

Ayman Mahfuz

ayman.afeef@gmail.com | linkedin.com/in/aymanmahfuz | aymanmahfuzportfolio.com

EDUCATION

The University of Texas at Austin

Austin, TX

Bachelor of Science in Computer Science

Aug 2023 – May 2027

- **Concentration:** AI and ML | **GPA:** 3.5/4.0 | **Courses:** HPC, NLP, Generative Visual Computing, Energy-Efficient Computing

SKILLS

ML/AI: PyTorch, Transformers, Hugging Face, LoRA/PEFT, LLMs, Reinforcement Learning (PPO, MAPPO), RLHF/DPO, Diffusion Models

Systems/Compute: Python, C++, SLURM, Apptainer, Docker, Linux, Git, WandB, AWS, GCP, TACC/HPC

Web/Data: TypeScript, React, FastAPI, Flask, PostgreSQL, BigQuery, Pandas, NumPy, REST APIs

EXPERIENCE

Arm

May 2025 – Present

ML Research Engineer Intern

- Reduced validation cost by ~100x, achieving 99.8th-percentile failure coverage with <1% of the original test budget; shipped a modular Python platform with extensible acquisition-function/kernel APIs, filed a patent, and presented to SVP/director-level leadership.
- Architected an ML-driven test prioritization platform for Arm server CPU validation, using Bayesian optimization with Gaussian-process surrogates to actively select stress-test configurations most likely to surface novel CPU/memory failures.

Texas Robotics, UTCS AI Lab

Jan 2025 – Present

Undergraduate Research Assistant

- Optimized a 400K+ LOC C++ simulator/training stack, cutting RL training time 67%; scaled curriculum training to 5M+ episodes with reproducible GPU configs and CI/CD tests.
- Deployed hierarchical multi-agent policies (dribble, pass, shoot, defend) on NAO robots, placing 3rd globally at RoboCup 2025 SPL.
- Training a Qwen3-4B LoRA advisor for Seq-Advisor, a NeurIPS-targeted Human-AI/RL project; built a dynamic-programming CPR heuristic improving Team 1 surplus from 136 to 259.

UT Austin - Center for Media Engagement

Aug 2023 – May 2025

Software Engineer, Undergraduate Research Assistant

- Engineered ETL/scraping pipelines feeding a 250M+ row BigQuery corpus; deployed BERT and LLM-based classifiers for clickbait and narrative detection.
- Built React/Flask/Firebase experimental platform with randomized assignment and action logging for 1,000+ participants in published behavioral studies.

UT Austin - Oden Institute

Aug 2023 – Jan 2025

Undergraduate Research Assistant

- Built and benchmarked CNN/Transformer 3D MRI segmentation pipelines on TACC H100s across 1,000+ scans, improving Dice by 12% while debugging mixed-precision, memory, and I/O bottlenecks on large-scale GPU systems.

PROJECTS

HelmPM (helmpm.app) – LLM-Native PM App | *OpenAI/Anthropic, React, FastAPI, PostgreSQL*

- Built and launched a multi-tenant LLM-native product management platform used by 4 teams; integrated OpenAI/Anthropic models with retrieval over documents, tickets, and codebases, automated meeting summaries, task generation, and Linear/Jira synchronization.

Tool-Call Mechanistic Interpretability | *PyTorch, TransformerLens, GPT-2, SAE*

- Fine-tuned GPT-2 Small on a 10K-example confound-balanced tool-use task; activation patching/DLA/SAE ablations across 144 heads found a distributed tool/no-tool mechanism.

Modern Transformer LLM with RLHF Alignment | *PyTorch, HPC*

- Built a 253M-parameter decoder-only LLM from scratch with RoPE/RMSNorm/SwiGLU/attention sinks; pretrained on 600M tokens and implemented SFT/DPO alignment.

Attention-Based Diffusion Image Generator | *PyTorch, HPC*

- Built DDPM/DDIM image generation from scratch with a U-Net + transformer-attention backbone; trained on CIFAR and improved sampling speed 20x while preserving generation quality.