worldwide  
1641 ShippingEstimated Delivery Time: In the United States 2-4 Business Days. About Urban Jewelry  
1642 We have a passion for fashion. Our goal is to create a jewelry haven where you will find great  
1643 quality, affordable prices and trendy pieces. Urban Jewelry is an exclusive brand specializing in  
1644 upscale stainlesssteel silver and leather accessories for women, men and teenagers. Urban Jewelry  
1645 is locatedin New York City and ships worldwide. From the runway to your home. Urban Jewelry  
1646 collection features thelatest styles, unique pieces, which will make you happy and your loved ones  
1647 as well..Eye Catching Women Leather Bracelet Silver Color Beads Cuff Jewelry with Magnetic  
1648 Stainless Steel Clasp 7.5(Black)"  
16491650 **Price:** \$9.90  
16511652 **User Rating:** 4.0  
16531654 5. **Title:** Vikoros Women Flowy Lace Overlay Adjustable Strap Crop Top Tank Bustier  
16551656 **Item description:** "Sell by Vikoros Store, pls refer the specification and picture details"  
16571658 **Price:** unknown  
16591660 **User Rating:** 2.0  
16611662 6. **Title:** FRYE Women's Molly D Ring Short Boot  
16631664 **Item description:** "Be the ring leader in the Molly D Ring Short boots by Frye. Hammered  
full grain leather upper. Two buckle accents for a vintage look. Side zipper closure for easy on  
and off. Soft leather lining. Cushioned leather footbed for all-day comfort. Durable leather and

1645 rubber outsole for added traction. Imported. Measurements: Heel Height: 1 in Weight: 1 lb 1 oz  
 1646 Shaft: 5 1/2 in Product measurements were taken using size 8, width B - Medium. Please note  
 1647 that measurements may vary by size.", "The Frye Company is the oldest continuously operated  
 1648 shoe company in the United States. Founded in 1863 by John A. Frye, a well-to-do shoemaker  
 1649 from England, and family-run until 1945, Frye products have a long and illustrious history. Frye  
 1650 boots were worn by soldiers on both sides of America's Civil War, soldiers in the Spanish-American  
 1651 war, and by Teddy Roosevelt and his Rough Riders. When home-steading drew adventurous New  
 1652 England families to the West during the mid and late 1800's many of the pioneers wore Frye Boots  
 1653 for the long journey. Today Frye remains true to its roots with its line of heritage boots, but continues  
 1654 to innovate as it introduces chic new handbags, pumps, and sandals to its collection."

1655 **Price:** \$179.00 - \$361.41

1656 **User Rating:** 4.0

1657 7. **Title:** totes Women's Zelus Snow Boot

1658 **Item description:** "If you are looking for a premium fashionable cold weather boot, look no further  
 1659 than the Totes Zelus. A waterproof rubber shell keeps your feet warm and dry in all types of weather.  
 1660 Top of the line Metro fleece lined and a thick collar for that snuggly fit. Plus a front speed lace  
 1661 system and colored trim gives you the best of the best- The Totes Zelus"

1662 **Price:** unknown

1663 **User Rating:** 5.0

1664 8. **Title:** totes Eric Black

1665 **Item description:** "- Item: Totes Eric. - Upper: Polyurethane. - Style: Mens winter boot featuring a  
 1666 polyurethane upper - rear pull tab and full length zippers for easy entry - faux leather overlays - faux  
 1667 fur lining with padded upper - Thermolite insole. - Sole: Durable rubber lug outsole for traction. -  
 1668 Heel: 1 inch. - Lining: Fully faux fur lined for warmth. - Origin: Imported. - Color: Black."

1669 **Price:** unknown

1670 **User Rating:** 5.0

1671 9. **Title:** Lonestar Jewelry Lacrosse Stick &Necklace

1672 **Item description:** "Exclusive Allan Robin Lacrosse Necklace with Lacrosse Stick Charm in Ball  
 1673 Chain Known to be the best gift for a Lacrosse player or enthusiast. Surprise your family with this  
 1674 perfect gift. Beautiful Lacrosse gift for girls and it comes in an elegant cotton-filled jewelry box so  
 1675 pretty easy to store. It's lead & nickel free so no irritation to sensitive skin. They can wear it at any  
 1676 occasion or casually while doing day-to-day activities. This piece of modern fashion jewelry is  
 1677 unique enough to show your loved ones passion to the world, without being too flashy. A beautiful  
 1678 Lacrosse Stick Charm which Measures 1 1/4 x 3/8 Not Only Is This Lacrosse Jewelry A Great Gift  
 1679 From Parents or Grandparents, But Would Also Be A Wonderful Gesture From A Coach To Their  
 1680 Students! Order Yours Now And Dont Forget To Check Out The Hundreds Of Other Unique Sports  
 1681 Gifts That We Offer For All Of The Athletes In Your Life!"

1682 **Price:** unknown

1683 **User Rating:** 5.0

1684 10. **Title:** M&c Women's | Boyfriend Style & Chronograph Look | FC0385

1685 **Item description:** "The Marciano & Co. house brand M&c offers customers superior value in an  
 1686 increasingly competitive market. The M&c product lines draw inspiration from the most popular  
 1687 designs and are available at a fraction of designer prices. The Marciano & Co. mission remains  
 1688 consistent; to provide customers with products that are novel in design and affordable without  
 1689 compromising quality."

1690 **Price:** \$14.99

1691 **User Rating:** 3.0

1692 **User Profile:** This user demonstrates varied shopping interests, with a clear preference for practical and  
 1693 comfortable footwear and accessories for themselves and potentially others. A recurring theme is a fondness  
 1694 for items emphasizing ease and warmth, evidenced by the 5-star ratings for snow boots (Totes Zelus & Eric),  
 1695 walking shoes (New Balance), and a stretch bracelet (Simulated Pink Pearl).

1696 While appreciating affordability (as revealed by purchases from brands like totes and M&c), the user isn't  
 1697 averse to investing in quality when it aligns with their needs, indicated by the 4-star rating for Frye boots.  
 1698 Style leans towards sporty-casual and practical, with a tolerance for some embellishments (beaded bracelet,  
 1699 lacrosse necklace – likely a gift). Negative ratings suggest an unwillingness to purchase ill-fitting (Dickies  
 1700 pants) or poorly-performing (Vikorous top) items. An unclear gender preference outside of sale of women's  
 1701 apparel.

1702

1703 **H.2 AMAZON PROFILE EXAMPLE #2**

1704

1705 **User History:**

1706

1707 1. **Title:** Hand Crocheted Beanie (02)-Rasta

1708

1709 **Item description:** "This beautifully thick, hand crocheted cotton beanie will brighten up your cold  
 1710 days with its unbeatable style and insulating warmth. The clean, round crown shape is achieved by 1  
 1711 panel of tightly crochet cotton, of repeating black, green, yellow and red stripes. The base of the  
 1712 crown is made of an open crochet pattern with the band at the opening crocheted with elastic to give  
 1713 it a lot of stretch for a perfect fit around any head. This hat is available in other stripe patterns as  
 1714 well as in solid colors to meet your particular taste. Made of 100% cotton. Hand crocheted item.  
 1715 Crown measures 6 deep, 7 wide. ONE SIZE fits most, from sizes 6 - 7 5/8. Available in an array of  
 1716 colors. Imported."

1717

1718 **Price:** \$4.99

1719

1720 **User Rating:** 3.0

1721

1722 2. **Title:** Intimo Men's Tricot Travel Pajama Set - Big Man Sizes

1723

1724 **Item description:** Every man should have a comfortable Nylon Travel Boxers. Wrinkle- Free."

1725

1726 **Price:** unknown

1727

1728 **User Rating:** 5.0

1729

1730 3. **Title:** Men Purple Mesh Pocket Shorts Inner Drawstring Avail Size S-5X Item description: unknown

1731

1732 **Price:** unknown

1733

1734 **User Rating:** 5.0

1735

1736 4. **Title:** Breda Men's 1627-Gold Mitchell Multi Time Zone Watch

1737

1738 **Item description:** "A great-looking timepiece from Breda, this watch utilizes an excellent blend  
 1739 of materials and mechanics to create a functional accessory with a stylish flair. With the perfect  
 1740 mix of style and comfort, this watch will quickly become one of the most popular members of your  
 1741 watch collection.", "The Mitchel Collection", "This three time zone, large face men 2019s watch is  
 1742 available in three colors.", "Breda Watches", "Breda. Original style. Non-singular aesthetic.", "A  
 1743 creative collective with a shared appreciation for design that tells more than one story. We believe  
 1744 in the freedom of self-expression through style. Breda was born when we poured our collective  
 1745 imagination and expertise into designing watches inspired by a global lens. An eclectic unit of  
 1746 ambitious artists, designers, business-brains, photographers, innovators and style rebels, we've come  
 1747 together to create pieces that intrigue, inspire and challenge the expected.", "Our process of creation  
 1748 is our own unique alchemy. A key principle is that of discovery: we live contemporary culture,  
 1749 explore the past and dream up the future, scour the world's fashion stages and streets to challenge  
 1750 and inspire each other's imaginations. Then we design.", "We work with global partners to source  
 1751 the latest materials. With meticulous attention to detail, each innovative product is born with the  
 1752 purpose of transcending the ordinary."

1739                   **Price:** unknown  
 1740                   **User Rating:** 5.0  
 1741  
 1742           5. **Title:** Breda Men's 1627-silver Mitchell Multi Time Zone Watch  
 1743           **Item description:**", "A great-looking timepiece from Breda, this watch utilizes an excellent blend  
 1744           of materials and mechanics to create a functional accessory with a stylish flair. With the perfect  
 1745           mix of style and comfort, this watch will quickly become one of the most popular members of your  
 1746           watch collection.", "The Mitchel Collection", "This three time zone, large face men2019s watch is  
 1747           available in three colors.", "Breda Watches", "Breda. Original style. Non-singular aesthetic.", "A  
 1748           creative collective with a shared appreciation for design that tells more than one story. We believe  
 1749           in the freedom of self-expression through style. Breda was born when we poured our collective  
 1750           imagination and expertise into designing watches inspired by a global lens. An eclectic unit of  
 1751           ambitious artists, designers, business-brains, photographers, innovators and style rebels, we've come  
 1752           together to create pieces that intrigue, inspire and challenge the expected.", "Our process of creation  
 1753           is our own unique alchemy. A key principle is that of discovery: we live contemporary culture,  
 1754           explore the past and dream up the future, scour the world's fashion stages and streets to challenge  
 1755           and inspire each other's imaginations. Then we design.", "We work with global partners to source  
 1756           the latest materials. With meticulous attention to detail, each innovative product is born with the  
 1757           purpose of transcending the ordinary."  
 1758           **Price:** unknown  
 1759           **User Rating:** 2.0  
 1760  
 1761           6. **Title:** Breda Men's 1627-black Mitchell Multi Time Zone Watch  
 1762           **Item description:**", "A great-looking timepiece from Breda, this watch utilizes an excellent blend  
 1763           of materials and mechanics to create a functional accessory with a stylish flair. With the perfect  
 1764           mix of style and comfort, this watch will quickly become one of the most popular members of your  
 1765           watch collection.", "The Mitchel Collection", "This three time zone, large face men2019s watch is  
 1766           available in three colors.", "Breda Watches", "Breda. Original style. Non-singular aesthetic.", "A  
 1767           creative collective with a shared appreciation for design that tells more than one story. We believe  
 1768           in the freedom of self-expression through style. Breda was born when we poured our collective  
 1769           imagination and expertise into designing watches inspired by a global lens. An eclectic unit of  
 1770           ambitious artists, designers, business-brains, photographers, innovators and style rebels, we've come  
 1771           together to create pieces that intrigue, inspire and challenge the expected.", "Our process of creation  
 1772           is our own unique alchemy. A key principle is that of discovery: we live contemporary culture,  
 1773           explore the past and dream up the future, scour the world's fashion stages and streets to challenge  
 1774           and inspire each other's imaginations. Then we design.", "We work with global partners to source  
 1775           the latest materials. With meticulous attention to detail, each innovative product is born with the  
 1776           purpose of transcending the ordinary."  
 1777           **Price:** unknown  
 1778           **User Rating:** 3.0  
 1779  
 1780           7. **Title:** Camouflage Cap, Color: Orange Camo, Size: One Size Item description: unknown Price:  
 1781           \$8.66  
 1782           **User Rating:** 5.0  
 1783  
 1784           8. **Title:** LRG Men's Big-Tall Core Collection Nine Camo T-Shirt  
 1785           **Item description:**The Core Collection Nine Camo Tee is constructed of Jersey featuring a contrasting  
 1786           pattern and custom camo print expertly rendered throughout with custom tag on neckline and hem."  
 1787           **Price:** unknown  
 1788           **User Rating:** 5.0  
 1789  
 1790           9. **Title:** Rocawear Men's Big-Tall Gradient Black Hearts Item description: unknown Price: unknown  
 1791           **User Rating:** 5.0

1786 10. **Title:** Kangol Men's Tropic Galaxy Cap  
 1787 **Item description:** The galaxy is a kangol interpretation of a classic 8-quarter cap"  
 1788 **Price:** \$37.40  
 1789 **User Rating:** 5.0  
 1790

1791 **User Profile:** This user demonstrates a strong preference for men's apparel, particularly casual wear. They  
 1792 consistently rate items favorably (mostly 5-star) suggesting a generally positive shopping experience and  
 1793 valuing style. Key interests include:  
 1794

- 1795 • **Streetwear & Camouflage:** A recurring theme of camouflage patterns and brands like LRG and  
 1796 Rocawear indicates a liking for streetwear aesthetics.
- 1797 • **Caps & Hats:** Several purchases of caps (Kangol, camouflage, orange camo) indicate a strong  
 1798 interest in headwear.
- 1799 • **Timepieces:** The purchase of the Breda watches, despite a mixed rating experience reveals an  
 1800 interest in watches.
- 1801 • **Comfortable Loungewear:** The purchase of pajama sets highlights a preference for comfort.
- 1802 • **Price Sensitivity:** The moderate rating on the beanie suggests a mindful of price.

1803

1804 The user appears to be male, likely younger, and influenced by contemporary trends. They are open to  
 1805 experimenting with different styles while maintaining a casual, streetwear-inspired aesthetic.

1806

1807 H.3 AMAZON PROFILE EXAMPLE #3

1808

1809 **User History:**

1810

- 1811 1. **Title:** Rubik's Cube - Mens Hip To Be Square Soft T-shirt Small Off-white  
 1812 **Item description:** "From Rubiks Cube comes this soft cotton short sleeve T-Shirt in off-white that  
 1813 features The Rubiks Cube and says Hip To Be Square. Great T-Shirt for anyone who loves The  
 1814 Rubiks Cube."  
**Price:** unknown  
**User Rating:** 3.0
- 1815 2. **Title:** Intimo Women's Printed Microfleece Pajama Pant  
**Item description:** unknown  
**Price:** \$18.00  
**User Rating:** 5.0
- 1816 3. **Title:** Minecraft Boys' Adventure Youth Tee  
**Item description:** "The first time you find yourself staring out at the expansive world of blocks  
 1817 laid out before you, the possibilities are truly endless. What will you choose to do in this biome,  
 1818 generated just for you you can build soaring castles, dig sprawling underground complexes, or  
 1819 maybe you'll set out to see what is just over the horizon. Whichever way you choose to play, each  
 1820 adventure will be your very own."  
**Price:** \$14.50 - \$28.99  
**User Rating:** 4.0
- 1821 4. **Title:** Hollywood Star Fashion Casual Basic Women's Semi-Crop Camisole Cami Tank Top with  
 1822 Adjustable Straps  
**Item description:** "This is a long length Tank top Versatile Basic Spaghetti Strap Satin Trim Stretch  
 1823 Camisole Tank Yoga Everyday Active Adventure Travel Fitted Scoop neckline, adjustable shoulder  
 1824 straps Satin Trim Fully stretchable Please note: this top dose NOT feature a built-in shelf bra Body  
 1825

1826

1827

1828

1829

1830

1831

1832

1833 length in size medium: 28" 95% Cotton, 5% Spandex Imported. Satisfaction guaranteed Returns  
 1834 accepted. We ship worldwide"  
 1835  
 1836 **Price::** \$5.00 - \$19.99  
 1837 **User Rating:** 5.0  
 1838 5. **Title:** MANDI HOME Hot Sale Wedding Fashion 925 Silver Plated Jewelry Set Big Hand Chain  
 1839 Bracelet Necklace Ring Stud Earings Eardrop Water Drops  
 1840 **Item description:** "1pcs for each!"  
 1841 **Price:** \$6.35  
 1842 **User Rating:** 2.0  
 1843 6. **Title:** Fruit of the Loom Girls' Cotton Spaghetti Strap Sport Bra  
 1844 **Item description:** "Fruit of the Loom girls' spaghetti strap sport bras provide comfort and support  
 1845 for any activity. The cotton and lycra materials give them a soft feel and also make them machine-  
 1846 washable. Give the girl in your life a good start with Fruit of the Loom. Designed to give her support  
 1847 while giving her room to grow, tried and true Fruit of the Loom underclothes are the perfect choice  
 1848 for your growing girl"  
 1849 **Price:** \$7.53 - \$12.68  
 1850 **User Rating:** 5.0  
 1851 7. **Title:** Fruit Of The Loom Women's Breathable Underwear Multipack (Assorted)  
 1852 **Item description:** "Our breathable cotton collection is designed with a soft, cotton-mesh fabrication  
 1853 that not only supports the flow of air but wicks away moisture to help you stay cool and dry all day  
 1854 long."  
 1855 **Price:** \$4.50 - \$45.00  
 1856 **User Rating:** 5.0  
 1857 8. **Title:** Froomer Womens Winter Outwear Light Down Coat Hooded Jackets  
 1858 **Item description:** unknown  
 1859 **Price:** unknown  
 1860 **User Rating:** 5.0  
 1861 9. **Title:** Rockland Luggage 17 Inch Rolling Backpack, Pink Dot, Medium  
 1862 **Item description:** "Rolling backpack with double skate wheels. Measures 13 inch l x 10 inch w x 17  
 1863 inch h. Features include 2 zippered pockets on the front of the bag with one having a built in organizer.  
 1864 Zippered main compartment opens to large storage opening, and two mesh side pockets. Fabric  
 1865 handle at the top and telescoping pull handle. Bag made of durable polyester with metal telescoping  
 1866 handle and heavy duty plastic stand on bottom to keep the bag standing.", "R01-PINKGIRAFFE  
 1867 shown.", "", "", "", ""  
 1868 **Price:** \$3.24  
 1869 **User Rating:** 5.0  
 1870 10. **Title:** Bags for Less Large Clear Vinyl Tote Bags Shoulder Handbag (Royal Blue)  
 1871 **Item description:** unknown  
 1872  
 1873 **User Profile:** This user demonstrates a preference for affordable, practical, and comfortable clothing and  
 1874 accessories, primarily for women and children. While they appreciate value, ratings suggest a focus on  
 1875 functionality over high-end brands, as evidenced by consistently positive reviews for items like Fruit of the  
 1876 Loom underwear and Fruit of the Loom sport bras. The 5-star rating for a Rockland luggage rolling backpack  
 1877 indicates an appreciation for travel-friendly and durable goods.  
 1878 They show an interest in basic apparel and accessories (camisoles, pajamas). There's a slight openness to  
 1879 novelty items (Rubik's Cube shirt, Minecraft tee), but overall, their choices lean towards everyday essentials.

1880 The lower rating on the jewelry set suggests a less enthusiastic response to more elaborate items. The user  
 1881 appears to gravitate toward comfort, utility, and reasonable pricing.  
 1882

1883 **H.4 AMAZON PROFILE EXAMPLE #4**  
 1884

1885 **User History:**  
 1886

1887 1. **Title:** Naturalizer Women's Ringo Sandal

1888 **Item description:** "The Ringo is a slingback sandal that features an N5 Comfort System and a  
 1889 manmade outsole.", "Naturalizer was one of the first shoe brands that women could turn to for the  
 1890 feminine style they coveted and the comfort they thought was impossible to attain. Naturalizer's  
 1891 fresh, unpretentious designs are a smooth fit with your wardrobe, your life and your own unique  
 1892 style. Naturalizer promises style that makes you look good and feel good - always."

1893 **Price:** unknown

1894 **User Rating:** 4.0

1895 2. **Title:** Aerosoles Women's Tapestry

1896 **Item description:** "Every girl needs a timeless pump like the Aerosoles Tapestry. Featuring a 3  
 1897 covered heel, softly angled toe and classic lines for a tried-and-true look. Stunningly soft memory  
 1898 foam insole is stitched and cushioned for your comfort, while the flexible rubber sole with diamond  
 1899 pattern drinks up hard impact. You'll feel great all day long!", "Destined to be your new favorite,  
 1900 Tapestry from Aerosoles Women's offers professional polish for the office and beyond. Showcasing  
 1901 a classic silhouette, this pretty pump features a leather or fabric upper that slips on to reveal the  
 1902 unbelievable comfort from Aerosoles that you have come to know and love. A flexible rubber outsole  
 1903 and a modest heel add a tasteful touch to any ensemble."

1904 **Price:** \$64.99

1905 **User Rating:** 2.0

1906 3. **Title:** Eagle Creek Travel Gear Undercover Money Belt (Khaki)

1907 **Item description:** "Looking for money belts? Look no further than this simple waist-worn under-  
 1908 clothing solution. Keep important travel documents and personal identification items out of sight  
 1909 in this money belt. It's made of durable and lightweight rip-stop fabric with a moisture-wicking  
 1910 and breathable back panel. Complete with zippered pocket for secure organization and soft elastic  
 1911 waistband with strap keeper. When you're not wearing it, simply tuck the strap into the slip pocket on  
 1912 the back, which was conveniently created for waist strap storage. Travel solutions that make sense."

1913 **Price:** \$15.85

1914 **User Rating:** 5.0

1915 4. **Title:** Maidenform Women's Comfort Devotion Demi Bra

1916 **Item description:** "Magnificently smooth and supportive. Maidenform's Comfort Devotion Demi  
 1917 Bra features foam, contour underwire cups made of plush fabric and smoothing wings with super  
 1918 soft fabric on the inside. Line dry or lay flat to dry"

1919 **Price:** \$15.80 - \$94.21

1920 **User Rating:** 5.0

1921 5. **Title:** Champion Women's Jersey Pant

1922 **Item description:** "Champion Jersey Pant with a rib waistband is just the right fit and look for  
 1923 everyday wear."

1924 **Price:** \$13.20 - \$72.06

1925 **User Rating:** 3.0

1926 6. **Title:** Men's Cotton Casual Ankle Socks

1927 **Item description:** "", ""

1927                   **Price:** \$15.90  
 1928                   **User Rating:** 5.0  
 1929  
 1930                   7. **Title:** Champion Absolute Sports Bra With SmoothTec Band  
                         **Item description:** "The Absolute workout bra solids and prints at a great value. This bra has a patented smooth tec band for the ultimate in chafe resistance and comfort. A must have for any gym bag."  
                         **Price:** \$7.93 - \$48.00  
                         **User Rating:** 5.0  
 1931  
 1932  
 1933  
 1934  
 1935  
 1936                   8. **Title:** uxcell Men Point Collar Button Down Long Sleeves Plaid Detail Slim Fit Shirts  
                         **Item description:** "Description:One mock pocket point collar button down long sleeves plaid detail slim fit shirt. Feature mock pocket, plaids detail for build up your special character. Buttoned point collar for standard button down shirt. Buttoned cuffs is fused to keep a crisp, dressy appearance. Soft touch, comfortable fabric which is comfort to wear in all season. Suitable for date, daily work, travel and everyday wear. Match with formal trousers or stylish denim pants to build up fashion casual look. Please check your measurements to make sure the item fits before ordering. Body Size Chart (in inches)International  
                         **Price:** \$12.24 - \$19.81  
                         **User Rating:** 2.0  
 1937  
 1938  
 1939  
 1940  
 1941  
 1942  
 1943  
 1944  
 1945                   9. **Title:** Mens Colorful Dress Socks Argyle - HSELL Men Multicolored Argyle Pattern Fashionable Fun Crew Socks Item description: unknown  
                         **Price:** \$11.99  
                         **User Rating:** 5.0  
 1946  
 1947  
 1948  
 1949                   10. **Title:** Marvel Comics Crew Socks Mens Mix Hulk Spiderman, The Avengers Large With Bonus One XWi Wrist Sweatband  
                         **Item description:** "A simply Marvel-ous way to get your man to change socks regularly, each set of socks feature classic artwork from the comics themselves, with full color, detailed woven images of the likes of Avengers past and present, including The Hulk, Spider-Man, Iron Man and Captain America Designs: Assorted Contents: 98% polyester, 2% elastane Our guarantee: 100% official merchandise Care instructions: Machine washable Packaging: Hanger Brand: Marvel Comics Size: USA(7-12) UK(6-11) Includes One Xportsworld Innovations Wrist Sweatband (Color chosen randomly)Composition: 80% Polyester, 20% Cotton Size: Width 8cm x Length 9cm ( W 3.14In x L 3.54In)"  
                         **Price:** unknown  
                         **User Rating:** 4.0  
 1950  
 1951  
 1952  
 1953  
 1954  
 1955  
 1956  
 1957  
 1958  
 1959  
 1960  
 1961                   **User Profile:** This user demonstrates a practical and value-focused shopping style with varied interests. They show a clear preference for comfortable and practical items, consistently rating items highly that prioritize these aspects. A strong inclination towards athletic wear and socks (Champion, Marvel socks) is evident, suggesting an active lifestyle or interest in sports-related apparel. They appreciate quality, as shown by the 4-star rating for Naturalizer sandals.  
 1962  
 1963  
 1964  
 1965  
 1966                   While primarily focused on apparel, the purchase of a money belt suggests travel interests. The mixed reviews (2-star for Aerosoles, 2-star for uxcell shirt) indicate potential sensitivity to fit or style outside their comfort zone. Maidenform bra and women's socks suggest occasional purchases of items for a female partner. The user appears to find good deals and prioritizes functionality over high-fashion trends.  
 1967  
 1968  
 1969  
 1970  
 1971                   H.5 MOVIELENS EXAMPLE #1  
 1972  
 1973                   **User History:**

1974 1. **Title:** Star Wars: Episode IV - A New Hope (1977)  
**User Rating:** 5.0

---

1975 2. **Title:** Star Wars: Episode V - The Empire Strikes Back (1980)  
**User Rating:** 5.0

---

1976 3. **Title:** Raiders of the Lost Ark (Indiana Jones and the Raiders of the Lost Ark) (1981)  
**User Rating:** 4.5

---

1977 4. **Title:** Matrix, The (1999)  
**User Rating:** 5.0

---

1978 5. **Title:** Aliens (1986)  
**User Rating:** 3.0

---

1979 6. **Title:** Usual Suspects, The (1995)  
**User Rating:** 4.0

---

1980 7. **Title:** Léon: The Professional (a.k.a. The Professional) (Léon) (1994)  
**User Rating:** 4.0

---

1981 8. **Title:** Star Wars: Episode VI - Return of the Jedi (1983)  
**User Rating:** 5.0

---

1982 9. **Title:** Dark Knight, The (2008)  
**User Rating:** 3.5

---

1983 10. **Title:** Godfather, The (1972)  
**User Rating:** 4.0

---

2004 **User Profile:** This user enjoys intellectually stimulating and visually impressive films, favoring narratives with depth and complex characters. They appreciate cinematic excellence and are willing to engage with challenging themes, demonstrating a sophisticated taste. A strong preference exists for science fiction and action, particularly those with philosophical underpinnings, as evidenced by high ratings of “Star Wars” films and “The Matrix”.

2005 They appreciate well-crafted thrillers and crime dramas, like “The Usual Suspects” and “Léon: The Professional”, valuing intricate plots and suspenseful storytelling. Classic films, notably “The Godfather”, are also enjoyed, suggesting an appreciation for cinematic history.

2006 The user isn’t limited to specific genres, embracing a diverse range of eras and cultural influences, with a noticeable fondness for European cinema. They are comfortable with slower pacing and unconventional structures, valuing atmosphere and character development. Overall, this individual seeks movies that provoke thought and deliver lasting impressions.

2007

2008 H.6 MOVIELENS EXAMPLE #2

2009

2010 **User History:**

2021 1. **Title:** Last of the Mohicans, The (1992)  
 2022 **User Rating:** 3.0

---

2023 2. **Title:** Ghostbusters (a.k.a. Ghost Busters) (1984)  
 2024 **User Rating:** 4.0

---

2025 3. **Title:** Star Wars: Episode I - The Phantom Menace (1999)  
 2026 **User Rating:** 5.0

---

2027 4. **Title:** Police Academy 3: Back in Training (1986)  
 2028 **User Rating:** 3.0

---

2029 5. **Title:** E.T. the Extra-Terrestrial (1982)  
 2030 **User Rating:** 3.0

---

2031 6. **Title:** 13th Warrior, The (1999)  
 2032 **User Rating:** 4.0

---

2033 7. **Title:** Being John Malkovich (1999)  
 2034 **User Rating:** 4.0

---

2035 8. **Title:** Blair Witch Project, The (1999)  
 2036 **User Rating:** 3.0

---

2037 9. **Title:** Bug's Life, A (1998)  
 2038 **User Rating:** 4.0

---

2039 10. **Title:** Clerks (1994)  
 2040 **User Rating:** 5.0

---

2041 **User Profile:** This user demonstrates a preference for films that are both engaging and thought-provoking.  
 2042 They appreciate a balance between entertaining narratives and deeper themes, often favoring stories with  
 2043 complex characters and layered plots. A strong interest in both action and comedy is evident, alongside a  
 2044 clear appreciation for science fiction and suspenseful thrillers.  
 2045  
 2046 The user enjoys films with strong directorial vision and is not afraid to explore unconventional storytelling  
 2047 techniques. They value realistic characters and are drawn to films that delve into the human condition, often  
 2048 with a focus on relationships and identity.  
 2049  
 2050 Despite a love for contemporary cinema, they also appreciate classic films, signaling a broad appreciation  
 2051 for cinematic history and diverse cultural influences. They are open to a range of emotional tones, from  
 2052 lighthearted to dramatic, and demonstrate a tolerance for challenging subject matter. Ultimately, this user  
 2053 seeks out movies that are memorable, intellectually stimulating, and leave a lasting impression.

2054

2055

2056

2057

2058

2059

2060

2061

2062

2063

2064

2065 H.7 MOVIELENS EXAMPLE #3

2066

2067 **User History:**

2068 1. **Title:** Batman (1989)  
 2069 **User Rating:** 4.0

---

2071 2. **Title:** Apollo 13 (1995)  
 2072 **User Rating:** 3.0

---

2074 3. **Title:** Pulp Fiction (1994)  
 2075 **User Rating:** 5.0

---

2078 4. **Title:** True Lies (1994)  
 2079 **User Rating:** 5.0

---

2081 5. **Title:** Die Hard: With a Vengeance (1995)  
 2082 **User Rating:** 5.0

---

2084 6. **Title:** Aladdin (1992)  
 2085 **User Rating:** 3.0

---

2087 7. **Title:** Ace Ventura: Pet Detective (1994)  
 2088 **User Rating:** 4.0

---

2090 8. **Title:** Batman Forever (1995)  
 2091 **User Rating:** 5.0

---

2094 9. **Title:** Shawshank Redemption, The (1994)  
 2095 **User Rating:** 3.0

---

2097 10. **Title:** Fugitive, The (1993)  
 2098 **User Rating:** 4.0

---

2100 **User Profile:** This user enjoys action-packed, engaging films with a strong sense of adventure and spectacle.  
 2101 They have a pronounced taste for thrillers and superhero movies, evident in their high ratings for iconic films  
 2102 like **Batman** and **Die Hard**. A preference for visually impressive and exciting narratives is clear.  
 2103

2104 They appreciate well-constructed plots and fast-paced storytelling, enjoying films that deliver immediate  
 2105 entertainment. However, the user also values character-driven narratives, demonstrated by the consistent 4.0  
 2106 rating for **The Shawshank Redemption**.

2107 There's a clear appreciation for entertainment that is accessible and exciting. They generally gravitate towards  
 2108 movies with a positive tone and fun energy, similar to **Aladdin** and **Ace Ventura**. The user isn't overly  
 2109 concerned with complex themes, valuing thrills and excitement over deeper philosophical explorations. A  
 2110 preference for blockbuster entertainment is evident, alongside a fondness for the superhero genre.

2111 **H.8 MOVIELENS EXAMPLE #4**

2112 **User History:**

---

2115 1. **Title:** Godfather: Part III, The (1990)  
 2116 **User Rating:** 3.0

---

2118 2. **Title:** Stepmom (1998)  
 2119 **User Rating:** 3.0

---

2121 3. **Title:** Blair Witch Project, The (1999)  
 2122 **User Rating:** 3.0

---

2124 4. **Title:** Haunting, The (1963)  
 2125 **User Rating:** 5.0

---

2127 5. **Title:** Abbott and Costello Meet Frankenstein (1948)  
 2128 **User Rating:** 3.0

---

2130 6. **Title:** Black Sabbath (Tre volti della paura, I) (1963)  
 2131 **User Rating:** 5.0

---

2133 7. **Title:** General, The (1926)  
 2134 **User Rating:** 5.0

---

2136 8. **Title:** Wings (1927)  
 2137 **User Rating:** 4.0

---

2139 9. **Title:** Jail Bait (1954)  
 2140 **User Rating:** 3.0

---

2142 10. **Title:** M (1931)  
 2143 **User Rating:** 5.0

---

2145 **User Profile:** This user enjoys films with a strong narrative and appreciates classic cinema, particularly those from the mid-20th century. They favor dramas and thrillers, often with complex characters and thought-provoking themes, demonstrating a sophisticated taste. A significant portion of their ratings fall within the 3.0 – 5.0 range, indicating an ability to appreciate both well-executed and critically acclaimed movies.

2150 They are open to diverse genres and cultural origins, showing a preference for European and international films. While they enjoy suspenseful and eerie stories like “The Haunting,” they also appreciate lighter, 2151 comedic fare like “Abbott and Costello Meet Frankenstein.” The user values strong directorial vision and 2152 cinematic artistry, favoring films that build atmosphere and explore deeper human experiences.

2154 A tolerance for slower pacing and unconventional storytelling is evident. They’re not afraid to give a lower 2155 rating to films that don’t meet their expectations. The user consistently seeks out films that are visually 2156 engaging and emotionally resonant, demonstrating a deep appreciation for the art of filmmaking.