# SHINJINI CHAKRABORTY

Phone: +91 62890 73166 Kolkata, West Bengal, India
Email: shinjinic05@gmail.com LinkedIn: linkedin.com/in/shinjini-chakraborty
GitHub: github.com/shinjinihehe Website: shinjinichakraborty.com

## EDUCATION

**Institute of Engineering & Management, Kolkata, India** Jul 2023 – Jul 2027 (expected)
BTech, Computer Science and Engineering (AI and Machine Learning) CGPA 9.2

**South Point High School, Kolkata** 2023
Class XII; 90.0%

**South Point High School, Kolkata** 2021
Class X; 98.4%

## PUBLICATIONS

**When Explainability Goes Silent: Diagnosing XAI Failure Modes in Medical Imaging**
*IEEE IPTA 2026* (to appear; first author)

- Faithfulness test of 6 explanation methods (Grad-CAM, Grad-CAM++, Integrated Gradients, LIME, SHAP, Occlusion) on 7 CNN, Transformer and VLM models (incl. ResNet-50, DenseNet-121, ViT-B/16, BiomedCLIP) over BUSI and Kvasir-SEG.
- Only Integrated Gradients and SHAP keep Deletion AUC below 0.43 on every classifier; Grad-CAM on ViT and VLM backbones scores a misleadingly good ≈0.24, and LIME collapses (0.84–0.94) on supervised backbones but not on BiomedCLIP (0.45).
- Method rankings did not carry over to ISIC-2018 or DRIVE. Released the PyTorch implementation and benchmark at github.com/shinjinihehe/xai-failure-modes.

**The Quantized Blind Spot: Task-Asymmetric Collapse in Foundation Vision–Language Models for Medical Imaging**
*IEEE IPTA 2026* (to appear)

- Evaluation harness for 7 open VLMs (3–26B parameters) at 5 bitsandbytes precisions on 10 medical datasets; for segmentation, the VLM's boxes prompt a frozen SAM2.
- At 4-bit (NF4), classification lost at most 3.7 accuracy points but segmentation Dice fell by up to 11.8 points; bounding-box parse-failure rate tracked Dice loss across 24 model-precision cells (Pearson $r = 0.84$).

**SAFE-Bench: A Sensor-Aware Framework for Evaluating Detection and Vision-Language Models**
*IEEE IPTA 2026* (to appear)

- Detectors and VLMs (YOLO family, SAM, Grounding DINO, YOLO-World) see only RGB frames while radar, GNSS and IMU signals enrich the evaluation, across 3 sensor configurations on the TIAND driving dataset.

**HiLEx: Image-Based Hierarchical Layout Extraction from Question Papers**
*ICDAR 2025, Springer LNCS*

- Layout benchmark of 1,965 exam-paper pages from 8 competitive exams, with 18,152 boxes in 6 hierarchical classes, released in COCO and YOLO formats at hilex-dla.github.io/HiLEx; pp. 485–505, doi:10.1007/978-3-032-04627-7_28.

**CultureHaystack: Probing Per-Image State and Attribute Binding in Multi-Image VLMs**
*WACV 2027 Workshop on Latent Visual Reasoning (LVR)* (poster, to appear)

- Benchmark of 160 targets drawn from an 890-image pool covering 123 countries: 6 open VLMs (3–12B) see up to 5 photos and must name one photo's item and country (36,576 model calls); dataset at kaggle.com/datasets/shinjinichakraborty/world-cultural-items-1k.
- 7B-plus models stay at 0.69–0.86 joint accuracy at every haystack size, while 3–4B models keep only 53–65% of the answers they get right on one photo; 43–80% of wrong answers attach one true attribute to the wrong photo.
- Ran the post-freeze probes on RunPod A40 GPUs: activation patching on LFM2.5-VL and Qwen3-VL (restoring the target's clean image tokens never significantly helped), options-first prompting, image tags, and a 2×2 collage that lowered accuracy in 5 of 6 models.

## RESEARCH EXPERIENCE

**Data Science Research Intern** Jul 2025 – Jun 2026
Centre of Excellence for Data Science, IEM Kolkata Kolkata (hybrid)

- Co-authored HiLEx (ICDAR 2025); released the dataset in COCO and YOLO formats (1,378 / 388 / 199 train / valid / test pages) and built and maintained the project website (15 commits), adding the Springer link and citation after acceptance.
- Built and maintained the centre's website across 3 repositories (Next.js, TypeScript, Tailwind, shadcn/ui): mentors, students, publications, projects and events pages driven by JSON content; live at ceds-main-rho.vercel.app.

**Multi-agent driving without traffic signals** Oct 2026
Supervised research, Institute of Engineering & Management, Kolkata; Vega HPC cluster

- Built a 41-junction city environment on MetaDrive and an independent per-agent PPO-Lagrangian trainer, then ran two pre-registered studies with Slurm and Apptainer.
- Across 10 multi-agent PPO variants on 7 junction scenes (350 runs), MAPPO-Lagrangian reached a 0.906 mean arrival rate against 0.705 for independent PPO; independent PPO-Lagrangian froze in 78 of 105 runs.
- Safety budgets cut first-come-first-served accuracy at an isolated crossing from 0.94 to 0.83–0.87, and no fleet switched to a different convention.

**Junction negotiation with constrained multi-agent RL** Sep 2026 – Present
Lead researcher, Institute of Engineering & Management, Kolkata; FRIDA HPC cluster

- Sole author of a safe-RL environment suite (22 commits): 8 three-dimensional driving scenarios on real road geometry with a lane graph, a 5-vehicle fleet with per-type cost budgets, actuator lag and sensor noise, and every cost channel (wrong-way, lane-keeping, jerk, time-to-collision) bounded to $[0, 1]$ per step; made it $\sim 2.1\times$ faster on one core and added a multi-process wrapper.
- Built junction-env on MetaDrive: random unsignalised T and X junctions fed by parking lots, 2–3 lanes per direction, 45–90° crossing angles and 4 real-car vehicle classes, trained with shared-network PPO under a traffic-density curriculum.
- Trained the shared policy for 300M car-steps on a Slurm cluster (8.5 h) to 84% arrivals at the densest traffic level, after a trip-level diagnosis traced an earlier 38% plateau to a policy that only drove straight.
- Now leading a pre-registered follow-up on 100 procedural junctions (80 train, 20 held out) with onboard-sensor agents (lidar and camera), comparing MAPPO and MAPPO-Lagrangian.

**Student Researcher** Aug 2024 – Jun 2025
IIC – IEDC, Institute of Engineering & Management Kolkata (on-site)

- Built and deployed a disease-diagnosis chatbot with retrieval-augmented generation (RAG) and NLP models: 85% diagnostic accuracy, tested by 500+ users.

## INDUSTRY EXPERIENCE

**Full Stack Developer Intern** Jul 2026 – Present
Tech for Social Good Foundation (freelance developer from Jun 2026) Kolkata

- Core engineer on Appo Workspace, a Zoho clone being built for NGOs: a multi-tenant HRMS/ERP suite of 12 apps (People, Payroll, Recruit, Performance, Expense, CRM, Finance, Projects, Help Desk, Policies, e-Sign, Workflows) with web and mobile apps; author of about 2,000 commits (about 60% of codebase activity) across Laravel, a Next.js mobile web app and a Flutter shell.
- Built 3 apps from scratch: a no-code Workflows approval builder, Surveys/Forms on the same engine, and Monitor for activity tracking.
- Led the Expense suite redesign (expenses, claims, trips, advances, multi-level approvals, status corrections with an audit trail) and built the e-Sign app with CA-issued digital signatures on PDFs.
- Designed and rolled out a simpler access model (primary admin, per-app admins, employee self-service) across 12 apps and the mobile API, replacing a legacy role and permission system.
- Ran a security audit of the web and mobile apps and fixed every finding: unauthenticated code execution, authentication bypass, SQL injection, public data exposure and an open proxy.
- Built a super-admin email editor that makes 171 of 249 system emails editable without code, with preview, test send and history.
- Shipped HR features end to end (leave rules, attendance regularisation, flexible and week-pattern shifts,

geolocation attendance report, configurable employee IDs, subscription lifecycle) and owned releases: GitHub Actions deploys to staging and production and versioned release notes.

**Software Developer Intern** Mar 2026 – Apr 2026
Deep Duo Foundation Kolkata (full-time)

- Front end for MarketOS, an 18-node LangGraph marketing pipeline in which a supervisor agent turns one campaign brief into a plan that 17 specialist agents carry out, from copywriting to WhatsApp and voice sends; agents talk over Kafka behind a FastAPI service.
- Imported the product's neo-brutalist design system from Google Stitch and shipped the MarketOS landing page v1.0 in Next.js.
- Split the app into separate marketing and application layouts, added routing for the agent console, and fixed the build errors blocking Railway deployment.

**Software Developer Intern, TradeSim** Dec 2025 – Mar 2026
IEM Research Foundation (IEMLabs) Kolkata (on-site)

- Lead front-end developer of TradeSim, an AI-assisted paper-trading platform on live NSE prices (about 160 commits across 3 repositories; Next.js 14, TypeScript, Tailwind).
- Built trading rooms with live market data on TradingView Lightweight Charts, an order book, and buy/sell flows including short selling, stop-loss and stop-limit orders.
- Co-built the AI stock-analysis layer with the team from the planning phase: a FastAPI service pulls 5 years of weekly NSE candles from Upstox, computes TA-Lib indicators and ATR-based risk levels, adds news headlines and asks Gemini 2.5 Flash for a report fitted to the investor's profile.
- Separated realised and unrealised P&L, built leaderboards, admin views to monitor participants, Zod-validated authentication, and the AI analysis page with investor snapshots and deep-dive cards.

**Software Developer Intern** Jul 2025 – Mar 2026
Young Architects Remote

- Two 300-hour tenures on client projects; named 2nd top performer of the batch.
- Lead contributor to Shobdodhara Prakashani, a Bengali publisher's book store: storefront (89 commits) with cart, stock handling, free shipping, platform and delivery fees, cash on delivery, Shiprocket order tracking and catalogue filters and search.
- Built the store's admin panel (46 commits): product CRUD with image upload, coupons, order management with shipping-label generation, and React Final Form validation.
- Built a configurable CRM for legal and e-commerce workflows (case management, client notifications, document storage) and landing pages for client sites.

**Intern – Frontend Development** Jul 2025 – Aug 2025
BTR Communication Kolkata (hybrid)

- Built the React front end of a learning management system for the West Bengal State Police Academy.
- Refactored the UI of a private hospital's Laravel/PHP website for consistency and front-end performance.

## STARTUP AND LEADERSHIP

**Core Team Member** Jun 2025 – Mar 2026
Solyrix, a student startup, Kolkata

- Autonomous AI calling agents for banking, financial-services and insurance firms; teammates built the voice agent and call backend.
- Revamped the solyrix.in landing site: UI v1.0, feature and FAQ sections, brand fonts and a dark-mode visibility fix.

**Founding Team Member** 2025 – 2026
TradeSim, IEMLabs (faculty-founded)

- Co-built a faculty-hosted paper-trading competition platform on live NSE prices from the planning phase, leading the trading interface and contributing to the AI stock-analysis layer.
- Pitched TradeSim at the IIM Calcutta Seva First Innovation Challenge (2026).

## SELECTED PROJECTS

**Zero-shot long-tail video classification** Aug 2026

Reproduction study on UCF101; X-CLIP and FROSTER (ICLR 2024)

- Zero-shot X-CLIP on the UCF101 test split: 72.5% top-1, 91.4% top-5 and 68.8% macro F1, with accuracy reported separately for many-, medium- and few-shot classes (78.9 / 70.9 / 72.6%).
- Reimplemented FROSTER's frozen-CLIP teacher setup: 69.5% top-1 and 92.5% top-5, 3.9 top-1 points below X-CLIP in the same pipeline, with the largest drop on few-shot classes (64.3%).

**IndoExamBench** Jun 2026 – Present

Multilingual exam benchmark for LLMs (in progress)

- Building question–answer pairs from CBSE Class XII board papers and their official marking schemes in 6 Indian languages (Bengali, Hindi, Gujarati, Marathi, Tamil, Telugu) to evaluate LLMs on exam reasoning outside English.

**Do vision-language models read the edges?** Oct 2026 – Present

Study of relational understanding in VLMs (in progress)

- Verified 32 papers (ICML, ACL, ECAI and others) on diagram and graph understanding, collecting evidence that VLMs rely on node text and fail on edges, arrows and paths, to set up probing experiments on open VLMs.

**Survey of vision datasets for waste and defect imaging** Oct 2026 – Present

Survey paper in preparation (ACM format)

- Wrote a Semantic Scholar mining, matching and enrichment pipeline that screened 813 candidate papers down to 193, coded by waste domain, task, modality, datasets and CORE venue rank.

**AstraLink** Jan 2026 – Apr 2026

ZK-KYC for tokenised assets; Stellar testnet prototype github.com/DeepanIsCool/Solyrix-AstraLink

- Wrote the Soroban (Rust) RWA asset-token contract: KYC-gated transfers that query a teammate's identity soulbound token, account freeze, multisig governance, and yield deposit and claim.
- Built the Next.js front end with Freighter wallet integration, then a full refactor for v2.

## HACKATHONS

**COMSYS 2025 Hackathon**, Team Solyrix, Task A Jun 2025 – Jul 2025

- Team member on a face gender-classification entry that reached 94.08% accuracy and 0.90 macro F1 (422 validation images): ConvNeXt-Base fine-tuned in PyTorch Ignite with class-weighted loss.

**Smart India Hackathon 2024**, deepfake detection 2024

- Ran video analysis and model evaluation for a deepfake-detection pipeline (CNN spatial features with LSTM temporal modelling).

**Diversion 2K25 (MLH)**, BloodLink 2025

- Built the React front end of a blood-bank management system.

**Also took part in:** Accenture GenAI Hackathon; RiseIn Web3 hackathons on Aptos (Build With Aptos 2025, Solyrix Arena), Algorand (DecentralAI) and Stellar (AstraLink)

## SKILLS

**Programming:** TypeScript, JavaScript, PHP, Python, SQL, C, Java; contracts in Rust (Soroban)
**Front end:** Next.js, React, Flutter, Tailwind CSS, shadcn/ui, Vite, React Final Form, TradingView Lightweight Charts
**Back end and data:** Laravel 12 (PHP 8), Sanctum, PDF signing, Node.js and Express, FastAPI, REST API design, MySQL, PostgreSQL, MongoDB
**Machine learning and research:** PyTorch, explainability (Grad-CAM, Integrated Gradients, SHAP, LIME), constrained multi-agent RL (PPO, PPO-Lagrangian), MetaDrive, document layout analysis, RAG
**DevOps:** GitHub Actions CI/CD, Docker Compose, Vercel, Railway, PM2, SSH deploys
**Web3:** Stellar/Soroban, Freighter, Algorand, Aptos
**Tools:** Git, LaTeX, Google Colab, MATLAB

## MEMBERSHIPS, CERTIFICATIONS AND LANGUAGES

**Memberships:** IEEE (Student Member); INSTICC (2026)
**Certifications:** Information Theory, CUHK (Coursera); Cybersecurity Fundamentals (Coursera); Advanced System Security Topics (Coursera); Cisco Network Security, 8-course series (LinkedIn Learning); Foundations of Azure AI, Machine Learning with Python Foundations, Introduction to Prompt Engineering for Generative AI (LinkedIn Learning); Python for Data Science (Udemy); Microsoft Learn achievements (2025–2026)
**Languages:** Bengali (native), English (professional working proficiency), Hindi (working proficiency), German (elementary)