BOJIAN HOU

(also Bo-Jian Hou) 4036 Walnut St, Philadelphia, PA, 19104
Tel: (+1)814-826-8040 ⋄ Email: bojianh@upenn.edu

EDUCATION

Nanjing University (NJU) **Ph.D.**
Department of Computer Science and Technology *Sep.2014-Jun.2020*
Supervisor: Prof. Zhi-Hua Zhou

Nanjing University (NJU) **Bachelor of Science**
Department of Computer Science and Technology *Sep.2010-Jun.2014*

WORK EXPERIENCE

University of Pennsylvania *Jan.2022-Present* **Postdoctoral Researcher**

Cornell University *May.2021-Jan.2022* **Postdoctoral Associate**

4th Paradigm Co. *Aug.2020-Apr.2021* **Research Scientist**

RESEARCH INTEREST

My research interests mainly focus on trustworthy AI, optimization for AI and AI for science. Specifically, they include:

- **Fairness Learning**: developing fair and unbiased machine learning algorithms.
- **Interpretability**: studying the interpretability of black-box machine learning models.
- **Feature Evolvable Learning**: studying learning scenarios where data features evolve.
- **Semi-Supervised Learning**: learning models from both labeled and unlabeled data.
- **Online Learning**: learning models from a continuous stream of data.
- **Natural Language Processing**: studying and applying large language models.
- **Biomedical Data Mining**: applying machine learning to analyze biomedical data.

PUBLICATIONS

Conference

1. **Bo-Jian Hou**, Lijun Zhang, and Zhi-Hua Zhou. *Learning with Feature Evolvable Streams.* **In: Advances in Neural Information Processing Systems 30 (NIPS'17), Long Beach, CA**, 2017, 30: 1417-1427. (Ranked #1 in Artificial Intelligence by Google Scholar)

2. **Bo-Jian Hou**, Lijun Zhang, and Zhi-Hua Zhou. *Storage Fit Learning with Unlabeled Data.* **In: Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI'17), Melbourne, Australia**, 2017, 1844-1850. (Ranked #9 in Artificial Intelligence by Google Scholar)

3. **Bo-Jian Hou**, Yu-Hu Yan, Peng Zhao and Zhi-Hua Zhou. *Storage Fit Learning with Feature Evolvable Streams.* **In: Proceedings of the AAAI Conference on Artificial Intelligence (AAAI'21), Virtual Conference**, 2021, 35(9), 7729-7736. (Ranked #4 in Artificial Intelligence by Google Scholar)

4. Zhuoping Zhou*, Davoud Ataee Tarzanagh*, **Bojian Hou***, Boning Tong, Jia Xu, Yanbo Feng, Qi Long, and Li Shen. *Fair Canonical Correlation Analysis.* **In: Proceedings of the 37th Annual Conference on Neural Information Processing Systems (NeurIPS'23), New Orleans, LA**, 2023, 36. (Ranked #2 in Artificial Intelligence by Google Scholar, * means equal contribution or co-first authors)

5. Davoud Ataee Tarzanagh*, **Bojian Hou***, Boning Tong*, Qi Long, and Li Shen. *Fairness-Aware Class Imbalanced Learning on Multiple Subgroups.* **In: Proceedings of the 39th Conference on Uncertainty in Artificial Intelligence (UAI'23), Pittsburgh, PA**, 2023, 2123-2133. (* means equal contribution or co-first authors)

6. Davoud Ataee Tarzanagh, Parvin Nazari, **Bojian Hou**, Li Shen, and Laura Balzano. *Online Bilevel Optimization: Regret Analysis of Online Alternating Gradient Methods.* **In: Proceedings of the 27th International Conference on Artificial Intelligence and Statistics (AISTATS'24), Palacio de Congresos de València, València SPAIN**, 2024, 2854-2862.

7. Yi He, Jiaxian Dong, **Bo-Jian Hou**, Yu Wang, and Fei Wang. *Online Learning in Variable Feature Spaces with Mixed Data.* **In: Proceedings of the 21st IEEE International Conference on Data Mining (ICDM'21), Auckland, New Zealand**, 2021, 181-190. (Ranked #6 in Data Mining & Analysis by Google Scholar)

8. Heng Lian, John S. Atwood, **Bo-Jian Hou**, Jian Wu, and Yi He. *Online Deep Learning from Doubly-Streaming Data.* **In: Proceedings of the 30th ACM International Conference on Multimedia (ACMMM'22), Lisbon, Portugal**, 2022, 3185-3194. (Ranked #4 in Multimedia by Google Scholar)

9. Zhuoping Zhou, Boning Tong, Davoud Ataee Tarzanagh, **Bojian Hou**, Andrew J. Saykin, Qi Long, and Li Shen. *Multi-Group Tensor Canonical Correlation Analysis.* **In: Proceedings of the 14th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM BCB, 2023), Houston, TX**, 2023, 1-10. (This paper won the best paper award.)

10. **Bo-Jian Hou** and Yuan Jiang. *Learning Interpretability from RNN with Feature Evolving.* **In: CCF Conference on Artificial Intelligence (CCFAI'19), Xuzhou, China, 2019**. (Recipient of the CCFAI Outstanding Student Paper Award)

11. Mingquan Lin, Yuyun Xiao, **Bojian Hou**, Tingyi Wanyan, Mohit Manoj Sharma, Zhangyang Wang, Fei Wang, Sarah Van Tassel, and Yifan Peng. *Evaluate Underdiagnosis and Overdiagnosis Bias of Deep Learning Model on Primary Open-Angle Glaucoma Diagnosis in Under-Served Populations.* **In: Proceedings of the AMIA 2023 Informatics Summit, Seattle, WA**, 2023, 370.

12. **Bojian Hou**, Hongming Li, Zhicheng Jiao, Zhen Zhou, Hao Zhang, and Yong Fan. *Deep Clustering Survival Machines with Interpretable Expert Distributions.* **In: Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI'23), Cartagena de Indias, Colombia**, 2023, 1-4.

13. Boning Tong, Zhuoping Zhou, Davoud Ataee Tarzanagh, **Bojian Hou**, Andrew J. Saykin, Jason Moore, Marylyn Ritchie, and Li Shen. *Class-Balanced Deep Learning with Adaptive Vector Scaling Loss for Dementia Stage Detection.* **In: Proceedings of the 14th International Workshop on Machine Learning in Medical Imaging (MLMI'23), Vancouver, Canada**, 2023, 144-154.

14. Weiqing He*, **Bojian Hou***, George Demiris, and Li Shen. *Interpretability Study for Long Interview Transcripts from Behavior Intervention Sessions for Family Caregivers of Dementia Patients.* **In: Proceedings of AMIA 2024 Informatics Summit, Boston, MA**, 2024, 201. (* means equal contribution or co-first authors)

15. Ruiming Wu, Bing He, **Bojian Hou**, Andrew J Saykin, Jingwen Yan, and Li Shen. *Cluster Analysis of Cortical Amyloid Burden for Identifying Imaging-driven Subtypes in Mild Cognitive Impairment.* **In: Proceedings of AMIA 2024 Informatics Summit, Boston, MA**, 2024, 439.

16. **Bojian Hou**, Andrés Mondragón, Davoud Ataee Tarzanagh, Zhuoping Zhou, Andrew J Saykin, Jason H Moore, Marylyn D Ritchie, Qi Long, and Li Shen. *PFERM: A Fair Empirical Risk Minimization Approach with Prior Knowledge.* **In: Proceedings of AMIA 2024 Informatics Summit, Boston, MA**, 2024, 211.

17. Jia Xu*, Tianyi Wei*, **Bojian Hou***, Patryk Orzechowski, Shu Yang, George Demiris, Li Shen. *MentalGPT: Harnessing AI for Compassionate Mental Health Support.* **In: Proceedings of AMIA 2024 Annual Symposium, San Francisco, LA**, 2024, in press. (* means equal contribution or co-first authors)

18. Yanbo Feng*, **Bojian Hou***, Ari Klein, Karen O'Connor, Jiong Chen, Andrés Mondragón, Shu Yang, Graciela Gonzalez-Hernandez, Li Shen. *Analyzing Dementia Caregivers' Experiences on Twitter: A Term-Weighted Topic Modeling Approach.* **In: Proceedings of AMIA 2024 Annual Symposium, San Francisco, LA**, 2024, in press. (* means equal contribution or co-first authors)

19. Boning Tong, Travyse Edwards, Shu Yang, **Bojian Hou**, Davoud Ataee Tarzanagh, Ryan J. Urbanowicz, Jason H. Moore, Marylyn D. Ritchie, Christos Davatzikos, Li Shen. *Ensuring Fairness in Detecting Mild Cognitive Impairment with MRI.* **In: Proceedings of AMIA 2024 Annual Symposium, San Francisco, LA**, 2024, in press.

Journal

20. **Bo-Jian Hou**, Lijun Zhang, and Zhi-Hua Zhou. *Prediction with Unpredictable Feature Evolution.* **IEEE Transactions on Neural Networks and Learning Systems (TNNLS)**, 2021, 33(10): 5706-5715. (Impact factor: 14.255, Ranked #7 in Artificial Intelligence by Google Scholar)

21. **Bo-Jian Hou** and Zhi-Hua Zhou. *Learning with Interpretable Structure from Gated RNN.* **IEEE Transactions on Neural Networks and Learning Systems (TNNLS)**, 2020, 31(7): 2267-2279. (Impact factor: 14.255, Ranked #7 in Artificial Intelligence by Google Scholar)

22. **Bo-Jian Hou**, Lijun Zhang, and Zhi-Hua Zhou. *Learning with Feature Evolvable Streams.* **IEEE Transactions on Knowledge and Data Engineering (TKDE)**, 2019, 33(6): 2602-2615. (Impact factor: 9.235, Ranked #2 in Data Mining & Analysis by Google Scholar)

23. Jie Ren, **Bojian Hou**, and Yuan Jiang. *Deep Forest for Multiple Instance Learning.* **Journal of Computer Research and Development**, 2019, 56(8): 1670-1676. (Impact factor: 1.043)

24. Mingquan Lin, **Bojian Hou**, Lei Liu, Mae Gordon, Michael Kass, Fei Wang, Sarah H. Van Tassel, and Yifan Peng. *Automated diagnosing primary open-angle glaucoma from fundus image by simulating human's grading with deep learning.* **Scientific Reports**, 2022, 12(1): 14080. (Impact factor: 4.379)

25. Mingquan Lin, **Bojian Hou**, Swati Mishra, Tianyuan Yao, Yuankai Huo, Qian Yang, Fei Wang, George Shih, and Yifan Peng. *Enhancing thoracic disease detection using chest X-rays from PubMed Central Open Access.* **Computers in Biology and Medicine**, 2023: 106962. (Impact factor: 6.698)

26. Heng Lian, Di Wu, **Bo-Jian Hou**, Jian Wu, and Yi He. *Online Learning from Evolving Feature Spaces with Deep Variational Models.* **IEEE Transactions on Knowledge and Data Engineering (TKDE)**, 2023, 36(8): 4144-4162. (Impact factor: 9.235, Ranked #2 in Data Mining & Analysis by Google Scholar)

27. Zexuan Wang, Qipeng Zhan, Boning Tong, Shu Yang, **Bojian Hou**, Heng Huang, Andrew J. Saykin, Paul M. Thompson, Christos Davatzikos, and Li Shen. *Distance-weighted Sinkhorn loss for Alzheimer's disease classification.* **iScience**, 2024, 27(3): 109212. (Impact factor: 4.6)

28. Jing-Xiao Liao, **Bo-Jian Hou**, Hang-Cheng Dong, Hao Zhang, Jinwei Sun, Shiping Zhang,

and Feng-Lei Fan. *Quadratic Neuron-empowered Heterogeneous Autoencoder for Unsupervised Anomaly Detection*. **IEEE Transactions on Artificial Intelligence** in press, 2024. (Impact factor: 4.9)

29. **Bojian Hou***, Zixuan Wen*, Jingxuan Bao, Richard Zhang, Boning Tong, Shu Yang, Junhao Wen, Yuhan Cui, Jason H Moore, Andrew J. Saykin, Heng Huang, Paul M. Thompson, Marylyn D. Ritchie, Christos Davatzikos, and Li Shen. *Interpretable Deep Clustering Survival Machines for Alzheimer's Disease Subtypes Discovery*. **Medical Image Analysis**, 2024, 97:103231. (Impact factor: 10.7, * means equal contribution or co-first authors)

Manuscript

30. **Bojian Hou**, Hao Zhang, Gur Ladizhinsky, Ali Kayyal, Stephen Yang, Volodymyr Kuleshov, Fei Wang, and Qian Yang. *Clinical Evidence Engine: Proof-of-Concept For a Clinical-Domain-Agnostic Decision Support Infrastructure*. **arXiv preprint** arXiv:2111.00621, 2021.

31. Dayang Wang, Feng-Lei Fan, **Bo-Jian Hou**, Hao Zhang, Rongjie Lai, Hengyong Yu, and Fei Wang. *Manifoldron: Direct Space Partition via Manifold Discovery*. **arXiv preprint** arXiv:2201.05279, 2022.

32. Yan Ma, Weicong Liang, Yiduo Hao, **Bojian Hou**, Xiangyu Yue, Chao Zhang, and Yuhui Yuan. *Revisiting DETR Pre-training for Object Detection*. **arXiv preprint** arXiv:2308.01300, 2023.

33. Dawei Li, Shu Yang, Zhen Tan, Jae Young Baik, Sukwon Yun, Joseph Lee, Aaron Chacko, **Bojian Hou**, Duy Duong-Tran, Ying Ding, Huan Liu, Li Shen, Tianlong Chen. *DALK: Dynamic Co-Augmentation of LLMs and KG to answer Alzheimer's Disease Questions with Scientific Literature*. **arXiv preprint** arXiv:2405.04819, 2024.

34. Ruochen Jin, **Bojian Hou**, Jiancong Xiao, Weijie Su, Li Shen. *Fine-Tuning Linear Layers Only Is a Simple yet Effective Way for Task Arithmetic*. **arXiv preprint** arXiv:2407.07089, 2024.

AWARDS & HONORS

PennAITech Innovation Fellow 2024

Best Paper Award in the 14th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics 2023

Excellent Doctoral Dissertation Award of Jiangsu Province 2021

Excellent Doctoral Dissertation Award of Nanjing University 2021

JSAI Excellent Doctoral Dissertation Award 2020

CS Excellent Doctoral Dissertation Award of Nanjing University 2020

Outstanding Graduate Student Award of Nanjing University 2020

CCFAI Outstanding Student Paper Award. 2019

The Program A for Outstanding PhD Candidate of Nanjing University 2019

National Scholarship for Ph.D. 2017

NeurIPS Volunteer Award 2017

IJCAI Travel Award 2017

Outstanding Undergraduate Student Award of Nanjing University 2014

National Endeavor Scholarship 2012

PROFESSIONAL ACTIVITIES

Program Committee (PC) Member or Reviewer of Conferences

- PC Member of NeurIPS'24 (The 38th Annual Conference on Neural Information Processing Systems)
- PC Member of ICML'24 (The 41th International Conference of Machine Learning)
- PC Member of NeurIPS'23 (The 37th Annual Conference on Neural Information Processing Systems)
- PC Member of NeurIPS'22 (The 36th Annual Conference on Neural Information Processing Systems)
- PC Member of ICML'22 (The 39th International Conference of Machine Learning)
- PC Member of NeurIPS'21 (The 35th Annual Conference on Neural Information Processing Systems)
- SPC Member of IJCAI'21 (The 30th International Joint Conference on Artificial Intelligence)
- PC Member of ICML'21 (The 38th International Conference of Machine Learning)
- PC Member of AAAI'21 (The 35th AAAI Conference on Artificial Intelligence) (**I was selected as one of the top 25% PC members at this conference.**)
- Reviewer of KDD'20 (The 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining)
- PC Member of NeurIPS'20 (The 34th Annual Conference on Neural Information Processing Systems)
- PC Member of AISTATS'20 (The 23rd International Conference on Artificial Intelligence and Statistics)
- PC Member of ICLR'20 (The 8th International Conference on Learning Representations)
- PC Member of AAAI'20 (The 34th AAAI Conference on Artificial Intelligence)
- PC Member of DFM'19 (The 1st ICDM Workshop on Dynamic Feature Mining)
- Reviewer of PRICAI'19 (The 16th Pacific Rim International Conference on Artificial Intelligence)
- PC Member of CCML'19 (The 17th China Conference on Machine Learning)
- PC Member of NeurIPS'19 (The 33rd Annual Conference on Neural Information Processing Systems)
- PC Member of ICML'19 (The 36th International Conference of Machine Learning)
- PC Member of AISTATS'19 (The 22nd International Conference on Artificial Intelligence and Statistics)
- PC Member of ICLR'19 (The 7th International Conference on Learning Representations)
- PC Member of AAAI'19 (The 33rd AAAI Conference on Artificial Intelligence)
- PC Member of NeurIPS'18 (The 32nd Annual Conference on Neural Information Processing Systems)

Reviewer of Journals

- Reviewer of TPAMI (IEEE Transactions on Pattern Analysis and Machine Intelligence)
- Reviewer of TNNLS (IEEE Transactions on Neural Networks and Learning Systems)
- Reviewer of MIA (Medical Image Analysis)
- Reviewer of TII (IEEE Transactions on Industrial Informatics)
- Reviewer of Nature Methods
- Reviewer of Machine Learning
- Reviewer of TKDD (ACM Transactions on Knowledge Discovery from Data)
- Reviewer of Scientific Reports.
- Reviewer of KIS (Knowledge and Information Systems)
- Reviewer of FCS (Frontiers of Computer Science)

Other Services

- Web Chair of MLA'16 (The 14th Chinese Workshop on Machine Learning and Applications)
- Web Chair of MLA'15 (The 13rd Chinese Workshop on Machine Learning and Applications)
- Web Chair of LAMDA Group from Sept. 2014 to Sept. 2017 (Including updating each page of LAMDA website, maintaining mail server, maintaining course FTP etc.)

TEACHING ASSISTANTS

Advanced Methods and Health Applications in Machine Learning (for graduate students) Spring, 2024

Advanced Methods and Health Applications in Machine Learning (for graduate students) Spring, 2023

Computational Thinking (for undergraduate students) Fall, 2016

Introduction to Java (for undergraduate students) Fall, 2014

TECHNICAL STRENGTHS

Programming: Python, Java, C/C++, MATLAB, Bash, LaTex, HTML, CSS

Machine Learning Platforms: PyTorch, TensorFlow, Keras, Pandas, Scikit-Learn, SciPy

Other Tools: Word, Powerpoint, Excel, Origin, Mendeley

LEADERSHIP

2014 Class Representative of Department of Computer Science and Technology at Nanjing University Sep. 2010 - Jun. 2014

President of Graduate English Club of Nanjing University Sep. 2014 - Jun. 2016