Raavi Gupta

Github: raavi02 | raavi.g@columbia.edu | +1(646)464-0221 | LinkedIn: raavi-gupta

Education

Columbia University (CU) New York, NY
*MS in Computer Science (GPA: **4.08/4.00**)* Aug 2024 - (expected) Dec 2025
Graduate Research Assistant: Among 1 of 3 from 400+ students selected for funded research with tuition waiver
Courses: Generative Models for Code, High Performance Machine Learning, Algorithms

Indian Institute of Technology (IIT) Bombay Mumbai, IN
*B.Tech. in Electrical Engineering with Minor in AI (GPA : **9.22/10.00**)* Nov 2020 - May 2024
Courses: Optimization in Machine Learning, Image Processing, Adv. Signal Processing, Reinforcement Learning
Awards: Undergraduate Research Award (5/200+ students)

Technical Skills

Programming Languages Python, C/C++, MATLAB, Bash, SQL, HTML, CSS, JavaScript, VHDL, ARM
Technologies/Frameworks PyTorch, Keras, TensorFlow, NumPy, Pandas, Transformers, Docker, Git, OpenCV

Reseach Experience

IBM Research | Research Intern May 2025 - Present
- Developed a novel large-scale multi-turn, multi-hop, real-time tool-invocable benchmark of **8k+** dialogues to evaluate reasoning capabilities of LLM-based agents, enabling robust testing of tool invocation and context retention
- Augmented Wiki-derived knowledge graph of **4.5M+** entities with tool-use capabilities to support RAG in benchmark
- Designing a public leaderboard to evaluate agents thus fostering standardized benchmarking and community engagement

Research and Technical Projects

Drug Overdose Prediction using Investigative Reports Jan 2025 - Present
Advisor: Prof. Smaranda Muresan | Graduate Research Assistant *Computer Science, CU*
- Designing a human-in-the-loop AI system to detect drug overdose cases from death investigative reports
- Finetuned LLaMA-8B on custom reasoning data thus reducing false negative cases by **80%** compared to manual review
- Accepted for Oral presentation at American Public Health Association 2025 Annual Meeting & Expo

Hallucination Detection of Large Language Models (with Adobe Inc.) Aug 2023 - Aug 2024
Advisor: Prof. Ganesh Ramakrishnan | Bachelor Thesis *Computer Science and Engineering, IIT Bombay*
- **Designed** an algorithm for detecting factual inaccuracies in LLM responses without relying on external databases
- Surpassed the SOTA AUC-ROC by **12%** (among other results) for Mistral-7B answers evaluated on the NQ Open dataset
- Attained **10X** faster hallucination detection per sentence with reduced computational resources compared to SOTA
- First authored a **research paper** | Submitted key findings to a top artificial intelligence conference (under review)

Algorithmic Construction of Lyapunov Functions Jun 2022 - Aug 2023
Advisor: Prof. Debasish Chatterjee | Research Project *Systems & Control Engineering, IIT Bombay*
- Devised a novel method to algorithmically construct Lyapunov functions for nonlinear vector fields
- **Outperformed** the **SOTA** `SOSTOOLS` library in handling non-polynomial continuous black-box vector fields
- Presented a **research paper** at the 2024 Australian and New Zealand Control Conference to **100+ attendees**

InterIIT Tech Meet 12.0: Tooling up for Success Nov 2023 - Dec 2023
Team Lead | DevRev *IIT Bombay*
- **Spearheaded** a team of **12 members** for creating an LLM-planner achieving **0.97** precision in tool-calling
- Implemented **10+ research papers** on fine-tuning, prompt engineering, and data generation to find the optimal agent
- Secured **third position** overall in the competition among **21 IITs** across India with **1000+ participants**

Professional Experience

Airlitz | Columbia Build Lab | Machine Learning Intern Sep 2024 - Dec 2024
- Fine-tuned CLIP on 12.5K+ scraped IMDb images for TV show recognition, achieving **80%** accuracy
- Boosted user engagement by extracting images and product links from web alongside summaries thus enabling visuals
- Utilized Google Vision API to filter appropriate images with **95%+** recall during upload to Airlitz platform

Piramal Capital and Housing Finance Limited | Machine Learning Intern May 2023 - Jul 2023
- Achieved **20.4%** improvement in transaction categorization accuracy over rule-based pipeline by using **LLaMA-7B**
- Augmented knowledge of LLaMA by leveraging web-scraped information of **1M**+ companies thus aiding categorization