Raha **Moraffah**

Causal Responsible Machine Learning, Data Science & Causal Learning

+1-480-652-3040 | Raha.Moraffah@asu.edu | https://rmoraffa.github.io | Google Scholar

Education

Arizona State University — *Advisor: Professor Huan Liu*

Doctor of Philosophy, Computer Science — *Jan. 2017 - Present*

Professional Experience

Data Mining and Machine Learning Lab, ASU — *Graduate Research Assistant*

Advisor: Professor Huan Liu — *Mar. 2017 - Present*

Army Research Lab (ARL) — *Graduate Researcher*

Supervisor: Dr. Adrienne Raglin — *May 2018 - Present*

USC Information Sciences Institute, USC — *Research Intern*

Supervisor: Dr. Kristina Lerman — *May 2018 - Sep. 2018*

Speech Processing Laboratory (SPL), Sharif University of Technology — *Undergraduate Researcher*

Supervisor: Dr. Hossein Sameti — *Aug. 2013 - Aug. 2015*

Research Interests

Responsible Machine Learning (esp. security and robustness)

- Generative AI and Foundation models
- Discriminative Deep Models

Causal Inference (causality for responsible Machine Learning)

- Causality for security and robustness
- Causal Machine Learning

AI for Social Good

- AI-generated text identification, hate speech detection, and disaster relief
- Cybersecurity
- Health care

Publications

Raha Moraffah has co-authored over 30 papers in top-tier data mining, machine learning, and Natural Language Processing venues. As of 12/2023, her work has been cited over 430 times, with h-index = 8. Below are her selected publications: † denotes the author is her mentee; ∗ indicates an equal contribution.

Books

[1] **Raha Moraffah**, Saketh Vishnubhatla, and Huan Liu. "Causal Robust Machine Learning". *Spring Briefs* (2023).

[2] **Raha Moraffah**, Shu Wan, and Huan Liu. "Machine Learning for Causal Inference". *Springer Nature* (2023).

Conferences , Journals , and Submissions

* indicates equal contributions

† indicates the author is her mentee

[3] **Raha Moraffah**, Shubh Khandelwal†, Amrita Bhattacharjee†, and Huan Liu. "Adversarial Text Purification: A Large Language Model Approach for Defense". *PAKDD* (2024).

[4] **Raha Moraffah** and Huan Liu. "A Generative Approach to Surrogate-based Black-box Attacks". *Submitted to IJCAI* (2024).

[5] **Raha Moraffah** and Huan Liu. "Exploiting Class Probabilities for Black-box Sentence-level Attacks". *EACL* (2024).
[6] **Raha Moraffah**, Paras Sheth[*†], Saketh Vishnubhatla[*†], and Huan Liu. "Causal Feature Selection for Responsible Machine Learning". *Submitted to IJCAI* (2024).
[7] **Raha Moraffah**, Suraj Jyothi Unni, and Huan Liu. "Domain Generalization for VQA: A Cross-prompt Approch". *Submitted to ECCV* (2024).
[8] **Raha Moraffah**, Isabel Valera, and Huan Liu. "Adversarial Transferability Through the Lens of Causality". *Submitted to ICML* (2024).
[9] **Raha Moraffah**, Chaowei Xiao, and Huan Liu. "What Should They Look Like? Reinforcing Surrogate-based Black-box Attacks with Distribution Feedback". *Submitted to CVPR* (2024).
[10] Paras Sheth[†], **Raha Moraffah**, Tharindu Kumarage[†], Aman Chadha, and Huan Liu. "Causality Guided Disentanglement for Cross-Platform Hate Speech Detection". *ACM International Conference on Web Search and Data Mining (WSDM)* (2024).
[11] Suraj Jyothi Unni[†], **Raha Moraffah**, and Huan Liu. "VQA-GEN: A Visual Question Answering Benchmark for Domain Generalization". *Submitted to SIGIR* (2024).
[12] Amrita Bhattacharjee[†], Tharindu Kumarage[†], **Raha Moraffah**, and Huan Liu. "ConDA: Contrastive Domain Adaptation for AI-generated Text Detection". *IJCNLP-AACL* (2023), Outstanding Paper Award.
[13] Amrita Bhattacharjee[†], **Raha Moraffah**, Joshua Garland, and Huan Liu. "Towards LLM-guided Causal Explainability for Black-box Text Classifiers". *AAAI-ReLM* (2023).
[14] Amrita Bhattacharjee[†], **Raha Moraffah**, Joshua Garland, and Huan Liu. "Zero-shot Counterfactual Explanation Generation with LLMs for Text Classifiers". *Submitted to LREC-Colling* (2023).
[15] Tharindu Kumarage[†], Paras Sheth[†], **Raha Moraffah**, Joshua Garland, and Huan Liu. "How Reliable Are AI-Generated-Text Detectors? An Assessment Framework Using Evasive Soft Prompts". *EMNLP* (2023).
[16] **Raha Moraffah**, Amir-Hossein Karimi, Adrienne Raglin, and Huan Liu. "Socially Responsible Machine Learning: A Causal Perspective". *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)* (2023).
[17] Paras Sheth[*†], Tharindu Kumarage[*†], **Raha Moraffah**, Aman Chadha, and Huan Liu. "Peace: Cross-platform hate speech detection-a causality-guided framework". *Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML-PKDD)* (2023).
[18] Lu Cheng, Ruocheng Guo, **Raha Moraffah**, Paras Sheth, K Selçuk Candan, and Huan Liu. "Evaluation methods and measures for causal learning algorithms". *IEEE Transactions on Artificial Intelligence* (2022).
[19] **Raha Moraffah** and Huan Liu. "Query-Efficient Target-Agnostic Black-Box Attack". *IEEE International Conference on Data Mining (ICDM)* (2022).
[20] **Raha Moraffah**, Paras Sheth[†], and Huan Liu. "Exploring the Target Distribution for Surrogate-Based Black-Box Attacks". *IEEE International Conference on Big Data (Big Data)* (2022).
[21] **Raha Moraffah**, Suraj Jyothi Unni[†], Adrienne Raglin, and Huan Liu. "Causal Data Fusion for Multimodal Disaster Classification in Social Media". *SBP-BRiMS* (2022).
[22] Paras Sheth[†], **Raha Moraffah**, K Selçuk Candan, Adrienne Raglin, and Huan Liu. "Domain Generalization–A Causal Perspective". *arXiv preprint arXiv:2209.15177* (2022).
[23] **Raha Moraffah**, Paras Sheth[†], Mansooreh Karami[†], Anchit Bhattacharya, Qianru Wang[†], Anique Tahir, Adrienne Raglin, and Huan Liu. "Causal inference for time series analysis: Problems, methods and evaluation". *Knowledge and Information Systems (KAIST)* (2021).
[24] Bahman Moraffah, Christ Richmond, **Raha Moraffah**, and Antonia Papandreou-Suppappola. "Metric-bayes: Measurements estimation for tracking in high clutter using bayesian nonparametrics". *Asilomar Conference on Signals, Systems, and Computers* (2020).
[25] Bahman Moraffah, Christ Richmond, **Raha Moraffah**, and Antonia Papandreou-Suppappola. "Use of bayesian nonparametric methods for estimating the measurements in high clutter". *arXiv preprint arXiv:2012.09785* (2020).
[26] **Raha Moraffah**, Mansooreh Karami[†], Ruocheng Guo, Adrienne Raglin, and Huan Liu. "Causal interpretability for machine learning-problems, methods and evaluation". *ACM SIGKDD Explorations* (2020).
[27] **Raha Moraffah**, Bahman Moraffah, Mansooreh Karami, Adrienne Raglin, and Huan Liu. "CAN: A Causal Adversarial Network for Learning Observational and Interventional Distributions". *https://arxiv.org/abs/2008.11376* (2020).

[28] Adrienne Raglin, **Raha Moraffah**, and Huan Liu. "Causality and Uncertainty of Information for Content Understanding". *IEEE International Conference on Cognitive Machine Intelligence (CogMI)* (2020).

[29] Lu Cheng, Ruocheng Guo*, **Raha Moraffah***, K Selçuk Candan, Adrienne Raglin, and Huan Liu. "A practical data repository for causal learning with big data". *Benchmarking, Measuring, and Optimizing: Second BenchCouncil International Symposium, Bench* (2019).

[30] **Raha Moraffah**, Kai Shu, Adrienne Raglin, and Huan Liu. "Deep causal representation learning for unsupervised domain adaptation". *arXiv preprint arXiv:1910.12417* (2019).

[31] Vineeth Rakesh*, Ruocheng Guo*, **Raha Moraffah**, Nitin Agarwal, and Huan Liu. "Linked causal variational autoencoder for inferring paired spillover effects". *ACM International Conference on Information and Knowledge Management (CIKM)* (2018).

[32] Mohamed Sarwat, **Raha Moraffah**, Mohamed F Mokbel, and James L Avery. "Database system support for personalized recommendation applications". *International Conference on Data Engineering (ICDE)* (2017).

[33] Jia Yu, **Raha Moraffah**, and Mohamed Sarwat. "Hippo in action: Scalable indexing of a billion new york city taxi trips and beyond". *IEEE 33rd International Conference on Data Engineering (ICDE)* (2017).

Peer-reviewed Posters (with poster papers and abstracts)

[34] **Raha Moraffah**, Suraj Jyothi Unni, Adrienne Raglin, and Huan Liu. "CAUSEMMD: A multi-modal classification platform for disaster relief". *SBP-BRiMS* (2022).

[35] **Raha Moraffah**, Lu Cheng, Roucheng Guo, and Huan Liu. "CAUSE: A data repository for causal inference from observational data". *SBP-BRiMS* (2019).

Proposal Writing

DISCOVER: Decision Intelligence System for Causal Exploration and Verification

PI: Dr. Huan Liu

- Designed and wrote two out of three tasks: "Causal Decision Exploration before Action" and "Causal Action Analysis after Action"
- Submitted to AMC in 2023

Elements: CausalBench: A Cyberinfrastructure for Causal-Learning Benchmarking for Efficacy, Reproducibility, and Scientific Collaboration

PI: Dr. Selcuk Candan, Co-PI: Dr. Huan Liu

- Designed and wrote one task out of 4 tasks "Open Sourcing and Validation"
- NSF (#2311716)

High-Performance Computing Cluster for NLP and Deep Learning Approaches to Understanding Social, Psychological, and Rhetorical Dimensions of Influence in Cross-Cultural Contexts

PI: Dr. Scott Ruston, Co-PI: Dr. Huan Liu

- Design and wrote one task out of 3 for "Benchmarking Causal Discovery"
- DURIP ONR

Characterizing and Understanding of Dynamic Time-Series Information for Decision Making: A Causal Learning Approach

PI: Dr. Huan Liu

- Designed and wrote three tasks out of five tasks "Causal discovery for time series", "Causal inference for time series", and "Causal Interpretable models for time series prediction"
- AMC (ARL W911NF2020124)

Knowing the Unknown: Exploring the Space of Adversarial Attacks via Causal Learning

PI: Dr. Huan Liu

- Designed and wrote 2 out of 4 tasks "Causal Learning for Exploring the Space of Adversarial Examples" and "Coverage of Causal Learning and Transferability of Adversarial Examples"
- AMC (ARO W911NF2110030)

Presentation

Invited Talks at Universities and Summer Schools

- Pragmatic Responsible ML: Security, Robustness, and Causality, Worcester Polytechnic Institute — Jan. 2023
- Pragmatic Responsible ML: Security, Robustness, and Causality, University of Texas at San Antonio — Dec. 2023
- Causal Adversarial Attacks: A New Frontier, Max Plank Institute — Oct. 2023
- Pragmatic Responsible ML: From Foundation to Application, Sharif University of Technology — Aug. 2023
- A mini-lecture series on Introduction to Causal Inference, ASU — Aug. 2023
- Causal Machine Learning: Issues, Purposes, and Challenges, Sharif University of Technology — Jul. 2022
- Lecture series on Causal Machine Learning for Adversarial Robustness, ASU — Nov. 2023
- A mini-lecture series on Causal Learning from Observational Data, ASU — Oct. 2023
- Lectures on Causal Inference, Army Research Lab — Mar. 2018- Present

Keynote/Invited Talks at Conferences and Workshops

- Responsible for Responsible Generative AI and Causality Discussion Sessions, Workshop on Generative AI and Large Language Models: Opportunities and Challenges for Healthcare and Medicine, ASU-Mayo Clinic — Nov. 2023
- Tutorial on Socially Responsible Machine Learning: A Causal Perspective, KDD 2023 — Aug. 2023
- PhD Doctrol Forumn, SDM 2023 — Nov. 2023
- Tutorial on Machine Learning for Causal Inference, SBP-BRiMS 2022 — Sep. 2022

Students and Mentees

Brian Vincent, Ph.D. student

Data Mining and Machine Learning Lab, ASU — *Aug. 2023- Present*

- Research: Causal Feature Selection with Application in E-commerce

Shu Wan, Ph.D. student

Data Mining and Machine Learning Lab, ASU — *July 2023- Present*

- Research: Large language models for causal discovery: A data-driven approach

Paras Sheth, Ph.D. student

Data Mining and Machine Learning Lab, ASU — *Sep. 2022- Present*

- Co-authored 4 papers, one accepted to IEEE Bigdata 2022, one accepted to ECML-PKDD 2023, one accepted to WSDM 2024, and one submitted to JAIR.

Kumarage Tharindu Kumarage, Ph.D. student

Data Mining and Machine Learning Lab, ASU — *Sep. 2022- Present*

- Co-authored "How Reliable Are AI-Generated-Text Detectors? An Assessment Framework Using Evasive Soft Prompts", Accepted to EMNLP 2023

Amrita Bhattacharjee, Ph.D. student

Data Mining and Machine Learning Lab, ASU — *Sep. 2022- Present*

- Co-authored 4 papers, one accepted and won the outstanding paper award at IJCNLP-AACL 2023, one accepted to AAAI-ReLM, one submitted to Lrec-Coling 2024, and one submitted to PAKDD
- Won best paper award

Ali Mohammad Beigi, Ph.D. student

Data Mining and Machine Learning Lab, ASU — *Sep. 2022- Present*

- Causal feature selection

Mansooreh Karami, Ph.D. student

Data Mining and Machine Learning Lab, ASU — *Jan 2019-Nov.2023*

- Co-authored 3 papers, one accepted at SigKDD Explorations, one accepted to KAIST 2021
- Joined Microsoft Research as a Data Scientist

Qianru Wang, Visiting Ph.D. student

School of Computer Science, Northwestern Polytechnical University — *Sep. 2019-Sep. 2020*

- Co-authored one paper accepted to KAIST 2021

Anchit Bhattacharya, Ph.D. student

Data Mining and Machine Learning Lab, ASU — *Jun. 2019-Sep. 2020*

- Co-authored one paper accepted to KAIST 2021
- Joined University of Kentucky as a Ph.D. Student

Saketh Vishnubhatla, Master's student

Data Mining and Machine Learning Lab, ASU — *Sep. 2023- Present*

- Research: Causal feature selection for responsible ML
- Couatheored one paper to be submitted to IJCAI 2024

Somayeh Khazaei, Visiting Master's student

University of Luxembourg — *Mar. 2023- Present*

- Research: Causal adversarial attacks

Suraj Jyoti Unni, Master's student

Data Mining and Machine Learning Lab, ASU — *Feb. 2019- Present*

- Co-authored two papers, one submitted to SIGIR 2024, and one submitted to ECCV 2024
- Successfully defended his master's and cofounded a company

Ishita Mathur, Undergraduate student

California Institute of Technology (Caltech) — *Jan. 2023- Present*

- Research: Textual adversarial attacks

Shubh Khandelwal, Undergraduate student

Arizona State University (ASU) — *Mar. 2023- Present*

- Research: Textual adversarial defense
- Coauthered a paper submitted to PAKDD 2024

Ryan Schwartz, Undergraduate student

Arizona State University (ASU) - NSF REU — *Jan. 2019- Sep. 2020*

- Research: Benchmarking adversarial attack

Qiang Fu, Undergraduate student

Arizona State University (ASU) - Furi — *Jan. 2019- Sep. 2019*

- Received the ASU Fulton School of Engineering Impact Award for his dissertation on "Benchmarking adversarial attack"
- Joined Google as a Software Engineer

Teaching Experience

CSE 472: Social Media Mining — *Arizona State University*

Guest Lecturer - Taught 7 lectures on causal inference in ML — *Fall 2020- Fall 2023*

Socially Responsible Machine Learning: A Causal Perspective — *KDD 2023*

Conference Tutor — *Fall 2023*

Machine Learning for Causal Inference — *SBP-BRiMS 2022*

Conference Tutor — *Fall 2023*

CSE 472: Social Media Mining — *Arizona State University*

Teaching Assistant — *Fall 2019- Fall 2023*

CSE 340: Principles of Programming Languages — *Arizona State University*

Teaching Assistant — *Fall 2017- Fall 2019*

Introduction to Java — *Sharif University of Technology*

Instructor — *Fall 2013-2015*

Editorial Boards and Reviewing Activities

Senior Program Committee

Association for the Advancement of Artificial Intelligence (AAAI, Main Track) (2024), ACM International Conference on Web Search and Data Mining (WSDM) (2023)

Program Committee Member

ICLR(2020-2024), NeurIPS(2020-2023), ICML(2021-2024), CVPR (2022-2023), AAAI (2020 - 2023), ACM SIGKDD (2020-2023), ICDM (2020-2023), IEEE Big data (2020-2023), (ECML-PKDD) (2020-2023), (SBP-BRiMS) (2019-2023), IEEE CogMI (2021-2023), ACL ARR (2022 - 2023), IJCAI (2021 - 2023), ACM MM (2021 - 2022), (CIKM) (2019-2020), WSDM (2022-2023), ICRA (2022), ECCV (2022), ICME (2022), EMNLP (2022), ICCV (2021), ACM SIGIR (2020), etc.

Journal Review

JMLR, TPAMI, TIP, TCSVT, TNNLS, TMLR, TC, IPM, MACH, TKDE, ISIT, IEEE Access, Computers & Security (COSE), IEEE Transactions on Computational Social Systems (TCSS), IEEE Transactions on Artificial Intelligence (TAI), ACM Computing Survey (CSUR), ACM Transactions on Knowledge Discovery from Data (TKDD), Data & Policy, etc.

Awards & Honors

2023	Best Paper Award , ASU SCAI Symposiom	*ASU*
2023	Outstanding Paper Award , IJCNLP-AACL	*Bali, Indonasia*
2022-2023	Doctoral Fellowship award , ASU SCAI	*ASU*
2022-2023	Travel award , ASU SCAI	*ASU*
2023	Travel Grant Award , KDD	*Long Beach, CA*
2023	Doctoral Forum Poster First Runner Up , SDM	*Minneapolis, MN*
2023	Travel Grant Award , SDM	*Minneapolis, MN*
2022	Travel Grant Award , ICDM	*Orlando, FL*
2022-2023	GPSA Individual Travel Funding	*ASU*
2022	Travel Grant Award , SBP-BRiMS	*Pittsburgh, PA*
2020	Graduate Outstanding Research Award	*ASU*
2019	Travel Grant Award , SBP-BRiMS	*Washington, D.C.*
2018	Grace Hopper Celebration Scholarship	*ASU*
2017	Travel Grant Award , ICDE	*San Diego, CA*
2014-2015	Outstanding Undergraduate Award	*Tehran, IR*
2011-2015	Undergraduate Scholarship	*Tehran, IR*
2010	Ranked 12^{th} out of 600,000 in Mathematics Nationwide Entrance Qualification Exam (Konkoor)	*Tehran, IR*
2009	Silver Medalist of Iranian National Mathematics Olympiad	*Tehran, IR*
2008-Present	Honorary Member of the Council of Young Mathematician Society	*Tehran, IR*
2008	Bronze Medal in Fryer, Galois, and Hypatia Mathematics Contests , University of Waterloo	*Waterloo, CA*
2007-Present	Member of Young Mathematician Society	*Tehran, IR*

References

Professor Huan Liu, Fellow of ACM, AAAI, AAAS, IEEE, ASEE, and SIAM

- Regents professor, Computer Science, Arizona State University
- Phone: +1-480-727-7349
- Email: huanliu@asu.edu
- Homepage: http://www.public.asu.edu/ huanliu/

Professor Kathleen M. Carley

- Professor, Computer Science, courtesy appointments at Electrical and Computer Engineering, Carnegie Mellon University
- Phone: +1-412-268-6016
- Email: kathleen.carley@cs.cmu.edu
- Homepage: http://www.casos.cs.cmu.edu/bios/carley/carley.html

Professor Chaowei Xiao

- Assistant Professor, University of Wisconsin at Madison, Research Scientist, NVIDIA
- Phone: +1-608-263-2900
- Email: cxiao34@wisc.edu
- Homepage: https://xiaocw11.github.io/

Dr. Adrienne Raglin

- ArtIAMAS co-Lead of the AI-Enabled Decision Making in Multiple Domains Capability, Team Lead for the Artificial Reasoning team, Army Research Laboratory
- Phone: +1-301-394-0210
- Email: adrienne.raglin2.civ@army.mil
- Homepage: https://artiamas.umd.edu/adrienne-raglin/