Zhanpeng Zhou Ph.D.

(+86) 177-1755-7410
zzp1012@sjtu.edu.cn
https://zzp1012.github.io

Education

School of Computer Science, Shanghai Jiao Tong University Shanghai, China
Ph.D. in Computer Science and Engineering 2022 - 2027 *(expected)*

- Advisor: Prof. Junchi Yan
- Research area: Theory of Deep Learning, Science of Large Language Models

UM-SJTU Joint Institute, Shanghai Jiao Tong University Shanghai, China
B.E. in Electrical and Computer Engineering 2018 - 2022

- GPA (before post-grad recommendation): 3.74/4.00, Rank: 9/158.

Publications

(* indicates equal contributions; †indicates correspondence.)

1. Jinbo Wang*, Mingze Wang*, **Zhanpeng Zhou***, Junchi Yan, Weinan E, Lei Wu. The Sharpness Disparity Principle in Transformers for Accelerating Language Model Pre-Training. **In submission**.
2. Zijun Chen*, **Zhanpeng Zhou***†, Bo Zhang, Weinan Zhang, Xi Sun, Junchi Yan. SE-Merging: A Self-Enhanced Approach for Dynamic Model Merging. **In submission**.
3. Tongcheng Zhang*, **Zhanpeng Zhou***†, Mingze Wang, Andi Han, Wei Huang, Taiji Suzuki, Junchi Yan. Going Richer and Sparser: The Learning Dynamics of Label Noise SGD. **In submission**.
4. Andi Han*, Wei Huang*, **Zhanpeng Zhou***†, Gang Niu, Wuyang Chen, Junchi Yan, Akiko Takeda, Taiji Suzuki. On the Role of Label Noise in the Feature Learning Process. **In submission**.
5. Tongtian Zhu, Tianyu Zhang, Mingze Wang, **Zhanpeng Zhou**†, Can Wang. A Single Global Merging Suffices: Recovering Centralized Learning Performance in Decentralized Learning. **ICLR 2025 Workshop Weight Space Learning**.
6. **Zhanpeng Zhou**†, Yongyi Yang, Jie Ren, Mahito Sugiyama, Junchi Yan. On the Cone Effect in the Learning Dynamics. **ICLR 2025 Workshop DeLTa**.
7. **Zhanpeng Zhou***†, Mingze Wang*, Yuchen Mao, Bingrui Li, Junchi Yan. Sharpness-Aware Minimization Efficiently Selects Flatter Minima Late in Training. **ICLR 2025** (**Spotlight**).
8. Bingrui Li, Wei Huang, Andi Han, **Zhanpeng Zhou**, Taiji Suzuki, Jun Zhu, Jianfei Chen. On the Optimization and Generalization of Two-layer Transformers with Sign Gradient Descent. **ICLR 2025** (**Spotlight**).
9. **Zhanpeng Zhou***, Zijun Chen*, Yilan Chen, Bo Zhang, Junchi Yan. On the Emergence of Cross-Task Linearity in the Pretraining-Finetuning Paradigm. **ICML 2024**.
10. Yiting Chen, **Zhanpeng Zhou**, Junchi Yan. Going Beyond Neural Network Feature Similarity: The Network Feature Complexity and Its Interpretation Using Category Theory. **ICLR 2024**.
11. **Zhanpeng Zhou**, Yongyi Yang, Xiaojiang Yang, Junchi Yan, Wei Hu. Going Beyond Linear Mode Connectivity: The Layerwise Linear Feature Connectivity. **NeurIPS 2023**.
12. Ling Tang*, Wen Shen*, **Zhanpeng Zhou**, Quanshi Zhang. Defects of Convolutional Decoder Networks in Frequency Representation. **ICML 2023**.
13. **Zhanpeng Zhou***, Wen Shen*, Huixin Chen*, Ling Tang, Quanshi Zhang. Batch Normalization Is Blind to the First and Second Derivatives of the Loss. **AAAI 2024** (**Oral**).
14. Jie Ren, **Zhanpeng Zhou**, Qirui Chen, Quanshi Zhang. Optimizing Logistics and Supply Chain Networks Using Machine Learning Techniques. **ICLR 2023**.

Publications (continue)

15. Jie Ren*, Die Zhang*, Yisen Wang*, Lu Chen, **Zhanpeng Zhou**, Yiting Chen, Xu Cheng, Xin Wang, Meng Zhou, Jie Shi, Quanshi Zhang. A Unified Game-Theoretic Interpretation of Adversarial Robustness. **NeurIPS 2021**.

Internships

National Institute of Informatics. | Tokyo, Japan 2023.09 - 2024.03
- Research Intern, Advised by Prof. Mahito Sugiyama.

Mila Quebec. | Montreal, Canada 2021.03 - 2021.06
- Research Intern, Advised by Prof. Jian Tang.

Awards and Honors

- **National Scholarship (top 0.2%)**, Ministry of Education 2024.11
- **Top Internship Evaluation**, National Institute of Informatics 2024.03
- **Outstanding Graduate Student**, Shanghai Jiao Tong University, 2022.05
- **Yu Liming Scholarship**, Shanghai Jiao Tong University 2021.11
- **John Wu & Jane Sun Scholarship**, Shanghai Jiao Tong University 2020.11
- **Best Technology Award in Summer Expo**, Shanghai Jiao Tong University 2019.08

Academic Services

Conference Reviewer: *ICML ('22-25), NeurIPS ('22-25), ICLR ('24-25), AISTATS '25*

Journal Reviewer: *T-PAMI, Intelligent Computing (Science Partner)*