Sida (Star) Li

(510)-847-2494 ◇ listar2000@berkeley.edu ◇ http://www.starli.xyz

EDUCATION

University of California, Berkeley, Berkeley, CA
Bachelor of Arts, Statistics & Computer Science, May 2022 (Expected) GPA: 4.00
Selected coursework: Theoretical Statistics§, Analysis of Time Series§, Bayesian Statistics§, Stochastic Processes, Linear Modelling, Deep Reinforcement Learning§, Real Analysis 1 & 2, Principles of Data Science
§: *graduate-level course*

RESEARCH PROJECTS

Comparing and Improving Approximate Bayesian Inference Methods
Mentor: Adityanand Guntuboyina *May 2021 - Present*
Performed empirical comparisons between variational and MCMC methods on Gaussian mixture models and LDA. Working on improving the efficiency of posterior approximation by combining variational inference into Monte-Carlo sampling algorithms.

Deep Reinforcement Learning for Autonomous Racing
Mentor: Allen Y. Yang *August 2020 - May 2021*
Developed a car racing simulator environment based on CARLA and applied model-based deep reinforcement learning algorithms to vehicle controllers for autonomous racing.

Recovering Gene Expression Datasets' True Dimensionality
Mentor: Hector Roux de Bezieux and Koen Van den Berge *January - May 2020*
Investigated into how initialization affects unsupervised dimensionality reduction methods such as UMAP and t-SNE in scRNAseq data, with an emphasis on the persevation of global structures in low dimensional space.

Directed Reading Program: Numerical Linear Algebra and Optimization
Mentor: Zhen Huang *August 2021 - Present*
Participate in weekly reading group on the application of numerical methods and convex optimization to statistical problems. Example topics include rank-deficient least square problems and regularized distribution estimation.

WORK EXPERIENCE

Software Engineer Intern, Duolingo, Pittsburgh, PA *May-August 2021*
Implemented internal tools in the ETL data pipeline that support efficient querying and computation on key metrics (e.g. daily bookings, active users); revised the A/B testing framework by enabling auto-correction in confidence intervals for ad-hoc metrics.

Data Consulting Intern, Concha Inc., Berkeley, CA *January-May 2020*
Worked on predicting customer's hearing loss curve based on response data from online testing. Experimented with and evaluated current machine learning methods such as regression tree and RNN. Prototyped and reported a reinforcement learning solution that enhanced prediction accuracy by 20% when data is sparse.

PAPERS

Michael Estrada, **Sida Li**, Xiangyu Cai. "Feedback Linearization of Car Dynamics for Racing via Reinforcement Learning." **arXiv, 2021**. [Here]

Sida Li, Adityanand Guntuboyina. "Variational Inference: a Tutorial for Undergraduates." Term report for the honor thesis program, **2021.** [Here]

PRESENTATION

"An introduction to variational inference." Final presentation for special seminar course. [Here]

SKILLS

Languages: English, Mandarin, Cantonese
Programming: Python, R, C++, Java, Javascript, Ruby, LaTeX

TEACHING

CS 198-097 Robot Autonomous Racing DeCal (Head Instructor)	Fall 2021
CS 198-097 Robot Autonomous Racing DeCal (Instructor)	Fall 2020
STAT 134 Probability Theory (Tutor)	Spring 2020
STAT 134 Probability Theory (Tutor)	Fall 2019
MATH 32 Precalculus (Tutor)	Summer 2019