

Aniket Ghosh

ghosh.anik@northeastern.edu • Boston, MA • Portfolio • LinkedIn • GitHub

Education

M.S. in Artificial Intelligence, ML Concentration – Northeastern University – GPA: 4.0/4.0 Aug 2025 – May 2027
Courses: Actionable Interpretable Methods, Applied Programming for AI, Foundations of AI, Algorithms.

B.Tech in CS & Business Systems – Institute of Engineering & Management – GPA: 4.0/4.0 Jul 2020 – Jun 2024
Key Courses: Neural Networks, NLP, Data Structures, Database Management, Operating Systems

Technical Skills

ML Frameworks: PyTorch, Hugging Face, TensorFlow, scikit-learn, OpenCV, OpenAI Gym

GenAI & Agents: RAG, LangChain, LangGraph, CrewAI, OpenAI Agents SDK, MCP, ChromaDB, LoRA, QLoRA

MLOps: Docker, Kubernetes, AWS, GCP, GitHub Actions, CI/CD, MLflow, FastAPI, Terraform, SLURM

Core Domains: Mechanistic Interpretability, AI Safety, Reinforcement Learning, CV, NLP, Multi-Agent Systems, Medical Imaging

Languages: Python, C/C++, SQL, Bash

Experience

AI/ML Engineer – Varosync, Inc., Boston, MA May 2026 – Present

- **Sole ML engineer** on an early-stage drug-discovery platform, building the stack from scratch, reporting directly to the CEO/CTO; architected a **biomedical knowledge graph spanning millions of molecules**.
- Trained **molecular-similarity embedding models** over **millions of molecules**, owning the full pipeline from **gold-label curation** through **business logic** to the **user query layer** (featurization, embedding training, evaluation, and **vector indexing**) on a hybrid cloud stack, with **agentic coding tools** (Claude Code, Cursor) in the daily loop.

AI Researcher – CMATER Lab, Jadavpur University, Kolkata Jun 2023 – May 2025

- Built an unsupervised ML pipeline for histopathological cell segmentation using **color quantization** and **DBSCAN clustering**, achieving **85%+ accuracy** and **30% fewer false positives** vs. a Mask R-CNN baseline — enabling faster, annotation-free pathology screening.

Computer Vision Research Intern – North-Eastern Hill University (Remote) Feb 2023 – Jul 2023

- Conducted a **50+ paper literature review** on sign language recognition to define annotation and quality standards for a novel **Indian Sign Language benchmark dataset**, adapting best practices from Microsoft ASL and MNIST to underrepresented language contexts.

Projects

Sufficient Cause Disambiguation (SCD) Mechanistic Interpretability • CS 7180 (PhD-level) • GitHub

- Proved probing accuracy and causal relevance are distinct properties on Llama-3-8B: separator tokens hit 1.000 LDA accuracy across all 32 layers yet drive 0% flip rate under causal patching.
- Built a mechanistic interpretability pipeline on **Llama-3-8B** (LDA + gradient attribution + causal patching); achieved **448× feature compression** at **99.6% accuracy** and **100% prediction flip rate** at $\alpha=2$.
- Audited **LLM resume screening for demographic bias**: name-swapped resumes triggered a **95.8% prediction flip rate** and accounted for **62% of the model's highest-uncertainty Fit decisions** on otherwise-identical content, surfacing systematic name-based bias.

Biomedical Knowledge Graph Link Prediction Healthcare AI • GitHub

- Engineered graph features (PageRank, structural metrics) on the **BioRED** biomedical corpus to train a **Random Forest** link prediction model achieving **0.94 ROC-AUC** — a **23%+ gain** over learned embeddings — and implemented **5-hop reasoning** for relationship explainability.
- Ran a rigorous **classical ML vs. TransE KG embedding** benchmark (0.94 vs. 0.61 ROC-AUC), demonstrating that hand-crafted graph features dominate learned representations on low-resource biomedical corpora — a practical finding for data-scarce production settings.

Autonomous Multi-Agent Trading Simulation GitHub

- Designed and deployed a **multi-agent trading simulation** using the **OpenAI Agents SDK**: **4 autonomous AI traders** coordinated across **6 MCP servers** with **44 distinct tools**, consuming live market data via **Polygon.io** for end-to-end autonomous portfolio management.

Publications & Patents

Intelligent Traffic Sign Detection Using YOLOv8 with Hybrid CNN Filtering *First Author*

Published – AISC 2024 • Springer “Information and Communication Technologies” series

Lightweight Hybrid DNN-GNN for Network Intrusion Detection with Adaptive Late Fusion *Co-author*

Published – IEEE • Patent Filed by SRM Institute of Science and Technology

Achievements & Leadership

- **Selected, BlueDot Impact Technical AI Safety:** Cohort program (2026).
- **Teaching Co-Lead, AIDE Program:** Designing a **12-session ML lab** for philosophy PhD students at Northeastern's AI and Data Ethics Summer Program (Jun–Aug 2026).
- **Graduate TA:** Lead PyTorch and deep learning labs for **40+ graduate students** in NLP at Northeastern University.
- **Academic Excellence:** Ranked **2nd out of 180+ students** in B.Tech; awarded the **Director's Award**.
- **Volunteer Educator:** **7 years** teaching math & CS to **underprivileged students** at The Hope Foundation.
- **Graduate Student Senator:** Elected to Northeastern's **Graduate Student Government**.