

Aritra Bandyopadhyay

📍 India ✉ aritraxban@gmail.com ☎ +917890044771 🌐 Personal Website in aritraban 📺 Techie5879

Education

Indian Institute of Engineering Science and Technology (IIST)

Nov 2021 – June 2025

Bachelor of Technology in Computer Science

- GPA: 9.87/10.0 (Upto 7th Semester)
- **Coursework:** Machine Learning, Data Science, Theory of Computation, Discrete Structures, Algorithm Design

Experience

Research Fellow

Berkeley, USA (Remote)

Supervised Program for Alignment Research

Feb 2025 – Present

- Accepted as SPAR Spring'25 Research Fellow
- Working under Shashwat Goel (Max Planck Intelligent Systems) in post-training interventions and model differences
- Working under Jacek Karwowski (Oxford University) in understanding and verifying the autoregressive conditioning hypothesis for LLMs in chess-playing

Machine Learning Engineer

California, USA (Remote)

Brighterway (YC S24)

Dec 2024 – Present

- Engineering OCR pipelines for medical-legal records, improving information extraction from patient-case records
- Fine-tuning LLMs and VLMs for medical-legal document layout recognition, optimizing specialty-specific data extraction

Machine Learning Research Intern

SC, USA (Remote)

Artificial Intelligence Institute of University of South Carolina (AIISC)

Sept 2024 – Present

- Investigating Safety Alignment & adversarial robustness of Text-to-Image models through training-free model editing and fine-tuning methods (LoRA, DPO)

ML and Computer Vision Research Intern (DAAD-WISE'24)

Munich, Germany

Technical University of Munich (Klinikum rechts der Isar)

May 2024 – Sep 2024

- Developed a novel point cloud registration technique for multi-modal (ultrasound-tactile) data, using coarse-to-fine alignment and learned similarity metrics (LC^2 , DISA) for improved 3D registration
- Investigated the effectiveness of different registration baselines (ICP, CPD) for improving 3D path-transfer accuracy between ultrasound and other modalities (Tactile, CT, MRI)

Machine Learning Engineer

California, USA (Remote)

Mercor

May 2023 – Aug 2024

- Architected and deployed LLM-driven agents, including an AI Interviewer and a Github Analyzer, driving a 150% increase in quarterly revenue and a 10x expansion of the customer base
- Led development of hybrid Information Retrieval pipeline boosting search relevance by 60% and reducing latency by 40%
- Built SFT & evaluation frameworks for Vision-Language Models, with semantic chunking for 30% higher parsing accuracy

Machine Learning Research Intern

Kharagpur, India

Indian Institute of Technology (IIT), Kharagpur

May 2023 – Nov 2023

- Worked on Time-Series analysis & fraud detection in Smart Grids, focusing on explainable AI techniques like LIME and SHAP, achieving 97% agreement with domain experts
- Implemented Attention Pattern Visualizations for justifications of predictions for Transformer-based models

Quantum Computing Research Intern

Kolkata, India

Indian Statistical Institute (ISI), Kolkata

July 2023 – Sept 2023

- Implemented quantum algorithms for Cryptography, Fourier analysis, and Iterative HHL with Quantum Phase Estimation

Publications

Differential Evolution Algorithm Based Hyper-Parameters Selection of Convolutional Neural Network for Speech Command Recognition *In Proceedings of the 15th International Joint Conference on Computational Intelligence (IJCCI-ECTA 2023)*

S Dhar, A Sen, **A Bandyopadhyay**, ND Jana, A Ghosh, Z Sarayloo [10.5220/0012251500003595](https://doi.org/10.5220/0012251500003595) [🔗](#)

Open Source Research

Research Scientist

USA

Manifold Research Group (Open Source Research Lab)

Nov 2024 – Present

- Member of Metacognition Core Team (working closely with 3 other members) on self-evaluation of multimodal models (LLMs, VLMs, VLAs) and improving reasoning at test-time

Projects

BrainQuery

[BrainQuery](#) 

- Full-stack semantic search web-app for arXiv, implementing hybrid search (dense and SPLADE sparse embeddings) for enhanced retrieval accuracy (Pinecone, Flask, ReactJS)

MOVIE++

[MOVIE++](#) 

- Full-stack movie recommendation web-app that leverages item-based collaborative filtering and SVD matrix factorization trained on the MovieLens 25M dataset (Surprise, Pandas, Scikit-learn, Flask, ReactJS)

Achievements

DAAD WISE 2024 Scholar: One among 200 in India to receive the prestigious DAAD WISE Fellowship

Perfect 10/10 Grade in 2 Semesters: Scored a perfect 10/10 SGPA in my 6th and 3rd Semesters

Code For Good'24 - JPMC: Among the 3,000 in India selected for JPMC's Code For Good event

1st in BrainDead: Team secured 1st place in IEST's Inter-University ML Hackathon (Animated Emotion Recognition)

5th in RootAccess: Team secured 5th place in IEST's Inter-University Capture-the-Flag (CTF)

Research Interests & Technical Skills

Research Domains: Mechanistic Interpretability, Alignment of Foundation Models, Vision-Language Models (VLMs), Model Editing, Agentic Systems, Computer Vision (3D Scene Understanding, Semantic Segmentation), Continual Learning

Technical Skills: PyTorch, Transformers (Hugging Face), TransformerLens, TensorFlow, Keras, OpenCV, Scikit-learn, SciPy, Pandas, NumPy, GCP (App Engine, Cloud Run, Cloud Functions), AWS (EC2, S3), Azure, SQL, C, C++, Linux