

Rahul Chand

650-509-9551 | rahulsc@stanford.edu | chandrahul0320@gmail.com

Website: rahulschand.github.io | LinkedIn: [/RahulSChand](https://www.linkedin.com/in/RahulSChand) | GitHub: [/RahulSChand](https://github.com/RahulSChand)

EDUCATION

Stanford University

2024-2026

Masters in Computer Science | 4.0 GPA

Birla Institute of Technology, Pilani

2015-2019

Bachelor in Computer Science | 9.7 GPA

INDUSTRY AND RESEARCH EXPERIENCE

NVIDIA | Intern @ GEAR Lab

Sept 2025 -

- Working in the GEAR lab on foundational models for humanoid robots (Gr00t N1).

NVIDIA | Intern @ Nemotron Post-training

June 2025 - Sept 2025

- Worked in the Nemotron LLM post-training team on exploring latent reasoning and async RL for nemo-rl.

Stanford AI Lab (ILIAD)

Nov 2024 -

- Working as a grad research student

Microsoft Research | Pre-Doctoral Fellow

July 2021 - Nov 2023

- Worked on Extreme Multi Label Learning (multi-label classification with million+ labels). I primarily worked on improving performance of extreme classifiers via distillation. [ICLR'24](#)
- Worked with Turing Team on factorization and sparsity methods for Transformer compression. [EMNLP 2022](#)

Arcesium (De-Shaw) | SWE

Sept 2019 - March 2021

- Developed microservices & frontend using Kotlin, React & SQL for handling large volumes (>100k) of trades daily.

Indian Institute of Science (IISc) | Intern

Jan 2019 - July 2019

- Worked at Video Analytics Lab (VAL), IISc on using capsule networks for optical flow estimation. [arxiv](#)

PUBLICATIONS

- Importance Weighted Retrieval for Few-shot Imitation Learning (**CorL 2025 Oral**) ([Link](#))
*Amber Xie, **Rahul Chand**, Dorsa Sadigh, Joey Hejna*
- Enhancing Tail Performance in Extreme Classifiers by Label Variance Reduction (**ICLR 2024**) | ([Link](#))
*Anirudh Buvanesh, **Rahul Chand (co-first author)**, Yashoteja Prabhu, Manish Gupta*
- DSFormer: Effective Compression of Text-Transformers by Dense-Sparse Weight Factorization ([Link](#))
***Rahul Chand**, Yashoteja Prabhu, Pratyush Kumar.*
- CapsFlow: Optical Flow Estimation with Capsule Networks ([Link](#))
***Rahul Chand**, Rajat Arora, Ram Prabhakar, Venkatesh Babu.*

OPEN SOURCE CONTRIBUTIONS

- Compute requirements for LLMs** ([Github](#)) | **1400+** ★
 - Tool to check GPU vRAM requirement and token/s for training & inference of any LLM. Supports frameworks like HuggingFace, vLLM, llama.cpp, and quantization (bitsandbytes, GGML & QLoRA).
 - The tool has been used over **150k+** times by **25k+** users.
- llama2.c for dummies** ([Github](#)) | **200+** ★
 - Walkthrough tutorial of [llama2.c](#) written as a starter reference for LLM inference.
- Nemo-RL** ([Github](#)) | **1000+** ★ (contributor)
 - Contributed to the Nemo-RL post-training library during my internship in the Nemotron LLM team. I worked on async rl
- Nemo-Gym** ([Github](#)) | **70+** ★ (contributor)
 - Contributed to the Nemo-Gym RL environment library

TEACHING EXPERIENCE

TA (Stanford, MS)

- **Winter 2025:** CS153

TA (BITS Pilani, BS)

- **Fall 2018:** Data Mining, Principles of Programming Languages, C Programming
- **Spring 2018:** Data Structures & Algorithms, Database Systems
- **Fall 2017:** Logic in Computer Science