Tracy (Yixin) Zhu

Email | Google Scholar | Website

Education

University of Chicago Chicago, IL

M.S. in Statistics *Sept 2023 - May 2025*

- Current GPA: 3.8
- Research Interest: Generative models, Multimodal Learning, Self-Supervised Learning, Computer Vision, Deep Learning
- Courses taken: Introduction to Computer Vision, High Dimensional Probability, Topics in Computer Vision, etc.

New York University New York, NY

B.A. in Data Science and Mathematics *Sept 2019 - May 2023*

- Honors: Cum Laude; 4 PhD-level courses; Dean's List of Year 2023 and 2022

Peer-Reviewed Papers

* indicates equal contributions.

Peer-Reviewed Papers

[1] **Tracy Zhu***, Amitabh Mahapatra*, Zhiyan Wang*, Svetlana Lazebnik, David Forsyth, Anand Bhattad. **Improving Projective Geometry in Image Generation**.
Under review

[2] Yukai Yang*, **Tracy Zhu***, Marco Morucci, Tim G.J. Rudner. **Weak-to-strong Confidence Prediction**.
Workshop on Statistical Foundations of Large Language Models, Attributing Model Behavior at Scale, Safe Generative AI, and Regulatable ML. (**NeurIPS Workshop**), 2024

[3] Hongyi Zheng, **Tracy Zhu**, Lavender Yao Jiang, Kyunghyun Cho, Eric Karl Oermann. **Making the Most Out of the Limited Context Length: Predictive Power Varies with Clinical Note Type and Note Section**.
ACL Student Research Workshop, 2023

Academic Experience

Reviewer Oct 2024

Remote

- Served as a reviewer for NeurIPS 2024

Student Researcher (Remote) Feb 2024 - Present

Center for Data Science, New York University *New York, NY*

- Conducted experiments with a linear probe to evaluate LLM uncertainty using representations from white-box LLMs in a generalizable evaluation framework and analyzed the learned information
- Drafted manuscripts and created visualizations, including plots and tables, for a resulting workshop paper

Student Research Assistant Jun – Sept 2023

Center for Data Science, New York University *New York, NY*

- Implemented and experimented with heuristics functions in active learning for image classification for social science
- Mentored two undergraduate students from the Center for Data Science Undergraduate Research Program at NYU

Teaching Assistantship Jun 2022 - May 2023

Center for Data Science, New York University *New York, NY*

- 2023: DS-UA 301 Advanced Topics in Data Science: Techniques in Deep Learning, Jan 2023 - May 2023
- 2022: DS-UA 201 Causal Inference, Jun 2022 - Aug 2022

Projects

Improving Projective Geometry in Generative Vision Models | *Generative models, Diffusion models*

- Advised by Dr. Anand Bhattad, Prof. David Forsyth, and Prof. Svetlana Lazebnik
- Improved generated images in Diffusion models by conditioning on projective geometry cues
- Compared against several SOTA models, including a Vanishing Point Loss Constraint Finetuned SD2, other conditional diffusion models, and ControlNet
- Developed an Evaluation framework using MMD- and Relative Density to Ratio-based metrics

Weak-to-Strong Confidence Prediction | *LLM Safety, Representation Learning*

- Advised by Dr. Tim G.J. Rudner and Dr. Marco Morucci
- Evaluated question-answering ability of larger, black-box LLMs using representations from smaller, white-box LLMs
- Explored a widely applicable setting to train a probe where it does not need to know the answer to the question
- Applied LoRA to finetune white-box LLMs for better representations and compared performance with the probe

Interest Point Detection in Generative Models with SIFT | *Interest Point Detection, SIFT, Generative Models*

- Adapted SIFT to label ground truth images with distributions of interest points
- Trained offsets in StyleGAN with interest points labeled both point-wise and distribution-wise
- Optimized the attention maps in text-conditioned Diffusion Models to detect interest points
- Combined k-means Clustering with SIFT to improve quality of labeled interest points for the Diffusion model

Using Function-Space-VI in Active Learning | *Image Classification, Active Learning*

- Implemented AL heuristics with a function space variational inference model
- Implemented a HARA-based heuristics function in active learning and compared the performance
- Applied informative Gaussian priors on deep Bayesian models to select the most informative protest images

Investigating Predictive Power distributions with Clinical Notes | *LLM in Medical Application*

- Implemented a framework to analyze medical notes with ClinicalBERT model
- Investigated on predictive power distributions of clinical notes between different sections and their combinations

Grants

Summer Research Grant - $5000 Jun-Aug 2024
Center for Data Science, New York University *New York, NY*

- Award supported by Dr. Tim G. J. Rudner
- Contributed to a paper on Weak-to-Strong Confidence Prediction of Large Language Models

NYU Dean's Undergraduate Research Fund - $1000 Jan-May 2022
Wasserman Center for Career Development, New York University *New York, NY*

- Research grant to support undergraduate research
- Studied active learning with entropy-based heuristic for vision models