

CS329H: Machine Learning from Human Preferences

Sanmi Koyejo, Sang Truong

Autumn 2024

[Export PDF](#)

Welcome to CS329H

1. Introduction

2.1. Choice Models

2.2. Choice Models (continue)

3.1. Active Learning

3.2. Metric Elicitation

4.1. Preference Optimization via Dueling Bandits

4.2. Preferential Bayesian Optimization