Title: Towards Adaptive System for Webpages: Learning Human Behaviors via Mouse Movement Patterns and Inverse Reinforcement Learning

1 Introduction

The new rules of social distancing and quarantining during COVID-19 have transformed every aspect of our lives, especially our behavior and routines. These changes have been drastically impacting the way we spend time (e.g., we spend more time at home and more time with computers or smart devices). With massive amounts of sensor data being collected everywhere (e.g., mouse movements and smart phones), we now can develop innovative algorithms to have a much greater understanding of complex behavior and behavior changes. The findings will open up the possibility to design adaptive systems that will reconfigure each web page to provide personalized framework and content.

The project provides a unique perspective to investigate behavioral norms at a population level and compare these norms among population subgroups under the current or transferred environment (e.g., switching between Gmail web page and Outlook webpage). Once differences in patterns are discovered, they can be used to better understand the impact of personal characteristics, such as age, health conditions, and education on daily routines. They can also be used to automate diagnoses and predict additional behavioral features of individuals within a group. The results will also help researchers in the fields of sociology, psychology, and anthropology to align their theories more closely with actual human behavior.

2 Methodology

This proposal focuses on understanding human behavior and its preferences based on mouse movements. The contributions are as follows: 1) first study of inverse reinforcement learning to model mouse movement patterns of an individual and its location preferences on a web page; 2) population level behavior analysis via mouse movement data; 3) proposing a new framework with a stack of multiple Long-Short-Term-Memory networks and a Power law mixture model, LSTM-PLMM, to synthesize data for small sample size problems; 4). utilizing cumulative fairness adaptive boosting method to achieve fairness and resolve imbalance problem.

2.1 Inverse Reinforcement Learning for Mouse Movement Patterns

When data are combined from multiple sources, such as multiple heterogeneous features, a challenge arises in creating a model that can include all available information [1]. For our study, we wanted to construct a model that combines temporal and spatial information of mouse movement. The selected model also had to be able to learn behavior preferences and distinguish behavior strategies between individuals and population groups. One data-driven approach to behavior modeling that meets these constraints is inverse reinforcement learning (IRL) [2].

We model a web browser as a 2D-grid composed of cells and mouse movement sequential patterns as a Markov Decision Process (MDP). We consider the problem as a study of inferring human intent (reward function) on a web-page setting and prefer a model-free approach, relative entropy IRL (RelEnt-IRL) [3], since for many real-world problems such as mouse movement patterns, reliable priors are not provided.

We assume that some "true" reward function R exists that can be expressed as a linear combination of some feature vector ϕ with a corresponding weight vector θ. Given a set of mouse movement trajectories $\mathcal{T}$ ($\tau \in \mathcal{T}$), a probability distribution over trajectories $P(\tau)$, and distribution induced by a baseline policy $Q(\tau)$, RelEnt-IRL minimizes the relative entropy (using the Kullback–Leibler divergence) between $P(\tau)$ and $Q(\tau)$. The problem can be formulated as:

$$\min_{\theta} \sum_{\tau \in \mathcal{T}} P(\tau) \ln \frac{\mathrm{P}(\tau)}{\mathrm{Q}(\tau)}, \ s.t. \left\{ |\sum_{\tau \in \mathcal{T}} P(\tau|\theta) \cdot \phi_i^{\tau} - \mu_i(\tau)| \leq \epsilon_i, \forall i \in 1,\dots,k; \sum_{\tau \in \mathcal{T}} P(\tau|\theta) = 1; P(\tau|\theta) \geq 0 \right\} \tag{1}$$

where k is the number of features, ϕ_i^{τ} is the i^{th} feature in the feature vector ϕ^{τ} that is extracted from a trajectory τ, $\mu_i(\tau)$ is the i^{th} feature expectation in $\mu(\tau)$, and ϵ_i is a threshold [4].

This constrained problem can be solved as an optimization problem with Lagrange multipliers and a Lagrange dual function via importance sampling to learn the preference vector θ. The learned θ will help us understand the location preference on a web-page for individuals and populations. The findings would open up the possibility to re-configure a webpage to adapt to each user's preferences instead of a fixed framework for everyone.