

Conrad Feng

205-200-8295 | conrad.feng@emory.edu | [linkedin.com/in/conradfeng](https://www.linkedin.com/in/conradfeng) | github.com/conrad204 | conradfeng.ai

EDUCATION

Emory University

B.S. in Computer Science, Expected May 2029

Atlanta, GA

Aug. 2026 – Present

University of Alabama at Birmingham (UAB)

Coursework in Biomedical Engineering; GPA: 3.92/4.0

Birmingham, AL

Aug. 2025 – May 2026

Relevant Coursework: Data Structures & Algorithms (IP), Computer Architecture (IP), Object-Oriented Programming, Linear Algebra, Mathematical Statistics (IP), Calculus III (*IP = in progress*)

EXPERIENCE

Undergraduate Researcher

Emory University – Advised by Prof. Ali Emami

August 2026 – Present

Atlanta, GA

- Building experimental infrastructure in Python/PyTorch for robust data generation pipelines, then extracting and analyzing internal representations across large open-weight language models.
- Research focus: cross-lingual activation probing to test whether evaluation-awareness signals in LLMs are language-dependent.

Machine Learning Researcher

Algoverse AI Summer Research Program

May 2026 – August 2026

Remote

- Led hypothesis generation and built an 11,520-prompt dataset, extracted activations from three Llama models up to 405B via a custom remote-inference (NDIF) pipeline, and designed study to demonstrate that the standard probing setup learns a spurious surface shortcut rather than the target concept of evaluation awareness.

Research Assistant

UAB Division of Nephrology – Dr. Paul Sanders' Lab

May 2023 – August 2026

Birmingham, AL

- Co-authored a **Kidney International** paper (IF 21.8) and **first-authored** an abstract in *Physiology* (APS Summit 2024) by building R/Python pipelines that processed raw RNA-seq from 40+ rat samples and surfaced 100+ significant pathways via DESeq2 and GSEA.

PROJECTS

NullMap | *Python, FastAPI, Elasticsearch, React, TypeScript*

HackMIT 2026

- **2nd place in Elastic's track** - Built a clinical-evidence search engine that surfaces unreported results, indexing **2.05M papers** *OpenAlex* and *ClinicalTrials.gov* into Elastic Cloud via a checkpointed DuckDB-over-S3 pipeline.
- Reduced per-query LLM cost to 0 by precomputing result direction, MiniLM embeddings and effect sizes at index time and answering verdict counts with one ES aggregation.

Emory Course Atlas Model Context Protocol (MCP) Server | *Python, MCP, httpx*

2026

- Reverse-engineered Emory's undocumented course-search API and built an MCP server exposing 5 tools that let Claude answer natural-language questions about courses, prerequisites, and live seat availability.
- Designed a standalone, thread-safe API client with response caching and rate limiting; parsed unstructured HTML/embedded JS into structured data for LLM consumption. Expanding to other colleges using similar APIs.

MNIST From Scratch | *Python, NumPy*

2026

- Built a 3-layer fully-connected neural network in **pure NumPy**—no PyTorch, TensorFlow, or autograd—achieving ~95% test accuracy on 10,000 held-out MNIST images.
- Implemented backpropagation from first principles, hand-deriving gradients for linear, ReLU, and loss layer via the chain rule and caching forward-pass activations for efficient reuse in the backward pass.

Automated RNA-seq Processing Pipeline | *Python, Slurm Docker, DESeq2*

2025

- Built an RNA-seq processing pipeline on UAB's Cheaha HPC cluster, using Slurm job arrays to parallelize STAR alignment across 40+ samples, replacing outsourced analysis (\$1k/dataset)
- Cut per-dataset cost by **98%** (\$1000→\$20) and turnaround from days to under 24 hours by scaling the workflow to 180 GB datasets on AWS EC2.

TECHNICAL SKILLS

Languages: Python, Java, R, C/C++, SQL, TypeScript

Tools & Frameworks: PyTorch, NumPy, pandas, Matplotlib, Plotly, Node.js, Flask, FastAPI, React

Developer Tools: Linux, Git, Docker, Slurm, AWS (EC2), OpenRouter, DESeq2, GSEA