

Pei Sim

UC Berkeley EECS graduate with 1.3 years of experience looking for full-time opportunities in **software engineering (backend/data/machine learning)** starting **Nov 2018**. All personal projects available at www.pysim.me/projects.

EXPERIENCES

Rakuten, San Mateo, CA — Software Engineer

Jun 2017 - Present

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

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%. This is used by [Citibank Price Rewind](#) and [Ebates Price Magic](#).
2. Rebuilt the backend to become a **multi-tenant, multi-service platform** (clients include Ebates, United, 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 data & set up **analytics/monitoring/alerting mechanisms**.
4. Built a distributed streaming pipeline that provides **near real-time price updates**.
5. **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 (JVM) 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. **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

Berkeley, CA, 94704

(831) 428-3525

py@pysim.me

www.pysim.me

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, Jupyter Notebook, Jenkins (CI/CD), Maven, AWS, Kibana, New Relic, Bash scripting, Docker, Hive, Zookeeper, Sensu

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