Hongxin Wei

PERSONAL INFORMATION

Assistant Professor , Southern University of Science and Technology, Shenzhen.

Homepage: https://hongxin001.github.io/

Google Scholar Profile: https://scholar.google.com/citations?user=cABH034AAAAJ&hl=en

Email: weihx@sustech.edu.cn, hongxin001@e.ntu.edu.sg

Phone: (+86)18818550154

Date of Birth: December 23, 1995

Nationality: China

Gender: Male

EDUCATION

- **Nanyang Technological University, Singapore** Jan 2020 - May 2023
  Ph.D. in Computer Science. Supervised by Prof. Bo An.

- **Huazhong University of Science and Technology, Wuhan, China** Sep 2012 - Jul 2016
  B.E. in Software Engineering.

RESEARCH APPOINTMENTS

- **Assistant Professor** (May 2023 - Present): Department of Statistic and Data Science, Southern University of Science and Technology, China.
- **Visiting Scholar/Intern** (Dec 2021 - Nov 2022): Department of Computer Sciences, University of Wisconsin Madison, US. Supervised by Prof. Yixuan Li.
- **Project Officer** (Aug 2018 - May 2023): School of Computer Science and Engineering, Nanyang Technological University, Singapore.
- **Research Intern** (Jun 2019 - Sep 2019): WeBank (Tencent), China.
- **Research Assistant** (Aug 2017 - June 2018): Institute for Interdisciplinary Information Sciences, Tsinghua University, China.

RESEARCH INTERESTS

Reliable Machine Learning, **AI Safety and Alignment**, **Data Optimization and Privacy**.

Research topics that I am currently working on include:

- LLM Uncertainty Estimation
- LLM Data Synthesis
- LLM Benchmark Contamination
- Conformal Prediction/Selection

PUBLICATIONS (†equal contribution; ⋆corresponding author)

Published Papers

46. Jianguo Huang, Jianqing Song, Xuanning Zhou, Bingyi Jing, **Hongxin Wei**⋆. TorchCP: A Python Library for Conformal Prediction. *The Journal of Machine Learning Research* (**JMLR**).

45. Huajun Xi, Kangdao Liu, Hao Zeng, Wenguang Sun, **Hongxin Wei**⋆. Exploring the Noise Robustness of Online Conformal Prediction. *Proceedings of the 39th Annual Conference on Neural Information Processing Systems* (**NeurIPS'2025**).

44. Beier Luo, Shuoyuan Wang, Yixuan Li, **Hongxin Wei**⋆. Your Pre-trained LLM is Secretly an Unsupervised Confidence Calibrator *Proceedings of the 39th Annual Conference on Neural Information Processing Systems* (**NeurIPS'2025**).

43. Junlei Zhou, Jiashi Gao, Xinwei Guo, Haiyan Wu, Quanying Liu, Xiangyu Zhao, **Hongxin Wei**, Xin Yao, Xuetao Wei. Does Confidence Calibration Help Conformal Prediction? *Proceedings of the 33rd ACM International Conference on Multimedia* (**MM'2025**).

42. Yufei Zhang, Yicheng Xu, **Hongxin Wei**, Zhiping Lin, Xiaofeng Zou, Cen Chen, Huiping Zhuang. Analytic Continual Test-Time Adaptation for Multi-Modality Corruption *Proceedings of the 33rd ACM International Conference on Multimedia* (**MM'2025**).

41. Huajun Xi, Jianguo Huang, Kangdao Liu, Lei Feng, **Hongxin Wei**⋆. Does Confidence Calibration Help Conformal Prediction? *Transactions on Machine Learning Research* (**TMLR**).

40. Shuoyuan Wang, Yixuan Li, **Hongxin Wei**⋆. Understanding and Mitigating Miscalibration in Prompt Tuning for Vision-Language Models. *Proceedings of the 42st International Conference on Machine Learning* (**ICML'25**).

39. Hao Zeng, Kangdao Liu, Bingyi Jing, **Hongxin Wei**⋆. Parametric Scaling Law of Tuning Bias in Conformal Prediction. *Proceedings of the 42st International Conference on Machine Learning* (**ICML'25**).

38. Hyeong Kyu Choi, Maxim Khanov, **Hongxin Wei**, Yixuan Li. How Contaminated Is Your Benchmark? Measuring Dataset Leakage in Large Language Models with Kernel Divergence. *Proceedings of the 42st International Conference on Machine Learning* (**ICML'25**).

37. Kangdao Liu, Hao Zeng, Jianguo Huang, Huiping Zhuang, Chi-Man Vong⋆, **Hongxin Wei**⋆. C-Adapter: Adapting Deep Classifiers for Efficient Conformal Prediction Sets. *Proceedings of the 28th European Conference on Artificial Intelligence* (**ECAI'25**).

36. Wenyu Jiang, Zhenlong Liu, Zejian Xie, Songxin Zhang, Bingyi Jing, **Hongxin Wei**⋆. Exploring Learning Complexity for Downstream Data Pruning. *Proceedings of the 13th International Conference*

on Learning Representations (**ICLR'25**).

35. Hengxiang Zhang, Songxin Zhang, Bingyi Jing, **Hongxin Wei***. Fine-tuning can Help Detect Pretraining Data from Large Language Models. *Proceedings of the 13th International Conference on Learning Representations* (**ICLR'25**).

34. Fanhu Zeng, Zhen Cheng, Fei Zhu, **Hongxin Wei**, Xu-Yao Zhang. Enhancing Outlier Knowledge for Few-Shot Out-of-Distribution Detection with Extensible Local Prompts. *Proceedings of the 13th International Conference on Learning Representations* (**ICLR'25**).

33. Shengjie Zhou, Senlin Shu, Haobo Wang, **Hongxin Wei**, Tao Xiang, Beibei Li. Multiple-Instance Learning from Pairwise Comparison Bags. *ACM Transactions on Intelligent Systems and Technology* (**TIST**), 2024.

32. Hongfu Gao, Feipeng Zhang, Wenyu Jiang, Jun Shu, Feng Zheng, **Hongxin Wei***. On the Noise Robustness of In-Context Learning for Text Generation. *Proceedings of the 38th Annual Conference on Neural Information Processing Systems* (**NeurIPS'24**).

31. Huiping Zhuang, Yizhu Chen, Di Fang, Run He, Kai Tong, **Hongxin Wei**, Ziqian Zeng, Cen Chen. GACL: Exemplar-Free Generalized Analytic Continual Learning. *Proceedings of the 38th Annual Conference on Neural Information Processing Systems* (**NeurIPS'24**).

30. Shuoyuan Wang, Jindong Wang, Guoqing Wang, Bob Zhang, Kaiyang Zhou, **Hongxin Wei***. Open-Vocabulary Calibration for Vision-Language Models. *Proceedings of the 41st International Conference on Machine Learning* (**ICML'24**).

29. Jianguo Huang, Huajun Xi, Linjun Zhang, Huaxiu Yao, Yue Qiu*, **Hongxin Wei***. Conformal Prediction for Deep Classifier via Label Ranking. *Proceedings of the 41st International Conference on Machine Learning* (**ICML'24**).

28. Zhenlong Liu, Lei Feng, Huiping Zhuang, Xiaofeng Cao, **Hongxin Wei***. Mitigating Privacy Risk in Membership Inference by Convex-Concave Loss. *Proceedings of the 41st International Conference on Machine Learning* (**ICML'24**).

27. Xiaobo Xia, Jiale Liu, Shaokun Zhang, Qingyun Wu, **Hongxin Wei**, Tongliang Liu. Towards Minimal Coreset Size under Model Performance Constraints . *Proceedings of the 41st International Conference on Machine Learning* (**ICML'24**), (Spotlight).

26. Chuanyang Zheng, Yixuan Wang, Yuqi Cheng, Xuesong Wang, **Hongxin Wei**, Irwin King, Yu Li. scNovel: a scalable deep learning-based network for novel rare cell discovery in single-cell transcriptomics. *Briefings in Bioinformatics* , (**BiB**), 2024.

25. Senlin Shu, Deng-Bao Wang, Suqin Yuan, **Hongxin Wei**, Jiuchuan Jiang, Lei Feng, Min-Ling Zhang.

Multiple-instance Learning from Triplet Comparison Bags. *ACM Transactions on Knowledge Discovery from Data* (**TKDD**), 2024.

24. Shiyu Tian, **Hongxin Wei**, Yiqun Wang, Lei Feng. Consistent Multi-Class Classification from Multiple Unlabeled Datasets. *Proceedings of The IEEE/CVF Conference on Computer Vision and Pattern Recognition* (**CVPR 2024**). (Oral).

23. Zixi Wei, Senlin Shu, Yuzhou Cao, **Hongxin Wei**, Bo An, Lei Feng. Consistent Multi-Class Classification from Multiple Unlabeled Datasets. *Proceedings of the 12th International Conference on Learning Representations* (**ICLR 2024**). (Spotlight).

22. Hao Chen, Jindong Wang, Ankit Shah, Ran Tao, **Hongxin Wei**, Xing Xie, Masashi Sugiyama, Bhiksha Raj. Understanding and Mitigating the Label Noise in Pre-training on Downstream Tasks. *Proceedings of the 12th International Conference on Learning Representations* (**ICLR 2024**). (Spotlight).

21. Wenyu Jiang, Hao Cheng, MingCai Chen, Chongjun Wang, **Hongxin Wei***. DOS: Diverse Outlier Sampling for Out-of-Distribution Detection. *Proceedings of the 12th International Conference on Learning Representations* (**ICLR 2024**).

20. Shuoyuan Wang, Jindong Wang, HuaJun Xi, Bob Zhang, Lei Zhang, **Hongxin Wei**. Optimization-Free Test-Time Adaptation for Cross-Person Activity Recognition. *The ACM international joint conference on Pervasive and Ubiquitous Computing* (**IMWUT/Ubicomp 2024**).

19. Renchunzi Xie, **Hongxin Wei***, Lei Feng, Yuzhou Cao, Bo An. On the Importance of Feature Separability in Predicting Out-Of-Distribution Error. *Proceedings of the 37th Annual Conference on Neural Information Processing Systems* (**NeurIPS'23**).

18. Yuzhou Cao, Hussein Mozannar, Lei Feng, **Hongxin Wei**, Bo An. In Defense of Softmax Parametrization for Calibrated and Consistent Learning to Defer. *Proceedings of the 37th Annual Conference on Neural Information Processing Systems* (**NeurIPS'23**).

17. Xin Cheng, Yuzhou Cao, Haobo Wang, **Hongxin Wei**, Bo An, Lei Feng. Regression with Cost-based Rejection. *Proceedings of the 37th Annual Conference on Neural Information Processing Systems* (**NeurIPS'23**).

16. **Hongxin Wei**, Huiping Zhuang, Renchunzi Xie, Lei Feng, Gang Niu, Bo An, Yixuan Li. Mitigating Memorization of Noisy Labels by Clipping the Model Prediction. *Proceedings of the 40th International Conference on Machine Learning* (**ICML'23**).

15. Eng Aik Chan, Carolina Rendón-Barraza, Benquan Wang, Tanchao Pu, Jun-Yu Ou, **Hongxin Wei**, Giorgio Adamo, Bo An, Nikolay I. Zheludev. Counting and mapping of subwavelength nanoparticles from a single shot scattering pattern. **Nanophotonics**.

14. Lei Feng, Senlin Shu, Yuzhou Cao, Lue Tao, **Hongxin Wei**, Tao Xiang, Bo An, Gang Niu. Multiple-Instance Learning from Unlabeled Bags with Pairwise Similarity. Transactions on Knowledge and Data Engineering (**TKDE**).

13. Senlin Shu, Shuo He, Haobo Wang, **Hongxin Wei**, Tao Xiang, Lei Feng. A Generalized Unbiased Risk Estimator for Learning with Augmented Classes. *Proceedings of the 37th AAAI Conference on Artificial Intelligence* (**AAAI'23**).

12. Lue Tao, Lei Feng, **Hongxin Wei**, Jinfeng Yi, Shengjun Huang, Songcan Chen. Can Adversarial Training Be Manipulated By Non-Robust Features? *Proceedings of the 36th Annual Conference on Neural Information Processing Systems* (**NeurIPS'22**).

11. Huiping Zhuang, Zhenyu Weng, **Hongxin Wei**, Renchunzi Xie, Toh Kar-Ann, Zhiping Lin. Analytic Class-Incremental Learning with Absolute Memorization and Privacy Protection. *Proceedings of the 36th Annual Conference on Neural Information Processing Systems* (**NeurIPS'22**).

10. **Hongxin Wei**, Renchunzi Xie, Hao Cheng, Lei Feng, Bo An, Yixuan Li. Mitigating Neural Network Overconfidence with Logit Normalization. *Proceedings of the 39th International Conference on Machine Learning* (**ICML'22**), pp.23631-23644, 2022.

9. **Hongxin Wei**, Lue Tao, Renchunzi Xie, Lei Feng, Bo An. Open-Sampling: Exploring Out-of-Distribution Data for Re-balancing Long-tailed Datasets. *Proceedings of the 39th International Conference on Machine Learning* (**ICML'22**), pp.23615-23630, 2022.

8. **Hongxin Wei**, Renchunzi Xie, Lei Feng, Bo Han, Bo An. Deep Learning from Multiple Noisy Annotators as A Union. *IEEE Transactions on Neural Networks and Learning Systems* (**TNNLS**).

7. Renchunzi Xie, **Hongxin Wei***, Lei Feng, Bo An. GearNet: Stepwise Dual Learning for Weakly Supervised Domain Adaptation. *Proceedings of the 36th AAAI Conference on Artificial Intelligence* (**AAAI'22**), 2022.

6. **Hongxin Wei**, Lue Tao, Renchunzi Xie, Bo An. Open-set Label Noise Can Improve Robustness Against Inherent Label Noise. *Proceedings of the 35th Annual Conference on Neural Information Processing Systems* (**NeurIPS'21**), pp.7978-7992, 2021.

5. Lei Feng, Senlin Shu, Yuzhou Cao, Lue Tao, **Hongxin Wei**, Tao Xiang, Bo An, Gang Niu. Multiple-Instance Learning from Similar and Dissimilar Bags. *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data* (**KDD'21**), pp.374-382, 2021.

4. Ziqi Zhang, Yuexiang Li, **Hongxin Wei**, Kai Ma, Tao Xu, Yefeng Zheng. Alleviating Noisy-label Effects in Image Classification via Probability Transition Matrix. *Proceedings of the 32nd British Machine Vision Conference* (**BMVC'21**), 2021.

3. Rundong Wang†, **Hongxin Wei**†, Bo An, Zhouyan Feng, Jun Yao. Commission Fee is not Enough:

A Hierarchical Reinforced Framework for Portfolio Management. *Proceedings of the 35th AAAI Conference on Artificial Intelligence* (**AAAI'21**), pp.626-633, 2021.

2. Lei Feng, **Hongxin Wei***, Qingyu Guo, Zhuoyi Lin, Bo An. Embedding-Augmented Generalized Matrix Factorization for Recommendation with Implicit Feedback. IEEE Intelligent Systems (IEEE-IS), in press.

1. **Hongxin Wei**, Lei Feng, Xiangyu Chen, Bo An. Combating noisy labels by agreement: A joint training method with co-regularization, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (**CVPR'20**), pp.13726-13735, 2020.

GRANTS

- Funding source: Shenzhen Science and Technology Innovation Commission. 2023.11 2026.11, **PI**

  Title: Research on data quality optimization for large foundation models

  Amount: 300K CNY

- Funding source: Guangdong Basic and Applied Basic Research Fund Committee. 2026 2029, **PI**

  Title: Research on data privacy protection of LLM post-training

  Amount: 100K CNY

- Funding source: Joint research lab with a company, 2024.01-2024.12, **Co-PI**

  Amount: 400K CNY

OPENSOURCE PROJECTS

- **TorchCP** (Downloads: 19.39K, published on **JMLR**)

  A Python toolbox for conformal prediction on deep learning models.

COURSE TEACHING

- Artificial Intelligence B: 2023 Fall
- Data Science Practice: 2024 Spring, 2025 Spring
- Selected Topics in Frontiers of Statistics II: 2025 Fall
- Mathematical Foundations of Artificial Intelligence: 2025 Fall (SLAI)

ACADEMIC SERVICES

Conference Area Chair:

- Neural Information Processing Systems (NeurIPS): 2024, 2025
- International Conference on Machine Learning (ICML): 2025
- International Conference on Learning Representations (ICLR): 2026

Conference Program Committee Member (Reviewer):

- Neural Information Processing Systems (NeurIPS): 2021, 2022, 2023
- IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR): 2022, 2023, 2024
- International Conference on Machine Learning (ICML): 2021, 2022, 2023, 2024
- International Conference on Computer Vision (ICCV): 2023
- ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD): 2021, 2022, 2023, 2024
- International Conference on Learning Representations (ICLR): 2022, 2023, 2024, 2025
- International Joint Conference on Artificial Intelligence (IJCAI): 2023
- SIAM International Conference on Data Mining (SDM): 2022

Conference Service Award:

- NeurIPS 2021 Outstanding Reviewer Award
- NeurIPS 2022 Top Reviewer Award

Journal Reviewer:

- Journal of the American Statistical Association (JASA)
- Journal of Machine Learning Research (JMLR)
- IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)
- International Journal of Computer Vision (IJCV)
- Pattern Recognition (PR)
- ACM Computing Surveys