Varun Mulchandani

vmmulcha@ncsu.edu | varun-mulchandani.github.io

Location: Raleigh, NC, USA ; Citizenship: U.S. Citizen

Education

North Carolina State University — Raleigh, NC, USA
PhD in Computer Science, Advisor: Dr. Jung-Eun Kim — *Aug 2024 -*

North Carolina State University — Raleigh, NC, USA
Master of Science in Computer Science, Thesis Advisor: Dr. Jung-Eun Kim — *Aug 2022 – Jul 2024*

- Full Tuition Waiver (3/4 semesters), GPA: 4.0/4.0

Vellore Institute of Technology — Vellore, TN, India
B.Tech in Computer Science and Engineering — *July 2018 – May 2022*

- GPA: 9.34/10

Publications

Varun Mulchandani and Jung-Eun Kim. Severing Spurious Correlations with Data Pruning. In International Conference on Learning Representations (ICLR), 2025 (Spotlight, $< 5.1\%$).

Experience

Lawrence Livermore National Laboratory — Livermore, CA, USA
Computing Graduate Student Intern — *June 2025 – Present*
Mentors: Dr. Bhavya Kailkhura and Dr. Brian Bartoldson

- Working on improving reasoning abilities of large language models through reinforcement learning, continual pre-training and model compression.

North Carolina State University — Raleigh, NC, USA
Graduate Research Assistant; Advisor - Dr. Jung-Eun Kim — *September 2022 – May 2025*

- Discovered that spurious correlations in deep neural networks are learned due to only a handful of all samples containing spurious features.
- Illustrated that attaining information regarding spurious features is often difficult without human intervention, rendering existing state-of-the-art techniques as ineffective.
- Created a data pruning technique to overcome spurious correlations without any domain knowledge or human intervention.
- **Current Research Directions**: Improving scaling laws of deep neural networks in language based arithmetic tasks such as GSM8K and MATH bechmarks; Studying the role of data availability on spurious feature reliance; Identifying the impact of model compression on out-of-distribution generalization.

Sandia National Laboratories — Albuquerque, NM, USA
Graduate Research and Development Intern; Mentor - Dr. Carter Jameson — *May 2023 – August 2023*

- Built language models to identify occurrences of classified information in official government documents.
- Improved existing rule-based entity-linkers deployed within Sandia National Laboratories using Transformer-based language models.
- Utilized Transformer based language models and Question-Answering data from SQuAD2.0 to build robust classifiers.
- Leveraged classical machine learning techniques to build lightweight topic agnostic classifiers.

RoboTutor, Carnegie Mellon University Pittsburgh, PA, USA(Remote)
Undergraduate Research Intern; Advisor - Dr. Jack Mostow *January 2021 – May 2022*

- Reordered the Instructional Sequence of an Intelligent Tutoring System to enhance student learning and engagement with the help of metaheuristic optimization algorithms and machine learning.

UBS Hyderabad, TG, India(Remote)
Summer Analyst Intern *June 2021 – August 2021*

- Built tools to automate Data Engineering tasks that were being performed manually daily.

Awards

NC State Graduate Merit Award, 2024 (3000$)
NC State Travel Grant, 2025 (1000$)
Chancellor Merit Scholarship Award, 2019

Technical Skills

Proficient: Python, PyTorch, NumPy, Bash
Intermediate: Java, SQL, Scikit-Learn, Flask, spaCy