

# YIFAN CHEN

✉ [yifanc@comp.hkbu.edu.hk](mailto:yifanc@comp.hkbu.edu.hk) · 🏠 <https://ychen-stat-ml.github.io/> · 📅 Last updated: August 29, 2023

## RESEARCH INTERESTS

---

I am broadly interested in the general area of **efficient machine learning**, aiming to understand the statistical structures of modern machine learning algorithms and apply these insights to real-world computational challenges.

I especially focus on **non-parametric models** and neural networks with intensive matrix operations, such as **Transformers** (language models) and graph neural networks (**GNNs**).

## ACADEMIC POSITIONS AND EDUCATION

---

<b>Hong Kong Baptist University</b> , Hong Kong, China	Aug 2023 –
• Assistant Professor in Computer Science and Mathematics (affiliate)	
<b>University of Illinois Urbana-Champaign (UIUC)</b> , Illinois, United States	Aug 2018 – Aug 2023
• Ph.D. in Statistics, advisor: Prof. Yun Yang	
• M.S. in Statistics	
	May 2022
<b>Fudan University</b> , Shanghai, China	Sept 2013 – July 2018
• B.S. in Statistics, School of Management	
<b>The University of British Columbia</b> , British Columbia, Canada	Sept 2016 – Dec 2016
• Exchange Student at the UBC Sauder School of Business	

## HONORS, AWARDS, AND GRANTS

---

Start-up Grant, Hong Kong Baptist University	2023
🏆 <i>Dissertation Completion Fellowships</i> (declined due to graduation), USD <b>\$25,000</b> , U of I Graduate College	2023
The Fortieth International Conference on Machine Learning (ICML 2023) Grant Award, USD <b>\$1,500</b>	2023
🏆 Graduated with distinction: <i>Shanghai Outstanding Graduate</i> , Shanghai Municipal Education Commission	2018
Singapore Technologies Engineering Ltd Scholarship (top 5%), CNY <b>¥15,000</b> , Fudan University	2015 – 2017
Sumitomo Corporation Scholarship (top 5%), CNY <b>¥3,000</b> , Fudan University	2013 – 2014

## INDUSTRY EXPERIENCES

---

<b>Microsoft</b> , Washington, United States	May 2022 – Aug 2022
<b>Research Intern</b> — Mentor: Ritchie Zhao, Bitu Darvish Rouhani	
Study the compression of Mixture-of-Experts Transformers. We reconstruct expert MLPs through optimal transport. The work here motivates the follow-up project [2], published at ICML 2023.	
<b>Amazon</b> , California, United States	Aug 2021 – Dec 2021
<b>Applied Scientist Intern</b> — Mentor: Di Jin, Dilek Hakkani-Tur	
Do research on parameter-efficient transfer learning, with [4], [6] accepted to EMNLP 2022 Oral, NAACL 2022 Findings respectively. We explore the connection between attention and kernel estimators to guide the parameter assignments in adapters.	

**Summary** (2021-present): **13** first-authored/supervised papers, including **9** published/accepted papers and **4** preprints

### Peer-reviewed conference and journal papers

- [1] **A Gromov–Wasserstein Geometric View of Spectrum-Preserving Graph Coarsening** 🔗  
Yifan Chen, Rentian Yao, Yun Yang, Jie Chen  
*The Fortieth International Conference on Machine Learning (ICML 2023)*
- [2] **NTK-approximating MLP Fusion for Efficient Language Model Fine-tuning** 🔗  
Tianxin Wei\*, Zeming Guo\*, Yifan Chen\* ✉, Jingrui He ✉ (**Role**: main idea, proof, writing, co-mentoring)  
*The Fortieth International Conference on Machine Learning (ICML 2023)*
- [3] **Calibrate and Debias Layer-wise Sampling for Graph Convolutional Networks** 🔗  
Yifan Chen\*, Tianning Xu\*, Dilek Hakkani-Tur, Di Jin, Yun Yang, Ruoqing Zhu  
*Transactions on Machine Learning Research (TMLR)*, 2023
- [4] **Inducer-tuning: Connecting Prefix-tuning and Adapter-tuning** 🔗  
Yifan Chen\*, Devamanyu Hazarika\*, Mahdi Namazifar, Yang Liu, Di Jin, Dilek Hakkani-Tur  
*The 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP 2022)* **Oral**
- [5] **Sketching as a Tool for Understanding and Accelerating Self-attention for Long Sequences** 🔗  
Yifan Chen\*, Qi Zeng\*, Dilek Hakkani-Tur, Di Jin, Heng Ji, Yun Yang  
*2022 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2022)* **Oral**
- [6] **Empowering parameter-efficient transfer learning by recognizing the kernel structure in self-attention** 🔗  
Yifan Chen\*, Devamanyu Hazarika\*, Mahdi Namazifar, Yang Liu, Di Jin, Dilek Hakkani-Tur  
*2022 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2022) Findings*
- [7] **Skyformer: Remodel Self-Attention with Gaussian Kernel and Nyström Method** 🔗  
Yifan Chen\*, Qi Zeng\*, Heng Ji, Yun Yang  
*Thirty-Fifth Conference on Neural Information Processing Systems (NeurIPS 2021)*
- [8] **Accumulations of Projections—A Unified Framework for Random Sketches in Kernel Ridge Regression** 🔗  
Yifan Chen, Yun Yang  
*The 24th International Conference on Artificial Intelligence and Statistics (AISTATS 2021)*
- [9] **Fast Statistical Leverage Score Approximation in Kernel Ridge Regression** 🔗  
Yifan Chen, Yun Yang  
*The 24th International Conference on Artificial Intelligence and Statistics (AISTATS 2021)*
- [10] **Narrowing the Gap between Professionalism and Popularity: An Empirical Investigation on Community QA** 🔗  
Chenghong Zhang, Yifan Chen, Hongyue Lan, Yicheng Zhang, Tian Lu  
*American Conference on Information Systems (AMCIS 2017) Session on TREOS*

### Preprints and submissions

- [11] **Sampling-based Randomized Sketching for Approximate Matrix Multiplication**  
Yifan Chen, Yun Yang
- [12] **Connecting Cross-Domain Representations: A Ladder for Domain Generalization**  
Tianxin Wei\*, Yifan Chen\*, Xinrui He, Jingrui He (**Role**: proof, writing, co-mentoring)
- [13] **Statistical Leverage Score Approximation for Penalized Kernel Empirical Risk Minimization**  
Yifan Chen, Yun Yang
- [14] **Hypervolume Maximization: A Geometric View of Pareto Set Learning**  
Xiaoyuan Zhang, Yifan Chen ✉, Bo Xue, Xi Lin, Qingfu Zhang ✉ (**Role**: proof, writing, co-mentoring)

### Patents

- [1] **Sanitary wastewater reuse system**  
Yifan Chen  
*China Patent CN202187397U, published 2012-04-11*

## TEACHING EXPERIENCE

---

### Discussion Leader

- STAT 400: Statistics and Probability I Spring 2022, Fall 2022  
Instructor: Albert Yu

### Teaching Assistant

- STAT 425: Statistical Modeling I (Upper undergraduate level) Spring 2023  
Instructor: Prof. Trevor H. Park
- STAT 576: Empirical Process Theory and Weak Convergence (Ph.D. core course) Spring 2021  
Instructor: Prof. Sabyasachi Chatterjee
- CSE 428: Statistical Computing (Upper undergraduate level) Fall 2019, Fall 2020  
Instructor: Uma Ravat, Prof. Shulei Wang
- STAT 510: Mathematical Statistics (Ph.D. core course) Spring 2020  
Instructor: Prof. Yun Yang
- STAT 410: Statistics and Probability II Spring 2019, Summer 2019  
Instructor: Prof. Yun Yang, Alexey G Stepanov
- STAT 400: Statistics and Probability I Fall 2018  
Instructor: Prof. Hyoeun Lee

## MENTORING EXPERIENCE

---

- Zeming Guo, Undergraduate at UIUC (now Master at Cornell).  
Topic: Efficient Transformers. **had a first-author paper published at ICML 2023.**

## PROFESSIONAL SERVICES

---

### Program Committee

- International Conference on Machine Learning (ICML) 2022, 2023
- Neural Information Processing Systems (NeurIPS) 2022, 2023
- AAAI Conference on Artificial Intelligence (AAAI) 2023
- International Conference on Artificial Intelligence and Statistics (AISTATS) 2023

### Journal Reviewer

- Statistica Sinica
- Computational Statistics & Data Analysis
- IEEE Transactions on Information Theory
- Journal of the American Statistical Association

### External Conference Reviewer

- The Conference on Information and Knowledge Management (CIKM) 2022

## PROFESSIONAL TALKS

---

- A Gromov–Wasserstein Geometric View of Spectrum-Preserving Graph Coarsening  
Department of Math at Hong Kong Baptist University, Aug 2023
- NTK-approximating MLP Fusion for Efficient Language Model Fine-tuning  
At the 1st International Conference on AI-generated Content (AIGC2023), Remote, Aug 2023
- Inducer-tuning: Connecting Prefix-tuning and Adapter-tuning  
At EMNLP 2022, Remote, Dec 2022

- One Expert with Multiple Instruments  
Microsoft Azure, Remote, Aug 2022
- Sketching as a Tool for Understanding and Accelerating Self-attention for Long Sequences  
AI Time PhD NAACL Special Session, Remote, Aug 2022
- Sketching as a Tool for Understanding and Accelerating Self-attention for Long Sequences  
At NAACL 2022, Seattle, Washington, United States, July 2022
- Empowering parameter-efficient transfer learning by recognizing the kernel structure in self-attention  
Amazon Alexa AI, Sunnyvale, California, United States, Dec 2021
- Skyformer: Remodel Self-Attention with Gaussian Kernel and Nyström Method  
At NeurIPS 2021, Remote, Dec 2021
- Fast Statistical Leverage Score Approximation in Kernel Ridge Regression  
At AISTATS 2021, Remote, Apr 2021