ZIRUO ZHAO

(323) 528-0850 | ziruo.rosie.zhao@gmail.com

EDUCATION

Stevens Institute of Technology   **NJ, US**
Ph.D. in Computer Science   *Sep. 2023 – Present*
Coursework: Machine Learning, Explainable NLP, Algorithm Design, Algorithmic Complexity

University of Southern California   **CA, US**
M.S. in Spatial Economics and Data Analysis   *Jan. 2021 – Dec. 2022*
Coursework: Spatial Databases, Time Series Analysis, Big Data Econometrics, Spatial Econometrics, Spatial Analysis and Modeling, Practice of Econometrics, Microeconomic Analysis and Policy

Imperial College London   **London, UK**
B.Sc. in Mathematics with Statistics   *Aug. 2016 – Aug. 2019*
Coursework: Methods for Data Science, Scientific Computation, High Performance Computing, Statistical Modeling, Introduction to Option Pricing, Probability and Statistics, Algebra, Analysis, Mathematics methods

RESEARCH EXPERIENCE

Privacy Vulnerabilities in LLMs   **NJ, US**
SIT Department of Computer Science   *Jun. 2025 – Present*
- Investigating privacy vulnerabilities of large language models (LLMs) under membership inference attacks (MIA).
- Developing experimental pipelines for document-level MIA, focusing on semantic similarity, perplexity-based detection, and loss-based feature analysis, motivated by recent research from USENIX Security and ICLR.
- Aiming to evaluate the security implications of LLM memorization and generalization behavior.

Theoretical Analysis of Noisy Contrastive Learning   **NJ, US**
SIT Department of Computer Science   *Jun. 2024 – Jan. 2025*
- Analyzed the theoretical limits of contrastive learning under adversarial noise.
- Conducted generalization analyses of contrastive learning.
- Preprint: link

Ghost Smelter: Environmental Justice Analysis of Historical Lead Smelters   **CA, US**
USC Spatial Sciences Institute   *Jun. 2022 – Oct. 2022*
- Analyzed environmental injustice related to the historical lead smelters in California.
- Analyzed and categorized census blocks using polygon centroids from U.S. Census Bureau's TIGER geographic data.
- Conducted t-tests and regression analyses to assess the impact of historical lead smelters using American Community Survey socioeconomic data.
- The project result was presented on the conference of the Association of Pacific Coast Geographers (APCG).

WORK EXPERIENCE

NetEase Network Co., Ltd.   **Hangzhou, China**
Intern, Media Technology department.   *Jul. 2021 – Oct. 2021*
- Produced three strategic industry reports that informed product development and positioning.
- Designed and implemented a UX quality evaluation framework adopted for social media product optimization.
- Completed daily product experience evaluation for AB Test.

PROJECTS

Individual Project, Explainable LLM
- Investigated shortcut learning behavior in BERT-based LLMs on Natural Language Understanding tasks, focusing on lexical and overlap biases that impact model generalization.
- Applied Integrated Gradients method and long-tailed distribution to interpret and quantify bias reliance.

Individual Project, Spatial Data Analysis
- Performed statistical and spatial regression analyses on housing sales data.
- Presented a project report using ArcGIS Story Map: link

SKILLS

Technical Skills: Python, R, SQL, ArcGIS, STATA, MATLAB, Fortran, LaTeX
Language Skills: Mandarin, English