

TATHAGATA DEBNATH

PhD Candidate in Computer Science

(575) 621-1224 · tirtha.debnath@gmail.com · [GitHub](#) · [LinkedIn](#) · [Google Scholar](#) · [Website](#)

SUMMARY

PhD candidate with hands-on experience training and fine-tuning LLMs using DPO/RLHF on multi-GPU clusters, building multi-agent orchestration systems, and designing RAG pipelines. Published algorithm designer (IEEE TPAMI, IEEE/ACM TCBB; 81+ citations) with 2 CRAN R packages. Seeking AI/ML engineering role for summer 2026.

TECHNICAL SKILLS

- **Languages:** Python, R, C/C++, JavaScript, SQL, Bash
- **ML/DL:** PyTorch, DPO/RLHF Fine-Tuning, Hugging Face Transformers & TRL, vLLM, scikit-learn, XGBoost, SHAP
- **LLM/Agents:** LangGraph, LangChain, Claude API, OpenAI API, RAG, FAISS, MCTS for Code Generation, Prompt Engineering
- **Infrastructure:** Docker, Git, Weights & Biases, LangSmith, FastAPI, Streamlit, Multi-GPU Training (H100)

PROJECTS

CodeQ: Self-Improving Code Debugging Agent

[*PyTorch, DPO, MCTS, vLLM, H100*]

- Built an autonomous bug-repair agent combining Monte Carlo Tree Search (MCTS) with Direct Preference Optimization (DPO) self-improvement on Qwen2.5-Coder-7B-Instruct across two shared H100 machines (inference on Machine A, DPO training on Machine B).
- Achieved +2 percentage-point improvement to 38.0% solve rate after Round 1 DPO training, with Round 2 incorporating early stopping, step-150 checkpoint selection, and few-shot prompting for further gains.
- Implemented full training pipeline: MCTS rollout collection, preference pair generation from pass/fail outcomes, DPO fine-tuning with Weights & Biases logging, and automated evaluation on SWE-bench Lite.

Parallel Multi-Agent Code Generation System

[*Python, LangGraph, Claude API, Docker*]

- Engineered a DAG-based multi-agent system with concurrent coder workers orchestrated via LangGraph, achieving 79/79 test pass rate with Docker-sandboxed pytest validation and automatic failure-driven revision loops.
- Extended to a self-evolving variant with LLM-as-Judge prompt evolution: automated scoring of generated code quality drives iterative prompt refinement without human intervention.
- Integrated LangSmith observability for full agent reasoning traces, token usage tracking, and latency profiling.

NutriBot RAG System

[*FAISS, BM25, Reciprocal Rank Fusion, FastAPI*]

- Built a domain-specific RAG system with hybrid search (FAISS semantic + BM25 lexical + RRF reranking), reducing LLM API costs ~50% through MD5 caching, score-threshold gating, chunk filtering, and 512-token output budgeting.
- Designed a 30-question evaluation pipeline measuring Keyword Hit Rate, MRR, and Answer Keyword Coverage with free retrieval-only tuning mode for cost-free iteration.

PUBLICATIONS & SOFTWARE

- Y. Chen, **T. Debnath**, A. Cai, M. Song. "Circular Silhouette and a Fast Algorithm." *IEEE TPAMI*, 2023. [**CRAN: CircularSilhouette**]
- **T. Debnath**, M. Song. "Fast Optimal Circular Clustering on Round Genomes." *IEEE/ACM TCBB*, 2021. [**CRAN: OptCirClust**]
- 10 publications total, 81+ citations — full list at [Google Scholar](#).

RESEARCH EXPERIENCE

Graduate Research Assistant

2020–Present

New Mexico State University ·

- Designed algorithms for circular data clustering (IEEE TPAMI, IEEE/ACM TCBB) and released two CRAN R packages with C++ backends (OptCirClust, CircularSilhouette).
- Developing Pohori, a model-free statistical method for detecting alternative splicing in RNA-seq data.
- Contributed bioinformatics analysis to two biology collaborations (radiation-induced splicing in *Aedes aegypti*, phosphoproteomics), resulting in a *Scientific Reports* 2025 publication.

EDUCATION

- **Ph.D. in Computer Science** Expected 2026
New Mexico State University · GPA: 4.00
- **M.S. in Computer Science** Spring 2026
New Mexico State University · *Focus: AI/ML*
- **M.Tech in Computer Science and Engineering** 2017
Tripura University · *Focus: Computer Vision* · GPA: 9.83/10 · Gold Medalist