

A. K. M Foyshal Sarkar

Research Assistant / Research Engineer | Machine Learning for Brain Networks (fMRI) and NLP | PyTorch

Dhaka, Bangladesh | +880 1706-929192 | akmfoyshalsarkar@gmail.com
portfolio-foyshal-sarkar.vercel.app | github.com/pialXVII | linkedin.com/in/foyshal-sarkar

SUMMARY

Computer Science graduate applying machine learning to brain-network and language data. Co-authored a 4-person B.Sc. thesis adding a learnable functional-connectivity bias to transformer attention for 7-class HCP task-fMRI decoding ($97.67 \pm 0.08\%$, 3 seeds, three-way ablation), and builds tested Python data and API tooling.

EDUCATION

B.Sc. in Computer Science and Engineering, BRAC University, Dhaka Sep 2022 - Sep 2026
CGPA 3.71 / 4.00 | University Merit List Scholarship | Coursework: Machine Learning, Artificial Intelligence, NLP, Algorithms
HSC (2021) and SSC (2019), Science: GPA 5.00 / 5.00 in both

RESEARCH

GC-JEPA: Graph-Conditioned Attention for Task-fMRI Classification - B.Sc. thesis (team of 4) Oct 2025 - Jun 2026
PyTorch, Hydra | Supervisor: Prof. Chowdhury Mofizur Rahman; co-supervisor: Zayed Humayun, BRAC University | Code: github.com/shoumikarc/jepa-research (team repository)

- Question: can a transformer use a subject's functional-connectivity matrix directly inside attention, and does self-supervised pretraining reduce the labels it needs?
- Method: $\text{softmax}(\mathbf{QK}^T/\mathbf{Vd} + \alpha\mathbf{G})\mathbf{V}$ with $\mathbf{G}$ the 400-region Pearson connectivity matrix and one learnable α per layer (4 extra parameters); GC-JEPA adds connectivity-prediction pretraining before fine-tuning. Data: HCP via NeuroGraph, 7,443 graphs, 7 tasks.
- Results: $97.67 \pm 0.08\%$ test accuracy over 3 seeds (supervised), above the best NeuroGraph-benchmark result (96.25% , not re-run on our split); GC-JEPA $96.73 \pm 0.17\%$. All variants reached about 95% with 25% of labels, so pretraining gave no significant label-efficiency gain on HCP.
- Evaluation: three-way ablation, learned- α analysis per layer, label-efficiency (10-100%) and noise-robustness tests, validation-based checkpoint selection, seeded runs with deterministic cuDNN.
- Open questions we reported: no supervised $\alpha = 0$ control under the same schedule, a single graph-level split, and a small self-supervised margin that did not replicate in an independent re-run.
- My part: added the connectivity bias to attention, analysed region relationships, ran hyperparameter tuning and fine-tuning, and typeset the thesis in LaTeX.

PROJECTS

Question Topic Classification - NLP benchmark (course project) Sep 2025 - Dec 2025
TensorFlow/Keras, scikit-learn, TF-IDF, Word2Vec | github.com/pialXVII/Natural-Language-Processing

- Compared 18 text classifiers (9 architectures, baseline and grid-tuned) on 153k questions in 10 classes; tuned on a held-out validation split, the best (GRU) reached 0.666 macro-F1 on 60k test questions vs 0.655 for TF-IDF logistic regression.

Phone Advisor - Local RAG and Multi-Agent System (take-home assessment) Aug 2026 - Sep 2026
Python, sentence-transformers, FAISS, Qwen2.5-1.5B, CrewAI, FastAPI | github.com/pialXVII/SamsungAI-PhoneAdvisor

- Built a fully local RAG system with intent routing (SQL rankings, two-record comparisons, aspect-filtered FAISS lookups), grounded a CrewAI crew in verified database records, and measured it on a 188-question held-out evaluation (intent accuracy 78% to 96%).

ENGINEERING EXPERIENCE

Freelance Full-Stack Developer | Repair By Engineer, Dhaka Sep 2026
Next.js 16, PostgreSQL, Vercel | Live: repair-by-engineer-inky.vercel.app

- Built and deployed a repair shop's booking site and staff console on a 10-table PostgreSQL schema with server-side role checks and an audit log.

SKILLS

Machine learning: PyTorch, TensorFlow/Keras, scikit-learn, Hugging Face Transformers, sentence-transformers, FAISS, NumPy, pandas, Jupyter

Methods: self-supervised pretraining, attention mechanisms, graph-structured data, ablation design, multi-seed evaluation

Engineering: Python, SQL, TypeScript, FastAPI, PostgreSQL, Git, GitHub, unittest, Hydra configs

Spoken: Bengali (native), English (professional), Mandarin Chinese (basic)

References: available on request