

Shuli Jiang

✉ shulij@andrew.cmu.edu

🐙 @11hifish




🔍 Google Scholar

🌐 <https://www.andrew.cmu.edu/user/shulij/>

Research Interests

I am a fourth-year Ph.D. student at the School of Computer Science, Carnegie Mellon University, working with Prof. Gauri Joshi. My interests lie broadly in theory and applications of federated learning, communication-efficient distributed learning algorithms, differential privacy and security problems of large foundation models.

Education

- August 2020 – present  **Carnegie Mellon University, Pittsburgh, PA, USA**
Ph.D. student at Robotics Institute, School of Computer Science
Advisor: Prof. Gauri Joshi
Expected Graduation Date: June 2025
- May 2019 – May 2020  **Carnegie Mellon University, Pittsburgh, PA, USA**
M.S. in Computer Science
Thesis title: *Deep Multi-view Clustering Using Local Similarity Graphs*
Advisor: Prof. Artur Dubrawski
- August 2015 – May 2019  **Carnegie Mellon University, Pittsburgh, PA, USA**
B.S. in Computer Science, University Honor
Minor: Electrical and Computer Engineering


Research Publications

(α/β : alphabetical order, **: contribution order)





In Submission

1. (**) [Shuli Jiang](#), Qiuyi Richard Zhang, Gauri Joshi
Optimized Tradeoffs for Private Majority Ensembling
In submission to the Twelfth International Conference on Learning Representations (ICLR 2024)

Workshop Proceedings

1. (**) [Shuli Jiang](#), Swanand Kadhe, Yi Zhou, Ling Cai, Nathalie Baracaldo
Forcing Generative Models to Degenerate Ones: The Power of Data Poisoning Attacks  [Link](#)
NeurIPS 2023 Workshop on Backdoors in Deep Learning - The Good, the Bad, and the Ugly (Best Poster Award)

Conference Proceedings

1. (**) [Shuli Jiang](#), Pranay Sharma, Gauri Joshi
Correlation Aware Sparsified Mean Estimation Using Random Projection  [Link](#)  [Code](#)
The Thirty-seventh Conference on Neural Information Processing Systems (NeurIPS 2023)
2. (**) [Shuli Jiang](#), Robson Leonardo Ferreira Cordeiro, Leman Akoglu
D.MCA: Outlier Detection with Explicit Micro-Cluster Assignments  [Link](#)  [Code](#)
The Twenty-second IEEE International Conference on Data Mining (ICDM 2022)

3. ($\alpha\beta$) [Shuli Jiang](#), Hai Thanh Pham, David P. Woodruff, Qiuyi Richard Zhang
Optimal Sketching for Trace Estimation [Link](#) [Code](#)
The Thirty-fifth Conference on Neural Information Processing Systems (NeurIPS 2021 Spotlight)
4. ($\alpha\beta$) [Shuli Jiang](#), Dongyu Li, Irene Mengze Li, Arvind V. Mahankali, David P. Woodruff
Streaming and Distributed Algorithms for Robust Column Subset Selection [Link](#) [Code](#)
The Thirty-eighth International Conference on Machine Learning (ICML 2021)
5. (**) Bohan Zhang, Dana Van Aken, Justin Wang, Tao Dai, [Shuli Jiang](#), Jacky Lao, Siyuan Sheng, Andrew Pavlo, Geoffrey J. Gordon
A Demonstration of the OtterTune Automatic Database Management System Tuning Service [Link](#) [Code](#)
The VLDB Endowment, Vol. 11, No. 12 (VLDB 2018)

Technical Reports

1. ($\alpha\beta$) Theresa Gebert, [Shuli Jiang](#), Jiaxian Sheng
Characterizing Allegheny County Opioid Overdoses with an Interactive Data Explorer and Synthetic Prediction Tool [Link](#) [Code](#)
HackAuton Best Show Prize, 2018
2. [Shuli Jiang](#)
Deep Multi-view Clustering Using Local Similarity Graphs [Link](#)
Master's Thesis, 2020, Advisor: Prof. Artur Dubrawski

Work Experience

- May 2023 - August 2023 ■ **IBM Research**, Almaden, CA, USA
 Research Summer Intern (AI Security and Privacy Solutions)
 Advisor: Swanand Kadhe, Manager: Nathalie Baracaldo
 Investigate security vulnerabilities of large language models (LLMs) in terms of data poisoning attacks targeting natural language generation (NLG) tasks, including text summarization, text completion, table-to-text generation, etc. Design and develop defense strategies to counter-attack those types of security threats to LLMs.
- June 2018 - August 2018 ■ **Morgan Stanley**, New York City, NY, USA
 Technology Analyst (Application Development)
 Develop a data quality management system which collects real-time trading data from multiple source databases, detects potential anomalies to ensure data quality and visualizes anomalous data.
- June 2017 - August 2017 ■ **PreSenso Ltd.**, Haifa, Israel
 Software Engineer Intern
 Develop an anomaly detection benchmark for evaluating and comparing the performances of different anomaly detection algorithms on various patterns of anomalies.


Public Talks


- September 2023 ■ AI-EDGE Students and Postdocs gathering for AI Research and Knowledge Sharing (AI-EDGE SPARKS)
 Topic: Federated Learning and Distributed Vector Mean Estimation


Public Talks (continued)

May 2023  CMU Robotics Institute Ph.D. Speaking Qualifier Public Talk
Topic: Differential Privacy and Private Majority Ensembling


Service


Conference/Workshop Reviewer  SODA 2022, SIGKDD 2023, AAAI The First Workshop on DL-Hardware Co-Design for AI Acceleration 2023, NeurIPS 2023, ICLR 2024, AISTATS 2024, SDM 2024


Journal Reviewer  IEEE/ACM Transactions on Networking 2023

Department  CMU Robotics Institute Ph.D. Admission Committee 2023

Teaching Assistantship

Fall 2022  **16-831 Statistical Techniques in Robotics**, @ Carnegie Mellon University

Fall 2020  **10-725 Convex Optimization**, @ Carnegie Mellon University

Fall 2017  **17-214 Principles of Software Construction**, @ Carnegie Mellon University

Technical Skills

Programming  Python, Java, Matlab (Basic), C (Basic)

Software Tools  Tensorflow, PyTorch, Pandas, LaTeX

Awards

2023  NeurIPS 2023 Scholar Award

2022  IEEE ICDM 2022 Student Travel Award (\$ 700)

 Graduate Student Assembly/Provost Conference Travel Grant (\$ 750)

2019  Carnegie Mellon University Undergraduate University Honor

2015 – 2019  Carnegie Mellon University Undergraduate Dean's List

2018  HackAuton Best Show Prize

2017 – 2019  Carnegie Mellon University Innovation Scholar

2017  Buncher Entrepreneurship Award (\$ 10,000)