Eddy Yeo

eddy.yeo@gmail.com

EDUCATION

Carnegie Mellon University - *BS Computer Science, Additional Major in Statistics*   **2012 - 2016**

- GPA: 4.0
- Notable Coursework: Operating Systems, Compiler Design, Distributed Systems, Probability and Computing, Machine Learning, Data Analysis
- SAT (college admission): 2400 / 2400

AI EXPERIENCE

Llama-Adapter   **April 2023**

- Contributed to the Hugging Face Parameter-Efficient Fine-Tuning (peft) library.
- Implementation of the Llama-Adapter by Zhang et al. that finetunes Llama using a set of learnable gated adaption prompts.
- https://github.com/huggingface/peft/pull/268

RobustFill   **Winter 2019 & 2023**

- Implemented a neural network by Devlin et al. that synthesizes a program in a string manipulation DSL based on a few examples of input-output string pairs.
- Trained on 4 x NVIDIA Tesla P4 using PyTorch's Distributed Data Parallel.
- https://github.com/yeoedward/Robust-Fill
- https://huggingface.co/eddyyeo/robustfill

Neural Turing Machine   **Spring 2016**

- Implemented a type of recurrent neural network that makes use of explicit memory and attentional processes (by Graves et al.) using TensorFlow.
- https://github.com/yeoedward/Neural-Turing-Machine

DATA EXPERIENCE

Blockchain Data Engineering
Founder   **March 2022 - Jan 2023**

- **Mixpanel for web3**: Data analytics platform for protocol developers to auto-generate metrics dashboard based on blockchain schema.
  - Built for terabytes of Solana data with the following systems:
    - BigQuery as the data warehouse
    - Prefect + Kubernetes for orchestration
- **Data3.Wiki**: Catalog of web3 projects and smart contract metadata.
  - Bootstrapped, hired and managed team of 3:
    - Frontend and backend engineers
    - Designer

Ricci.sg   **January 2021 - Present**
Founder

- News service that scrapes and aggregates news published in the Chinese language from a wide variety of sources with learning aids for people who are studying Chinese.

Two Sigma - New York, NY   **August 2016 - July 2020**
Software Engineer

- Worked on data engineering:
  - Data pipelines for the alternative data group.
- Worked on greenfield data platform projects:
  - Bitemporal knowledge graph platform
    - Service layer over SQL database to version graph data along two dimensions of time.
  - Entity linking service
    - Mapping various entities of interest like companies, people, etc. to canonical identifiers.
  - Data pipeline orchestration system
    - Encourage data scientists to share data by making it easy to deploy pipelines and share the resulting data.

Facebook - New York, NY   **Summer 2015**
Software Engineering Intern

- Worked with the newsfeed data team to build an internal user clustering tool.

Google - Mountain View, CA   **Summer 2014**
Software Engineering Intern

- Worked on the Similar Hotels feature in Google's Hotel Finder. Code was pushed to production as an experiment at the end of the summer.

Infosys - Bangalore, India   **Summer 2012**
Software Engineering Intern

- Worked on a mobile web application for monitoring electrical usage in office buildings.

SYSTEMS EXPERIENCE

Operating System Kernel   **Fall 2014**

- Wrote a Unix-inspired kernel from scratch, but excluding the file system. Was provided with the boot loader and various utility code like kernel malloc.

Compiler   **Fall 2015**

- Implemented a compiler for c0, a safe subset of C, with a semi-conservative "stop-the-world" mark-and-sweep garbage collector. The compiler was written in OCaml and the runtime in C.

Distributed Systems   **Spring 2014**

- Implemented a client/server protocol on top of UDP for distributed hashing. Implemented a distributed key-value store with consistent hashing. Implemented a distributed music playlist using paxos, with synchronized music playback. The playlist was voted best final project.

N-Store: Database for Non-Volatile Memory   **Spring 2015**

- Worked with a graduate student in CMU and contributed to the design and early implementation of the tile-based execution engine.