Reagan Lee

(626) 802-8933 • reaganjlee@berkeley.edu
linkedin.com/in/reaganjlee/ • github.com/reaganjlee

EDUCATION AND SKILLS

University of California, Berkeley **Dec 2023**
B.A. Data Science and Economics GPA: 3.7

Relevant Coursework

Natural Language Processing, Introduction to Machine Learning, Probability for Data Science, Data Mining & Analytics, Data Structures, Principles & Techniques of Data Science, Statistics and Mathematical Probability in Data Science, Game Theory, Linear Algebra

Skills

Software and Paradigms - Python, Java, SQL, Javascript, Latex, Computer Science, Software Engineering, AI/ML, Data Processing
Technical Skills - Numpy, Pytorch, Scikit-learn, Seaborn, Pandas, Matplotlib, Kubernetes, AWS, Jupyter, Go, C++, Git, Linux

RESEARCH EXPERIENCE

Breadth-Focused Red-Teaming for Language Models **Sep 2023 – Present**

- Develop new breadth-focused red-teaming methods for domain-agnostic evaluation of LLMs, advised by Stanford PhD Anka Reuel
- Collaborate with a team of 5 researchers by discussing relevant literature, ideas, and applications of adversarial robustness techniques
- Implement optimization-based red teaming using libraries such as Pytorch to evaluate performance, and experiment with new methods

Consistent Representations of Truth by Contrast-Consistent Search (CCS) **Feb 2023 – Aug 2023**

- Investigated capabilities of interpretability methods such as Contrast-Consistent Search, advised by MATS scholar Walter Laurito
- Implemented visualization modules in the Eliciting Latent Knowledge (ELK) library requiring knowledge of Pandas and Seaborn
- Ran LLM inference on CoreWeave compute clusters to both create new datasets and train linear probes on transformer activations

PROFESSIONAL EXPERIENCE

Pytest **June 2023 – Present**
Open-Source Contributor *Berkeley, CA*

- Shipped 'breaking'-type feature in pytest.warns in Pytest 8.0, Python's most popular testing library used by ~50% of all Python users
- Developed feature for dynamic assertion rewriting for set-type relational operators as part of library's core assertion functionality
- Collaborated with 5 maintainers on API design to refactor complex modules to improve readability and future functionality

HypothesisWorks **Feb 2023 – Present**
Open-Source Contributor *Berkeley, CA*

- Streamlined user-facing Healthcheck configurations for Hypothesis, a property-based testing library used by ~5% of all Python users
- Built pipeline to auto-refactor deprecated functions using Instagram's libCST, a syntax tree parser library, for downstream portability
- Optimized runtime performance on Hypothesis's filtered text strategy by 66% by utilizing strategy limitations set by filter conditions

LightningAI **Dec 2022 – Jan 2023**
Open-Source Contributor *Berkeley, CA*

- Standardized Coatue, Bain Capital-backed Series B startup's entire metrics codebase, requiring extensive knowledge of ML and Sphinx
- Collaborate with 3 open-source maintainers on codebase guidelines to maximize accessibility for users within extensive codebase

PROJECTS AND LEADERSHIP

Berkeley Supervised Program for Alignment Research **Aug 2023 – present**
Program Lead *Berkeley, CA*

- Lead team of 10 organizers to connect 90+ top technical students with 20+ mentors from Stanford, Berkeley, Columbia, etc.
- Organize program resources including $13k+ in compute credits and cluster access, and academic units to conduct AI safety research

Poker at Berkeley **Aug 2022 – present**
Cofounder, Head of External *Berkeley, CA*

- Lead team of six to organize 300+ participant, weekend-long poker tournaments, sponsored by trading firms Citadel and SIG
- Managed 1000+ user discord server, leading server-wide policies and as liaison between Poker@Berkeley and poker community at Cal

"Attention is All You Need" Paper Implementation **Aug 2022**

- Reimplemented the foundational 2017 transformer architecture paper "Attention Is All You Need" from scratch using Pytorch, Torchtext, and Spacy, requiring knowledge of the attention mechanism, encoder/decoder architectures, optimizers, and regularizations