Arnesh Banerjee

Machine Learning and AI Engineer

+91 8902037524 | arneshbanerjee24@gmail.com | arnesh24 | ArneshBanerjee | portfolio

Education

Heritage Institute of Technology **2023 – 2027**
B.Tech in Computer Engineering (Data Science), CGPA: 8.9 *Kolkata, India*

Research Experience

Indian Institute of Technology Kharagpur **May 2026 – Present**
Summer Research Intern, with Dr. Sourangshu Bhattacharya *Kharagpur, India*

- Architected an improved MoXGATE multi-omic network (**0.827** to **0.927** accuracy) and a LightGBM pipeline from scratch, training **2,040** models each across all **63** combinations of **6** TCGA data types from **528** cancer patients.
- Ran ablations showing DNA methylation alone predicts oral vs. non-oral tumour origin at **0.937** accuracy, then designed confound tests ruling out HPV and hospital effects (**16/16** HPV-negative laryngeal tumours correct).
- Implemented TreeSHAP attribution tracing **87%** of the model's decision to methylation probes in the developmental HOX and PITX2 genes, then engineered a compact **50**-probe signature that retains **0.935** accuracy.

Jadavpur University **Nov 2025 – May 2026**
Research Intern, CMATER Lab, with Prof. Debotosh Bhattacharjee *Remote*

- Designed and coded a hybrid CNN-Transformer (ResNet-34, ASPP, Transformer bottleneck, SE-gated skips) with a novel differentiable Chan-Vese loss, reaching **0.9716** Dice / **0.9463** IoU on DMR-IR (**357** images, **119** patients, **5**-fold CV).
- Implemented and trained Attention U-Net, UNet++, DeepLabV3+ and TransUNet baselines, then proved all converge to ∼**0.97** Dice ($\mathbf{p > 0.05}$) with Wilcoxon tests and bootstrap CIs, exposing an annotation-quality ceiling in labels.

New Jersey Institute of Technology **June 2025 – Nov 2025**
Summer Research Intern, Learning-based Decision Making Lab, with Dr. Arnob Ghosh *Newark, USA*

- Curated a **2,500**-pair safety dataset of benign, context-sensitive and jailbreak prompts, and fine-tuned the final **6** layers of LLaMA-2-7B-chat into a semantic cost model reaching **92%** agreement with human safety judgments.
- Built the evaluation pipeline: scored **0.91**–**0.96** on XS-Test vs. **0.07**–**0.32** for Safe-RLHF; across **1,000**-prompt and **1,172**-pair evaluations the policy flagged unsafe responses **5**× more efficiently and won **60%** (**+70** Elo) vs. SACPO.

Heritage Institute of Technology **Oct 2024 – May 2025**
Undergraduate Researcher, Dept. of CSE (Data Science), with Prof. Arpita Talukdar *Kolkata, India*

- Built a dual-stage ML pipeline for breast cancer diagnosis (WDBC, **569** patients) and recurrence prognosis (WPBC, **198** patients), implementing RFE and SFS wrapper selection for **5** classifiers tuned with **10**-fold GridSearchCV.
- Raised recurrence accuracy from **75.71%** to **93.67%** (SVM + RFE, **18** features) and diagnosis to **97.77%** (Logistic Regression, **13** of **30** features), applying train-only SMOTE to lift minority recall **8**–**12%** without data leakage.

Projects

GPT-2 in Assembly | AArch64 Assembly, NEON SIMD, Apple Silicon **Code**

- Wrote GPT-2 small (**124**M) inference in ∼**1,800** lines of hand-written ARM64 assembly with no C or libc, loading Hugging Face safetensors via `mmap` and hand-coding the BPE tokenizer, KV-cached attention and exp/GELU/softmax.
- Reached ∼**100** tokens/s on one M1 core (**2.5**× PyTorch on the same CPU) with 8-accumulator NEON kernels and startup weight transposes, near the core's memory-bandwidth limit, while matching PyTorch on **160/160** greedy tokens.

Kimi K3 For All | Python, PyTorch, CUDA/Triton **Code**

- Rebuilt the training pipeline for Kimi K3, a **2.8T**-parameter MoE released inference-only, fixing **4** training blockers, and trained a **1.27**B (**364**M active) model on **4.6**B tokens on one H100, cutting validation perplexity from **55.6** to **13.89**.

MLCode | Python, PyTorch, NumPy, Jupyter **mlcode.tech**

- Launched a free ML learning platform and curated its **188** coding problems across **21** topics (**59** easy, **96** medium, **33** hard), from linear algebra to GNNs and LLMs, plus candidate-reported company interview questions.

Publications

- A. Bhattacharjee, **A. Banerjee**, A. Talukdar. *Recursive and Sequential Wrapper-Based Feature Selection for Breast Cancer Diagnosis and Prognosis.* **AGC 2026**. [Accepted & Presented]

Technical Skills

Languages: Python, C++, Go, Rust, TypeScript, JavaScript, SQL, Java
Frameworks & Tools: PyTorch, TensorFlow, Scikit-learn, LightGBM, CUDA/Triton, FastAPI, Next.js, React, Docker, Git
Open Source: Merged PRs to CPython, Rust, Zed, OpenCV and Kubescape (all PRs)