Edward Gan

edgan8@gmail.com edgan8.github.io

Education

- **Stanford University** — Stanford, CA
  PhD Student, Computer Science, Expected Graduation Spring 2020 — *2015 - Present*
  - Research with Prof. Peter Bailis on scalable statistics and data analytics
- **Stanford University** — Stanford, CA
  M.S. in Computer Science, GPA 4.1/4.0 — *Sep 2017*
- **Harvard University** — Cambridge, MA
  A.B. Summa Cum Laude in Computer Science and Mathematics — *May 2013*

Experience

- **Google** — Mountain View, CA
  Research Intern: Tensorflow Data Validation — *Summer 2019*
  - Developed customized data summarization algorithms for monitoring feature statistics
  - Improved end to end production data validation times by 20%
- **Airbnb** — San Francisco, CA
  Engineering Intern: Price Modeling — *Summer 2016*
  - Developed calibrated models in Spark to target users for customized price suggestions
  - Set up framework for variability analyses on the pricing model to prioritize improvements
- **Facebook** — Menlo Park, CA
  Software Engineer: Data Tools — *2013 - 2015*
  - Designed and built an execution engine and UI for backfilling data pipelines
  - Developed remediation and monitoring systems for dynamic product ads

Selected Publications

- **CrossTrainer: Practical Domain Adaptation with Loss Reweighting** — DEEM, ACM SIGMOD
  *Justin Chen, **Edward Gan**, Kexin Rong, Sahaana Suri, Peter Bailis* — *June 2019*
  - Robust techniques for automatic transfer learning across datasets.
- **Moment-Based Quantile Sketches for efficient ... Aggregation Queries** — SysML, VLDB
  ***Edward Gan**, Jialin Ding, Kai Sheng Tai, Vatsal Sharan, Peter Bailis* — *Aug 2018*
  - Distributed quantile estimation using a maximum entropy model, incorporated into Apache Druid.
- **Scalable Kernel Density Classification via Threshold-Based Pruning** — ACM SIGMOD
  ***Edward Gan**, Peter Bailis* — *May 2017*
  - Unsupervised, non-parametric outlier classification, outperforming scikit-learn.
- **MacroBase: Prioritizing Attention in Fast Data** — ACM SIGMOD
  *P. Bailis, **E. Gan**, S. Madden, D. Narayanan, K. Rong, S. Suri* — *May 2017*
  - System for anomaly explanation on multi-dimensional event log data.

Skills + Awards

- Proficient with Python and Java. Familiar with C++, NumPy / Pandas, PyTorch, Spark.
- NSF Graduate Research Fellowship 2015-2020.