Vinay Khosya

+91-9996303072 vinay.khosya.ug23@nsut.ac.in github.com/vinaykhosya linkedin.com/in/vinaykhosya

Professional Summary

Applied Machine Learning Engineer with **1.5+ years of hands-on industry experience** building and deploying production-grade ML systems across industrial automation, healthcare AI, and cybersecurity domains. Strong focus on model reliability, inference optimization, backend integration, and real-world deployment constraints.

Experience

- **AI Engineer – ThirdEye AI (JBM Group)** *Gurugram, India*
  Automotive AI & Smart Manufacturing Aug 2023 – Present
  – Developed applied ML solutions for industrial automation and smart manufacturing workflows.
  – Optimized inference pipelines for edge environments under memory and latency constraints.
  – Integrated ML models with backend services and embedded systems for production deployment.
- **AI Systems Engineer – ElectraWireless (Healthcare AI Startup)** *Remote*
  Applied AI Systems & Backend Engineering Jan 2024 – Present
  – Built end-to-end ML pipelines for iris-based health analysis and contactless vitals monitoring.
  – Designed backend services for real-time inference, data ingestion, and system reliability.
  – Worked closely with stakeholders to align AI systems with product and operational goals.
- **AI Engineer Intern – Atharvva AI** *Remote*
  Industrial AI & OCR Systems May 2024 – Aug 2024
  – Designed and deployed an OCR pipeline for industrial document and image processing using OpenCV and deep learning-based text recognition.
  – Built backend APIs for model inference, database integration, and workflow automation.
  – Developed frontend interface for real-time document scanning and end-to-end system integration.
- **Machine Learning Engineer – Gurugram Police Cyber Security Division** *Remote*
  Cybersecurity & Applied AI Apr 2024 – Jul 2024
  – Designed multi-model malware detection pipeline using PyTorch and anomaly-based learning.
  – Implemented FAISS-based similarity search improving detection accuracy by **15%**.

Selected Projects

- **GuardEye – Offline Android Malware Detection** GitHub
  PyTorch • FAISS • OCR
  – Multi-stage malware detection using OCR, permission analysis, and vector similarity.
  – Trained on **60GB+ APK dataset**, achieving **98% accuracy**.
- **Hospital Resource Optimization**
  Time Series • MaxEnt IRL • PPO
  – Forecasted patient influx using Conv1D-based time-series modeling.
  – Implemented MaxEnt IRL + PPO for hospital resource allocation optimization.
- **Contactless BP Monitoring (rPPG)**
  Computer Vision • Signal Processing
  – Developed rPPG-based blood pressure estimation from facial video streams.
  – Achieved **89% accuracy** under controlled evaluation.
- **Second Life – Local AI Memory Assistant** GitHub
  LLM • FAISS • SQLite
  – Built privacy-preserving semantic retrieval system with sub-second latency.
  – Managed **10k+ local conversations and files**.

Technical Skills

Core: Applied ML, Computer Vision, ML System Design, AI Deployment
Languages: Python, C++, SQL
ML Frameworks: PyTorch, TensorFlow, scikit-learn, Transformers
Systems & Tools: FAISS, Docker, Git, SQLite, FastAPI, gRPC
Domains: Industrial AI, Healthcare AI, Cybersecurity, OCR, NLP

Education

B.Tech in Electronics & Communication Engineering (AI & ML) *2023 – 2027*
Netaji Subhas University of Technology, Delhi