



Purvi Verma

purviverma2026@gmail.com

+91-9399723537

AI Engineer | LLMs | Multimodal Systems | GenAI Platforms | LLMOps
Github | Portfolio

linkedin | Leetcode

EDUCATION

Degree/Certificate	Institute/Board	CGPA/Percentage	Year
M.Tech (AI)	IIT, Roorkee		2026 - 2028
B.Tech	IIT, Bhubaneswar	8.2	2020 - 2024
XII	CBSE Board	92.2%	2020
X	CBSE Board	91.6%	2018

PROFESSIONAL EXPERIENCE

- Publicis Sapient** July 2024 - July 2026
Senior Associate Data Science Bangalore
 - Architected and deployed end-to-end ML pipelines for AI safety guardrails on Bodhi integrating computer vision and LLMs; achieved production-grade reliability improvements with measurable reduction in harmful output rates.
 - Designed and implemented a multi-turn conversational LLM agent with tool calling, modular guardrails, and context-aware dynamic workflows — increasing task success rate and enabling A/B testing of model variants in production.
 - Integrated MCP tools for fine-grained policy enforcement and user-specific safety adaptation; led cross-functional discussions with product and engineering stakeholders to align on solution architecture.
 - Built a multimodal AI pipeline (CV + LLM) to extract structured insights from images, videos, and animations; evaluated across 500+ assets for 12 brands of a Pharmaceutical Company — achieved 90.3% average accuracy, 87% semantic consistency, and zero production failures.
 - Championed LLMOps best practices including model monitoring, versioning, and retraining triggers; presented AI platform strategy and results to senior stakeholders, articulating trade-offs between latency, cost, and accuracy.
 - Explored model fine-tuning and optimization using parameter-efficient techniques (LoRA/PEFT) to reduce resource requirements; experimented with vLLM for high-throughput LLM inferencing at scale.
- Blockatena** March 2023 - May 2023
Artificial Intelligence Intern Remote
 - Developed a hybrid GAN + LSTM architecture for video frame prediction, improving temporal consistency in generated sequences through sequence modeling.
 - Delivered end-to-end solution from data pipeline design to model evaluation, communicating progress and findings to the technical lead.

PROJECTS

- CodeMap** <https://github.com/purvi508/CodeMap>
 - Built an AI-powered system to analyze codebases and generate architecture diagrams, dependency graphs, and documentation using static analysis + LLMs — enabling faster onboarding for complex/legacy systems.
 - Designed scalable API layer for model inferencing; applied prompt engineering and semantic chunking for high-accuracy architectural summaries.
- AI-Powered Agricultural Assistant** Google Cloud GenAI Hackathon
 - Built a multimodal AI assistant for Indian farmers supporting voice, image, and text inputs, using LangGraph + Gemini (Vertex AI) for task-specific agentic workflows.
 - Implemented RAG pipelines for personalized government subsidy recommendations; integrated real-time crop advisory, weather insights, and disease detection features.
 - Deployed scalable backend on Kubernetes + Google Cloud + Firebase; designed for reliability and low-latency inferencing under variable load.
- Safeguard Toolkit — Open-Source Python Security Scanner** pip install safeguard_toolkit
 - Developed a modular static analysis tool for detecting security vulnerabilities in Python codebases (eval/exec abuse, permission misconfigurations). (eval/exec abuse, permission misconfigurations)
 - Packaged and distributed via PyPI for seamless CI/CD integration; demonstrates commitment to open-source engineering and software reliability.

SKILLS SUMMARY

- Languages:** Python, C++, Julia
- ML/AI:** LLMs, NLP, Computer Vision, Multimodal AI, Model Fine-Tuning (LoRA, PEFT)
- ML Platforms:** Vertex AI (Gemini), Azure AI, MLflow — experiment tracking, model registry, A/B testing
- LLM Stack:** LangChain, LangGraph, RAG, MCP
- Frameworks:** PyTorch, TensorFlow, Hugging Face, DeepSpeed (distributed training)
- MLOps / LLMOps:** Docker, Kubernetes, CI/CD pipelines, model monitoring, retraining workflows, LLMOps practices

OPEN SOURCE CONTRIBUTIONS

- google bespoke**, Improved system debuggability and failure resilience in production AI pipelines.
- video-keyframe-detector**, Contributed enhancements to keyframe detection logic in a Python-based video processing tool.
- CRAFT-pytorch**, Enhanced inference stability in CRAFT-pytorch OCR pipeline.

SCHOLASTIC ACHIEVEMENTS

- 🏆, Finalist Google Cloud Agentic AI Day Hackathon (Top 700 out of 9,100 teams)
- 1st place**, Hackathon held at GC, IIT Bhubaneswar (Real-time Age & Gender Detection System)