Xinyan DAI

homepage github scholar linkedin

Email : xinyan.dai@outlook.com
Mobile : (+86)18795855867

Education

- **The Chinese University of Hong Kong**
  Doctor of Philosophy, supervised by **James Cheng** — *August 2018 - Present*
- **NanJing University** — Top 10%
  Bachelor of Engineering, Software Institute — *Sept 2014 - June 2018*

Research Interests

Similarity Search and Machine Learning. Typically, I am interested in applying similarity search techniques (Locality Sensitive Hashing, Sketching, Vector Quantization) on large scale machine learning.

Programming Skills

- **Languages**: C++, Python, Java, MPI, Hadoop, Tensorflow, Pytorch **Technologies**: Git, Linux

Projects

- **Gradient Quantization** : In this project we first apply Vector Quantization on gradient compression, i.e. **HSQ**, and compare it with classic scalar quantization SignSGD, QSGD, TernSGD.
- **Similarity Search** : A framework for index based similarity search. Where almost all classic similarity techniques, E2LSH, SRP, ITQ, PQ, IMI, Cross-LSH, indexing techniques for maximum inner product search, L2-ALSH, Sign-ALSH, Simple-LSH, and our papers **Norm-Ranging LSH, Rational-LSH** , are implemented and can be compared fairly.
- **Vector Quantization** : A framework for vector quantization (PQ, RQ, AQ, and our paper **NEQ**).
- **MPI-TensorFlow** : A library for tensorflow with MPI-support for distributed machine learning.
- **Tensor** : A numpy like computation library for c++.

Publication & working papers

- Norm-Explicit Quantization: Improving Vector Quantization for Maximum Inner Product Search
  Xinyan Dai**, Xiao Yan***, Kelvin K. W. Ng, Jie Liu, James Cheng [arxiv] [github]* [**AAAI 20**] [**Oral**]
- Hyper-Sphere Quantization: Communication-Efficient SGD for Federated Learning
  Xinyan Dai*, Xiao Yan, Kaiwen Zhou, Kelvin K. W. Ng, James Cheng [arxiv] [github]*
- Norm-Ranging LSH for Maximum Inner Product Search
  Xiao Yan, Jinfeng Li, **Xinyan Dai***, Hongzhi Chen, and James Cheng [arxiv] [github]* [**NeurIPS 18**]
- Norm-Range Partition: A Universal Catalyst for LSH based MIPS
  Xiao Yan, **Xinyan Dai***, Jie Liu, Kaiwen Zhou, James Cheng [arxiv] [github]*
- Understanding and Improving Proximity Graph based Maximum Inner Product Search
  Jie Liu**, Xiao Yan***,* **Xinyan Dai***, James Cheng, Ming-Chang Yang [arxiv]* [**AAAI 20**]

Awards

- **NeurIPS Travel Award** 2018
- **Chinese National Endeavor Scholarship** 2015
- **NJU People's Scholarship** 2016
- **NJU Sun Dalun Hong Kong, Macao and Taiwan Exchange Scholarship** 2016

Teaching Assistant

- **Problem Solving By Programming (ENGG1110)** 2018-2019 Term1 & 2, 2019-2020 Term 1