Saurabh Shah

saurabhs.site saurabhs@allenai.org LinkedIn GitHub Google Scholar

Education

University of Pennsylvania **Philadelphia, PA**

Master of Science in Engineering in Computer Science 3.91 GPA *Jan 2021 - May 2023*

- Focused on Algorithmic Theory, Artificial Intelligence, Machine Learning, and Natural Language Processing (NLP)

Bachelor of Science in Engineering in Networked and Social Systems Engineering 3.90 GPA *Aug 2019 - May 2023*

- Major combines Computer Science, Systems Engineering, and Economics. Minors in Data Science and Mathematics

Experience

Allen Institute for AI (Ai2) **Seattle, WA**

Research Engineer *Feb 2025 - Present*

- Olmo team. Training open language models (Olmos) to write code, use tools, and reason

Apple **Seattle, WA**

Machine Learning Engineer *Oct 2023 - Feb 2025*

- Siri Natural Language Understanding (NLU). Helped build an agentic Siri planner powered by Apple Intelligence

Allen Institute for AI (Ai2) **Seattle, WA**

Research Engineering Intern (paper) *Aug 2023 - Oct 2023*

- Tried pretraining Olmo with ReLoRA, a parameter-efficient *pretraining* method. Learned lots about PyTorch/FSDP
- Accepted into the Association of Computational Linguistics (ACL) 2024 Main Conference - Theme Paper Award

University of Pennsylvania **Philadelphia, PA**

Researcher (paper) (talk) *Aug 2022 - May 2023*

- Explored using free-text explanations for improving the robustness of LLMs to spurious cues in training data
- Accepted into the Association of Computational Linguistics (ACL) 2023 Main Conference

Apple **Seattle, WA**

Machine Learning Engineering Intern *May 2022 - Aug 2022*

- Siri NLU. Built an internal iOS app in Swift to help test different natural language text-to-intent parses and streamline the counterfactual evaluation flow of the NLU system. Used by annotators and QA testers

Amazon **Nashville, TN**

Software Development Engineering (SDE) Intern *May 2021 - Aug 2021*

- Robotics-AI Computer Vision. Built a web app with React and AWS to configure, search, and view over 300,000 cameras

Personal Projects

The Learning Curve (link) **March 2025 - Present**

- A blog where I talk about machine learning research and engineering

Griffin LM + CUDA (link) **May 2024 - August 2024**

- I learned some cuda (link) and tried to implement Griffin from scratch in PyTorch with a cuda extension for the scan

Concept Space Embeddings (link) **Feb 2023 - Apr 2023**

- Worked with a team of 2 to create a novel method for interpretable embeddings of arbitrary text using LLMs and Decision Trees. Works for classification, regression, clustering, and post hoc explanation of black box models

Compass (Penn Course Recommendation) (link) **Jan 2023 - Apr 2023**

- Group of 4. Course recommendation web app. I built the recommendation system with (1) collaborative filtering and (2) text embeddings to recommend courses to students based on (1) perceived difficulty and (2) natural language interests

Poké-GANs (Pokémon Generator) (link) **Mar 2022 - Apr 2022**

- Generated complete Pokémon from names. Fine-tuned GPT-3 for types, stats, abilities; CLIP+VQGAN for images from generated text. Trained custom LSTM and GANs from scratch and compared results. Worked with partner.

Comedy Bot (link) **July 2020**

- Experimented with ML models to recognize and rate jokes I write and perform for crowds of 150+. Joke datasets from Kaggle. Experimented with Bag of Words/Naïve Bayes and LSTM models. Built with PyTorch

Technical Skills

Languages: Python, TypeScript/JavaScript, Go, CUDA/C++, Java, Haskell, Coq, Swift
Technologies/Frameworks: PyTorch/FSDP, LLMs, HuggingFace, AWS, React, Pandas