Ruijiang Gao

(734)272-9935 • ruijiang@utexas.edu• 1628 W 6th St • Austin, TX 78703

EDUCATION

- PhD in Information, Risk and Operation Management, University of Texas at Austin 2018 - 2023(Expected)
- Master of Statistics, University of Michigan 2016-2018
- B.S. Statistics (School of the Gifted Young), University of Science and Technology of China 2012-2016

PUBLICATIONS (MANUSCIPTS WILL BE SHARED UPON REQUEST)

1. Ruijiang Gao, Max Biggs, and Wei Sun. Loss functions for contextual pricing. Submitted, 2021
2. Ruijiang Gao and Maytal Saar-Tsechansky. Active incentive learning. Submitted, 2021
3. Ruijiang Gao, Maytal Saar-Tsechansky, Maria De-Arteaga, Ligong Han, Min Kyung Lee, and Matthew Lease. Human-AI Collaboration with Bandit Feedback. *IJCAI*, 2021
4. Ruijiang Gao and Han Feng. Identifying best fair intervention. Submitted, 2021
5. Ruijiang Gao, Max Biggs, Wei Sun, and Ligong Han. Enhancing counterfactual classification performance via self-training. Submitted, 2021
6. Cost-Accuracy Aware Adaptive Labeling for Active Learning, author=Gao, Ruijiang and Saar-Tsechansky, Maytal, booktitle=Proceedings of the AAAI Conference on Artificial Intelligence, volume=34, number=03, pages=2569–2576, year=2020
7. Ligong Han, Ruijiang Gao, Mun Kim, Xin Tao, Bo Liu, and Dimitris N Metaxas. Robust Conditional GAN from Uncertainty-Aware Pairwise Comparisons. In *AAAI*, pages 10909–10916, 2020
8. Ligong Han, Yang Zou, Ruijiang Gao, Lezi Wang, and Dimitris Metaxas. Unsupervised Domain Adaptation via Calibrating Uncertainties. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 99–102, 2019
9. Ligong Han, Martin Renqiang Min, Anastasis Stathopoulos, Yu Tian, Ruijiang Gao, Asim Kadav, and Dimitris Metaxas. Dual projection generative adversarial networks for conditional image generation. *arXiv preprint arXiv:2108.09016*, 2021

PROFESSIONAL EXPERIENCE

- **IBM: Research Internship** Yorktown Heights, 2021/06-2021/08
  - Developed Human-AI algorithms considering human overriding behaviors.
  - Bridged gap between causal inference, learning from supervision and proposed new estimators for contextual / personalized pricing.
- **IBM: Research Internship** Yorktown Heights, 2020/06-2020/08
  - Developed novel algorithm based on self-training for counterfactual inference given only observational data for applications like pricing, precision medicine and ads placement.
  - Used theoretical analysis to demonstrate how self-training helps counterfactual learning.
  - Showed state-of-the-art performance on synthetic and real datasets.
  - Applied domain knowledge like monotonicity to further improve our algorithm.
- **Tencent: Data Scientist Internship** Shenzhen, 2018/04-2018/07
  - Worked at Tencent Social Network Group using machine learning algorithms to learn better about customers.
  - Built retention models for Tencent ESports users.
  - Used emoji and bullet screen to cluster short videos for auto-tagging.
- **Amazon: Business Intelligence Internship** Seattle, 2017/06-2017/09
  - Worked at Amazon Prime BI team using machine learning algorithms to learn better about customers.
  - Used Gaussian Mixture Model to study customers' behaviors and clustered customers into hierarchical structures.

- **University of Texas at Austin: Research Assistant** Austin, 2018/09-Present
  - **Human-AI Collaboration with Bandit Feedback**
    - Propose and develop a solution for a novel human-machine collaboration problem in a bandit feedback setting.
    - Extend our approach to settings with multiple human decision makers.
    - Demonstrate the effectiveness of our proposed methods using both synthetic and real human responses.
  - **Identifying Best Fair Intervention**
    - Define a counterfactual fairness on revenue with respect to a binary sensitive attribute.
    - Find the best (soft) intervention in a given causal graph meeting the fairness constraint required.
    - Theoretically prove the exponential decrease rate of probability of error.
    - Empirically examined the effectiveness of proposed method using synthetic and real datasets.
  - **Adaptive Labeling Payment Selection from Noisy Workers with Varying Cost**
    - Select payment and worker pair for active learning in crowdsourcing or dataset construction to improve model performance under a budget constraint.
    - Identify the shortcoming of existing method.
    - Use generalization bound for learning from noisy labels to guide the selection process to select payment more efficiently.
  - **Active Incentive Learning**
    - Select payment for active learning in crowdsourcing platform to improve auxiliary model performance under a budget constraint.
    - Use expected error reduction to estimate payment's effect on generalization error using loss correction from learning from noisy supervision literatures.
    - Propose a novel payment utility estimation method to calibrate biased estimation in existing method.
  - **Conditional GAN from Uncertainty-Aware Pairwise Comparisons**
    - Use weak supervision (pairwise comparisons) to learn attribute strength (e.g. age) from images.
    - Utilize robust conditional GAN for image attribute editing.
    - Qualitative and quantitative results show proposed method has a comparable performance with fully-supervised methods and is better than unsupervised methods.
  - **Unsupervised Domain Adaptation via Calibrating Uncertainties**
    - Propose Renyi Entropy regularization for unsupervised domain adaptation.
    - Propose Gradient Variance Regularization for entropy regularization.
    - Empirically show the effectiveness of proposed method on MNIST, USPS, SVHN and VisDA17.

TEACHING EXPERIENCE

- Teaching Assistant for INTRODUCTION TO DATA SCIENCE. Fall 2020, Spring 2021
- Teaching Assistant for DATABASE MANAGEMENT. Spring, 2020
- Teaching Assistant for PREDICTIVE ANALYSIS AND DATA MINING. Spring, 2019
- Teaching Assistant for STRATEGIC INFORMATION TECHNOLOGY MANAGEMENT. Fall, 2018
- Teaching Assistant for STRATEGIES FOR NETWORKED ECONOMY. Fall, 2018

COMPUTER SKILLS

- Python, R, Matlab, SQL, SAS, Mathematica, Spark, Hive, Hadoop, Linux, Excel, imacros, LaTeX, Bloomberg