Gary Cheng

2455 Hilgard Ave, Apt 11, Berkeley, CA 94709 • www.garycheng.me • garycheng@berkeley.edu

EDUCATION

University of California, Berkeley   **GPA: 4.0**

B.A. Computer Science | Anticipated Graduation: May 2019

- *Coursework*: Machine Learning, Time Series, Statistics, Probability & Random Processes, Optimization, Signal Processing, Algorithms, Data Structures, Operating Systems, Real Analysis, Linear Algebra
- *Awards*: Sutardja Center's Certificate in Entrepreneurship, Summer 2016; Cal Alumni Association Leadership Award Scholar 2015; Kraft Award 2015

PUBLICATIONS

Gary Cheng, Armin Askari, Kannan Ramchandran, and Laurent El Ghaoui. Frank-Wolfe Algorithm for Exemplar Selection. Presented poster at BayLearn 2018 and submitted to ICML 2019. Nov. 2018. *Preprint*: https://arxiv.org/abs/1811.02702

Gary Cheng, Kabir Chandrasekher, and Jean Walrand. Static and Dynamic Appointment Scheduling with Stochastic Gradient Descent. American Control Conference 2019. Sept. 2018. *Preprint*: https://bit.ly/2QST1IN

EXPERIENCE

Research Assistant   **Jan '18 – Present**

Laurent El Ghaoui and Kannan Ramchandran, BAIR Lab, UC Berkeley EECS

- Applying optimization techniques to find best exemplars that summarize a dataset. Instead of using entire dataset, exemplars can be used to train model. Has promise to reduce computational overhead.
- Designed a Frank Wolfe algorithm to greedily select exemplars. Proved that Algorithm has a linear convergence rate. Empirically outperforms other exemplar selection algorithms
- Supported by a **NSF REU** with Kannan Ramchandran; Presented poster at **BayLearn 2018**; Submitted paper as first author to **ICML 2019**

Research Assistant   **Aug '16 – Present**

Jean Walrand, BLISS Lab, UC Berkeley EECS

- Applying a stochastic gradient descent approach to the classic static scheduling problem, aiming to minimize idle, overtime, and wait times by altering when patients arrive at the hospital
- Developed novel dynamic appointment scheme which provably does just as good, if not better, than classic static scheme, dynamic scheme empirically outperforms classic static scheme
- Outstanding Presentation Award at **GCURS 2017** at Rice University; first author paper accepted to **American Control Conference 2019**

Research Assistant   **Aug '17 – Dec '17**

Laurent El Ghaoui, BAIR Lab, UC Berkeley EECS

- Tasked with speeding up a critical Non-Negative Least Squares sub-problem in one of Professor El Ghaoui's graduate student's projects
- After doing literature review and empirical testing, I implemented the fastest algorithm I found which halved the training time

Software Development Intern   **May '17 – Aug '17**

Amazon.com, Seattle

- Full stack developer on the Forecasting team in Supply Chain Optimization, used data analysis techniques to analyze runtime and success rates for the Forecasting Workflow Service Platform

- Implemented graph algorithms to give insight into the longest running components of forecasting calculations, created new Java APIs and integrated them into a Ruby on Rails front-end, researched customer usage patterns to determine the most useful metrics to develop

TEACHING

Teaching Assistant – Probability & Random Processes (EECS126)   **Jan '18 – May '18**
UC Berkeley EECS, Berkeley   **Jan '19 – May '19**
- Teaching a section of ~30 students about the following topics: Markov Chains, Poisson Processes, Entropy, Moment Generating Functions, Maximum Likelihood Estimation
- Holding office hours and a weekly homework help session, helping the professor develop exam/homework questions

Teaching Assistant – Algorithms (CS170)   **Aug '17 – Dec '17**
UC Berkeley EECS, Berkeley
- Teaching a section of ~30 students every week about the following topics: Dynamic Programming, Linear Programming, Streaming Algorithms, Divide and Conquer, and NP-Completeness
- Helping the professor develop exam/homework questions, holding office hours to help students learn

Teaching Assistant – Data Structures (CS61b)   **Jan '17 – May '17**
UC Berkeley EECS, Berkeley
- Teaching a section and lab of ~30 students every week, preparing supplemental discussion slides to aid student understanding in learning about data structures, working in office hours to resolve student issues
- Working on the Data Analysis team, finding trends in office hour, grading, and survey data

SIDE PROJECTS

Terrorist Attack Prediction *(github.com/garyxcheng/terrorism_prediction)*
- Used SARIMA Time Series models in R to predict the outcome and number of future terrorist attacks in the next year. Performed data interpolation and outlier removal operations to preprocess data from the Global Terrorism Database

Taxi Cab Markov Chain *(github.com/garyxcheng/pafnutys-taxi)*
- Used Python, Pandas DataFrames, Numpy to model NYC taxi cab movement/fares as a Markov Chain, Used k-means algorithm and elbow method to generate state space of Markov Chain. Found steady state distribution of the graph as well as the locations in NYC that maximize profit for taxi cab drivers.

LANGUAGES/FRAMEWORKS

Proficient: Python, Pandas, Numpy, Java, C, SQL
Familiar: PyTorch, TensorFlow, R, Ruby on Rails, Swift, HTML, CSS, JavaScript, Node.js