Darren Huang

darren.y.huang@gmail.com | github.com/darren-huang | +1(510) 574-5553

Experience

Bloomberg - *Sr. Software Engineer | Artificial Intelligence Group* *Sept 2020 - April 2023*

- Built, managed, and maintained core document ranking pipeline used for all client financial document queries
- Utilized large amounts of data and cutting edge technology to train low latency search relevancy models
  - Built scalable document re-ranking service easily handling 200K+ requests per day
  - Reduced model inference latency from ~1s to <200ms by optimizing featurization pipeline
  - Led migration of model inference to Kfserving/Kserve and designed deployment & remediation plans
  - Streamlined Apache Spark training data pipeline, reducing processing time by ~40% for terabytes of log data
  - Re-designed training pipeline to enable defining and iterating on new models with ~500 less lines of code
- Automated core training, testing, deployment, and monitoring processes
  - Designed a continuous training pipeline with Argo Workflows for periodic model retraining and deployment
    - Automates 8 manual steps from data collection to regression testing
  - Created Splunk, Humio, and Grafana dashboards to consolidate monitoring for 3 services across 20+ machines
  - Built continuous integration & deployment pipelines with Jenkins, Kserve, Docker, and Pytest

Facebook - *Software Engineering Intern | Infrastructure Team* *May 2019 - Aug 2019*

- Built scalable pipelines to analyze performance regressions on multiple internal platforms
- Recommend root cause investigation paths for any given regressions
- Designed robust end-to-end testing practices and made continuous integration tests for early bug detection
- Created an efficient but flexible API for performance engineers to customize recommendations

Institute for Human & Machine Cognition - *Software Intern | Robotics Lab* *May 2018 - Aug 2018*

- Extended existing reinforcement learning (RL) research from a discrete domain to a continuous one
- Created a framework that can augment the exploration of any RL algorithm with a state evaluation function
- Identify most promising states states to explore via UCB1 Algorithm & base RL algorithm's state evaluation
- Navigate to these identified states with a neural network dynamics estimator used for model predictive control

Alarm.com - *Software Engineering Intern | Infrastructure Team* *June 2017 - Aug 2017*

- Built a machine learning pipeline to predict server failures with visualizations of basic root cause analysis
- Detected real-time anomalies in timeseries data with supervised ensemble learning methods
- Track 20,000+ metrics and made a web dashboard to display the current most problematic metrics/servers

Organizations

Computer Science Mentors @ UC Berkeley - *Course Coordinator | CS61B* *Jan 2018 - May 2020*

- Lead meetings to coordinate 50+ mentors, 200+ students for UC Berkeley's Data Structures class
- Create educational materials covering data structures, algorithms, and software engineering fundamentals

DJI Robomaster AI Comp Team @ UC Berkeley - *Path & Strategies Lead* *Dec 2018 - May 2020*

- Lead sub-team responsible for autonomous robot path planning and defensive/offensive positioning

Education

University of California, Berkeley **Major GPA: 3.9**

Computer Science (B.A.), Applied Mathematics (B.A.) *[double major]* *Sept 2016 - May 2020*

- **Affiliations:** Upsilon Pi Epsilon, Undergraduate Student Instructor for Data Structures Course

Skills

- **Languages:** Python, Java, C, Bash, SQL/MySQL, MatLab, LaTeX
- **Technologies:** Kubernetes, Kfserving/Kserve, Apache Spark, Solr, Docker, Jenkins, Tensorflow
- **Machine Learning:** Learn-To-Rank, Reinforcement Learning, Deep Learning