

# BOJIAN HOU

(also Bo-Jian Hou) 4014 Pine St, Philadelphia, PA, 19104

Tel: (+1)814-826-8040 ◊ Email: bojian.hbj@gmail.com

## EDUCATION

---

**Nanjing University (NJU)**

*Department of Computer Science and Technology*

Supervisor: Prof. Zhi-Hua Zhou

**Ph.D.**

*Sep.2014-Jun.2020*

**Nanjing University (NJU)**

*Department of Computer Science and Technology*

**Bachelor of Science**

*Sep.2010-Jun.2014*

## RESEARCH INTEREST

---

I have broad interests in **machine learning** and **data mining**, and their potential applications to **biomedical data** such as medical images, medical literature, and electronic health records (EHR).

During my doctoral studies, I developed a novel learning scenario known as feature evolvable learning, where data features would evolve in an open and dynamic environment. The goal was to keep optimal online learning performance in dynamic feature space. I also studied semi-supervised learning and interpretability problems, such as storage-fit learning with unlabeled data and learning the interpretable structure from RNNs, respectively.

At my postdoctoral position, I mainly conducted multimodal survival analysis for medical images and clinical data, built natural language processing models to do medical literature mining, and investigated the potential issues of the interpretability methodologies for medical data.

In summary, my research interests include:

- **Interpretability**: studying the interpretability of black-box machine learning models.
- **Feature Evolvable Learning**: studying learning scenarios where data features evolve.
- **Multimodal Learning**: learning models by combining diverse data with different modalities.
- **Semi-Supervised Learning**: learning models from both labeled and unlabeled data.
- **Online Learning**: learning models from a continuous stream of data.
- **Natural Language Processing**: using pre-trained models to understand natural language.
- **Deep Learning**: leveraging deep neural networks to handle complex spatial and temporal data.

## WORK EXPERIENCE

---

**University of Pennsylvania**

*Center for Biomedical Image Computing and Analytics*

**Postdoctoral Researcher**

*Jan.2022-Present*

- Multimodal Survival Analysis for Lung Cancer Patients
  - Built a fully-parametric deep learning model on 3D lung cancer CT images as well as clinical data to predict risks of time-to-event recurrence and to cluster patients automatically
  - Improved the prediction performance, measured by concordance index, from 87% to 90%
  - Improved the clustering performance, quantified by log-rank, from 12.08 to 15.36

**Weill Cornell Medical College**

*Department of Population Health Sciences*

**Postdoctoral Associate**

*May.2021-Jan.2022*

- Natural Language Processing for Medical Literature

- Deployed a BERT-based literature mining system for clinical decision support
- Improved the relevant literature retrieval accuracy from 95% to 99%
- Improved the F1 score for information extraction from 68% to 73%
- Interpretability Study for Structured Clinical Report
  - Investigated unstable and inaccurate issues of Shapley value-based interpretable methodologies
  - Addressed possible confusions in medical guidance caused by the Shapley value flipping from 100 to -100 for repeated runs

#### 4th Paradigm Co.

*Department of Science and Technology*

**Research Scientist**

*Aug.2020-Apr.2021*

- Proposed a new learning paradigm called “Environment Learning”
  - Defined the virtual environment by several parameters and learned the virtual environment by data and knowledge hybridly for decision optimization
  - Reduced the designed validation loss from 13.19 to 2.67 in estimating the number of asymptomatic COVID-19 infection of Beijing

## 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 #2 in Artificial Intelligence by Google Scholar)
2. **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)
3. **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)
4. 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)
5. Heng Lian, John S. Atwood, **Bo-Jian Hou**, Jian Wu, Yi He. *Online Deep Learning from Doubly-Streaming Data*. In: **Proceedings of the 30th ACM International Conference on Multimedia (ACMMM’22)**, Lisbon, Portugal, 2022, in press. (Ranked #4 in Multimedia by Google Scholar)
6. **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 CCF AI Outstanding Student Paper Award)

### Journal

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

7. **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)
8. **Bo-Jian Hou**, Lijun Zhang, 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)
9. 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)
10. Mingquan Lin, **Bojian Hou**, Lei Liu, Mae Gordon, Michael Kass, Fei Wang, Sarah H. Van Tassel, Yifan Peng. *Automated diagnosing primary open-angle glaucoma from fundus image by simulating human's grading with deep learning*. **Scientific Reports**, 2022, in press. (Impact factor: 4.379)

### Manuscript

11. **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*. Submitted to **the annual ACM Conference on Intelligent User Interfaces (ACM IUI'22)**
12. Jing-Xiao Liao, **Bo-Jian Hou**, Hang-Cheng Dong, Hao Zhang, Jianwei Ma, Jinwei Sun, Shiping Zhang, Feng-Lei Fan. *Heterogeneous Autoencoder Empowered by Quadratic Neurons*. Submitted to **the 31st International Joint Conference on Artificial Intelligence (IJCAI'22)**
13. Dayang Wang, Feng-Lei Fan, **Bo-Jian Hou**, Hao Zhang, Rongjie Lai, Hengyong Yu, Fei Wang. *Manifoldron: Direct Space Partition via Manifold Discovery*. Submitted to **IEEE Transactions on Neural Networks and Learning Systems (TNNLS)**

### AWARDS & HONORS

---

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'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 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

---

- |   |            |
|---|------------|
| <b>Computational Thinking</b> (for undergraduate students). | Fall, 2016 |
| <b>Introduction to Java</b> (for undergraduate students).   | Fall, 2014 |

## TECHNICAL STRENGTHS

---

- |                                    |   |
|------------------------------------|---|
| <b>Programming:</b>                | Python, Java, C/C++, MATLAB, Bash, LaTeX, HTML, CSS     |
| <b>Machine Learning Platforms:</b> | PyTorch, Tensorflow, Keras, Pandas, Scikit-Learn, SciPy |
| <b>Other Tools:</b>                | Word, Powerpoint, Excel, Origin, Mendeley               |

## **LEADERSHIP**

---

<b>2014 Class Representative of Department of Computer Science and Technology at Nanjing University</b>	Sep. 2010 - Jun. 2014
<b>President of Graduate English Club of Nanjing University</b>	Sep. 2014 - Jun. 2016