ASWANI KUMAR SAHOO

aswanisahoo227@gmail.com | linkedin.com/in/aswani-sahoo | github.com/AswaniSahoo | Rourkela, India

Education

National Institute of Technology, Rourkela **Jul 2023 – May 2027**
*Bachelor of Technology in Ceramic Engineering, CGPA **8.02** / 10* Rourkela, Odisha, India

Experience

Shipd by Datacurve **Sep 2026 – Present**
ML Benchmark and Agent-Evaluation Contributor (Freelance) Remote

- Solve **computer vision** (aerial imagery), **NLP** (text and dialogue repair) and structured-data reconstruction challenges with fine-tuned **CNNs and transformers** and **gradient-boosted models**, choosing versions on k-fold, shift-aware validation: **1st on the private leaderboard in 4 challenges**.
- Author benchmark tasks that **AI coding agents** attempt: dataset builds, **hidden-test splits** and **deterministic graders**, with a reference-solution ladder proving each task is learnable but not trivial.
- Add **leakage and saturation audits** plus **planted-defect tests** proving each verifier gate catches an injected defect.

Open Climate Fix (graph_weather, open-source weather forecasting library) **May 2026 – Jul 2026**
Open-Source Software Engineer (mentored by core maintainer) Remote

- Picked up a **4-year-old open issue** (adaptive meshing): read **14 regional weather datasets** on mismatched grids before any code, then shipped **19 merged PRs** (16 from that issue); named contributor on **v1.0.132**. **[Merged PRs]**
- Traced a defect that had silently collapsed the model to its no-change baseline, lifting measured skill from **4% to 20%** above baseline with **zero added parameters**, then found the same defect in a second component.
- Retracted my own invalid ablation in public (issue #238) and replaced it with a **5-seed** control: message passing adds no skill.

Projects

Climate-Risk Analyst Agent **Jun 2026 – Present**
Python, LangGraph, MCP, BM25 + dense retrieval, Gemini, FastAPI, Google Cloud Run **[Live] [GitHub] [MCP Registry]**

- Built a **five-node LangGraph agent** over **ERA5 extreme-value (GEV) statistics** and IPCC AR6 retrieval that refuses questions outside its data and degrades to a citation-less report on failure rather than fabricating.
- On a **SHA-256-frozen 105-question held-out set**: **R@3 87%** on the 70 answerable questions, **96% citation validity**, **zero false answers**, and all 35 refusal traps refused.
- Deployed on **Cloud Run** with per-request cost and latency telemetry at about **$0.0025 per question**, with **1,138 passing tests**; the retrieval layer is one of **2 MCP servers**, **listed on the official MCP Registry**.

Incident Evidence Compiler **Jul 2026 – Present**
Python 3.12 (mypy strict), PostgreSQL, FastAPI, Gemini, Prometheus, hexagonal architecture **[GitHub]**

- Built a root-cause service on the rule that **the model proposes and deterministic code decides**: hypotheses restricted to a signal allow-list, and a verifier returning SUPPORTED, REFUTED or UNKNOWN against a content-addressed evidence ledger.
- On **RCAEval**, run once on the **sealed held-out split** under a frozen configuration: **top-1 0.932** on 88 dev cases and **0.767** on 90 unseen cases, cited evidence must exist in the ledger by design; **363 tests** in CI, including integration tests on a live PostgreSQL.

Weather Transformer from Scratch **Feb 2026**
PyTorch, xarray, zarr, ERA5 / WeatherBench2, Google Cloud Storage, pytest **[GitHub]**

- Wrote a **4.8M-parameter Vision Transformer** with no library attention primitive (6 pre-norm blocks, 8-head attention) and its ERA5 pipeline: on held-out 2020 data, 10 m wind error is **22% (u10) and 35% (v10)** below persistence; temperature and pressure still trail anomaly persistence.

Open Source

35 merged pull requests across 6 open-source projects, every one through external maintainer review:

- **PyTorch / ExecuTorch (Meta)**: **9 merged PRs** (docs). | **Open Climate Fix**: **19 merged PRs**. **[PyTorch] [ExecuTorch]**
- **krkn-chaos (CNCF)**: **4 merged PRs**. | **MalariaGEN**: **2 merged PRs** (genomics). | **Brightband**: **1 merged PR** into ExtremeWeatherBench. **[View PRs]**

Technical Skills

- **LLMs and AI agents**: LangGraph, agent orchestration and tool calling, Model Context Protocol (MCP), structured outputs, retrieval-augmented generation (RAG), hybrid dense and BM25 retrieval, embeddings and vector search, LLM APIs (Gemini)
- **LLM evaluation (evals)**: frozen held-out and sealed-split protocols, refusal and citation checks, hidden-test graders, bootstrap confidence intervals
- **ML**: PyTorch, scikit-learn, pandas, NumPy, xarray, zarr, graph neural networks
- **MLOps and systems**: GCP (Cloud Run, Vertex AI), Docker, GitHub Actions CI/CD, Prometheus, PostgreSQL, FastAPI, Git
- **Languages and practice**: Python, SQL, C++, Bash; pytest, test-driven development, `mypy` strict, Claude Code, GitHub Copilot

Certifications and Recognition

- **PyTorch Docathon 2026, Honorable Mention (First-timer)**: **1 of 7** from **30+ active participants**. **May 2026**
- **Oracle Cloud Infrastructure 2025**: Generative AI Professional, Data Science Professional, AI Foundations Associate. **Oct 2025**