

Harsh Shah

🌐 harsh-sensei.github.io | **in** connecthshah |  Harsh Shah |  hshah2@cs.cmu.edu |  +1 412-708-8829

Tech enthusiast with expertise in machine learning and statistics, driven by collaboration and impactful innovation

Education

Carnegie Mellon University, Pittsburgh, PA

(December 2025)

Master's in Machine Learning

Key Coursework : Deep Reinforcement Learning and Control, Intermediate Statistics, Probabilistic Graphical Models

Indian Institute of Technology Bombay (IITB), Mumbai, India

(May 2024)

Bachelor of Technology in Computer Science and Engineering, Minor in Artificial Intelligence

CPI - **9.43/10**

Key Coursework : Adv. Information Retrieval, Adv. Machine Learning, Financial Engineering, Database Systems

Teaching Assistant : Computer Systems, Reinforcement Learning, Statistical Image Processing

Skills

Programming Languages : C, C++, Python, Prolog, VHDL, Assembly Language (ARM), Java, Javascript

Softwares and Tools : Matlab, Git, Robot Operating System, Docker, Scilab, Flutter, Unity Game Engine, Android studio

Libraries and Packages : PyTorch, OpenGL, Gymnasium, TensorFlow, Flax, SFML, Django, OpenCV, Scikit-Learn

Work Experience

Phynx Labs

(July 2024-August 2024)

ML Engineer

- Developed a system for hierarchical layout analysis of arbitrary flat pdfs along with chunking for multilingual setting
- Built retrieval engine for Retrieval Augmented Generation pipeline incorporating search in hierarchical pdf structure

Quadeye Securities LLP [certificate]

(May 2023-July 2023)

Quant Systems Engineer

- Achieved a **1.3%** performance boost in low-latency **Network Interface Card (NIC)** drivers by implementing **Advanced Vector Extensions (AVX)** instructions and **optimizing cache prefetching** inside linux kernel modules
- Accelerated DNS resolution efficiency by integrating asynchronous DNS lookups, eliminating threading overhead

AFAR lab, University of Cambridge [certificate]

(March 2022-July 2022)

Robotic Systems Developer

- Enhanced open-source development of HARMONI, optimizing the robot interaction pipeline in the ROS framework
- Engineered **behavior trees**, integrating **concurrent multi-threaded modules** for efficient asynchronous processing

Agency for Science, Technology and Research (A*STAR) [certificate]

(March 2022-July 2022)

NLP Research Intern

- Enhanced text-to-SQL models to get **98%** count prediction accuracy by gathering diverse SQL data to finetune BERT
- Accelerated model training using distributed training in pytorch functionalities and CUDA on remote GPU clusters

Projects & Leaderships

CoReader [app] - Flutter application with a growing user base of 100+ active users

Spherical Harmonics Estimation for Mixed Reality [code] - AR object rendering application built in Unity

Team Leader, Mars Rover Team (IITB) [website]

Led a team having **60+** members to achieve *Best Navigation Award* in international competitions (IRC2023, ERC2022)

- **Management** : Managed team and planned funds over **18k USD** to improve rover's **structure** and **autonomy**
- **Technical** : Devised **vision perception system** of autonomous rover and fused with other sensors for localization

Publications

Group Testing for Accurate and Efficient Range-Based Near Neighbor Search for Plagiarism Detection [paper]

- Published first-author research at prestigious A* conference, European Conference on Computer Vision (ECCV) 2024
- Engineered a novel algorithm that optimized insertion time by over 100x, surpassing existing graph ANN methods

Awards

- Secured **national rank 11 (top 0.01 percentile)** in Joint Entrance Examination (Advance), **1 million+** aspirants (2020)
- Among **Nation's top 1%** in National Standard Examination in Physics (NSEP) given by **50K+** candidates (2019)
- Qualified for **Indian National Mathematics Olympiad (INMO)** upon clearing two stages of the olympiad (2018)