Daiwei Chen

daiwei.chen@wisc.edu @daiweichen99 daiwei-chen
https://chendaiwei-99.github.io/

Education

2023-Present ◇ **Ph.D. Student, University of Wisconsin-Madison**
Research Topics: *Pluralistic Alignment, Representation Learning.*
Advisor: Ramya Korlakai Vinayak

2021-2023 ◇ **M.S. Electrical and Systems Engineering, University of Pennsylvania**
Research Topics: *Machine Learning Theory, PAC-Bayesian framework, Generalization.*
Advisor: Pratik Chaudhari

2017-2021 ◇ **B.S. Psychology, Zhejiang University**
Thesis title: *Prediction of Human Musical Emotion Perception Using Machine Learning.*
Advisor: Xiuying Qian, Yongchun Cai

Research Interests

Pluralistic Alignment; Foundation Models; In-Context Learning; Representation Learning; Preference Learning.

Research Projects

12/2023-Present ◇ **Pluralistic Alignment Framework** – *Research Assistant, MLOPT Lab, UW-Madison*

- Developed the PAL framework to address AI pluralistic alignment using latent variables and mixture modeling techniques.
- Demonstrated the PAL captures the diversity of user preferences while learning a shared latent preference space capable of few-shot generalizing to new users.
- Showcased PAL's competitive reward model accuracy in LLM tasks and image generation benchmarks, outperforming strong baseline models.

06/2022-06/2023 ◇ **Estimate the learning capacity of DNNs** – *Research Assistant, GRASP Lab, UPenn*

- Proposed a novel effective dimension estimation algorithm for both parametric and non-parametric models and discovered the phase transition phenomenon of the DNNs' effective dimension
- Analyzed the geometry property of DNNs' manifold by estimating the fisher information matrix and found that as the sample size increases, the DNNs model manifold becomes smoother.

09/2022-12/2022 ◇ **Generalization Performance of Lottery Ticket Hypothesis** – *Team Leader, UPenn*

- Implemented Lottery Ticket Hypothesis on different datasets/ model architectures/ optimizers/ pruning methods.
- Demonstrated that the winning ticket obtained on a specific dataset/ optimizer/ pruning method generalize well to other datasets/ optimizers.
- Although it's almost impossible to find the winning ticket before training, we found that the winning ticket obtained on a small sample size can generalize well on the whole dataset.

Research Publications

- **PAL: Pluralistic ALignment Framework for Learning from Heterogeneous Preferences**
  Daiwei Chen, Yi Chen, Aniket Rege, Ramya Korlakai Vinayak
  *MFHAIA workshop @ International Conference on Machine Learning (**ICML**), 2024* **(Oral)**
  *TF2M workshop @ International Conference on Machine Learning (**ICML**), 2024*
  Under Review, 2024
- **Unraveling The Impact of Training Samples**
  Daiwei Chen, Jane Zhang, Ramya Korlakai Vinayak
  *Blogpost @ International Conference on Learning Representations (**ICLR**), 2024*
- **Learning Capacity: A Measure of the Effective Dimensionality of a Model**
  Daiwei Chen*, Weikai Chang*, Pratik Chaudhari
  arXiv, preprint, 2023

Work Experience

05/2021-08/2021 ◇ **User Researcher, UX.** *NETEASE*, Interactive Entertainment
Conducted data analysis and user interviews to understand user sentiments and interactions with the product, aiming to enhance user experience.
Advisor: Fan Wu

07/2020-09/2020 ◇ **Machine Learning Engineer.** *HUAWEI*, Hangzhou Research Center
Analyzed user behavioral data using machine learning methods to develop and deliver personalized coupon offers efficiently.
Advisor: Xiuying Qian

Service and Organization

Summer 2022 ◇ **Volunteer Session Manager.** International Conference on Machine Learning (***ICML***).
◇ **Volunteer.** International Conference on Robotics and Automation (***ICRA***).
Spring 2022 ◇ **Graduate Teaching Assistant.** CS 350 Software Design & Engineering, ***UPenn***.
Fall 2021 ◇ **Graduate Teaching Assistant.** ESE 542 Statistic for Data Science, ***UPenn***.

Skills

AI Framwork ◇ PyTorch, TensorFlow.
Coding ◇ Java, C, Python, R, SQL, MatLab, LaTeX, Html, Css, JavaScript, …
Languages ◇ Strong reading, writing and speaking competencies for English, Mandarin Chinese.
Misc. ◇ Academic research, Teaching, Badminton, Running.

Awards and Achievements

2023 ◇ **Outstanding Research Award**, University of Pennsylvania.
2021 ◇ **Zhejiang University Scholarship**, Zhejiang University.
2020 ◇ **Academic Excellence Award**, Zhejiang University.
2019 ◇ **Title of School Outstanding Student**, Zhejiang University.