Xiao Shou, PhD
xshou01@gmail.com | (614) 695-0856

Education

Rensselaer Polytechnic Institute, Troy NY
- School of Science: PhD in Applied Mathematics August 2023
- Thesis: Learning from Event Sequences
- **RPI – IBM AI Scholar (2020-2023)**
- Advisor: Kristin P. Bennett
- IBM Mentors: Dharmashankar Subramanian and Tian Gao
- Committee Members: Yangyang Xu & Peter Kramer

Rensselaer Polytechnic Institute, Troy NY
- School of Science: MS in Computer Science May 2023
- Advisor: Kristin P. Bennett

The Ohio State University, Columbus Ohio
- College of Arts & Sciences: MS in Chemistry August 2014
- Advisor: Heather C. Allen
- Thesis: Low Frequency and Total Internal Reflection Raman Spectroscopic Study of Ions in Bulk and at the Silica/Aqueous Interface

Wittenberg University, Springfield Ohio
- BA in Chemistry with highest distinction May 2012
- Minor in Applied Mathematics, Japanese and East Asia Studies

Past Experiences

Visiting Scholar Dr. Jianxi Gao Lab CS@Rensselaer Polytechnic Institute September 2023 – Present
- Research on network science and dynamics

AI Scholar IBM-AI Research August 2020 – August 2023
- Conducted machine learning research in graphical event models

Research Assistant Rensselaer Polytechnic Institute January 2019 – August 2020
- Conducted health informatics research on population health

Chemist Precision Labs April 2015 – June 2018
- Developed statistical & quality control models of quantifying drug level on medical devices

Product Safety Coordinator LBrands November 2014 – April 2015
- Performed information retrieval from databases of personal care products

Research Interests

My primary interests lie in the intersection of probabilistic machine/deep learning and dynamic systems, particularly in the domains of time series and temporal point processes. Additionally, I am also interested in causal machine learning, applied both in tabular and sequential data contexts, encompassing fields such as healthcare, recommendation systems, natural language understanding, and genomics, among others. Beyond modeling, I am equally intrigued by the challenges of structure learning and decision-making processes when dealing with sequential data, particularly for health systems.

Awards

- RPI-IBM AI Research Collaboration (AIRC) Scholar
- RPI-IBM AIRC Fellowship (2020-2023)
- Schmidt Science Fellowship Nomination (by RPI)
- ACM BCM 21 Best Student Paper Award

- Patterson Award for Outstanding Junior in Chemistry
- Midwest Regional College Math Competition Team First Place

Publications (Details @ https://shou-xiao.github.io/)

- **Shou, X**., Gao, T, Subramanian, D., Bhattacharjya, D. & Bennett, K. P. Pairwise Causality Guided Transformers for Event Sequences. NeurIPS 23.

- **Shou, X**., Bhattacharjya, D. Gao, T., Subramanian, D., Hassanzadeh, O., & Bennett, K. P. Probabilistic Attention-to-Influence Neural Models for Event Sequences. ICML 2023.

- Bhattacharjya D., Gao T., Subramanian, D., & **Shou, X**. Score-Based Learning of Graphical Event Models with Background Knowledge Augmentation. AAAI 23.

- **Shou, X**., Gao, T., Subramanian, D., Bhattacharjya, D. & Bennett, K. P. Multi-Label Event Prediction in Continuous Time. AAAI 23. (Oral)

- **Shou, X**., Gao, T., Subramanian, D., Bhattacharjya, D. & Bennett, K. P. Influence-Aware Attention for Multi-dimensional Temporal Point Process. Causal Learning and Reasoning Conference (CLeaR) 2023.

- Mavroudeas, G., Magdon-Ismail M., **Shou, X.,** and Bennett K. P. HMM-Boost: Improved Time Series State Prediction Via Supervised Hidden Markov Models: Case Studies in Epileptic Seizure and Complex Care Management.
    - Workshop on Data Mining in Biomedical Informatics and Healthcare 2022 (ICDM-DMBIH'22).
    - IEEE International Conference on Knowledge Graph (ICKG), 2022.

- Gao, T., Subramanian, D., Bhattacharjya, D., **Shou, X**., Mattei, N., & Bennett, K. Causal Inference for Event Pairs in Multivariate Point Processes. NeurIPS 2021.

- **Shou, X**., Gao, T., Subramanian, D., & Bennett, K. P. Match2: hybrid self-organizing map and deep learning strategies for treatment effect estimation. *ACM Conference on Bioinformatics, Computational Biology, and Health Informatics* (ACM BCB) 2021**. (*Best Student Paper Award*)**

- Mavroudeas, G., Neehal, N., **Shou, X.,** Magdon-Ismail, M., Kuruzovich, J., and Bennett K. P. Predictive Modeling for Complex Care Management. *IEEE International Conference on Bioinformatics and Biomedicine (*IEEE BIBM*)*, 2021.

- **Shou, X.,** Mavroudeas, G., Magdon-Ismail, M., Figueroa, J., Kuruzovich, J. N., & Bennett, K. P. Supervised mixture of expert models for population health. *Methods*, *179*, 101-110, 2020.

- **Shou, X.,** Mavroudeas, G., New, A., Arhin, K., Kuruzovich, J. N., Magdon-Ismail, M., & Bennett, K. P. Supervised mixture models for population health. *IEEE International Conference on Bioinformatics and Biomedicine (*IEEE BIBM*)*, 2019.

Manuscripts Under Review

- **Shou, X**., Subramanian, D., Bhattacharjya, D., Gao, T. & Bennett, K. P. Self-Supervised Contrastive Pre-Training for Multivariate Point Processes. Under review.

Manuscripts In Preparation

- Xiao Shou. Logical Graphical Event Models.
- Xiao Shou. Causal Transformer for Well-Timed Sequential Recommendation.

Teaching Experience

Department of Chemistry and Biochemistry, The Ohio State University

- Physical Chemistry I: August 2012
- General Chemistry I & II: January 2013 – August 2013

Department of Mathematics, Rensselaer Polytechnic Institute

- Calculus I: Fall 2018
- Introduction to Data Mathematics (with R): Spring 2019

Professional Development

- Mentor services:
    - Co-Mentor (with Prof. Kristin P. Bennett) Hannah Power (Accelerated BS/PhD, RPI) & Marguerite Demasi (Undergraduate, RPI)
      Project: Visualizing matched representations for causal inference via R shiny Fall 2021,
    - Aaron Green (RPI PhD)
      Project: clustering event streams, Spring 2022, RPI.

- Review services: AMIA 22, AISTATS 23, NeurIPS 23, TNNLS, ICLR 23, CLeaR 24.

Presentations & Talks

Invited Talks

- Department of AI Automation, Planning, and Performance. IBM AI Research, Yorktown Heights. Pairwise Causality Guided Transformers for Event Sequences. October 26, 2023. (Virtual Presentation)
- Dumitrascu Lab. Causal Pairs in Event Sequences. Department of Statistics & Irving Institute for Cancer Dynamics, Columbia University, October 10, 2023

IBM Workshops

- IBM-RPI AI Research Collaborations (AIRC) Scholar and Project Talk Series: Learning from Event Streams. May 12, 2023. (Virtual Presentation)
- IBM RPI AIRC Fall Workshop Poster Presentation: Neural Temporal Point Processes - A Self-Supervised Learning Paradigm. IBM Yorktown Heights Research Center, November 16, 2022.
- IBM RPI AIRC scholarly presentation: Learning and Causal Inference in Marked Temporal Point Processes. Virtual event. November 18, 2022.

RPI Campus Presentation Talks

- RPI Accepted Students Day Poster Presentation: Learning and Inference from Temporal Event Sequences. March 24, 2023.
- IDEA Community Talk: Learning and Inference of Temporal Event Sequences. Rensselaer Polytechnic Institute, January 31 2023.
- RPI CS Poster Presentation: Event-former: A Self-supervised Learning Paradigm for Temporal

Point Processes, Rensselaer Polytechnic Institute, December 2, 2022.

Conference Presentations & Tutorials

- IJCAI 23. Tutorial on Graphical Event Models. Macao, China. August 19-25, 2023.
- ICML 23. Probabilistic Attention-to-Influence Neural Models for Event Sequences. Honolulu, HI. July 23-29.
- CLeaR 23: Influence-Aware Attention for Multi-dimensional Temporal Point Process. Tübingen, Germany, April 11-14, 2023. (Virtual Presentation)
- AAAI 23: Multi-Label Event Prediction in Continuous Time. Washington DC. February 7-14, 2023.
- ACM BCB 21: Match2: hybrid self-organizing map and deep learning strategies for treatment effect estimation. August 3, 2021. (Virtual Presentation)

Programming Languages

- Mathematical Programming: AMPL & Matlab
- Logical Programming: Prolog
- Machine Learning Software: Python & C & C++
- Deep Learning Software: Pytorch and Tensorflow
- Statistical Software: R & SAS (Base, Statistical Business Analyst, Enterprise Miner Certificate)
- Functional Programming: Scheme & Haskell

Patents

Match2: Supervised Similarity Learning for Covariate Matching and Treatment Effect Estimation via Self-Organizing Maps
patent application number:17348492
publication date: 2022/12/15
patent URL: Patent

Collaborators

- Industry: Dr. Dharmashankar Subramanian (IBM), Dr. Tian Gao (IBM), Dr. Debarun Bhattacharjya (IBM), Dr. Georgios Mavroudas (Meta, USA)
- Academia: Dr. Kristin P. Bennett (RPI), Dr. Malik Magdon-Ismail (RPI), Dr. Feng-lei Fan (The Chinese University of Hong Kong), Dr. Yangyang Xu (RPI), Dr. Jianxi Gao (RPI)