

Yangxinyu Xie

<https://xieyangxinyu.github.io/>

+1 (737) 207 2130

xinyux@wharton.upenn.edu

Education

- 2022 – **Ph.D. Student**, Wharton School, University of Pennsylvania, *Statistics*.
Advisor: Weijie J. Su
- 2024 – **Visiting Student**, *Mathematics and Computer Science*, Argonne National Laboratory.
Advisor: Tanwi Mallick
- 2021 – 2022 **Master of Science**, University of Texas at Austin, *Computer Science*.
Advisors: Anna Gál and Ngoc Mai Tran
Thesis: *Matrix Rigidity: A Survey*
- 2017 – 2021 **Bachelor of Science**, University of Texas at Austin, *Mathematics, Computer Science*.

Publications

- In Preparation
- Xie, Y., Li, X., Mallick, T., Su, W. J. and Zhang, R. Unbiased Watermark for Large Language Models via Importance Sampling.
 - Bhattacharya, B. and Xie, Y. Joint Convergence of Monochromatic Edges in Multi-plex Hypergraphs.
- Preprint
- Baptista, R., O'Reilly, E. and Xie, Y. TrIM: Transformed Iterative Mondrian Forests for Gradient-based Dimension Reduction and High-Dimensional Regression. arXiv preprint arXiv:2407.09964.
 - Jiang, B., Xie, Y., Wang, X., Su, W.J., Taylor, C.J. and Mallick, T., 2024. Multi-Modal and Multi-Agent Systems Meet Rationality: A Survey. arXiv preprint arXiv:2406.00252.
 - Xie, Y., Jiang, B., Mallick, T., Bergerson, J., Hutchison, J., Verner, D., Branham, J., Alexander, R., Ross, R., Feng, Y., Levy, L.-A., Su, W. J. and Taylor, C. J. WildfireGPT: Tailored Large Language Model for Wildfire Analysis. arXiv preprint arXiv:2402.07877.
 - Xie, Y. and Mallick, T. A Comparative Study of Loss Functions: Traffic Predictions in Regular and Congestion Scenarios. arXiv preprint arXiv:2308.15464.
- 2024
- Jiang, B., Xie, Y., Hao, Z., Wang, X., Taylor, C. J., Su, W. J., Mallick, T. and Roth, D. A Peek into Token Bias: Large Language Models Are Not Yet Genuine Reasoners. In Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP).

- 2023 ○ Krenn, M., Buffoni, L., Coutinho, B., Eppel, S., Foster, J.G., Gritsevskiy, A., Lee, H., Lu, Y., Moutinho, J.P., Sanjabi, N., Sonthalia, R., Tran, N., Valente, F., Xie, Y., Yu, R., and Kopp, M. Forecasting the future of artificial intelligence with machine learning-based link prediction in an exponentially growing knowledge network. *Nature Machine Intelligence*, pp.1-10. <https://doi.org/10.1038/s42256-023-00735-0>
- 2022 ○ Sun, A.Y., Jiang, P., Yang, Z.L., Xie, Y. and Chen, X. A graph neural network (GNN) approach to basin-scale river network learning: the role of physics-based connectivity and data fusion. *Hydrology and Earth System Sciences*, 26(19), pp.5163-5184.
- 2021 ○ Tran NM, Xie, Y. Improving random walk rankings with feature selection and imputation. In 2021 IEEE International Conference on Big Data (Big Data) 2021 Dec 15 (pp. 5824-5827). IEEE.
- Xie, Y., Kulpanowski, D., Ong, J., Nikolova, E., Tran, NM. Predicting Covid-19 emergency medical service incidents from daily hospitalization trends. *Int J Clin Pract.* 2021; 00:e14920. doi:10.1111/ijcp.14920
- Xie, Y. On 2×2 Tropical Commuting Matrices. *Linear Algebra and its Applications*, 620, 92-108.

Public Scholarship

- 2024 ○ Xie, Y., Pizzurro, J., Santosa, F. and O'Reilly, E., 2024. An assessment of the impact of the Baltimore Red Line on job accessibility. Press release, Johns Hopkins University, 7 June. Available at: <https://hub.jhu.edu/2024/06/07/red-line-job-access-report/>

Competitions and Awards

- 2024 **Load Disaggregation Challenge: Energy Use in Buildings, 2nd Place, ADRENALIN Energy Competitions.**
Team Members: Ngoc Mai Tran, Yangxinyu Xie
- 2023 **CityLearn Challenge Forecasting Track, 2nd Place, NeurIPS.**
Team Members: Tanwi Mallick, Ngoc Mai Tran, Yangxinyu Xie
- 2022 **Longhorn Research Poster Session Award, 3rd Place, Research Week, University of Texas at Austin.**
- 2021 **Science4cast Competition, 2nd Place, IEEE BigData Cup Challenges, Institute of Advanced Research in Artificial Intelligence, Vienna, Austria.**
Team Members: Ngoc Mai Tran, Yangxinyu Xie
- 2021 **Smart City Hackathon, 1st Place, Good Systems, University of Texas at Austin.**
Team Members: Jacob Helwig, Jack Si, Yangxinyu Xie

Presentations

Research Talks

- 2025
 - Debiasing Watermarks for Large Language Models via Importance Sampling. AMS Special Session on Trustworthy AI Applications Including Machine Learning, PINN, and Inverse Problems. Joint Mathematics Meetings. Seattle, WA, January, 2025.
 - Unveiling Large Language Model Vulnerabilities: A Geometric and Topological Approach to Adversarial Robustness. AMS Special Session on MRC Mathematics of Adversarial, Interpretable, and Explainable AI. Joint Mathematics Meetings. Seattle, WA, January, 2025.

- 2024
 - Statistical Gems in AI Safety. Data-driven Decision-making: Personalization, Practicality, Perspicuity. INFORMS Annual Meeting. Seattle, WA, October, 2024.
 - WildfireGPT: Tailored Large Language Model for Wildfire Analysis. The SIGIR 2024 Workshop on Information Retrieval for Climate Impact. Washington D.C. July, 2024.
 - Unbiased Watermark for Large Language Models via Importance Sampling. Lawrence D. Brown Student Workshop. University of Pennsylvania. March, 2024.
 - Joint Convergence of Monochromatic Edges in Multiplex Hypergraphs. UNL Discrete Seminar (DisCo). University of Nebraska-Lincoln. February, 2024.
 - Joint Convergence of Monochromatic Edges in Multiplex Hypergraphs. AMS Contributed Paper Session on Combinatorics, IV. Joint Mathematics Meetings. San Francisco, CA, January, 2024.

- 2023
 - A Comparative Study of Loss Functions: Traffic Predictions in Regular and Congestion Scenarios. Summer Argonne Student Symposium (Sassy). Virtual, August, 2023.

- 2021
 - Random walk rankings with feature selection and imputation. IEEE International Conference on Big Data, Virtual, December, 2021.
 - Optimizing Ambulance Allocation and Routing During Extreme Events. Smart Cities Consortium, Good Systems, University of Texas at Austin, Austin, TX, January, 2021.

- 2020
 - Matrix Rigidity, Algebraic and Combinatorial Techniques. 3rd Bringing Young Mathematicians Together Conference, Universitat de Valencia, Virtual, December, 2020.

- 2019
 - On 2 by 2 Tropical Commuting Matrices. Southeastern Undergraduate Mathematics Workshop, Georgia Institute of Technology, Atlanta, GA, August, 2019.

Poster Presentations

- 2024
 - WildfireGPT: Tailored Large Language Model for Wildfire Analysis. Tackling Climate Change with Machine Learning Workshop, NeurIPS, Vancouver, BC, December, 2024.
 - A Peek into Token Bias: Large Language Models Are Not Yet Genuine Reasoners. Workshop on Statistical Frontiers in LLMs and Foundation Models, NeurIPS, Vancouver, BC, December, 2024.

- 2022
 - Towards Label Driven Mondrian Forests. The Southern Regional Council on Statistics, 57th Summer Research Conference, Jekyll Island, GA, October, 2022.
 - Predicting Covid-19 EMS Incidents from Daily Hospitalization Trends. Longhorn Research Poster Session, UT Austin, April, 2022. *Poster Presentation Award, Third Place.*
- 2021
 - Optimizing EMS Responses during Extreme Events. MAA Undergraduate Poster Session, Joint Mathematics Meetings, January, 2021. *Honorable Mention Poster.*
- 2020
 - On Tropical Commuting Matrices. MAA Undergraduate Poster Session, Joint Mathematics Meetings, Denver, CO, January, 2020.

Public Scholarship

- 2025
 - A Data-Driven Analysis of the Surge in Anti-Transgender Bills in U.S. State Legislatures, 2018-2023. Spectra Special Session. Joint Mathematics Meetings. Seattle, WA, January, 2025.
- 2023
 - Modeling the Impact of the Baltimore Red Line on Job Accessibility, GIS DAY 2023, University of Pennsylvania, November 2023.
 - All Aboard! Assessing Transit Equity in Baltimore, MD, Impact Summit, August 2023.

Professional Service

Reviewer

Journals ○ *Stat*

- Workshops ○ NLP for Positive Impact Workshop at EMNLP 2024
 ○ NeurIPS 2024 Workshop on Behavioral ML

Scholarships, Fellowships and Grants

- 2024 **AMS Graduate Student Travel Grants for Joint Mathematics Meeting.**
- 2023 **Sustainable Research Pathways**, *Sustainable Horizons Institute & U.S. Department of Energy Exascale Computing Project.*
- 2022 **Boyd Harshburger Travel Award**, *The Southern Regional Council on Statistics.*
- 2021 **Graduate School Application Financial Aid Program**, *oSTEM, Queer in AI.*
- 2021 **Student Fellow**, *College of Undergraduate Studies*, University of Texas at Austin.
- 2021, 2020 **AMS Undergraduate Student Travel Grants for Joint Mathematics Meeting.**

Recognition and Distinctions

- 2022 **Dean's Honored Graduate**, *College of Natural Sciences*, University of Texas at Austin.

- 2022 **Research Distinction**, *College of Natural Sciences*, University of Texas at Austin.
- 2021 **UT Austin Member**, *Academy of Undergraduate Researchers Across Texas*.
- 2021 **George H. Mitchell Student Awards Semi-Finalist**, *Student Success Initiatives*, University of Texas at Austin.
- 2021 **Distinguished College Scholar**, *College of Natural Sciences*, University of Texas at Austin.
- 2021 **Cactus Standout Award**, *Cactus Yearbook*, University of Texas at Austin.
- 2019 **Second Year Excellence Award**, *College of Natural Sciences*, University of Texas at Austin.
- Fall 2017 **Dean's List**, *College of Undergraduate Studies*, University of Texas at Austin.

Professional Development

- Summer 2024 **Mathematics Research Communities (MRC)**, *Mathematics of Adversarial, Interpretable, and Explainable AI*, American Mathematical Society.

Teaching

Certification

- Spring 2024 **The Carpentries Coding Instructor**, *Certified by: Karen Word*.

Workshops

- March 2024 **Data Carpentries: Python for Social Scientists**, *Research Data & Digital Scholarship*, University of Pennsylvania Libraries.

Teaching Assistantship - University of Pennsylvania

- Fall 2024 **Introductory Statistics**, *Undergraduate*.
- Fall 2024 **Stat Computing with R**, *Undergraduate*.
- Fall 2023 **Statistical Methodology**, *Graduate*.
- Fall 2022 **Applied Econometrics I**, *Graduate*.

Teaching Assistantship -University of Texas at Austin

- Spring 2020 **Algorithms and Complexity: Honors**, *Undergraduate*.
- Spring 2020 **Predictive Analytics**, *Undergraduate*.
- Fall 2019 **Applied Regression and Time Series**, *Undergraduate*.
- Fall 2019 **Discrete Mathematics**, *Undergraduate*.

Mentorship

- Spring 2023 **Graduate Student Mentor**, *Undergraduate Research Program in Probability and Statistics*, University of Pennsylvania.
- Fall 2022 **Graduate Student Mentor and Panelist**, *3rd annual Statistics Undergraduate Research Experience (SURE)*, The Southern Regional Council on Statistics.
- Spring 2021 – Spring 2022 **Mentor**, *Directed Reading Program, ACM For Change*, University of Texas at Austin.

Leadership

- 2024 **Board Member**, *Independence Business Alliance*, LGBTQ+ Chamber of Commerce for Greater Philadelphia.
- Fall 2022 – **Board Member**, *Wharton Doctoral Council*, University of Pennsylvania.
current
- 2023 **Participant**, *Michael S. Hinson, Jr. LGBTQ+ Leadership Pipeline Program*, Philadelphia Office of LGBT Affairs.
- Spring 2023 **Volunteer**, *Community Engagement Program*, University of Pennsylvania.

Professional Experience

- Summer 2024 **Givens Associate**, *Argonne National Laboratory*, Lemont, IL.
- Summer 2023 **Visiting Student**, *Argonne National Laboratory*, Lemont, IL.
- Summer 2019 **Software Engineering Intern**, *Sense Talent Labs*, San Francisco, CA.
- 2018–2019 **Data Science Innovation Fellow**, *UT OnRamps*, Austin, TX.
- 2018 – 2019 **ProjectLEAD**, *Leadership and Ethics Institute*, University of Texas at Austin.

Skills

- | | |
|--------------|----------------------------|
| Fluent | Python, R, \LaTeX |
| Intermediate | Java, C#, C, C++, SQL |
| Basic | HTML, Shell |