# TAO YU

Gates Hall G23, Cornell University, Ithaca, NY
Homepage: http://www.cs.cornell.edu/∼tyu/
Email: tyu@cs.cornell.edu

## EDUCATION

**Cornell University, Ithaca, NY, United States** *Sep. 2019 - June 2024 (anticipated)*
Ph.D. in Computer Science Dept. of Computer Science
**Shanghai Jiao Tong University, Shanghai, China** *Sep. 2015 - June 2019*
B.S. in Mathematics and Applied Mathematics (Honors) ZhiYuan college

## RESEARCH INTEREST

I am intrigued by the prospect of integrating data geometry into machine learning and NLP, as it helps capture diverse properties exhibited by data across various tasks. I'm also dedicated to developing algorithmic and library solutions to ensure the robust numerical computation of low-precision ML models. Additionally, my interests extend to LLMs, machine learning privacy and robustness, along with a curiosity for emerging cognitive learning paradigms such as hyperdimensional computing.

## PUBLICATIONS

**Tao Yu***, Toni J.B. Liu*, Albert Tseng, Christopher De Sa. "Shadow Cones: Unveiling Partial Orders in Hyperbolic Space" (Under review).

**Tao Yu**, Yichi Zhang, Zhiru Zhang, Christopher De Sa. "FedHDC: Secure and Private Federated Hyperdimensional Computing", (Under review).

Albert Tseng, **Tao Yu**, Toni J.B. Liu, Christopher De Sa. "Coneheads: Hierarchy Aware Attention" —In 37th Conference on Neural Information Processing Systems (NeurIPS 2023).

**Tao Yu**, Christopher De Sa. "HyLa: Hyperbolic Laplacian Features For Graph Learning" -In 11th International Conference on Learning Representations (ICLR 2023).

**Tao Yu***, Yichi Zhang*, Zhiru Zhang, Christopher De Sa. "Understanding Hyperdimensional Computing for Parallel Single-Pass Learning" —In 36th Conference on Neural Information Processing Systems (NeurIPS 2022).

**Tao Yu***, Wentao Guo*, Jianan Canal Li*, Tiancheng Yuan*, Christopher De Sa. "MCTensor: A High-Precision Deep Learning Library with Multi-Component Floating-Point". —In 39th International Conference on Machine Learning (ICML 2022), Workshop on Hardware Aware Efficient Training (HAET 2022).

**Tao Yu**, Eugene Bagdasaryan, Vitaly Shmatikov. "Salvaging Federated Learning by Local Adaptation" (Preprint).

**Tao Yu**, Christopher De Sa. "Representing Hyperbolic Space Accurately using Multi-Component Floats". —In 35th Conference on Neural Information Processing Systems (NeurIPS 2021).

**Tao Yu**, Christopher De Sa. "Numerically Accurate Hyperbolic Embeddings Using Tiling-Based Models". —In 33rd Conference on Neural Information Processing Systems (NeurIPS 2019). **Spotlight**.

**Tao Yu***, Shengyuan Hu*, Chuan Guo, Weilun Chao, Kilian Q. Weinberger. "A New Defense Against Adversarial Images: Turning a Weakness into a Strength". —In 33rd Conference on Neural Information Processing Systems (NeurIPS 2019).

Felix Wu, Tianyi Zhang, Amauri Holanda de Souza Jr., Christopher Fifty, **Tao Yu**, Kilian Q. Weinberger. "Simplifying Graph Convolutional Networks". —In 36th International Conference on Machine Learning (ICML 2019).

**Tao Yu**, Huan long, John Hopcroft. "Curvature-based Comparison of Two Neural Networks". —In 24th International Conference on Pattern Recognition (ICPR 2018).

Mengxiao Zhang, Wangquan Wu, Yanren Zhang, Kun He, **Tao Yu**, Huan Long, John E Hopcroft. "The local dimension of deep manifold" (Preprint).

## EMPLOYMENT

**Applied Research Intern**, Amazon. *Aug. 2023 - Dec. 2023*
Manager: Luke Huan; Mentor: Gaurav Gupta AWS AI Team
— We studied the impact of mixed-precision strategies on LLM training, specifically by making the following contributions:
1) Designed a metric to explain the difference between various precision strategies
2) Design adaptive precision strategies with best trafe-off between memory, throughput and accuracy.

**Research Intern**, Apple. *2020, 2021, 2022, 2023 summer*
Manager: Vojta Jina, Mona Chitnis, Ulfar Erlingsson PriML Team
Mentor: Martin Pelikan, Congzheng Song MLPT
— Developed solutions for asynchronous federated learning to improve the system latency and performance of adaptive optimizers. Our method achieves a 3× speedup without performance drop.
— Proposed and improved personalized federated learning with cohort adaptation for language models.
— Worked on federated learning with privacy guarantees, in particular examine the privacy vulnerabilities in private federated learning, evaluate different privacy mechanisms (e.g. DP, SeparatedDP) with practical inference and reconstruction attacks.
— Proposed data poisoning attacks to craft outliers and theoretically measure the lower bound of privacy leakage in federated learning.

**Research Intern**, Cornell University, Ithaca, NY, USA. *July. 2018 - Dec. 2018*
Supervisor: Kilian Q. Weinberger, Christopher De Sa Dept. of Computer Science
— Worked on defending against adversarial examples. We proposed an algorithm to detect white-box adversarial attacks efficiently and accurately based on boundary information.
— Worked on simplifing graph convolution networks. We proposed the SGC model to speed up training with comparable performances to graph convolutional networks on a variety of tasks.
— Worked on the precision problem of hyperbolic embeddings, we constructed an accurate representation of hyperbolic space, which can represent any point within a constant distance (provably). Empirically our model outperforms state-of-the-art results.

**Research Intern**, MSAR, BeiJing, CHINA. *March. 2019*
Supervisor: Jifeng Dai Visual Computing
— Empirically studied the spatial attention mechanism in different scenarios, in particular, the graph attention networks, we found that previous attention mechanism is redundant and not efficient, and we proposed a more efficient attention mechanism in graph networks.

## PROFESSIONAL ACTICITY

**Talks:**

- **NeurIPS 2019**, "Numerically Accurate Hyperbolic Embeddings Using Tiling-Based Models."
- **VSAONLINE 2022**, "Understanding hyperdimensional computing for parallel single-pass learning." Invited talk in Vector Symbolic Architectures and Hyperdimensional Computing Workshop.

**PC/Reviewer:** NeurIPS, ICML, AISTATS, ICLR, KDD, SDM
**Teaching:** TA for CS1110, Intro to Computing with Python, Aug. 2019 - May. 2020