

MAYANK SHRIVASTAVA

405 E. Stoughton, Champaign, Illinois, 61801

☎ 217-200-5374

✉ mayank98shri@gmail.com

🌐 [linkedin.com/in/mayank-shrivastava](https://www.linkedin.com/in/mayank-shrivastava)

🐙 github.com/mayank010698

Education

University of Illinois at Urbana-Champaign

Ph.D. in Computer Science

Specialization : Machine Learning

Aug. 2024 – 2028 (Expected)

Illinois, USA

University of Illinois at Urbana-Champaign

Master of Science (Thesis-Track) in Computer Science

Specialization : Machine Learning, GPA: 4.0/4.0

Aug. 2022 – May 2024

Illinois, USA

Indian Institute of Technology Kanpur

Bachelor of Technology, Electrical Engineering with minor in Machine Learning

GPA: 9.8/10, Department Rank : 2/135

Aug. 2016 – Aug. 2020

Kanpur, India

Publications

Breaking the Dimension Dependence in Sketching for Distributed Learning

NeuRIPS 2024

Mayank Shrivastava, Berivan Isik, Qiaobo Li, Sanmi Koyejo, Arindam Banerjee

Paper

38th Annual Conference on Neural Information Processing Systems, 2024

Max-Quantile Grouped Infinite-Arm Bandits

ALT'23

Ivan Lau, Yan Hao Ling, Mayank Shrivastava, Jonathan Scarlett

Paper

34th International Conference on Algorithmic Learning Theory, 2023

Research Experience

Breaking the Dimension Dependence in Sketching for Distributed Learning

July 2023 – March 2024

Research Assistant with Dr. Arindam Banerjee

UIUC

- Analyzed the convergence rates and communication efficiency of federated learning with linear sketching.
- Existing works inherit a dependence on ambient dimensionality - the size of the deep learning model.
- For overparametrized models, we derive dimension-free convergence rates for sketching-based federated learning.
- Extended our work to differentially private settings and heterogeneous clients, showing improvements in communication.

Max-Quantile Grouped Infinite-Arm Bandits

October 2020 – March 2022

Research Assistant with Dr. Jonathan Scarlett

NUS, Singapore

- Studied the problem of max-min grouped multi-arm bandits and extended it to the case of infinite arms.
- Developed a two-phase algorithm and derived instance-dependent sample complexity bounds.

Professional Experience

AI R&D Lab, Samsung Electronics

October 2020 – March 2022

Machine Learning Engineer, Automatic Speech Recognition team

Suwon, South Korea

- Worked on training and inference engine of Conformer Speech Model for Bixby's(virtual assistant) End-to-End ASR.
- Implemented neural model for External Language Model selection (shallow fusion) - improving inference accuracy by 3%
- Implemented RIR (Impulse Response) Augmentation for improving the accuracy on far-field ASR for IoT devices.

AI R&D Lab, Samsung Electronics, South Korea

May 2019 – July 2019

Summer Internship, Automatic Speech Recognition(ASR) Team

Suwon, South Korea

- Augmented female speech dataset by implementing Voice Conversion(VC) generative models on male speech dataset.
- Implemented generative models, including Conditional VAE, and Star-GAN VC in Pytorch- improving WER(Word Error Rate) of ASR engine by 2% for female speakers through the proposed augmentation pipeline.
- Received Pre-Placement Offer from Samsung as a result of exceptional performance and sincere effort.

Indian Institute of Science, Bangalore

May 2020 – Sep. 2020

Project Assistant with Dr. Himanshu Tyagi

Bangalore, India

- Developed an agent-based simulator for modeling COVID-19 spread in Karnataka, India.
- Modelled neighborhood movement as a Markov Chain using OpenCV to extract geographical location using Maps.
- Developed sampling-based testing strategies on the simulator to reduce total tests and identify outbreak clusters.
- Developed a web application to optimally allocate collected samples across testing centers based on payload and geographical constraints.

Selected Projects

- 🔗 **Generating Wikipedia infoboxes using LLMs** | *LLMs, Vector Databases, Natural Language Processing* **Dec 2023**
- Developed an LLM pipeline to generate tabular summaries of Wikipedia pages over existing manual process.
 - Scraped 2k+ articles to build a vector DB and proposed a novel similarity-based template retrieval system.
 - Reported an improvement of 97% in BLEU scores over zero-shot prompting and comparable accuracy to ground-truth on downstream QA tasks.
- 🔗 **Analysis of Langevin Algorithms** | *Diffusion Models, Sampling, Optimization* **May 2023**
- Studied and implemented various algorithms for reverse diffusion in Diffusion models.
 - Implemented a dimensional-independent version of Langevin Algorithm(ULA-PD) enabling dimension-free convergence.
 - Proposed and evaluated a metropolis version of ULA-PD for faster convergence of neural network optimization.
- 🔗 **Excess Food Distribution Application** | *MySQL, Flask, React* **Dec 2022**
- Developed a web application using Flask and React linking users & restaurants for food redistribution and analytics.
 - Implemented triggers, stored procedures, & indexing for superior system performance.

Technical Skills

Languages: Python, Java, C++, SQL

Technologies/Frameworks: PyTorch, MongoDB, Neo4j, MySQL, L^AT_EX, Flask, React