Pei Sim

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

UC Berkeley EECS graduate with one-year experience looking for full-time opportunities in **software engineering (backend/data/machine learning)** starting **Oct 2018**. All personal projects available at www.pysim.me/projects.

EXPERIENCES

Rakuten, San Mateo, CA — *Software Engineer*

Jun 2017 - Present

Role: Working on Product Catalog & Search team to enhance the ETL pipeline built using Big Data technologies (Mesos, Spark, Kafka, Cassandra, Elasticsearch, Redis, Kubernetes) and Java REST APIs. Work with PMs to identify business requirements, QA for deployment, TechOps for infrastructure management.

Projects (* currently migrating pipeline to Google Cloud)
1. Designing and integrating **Product Clustering** (built using various NLP and ML techniques) onto the pipeline which improved products matching by 5%.
2. Rebuilt the backend to become a **multi-tenant, multi-service platform** (clients include Ebates, Cartera, Citibank, Shopstyle). Provide Search, Price Tracker, Trending Products APIs as services. Platform serves >10 million engaged users.
3. Maintain systems **(Image, ML classification)** that enhance the product catalog from terabytes of raw crawled data & setting up **analytics/monitoring mechanisms** (visualizations and alert using Sensu, Kibana and New Relic).
4. **Fullstack BizOps app** that enables merchant onboarding and business requirements specification, and content curation to improve search results.

Synocate, Palo Alto, CA — *Software Engineer Intern*

July 2016 - Sep 2016

Role: Helped build tools to enhance college admissions experience (tech stack include Node.js, Express, Python, PostgreSQL).

Projects
1. **Web scraper & browser emulator** to gather essay prompts for student portal.
2. Summer programs **recommender system w/ text search engine**. A/B tests showed increased click-through rates in the group w/ recommendation.

University of California, Berkeley, CA — *Academic Intern*

Jan 2016 - May 2017

Role: Help students in CS 189 (Machine Learning) with homework and projects.

SELECTED PROJECTS

1. Compiler — Compiles a statically-typed Object Oriented language source code into Java bytecode-like language and then executes the compiled class files.
2. DBMS — SQL (relational) database management system that supports CRUD & Join operations, B+ trees indexing, query optimization and concurrency control.
3. NoSQL — Distributed Key-Value store that uses 2PC protocol for leader & follower servers coordination. It is fault tolerant and supports crash recovery.
4. PintOS — Unix operating system framework that supports kernel threads, user programs execution and file systems w/ buffer cache so that data is served faster.
5. PyChat — LAMP stack (MVC) web app that supports user management, instant messaging (w/ SSL on TCP/IP to encrypt, secure transmitted data) using event-driven architecture & A.I. ChatBot.

EDUCATION

University of California, Berkeley - *B.S. in EECS* (2017)

Coursework:
1. Fundamentals: Data Structures, Algorithms, Discrete Math, Systems Design I/II
2. Systems: Computer Architecture, Operating Systems, Networking (Internet Architecture), Compilers and Programming Languages, Databases
3. Data Analytics: Data Science, A.I., Machine Learning, Probability, Game Theory

SKILLS

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

Frameworks
1. Big Data: Spark, Hadoop
2. Web: Flask (Python), HTML, Bootstrap CSS

Libraries:
1. Data Analytics: TensorFlow, Scikit-learn, Numpy, Matplotlib (data visualization)
2. Web: jQuery, Scrapy

Other tools:
Git, Jenkins, Maven, Heroku Cloud, AWS, PostgreSQL, Kibana, New Relic, Bash scripting, Docker, Hive, Zookeeper

EXTRACURRICULARS

Data Science Society at Berkeley
Social Network Analysis on borrowers and lenders on Kiva's microlending platform.

Hack UCSC Semi-finalists in the hackathon where my team and I built cccPlan using Node.js and mongoDB.

LANGUAGES

Chinese, Cantonese, Malay