Varun Inamdar

(+91) 7517277551 | vninamdar03@gmail.com | linkedin.com/in/varun-inamdar | github.com/vnigoated | portfolio/varuninamdar

Education

Vishwakarma University — Pune, Maharashtra
B.Tech in Artificial Intelligence — CGPA: 8.7/10 — *Aug 2023 – May 2027*

Vishwakarma University — Pune, Maharashtra
Honors in Cybersecurity — CGPA: 9.0/10 — *Aug 2024 – May 2027*

Experience

AI Engineer Intern — Jun 2024 – Dec 2024
Payload Drone Project — Vishwakarma University — *Pune, Maharashtra*

- Designed a high-payload drone system with intelligent autonomous navigation for Southern Command, increasing flight stability by 22%.
- Built a GPS-independent visual SLAM pipeline integrating YOLOv5 object detection and MiDaS depth estimation, achieving 92% localization accuracy in indoor environments.
- Applied model quantization techniques to reduce inference latency by 40%, enhancing real-time decision-making on edge devices.

Security Analyst — Web Application VAPT — Apr 2025 – Sep 2025
Beeman & Nimka — *Pune, Maharashtra*

- Performed vulnerability assessments using BurpSuite and OWASP ZAP across 10+ web applications, identifying 15+ high-risk vulnerabilities.
- Discovered unencrypted form data, missing security headers, and SQL injection flaws, reducing critical exposure by 35% post-mitigation.
- Authored detailed VAPT reports aligned with OWASP Top 10 standards, including proof-of-concept exploits and remediation guidelines.
- Collaborated with developers to implement security hardening measures, improving application security scores by 28%.

Software Developer Intern — Oct 2025 – Nov 2025
Bootcoding Pvt. Ltd. — Remote

- Contributed to real-world product engineering in a structured corporate tech environment.
- Developed and deployed production-grade features using React.js, Node.js, FastAPI, and cloud infrastructure.
- Applied model quantization to reduce LLM inference latency by 38% and memory footprint by 60% on edge deployments.
- Gained hands-on experience in agile workflows, code reviews, and CI/CD pipelines, enhancing software delivery efficiency.
- Collaborated with cross-functional teams to design scalable solutions, improving system performance and user experience.

Technical Skills

Languages: Python, C, JavaScript, TypeScript, SQL
Frameworks: React, Next, Node, Express, Flask, FastAPI, TailwindCSS
AI/ML Stack: PyTorch, Scikit-learn, Pydantic, spaCy, Transformers, OpenCV, LangChain, LangGraph, LlamaIndex
Databases: PostgreSQL, MongoDB, Redis, FAISS, Pinecone
Specializations: Natural Language Processing, RAG, Agentic AI
DevOps & Tools: Git, Docker, AWS, GCP, Vercel, MLflow, BentoML, Langflow, Ollama
Cybersecurity: OWASP ZAP, BurpSuite, VAPT

Projects

Skoda AI/ML Based Building & Energy Management System | *FastAPI, PyTorch Transformer*

- Developing an intelligent BMS/EMS solution for Skoda manufacturing facilities integrating chiller, boiler, and HVAC optimization using predictive machine learning models.
- Implemented time-series forecasting models for energy consumption prediction, achieving 87% accuracy and reducing energy waste by 18%.
- Built predictive failure detection system using anomaly detection algorithms, identifying equipment faults 48 hours in advance with 91% precision.
- Designed centralized failure management dashboard with automated alert mechanisms, reducing maintenance downtime by 25%.

Digital and Intelligence Diet Plan RAG Agent | *Watsonx AI, Granite 8b-instruct, FAISS, LangFlow, Node.js, React*

- Developed a Retrieval-Augmented Generation system for personalized nutrition guidance using Watsonx Studio for model orchestration and deployment.
- Leveraged Granite models for both embedding generation and LLM-based response synthesis, ensuring domain-specific accuracy and efficiency.
- Integrated FAISS vector store for scalable semantic retrieval across 10K+ nutritional records, enabling context-aware and adaptive meal recommendations.

Career Connect | *FastAPI, Next.js, LangChain, Supabase, Docker, Vercel* January 2025

- Developed a cybersecurity learning platform with personalized learning paths, adaptive assessments, AI proctoring, and virtual labs.
- Created and embedded course materials into Supabase pgvector to provide explanations and curriculum guidance through a LangChain-driven RAG chatbot.
- Integrated LangChain with Google Gemini to deliver intelligent question answering and confidence-based MCQ scoring with automated integrity logs.
- Implemented Supabase for course storage and containerized the entire stack with Docker for consistent development and deployment.

CNN-based Diabetic Retinopathy Detection with Explainable AI | *PyTorch, Grad-CAM, Flask, Hugging Face*

- Developed a CNN model achieving 91% accuracy for diabetic retinopathy classification using retinal fundus images.
- Applied Grad-CAM for visual interpretability, improving clinician trust and model transparency.
- Deployed with Flask API and integrated LLM-generated natural language diagnostic reports.

Publications

AICCT 2025 (Accepted)
Explainable AI in Diabetic Retinopathy Diagnosis: CNN-Based Detection with Gradient-Weighted Class Activation Mapping

SPRINGER NATURE 2025 (Accepted)
Document Summarizer: A Machine Learning Approach to PDF Summarization

Achievements

1st Place — IBM Hackathon — Secured 1st place for developing a multi-agent AI solution using Watsonx and Cloud; presented project at IBM Summit, Delhi.

IBM Business Intelligence Specialization (SQL, ETL, Data Warehousing) — Coursera Certificate

Google Data Analytics Certificate — View Credential

3× Hackathon Winner — Recognized in innovation events hosted by Binghamton University and WILO for AI-driven automation solutions.

1st Place — Codeathon 2025 — Won against 20+ teams for a multimodal healthcare prediction platform using ML and LLMs.

Open Source Contributor — Contributed to 10+ GitHub repositories involving machine learning frameworks and educational AI tools.

Deployed AI Monitoring system at Shri Mahalaxmi Mandir, Pune during Navratri, counting 200,000+ people and boosting crowd management