Joseph Simonian

http://joseph.simonian.me
simonian.joseph@gmail.com | 805-709-1328

EDUCATION

UC BERKELEY

BA, Computer Science and Statistics (Double Major)
August 2017 | Berkeley, CA
Cum. GPA: 3.40
EECS GPA: 3.55

LINKS

Github:// **jsimonian**
LinkedIn:// **josephsimonian**
Website:// **joseph.simonian.me**

COURSEWORK

COMPUTER SCIENCE

Algorithms
Computer Security
Operating Systems
Computer Architecture
Database Systems
Artificial Intelligence
Machine Learning
Deep Neural Networks

STATISTICS

Statistical Learning Theory
Data Science
Linear Modeling

SKILLS

PROGRAMMING

Over 10000 lines:
Java • Python • JavaScript • Shell
Over 2000 lines:
C • C++ • HTML/CSS • PostgreSQL

EXPERIENCE

AMAZON LAB126 | Software Development Engineer

July 2017 – Present | Cupertino, CA

- Developed web-based data annotation platform for training of computer vision models in React/Node.js, processing millions of annotations per day from thousands of users.
- Migrated PostgreSQL data pipeline to a new platform using S3 for storage and SparkSQL for data processing, speeding up data processing jobs from hours to minutes on average.

PALO ALTO NETWORKS | Intern - Data Scientist

May 2016 – Aug 2016 | Santa Clara, CA

- Expanded coverage for URL filtering product by implementing additional web scraping methods, allowing operations on over 400 million additional pages.
- Implemented unsupervised clustering ML models in the Hadoop framework, speeding up data analysis 50x compared to the previous single-machine deployment.

UC BERKELEY EECS DEPARTMENT | Undergraduate Student Instructor (TA)

Jan 2016 – July 2017 | Berkeley, CA

- Designed course materials for new Introduction to Data Science (CS C8) and Intermediate Data Science (CS C100) courses.
- Led lab and discussion sections, designed and graded homework and exams, and supervised junior TAs (20 hours per week).

RESEARCH

MACHINE LEARNING AT BERKELEY (ML@B)

Continuous Control with Prioritized Experience Replay
Applied Prioritized Experience Replay, a strategy for improving the performance of discrete action space models such as Deep Q-Networks, to continuous action space models. Found improvements over uniform experience sampling and reduced training time.

PROJECTS

YELP SUDDEN SHIFT | Yelp Dataset Challenge 6

Utilized sentiment analysis and change point detection methods to determine the causes of sudden ratings changes for business in the yelp dataset.

REGEX-F | Personal

Chrome and Safari browser extension that extends standard Cmd-F functionality to allow page search for regular expressions.

GITLET++ | Personal

A distributed version control system, based on git. Originally based on **this project spec** , with lots of additions.