Stuart Sy

stuartsy@stanford.edu | (650) 868-9866
stuartsy.com | github.com/stuartsy | linkedin.com/in/stuartsy

EDUCATION

Stanford University – Stanford, CA — Sept 2014 – June 2018

- B.S. Computer Science 2018, Cumulative *GPA: 3.84*
- Related Coursework: Artificial Intelligence (CS221), Machine Learning (CS229), Computer and Network Security (CS155), Computer Networking (CS144), Databases (CS145), Web Applications (CS142)

WORK EXPERIENCE

Evernote – Redwood City, CA — June 2016 – Sept 2016

Data Engineering Intern

- Evaluated and integrated *Airflow*, a workflow management system, into the data pipeline as a replacement for the old *cron/bash* system. Migrated 90+ scripts worth of critical ETLs into python workflow definitions, allowing for more resilient and maintainable data pipeline management/deployment.
- Created a self-service analytics web portal for analysts to schedule recurring SQL reports and automate table creation/sync from Google Spreadsheets to the *ParAccel* analytics database.
- Set up a PoC machine learning pipeline through *Airflow* to periodically train on user activity metadata and feed predictions for user conversions into the data warehouse for marketing/promotions targeting.

Stanford Computer Science Department – Stanford, CA — Jan 2016 – Present

Section Leader

- Taught a weekly discussion section of Programming Methodology/Abstractions and held weekly office hours.
- Debugged, graded, and provided feedback on 106A and 106B students' Java and C++ programming assignments.

Focus Global Inc (Home Furnishings Distributor) – Taguig, Philippines — June 2015 – Aug 2015

Software Engineering Intern

- Revamped company's corporate website (focusglobalinc.com), and created a *Django*-powered furniture inventory catalog (ethanallen.com.ph) integrated with a RESTful API, which resulted in much faster inventory lookups.
- Automated collection of customer satisfaction surveys and created a dashboard for visualizing responses with *d3.js.*

PROJECTS

MapReduce Framework — Mar 2016

- A basic MapReduce implementation built to run on Stanford's myth computer clusters and shared AFS filesystem.
- Uses a threadpool and client/server sockets to spawn workers on separate machines to collectively operate on input.
- Used C++ mutexes and semaphores to ensure thread-safety and avoid concurrency related race-conditions.

MLB (Machine Learning Baseball) — Nov 2015

- A program which attempts to choose the best possible fantasy baseball team; achieves an average of 70% of best possible score in 10 seasons of test cases.
- Used the *pandas* data science library to clean, format, and extract relevant data from 100 years of baseball statistics.
- Predicts each individual player's score with linear regression algorithms from *scikit-learn* and models/solves the task of choosing the team as a constraint satisfaction problem.

Autonomous Robot Pacman — Sept 2015

- Three 40x40cm wheeled robots (two ghosts and pacman) programmed to autonomously play a game of Pacman.
- Robots communicate via a Bluetooth API (converted with pyobjc) which interfaces with the OSX CoreBluetooth stack.
- Algorithm uses an Expectimax game tree, a custom evaluation function for various states, and alpha/beta pruning.

TECHNICAL SKILLS

Languages: Python, Java, C, C++, SQL, HTML/CSS, Javascript, Bash, MATLAB, R

Tools/Technologies: Git, Rails, Django, Angular.js, d3.js, pandas/scikit-learn (ML)