Jeffrey Wu

jeffwu@alum.mit.edu wuthejeff.com
858-405-1376 github.com/WuTheFWasThat

Summary: Full-stack developer with research experience and wishing to do good. Especially interested in keeping the world secure as technologies such as machine learning become more powerful.

Education

Massachusetts Institute of Technology **Cumulative GPA**: 4.8/5
B.S. in Mathematics, **B.S.** in Computer Science May 2012
M.Eng. in Computer Science January 2013

Work Experience

- **OpenAI** Research engineer Aug 2018 – Present
  Did research on language model training and reward modeling; see Research section. Also led efforts to apply human-feedback techniques to GPT-3, resulting in the deployment of the instruct-series of models in our API product.

- **Google Research** Software engineer Oct 2016 – Aug 2018
  Built general infrastructure (data pipelines, libraries, custom Tensorflow ops, a fun Lisp-like DSL) for supporting models for personalization from cross-product user history (basically learning giant embedding spaces for all users, Chrome pages, Youtube videos, etc). Experimented with RNN models to replace bag-of-words models, and helped launch news feed personalization experiments. In 20% time, studied properties of generalization error.

- **Terminal.com** Founding engineer Jan 2013 – Oct 2016
  Building cloud-based container infrastructure, for scientific computing and online education. Helped design and implement many core systems across the stack and oversaw their security and scalability. Saw company grow from 2 to 12, and managed a small team of engineers. Interfaced with clients, including Crunchbase, Stanford University, Codecademy, and Udacity. Company was sold to Udacity.

- **Probabilistic Computing Project** Master's student Nov 2011 – Jan 2013
  Implemented a probabilistic programming language. Explored a new Gibbs sampling algorithm to make inference more efficient in very general settings. Work presented [at NIPS 2012 probabilistic programing workshop]. [Source code] and [thesis].

- **Google** Software Engineering Intern Jun 2012 – Aug 2012
  Improved landing page quality for Dynamic Search Ads. Wrote MapReduces in C++ for identifying and finding replacements for pages with lower quality clicks.

- **6.006 Introduction to Algorithms** Teaching assistant Spring 2012, Fall 2012

- **Applied Operations Research, Inc** Intern Jun 2010 – Aug 2010
  Developed algorithms and models in MATLAB for ship routing under uncertain weather. The simulations convinced the navy to adopt new weather prediction technologies.

Research

- **Large-scale language model training** I trained GPT-2 (blog) and the early iterations of GPT-3 (NEURIPS 2020 best paper award), scaling up OpenAI's largest models from 110M parameters to over 6B parameters. I also helped extensively with building an evaluation suite, making performance optimization, and building a web UI.
- **Reinforcement learning from human preferences** I then switched my efforts full-time to learning from human feedback. I first helped with analysis at the tail end of the project Fine-Tuning Language Models from Human Preferences (blog). Then after starting with a clean slate (and switching from Tensorflow to Pytorch), we worked on the paper I'm most proud of, Learning to Summarize with Human Feedback (NEURIPS 2020) (blog). I was involved in nearly every aspect of this paper, and I think our results were very strong and somewhat underappreciated.
- **Miscellaneous** I also helped out a little on some miscellaneous projects, such as iGPT (ICML 2020) and Scaling Laws for Neural Language Models.

Selected Side Projects

- **Vimflowy** Vim inspired outlining tool with many features. [Source] (Typescript) and [Demo].
- **Hanabi simulation** Game engine for simulating hanabi strategies, and state of the art bots. Cited in [DeepMind/Brain paper] and subsequently [interviewed for WSJ]. [Source] (Rust).
- **Send A Damned Message** A simple puzzle game written to learn ReasonML. (Play here!)
- **tapystry** A small library for handling side effects in python,inspired by redux-saga [source]
- **plotserver** A small app for plotting results of ML jobs [source]

Skills

- Machine learning and Tensorflow
- Algorithms and distributed systems design
- Devops, e.g. AWS/GCP, linux, containers
- Front end, e.g. React frameworks
- CS theory
- Mathematics (2006-2008 USAMO, 2010 Putnam top 200)
- Productivity
- Learning new skills
- Keeping an eye on the big picture
- Acting with integrity