Torsha Majumder

Kolkata, India • tmajumder.work@gmail.com • **GitHub** • **HF** • **LinkedIn** • **Portfolio** • *Relocation: Singapore, Toronto, Vancouver*

Professional Summary

Lead ML Architect & Research Scientist (NVIDIA Grant Recipient)
MS in Computer Science with a background in Computational Physics and 9+ years of professional experience in the USA and India. Spearheading NVIDIA-supported research to architect high-performance **SSL frameworks** for foundational-scale representation learning via **distributed Multi-GPU training (8x A100)**. Expert in architecting **multimodal Agentic Workflows (MAS)** and **DPO-aligned inference pipelines** (SageMaker/Docker). Specialized in bridging the gap between breakthrough scientific research and scalable industrial automation in high-stakes sectors.

Core Technical Skills

- **ML & AI (Research):** Self-Supervised Vision Transformers (DINOv2/ViT) • Representation Learning • DPO Alignment • Knowledge Distillation (BERT/MiniLM) • Anomaly Detection • NLP (spaCy, NLTK) • SciPy
- **Agentic AI & LLMs:** LangGraph • Stateful Cyclic MAS • Tool-Calling • Retrieval-Augmented Generation (RAG) • Prompt Engineering (Gemini 3.1/3.5, Llama 3, OpenAI)
- **Infrastructure & MLOps:** Distributed Training • Multi-GPU Orchestration • NCCL • Inference Optimization • PEFT • AWS (SageMaker, S3, Boto3) • Docker • NVIDIA Brev • MLflow • Zilliz (Milvus) • Supabase (pgvector) • Linux/Bash
- **Languages & Engineering:** Python (PyTorch, TensorFlow, NumPy, Pandas, Scikit-learn) • C • C++ • Java • SQL (PostgreSQL, Oracle) • Flask • React.js • JavaScript
- **Data & Specialized Tools:** PySpark • Dask • LSDB • Astropy • MongoDB • Matplotlib • Git/GitHub • LaTeX

Professional & Research Experiences

Nu Energy India | Lead ML Architect | Kolkata, India **Feb. 2026 – Present**

- **Architected** a proprietary **cyclic Multi-Agent System (MAS)** using **LangGraph** and **Gemini 3.1** to automate **self-correcting** industrial energy and safety auditing workflows.
- Developing a **multimodal ingestion pipeline** utilizing **Gemini Vision** and **Supabase (pgvector)** to automate entity extraction from technical manuals and instrument logs.
- Implementing a **deterministic calculation engine** using **NumPy** with a **Human-in-the-Loop (HITL)** layer, reducing audit report generation from **weeks to days** with a benchmarked **<0.5% error margin**.

The SNAD Team | Lead ML Architect (NVIDIA Grant Recipient) **Apr. 2024 – Present**

- **Architected ASTRA**, a modular SSL framework; engineered **Astra-CLR**, a multi-filter time-series Transformer **trained on 8x NVIDIA A100 80GB GPUs** and scaled to generate embeddings for **~70M light curves**.
- **Engineered** a **multi-view late-fusion architecture** and **"multi-crop"** paradigm within ASTRA; utilized a **PEFT** strategy to achieve **77% fine-tuned accuracy**; research first-authored and **submitted to A&C Journal (arXiv)**.
- Spearheading the **ASTRA ecosystem expansion** by developing a breakthrough architectural advancement targeting **top-tier ML Conferences (2027)** for a **universal scientific foundation model** for photometric data.
- Discovered **8 Superluminous Supernovae (SLSN)** candidates in the ZTF DR8 release via an **Active Learning pipeline** utilizing the **"Pineforest" algorithm**, resulting in a first-author publication in **PoS**.

Verizon | Data Scientist | Sunnyvale, CA, USA **Jun. 2021 – May 2023**

- **Engineered** a scalable vendor recommendation and forecasting engine on **AWS SageMaker**, increasing prediction accuracy by **25%** through **collaborative filtering** and rigorous **A/B testing**.
- **Architected** an end-to-end automation pipeline for document classification and metadata extraction **(NER)** using **ResNet** and **OCR**, utilizing **MongoDB** and **Boto3** for **S3** integration; reducing manual operational effort by **15%**.
- **Led a 5-person cross-functional analytics team** for the FUZE Regulatory platform, overseeing the delivery of data-driven solutions to optimize nationwide business outcomes.

MIT | Research Affiliate | Cambridge, MA, USA **Oct. 2021 – Nov. 2023**

- **Optimized the DASH spectral classification engine**, migrating the architecture from **TensorFlow to Keras** and implementing **batch processing** to support novel transient classes (**Kilonovae, TDEs, AGNs**).
- **Engineered unsupervised ML pipelines** for **anomaly detection** and noise removal across massive **TESS, Kepler, and PLAsTiCC** datasets to identify rare exoplanetary and transient signals.
- **Co-authored research** in astronomical deep learning published at **NeurIPS 2023**.

UT Dallas | Research Collaborator (Independent) | Richardson, TX, USA **Nov. 2021 – Mar. 2026**

- Engineered a **probabilistic neural network** to optimize **MCMC solver initialization** for exoplanetary orbital modeling; second-author research submitted to the **American Astronomical Society (AAS)**.

Selected AI Portfolio

- **Sentinel-Hub (Multi-Agent System): Architected** a **multimodal**, 4-agent autonomous **cyclic MAS** using **LangGraph** and **Gemini 3.1** to orchestrate cross-domain security forensics; integrated vision, transactional, and log-based **microservices** into a **stateful**, **self-correcting reasoning graph**.
- **DeepShield (AI Safety & DPO):** Aligned a **DINOv2** backbone via a **DPO-based PEFT approach** to prioritize human-perceived risk, achieving **97.5% sensitivity** on high-fidelity fakes and a **50.6% reduction** in False Positives.
- **LogSentinel (Infrastructure Security): Engineered** a **zero-day anomaly detector** using **distilled BERT** and **Zilliz (Milvus)** to perform **low-latency semantic analysis** on high-throughput HDFS logs; utilized **SSL representations** to identify semantic deviations from normal system behavior.
- **NeuralAudit (Financial SSL): Developed** an **SSL-based** tabular anomaly detector using a **bottleneck autoencoder** and **Supabase (pgvector)** to identify fraudulent transactions in **high-dimensional latent space** via Euclidean distance.

Education

MS in Computer Science (Thesis) | **The University of Texas at Dallas**, USA | 2018 – 2020
Thesis: Ensembles of Oblique Decision Trees (ML & Algorithm Design)

Graduate Research (Physics) | **The University of Lethbridge**, Canada | 2024 – 2025
Transition-Edge-Sensor (TES) optimization for Double Fourier Interferometry; Instrumentation Group.

B.Tech in Information Technology | **Maulana Abul Kalam Azad University of Technology**, India | 2013 – 2017
Project: Institute Library Management System (DBMS)

Publications

- **Majumder, T.**, et al. (In Prep). **Large-Scale Self-Supervised Pre-training for Representational Learning in Time-Domain Astronomy**. Targeting ***top-tier ML conferences (2027)***.
- **Majumder, T.**, Malanchev, K., Ishida, E. E. O. **Multi-Scale Contrastive Attention for Light-Curve Representation Learning**. Submitted to *Astronomy and Computing **(A&C) Journal***, July 2026. **arXiv:2606.31627**.
- Huang, H., ... **Majumder, T.**, et al. **Predicting the Age of Astronomical Transients from Real-Time Multivariate Time Series**. *Neural Information Processing Systems **(NeurIPS 2023)***. **arXiv:2311.17143**.
- **Majumder, T.**, et al. **Superluminous Supernova Search with PineForest**. *Proceedings of Science **(PoS)***, PoS(MULTIF2025)041. Accepted Jan 2026.
- Schussler, J., **Majumder, T.**, Penev, K. **Tidal Dissipation Leading to Circularization in Kepler Binaries**. Submitted to the *American Astronomical Society **(AAS)** Journal*, Apr 2026.
- **Majumder, T.** (2020). **Ensembles of Oblique Decision Trees**. *UTD Master's Thesis*. URI: hdl.handle.net/10735.1/8818.

Awards & Honors

- **NVIDIA Academic Grant Recipient** | NVIDIA | Oct. 2025 – Mar. 2026
- **Graduate Research Award (ULGRA)** | The University of Lethbridge | 2024
- **Associate Member (AMIE)** | Institute of Engineers (India), Computer Science & Engineering | 2017
- **Best Final-Year Project Award** | Institute Library Management System, STCET | 2017

Talks & Mentorships

- **Selected Talks: *MIT* TESS Science Talk**, Cambridge, MA (2022); **Int'l Meeting on Eclipsing Binaries**, China/Online (2024)
- **Mentorship:** Led **Data Analytics & FinTech bootcamp programs** at 2U Inc./edX for US/Australian universities (2020 – 2023).
- **Instruction: Graduate Teaching Assistant** for Biophysics and Engineering Mechanics at The University of Lethbridge (2024 – 2025).