Parv Mahajan

parvm@gatech.edu • https://www.linkedin.com/in/parv-mahajan • https://github.com/parviam • (470) 358-0186 • US Citizen

AI Safety researcher with extensive research program management experience. Regular partner with government entities and university defense laboratories. Interested in scaleable oversight, CBRN evaluations, and governance strategies.

EDUCATION

Georgia Institute of Technology - B.S. in Computer Science, Minor in German, **GPA: 3.94**   *Expected: May 2026*

Georgia Institute of Technology - M.S in Computer Science, Specialization in Machine Learning   *Expected: May 2028*

WORK EXPERIENCE

Counterproliferation and Counter-WMD Intern, Georgia Tech Research Institute   *September 2024 - Present*
- Managed experiment portfolio of $63,000 for affecting and detecting potential bio-cyber attack vectors, enhancing security frameworks for 5000+ bioinformatics projects globally through systematic hypothesis testing and validation
- Formulated and red-teamed catastrophic nuclear risk AI system evaluation methodology across 168 infohazard scenarios in coordination with SMEs, broadening defended attack space and reducing evaluation timelines by 31%
- Accelerated project timeline by 44% through contributions in strategic project planning in a fast-paced, detail-oriented research environment by effectively communicating complex findings to sponsors and lab leadership

Artificial Intelligence and Machine Learning Intern, Honeywell   *May 2025 - August 2025*
- Designed and conducted end-to-end ML experimentation including data preprocessing, model training, hyperparameter tuning, and performance evaluation on 7 industrial IoT datasets
- Used distillation and interpretability techniques from modern literature to optimize LLM pipelines for real-time inference on edge devices, reducing compute overhead by 28% while maintaining model accuracy

Artificial Intelligence and Computational Biology Researcher   *August 2023 - May 2024*
- Created novel AI system to synthetically augment biopsy slide segmentation, improving on current models by 3-10%
- Presented work at international biomedical conference attended by over 250 industry and academic researchers
- Communicated research with military officials and area experts at 3 state-wide symposiums and competitions

Director of Data Science, GSMST Counseling Department   *August 2022 - September 2023*
- Empowered 1,254 students by building an individualized grade predictor powered by tree-based gradient boosting
- Led independent research analyzing 54 months of unconventional textual and numerical academic data using time series analysis and presented reports to varied stakeholders

TECHNICAL SKILLS

Computing Languages: Proficient in **Java**, **Python**, **Julia**, **MATLAB/Octave, Bash**; familiar with **C++, Perl**

Technologies: PyTorch, TensorFlow, LLM fine-tuning, Basic mechanistic interpretability, ML experimental design

LEADERSHIP AND RESEARCH

Collaborative Initiatives Lead, AI Safety Initiative   *January 2025 - Present*
- Led cross-disciplinary collaboration between 11+ research partners, driving the development of AI governance frameworks submitted to the National Science Foundation and .
- Spearheaded the organization of regional AI safety workshops and symposia, attracting over 110 participants from academia, industry, and government sectors, to share cutting-edge research and foster partnerships.
- Operated $30,000 research program funding 6 concurrent projects and 26 students, including international collaborations with the Liberian Ministry of Health for a comprehensive AI biosecurity regulatory framework.

Classification on Latent Feature Activation for Detecting Adversarial Prompt Vulnerabilities   *Apart Studio*

Sandbagging Model Organisms for Evaluations of Model-diffing Methods   *Forthcoming*

ADDITIONAL SKILLS AND INTERESTS

Languages: Fluent in **Hindi** and **German**

Fine Arts: Award-winning classical composer; Tenor Saxophonist (7 years)

Interests: Reading; Medium-distance running; Badminton