Jing-Jing Li

jl3676@berkeley.edu
http://jl3676.github.io
Google Scholars

Education

University of California, Berkeley — Berkeley, CA
Ph.D. in Neuroscience with concentrations in computation and cognition, GPA: 3.94/4.00 — 2021–2026

Cornell University — Ithaca, NY
B.A. in Computer Science and Mathematics, Minor in Cognitive Science, GPA: 4.01/4.30 — 2017–2020

Work experience

Allen Institute for Artificial Intelligence (Mosaic Team) — Seattle, WA
PhD Research Intern — May 2024 - August 2024

- Developing a safeguarding mechanism for LLM outputs by modeling how they lead to harms to society.
- Performing prompt engineering, taxonomy development, natural language data generation, crowdsourcing, symbolic knowledge distillation, supervised fine-tuning, and evaluation.

Skills

- **Programming:** Python, Java, Julia, C, C++, Swift, Bash, Shell, OCaml
- **Data Science:** R, Numpy, SciPy, pandas, Matplotlib, Seaborn, MATLAB, SPM, FSL
- **Machine Learning:** TensorFlow, scikit-learn, PyTorch, OpenAI Gym, MuJoCo, CUDA, Kaggle, Google Colab
- **Natural Language Processing:** Large Language Models, Prompt Engineering, Fine-Tuning, Crowdsourcing
- **Experimental Design:** PsychoPy, Psychtoolbox, jsPsych, Amazon MTurk, EEGLAB, Persyst, EyeLink
- **Operating Systems:** Linux, Unix, Windows
- **Web Development:** HTML, CSS, JavaScript, Heroku
- **Database Management:** SQL, Microsoft Excel, RAID
- **Other:** LaTex, Adobe Illustrator, Adobe Photoshop, GitHub

Publications

[1] **J.-J. Li** and A. Collins, "An algorithmic account for how humans efficiently learn, transfer, and compose hierarchically structured decision policies", *PsyArXiv*, 2024.

[2] **J.-J. Li**, C. Shi, L. Li, and A. G. Collins, "Dynamic noise estimation: A generalized method for modeling noise fluctuations in decision-making", *Journal of Mathematical Psychology*, vol. 119, p. 102 842, 2024.

[3] **J.-J. Li**, C. Shi, L. Li, and A. G. Collins, "A generalized method for dynamic noise inference in modeling sequential decision-making", in *Proceedings of the Annual Meeting of the Cognitive Science Society*, 2023.

[4] C. McCafferty, B. F. Gruenbaum, R. Tung, **J.-J. Li**, X. Zheng, P. Salvino, P. Vincent, Z. Kratochvil, J. H. Ryu, A. Khalaf, *et al.*, "Decreased but diverse activity of cortical and thalamic neurons in consciousness-impairing rodent absence seizures", *Nature Communications*, vol. 14, no. 1, pp. 1–19, 2023.

[5] J. Ding, **J.-J. Li**, and M. Xu, "Classification of murmurs in pcg using combined frequency domain and physician inspired features", in *2022 Computing in Cardiology (CinC)*, IEEE, vol. 498, 2022, pp. 1–4.

[6] **J.-J. Li**, L. Xia, F. Dong, and A. G. Collins, "Credit assignment in hierarchical option transfer", in *Proceedings of the Annual Meeting of the Cognitive Science Society*, 2022.

Scholarships and Awards

- CogSci Conference Travel Grant 2023
- Milton I. and Florence Mack Neurology Research Fund 2021–2022
- Summer Undergraduate Research Fellowship (SURF), Caltech 2018

Conference Presentations

Talks

Dynamic noise modeling in decision-making Tahoe, CA
Berkeley Neuroscience Conference October 2023

A generalized method for dynamic noise inference Sydney, Australia
CogSci Conference July 2023

Credit assignment in the transfer of hierarchical options Toronto, Canada
CogSci Conference July 2022

Posters

Modeling how humans learn, transfer, and compose hierarchical policies Boston, MA
Cognitive Computational Neuroscience Conference August 2024

Credit assignment in the learning and transfer of hierarchical options San Francisco, CA
Cognitive Neuroscience Society Conference April 2022

Relevant Courses

- **Machine Learning:** Deep Unsupervised Learning, LLMs and Cognition, Deep Reinforcement Learning, Computer Vision, Intro to Machine Learning, Large-Scale Machine Learning, Computational Genetics
- **Software Engineering:** Object-Oriented Design and Data Structures (Honors), Algorithms, Computational Problem Solving, Operating Systems, Database Systems, Database Systems Practicum
- **Mathematics and Statistics:** Numerical Analysis, Biological Statistics, Basic Probability, Applicable Abstract Algebra, Linear Algebra (Honors), Multi-variable Calculus
- **Neuroscience:** Methods in Computational Modeling for Cognitive Science, Computational Psychology, Clinical Neuroscience, Developmental Psychology, Biopsychology, Cellular and Developmental Neuroscience

Research Experience

University of California, Berkeley, Dr. Anne Collins's Lab Berkeley, CA
PhD student August 2021 - Present

- Researching knowledge generalization in human reinforcement learning and decision-making.
- Collecting data for online behavioral experiments, conducting data analysis, and developing mathematical models to account for human behavior.

Yale University School of Medicine, Dr. Hal Blumenfeld's Lab New Haven, CT
Full-time Postgraduate Research Associate June 2020 - July 2021

- Researched the neuronal, electrophysiological, behavioral and hemodynamic changes underlying absence seizure severity in an awake rat model.
- Analyzed single unit recording data of thalamocortical neurons.

– Applied machine learning to classify seizures based on scalp EEG and single unit brain data.

University of Geneva, Dr. Daphne Bavelier's Lab Geneva, Switzerland

Full-time Visiting Undergraduate Research Student January - July 2019

– Researched the decoding of BOLD activation patterns of abstract symbols in visual cortex using fMRI and multivariate pattern analysis.

– Streamlined the preprocessing pipeline of fMRI images using MATLAB and SPM.

– Applied machine learning algorithms to predict visual stimuli based on brain activation patterns.

California Institute of Technology, Dr. Colin Camerer's Lab Pasadena, CA

Full-time Summer Undergraduate Research Fellow May - August 2018

– Researched the relationship between confidence and epistemic curiosity and the effects of curiosity on memory.

– Programmed a behavioral task and a follow-up survey in JavaScript using jsPsych.

– Conducted the experiment both on Amazon Mechenical Turk (n=104) and in lab (n=25) with eye-tracking.

– Performed preliminary data analysis in MATLAB and Python.