Keerthana Gurushankar

Email | Personal Website | GitHub | LinkedIn

EDUCATION

Carnegie Mellon University

PhD, Computer Science   Aug 2022 – (expected) May 2027
    Research focus: Performance Modeling for Machine Learning Systems

B.S., M.S., Mathematics, GPA: 3.80   Aug 2017 – May 2021

WORK EXPERIENCE

- **PhD AI Frameworks Intern, AMD**   May 2025 – Present
  Performance Engineering of Post-training optimization of Large Language models (LLMs)
  - 2x sped up LLM inference: profiled & fixed perf bottlenecks by kernel compilation & workload redistribution
  - Scaling LLM Quantization via model sharding, multithreading, multiprocessing & minimizing inter-device comm
- **Machine Learning Intern, CapSen Robotics**   May 2023 – Aug 2023
  Computer Vision models and software for robotics applications
  - Built & tuned segmentation-based ML models for 3D object detection & pose estimation for robot applications
  - Contributed performance improvements to production-level project, reducing failure rate from 1/1000 to 1/3000
  - Designed & developed new calibration method in C++/CUDA, reduced calibration time 6-fold
- **Pre-doctoral Researcher, CMU**   May 2021 – May 2022
  Neural modeling with Information theoretic decompositions
  - Modeled neural data with information theory to detect causal relations, leading to publications in top conferences
  - Deduced tractability results using probability, linear algebra, algorithmic methods & data visualization
  - Collaborated with and presented work to researchers of diverse technical backgrounds

RELEVANT RESEARCH PROJECTS

- **Data Integrity in Clinical Trials - Citadel Global PhD Datathon 3rd Place Winner**   Nov 2024 - Nov 2024
  - Produced a novel large-scale statistical analysis for data integrity in clinical trials (using Python/pandas/scipy)
  - Identified 1k trials with low integrity data among 22k published in 2014-24, using chi-square/Benford's law tests
  - Best among 30 submissions, 1k+ applicants worldwide, awarded $2.5k for "rigorous & highly practical insights"
- **Deep Learning Library Implementations**   Aug 2022 – Dec 2022
  - Built a complete deep learning library from scratch, with basic implementations of PyTorch and Numpy
  - Wrote efficient CPU/GPU backends using C++/CUDA, autodiff support, modules for optimizers & data loaders
- **Statistical Detector for Cortical Spreading Depressions**   Aug 2018 – Apr 2019
  - Designed & implemented (in MATLAB) a statistical detector for anomalous brain waves using ECoG data
  - Modeled & tested Maximum Likelihood Detection, to automate work currently done by expert inspection

SKILLS

Programming Languages: Python (pandas/PyTorch), C/C++/CUDA, SML

SELECTED COURSEWORK

Graduate Artificial Intelligence	Deep Learning Systems	Performance Modeling	Computer Networks
Stochastic Calculus	Probability & Computing	Game Theory	Programming Languages

HONORS/AWARDS

2025 Reproducibility Hackathon - Best Original Data Visualization in city-wide hackathon implementing recent ML paper

2022 Irwin Mark Jacobs & Joan Klein Jacobs Presidential Fellowship - presented prestigious fellowship/funding at MIT

2019 CMU Quantathon - member of winning team solving quantitative finance problem in university-wide competition

2018 Putnam Competition - ranked in top 500 nationally

2017 International Physics Olympiad Selection Camp - top 35 students in Physics nationally

2017 NIOS Senior Secondary Board Examination - Highest score among 300,000 students