Wen Kaiyue

wenky20@mails.tsinghua.edu.cn && kaiyuewen3@gmail.com
（86）13670156595
Tsinghua University
Beijing,China, 100084

EDUCATION

BS, Institute for Interdisciplinary Information, Tsinghua University Expected 07/2023
Overall GPA: 3.94/4.00

AWARDS AND HONORS

1st Prize in National High School Mathematics Olympics Competition	11/2019
1st Prize in National High School Mathematics Olympics Competition	11/2018
Comprehensive Merit Scholarship of Tsinghua	10/2021
Silver Medal in S.-T. Yau College Student Mathematics Contest on Probability and Statistics (rank 3)	05/2021
Bronze Medal in S.-T. Yau College Student Mathematics Contest Team Track	05/2021
Silver Medal in S.-T. Yau College Student Mathematics Contest on Probability and Statistics (rank 3)	09/2022
Silver Medal in S.-T. Yau College Student Mathematics Contest Team Track (rank 2)	09/2022

PUBLICATION

Kaiyue Wen, *Tengyu Ma, Zhiyuan Li, How Sharpness Aware Minimization Minimizes Sharpness?, under review of ICLR2023*

Kaiyue Wen*, *Jiaye Teng*, Jingzhao Zhang, Realistic Deep Learning May Not Fit Benignly, under review of ICLR2023*

Xiaozhi Wang,* ***Kaiyue Wen*,*** *Zhengyan Zhang, Lei Hou, Zhiyuan Liu, Juanzi Li, Finding Skill Neurons in Pre-trained Transformers via Prompt Tuning, April ARR AC Assessment 4/5, under review of EMNLP 2022*

Yusheng Su, Xiaozhi Wang, Yujia Qin, Chi-Min Chan, Yankai Lin, Huadong Wang, ***Kaiyue Wen****, Zhiyuan Liu, Peng Li, Juanzi Li, Lei Hou, Maosong Sun, Jie Zhou, On Transferability of Prompt Tuning for Natural Language Processing, NACCL 2021*

SELECTED RESEARCH EXPERIENCE

Research on the limiting dynamics of Sharpness Aware Minimization 06/2022 –09/2022
Core group member, Supervised by Zhiyuan Li and Tengyu Ma, Stanford

- Investigate the limiting flow of Sharpness Aware Minimization when the learning rate and perturbation radius converges to 0 under the assumption that global minimizers form a Riemannian manifold
- Provably show different implicit bias for Sharpness Aware Minimization in full-batch and stochastic settings
- Show the necessity of stop gradient step by proving that without which the method will have different bias

Research on over-parameterized classification with label noise 03/2022 –09/2022
Core group member, Supervised by Jiaye Teng and Jingzhao Zhang, Tsinghua IIIS

- Show fundamental difference label noise cause on gradient descent algorithm on sub-gaussian mixture model
- Conduct simulation experiment showing how result generalize to real world data

Research on understanding soft prompt tuning via neuron activation pattern 08/2021 – 04/2022
Core group member, Supervised by Xiaozhi Wang and Zhiyuan Liu, Tsinghua NLP Lab

- Discover highly predictive neurons in pretrained language model which consistently emerges in prompt tuning
- Show importance of these neurons to the performance under various ways of parameter efficient tuning

SKILLS

- **Computer skills:** **Programming languages:** C/C++, Python, R
  Deep Learning Framework: Pytorch
- **Math skills:** Familiar with mathematics analysis, measure theory, linear algebra and abstract algebra
  Skillful in probability theory, statistics and elementary mathematics
- **Other skills:** Graduate level understanding on Causal Inference