Alex Bie

| alexbie98.github.io | alexbie98@gmail.com | 438-878-9409

EDUCATION

UNIVERSITY OF WATERLOO
MMath in Computer Science
2021-2023
Grade: 94/100
Advisors: Gautam Kamath & Shai Ben-David
Thesis: Private distribution learning with public data

UNIVERSITY OF WATERLOO
BMath in Computer Science
2016-2021
Grade: 92/100
Minor: Pure Mathematics

SELECTED PAPERS

(*) denotes alphabetical order.

[1] S. Ben-David*, A. Bie*, G. Kamath*, T. Lechner*. *Distribution learnability and robustness.* NeurIPS 2023.

[2] S. Ben-David*, A. Bie*, C. Cannone*, G. Kamath*, V. Singhal*. *Private distribution learning with public data: The view from sample compression.* NeurIPS 2023 *(as spotlight).*

[3] A. Bie, G. Kamath*, G. Zhang*. *Private GANs, revisited.* TMLR, 2023 *(with survey certification).*

[4] A. Bie*, G. Kamath*, V. Singhal*. *Private estimation with public data.* NeurIPS 2022.

[5] T. Cao, A. Bie, A. Vahdat, S. Fidler, K. Kreis. *Don't generate me: Training differentially private generative models with Sinkhorn divergence.* NeurIPS 2021.

[6] A. Bie, B. Venkitesh, J. Monteiro, M.A. Haidar, M. Rezagholideh. *Fully quantizing Transformer-based ASR for edge deployment.* Hardware-aware efficient training @ ICLR 2021.

SKILLS

Languages: Python • C/C++ • Java • JavaScript • SQL • OCaml • R

Technologies: NumPy/PyTorch /TensorFlow • Node.js • Unity • Linux

Coursework: Operating Systems • Algorithms • Graphics • Computer Vision • Machine Learning • Real Analysis • Differential Privacy

EXPERIENCE

HUAWEI | Research Engineer
July 2023 - Now | Montreal

- Research and development on privacy-preserving and federated learning.
- Data work for improving reasoning capabilities of LLMs.

HUAWEI | Research Engineer Intern
May 2022 - Nov 2022, Sept 2019 - Jan 2020 | Montreal

- Developed and improved over state-of-the-art **GAN** approaches for generating **synthetic data** under differential privacy guarantees [3,🔗] (reducing error rate from 17% → 5%).
- Added functionality, wrote documentation, and performed code reviews for the development of our federated learning codebase.
- Implemented and evaluated techniques for speeding up model inference of **Transformer** deep learning systems for on-device **speech recognition** [6,🔗]. Used **PyTorch** + **fairseq** to implement model training and **quantization**.
- Collaborated with product team to prototype Chinese speech recognition systems on internal datasets, using **Tensorflow** + Huawei's **GPU cloud cluster** (ROMA).

NVIDIA | Research Intern
May 2020 – Oct 2020 | Toronto

- Proposed, implemented, and iteratively improved a novel **generative modelling** approach (post 🔗) for producing synthetic versions of privacy-sensitive datasets.
- Designed and wrote the **Python** + **PyTorch** codebase (repo 🔗) used to run experiments on NVIDIA's **GPU cloud cluster** (NGC).
- Wrote a paper [5, 🔗] based on this work, which was accepted for publication at **NeurIPS'21**. We also filed a **patent** [scholar 🔗].

NATIONAL RESEARCH COUNCIL | Research Assistant
Jan 2019 – May 2019 | Ottawa

- Implemented and benchmarked models from natural language processing (**NLP**) research papers applying **LSTM** sequence models with **attention** for style-conditioned natural language generation in **PyTorch** and **Tensorflow.**

TRADEREV | Software Developer Intern
May 2018 – September 2018 | Toronto

- Carved out and refactored financial system components living inside a legacy **Groovy** + **SQL** monolith into **Node.js** microservices.
- Implemented a banking integration mock service and utilized testing frameworks to **improve code-coverage** of financial systems in pre-production environments.
- Performed production debugging on **$2M/day-throughput** financial system.

PROJECTS

Compiler (C++, 🔗). Compiler for Joos 1W (subset of Java 1.3), supporting **control flow**, **classes**, and **inheritance**. Implements lexical/syntatic analysis, name resolution + type-checking, and code generation.

Text Editor (C++, 🔗). `vim`-like modal text editor designed with **object-oriented design principles** in mind. Supports a limited set of `vim` commands.

AWARDS

2023 **Scotiabank Data Science Challenge** (1st place out of 29 teams, $2,000)
2021 **Vector Scholarship in Artificial Intelligence** ($17,500)
2021 **David R. Cheriton Graduate Scholarship** ($20,000)
2016 **Governor General's Bronze Medal** (Highest grade in HS graduating class)
2012 **CEMC Gauss Math Contest** (1 of 27 perfect scores out of 12,000 students)