

1 Introduction

Optimizing predictive models datasets obtained from citizen science projects can be computationally expensive due to limited parallelism offered by Central Processing Units (CPUs). Since these crowd-sourced datasets grow with time as the researchers warrant more information, running models can prove increasingly difficult even on the fastest CPUs available in the market. However, Graphical Processing Units (GPUs), which offer multiple cores to parallelize computation, can outperform CPUs in computing such predictive models given the operations are arithmetically simple (though large in number) and boolean logic-free to the maximum extent.

One such predictive model is Avicaching [cite] in the eBird platform [cite], which aims to “maximize the utility and accessibility of the vast numbers of bird observations made each year by recreational and professional bird watchers” [cite website]. Since the eBird dataset is heterogeneous geographically, Avicaching tries to homogenize future observations by providing observer-citizens incentives to visit under-sampled locations [cite]. Therefore, the Avicaching game can be modeled as two separate problems - identifying parameters of citizens’ behavior on a specific set of rewards (incentives), and pooling rewards from a budget to locations such that the citizens’ predicted behavior is homogeneous.

1.1 Identification Problem

This subroutine will learn parameters of the change in citizens’ behavior due to rewards. Specifically, given datasets \mathbf{y}_t and \mathbf{x}_t of citizens’ visit densities with and without the rewards (r_t), we want to find weights \mathbf{w} that caused the change from \mathbf{x}_t to \mathbf{y}_t , factoring in possible influence due to environmental factors \mathbf{f} . Therefore, the model is:

$$\underset{\mathbf{w}}{\text{minimize}} \quad Z_I(\mathbf{w}) = \sum_t (u_t(\mathbf{y}_t - P(\mathbf{f}, \mathbf{r}_t; \mathbf{w})\mathbf{x}_t))^2 \quad (1)$$

where u_t is the total number of submissions at time t for a given citizen and elements $p_{u,v}$ of P are given as:

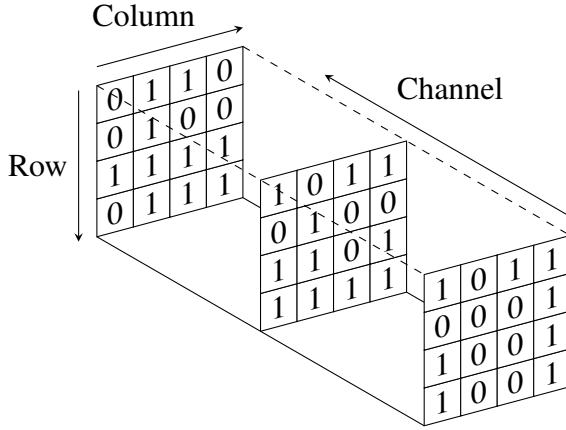
$$p_{u,v} = \frac{\exp(\mathbf{w} \cdot [\mathbf{f}_{\mathbf{u},\mathbf{v}}, r_u])}{\sum_{u'} \exp(\mathbf{w} \cdot [\mathbf{f}_{\mathbf{u}',\mathbf{v}}, r_{u'}])} \quad (2)$$

1.2 Pricing Problem

The Pricing subroutine will attempt to allocate rewards to all locations such that the predicted behavior is least heterogeneous, given the learned parameters from Section 1.1. This optimization problem can be written as:

$$\begin{aligned}
 & \underset{\mathbf{r}}{\text{minimize}} && Z_P(\mathbf{r}) = \frac{1}{n} \|\mathbf{y} - \bar{\mathbf{y}}\| \\
 & \text{subject to} && \mathbf{y} = P(\mathbf{f}, \mathbf{r}; \mathbf{w}) \mathbf{x} \\
 & && \sum_i r_i \leq \mathcal{R} \\
 & && r_i \geq 0
 \end{aligned} \tag{3}$$

where \mathcal{R} is total reward budget and P is defined as in Equation 2.



Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.