Shiyuan Zhang

Email: sz54@illinois.edu   **Homepage**: https://seanzh30.github.io   **Phone**: +1 (217) 377-5722

EDUCATION

08/2023–Current	**Master of Computer Science** University of Illinois at Urbana-Champaign **Cumulative GPA:** 4.00
08/2019–05/2023	**B.S Statistics** **Minor in Computer Science** University of Illinois at Urbana-Champaign **Cumulative GPA:** 3.96
	Previously enrolled in the joint Information & Computer Science Bachelor's program at University of Liverpool & XJTLU (non-degree coursework)

RESEARCH INTEREST

ML Explainability: Data Attribution Method on Generative Model
Trustworthy ML: Model Fairness, Model Robustness, Model safety
NLP: Societal Effects and Social Bias of Language Models

PUBLICATIONS

1. DENG*, J., LI*, T. W., **ZHANG, SHIYUAN**, LIU, S., PAN, Y., HUANG, H., WANG, X., HU, P., ZHANG, X., AND MA, J. *dattri*: A Library for Efficient Data Attribution. *Under Review*

2. GE*, Y., **ZHANG, SHIYUAN***, AND LIU, X. Nuanced Multi-class Detection of Machine-Generated Scientific Text. *Under Review*

3. DENG, J., **ZHANG, SHIYUAN**, AND MA, J. Computational Copyright: Towards A Royalty Model for AI Music Generation Platforms. *Under Review*
   (* Equal Contribution)

RESEARCH EXPERIENCE

02/2024–Current	**Trustworthy and Regulatable AI Systems Lab, UIUC** Mentor: Professor Jiaqi Ma

1. **Development of the Open-Source Library *dattri***

- Implemented core functions, including Leave-One-Out (LOO), Linear Data-modeling Score (LDS), and Inverse Hessian Vector Product (IHVP).
- Designed and implemented a unified API, simplifying the switch between different data attribution methods.
- Integrated nanoGPT trained on different datasets, MNIST+MLP, and MNIST+LR, into the benchmark framework to compare the performance and effectiveness of different data attribution methods.

2. **Data Attribution Method for Generative Music Models**

- Implemented TRAK and TracIn frameworks to calculate the influence score of training data points on Music Transformer models.
- Evaluated the effectiveness of data attribution models using linear data modeling scores.

- Designed an adversarial attack using repetition and copy-paste techniques on the music dataset to test the robustness of the data attribution model.

12/2023–05/2024   **Independent Research, UIUC**
Collaborator: Yubin Ge
Enhanced Text Classification with Contrastive Learning
- Collected and filterd text data using the Semantic Scholar API and generated text by using language models (GPT-3.5) to simulate generating methods such as paraphrasing, summarization, and title-based generation.
- Proposed and implemented a novel baseline model using contrastive learning based on SciBert, RoBerta, and DeBerta, significantly improving text classification accuracy and robustness.

Teaching Experience

05/2024–Current   **CEE 202 - Eng Risk & Uncertainty, UIUC**
Teaching Assistant   Professor Sotiria Koloutsou-Vakakis
- Conducted weekly office hours to support students programming assignments and questions about the lecture content.
- Led R programming tutorials, graded assignments, and responded to student inquiries on Campuswire.

09/2023–06/2024   **SIIP Program for CEE 202, CEE 330, CEE 340, UIUC**
Teaching Assistant   Professor Sotiria Koloutsou-Vakakis
- Developed automated grading systems on the PrairieLearn and Gradescope platform(Python).
- Designed Jupyter Notebook assignments, and created R/Python problem sets and corresponding answer sets.

05/2022–08/2022   **CS 411 - Database Systems, UIUC**
Course Assistant   Professor Abdussalam Alawini
- Proctored exams, graded assignments and provided homework assistance.
- Designed and checked SQL questions on PrairieLearn platform.

Skills

Programming   Python, C/C++, R

Tools   Pytorch, MySQL, MongoDB, Neo4j, Linux, Git, Docker, Notebook, Slurm, Ubuntu

Honors, Awards & Scholarships

08/2019-05/2023   UIUC Dean's List in all the semesters

05/2023   Graduated with Highest Distinction

Selected courses

ML   Advanced NLP, Applied ML, AI Algorithms, Computer Vision

System   Software Engineering, Database, Computer Network, Operating System

Stat&Math   Abstract Linear Algebra, Statistical Modeling, Adv Prob&Stat, Math Analysis