

Vinay Sharma

📍 U.P. ✉ vinayiet435@gmail.com ☎ +91 7379771194 in vinayiet 🌐 vinayiet

Experience

Technical Trainer (Machine Learning)

Oct 2024 – Present

W3Grads

- Trained over 270 students in core Python, OOP, SQL, file I/O, and foundational concepts in Generative AI including RAG and prompt engineering

Data Scientist Associate

Sep 2024 – Dec 2024

Blackcoffer

- Engineered serverless workflows using AWS Lambda and S3, achieving a 60% reduction in PDF data processing time
- Designed a LangChain-powered chatbot to handle client queries, lowering average resolution time by 40%
- Orchestrated Dockerized ETL tools and integrated them with AWS services for scalable deployment
- Implemented fine-tuned RAG models that improved document QA accuracy by 25%

Data Science & Engineering Intern

May 2024 – Sep 2024

Betaknot AI

- Oversaw Azure DevOps CI/CD pipelines for bulk and incremental ingestion, ensuring 98% pipeline success rate
- Enhanced data quality checks using SQL assertions, reducing error rates by 30%
- Created a recommendation system combining OpenAI models and scraped data, improving precision by 22%

Projects

RepoReviewAI – AI-Powered GitHub Code Auditor

Python, GitHub API,
LLMs, NLP, LangChain,
OpenAI

- Developed an AI-based tool that takes any public GitHub repository, analyzes its Python codebase, and summarizes the purpose and structure of the project
- Integrated LLMs to detect bugs, unsafe practices, deprecated code usage, and provide intelligent suggestions for fixes and improvements
- Automated extraction and analysis of files using the GitHub API, allowing fast auditing and understanding of open-source repositories

RAG Chatbot for Documents

Streamlit, HuggingFace,
Pinecone

- Developed a PDF-based RAG chatbot capable of answering context-rich queries
- Benchmarked 18+ models to optimize retrieval latency and generation coherence

Sensor Fault Detection System

Python, TensorFlow,
Scikit-learn

- Programmed an ML-based anomaly detector for IoT sensor networks with a 92% detection rate
- Embedded model in a Flask API for real-time monitoring

Automated DataSync Tool

Python, Regex, SQL

- Streamlined data ingestion from PDFs using regex and synced outputs to relational databases

Education

Institute of Engineering & Technology

B.E. in Computer Science & Engineering

- CGPA: 7.14/10

St. Michael's School

Intermediate (PCM)

- Percentage: 71%

Technical Skills

Programming Languages: Python, SQL, JavaScript

Generative AI: LangChain, LangGraph, OpenAI APIs, HuggingFace Transformers, RAG

ML/AI Frameworks: Scikit-learn, TensorFlow, PyTorch, CNNs, Transformers

Web Frameworks: Flask, FastAPI, Streamlit, React.js, Node.js

Databases: FAISS, MySQL, MongoDB

Cloud/DevOps: Azure, AWS, Docker, Azure DevOps, GitHub Actions

Tools: Git, Power BI, Jupyter

Competitions & Achievements

Kaggle Competitions

*Participant and
Contributor*

- Participated in structured data challenges and improved classification AUC scores by 12–15% across entries
- Applied Random Forest, Decision Trees, and ensemble modeling to diverse real-world datasets

Certifications

- Data Science with Generative AI — PW Skills, 2024
- Introduction to Python Programming — Harvard University, 2024