Pei Yong Sim

EECS student looking for full-time opportunities in software engineering after **graduating in May 2017**. (Projects are available at pysim.me/projects; Resume available at pysim.me/cv.pdf)

Berkeley, CA, 94704
(831) 428-3525
py@pysim.me
www.pysim.me

EXPERIENCES

Synocate, Palo Alto — *Intern*

July 2016 - Sep 2016

Role: Helped build tools to enhance college admissions experience.

Projects
1. **Web scraper and browser emulator** to gather essay prompts to be populated to our site (tech stack include Node.js, Express, Swagger, PostgreSQL).
2. OkCupid's **matching algorithm** to best match counselors with students.
3. Summer programs **recommender system** with search and TF-IDF ranking.

University of California, Berkeley — *Academic Intern*

Jan 2016 - Present

Role: Help students in CS 189 (Machine Learning) and CS 61A (Intro to CS) with homework, projects and labs.

SELECTED PROJECTS *(* denotes in progress)*

Networking
1. PyChat — A full-stack (MVC) webapp that supports user management (registration/email confirmation), instant messaging (w/ WebSocket protocol on top of TCP/IP) and AI ChatBot. Also provide a REST API (demo: *pychat.pysim.me*).
2. Routing — Implemented a learning switch and distance vector routing with split horizon and poisoned reverse for efficient packet forwarding.
3. WAN Optimizer — Middlebox application that optimizes the amount of transmitted data over a wide-area network.

Systems
1. DBMS* — A SQL database management system that supports CRUD & Join operations, B+ trees indexing for performance enhancement, optimized querying and concurrency control. Written in Java.
2. PintOS — A Unix operating system framework that supports kernel threads, user programs execution and inode-structured file systems.
3. Interpreters/Compilers* — Built interpreters for statically and dynamically typed languages. Used Java w/ ANTLR for parse tree generation from context-free grammars. Currently implementing a compiler w/ Stack Machines.

A.I. & Machine Learning
1. A.I. Pacman — Pacman that uses sensors to locate and eat invisible ghosts.
2. OCR — Built a neural network that does handwritten digit recognition.
3. Sentiment Analysis — Simple NLP application (demo: *sentiment.pysim.me*).
4. SONG — Trained a linear regression model to predict release year of song given a set of audio features. Used Spark and MLlib.

EDUCATION I

University of California, Berkeley - *B.S. in EECS*

Aug 2015 - May 2017, GPA: 3.2

Coursework *(* denotes in progress)*:

1. Fundamentals: CS 61A (SICP), CS 47B (Data Structures), CS 170 (Algorithms), EE 16A/B (Signals & Systems Design)
2. Systems: CS 61C (Computer Architecture), CS 162 (Operating System), CS 168 (Computer Networking), CS 164* (Compilers), CS 186* (Databases)
3. Data Analytics: CS C8 (Data Science), CS 188 (Artificial Intelligence), CS 189 (Machine Learning), EE 126* (Probability in EECS), Stat 155 (Game Theory)

SKILLS

Programming Languages
1. Proficient: Python, Java, C/C++, SQL
2. Familiar: JavaScript, Ruby, Node.js

Frameworks
1. Distributed Computing: Spark
2. Web: Flask, Express, HTML, Bootstrap CSS

Libraries:
1. Data Analytics: TensorFlow, Sklearn, Numpy, Matplotlib
2. Testing: JUnit, Selenium
3. Web: jQuery, Scrapy

EXTRACURRICULARS

Data Science Society at Berkeley
Identifying factors that influence popularity of a loan on Kiva's microlending platform. Use NetworkX for ER random graphs generation in social network analysis.

Cabrillo College Tutor Provided tutoring assistance to students at the Math Learning Center.

LANGUAGES

Chinese, Cantonese, Malay

EDUCATION II

Cabrillo College, Santa Cruz - *CS*

Aug 2013 - May 2015, GPA: 3.9

Completed transfer work in computer science, physics and math. Built a website called cccPlan using Node.js and mongoDB at Hack UCSC 2015.