

# WENXUAN BAO

Address: 614 E Daniel St, Champaign, IL 61820 ◊ Phone: (217)480-2150

Website: <https://baowenxuan.github.io> ◊ Email: [wbao4@illinois.edu](mailto:wbao4@illinois.edu)

## EDUCATION

**University of Illinois Urbana-Champaign**

**Champaign, IL, United States**

*Ph.D. Candidate in Computer Science*

*Jan 2021 - Present*

- GPA: 3.97/4.00, Thesis: Heterogeneous Machine Learning with Decentralized Data, Advisor: Prof. Jingrui He

**Tsinghua University**

**Beijing, China**

*B.Eng. in Automation, Minor in Statistics*

*Aug 2016 - Jun 2020*

- GPA: 3.91/4.00, Graduating Summa Cum Laude (top 2%), National Scholarship (top 2%)

## WORKING EXPERIENCE

**Jane Street Capital**

**New York, NY, United States**

*Quantitative Research Intern*

*May 2024 - Aug 2024*

- Used linear models and neural networks with BBO data to predict trading volumes of stocks across multiple time scales, supporting trading strategies.
- Applied large language models to extract and analyze key operational metrics from company current reports, automatically imputing missing data to support more informed investment decisions.

**ByteDance Ltd. | Pangle**

**Shanghai, China**

*Advertising System AI Intern*

*Jun 2020 - Dec 2020*

- Designed statistical models to improve the exposure of novel advertisements by 3 times.
- Optimized CVR prediction model to fasten the model calibration given limited historical data.
- Developed offline model evaluation tool to efficiently evaluate the model before deployment.

## RESEARCH INTERESTS

- **Trustworthy federated learning:** generalization, robustness, fairness, and privacy
- **Transfer learning:** test-time adaptation, domain adaptation, domain generalization

## PUBLICATIONS & PREPRINTS

### Conference

- [1] **AdaRC: Mitigating Graph Structure Shift during Test-Time**  
Wenxuan Bao, Zhichen Zeng, Zhining Liu, Hanghang Tong, Jingrui He  
The Thirteenth International Conference on Learning Representations (**ICLR 2025**)
- [2] **Ask, and It Shall Be Given: Turing Completeness of Prompting**  
Ruizhong Qiu, Zhe Xu, Wenxuan Bao, Hanghang Tong  
The Thirteenth International Conference on Learning Representations (**ICLR 2025**)
- [3] **Connecting Domains and Contrasting Samples: A Ladder for Domain Generalization**  
Tianxin Wei, Yifan Chen, Xinrui He, Wenxuan Bao, Jingrui He  
Proceedings of the 31th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (**KDD 2025**)
- [4] **BOBA: Byzantine-Robust Federated Learning with Label Skewness**  
Wenxuan Bao, Jun Wu, Jingrui He  
International Conference on Artificial Intelligence and Statistics (**AISTATS 2024**)

- [5] **FedSelect: Personalized Federated Learning with Customized Selection of Parameters for Fine-Tuning**  
Rishub Tamirisa, Chulin Xie, **Wenxuan Bao**, Andy Zhou, Ron Arel, Aviv Shamsian  
Conference on Computer Vision and Pattern Recognition (**CVPR 2024**)
- [6] **Adaptive Test-Time Personalization for Federated Learning**  
**Wenxuan Bao**, Tianxin Wei, Haohan Wang, Jingrui He  
Advances in Neural Information Processing Systems 36 (**NeurIPS 2023**)
- [7] **Optimizing the Collaboration Structure in Cross-Silo Federated Learning**  
**Wenxuan Bao**, Haohan Wang, Jun Wu, Jingrui He  
Proceedings of the 40th International Conference on Machine Learning (**ICML 2023**)
- [8] **Personalized Federated Learning with Parameter Propagation**  
Jun Wu, **Wenxuan Bao**, Elizabeth Ainsworth, Jingrui He  
Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (**KDD 2023**)
- [9] **Deep Grouping Model for Unified Perceptual Parsing**  
Zhiheng Li, **Wenxuan Bao**, Jiayang Zheng, Chenliang Xu  
Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR 2020**)

## Journal

- [10] **Privacy-Preserving Graph Machine Learning from Data to Computation: A Survey**  
Dongqi Fu\*, **Wenxuan Bao\***, Ross Maciejewski, Hanghang Tong, Jingrui He  
ACM SIGKDD Explorations Newsletter, Volume 25, Issue 1

## Preprint

- [11] **Understanding the Effect of Data Augmentation on Knowledge Distillation**  
Ziqi Wang, Chi Han, **Wenxuan Bao**, Heng Ji

## Tutorial

- [12] **Learning from Non-IID Data: Centralized vs. Federated Learning**  
Jun Wu, **Wenxuan Bao**, Jingrui He  
Proceedings of the 32nd International Joint Conference on Artificial Intelligence (**IJCAI 2023**)
- [13] **Learning from Non-IID Data: Centralized vs. Federated Learning**  
Jun Wu, **Wenxuan Bao**, Jingrui He  
2022 IEEE International Conference on Big Data (**IEEE BigData 2022**)

## RESEARCH EXPERIENCE

---

<b>University of Illinois Urbana-Champaign</b> <i>Research Assistant</i> (advisor: Prof. Jingrui He) · Thesis: Heterogeneous machine learning with decentralized data	<b>Champaign, IL, United States</b> <i>Jan 2021 - Present</i>
<b>University of Rochester</b> <i>Research Intern</i> (advisor: Prof. Chenliang Xu) · Project: Deep grouping model for unified perceptual parsing	<b>Rochester, NY, United States</b> <i>Jul 2019 - Dec 2019</i>
<b>Tsinghua University</b> <i>Research Assistant</i> (advisor: Prof. Hua Xu) · Project: Multimodal sentiment analysis	<b>Beijing, China</b> <i>Jul 2018 - Jun 2019</i>

## TEACHING EXPERIENCE

---

**University of Illinois Urbana-Champaign**

**Champaign, IL, United States**

*Teaching Assistant of CS 412: Introduction to Data Mining (instructor: Prof. Hanghang Tong) Jan 2023 - May 2023*

*Teaching Assistant of IS 577: Data Mining (instructor: Prof. Jingrui He) Aug 2024 - Dec 2024*

## SELECTED AWARDS

---

- NeurIPS 2023 Scholar Award *2023*
- ICML 2023 Grant Award *2023*
- IEEE BigData Tutorial Honorarium *2022*
- Summa Cum Laude, Tsinghua University (top 2%) *2020*
- Outstanding Graduates in Beijing (top 5%) *2020*
- National Scholarship (top 2%) *2017 and 2019*

## SERVICES

---

- Program committees of top-tier conferences: NeurIPS (2023, 2024), ICML (2024), ICLR (2024, 2025), CVPR (2024, 2025), AAAI (2024, 2025), AISTATS (2025), KDD (2021), CIKM (2021), BigData (2023), DSAA (2024), ICASSP (2025)
- Reviewers of top-tier journals: TKDD, TPAMI
- Student volunteer of ICML 2023

## SKILLS

---

- Languages: English (fluent), Chinese (native)
- Programming languages: Python, C/C++, R, MATLAB, SQL, Shell, Verilog
- Machine learning packages: PyTorch, PyG, TensorFlow, scikit-learn
- Data science packages: NumPy, SciPy, pandas, matplotlib, seaborn
- Tools: MapReduce, Git, Markdown, L<sup>A</sup>T<sub>E</sub>X
- Miscellaneous: I play bass guitar in student bands and sing baritone in Acapella groups