Minhal Khan

[redacted] | minhal@minhalkhan.com | linkedin.com/in/~minhal | github.com/uboaaaa | minhalkhan.com

Education

University of Maryland, College Park — College Park, MD
M.S. in Computer Science – GPA: 3.60 — Sep. 2025 – Dec. 2027 (expected)

- Relevant coursework: Database Management Systems, Algorithm Analysis

University of Pittsburgh — Pittsburgh, PA
B.S. in Computer Science, Minor in Mathematics – GPA: 3.72 — Jan. 2021 – Dec. 2024

- Teaching Assistant, Discrete Structures in CS (2022 – 2024)

Projects

Lento Audio Controls Extension | *JavaScript, Web Audio API, GitHub REST API* — Jun. 2024 – Present

- Published a browser extension that slows, reverbs, and bass-boosts music, with 600+ users and a 4.9-star rating.
- Optimized the extension's Web Audio pipeline, cutting idle CPU usage 35% by disconnecting unused reverb nodes and memory leakage by 98% by releasing dead media elements using WeakRef registries.
- Ported from Firefox to Chrome with builds linted and staged by GitHub Actions. Built an in-extension bug reporter filing Google Forms responses as labeled GitHub issues via GitHub REST API.
- Expanded support to 6+ major music-streaming sites, including Spotify, SoundCloud, and Apple Music.

Liquidity Arena Trading Bot | *Python, asyncio, pytest, Amazon EC2, MiniMax M3* — Jun. 2026 – Aug. 2026

- Created an LLM-driven crypto futures trading bot running autonomously on EC2, journaling 3,000+ trading decisions.
- Generated trading signals with a MiniMax M3 model fed by news and market data APIs. Implemented risk monitoring and trading strategy on an asyncio loop, offloading LLM calls to worker threads to avoid blocking risk checks.
- Identified and resolved 7 production incidents. Validated trading strategies through experiments and counterfactual replays against historical candle data, backed by 129 tests.

LLM-as-a-Judge for Audio Reasoning | *Python, PyTorch, TRL/PEFT, vLLM* — Oct. 2025 – Dec. 2025

- Built an AI judge to score audio-language-model answers, reaching 0.82 Pearson correlation with human ratings vs. a 0.21 BERTScore baseline.
- Fine-tuned Qwen3-0.6B using LoRA supervised fine-tuning and GRPO reinforcement learning with custom reward functions, reaching 67.8% exact-score accuracy on a question-disjoint held-out split.
- Assembled a training set of 3,396 human-rated answers from 15 audio-language models. Expanded it 5x with label-preserving GPT-4o-mini paraphrases via an async OpenAI pipeline running up to 200 concurrent requests.

Experience

Machine Learning Research Assistant — Pittsburgh, PA
ADMT Lab, University of Pittsburgh — Sep. 2024 – Feb. 2025

- Co-authored "RAMping up Data Postdiction" (VLDB 2026 submission), a pipeline replacing decayed data columns with regression models achieving 60–91% storage reduction at 1% error.
- Designed the regression-based split criterion adopted in the paper's train-and-split method, clustering 5–10x faster than k-means while improving storage reduction.
- Rebuilt the experiment pipeline into a config-driven benchmarking harness. Converted YAML configs to 500+ automated test runs across clustering algorithms, error thresholds, and split sizes.

Software Engineering Capstone Project — Pittsburgh, PA
Neuraville (sponsor) — Aug. 2024 – Dec. 2024

- Led a 3-person team to design a spiking neural network for real-time hand-gesture recognition for a local robotics startup.
- Encoded webcam videos into spiking activity on FEAGI, Neuraville's neuromorphic platform, building gesture signatures from downsampled frames and spatiotemporal tracking.
- Designed a supervised training circuit with excitatory and inhibitory links, tuning plasticity parameters to diagnose test-phase recognition failures.
- Organized weekly stakeholder meetings with the sponsor's CEO and engineers. Delivered a live demo and authored the project's technical post on the company's blog.

Skills

Languages: Python, JavaScript, SQL
ML and LLM: PyTorch, Hugging Face (TRL, PEFT), LoRA, GRPO, vLLM, scikit-learn, NumPy, Pandas
Infrastructure and Tools: Git, GitHub Actions (CI/CD), REST APIs, pytest, Amazon Web Services (EC2), Linux, PostgreSQL
AI-Assisted Development: Claude Code, agentic workflows, Claude Skills, MCP, LLM evaluation