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 **Prof. 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 (LSH, Sketching, Vector Quantization) on large scale machine learning, e.g. **HSQ** and machine learning based similarity search, e.g. **CNN-ED**.

Skills

- **Mathmatical**: matrix theory, probability theory, optimization
- **Programming**: C++(MPI/OpenMP/CBlas), Python(TF, Pytorch, Numpy,), Java, Matlab, Hadoop

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 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.
- **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++.

Selected Publication

- Convolutional Embedding for Edit Distance
  Xinyan Dai, *Xiao Yan, Kaiwen Zhou, Yuxuan Wang, Han Yang, James Cheng [arxiv] [github]* [**SIGIR 20**]
- 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**]
- Norm-Ranging LSH for Maximum Inner Product Search
  Xiao Yan, Jinfeng Li, **Xinyan Dai**, *Hongzhi Chen, and James Cheng [arxiv] [github]* [**NeurIPS 18**]

working papers

- Hyper-Sphere Quantization: Communication-Efficient SGD for Federated Learning
  Xinyan Dai, *Xiao Yan, Kaiwen Zhou, Kelvin K. W. Ng, James Cheng [arxiv] [github]*

Awards

- **AAAI Travel Award** 2020
- **NeurIPS Travel Award** 2018
- **NJU People's Scholarship** 2016
- **Chinese National Endeavor Scholarship** 2015