

Vaibhav Mahore

B.Tech, Mathematics and Computing, Indian Institute of Science (IISc), Bengaluru
mvaibhav@iisc.ac.in | +91-9284782375 | [LinkedIn](#) | [GitHub](#) | [LeetCode](#) | [Portfolio](#)

EDUCATION

Indian Institute of Science (IISc), Bengaluru

Aug 2023 - Present

B.Tech in Mathematics and Computing

Coursework: Data Structures and Algorithms, Linear Algebra, Probability, Machine Learning, Deep Learning, NLP

St. Ann's Senior Secondary School, CBSE

Class XII: 2023, Class X: 2021

EXPERIENCE

AI/ML Research Intern, Ericsson India Private Limited

Bengaluru, May 2026 - Jul 2026

- Benchmarked WiMamba (Mamba-based wireless model) vs a Transformer on AWS A10G GPUs: **10.7× lower latency** (30 ms vs 327 ms) and **42.6× less GPU memory** (113 MB vs 4.8 GB) at 4×4 patch size, with linear-time scaling.
- Evaluated transfer to a new geography (Florida to San Diego) across 5 tasks: **+16.6% composite score** over baseline and 7.4-7.7 dB lower error on channel estimation and interpolation.
- Fine-tuning with 75% of parameters frozen kept **98.5%** of full fine-tuning performance.

Machine Learning Engineer (AI Data Trainer), Alignerr and micro1

Remote, May 2025 - Aug 2025

- Alignerr: wrote reasoning traces and training data to improve LLM tool selection and problem solving.
- micro1: data annotation, model evaluation and validation, and data preprocessing (TensorFlow, scikit-learn).

PROJECTS

Efficient Sequence Modeling: From S4 to Mamba-3 (Python, PyTorch)

May 2026 | [GitHub](#)

- Implemented S4D, Mamba, Mamba-2 and Mamba-3 from scratch in PyTorch (Mamba-2 chunked path unit-tested against a sequential scan), with latency, memory and throughput benchmarks at 1K-32K lengths on a 6 GB RTX 4050.
- Reproduced the Mamba-3 paper's parity-tracking result: **98.3%** accuracy with complex-valued transitions vs at most 56% for real-valued state space models and Transformers; selective copying: 67% (selective) vs 6% (S4D).

AIRE: AI Reliability & Evaluation Engine (Python, pytest)

Aug 2026 | [GitHub](#)

- Built an evaluation framework for LLM/RAG systems: replayable runs, deterministic metrics (groundedness, citation coverage, retrieval recall and MRR), failure classification, and per-case regression flags between versions.
- Evaluated three RAG versions on a 26-case suite: it caught a deliberately introduced regression (groundedness **0.57 to 0.00**), and a live Gemini responder fixed abstention (0.0 to **1.0**) and grounding (**1.0**) at 15 s per case.

Autonomous Repository Engineer (Python, pytest, Git)

Sep 2026 | [GitHub](#)

- Built a coding agent around a deterministic plan-act-observe-verify loop: the LLM planner only proposes actions, and seven schema-validated tools run under path restrictions, allow/ask/deny permissions and command allowlists.
- Passed **4/4** SWE tasks with zero failed tool calls using a deterministic planner, and **2/2** with a live Gemini planner; a task counts as done only when tests pass, never on the model's claim.

SKILLS

Languages: Python, C++, SQL (basic)

ML: PyTorch, TensorFlow, scikit-learn, Transformers, state space models (Mamba)

LLM systems: RAG, LLM evaluation, tool-using agents

Tools: NumPy, Pandas, Hugging Face, pytest, Git/GitHub, Linux, Docker, AWS (EC2, S3)

CERTIFICATIONS

Oracle Certified: Generative AI Professional ([credential](#)) and Data Science Professional ([credential](#))

Aug 2025

LEADERSHIP

Sci-Tech Coordinator, Rhapsody Fest, IISc

2023 - 2026

- Organized three inter-college gaming tournaments and a coding contest for 200+ registered teams.
- Ran a 500+ member Discord community for registration and tech support.