DHANANJAY G. KHARKAR

Nagpur, Maharashtra, 441501

📞 +91-9595167618 ✉ dkharkar00@gmail.com 🔗 Dhananjay Kharkar ⚫ dhananjaykharkar

SUMMARY

Applied AI Engineer proficient in orchestrating LLMs and ML models into deployable applications using Python, Docker, and LangChain. Specialized in building end-to-end AI pipelines that solve real-world data problems with strict validation and production-grade reliability.

EDUCATION

JD College of Engineering & Management **Nov 2022 – Pursuing**
BTech in Artificial Intelligence **CGPA: 8.84 / 10** *Nagpur, Maharashtra*

PROJECTS

SmartFlow:AI-Powered Data Quality Pipeline ↗ | Python, MSSQL, FastAPI, Streamlit, Isolation Forest, LLM
- Architected an enterprise-grade ETL pipeline to solve "Garbage In, Garbage Out" by integrating Large Language Models (LLM) for unstructured text parsing and an Isolation Forest model for anomaly detection.
- Implemented a robust multi-stage validation engine enforcing SQL referential integrity, canonical normalization, and Idempotency checks to ensure database consistency and prevent duplicate transactions.
- Developed a full-stack data entry and monitoring solution using FastAPI and Streamlit, featuring a real-time dashboard to visualize data health, rejection rates, and anomaly scores.

MetricMind: Predictive Analytics & Decision Support ↗ | Python, Docker, Scikit-Learn, Streamlit, SQLite
- Built a production-ready predictive analytics system using Test-Driven Development (TDD) to forecast sales volume and simulate profitability scenarios based on historical data patterns.
- Developed an interactive "What-If" decision support simulator using Streamlit and Random Forest Regression, allowing users to dynamically optimize pricing and marketing strategies.
- Containerized the entire application using Docker to ensure cross-platform reproducibility and implemented a complete ETL workflow from SQLite seeding to serialized model inference.

DocuMinr: Intelligent Document Processing ↗ | Python, LangChain, Google Gemini API, Streamlit, Docker
- Engineered an automated document-to-dataset extraction pipeline that converts unstructured narrative text (PDFs) into structured, queryable JSON/CSV datasets using LangChain and Google Gemini 1.5.
- Implemented a deterministic Validation Layer using Pydantic to enforce strict user-defined schemas, ensuring type safety and automatically flagging data anomalies before export.
- Dockerized the application to create a portable, reproducible microservice and developed a Streamlit frontend for dynamic schema definition and real-time extraction monitoring.

PROFESSIONAL EXPERIENCE

CSE Dept., Visvesvaraya National Institute of Technology ↗ **June 2025 – July 2025**
IoT & Data Analytics Intern *Nagpur, Maharashtra*

- Worked hands-on with basic sensors including temperature (DHT11), ultrasonic, TDS, GPS, soil moisture, and LDR modules using Arduino UNO and Mega 2560 for foundational data acquisition and control systems.
- Developed a real-time IoT data pipeline using ESP32, MQTT, and Node-RED to simulate and process temperature and humidity readings; stored structured data in SQL Server and visualized trends via Grafana and Node-RED dashboards.
- Implemented threshold-based alerting via Email/SMS and integrated multiple sensor types for real-time monitoring.

TECHNICAL SKILLS

Languages: Python, SQL
AI & Machine Learning: LangChain, Scikit-Learn, Google Gemini API, OpenAI API
Data Engineering & Validation: Pandas, NumPy, Pydantic, SQLAlchemy
Deployment & Tools: Docker, Streamlit, Git, GitHub, VS Code
Databases: Microsoft SQL Server, SQLite