Where You Go is Who You Are: Behavioral Theory-Guided LLMs for Inverse Reinforcement Learning

Anonymous ACL submission

Abstract

Big trajectory data hold great promise for human mobility analysis, but their utility is often constrained by the absence of critical traveler attributes, particularly sociodemographic information. While prior studies have explored predicting such attributes from mobility patterns, they often overlooked underlying cognitive mechanisms and exhibited low predictive accuracy. This study introduces *SILIC*, short for Sociodemographic Inference with LLM-guided Inverse Reinforcement Learning (IRL) and Cognitive Chain Reasoning (CCR), a theoretically grounded framework that leverages LLMs to infer sociodemographic attributes from observed mobility patterns by capturing latent behavioral intentions and reasoning through psychological constructs. Particularly, our approach explicitly follows the Theory of Planned Behavior (TPB), a foundational behavioral framework in transportation research, to model individuals' latent cognitive processes underlying travel decision-making. The LLMs further provide heuristic guidance to improve IRL reward function initialization and update by addressing its ill-posedness and optimization challenges arising from the vast and unstructured reward space. Evaluated in the 2017 Puget Sound Regional Council Household Travel Survey, our method substantially outperforms state-of-the-art baselines and shows great promise for enriching big trajectory data to support more behaviorally grounded applications in transportation planning and beyond.

1 Introduction

Understanding human mobility patterns is critical for many fields, such as transportation engineering (Luca et al., 2021), marketing (Ghose et al., 2019), urban planning (Haraguchi et al., 2022), and emergency management (Yabe et al., 2016; Zhao et al., 2022). In recent years, researchers and practitioners have been leveraging real-world trajectory data generated by mobile devices (or synthetic trajectories (Zhu et al., 2023; Kim and Jang, 2024)) to analyze and model people's movements to facilitate decision-making (Ghose et al., 2019; Zhao et al., 2022). However, these trajectory datasets often lack traveler attributes, particularly sociodemographic information, limiting their usefulness for important applications such as causal analysis to inform transportation planning and policy (Holz-Rau and Scheiner, 2019), personalized incentives for targeted marketing (Zhong et al., 2015), and tailored crisis communication to enhance evacuation compliance (Leykin et al., 2016).

Some prior studies have explored inferring sociodemographic attributes from people's travel trajectories, e.g., Chen et al. (2024); Zhang et al. (2024a). These studies typically used classical machine learning models (e.g., Support Vector Machines, XGBoost) to predict attributes like gender, age, or income from features extracted from individual trajectories (Zhu et al., 2017; Wu et al., 2019; Bakhtiari et al., 2023). While insightful, these approaches often exhibit limited predictive performance, reducing their effectiveness in large-scale real-world applications.

A key reason for their poor performance is the oversimplification of the link between sociodemographic attributes and mobility, often ignoring *underlying cognitive processes*. Human mobility is complex and shaped by latent cognitive processes, often modeled using the Theory of Planned Behavior (TPB) (Ajzen, 1991), as shown in Figure 1 (upper section). TPB posits that background factors, such as sociodemographic and contextual attributes, inform three core beliefs: attitudes, subjective norms, and perceived behavioral control (Ajzen, 2020). These beliefs collectively shape individuals' intentions, which ultimately drive their observable behaviors (e.g., travel decisions). Therefore, predicting sociodemographic attributes directly from mobility patterns, without accounting

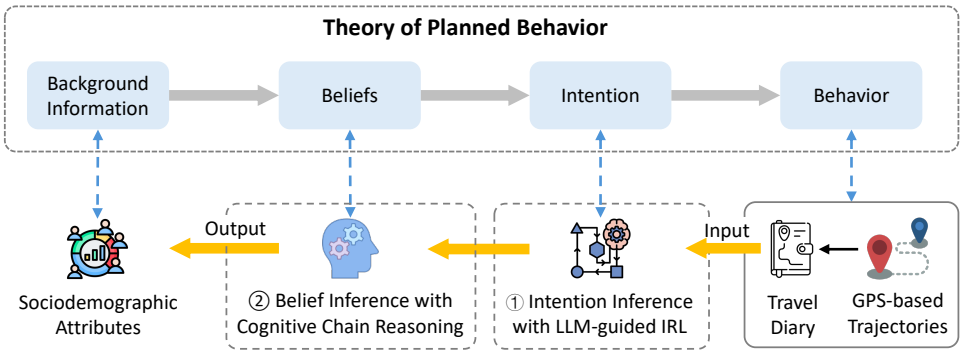

Figure 1: Overview of our proposed framework. We inversely follow the Theory of Planned Behavior (TPB) to predict sociodemographic attributes from travel trajectories. Our method first uses an LLM-guided IRL model to infer behavioral intentions, followed by a Cognitive Chain Reasoning strategy that predicts sociodemographic attributes via intermediate belief constructs.

for these mediating constructs, may lead to inaccurate or incomplete inferences.

Addressing this limitation requires models capable of reasoning over latent cognitive constructs that mediate behavior—a capacity demonstrated by large language models (LLMs) (Nguyen et al., 2024; Lee et al., 2024; Liu et al., 2025b). Unlike traditional machine learning models, LLMs exhibit strong reasoning and theory-of-mind (ToM) abilities (Street, 2024; Zhang et al., 2024b; Bandyopadhyay et al., 2025). However, predicting sociodemographic attributes from mobility patterns requires inversely following the TPB framework. Leveraging existing LLMs for this inverse cognitive modeling still face the following challenges: (1) **Misalignment with behavioral theory**: Without structural guidance (Zhou et al., 2023), LLMs often capture only partial reasoning patterns, leading to incomplete or theoretically inconsistent interpretations (Tjuatja et al., 2024); (2) **Lack of iterative behavioral refinement**: Although LLMs can perform inverse alignment to infer latent mental states from behavior, they often generate intentions in a single pass without feedback refinement (Sun et al.; Sun and van der Schaar, 2024); and (3) **Ambiguity in mapping mental states to identity**: Similar mental states can arise across different sociodemographic groups due to shared external influences, making LLMs prone to misclassification without contextual input (Chen et al., 2025).

Based on these challenges, we raise a key question: *How can we construct a principled framework that efficiently leverages LLMs to inversely follow the cognitive pathway in the TPB to recover sociodemographic attributes?*

To address this research question, we propose *SILIC*, a two-stage framework, as illustrated in Figure 1 (bottom section). In the first stage, we propose an LLM-guided Inverse Reinforcement Learning (IRL) approach to infer individualized reward functions (Ng et al., 2000) that reflect intentions (Liang et al., 2025) from travel trajectories (Figure 1, dashed box denoted as ①). Our core idea is to address the challenge of limited iterative behavioral refinement in LLMs by integrating IRL. In turn, we leverage structured domain knowledge from LLMs (Wu, 2024) to guide both the initialization and update of the IRL reward function (Ma et al., 2023; Kwon et al., 2023; Chu et al., 2023). Such guidance is essential for addressing key methodological limitations of IRL, including: (1) **Ill-posed nature** (Ng et al., 2000; Cao et al., 2021), where, in the context of travel behavior analysis, multiple latent intentions can explain the same observed travel patterns; (2) **Vast and unstructured reward space**, which hinders efficient exploration and increases the risk of converging to behaviorally implausible or non-generalizable solutions (Adams et al., 2022); and (3) **Suboptimal reward convergence**, particularly during the early training stages when lacking heuristic guidance (Wu, 2024).

In the second stage, we introduce a Cognitive Chain Reasoning (CCR) strategy to predict the final sociodemographic attributes from inferred intentions (Figure 1, dashed box denoted as ②). Leveraging LLMs' ToM capacities, This stage is designed to guide the LLM in aligning with the theoretical framework by first inferring belief constructs, and subsequently mapping them to sociodemographic attributes.. By incorporating relevant contextual factors (e.g., urbanicity, transit access) in the latter step, the model mitigates the ambiguity in LLM-based mappings and enables predictions that are both theory-aligned and context-

aware(Tjuatja et al., 2024; Petrov et al., 2024; Chen et al., 2025).

In summary, our major contributions are:

- We propose a novel methodological framework that inversely follows TPB to predict sociodemographic attributes from travel trajectories.
- We leverage LLMs to provide heuristic support for IRL reward function initialization and updates, enabling the inference of individualized and well-posed reward solutions.
- We introduce a CCR strategy that guides the LLM to make predictions in full alignment with the TPB, by explicitly modeling the mediating cognitive constructs.
- The experiments demonstrate that, compared to baselines, our model achieves substantially higher accuracy (e.g., a 30.93% improvement in gender prediction). By inferring sociodemographic attributes from trajectories, we can enrich mobile sensing datasets and support the development of more realistic, behaviorally grounded AI agents for simulation and decision-making.

2 Related Work

Sociodemographic Inference Based on Human Mobility Patterns. Recent studies leveraged machine learning methods to estimate sociodemographic attributes from mobility/activity features derived from trajectory data, often incorporating contextual signals such as semantic Points of Interest (POIs) (Zhong et al., 2015; Wu et al., 2019) or social network information (Zhong et al., 2015; Chen et al., 2024). For example, Chen et al. (2024) predicted housing prices, used as a proxy for income, by combining mobility embeddings learned via Word2Vec with social network characteristics derived from call detail records (CDRs). Others have applied models such as XGBoost (Wu et al., 2019), Support Vector Machines (SVM) (Zhang et al., 2024a), or CatBoost (Bakhtiari et al., 2023) to infer attributes like gender, age, or income status from activity-derived or mobility-derived features. Zhu et al. (2017) predicts sociodemographic attributes from mobility trajectories by modeling variability in individual mobility patterns.

LLMs for Modeling Cognitive Processes. Recent works show that LLMs can simulate human cognitive processes through Theory-of-Mind (ToM) reasoning (Gandhi et al., 2023; Li et al., 2023; Amirizaniani et al., 2024; Street, 2024). To ensure alignment with human mental states, LLMs are most effective when supported by deliberate prompting (Gu et al., 2024) and structured reasoning frameworks (Zhou et al., 2023). Additionally, LLMs have demonstrated strong potential in inferring latent cognitive constructs (Ali et al., 2024; Chen et al., 2025) and predicting sociodemographic attributes from behavioral cues (Orlikowski et al., 2025), laying the foundation for cognitively grounded modeling of human behavior.

Inverse Reinforcement Learning. The potential of IRL to uncover latent intentions and inform downstream classification tasks has been demonstrated in Hayes et al. (2011); Dadgostari et al. (2022). It has been increasingly applied to travel behavior analysis to uncover the underlying intentions. Methods include feature matching (Liu et al., 2022), maximum entropy (Koch and Dugundji, 2020; Okubo et al., 2024), and deep/adversarial IRL for high-dimensional settings (Zhao and Liang, 2023; Liu et al., 2025a). A recent work also exploreed model interpretability in mobility-focused IRL (Liang et al., 2025).

To address IRL's reward ambiguity (Ng et al., 2000), some recent studies shifted from identifying a single solution to learning feasible reward sets (Metelli et al., 2021, 2023; Zhao et al., 2023). Domain knowledge has long informed reward initialization (Liu et al., 2013), and more recently, LLMs have been used to guide both initialization and refinement (Ma et al., 2023; Kwon et al., 2023; Chu et al., 2023).

3 Methodology

Our methodological framework is grounded in the **Theory of Planned Behavior (TPB)** (see Figure 1). TPB posits that background factors, such as sociodemographic and contextual attributes, influence people's beliefs, which include attitude, subjective norm, and perceived behavioral control. Attitude refers to an individual's evaluation of the behavior, subjective norm reflects perceived social pressure from peers or influential figures, and perceived behavioral control represents one's perceived ability to perform the behavior. These beliefs collectively inform intentions, and ultimately drive observed behaviors (Ajzen, 1991, 2020).

To infer individuals' sociodemographic attributes from their observed mobility patterns, we propose to follow the TPB in reverse, and an overview of the proposed methodology is shown in Figure 2. Specifically, individual trajectories

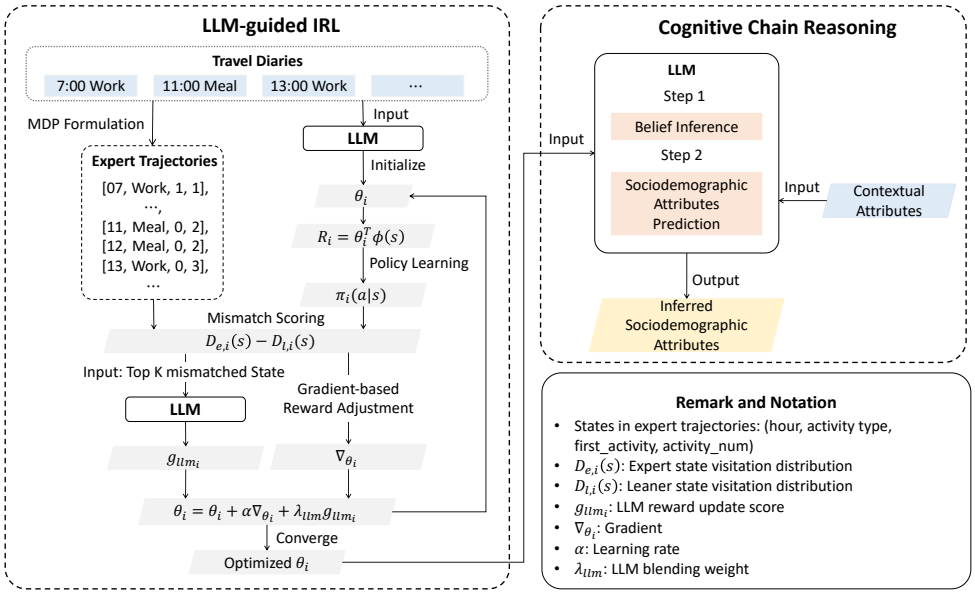

Figure 2: Methodological Framework.

are first converted into structured travel diaries that record detailed information on daily trips, such as activity types and departure times. These diaries are then represented as sequential activity patterns and modeled within a Markov Decision Process (MDP) framework. We subsequently introduce an LLM-guided inverse reinforcement learning (IRL) framework, in which the LLM provides heuristic guidance for reward initialization and updates to infer individuals' latent intentions. Building on this, we propose a Cognitive Chain Reasoning (CCR) strategy that guides the LLM to first infer belief constructs from learned intentions and subsequently predict sociodemographic attributes using both these beliefs and external contextual variables (e.g., urbanicity, home location population density).

3.1 MDP Framework

We begin by modeling travel diaries as a Markov Decision Process (MDP), where each agent represents an individual making sequential travel decisions over time. Each dairy captures daily travel behavior, including features such as activity type and departure time. To enable structured temporal reasoning, we discretize the day into 24 hourly intervals, assuming one action decision is made at the start of each hour (Liang et al., 2025). This is a reasonable assumption, as according to the National Household Travel Survey (NHTS), Americans made an average of approximately 3-4 trips per person per day (U.S. Department of Transportation, 2018).

An MDP is formally defined as $M = \{S, A, T, R, \gamma\}$, where S denotes the set of states and A the set of possible actions. The transition function $T(s, a, s\prime)$ specifies the probability of transitioning to state $s\prime \in S$ after taking action $a \in A$ in state $s \in S$. The reward function $R(s)$ assigns a scalar reward to each state s, and the discount factor $\gamma \in [0, 1]$ controls the agent's preference for immediate versus future rewards. Another key concept in the MDP framework is the policy $\pi(a|s)$, which defines the probability of an individual selecting action a when in state s.

The five components of the MDP are specified as follows in the context of mobility modeling:

State Space S: Each state $s \in S$ is represented as a feature vector $s = [h, a, f, n]$, where $h \in \{0, \cdots, 23\}$ denotes the hour of the day, $a \in \mathcal{A}$ indicates the activity type (i.e., home, work, education, escort and errand, leisure (Akar et al., 2012)), $f \in \{0, 1\}$ is a binary indicator of whether the activity is the first of the day, and $n \in R \geq 0$ captures the cumulative number of activities up to the current activity.

Action Space A: We define $A = \{\textbf{stay}, \textbf{travel}\}$, where **stay** represents the decision to continue the current activity, and **travel** indicates a transition to a new activity.

Transition Probability Function T: When $a =$ **stay**, the transition is deterministic: the hour increments by one while all other state components remain unchanged, resulting in $s\prime = [h+1, a, f, n]$. In contrast, when $a =$ **travel**, the transition is

stochastic and modeled using empirical transition frequencies derived from the individual's observed activity sequences. In this case, the hour also advanced by one ($h\prime = h + 1$); the first-activity indicator is updated such that $f\prime = 0$ if $f = 1$, and remains unchanged otherwise; and the cumulative activity count increases by one ($n\prime = n + 1$) as the activity changes.

Reward Function R: The reward function $R(s)$ is defined as a linear function of the state features. Specifically, the reward assigned to a state s is computed as:

$$R(s) = \boldsymbol{\theta}^\top \boldsymbol{\phi}(s) \tag{1}$$

where $\boldsymbol{\phi}(s)$ denotes the concatenated feature vector comprising the one-hot hour indicators, one-hot activity type indicators, the binary first-activity indicator f, and the numerical feature n. The vector $\boldsymbol{\theta}$ contains the corresponding reward weight associated with each feature dimension.

3.2 Intention Inference with LLM-guided IRL

Once the MDP is formulated, IRL aims to infer individuals' underlying intentions from observed mobility patterns. Specifically, it estimates reward weights such that the resulting policy produces trajectories, referred to as learner trajectories, that closely replicate the expert trajectories, i.e., the individuals' observed state sequences.

We introduce an IRL framework in which LLMs provide guidance for both the initialization and iterative update of reward weights.

Reward Initialization: Inspired by Ma et al. (2023), we use an LLM to initialize individual-specific reward weights given each individual's multi-day travel diaries. The prompt template is provided in Figure 5 (Appendix B). In addition to interpreting the input diaries, the LLM leverages prior knowledge of human mobility behavior (Liu et al., 2024) and inverse alignment capabilities (Sun and van der Schaar, 2024) to generate behaviorally meaningful and unique initializations. This mitigates the ill-posed nature of IRL (Ng et al., 2000; Cao et al., 2021), where multiple reward functions can explain the same observed behavior, and helps constrain the vast reward space (Adams et al., 2022) toward plausible solutions.

Policy Learning: Once the reward weights are initialized or updated, they are used to infer each individual's policy $\pi_i(a|s)$. To account for the inherent variability and bounded rationality in human decision-making, we adopt a stochastic policy framework rather than assuming deterministic action selection. Individuals' behavior is modeled using the maximum entropy principle (Ziebart et al., 2008), wherein actions with higher expected value are assigned higher probabilities. Specifically, the value of a state s is iteratively updated according to:

$$V_i(s) = R_i(s) + log \sum_{a \in \mathcal{A}} \exp(\gamma \sum_{s\prime} T_i(s\prime|s,a) V_i(s\prime)) \tag{2}$$

The updates proceed iteratively until the change in $V_i(s)$ across iterations falls below a predefined threshold ϵ. Upon convergence of $V_i(s)$, the corresponding stochastic policy $\pi_i(a|s)$ is computed as:

$$\pi_i(a|s) = \frac{\exp(R_i(s) + \gamma \sum_{s\prime} T_i(s\prime|s,a) V_i(s\prime))}{\sum_{a\prime} \exp(R_i(s) + \gamma \sum_{s\prime} T_i(s\prime|s,a\prime) V_i(s\prime))} \tag{3}$$

Mismatch Scoring: To iteratively refine reward weights, we compute a mismatch score as the discrepancy between learner and expert state visitation distributions. The expert state visitation distribution for individual i, denoted as $D_{e,i}(s)$, is estimated by counting the frequency of state visits across the individual's trajectories and normalizing the counts to obtain a probability distribution. The corresponding learner state visitation distribution, $D_{l,i}(s)$ is computed by simulating expected state occupancies under the learned policy $\pi_i(a|s)$ over a planning horizon of $\mathcal{T}(\mathcal{T} = 24)$ steps. The initial state visitation distribution is defined by the empirical distribution of starting activities from expert trajectories. The learner's state visitation distribution is then iteratively updated by propagating state occupancies according to the policy $\pi_i(a|s)$ and the transition dynamics $T_i(s\prime|s,a)$. At each time step t, the visitation distribution is updated as follows:

$$D_{l,i,t+1}(s\prime) = \sum_s \sum_a D_{l,i,t}(s) \cdot \pi_i(a|s) \cdot T_i(s\prime|s,a) \tag{4}$$

The top K states exhibiting the largest absolute mismatch score $|D_{e,i}(s) - D_{l,i}(s)|$ between expert and learner visitation frequencies are identified to use as inputs for the LLM-guided reward update procedure within the iterative learning process.

LLM-guided Reward Updates: In the early stages of training, the convergence of reward weights may be suboptimal, particularly given the bias introduced by traditional Maximum Entropy gradient ascent methods when applied to sparse

expert trajectories, such as multi-day activity sequences. To address this issue, an LLM is incorporated to provide heuristic reward weight adjustments during updates. Leveraging the LLM's prior knowledge of activity behavior patterns further mitigates the challenges posed by data sparsity and improves the robustness of the learning process. Specifically, the reward weight vector θ_i for individual i is iteratively updated according to:

$$\nabla_{\theta_i} = \frac{1}{n_i}\sum_{j=1}^{n_i}\sum_{s\in\tau_{i,j}}\phi(s) - \frac{1}{\mathcal{T}}\sum_{t=0}^{\mathcal{T}-1}\sum_{s} D_{l,i,t}(s)\cdot\phi(s) \quad (5)$$

$$\Delta\theta_i = \alpha\cdot\nabla_{\theta_i} + \lambda_{\text{LLM}}\cdot \mathbf{g}_{\text{LLM}_i} \quad (6)$$

$$\theta_i = \theta_i + \Delta\theta_i \quad (7)$$

Here, ∇_{θ_i} denotes the Maximum Entropy IRL gradient, n_i is the number of expert trajectories for individual i, and $\tau_{i,j}$ represents the j-th expert trajectory of individual i. The parameter α is the learning rate, λ_{LLM} is a tunable blending weight, and $\mathbf{g}_{\text{LLM}}$ denotes the LLM-based update direction.

For LLM-guided reward updates, the top-K states with the highest absolute mismatch scores are provided as input to the LLM. The LLM then suggests an update direction for each of the reward weights. The LLM's output for each weight is restricted to one of three discrete values: -1, 0, 1, corresponding to a decrease, no change, or an increase. The complete prompt template used for this process is presented in Figure 6 in Appendix B. The updates continue iteratively until convergence, which is determined by either the change in θ_i or the KL divergence between the expert and learner state visitation distributions dropping below a predefined threshold ϵ'. The KL divergence formulation is detailed in Appendix A.5.

3.3 Sociodemographic Attribute Prediction with Cognitive Chain Reasoning

After inferring individual-specific reward weights that represent latent intentions, we leverage these to predict individual-level sociodemographic attributes. Grounded in the TPB, sociodemographic and contextual attributes act as background variables that indirectly shape intentions through their influence on beliefs. Accordingly, we first infer latent beliefs from the learned reward weights and then predict sociodemographic attributes by integrating these beliefs with external contextual features.

To implement this approach, we design a CCR prompting template that guides the LLM to sequentially reason through three beliefs before making a prediction. This structure aligns the model's reasoning process with the TPB to enhance its theoretical consistency. Furthermore, the CCR framework capitalizes on the LLM's strengths in multi-step reasoning capabilities that are particularly important for mapping abstract reward weights into psychologically meaningful beliefs and, ultimately, individuals' sociodemographic profiles. Contextual attributes are also incorporated to aid sociodemographic prediction following belief inference, as they jointly influence belief formation alongside sociodemographic attributes. This accounts for the fact that similar belief patterns may correspond to different sociodemographic groups under varying contextual conditions. The CCR prompting template is provided in Figure 7 (Appendix B).

4 Experiment

4.1 Experiment Setup

Implementation Details. Please find the implementation details in Appendix A.1.

Baselines. We evaluate our proposed model against a set of baseline models for predicting sociodemographic attributes from mobility data, including SVM, XGBoost, CatBoost, and GPT-4o. The first three models are selected based on their demonstrated effectiveness in prior work (Zhu et al., 2017; Wu et al., 2019; Bakhtiari et al., 2023). In addition, we include GPT-4o as a baseline to explore the emerging capability of LLMs to infer sociodemographic attributes directly from mobility patterns in a zero-shot setting.

We evaluate the models using either extracted mobility features or direct travel diary data, both combined with contextual attributes. Specifically, extracted features, selected based on prior literature, capture key aspects of individual mobility behavior, such as averages and variations in departure times, activity types, trip distances, travel times, and trip frequencies. The extracted features are detailed in Table 5 in Appendix C.1. To reduce overfitting and ensure fair evaluation, feature selection is performed on the training set (Appendix A.3). The same set of selected features is also provided as input to GPT-4o to infer sociodemographic attributes from the test set. Direct travel diary data, containing multiple-day trip records, along with the contextual data, are converted into semanti-

cally descriptive inputs for GPT-4o. This setup is designed to evaluate GPT-4o's ability to directly infer sociodemographic attributes from mobility data without relying on feature engineering.

Datasets. To obtain both individuals' travel trajectories and their sociodemographic profiles, we use the 2017 Puget Sound Regional Council Household Travel Survey (National Renewable Energy Laboratory, 2025). It comprises data from 6,254 participants, covering a total of 52,492 trips. It includes two mobility components: a one-day household travel diary and a seven-day smartphone-based GPS diary. In our study, we only used the GPS diaries for analysis as they are derived from passive collected multi-day trajectories that better reflect naturalistic behavior. The GPS diaries provide detailed individual-level trip and activity information, such as travel day, departure and arrival times, activity type, and trip distance, which can be readily inferred from any raw GPS trajectory data.

To ensure data quality and analytical relevance, we applied a series of filtering criteria to the GPS diary data (Appendix A.4). After applying the filtering criteria, the final dataset consists of 617 qualified individuals, contributing a total of 11,964 trips for analysis. Additionally, to support contextualized reasoning in prediction, we incorporated contextual variables as auxiliary inputs. Details of contextual attributes are provided Appendix C.2.

Evaluation Metrics. We use KL divergence and L1 distance to quantify the discrepancy between expert and learner policies. To evaluate prediction performance, we reporte class-level Precision, Recall, and F1 score, as well as overall Accuracy and Weighted F1 score. Metric definitions are in Appendix A.5 and A.6.

4.2 Main Results

We evaluate our model on gender and age prediction (see Appendix D for income and employment), comparing it with baselines to assess the effectiveness of incorporating psychological theory and LLM prior knowledge without relying on additional training data.

As shown in Table 1, GPT-4o slightly outperforms traditional machine learning models using the same input features, demonstrating its ability to distinguish behavioral patterns through prior domain knowledge. Moreover, its comparable performance with direct travel diaries suggests that it can operate effectively without relying on explicit feature engineering. Our proposed model further advances GPT-4o's inference from observed mobility data, yielding over a 30% improvement in overall accuracy. It also achieves more balanced performance across classes, as evidenced by higher weighted and class-specific F1 scores.

Table 1: Model comparison for gender prediction. The proposed model outperforms all baselines in both overall and class-level performance.

Method	Class	Precision	Recall	F1-score	Accuracy	Weighted F1
SVM	Male	0.603	0.633	0.618	0.621	0.621
	Female	0.639	0.609	0.624		
XGBoost	Male	0.597	0.667	0.630	0.621	0.620
	Female	0.649	0.578	0.612		
CatBoost	Male	0.683	0.642	0.662	0.645	0.646
	Female	0.607	0.649	0.627		
GPT-4o	Male	0.627	0.700	0.661	0.653	0.653
	Female	0.684	0.609	0.645		
GPT-4o (Diaries)	Male	0.625	0.667	0.645	0.645	0.645
	Female	0.667	0.625	0.645		
SILIC	Male	0.920	0.767	0.836	**0.855**	**0.853**
	Female	0.811	0.938	0.870		

Table 2: Model comparison for age prediction. The proposed model outperforms all baselines in both overall and class-level performance, and effectively identifies age groups that baseline models struggle to classify.

Method	Class	Precision	Recall	F1-score	Accuracy	Weighted F1
SVM	18–44	0.752	0.989	0.854	0.742	0.644
	45–64	0.000	0.000	0.000		
	65+	0.333	0.200	0.250		
XGBoost	18–44	0.767	0.859	0.810	0.677	0.650
	45–64	0.250	0.148	0.186		
	65+	0.200	0.200	0.200		
CatBoost	18–44	0.792	0.870	0.829	0.718	0.700
	45–64	0.421	0.296	0.348		
	65+	0.250	0.200	0.222		
GPT-4o	18–44	0.870	0.870	0.870	0.710	0.732
	45–64	0.615	0.296	0.400		
	65+	0.000	0.000	0.000		
GPT-4o (Diaries)	18–44	0.863	0.891	0.877	0.734	0.740
	45–64	0.571	0.296	0.390		
	65+	0.067	0.200	0.100		
SILIC	18–44	0.887	0.989	0.935	**0.863**	**0.844**
	45–64	0.900	0.375	0.529		
	65+	0.500	0.800	0.615		

Table 2 presents the age prediction results for a multi-class task. While baseline models achieve decent overall accuracy, they struggle to identify individuals over the age of 45, particularly those over 65, as evidenced by their low class-level F1

scores. For machine learning models, this performance gap may be attributed to data imbalance, as individuals aged 45 and above, particularly those over 65, are underrepresented in the training set. Consequently, they are biased toward the majority class, resulting in diminished predictive performance for those age groups. The GPT 4o inference model, despite its general reasoning capabilities, may have difficulty distinguishing age-specific behavioral patterns in mobility without access to the underlying cognitive patterns. In contrast, our proposed model achieves superior performance compared to all baseline methods, with significantly higher overall prediction accuracy and weighted F1 score, while also demonstrating improved ability to distinguish adults across different age groups. This improvement is particularly important for downstream applications that rely on accurately identifying targeted sociodemographic groups.

Overall, the proposed model consistently achieves strong performance in both total prediction accuracy and class-level identification across all sociodemographic attributes. These results highlight the effectiveness of inversely following the well-established behavioral theory by capturing mediating cognitive processes for inferring sociodemographic characteristics from mobility patterns.

4.3 Ablation Study

We conducted ablation experiments using the GPT-4o model to evaluate the effectiveness of (1) LLM-guided reward weight initialization and updates and (2) the CCR strategy.

First, we compare the KL divergence and L1 distance (definitions in Appendix A.5) between expert and learner behaviors under different settings to evaluate whether the policy derived from the final reward weights effectively replicates expert trajectories. As shown in Table 3, removing both LLM-guided initialization and updates results in over a 41% increase in KL divergence and over a 10% increase in L1 distance, confirming their contribution. When only the LLM-guided updates are replaced with standard gradient ascent, the increase is less pronounced, highlighting the importance of LLM-guided initialization in constraining the reward space effectively. Furthermore, retaining only the LLM-guided reward weight updates still achieves better performance than removing both the initialization and update steps, indicating the effectiveness of LLM-provided heuristics in guiding optimization toward behaviorally plausible reward structures.

Table 3: Ablation Study of LLM-guided IRL. It demonstrates the effectiveness of both LLM-based reward initialization and iterative update guidance.

	KL Divergence	L1 Distance
Random Initialization + Gradient Ascent	0.594	0.722
LLM-guided Initialization + Gradient Ascent	0.447	0.696
Random Initialization + LLM-guided updates	0.543	0.704
LLM-guided Initialization + updates	**0.419**	**0.654**

Second, we compare the Accuracy and Weighted F1 of our proposed model against two variants that substitute the CCR module with either pure GPT-4o inference or Chain-of-Thought (CoT) inference, in order to assess the effectiveness of the CCR module. As shown in Table 4, the CoT strategy outperforms pure inference by leveraging the structured reasoning capabilities of the LLM. By further guiding the reasoning process with the well-established behavioral theory, the CCR module aligns more closely with human cognitive processes, leading to the best performance among all compared methods.

Table 4: Ablation study of CCR on gender and age prediction. CCR outperforms both pure inference and CoT.

Method	Gender		Age	
	Accuracy	**Weighted F1**	**Accuracy**	**Weighted F1**
IRL + Inference	0.718	0.708	0.815	0.797
IRL + CoT	0.766	0.766	0.823	0.812
IRL + CCR	**0.855**	**0.853**	**0.863**	**0.844**

5 Conclusion

This study proposes *SILIC*, a novel IRL+CCR framework to predict sociodemographic attributes from mobility patterns by inversely aligning with the TPB. We leverage IRL to infer individual-specific latent intentions, while addressing its key methodological challenges through heuristic guidance provided by LLMs. Building on inferred intentions, the CCR module guides LLM reasoning to sequentially infer belief constructs and predict sociodemographic attributes. Contextual features are incorporated to support informed predictions, as sociodemographic and contextual factors jointly influence behavioral beliefs. Extensive experiments across multiple prediction tasks demonstrate that our framework not only significantly outperforms established baselines, but also offers a solution for enriching large-scale real-world or synthetic trajectory datasets, with implications for various applications involving human behavior modeling.

Limitations

Although the proposed model shows strong potential for inferring sociodemographic attributes from trajectories, several limitations remain. First, while the model infers latent intentions and belief constructs from observed behavior, these internal variables lack ground truth for direct validation. Any misalignment in these inferred representations may propagate to downstream sociodemographic predictions. Second, the IRL component depends on heuristic support from LLMs, whose guidance—partially derived from general domain knowledge—may fail to fully capture behavioral nuances specific to certain geographic regions. Third, the current state space design may omit relevant mobility-related features (e.g., trip distance), limiting its capacity to fully represent human decision-making dynamics. Future work should address these challenges. For example, leveraging survey data that include cognitive or attitudinal variables to assess the validity of inferred intentions and beliefs, and incorporating human-in-the-loop calibration to refine LLM-generated reward priors or updates to better reflect regional characteristics.

References

Stephen Adams, Tyler Cody, and Peter A Beling. 2022. A survey of inverse reinforcement learning. *Artificial Intelligence Review*, 55(6):4307–4346.

I Ajzen. 1991. The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2):179–211.

Icek Ajzen. 2020. The theory of planned behavior: Frequently asked questions. *Human behavior and emerging technologies*, 2(4):314–324.

Gulsah Akar, Kelly J Clifton, and Sean T Doherty. 2012. Redefining activity types: Who participates in which leisure activity? *Transportation research part A: policy and practice*, 46(8):1194–1204.

Hassan Ali, Philipp Allgeuer, and Stefan Wermter. 2024. Comparing apples to oranges: Llm-powered multimodal intention prediction in an object categorization task. In *International Conference on Social Robotics*, pages 292–306. Springer.

Maryam Amirizaniani, Elias Martin, Maryna Sivachenko, Afra Mashhadi, and Chirag Shah. 2024. Can llms reason like humans? assessing theory of mind reasoning in llms for open-ended questions. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*, pages 34–44.

Ali Bakhtiari, Hamid Mirzahossein, Navid Kalantari, and Xia Jin. 2023. Inferring socioeconomic characteristics from travel patterns. *Journal of Regional and City Planning*, 34(1):122–136.

Dibyanayan Bandyopadhyay, Soham Bhattacharjee, and Asif Ekbal. 2025. Thinking machines: A survey of llm based reasoning strategies. *arXiv preprint arXiv:2503.10814*.

Haoyang Cao, Samuel Cohen, and Lukasz Szpruch. 2021. Identifiability in inverse reinforcement learning. *Advances in Neural Information Processing Systems*, 34:12362–12373.

Ruxiao Chen, Chenguang Wang, Yuran Sun, Xilei Zhao, and Susu Xu. 2025. From perceptions to decisions: Wildfire evacuation decision prediction with behavioral theory-informed llms. *arXiv preprint arXiv:2502.17701*.

Xiao Chen, Tao Pei, Ci Song, Hua Shu, Sihui Guo, Xi Wang, Yaxi Liu, and Jie Chen. 2024. Coupling human mobility and social relationships to predict individual socioeconomic status: A graph neural network approach. *Transactions in GIS*, 28(5):1412–1438.

Kun Chu, Xufeng Zhao, Cornelius Weber, Mengdi Li, and Stefan Wermter. 2023. Accelerating reinforcement learning of robotic manipulations via feedback from large language models. *arXiv preprint arXiv:2311.02379*.

Faraz Dadgostari, Samarth Swarup, Stephen Adams, Peter Beling, and Henning S Mortveit. 2022. Identifying correlates of emergent behaviors in agent-based simulation models using inverse reinforcement learning. In *2022 Winter Simulation Conference (WSC)*, pages 322–333. IEEE.

Kanishk Gandhi, Jan-Philipp Fränken, Tobias Gerstenberg, and Noah Goodman. 2023. Understanding social reasoning in language models with language models. *Advances in Neural Information Processing Systems*, 36:13518–13529.

Anindya Ghose, Beibei Li, and Siyuan Liu. 2019. Mobile targeting using customer trajectory patterns. *Management Science*, 65(11):5027–5049.

Yuling Gu, Oyvind Tafjord, Hyunwoo Kim, Jared Moore, Ronan Le Bras, Peter Clark, and Yejin Choi. 2024. Simpletom: Exposing the gap between explicit tom inference and implicit tom application in llms. *arXiv preprint arXiv:2410.13648*.

Masahiko Haraguchi, Akihiko Nishino, Akira Kodaka, Maura Allaire, Upmanu Lall, Liao Kuei-Hsien, Kaya Onda, Kota Tsubouchi, and Naohiko Kohtake. 2022. Human mobility data and analysis for urban resilience: A systematic review. *Environment and Planning B: Urban Analytics and City Science*, 49(5):1507–1535.

Roy Hayes, Jonathan Bao, Peter Beling, and Barry Horowitz. 2011. Use of inverse reinforcement learning for identity prediction. In *Selected Papers and Presentations Presented at MODSIM World 2010 Conference Expo*.

Christian Holz-Rau and Joachim Scheiner. 2019. Land-use and transport planning–a field of complex cause-impact relationships. thoughts on transport growth, greenhouse gas emissions and the built environment. *Transport Policy*, 74:127–137.

Jong Wook Kim and Beakcheol Jang. 2024. Privacy-preserving generation and publication of synthetic trajectory microdata: A comprehensive survey. *Journal of Network and Computer Applications*, page 103951.

Thomas Koch and Elenna Dugundji. 2020. A review of methods to model route choice behavior of bicyclists: inverse reinforcement learning in spatial context and recursive logit. In *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on GeoSpatial Simulation*, pages 30–37.

Minae Kwon, Sang Michael Xie, Kalesha Bullard, and Dorsa Sadigh. 2023. Reward design with language models. *arXiv preprint arXiv:2303.00001*.

Seungpil Lee, Woochang Sim, Donghyeon Shin, Wongyu Seo, Jiwon Park, Seokki Lee, Sanha Hwang, Sejin Kim, and Sundong Kim. 2024. Reasoning abilities of large language models: In-depth analysis on the abstraction and reasoning corpus. *ACM Transactions on Intelligent Systems and Technology*.

Dmitry Leykin, Limor Aharonson-Daniel, and Mooli Lahad. 2016. Leveraging social computing for personalized crisis communication using social media. *PLoS currents*, 8:ecurrents–dis.

Huao Li, Yu Quan Chong, Simon Stepputtis, Joseph Campbell, Dana Hughes, Michael Lewis, and Katia Sycara. 2023. Theory of mind for multi-agent collaboration via large language models. *arXiv preprint arXiv:2310.10701*.

Yuebing Liang, Shenhao Wang, Jiangbo Yu, Zhan Zhao, Jinhua Zhao, and Sandy Pentland. 2025. Analyzing sequential activity and travel decisions with interpretable deep inverse reinforcement learning. *arXiv preprint arXiv:2503.12761*.

Shan Liu, Ya Zhang, Zhengli Wang, Xiang Liu, and Hai Yang. 2025a. Personalized origin–destination travel time estimation with active adversarial inverse reinforcement learning and transformer. *Transportation Research Part E: Logistics and Transportation Review*, 193:103839.

Siyuan Liu, Miguel Araujo, Emma Brunskill, Rosaldo Rossetti, Joao Barros, and Ramayya Krishnan. 2013. Understanding sequential decisions via inverse reinforcement learning. In *2013 IEEE 14th International Conference on Mobile Data Management*, volume 1, pages 177–186. IEEE.

Yepeng Liu, Xuandong Zhao, Dawn Song, and Yuheng Bu. 2025b. Dataset protection via watermarked canaries in retrieval-augmented llms. *arXiv preprint arXiv:2502.10673*.

Yifan Liu, Xishun Liao, Haoxuan Ma, Brian Yueshuai He, Chris Stanford, and Jiaqi Ma. 2024. Human mobility modeling with limited information via large language models. *arXiv preprint arXiv:2409.17495*.

Yiru Liu, Yudi Li, Guoyang Qin, Ye Tian, and Jian Sun. 2022. Understanding the behavioral effect of incentives on departure time choice using inverse reinforcement learning. *Travel Behaviour and Society*, 29:113–124.

Massimiliano Luca, Gianni Barlacchi, Bruno Lepri, and Luca Pappalardo. 2021. A survey on deep learning for human mobility. *ACM Computing Surveys (CSUR)*, 55(1):1–44.

Yecheng Jason Ma, William Liang, Guanzhi Wang, De-An Huang, Osbert Bastani, Dinesh Jayaraman, Yuke Zhu, Linxi Fan, and Anima Anandkumar. 2023. Eureka: Human-level reward design via coding large language models. *arXiv preprint arXiv:2310.12931*.

Alberto Maria Metelli, Filippo Lazzati, and Marcello Restelli. 2023. Towards theoretical understanding of inverse reinforcement learning. In *International Conference on Machine Learning*, pages 24555–24591. PMLR.

Alberto Maria Metelli, Giorgia Ramponi, Alessandro Concetti, and Marcello Restelli. 2021. Provably efficient learning of transferable rewards. In *International Conference on Machine Learning*, pages 7665–7676. PMLR.

National Renewable Energy Laboratory. 2025. Transportation secure data center. https://www.nrel.gov/tsdc. Accessed May 13, 2025.

Andrew Y Ng, Stuart Russell, and 1 others. 2000. Algorithms for inverse reinforcement learning. In *Icml*, volume 1, page 2.

Thuy Ngoc Nguyen, Kasturi Jamale, and Cleotilde Gonzalez. 2024. Predicting and understanding human action decisions: Insights from large language models and cognitive instance-based learning. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, volume 12, pages 126–136.

Tomohiro Okubo, Akihiro Kobayashi, Daisuke Kamisaka, and Akinori Morimoto. 2024. Estimation of route-choice behavior along lrt lines using inverse reinforcement learning. *Inventions*, 9(6):118.

Matthias Orlikowski, Jiaxin Pei, Paul Röttger, Philipp Cimiano, David Jurgens, and Dirk Hovy. 2025. Beyond demographics: Fine-tuning large language models to predict individuals' subjective text perceptions. *arXiv preprint arXiv:2502.20897*.

Nikolay B Petrov, Gregory Serapio-García, and Jason Rentfrow. 2024. Limited ability of llms to simulate human psychological behaviours: a psychometric analysis. *arXiv preprint arXiv:2405.07248*.

Shaikh Shakeela, N Sai Shankar, P Mohan Reddy, T Kavya Tulasi, and M Mahesh Koneru. 2021. Optimal ensemble learning based on distinctive feature selection by univariate anova-f statistics for ids. *International Journal of Electronics and Telecommunications*, pages 267–275.

Winnie Street. 2024. Llm theory of mind and alignment: Opportunities and risks. *arXiv preprint arXiv:2405.08154*.

Hao Sun, Thomas Pouplin, Nicolás Astorga, Tennison Liu, and Mihaela van der Schaar. Improving llm generation with inverse and forward alignment: Reward modeling, prompting, fine-tuning, and inference-time optimization. In *The First Workshop on System-2 Reasoning at Scale, NeurIPS'24*.

Hao Sun and Mihaela van der Schaar. 2024. Inverse-rlignment: Inverse reinforcement learning from demonstrations for llm alignment. *arXiv preprint arXiv:2405.15624*.

Lindia Tjuatja, Valerie Chen, Tongshuang Wu, Ameet Talwalkwar, and Graham Neubig. 2024. Do llms exhibit human-like response biases? a case study in survey design. *Transactions of the Association for Computational Linguistics*, 12:1011–1026.

U.S. Census Bureau. 2017. TIGER/Line Shapefile, 2017, 2010 nation, U.S., 2010 Census Urban Area National. Accessed: 2025-05-13.

U.S. Census Bureau. 2025. U.s. census bureau. https://www.census.gov/. Accessed: 2025-05-13.

Federal Highway Administration U.S. Department of Transportation. 2018. 2017 national household travel survey: U.s. travel profile. Technical report, Federal Highway Administration. Accessed: 2025-05-18.

U.S. Environmental Protection Agency. 2025. Smart location mapping. https://www.epa.gov/smartgrowth/smart-location-mapping. Accessed: 2025-05-13.

Washington State Geospatial Open Data Portal. 2025. Washington state geospatial open data portal. Accessed: 2025-05-13.

Lun Wu, Liu Yang, Zhou Huang, Yaoli Wang, Yanwei Chai, Xia Peng, and Yu Liu. 2019. Inferring demographics from human trajectories and geographical context. *Computers, Environment and Urban Systems*, 77:101368.

Xiefeng Wu. 2024. From reward shaping to q-shaping: Achieving unbiased learning with llm-guided knowledge. *arXiv preprint arXiv:2410.01458*.

Takahiro Yabe, Kota Tsubouchi, Akihito Sudo, and Yoshihide Sekimoto. 2016. Estimating evacuation hotspots using gps data: What happened after the large earthquakes in kumamoto, japan. In *Proc. of the 5th International Workshop on Urban Computing*, volume 81, pages 1–5.

Bin Zhang, Soora Rasouli, and Tao Feng. 2024a. Social demographics imputation based on similarity in multi-dimensional activity-travel pattern: A two-step approach. *Travel Behaviour and Society*, 37:100843.

Yadong Zhang, Shaoguang Mao, Tao Ge, Xun Wang, Adrian de Wynter, Yan Xia, Wenshan Wu, Ting Song, Man Lan, and Furu Wei. 2024b. Llm as a mastermind: A survey of strategic reasoning with large language models. *arXiv preprint arXiv:2404.01230*.

Lei Zhao, Mengdi Wang, and Yu Bai. 2023. Is inverse reinforcement learning harder than standard reinforcement learning? a theoretical perspective. *arXiv preprint arXiv:2312.00054*.

Xilei Zhao, Yiming Xu, Ruggiero Lovreglio, Erica Kuligowski, Daniel Nilsson, Thomas J Cova, Alex Wu, and Xiang Yan. 2022. Estimating wildfire evacuation decision and departure timing using large-scale gps data. *Transportation research part D: transport and environment*, 107:103277.

Zhan Zhao and Yuebing Liang. 2023. A deep inverse reinforcement learning approach to route choice modeling with context-dependent rewards. *Transportation Research Part C: Emerging Technologies*, 149:104079.

Yuan Zhong, Nicholas Jing Yuan, Wen Zhong, Fuzheng Zhang, and Xing Xie. 2015. You are where you go: Inferring demographic attributes from location check-ins. In *Proceedings of the eighth ACM international conference on web search and data mining*, pages 295–304.

Pei Zhou, Aman Madaan, Srividya Pranavi Potharaju, Aditya Gupta, Kevin R McKee, Ari Holtzman, Jay Pujara, Xiang Ren, Swaroop Mishra, Aida Nematzadeh, and 1 others. 2023. How far are large language models from agents with theory-of-mind? *arXiv preprint arXiv:2310.03051*.

Lei Zhu, Jeffrey Gonder, and Lei Lin. 2017. Prediction of individual social-demographic role based on travel behavior variability using long-term gps data. *Journal of Advanced Transportation*, 2017(1):7290248.

Yuanshao Zhu, Yongchao Ye, Ying Wu, Xiangyu Zhao, and James Yu. 2023. Synmob: Creating high-fidelity synthetic gps trajectory dataset for urban mobility analysis. *Advances in Neural Information Processing Systems*, 36:22961–22977.

Brian D Ziebart, Andrew L Maas, J Andrew Bagnell, Anind K Dey, and 1 others. 2008. Maximum entropy inverse reinforcement learning. In *Aaai*, volume 8, pages 1433–1438. Chicago, IL, USA.

A Implementation Details

A.1 Implementation Details

We implement our full framework using GPT-4o as the backbone for both the LLM-guided IRL module and the CCR module. The data is split into 80% training and 20% testing. Evaluation is performed on the held-out test set. We set the hyperparameters as follows: $K = 30$, $\alpha = 2$, $\lambda_{\text{LLM}} = 0.002$, and convergence thresholds $\epsilon = \epsilon' = 10^{-4}$.

A.2 Activity Mapping

To reduce the state space, we group activities into five high-level categories according to the following mapping scheme:

- **Home**: Home
- **Work**: Work, Work-related
- **Education:** School
- **Escort and Errand:** Personal Business / Errand / Appointment, Escort, Change Model
- **Leisure:** Social / Recreational, Shopping, Meal, Other

A.3 Feature Selection

We compute ANOVA F-scores (Shakeela et al., 2021) to assess the statistical association between each feature and the target variable in the training set, rank the features accordingly, and retain only those within the top 40%, corresponding to scores above the 60th percentile threshold. This selection aims to optimize the performance of baseline machine learning models, ensuring a fair comparison with our proposed approach. This filtering process is applied to the combined set of extracted and contextual attributes. Figures 3 and 4 present line plots illustrating how F1 scores vary with different feature selection thresholds across the three baseline machine learning models for gender and age prediction, respectively, using five-fold cross-validation.

A.4 Data Selection

We applied data selection procedures to ensure both the quality and relevance of the dataset for our analysis. First, we retained only the household representatives (defined as the primary survey respondent (National Renewable Energy Laboratory, 2025)), as they typically provide the most complete and reliable information within each household. Next, we excluded individuals who did not complete the survey to ensure the completeness of sociodemographic labels and mobility records. We focused on weekday trips to capture routine mobility behavior, which is more informative for modeling stable mobility patterns. Finally, we retained only individuals with at least two days of GPS trajectory data to ensure sufficient behavioral observations for effective IRL model training.

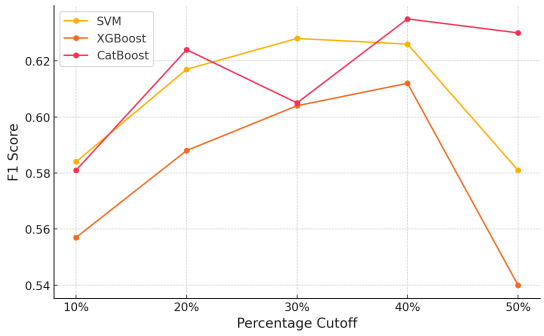

Figure 3: F1 score variation across feature selection thresholds for gender prediction using SVM, XGBoost, and CatBoost.

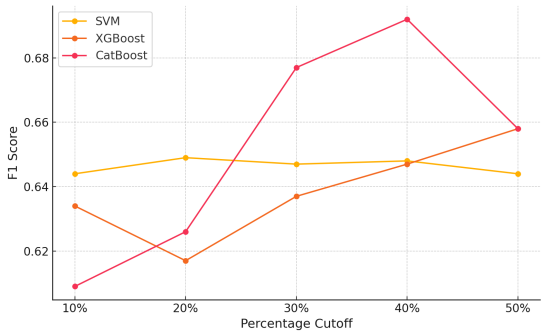

Figure 4: F1 score variation across feature selection thresholds for age prediction using SVM, XGBoost, and CatBoost.

A.5 Evaluation Metrics for IRL

To quantify the discrepancy between expert and learner policies, we adopt both KL divergence and L1 distance. For each individual i, KL divergence captures how much the learned policy $\pi_{l,i}$ deviates from the expert policy $\pi_{e,i}$, weighted by the empirical state visitation distribution. L1 distance computes the average absolute difference between $\pi_{l,i}$ and $\pi_{e,i}$ across states. The formal definitions of KL divergence and L1 distance are presented as follows:

$$KL(\pi_{e,i}||\pi_{l,i}) = \sum_s D_{e,i}(s) \sum_a \pi_{e,i}(a|s) \log(\frac{\pi_{e,i}(a|s)}{\pi_{l,i}(a|s)}) \tag{8}$$

$$L1(\pi_{e,i}, \pi_{l,i}) = \frac{1}{|S|} \sum_{s \in S} \sum_a |\pi_{e,i}(a|s) - \pi_{l,i}(a|s)| \tag{9}$$

A.6 Evaluation Metrics for Classification Tasks

To assess the effectiveness of the proposed model in predicting sociodemographic attributes from observed mobility patterns, we evaluate its performance against ground truth labels using standard classification metrics: class-level **Precision**, **Recall** and **F1-score**, and total-level **Accuracy** and **Weighted F1-score**.

Specifically, for each class $c \in C$, let TP_c, TN_c, FP_c, and FN_c represent the number of true positives, true negatives, false positives, and false negatives, respectively. Let N_c denote the total number of instances belonging to class c (i.e., the support), and let N denote the total number of instances across all classes. Based on these definitions, the evaluation metrics are defined as follows:

$$Precision_c = \frac{TP_c}{TP_c + FP_c} \tag{10}$$

$$Recall_c = \frac{TP_c}{TP_c + FN_c} \tag{11}$$

$$F1_c = \frac{2 \cdot Precision_c \cdot Recall_c}{Prevision_c + Recall_c} \tag{12}$$

$$Accuracy = \frac{\sum_{c \in C} TP_c + TN_c}{\sum_{c \in C} TP_c + TN_c + FP_c + FN_c} \tag{13}$$

$$Weighted\ F1 = \sum_{c \in C} \frac{N_c}{N} \cdot F1_c \tag{14}$$

B Prompt Design

Prompt Template for Reward Initialization

Task Description
You are an expert in travel behavior modeling and inverse reinforcement learning (IRL). You are provided with an individual's multi-day weekday travel diaries, presented as a chronological sequence of activities and their corresponding departure times.

Input:

- **Travel Diaries:** *[Travel Diaries]*
- **State Features:** $\boldsymbol{\phi}(s)$, a feature vector encoding relevant state attributes.

Objective
Estimate an initial reward weight vector $\boldsymbol{\theta}$ for an IRL model, where the reward function is defined as:

$$R(s) = \boldsymbol{\theta}^\top \cdot \boldsymbol{\phi}(s)$$

Instructions:

- Analyze the provided travel diaries to identify patterns in the individual's observed activity behavior.
- Apply general travel behavior knowledge (e.g., typical preferences for activity types and times of day).
- Assign meaningful weights to the corresponding features in $\boldsymbol{\phi}(s)$.
- For features with insufficient or ambiguous evidence, assign values near zero.

Output Format
Return the 31-dimensional vector $\boldsymbol{\theta}$ as a valid Python list of 31 float values. Each value must be between -2 and 2. No additional explanation or formatting should be included.

Figure 5: Prompt used to initialize reward weights for IRL

Prompt Template for Reward Weight Update

Task Description
You are assisting in tuning the reward weight vector $\boldsymbol{\theta}$ for a maximum entropy inverse reinforcement learning (IRL) model. The objective is to adjust $\boldsymbol{\theta}$ to better align the learner's state visitation distribution with that of the expert.

Input:

- **Current Reward Weights:** $\boldsymbol{\theta}$ = *[current θ values]*
- **State Mismatches:** A list of the top 30 state mismatches, where each state is represented as a 4-tuple $(\text{hour}, \text{activity_type}, \text{is_first_trip}, \text{activity_segment_count})$, along with the expert and learner visitation frequencies for each state.

Reward Function
The reward function is defined as:

$$R(s) = \boldsymbol{\theta}^\top \cdot \boldsymbol{\phi}(s)$$

where $\boldsymbol{\phi}(s)$ is a 31-dimensional feature vector:

- 5 one-hot indicators for activity type: [Home, Work, School, Errand and Escort, Leisure]
- 24 hour-of-day indicators: [hour_0 to hour_23]
- 1 binary indicator: is_first_trip
- 1 normalized numeric feature: activity_segment_count

Objective
Based on the provided state mismatches and your prior domain knowledge, suggest an **update direction** for each of the 31 reward weights to reduce the discrepancies between the expert and learner state visitation distributions.

Instructions:

- Analyze the provided state mismatches and determine whether each feature weight in $\boldsymbol{\theta}$ should be increased, decreased, or left unchanged.
- For each of the 31 reward weights, output an update direction constrained to {-1, 0, 1}, where -1 indicates decrease, 0 indicates no change, and 1 indicates increase.

Output Format
Return a Python list of 31 integers, each being -1, 0, or 1. No additional text or explanation should be included.

Figure 6: Prompt template used to generate reward update directions from the LLM.

Prompt Template for Sociodemographic Attribute Prediction with CCR

Task Description
You are an expert in travel behavior modeling. In this task, you will apply the Theory of Planned Behavior (TPB) in reverse to infer ***[age / gender / income level / employment status]*** from the individual's intention, as captured by the learned reward weights, and environmental context.

Input:

- **Reward Weights:** $\boldsymbol{\theta}$, a 31-dimensional vector representing the individual's inferred preferences over activities, time-of-day, and trip structure.
- **Environmental Context:** External attributes such as urban/rural, population density, distance to transit, and housing characteristics.

Objective
Predict the individual's ***[age / gender / income level / employment status]*** by interpreting the beliefs encoded in $\boldsymbol{\theta}$, supported by the environmental context.

Instructions:

1. **Step 1: Belief Inference**
   Analyze the reward weights to identify the individual's underlying attitude, subjective norm, and perceived behavioral control.
2. **Step 2: Sociodemographic Prediction**
   Combine the inferred beliefs with the provided environmental context to predict the individual's ***[age / gender / income level / employment status]***.

Follow the above two steps in order when generating your prediction.

Output Format
Return only the predicted label (e.g., `0`, `1`, or `2`) corresponding to the target category. No explanation or additional formatting should be included.

Figure 7: CCR prompt for predicting sociodemographic attributes from IRL reward weights and contextual attributes

C Contextual and Extracted Features

C.1 Extracted Features

Table 5: Extracted Features and Definitions.

Extracted Feature	Definition
Trip Distance	Average distance per trip (in miles)
Std Trip Distance	Standard deviation of trip distance (in miles)
Num Trips per Day	Average number of trips taken daily
Std Num Trips per Day	Standard deviation of daily trip counts
Destination Entropy	Diversity of unique destination locations visited
% Work Trips	Percentage of trips for work or work-related purposes
% School Trips	Percentage of trips for school or educational purposes
% Shopping Trips	Percentage of trips for shopping activities
% Social/Recreation Trips	Percentage of trips for social or recreational purposes
% Errand Trips	Percentage of trips for business, errands, or appointments
% Escort Trips	Percentage of escort-related trips
Travel Time	Average travel time per trip (in minutes)
Std Travel Time	Standard deviation of trip travel times
First Departure Time	Hour of the first activity's departure
Last Departure Time	Hour of the last activity's departure
Home Time	Average daily duration spent at home
Work Time	Average daily duration spent at work

C.2 Contextual Features

The contextual attributes include: urban/rural indicator, population and housing density from the U.S. Census Bureau (U.S. Census Bureau, 2017, 2025), housing type and residential area proportion from the Washington State Geospatial Open Data Portal (Washington State Geospatial Open Data Portal, 2025), and both transit accessibility and network density from the EPA Smart Location Mapping dataset (U.S. Environmental Protection Agency, 2025).

Table 6: Contextual Features, Definitions, and Summary Statistics

Feature	**Definition**	**Summary Statistics**
Urban Indicator	Binary variable indicating urban (1) or rural (0)	1: 0.987, 0: 0.013
Population Density	People per square mile	Mean: 16047.36 Std: 14835.13
Distance to Transit	Distance to nearest public transit stop (m)	Mean: 270.48 Std: 191.10
Network Density	Road length (km) per km^2 of area	Mean: 29.03 Std: 10.51
Housing Density	Housing units per square mile	Mean: 9771.24 Std: 11253.57
Residential Proportion	Residential land share in home census block group	Mean: 0.346 Std: 0.194
Commercial Proportion	Commercial land share in home census block group	Mean: 0.130 Std: 0.114
Educational Proportion	Educational land share in home census block group	Mean: 0.055 Std: 0.073
Recreational Proportion	Recreational land share in home census block group	Mean: 0.054 Std: 0.094
Housing Type	Type of residence structure	Residential Condominium: 0.235 Single Family Unit: 0.334 Multi-Unit (2–4): 0.059 Multi-Unit (>5): 0.372

D Additional Prediction Results

D.1 Employment Status Prediction

Similar to age prediction, employment status prediction is framed as a three-class classification task. As shown in Table 7, while baseline models achieve reasonable overall accuracy, they struggle to accurately identify underrepresented groups (e.g., retired and employed populations) in the dataset. In contrast, our model not only achieves superior overall accuracy and weighted F1 score but also demonstrates more balanced performance across all classes.

D.2 Income Prediction

The prediction results for household income are presented in Table 8. It is important to note that the Puget Sound Regional Survey provides household-level income rather than individual income. Although we selected household representatives for analysis, their characteristics may not fully reflect household income levels due to unobserved intra-household factors such as household size. As a result, the overall prediction performance may be somewhat constrained by this limitation. Despite this limitation, our proposed model still significantly outperforms the baseline models, and demonstrates reasonable prediction performance, indicated by overall accuracy, weighted F1 score, and class-level F1 scores.

Table 7: Model comparison for employment status prediction. The proposed model achieves the highest overall performance and demonstrates balanced classification across categories.

Method	Class	Precision	Recall	F1-score	Total Accuracy	Weighted F1
SVM	unemployed	0.167	0.125	0.143	0.887	0.879
	employed	0.930	0.955	0.943		
	retired	0.667	0.500	0.571		
XGBoost	unemployed	0.250	0.375	0.300	0.879	0.885
	employed	0.955	0.938	0.946		
	retired	0.500	0.250	0.333		
CatBoost	unemployed	0.250	0.250	0.250	0.887	0.886
	employed	0.938	0.946	0.942		
	retired	0.667	0.500	0.571		
GPT-4o	unemployed	0.276	1.000	0.432	0.831	0.870
	employed	1.000	0.821	0.902		
	retired	1.000	0.750	0.857		
GPT-4o (Diaries)	unemployed	0.600	0.375	0.462	0.911	0.890
	employed	0.924	0.982	0.952		
	retired	0.000	0.000	0.000		
SILIC	employed	1.000	0.375	0.545	**0.952**	**0.943**
	unemployed	1.000	0.750	0.857		
	retired	0.619	0.796	0.696		

Table 8: Model comparison for household income prediction. The proposed model achieves the best overall and class-level performance.

Method	Class	Precision	Recall	F1-score	Total Accuracy	Weighted F1
SVM	<50k	0.636	0.259	0.368	0.452	0.418
	50–100k	0.560	0.255	0.350		
	100k+	0.398	0.833	0.538		
XGBoost	<50k	0.419	0.481	0.448	0.435	0.431
	50–100k	0.529	0.327	0.404		
	100k+	0.390	0.548	0.455		
CatBoost	<50k	0.367	0.407	0.386	0.403	0.393
	50–100k	0.455	0.273	0.341		
	100k+	0.393	0.571	0.466		
GPT-4o	<50k	0.333	0.667	0.444	0.524	0.533
	50–100k	0.646	0.562	0.602		
	100k+	0.727	0.381	0.500		
GPT-4o (Diaries)	<50k	0.714	0.185	0.294	0.540	0.487
	50–100k	0.521	0.909	0.662		
	100k+	0.571	0.286	0.381		
SILIC	<50k	0.909	0.588	0.714	**0.677**	**0.678**
	50–100k	0.641	0.610	0.625		
	100k+	0.619	0.796	0.696		