

ARINDAM CHOWDHURY

Applied Scientist | LLM Post-Training & Optimization · Agentic AI Systems · Foundation Models · Production ML at Scale

Seattle, WA · +1 (832) 660-6791 · arindam.nitdgp14@gmail.com · [Website](#) · [GitHub](#) · [LinkedIn](#) · [Google Scholar](#) (h-index: 11, 550+ citations)

PROFESSIONAL SUMMARY

Applied Scientist (PhD, Data Science) building and leading agentic AI systems for fraud prevention at Amazon — multi-agent orchestration, LLM-powered automation, and production machine learning (ML) at scale. Research spans foundation models, graph ML, constrained RL, and distributed training; three granted US patents.

TECHNICAL SKILLS

LLM & Post-Training: LLM fine-tuning & adaptation, RAG pipelines, prompt engineering, synthetic data generation, evaluation & ablation design, vector databases (FAISS)

Agentic AI: Multi-agent orchestration, agent-harness design, tool-use & agentic coding workflows (Kiro)

Core ML: PyTorch, TensorFlow, PyG, DGL, Scikit-learn; GNNs, constrained RL, PU learning, sequence modeling

Languages & Systems: Python, C++, UNIX/Linux, Git

Infra & Scale: AWS SageMaker, S3, Lambda, Step Functions, CloudWatch; HPC: ORNL Frontier supercomputer

INDUSTRY RESEARCH EXPERIENCE

Applied Scientist | [Amazon – Buyer Risk Prevention](#) | Seattle, WA

Oct 2025 – Present

- **Lead** a multi-agent orchestrator (planner–executor) that cuts new-market launch effort for Amazon's ML fraud pipelines from weeks to hours: parses dense business-requirement documents into knowledge graphs for task planning, grounds decisions in a persistent cross-launch memory and a fraud-ML knowledge base, and drives a coding agent that generates production launch code with multi-step validation before code review.
- **Co-built** the ML-experimentation and governance components of an org-wide agentic “co-scientist” (orchestrator + multi-agent executor) serving multiple ML, engineering, and PM teams, automating routine scientist operations across experiment execution, Slack communication, and meeting attendance; own the agent that runs ML experiments on internal compute and the governance and behavior-steering layer that prevents leakage of sensitive information and steers agent outputs to match each scientist's communication style; designed to cut routine daily operational effort per person from ~8 hours to minutes.
- **Major contributor** to an org-wide agentic system for user-acceptance testing of ML fraud pipelines: on a PM's launch request, task-specialized agents create test accounts, drive realistic risk behavior, resolve internal authentication to place orders, and validate the pipeline's ML decisions end-to-end, auto-generating a validation report; reduced manual UAT effort by 88%.
- **Lead** a Positive-Unlabeled (PU) learning framework for proactive, early chargeback detection that decomposes label propensity across fraud channels, with an *autoresearch* loop optimizing propensity estimation against downstream AUC-ROC; +280 bps over the production baseline, validated via production A/B testing. [C1]
- **Own** end-to-end ML infrastructure for the worldwide, real-time, low-latency credit-card-testing detection pipeline serving Amazon retail businesses.
- Developing a multi-task sequential-learning model that transfers signal from the label-rich order stage and historical user behavior to the label-scarce payment-method-addition stage, targeting earlier fraud detection where ground-truth labels are sparse.

Applied Scientist Intern | [Amazon – Buyer Abuse Prevention](#) | San Diego, CA

Summer 2022 & 2023

- Built a heterogeneous temporal graph model for real-time fraud detection on large-scale, time-varying user-attribute knowledge graphs; improved recall from 0.95 to 0.98 at production latency SLAs, with ego-graph-based inference reducing wall-clock latency by ~80%.
- Developed attention score-based entity sampling for heterogeneous graph transformers under ~70:1 class imbalance; improved abusive-customer detection recall from 0.59 to 0.63, deployed to production.

ACADEMIC RESEARCH EXPERIENCE

Postdoctoral Research Associate | [Oak Ridge National Laboratory \(ORNL\)](#) | Tennessee

2024 – 2025

- Led a systematic evaluation of fine-tuning strategies for a scalable global-attention graph foundation model (HydraGNN) on leadership-class supercomputers (ORNL Frontier); ablation-driven benchmarking of post-training choices with methodology that transfers directly to LLM fine-tuning. [J2, J3]
- Built interpretable GNN ensemble surrogate models (model explainability + uncertainty quantification) for chemical reaction simulation; $\sim 22\%$ accuracy improvement over baseline. [J4]

Graduate Research Assistant | [Rice University](#) | Houston, TX

2019 – 2024

- Developed a deep constrained-RL framework for sequential resource allocation under episodic constraints; $\sim 17\%$ improvement over greedy baselines; multi-step decision-making and reward shaping directly inform agentic workflow design. [J1, C2]
- Built a GNN + algorithm-unfolding hybrid (WMMSE) for large-scale network optimization; $30\times$ faster inference, 25% throughput gain; low-latency distributed inference at scale. [J6, J7]
- Developed distributed GCN training via hidden-layer partitioning; $\sim 7\times$ speedup on large graphs. [J5]

SELECTED PUBLICATIONS & PATENTS

Journals

- [J1] Non-Myopic Power Allocation in Wireless Networks with Constrained Reinforcement Learning. In preparation, *IEEE Trans. Wireless Commun.*, 2026.
- [J2] A Systematic Evaluation of Fine-Tuning Strategies for Graph Foundation Models in Molecular Property Prediction. Under review, *ACM Trans. on AI for Science*, 2026.
- [J3] When Does Global Attention Help? A Unified Empirical Study on Atomistic Graph Learning. *Journal of Cheminformatics*, 2025.
- [J4] Automated Learning of GNN Ensembles for Predicting Redox Potentials with Uncertainty. Under review, *ACM Trans. on AI for Science*, 2025.
- [J5] GIST: Distributed Training for Large-Scale Graph Convolutional Networks. *J. App. and Comp. Topology*, 2024.
- [J6] Deep Graph Unfolding for Beamforming in MU-MIMO Interference Networks. *IEEE Trans. Wireless Commun.*, 2023.
- [J7] Unfolding WMMSE using Graph Neural Networks for Efficient Power Allocation. *IEEE Trans. Wireless Commun.*, 2021.

Conferences & Preprints

- [C1] RIPE: Risk Inference from Positive and Evolving Labels via Decomposed Propensity PU Learning. Under review, *Amazon Machine Learning Conference (AMLC)*, 2026.
- [C2] Learning Non-myopic Power Allocation in Constrained Scenarios. *Asilomar Conf. on Signals, Systems, and Computers*, 2023.
- [C3] Label Propagation across Graphs: Node Classification using Graph Neural Tangent Kernels. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022.
- [C4] Stability Analysis of Unfolded WMMSE for Power Allocation. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022.
- [C5] ChartNet: Visual Reasoning over Statistical Charts using MAC-Networks. *IEEE Intl. Joint Conf. on Neural Networks (IJCNN)*, 2019.
- [C6] An Efficient End-to-End Neural Model for Handwritten Text Recognition. *British Machine Vision Conference (BMVC)*, 2018.

Patents

- [P1] Digitization of Industrial Inspection Sheets by Inferring Visual Relations. *US Patent 10,970,531*.
- [P2] Information Extraction from Document Images using Conversational Interface and Database Querying. *US Patent 10,936,897*.
- [P3] Systems and Methods for End-to-End Handwritten Text Recognition using Neural Networks. *US Patent 10,839,246*.

EDUCATION

Ph.D., Data Science · Rice University, Houston, TX · GPA: 3.96/4.00

August 2024

Advisor: Dr. Santiago Segarra | Focus: Graph ML, Deep RL, Distributed Learning

M.Tech., Signal Processing · IIT Guwahati, India · CGPA: 9.32/10

June 2017

B.Tech., Electronics & Communications · NIT Durgapur, India · CGPA: 8.70/10

June 2014

GRANTS, AWARDS & SERVICE

- Co-PI, Scalable Generative Graph Foundation Models for Atomistic Materials Modeling, ORNL Laboratory Directorate R&D, US DOE (2024–2025)
- Co-PI, OLCF Frontier Supercomputer Allocation (200k node hours), US DOE Office of Science (2025)
- IEEE SPS Travel Grant, ICASSP 2022
- Reviewer (~ 20 manuscripts): IEEE TSP, TWC, TMC, TCOM, TNNLS, J-SAC; ICASSP, MLSys, ISIT, AMLC