

ANGSHUMAN CHAKRAVERTY

ML Systems & Data Science Engineer — End-to-end pipelines · Benchmarking frameworks · Intelligent agents

angshumanchakraverty2@gmail.com • (+91) 9885289622 • [GitHub](#) • [LinkedIn](#) • [Portfolio](#)

SKILLS

- **Languages:** Python, C++, SQL
- **ML / DL:** PyTorch, TensorFlow, Scikit-Learn, HuggingFace Transformers
- **Generative AI & LLMs:** LangChain, LangGraph, RAG Pipelines, Ollama
- **Data & Viz:** Pandas, NumPy, Matplotlib
- **MLOps / DevOps:** Docker, Kubernetes, Git & GitHub, CI/CD (AWS CodePipeline), Linux/Bash
- **Cloud & Deployment:** AWS (EC2, S3, SageMaker, Elastic Beanstalk), FastAPI, Flask, Streamlit

PUBLICATIONS

VESTIGE: A Knowledge-Guided Masking Strategy for Corruption-Aware Fine-Tuning of Genomic Transformers, Validated on Ancient DNA Reconstruction

Chakraverty, A., Maheshwari, R.

[arXiv:2607.27712 \[cs.LG\]](#) | 2026

- Introduced a parameter-free masking strategy that aligns training with an empirically measured per-position corruption profile, replacing the standard assumption that reconstruction difficulty is position-agnostic.
- **Cut validation cross-entropy 13 %** (3.274 vs. 3.757) and led standard MLM at every damage width tested ($\Delta = +4.18$ to $+10.35$ pp, all $p < 10^{-8}$), with ESMFold reconstructions holding **TM-score** > 0.95 even under damage amplified up to $30\times$ beyond authentic rates.

Benchmarking LLMs for Verilog Design Flows

Chakraverty, A., Koshti, R., Sharma, B.P., Chamola, V.

[arXiv:2607.22759 \[cs.AR\]](#) | 2026

- Built a reproducible benchmarking platform evaluating open-source LLMs on Verilog RTL generation across 50 curated tasks spanning combinational, sequential, FSM, and mixed designs, validated via Verilator compilation and Icarus simulation.
- **Raised syntax validity from 0 % to 70.4 % average** and **simulation pass rate to 51.8 %** across 1,610 runs on three open-source LLMs (Llama-3-8B, StarCoder2-7B, TinyLlama-1.1B); prepared for submission to IEEE Design & Test.

EXPERIENCE

AI Backend Engineering Intern | Asset Telematics

[GitHub](#) → 2026

- **Extended a Gemini-powered OCR microservice** with multi-language output (English, Arabic, Hindi, French, German) via prompt engineering and a per-request `output_lang` parameter across 10 document types; diagnosed and fixed a Unicode data-loss bug where Latin-only regex silently stripped Arabic/Devanagari — replaced with Unicode-category-aware cleaning preserving combining marks.
- **Built an async batch-processing pipeline** (`/extract_batch + /batch_status` polling API) with bounded parallelism via `asyncio` semaphore and per-file partial-failure isolation; engineered concurrency around a blocking vendor SDK using a multi-key round-robin thread-pool, validating graceful degradation under live API rate limits.
- **Integrated OpenRouter (Qwen3-VL)** as an isolated document Q&A chatbot module with upload-once/ask-many session design; built a benchmarking harness measuring throughput, latency, and extraction accuracy — achieving **~86 % name-extraction accuracy** on passports with structured validation rejecting incomplete documents.

PROJECTS

Anvil: Adversarial ML Red-Teaming | Python · PyTorch · LangGraph · FAISS · UMAP · HDBSCAN · FastAPI

[GitHub](#) → 2026

- **Implemented 4 adversarial attacks from scratch** (FGSM, PGD, Adversarial Patch, Semantic) in PyTorch autograd with no external attack libraries, paired with an autonomous patching engine testing up to 3 fix strategies per cluster against a composite safety gate (score ≥ 0.70 , accuracy drop $\leq 3\%$).
- **Clustered adversarial failure modes** by extracting penultimate-layer activations and projecting through UMAP + HDBSCAN without a fixed cluster count, separating structurally distinct vulnerability classes that gradient-only analysis cannot distinguish.
- **Built a RAG-grounded LangGraph agent** over a FAISS index of 10 adversarial ML papers to generate citation-backed per-cluster vulnerability explanations via Gemini 2.5 Flash with coherence-gated retry logic rejecting low-quality outputs.
- **Deployed the full 8-phase pipeline as an async FastAPI service** on HuggingFace Spaces via Docker with a background job queue, polling endpoint, streamed ReportLab PDF output, and a React frontend on GitHub Pages.

Argus: Spot-Resilient ML Training Orchestrator | Kubernetes · kopf · AWS EKS · Terraform · Prometheus · Grafana · PyTorch

[GitHub](#) → 2026

- **Built and deployed a Kubernetes operator** (custom `SpotResilientJob` CRD + `kopf` reconcile loop) on AWS EKS that survived a real EC2 Spot interruption end-to-end — draining in **10s**, checkpointing a live PyTorch job to S3 on `SIGTERM`, and resuming on a healthy node in **~75s** with only one epoch of work lost.
- **Provisioned a 5-service AWS stack** (EKS, S3, SQS, Lambda, IAM) as Terraform IaC with IRSA/OIDC for zero-static-credential pod access; the end-to-end predict → checkpoint → cordon → reschedule migration chain completes in **~112ms**.
- **Instrumented the operator with Prometheus + Grafana** (real-time risk → checkpoint visualization) and enforced an all-region “\$0 teardown” cost regime after a \$66 overrun — keeping total spend inside the \$100 credit.

- **Built an end-to-end sim-to-real retrieval system** — a multi-instrument transformer encoder, a noise-conditioned flow-matching posterior (fm4ar backbone), and an optimal-transport calibration step — attacking the sim-to-real gap from both the domain-adaptation and inference sides.
- **Diagnosed the real-data collapse on WASP-39b** as a radius/baseline degeneracy rather than forward-model fidelity (ruled out via controlled ablations); adding the radius parameter and calibrating against an independent nested-sampling reference drove best-fit χ^2/dof **301** → **0.06** and posterior coverage **2/7** → **7/7**.
- **Generalized the unchanged method across 3 targets** spanning planet class, instrument, and data provenance — including a spectrum **self-reduced end-to-end from raw JWST MAST data** (exoTEDRF + batman), recovering its water feature at 5.8σ — with NS-agreement coverage of 7/7, 7/7, and 6/7.
- Ran controlled ablations (CycleGAN domain translation vs. structured randomization; higher-resolution opacity grids) and reported the honest negative: noise-conditioning improved synthetic-benchmark calibration but did **not** transfer to real data.

OPEN SOURCE — ML4SCI / DeepLense

- [PR #194 \(#191\)](#) — Fixed a breaking **torchmetrics API incompatibility** affecting the full evaluation suite: audited all notebook cells, migrated metric instantiation to the correct API, and refactored shared imports to prevent regression.
- [PR #195 \(#192\)](#) — Replaced hardcoded W&B credentials with a **dynamic –entity CLI argument**; added config validation and standardised run-naming — enabling reproducible multi-user training runs.
- [PR #190](#) — Authored **dataset setup docs** (download links, folder structure) for Models I–IV in root and subfolder READMEs, removing the main contributor onboarding blocker.

EDUCATION

B.Tech CSE (Data Science) — SVKM's NMIMS, Hyderabad

Jul 2023 – May 2027

CGPA: 7.1 / 10 | Sem 5 GPA: 3.14 / 4.0

ACHIEVEMENTS & CERTIFICATIONS

- **1st Place, NMIMS Hackathon NMTF** — Team of 6, 24-hr sprint
- CS50's Introduction to Artificial Intelligence with Python (2024)
- NPTEL Python for Data Science (2024) ■ NPTEL Introduction to Machine Learning (2024)

2025

[GitHub Repo →](#)