Kulin Shah

kulin.shah98@gmail.com ⋄ kulin@cs.utexas.edu ⋄ Phone: +1 669 290 5808 ⋄ Webpage (🏠) ⋄ Github (🐙)

EDUCATION

University of Texas at Austin *August 2021 -*
Ph.D. in Computer Science
Advisor: Prof. Adam Klivans

International Institute of Information Technology, Hyderabad *August 2015 - July 2019*
B. Tech (Honors) in Computer Science and Engineering
Advisor: Prof. Naresh Manwani

RESEARCH INTEREST

Various aspects (e.g., robustness, reasoning, privacy, efficiency) of Large Language Models and Diffusion Models.

RESEARCH EXPERIENCE

Graduate Research Assistant, University of Texas at Austin Aug 2021 - Present
- Advisor: Prof. Adam Klivans
- Working on problems in understanding and improving the building blocks of modern generative models (diffusion models and autoregressive models).

Student Researcher, Google Research June 2023 - March 2024
- Manager: Dr. Rina Panigrahy
- Worked on problems in language modeling to improve its reasoning capabilities and efficiency of the architecture.
- Finished two projects on understanding reasoning and efficiency of the language models (see papers this and this).
- The dataset created in our reasoning NeurIPS'24 paper was used in BIG-Bench Extra Hard benchmark of language models and used to evaluate Gemma 3 models (Google's open language models).

Research Fellow, Microsoft Research, India Aug 2019 - July 2021
- Mentor: Dr. Navin Goyal and Dr. Amit Deshpande
- Worked on problems in generative models, representation learning, theory of deep learning.

Research Intern, Microsoft Research, India May 2019 - July 2019
- Mentor: Dr. Amit Deshpande and Prof. Chiranjib Bhattacharyya
- Worked on problems related to fairness in machine learning.

Research Intern, Indian Institute of Science (IISc), Bangalore May 2018 - June 2018
- Mentor: Prof. PS Sastry
- Worked towards understanding architecture and training dynamics of Capsule Network.

SELECTED RESEARCH PROJECTS

- **Robust training of diffusion models and its implication on memorization**
  - ◇ Proposed a diffusion-based framework to learn distributions using only highly corrupted samples, enabling generative modeling in scenarios without clean data and improving memorization (Paper at NeurIPS'23).
  - ◇ Developed a principled approach to mitigate memorization in diffusion models by leveraging training using corrupted data at large noise scales, improving *image-generation quality (fidelity) and memorization both simultaneously* (Paper at ICML'25).

- **Understanding and improving the reasoning capabilities of language models**
  - ◇ Investigated the reasoning and search capabilities of LMs to solve complex puzzles (e.g., Sudoku), demonstrating the logical token-generation order dramatically enhances reasoning abilities (Paper at NeurIPS'24).
  - ◇ Analyzing the benefit of Masked Diffusion Models (MDMs) in any order generation/reasoning on solving complex puzzles (e.g., Sudoku) and showing that MDMs *without supervision* about the logical token-generation order outperform autoregressive models *with supervision* (Paper at ICML'25).

PAPERS ($(\alpha - \beta)$ indicates the alphabetical ordering and * indicates equal contribution)

16. **Train for the Worst, Plan for the Best: Understanding Token Ordering in Masked Diffusions** [paper]
Jaeyeon Kim*, **Kulin Shah***, Vasilis Kontonis, Sham M. Kakade, Sitan Chen
International Conference on Machine Learning (**ICML**), 2025 (**Oral**)

15. **Does Generation Require Memorization? Creative Diffusion Models using Ambient Diffusion** [paper]
Kulin Shah, Alkis Kalavasis, Giannis Daras, Adam Klivans
International Conference on Machine Learning (**ICML**), 2025

14. **Learning general Gaussian mixtures with efficient score matching** [paper]
$(\alpha - \beta)$ Sitan Chen, Vasilis Kontonis, **Kulin Shah**
Conference on Learning Theory (**COLT**), 2025

13. **Causal Language Modeling Can Elicit Search and Reasoning Capabilities on Logic Puzzles** [paper]
Kulin Shah, Nishanth Dikkala, Xin Wang, Rina Panigrahy
Neural Information Processing Systems (**NeurIPS**), 2024

12. **Unrolled denoising networks provably learn optimal Bayesian inference** [paper]
Aayush Karan*, **Kulin Shah***, Sitan Chen, Yonina Eldar
Neural Information Processing Systems (**NeurIPS**), 2024

11. **Simple Mechanisms for Representing, Indexing and Manipulating Concepts** [paper]
$(\alpha - \beta)$ Yuanzhi Li, Raghu Meka, Rina Panigrahy, **Kulin Shah**
Preprint

10. **Learning Mixtures of Gaussians Using the DDPM Objective** [paper]
Kulin Shah, Sitan Chen, Adam Klivans
Neural Information Processing Systems (**NeurIPS**), 2023

9. **Ambient Diffusion: Learning Clean Distributions from Corrupted Data** [paper]
Giannis Daras, **Kulin Shah**, Yuval Dagan, Aravind Gollakota, Alexandros G. Dimakis, Adam Klivans
Neural Information Processing Systems (**NeurIPS**), 2023

8. **Debiased Dynamic Stochastic Gradient Aggregation for Learning with Multiple Objectives**
Mao Ye*, **Kulin Shah***, Qiang Liu
Preprint

7. **Learning and Generalization in Overparameterized Normalizing Flows** [paper]
Kulin Shah, Amit Deshpande, Navin Goyal
International Conference on Artificial Intelligence and Statistics (**AISTATS**), 2022.
Workshop on the Theory of Overparameterized Machine Learning (**TOPML**), 2021.

6. **RISAN: Robust Instance Specific Deep Abstention Network** [paper]
Bhavya Kalra, **Kulin Shah**, Naresh Manwani
Conference on Uncertainty in Artificial Intelligence (**UAI**), 2021 (**Oral**).

5. **Rawlsian Fair Adaptation of Deep Learning Classifiers** [paper]
Kulin Shah, Pooja Gupta, Amit Deshpande, Chiranjib Bhattacharyya
AAAI/ACM Conference on AI, Ethics, and Society (**AIES**), 2021.

4. **Online Active Learning for Reject Option Classifier** [paper]
Kulin Shah, Naresh Manwani
AAAI Conference on Artificial Intelligence (**AAAI**), 2020 (**Oral**).

3. **Sparse Reject Option Classifier using Successive Linear Programming** [paper]
Kulin Shah, Naresh Manwani
AAAI Conference on Artificial Intelligence (**AAAI**), 2019 (**Oral**) .

2. **PLUME: Polyhedral Learning Using Mixture of Experts** [paper]
Kulin Shah, PS Sastry, Naresh Manwani

1. **Ingredients for Happiness: Modeling Constructs via Semi-supervised Content Driven Inductive Transfer** [paper]
   Bakhtiyar Syed, V. Indurthi, **Kulin Shah**, Manish Gupta and Vasudeva Varma
   AAAI-19 Workshop on Affective Content Analysis, AFFCON-19 (**Runner-up** for CL-Aff shared task).

AWARDS AND ACHIEVEMENTS

- Awarded **Google** conference travel scholarship award in 2024.
- Awarded **NeurIPS** scholar award 2023.
- Awarded **Google**, **Microsoft Research** travel grant and **AAAI Student Scholarship** to attend **AAAI** 2019.
- Awarded **Research Award** for exceptional research work at IIIT Hyderabad.
- Awarded **Dean's List** award for excellent academic performance in 2016, 2017 and 2018.
- **34 rank** in India in online round of ACM ICPC programming contest, 2018 (Total 3000+ teams)
- **53 rank** in Amritapuri regional of ACM ICPC programming contest, 2017 (Total top 260 teams from India).

TALKS

- Presented our work on learning mixtures of Gaussians using diffusion models at a joint diffusion seminar between Harvard University, Caltech, and UT Austin. *2024*
- Presented our work on learning mixtures of Gaussians using diffusion models at Apple Machine Learning Research. *2024*
- Presented our work on learning in Normalizing Flows at a general meeting at Microsoft Research India. *2021*
- Presented our work on reject option classifier in AAAI Conference on Artificial Intelligence. *2019*

TECHNICAL SKILLS

Programming Languages	Python, Matlab, C, C++, Bash, Java
Libraries & Tools	PyTorch, TensorFlow, Jax, Huggingface, Keras, Scikit-learn, Git, Latex

RELEVANT COURSES

Generative Models & Multiobjective optimization
Topics in Machine Learning (Online Learning & Bandits)
Optimization Methods
Game Theory
Adv. Probability (Concentration, Stein's Method, Mean-field theory)
Reinfocement Learning
Statistical Methods in AI
Autonomous Robots
Computer Vision
Functional Analysis