

VIVEK KURMI

Data Scientist | AI Engineer

Email:- vivekdhanantar@gmail.com | [LinkedIn](#) | [Github](#) | Location:- Bhopal , M.P. Experience :- 2+ Years

Mobile: - +91 7828166109

PROFILE

AI Engineer with 2.4 years of experience designing, developing, and deploying AI-driven solutions and intelligent web-scraping pipelines. Specialized in LangChain, LangGraph, and Retrieval-Augmented Generation (RAG) for multi-agent orchestration and knowledge-centric applications. Experienced in scalable model deployment, inference optimization, and workflow automation. Strong background in data engineering, model fine-tuning, and production-grade AI integration.

EDUCATION

Indian Institute of Technology, Roorkee	June 2021– May 2023
Master of Technology	CGPA: 9.079
Maulana Azad National Institute of Technology, Bhopal	July 2016 – May 2020
Bachelor of Technology	CGPA: 7.030
Govt boys H S School Deori, Sagar, M.P.	July 2014– April 2015
Class XII	91.60 %

SKILLS SUMMARY

- Languages:** Python, SQL
- Frameworks:** Langchain, LangGraph, CrewAI, Pandas, Numpy, Scikit-Learn, TensorFlow, PyTorch, FastAPI
- Tools:** Power BI, Azure, AWS (Bedrock, Lambda), SQL Server, Docker, Git, Model Context Protocol (MCP)
- Platforms:** Visual Studio, Jupyter Notebook, PyCharm, Putty, VS Code, Colab
- Model Serving:** VLLM, Triton Inference Server, HuggingFace Transformers, Ollama
- Soft Skills:** Collaboration, Innovation, Decision-Making, Problem-Solving, Technical Communication

WORK EXPERIENCE

Data Scientist Tech Mahindra	July 2023 - Present
<ul style="list-style-type: none">Led development of multi-agent AI solutions for Cisco, automating corporate hierarchy mapping using LangChain, LangGraph, Serper.dev, Google Search, Crawl4AI, and pgvector-based RAG. Improved data accuracy by 90%, reduced processing time by 80%, and cut manual effort by 75%.Delivered synthetic data generation pipelines for Google (grants, employee profiles, succession mapping) using Azure OpenAI LLM and LangChain agents, ensuring scalability, accuracy, and client-driven iterations.In a client-focused project for Cisco, leveraged agentic AI (using Langchain) for automation to scrape and map complex company hierarchies. Custom operations and targeted modifications improved functionality, increased data accuracy by 20%, reduced processing time by 80%, and cut manual work by 55%—delivering a highly efficient, scalable solution tailored to client needsBuilt a robust Guardrails system using RoBERTa based model to enhance the security of large language models (LLMs) on our server infrastructure.Developed an AI companion with memory using LangMem, LangGraph and PostgreSQL.Engineered a high-performance inference server using the VLLM tool, integrating advanced models such as Phi-2 and Meditron-7B to optimize processing capabilities.Fine-tuned a Vision Transformer model for rail track damage detection, leveraging visual data to automate inspection processes and achieve 99% accuracy—significantly improving reliability, efficiency, and safety in infrastructure monitoring.	

- Contributed to **Microsoft's Purview R&D** project by implementing a unified data governance solution using Azure Purview and Data Factory for data lineage.
 - Collected, cleaned, and transformed **30+ macroeconomic** datasets using Databricks with Python, managing on-premises data via SQL Server.
 - Built interactive **Power BI** dashboards to support data-driven insights and decision-making.
-

PROJECTS

Inventory Forecasting | Dissertation, IIT Roorkee**January 2022- May 2023**

- Researched and implemented forecasting models (SARIMAX, LSTM, and other ML/DL approaches) for inventory management optimization.
- Collected and analyzed real-world inventory datasets from Thoucentric Consulting.
- Designed and executed data preprocessing pipelines using Pandas and NumPy
- Performed exploratory data analysis (EDA) with Seaborn and Matplotlib to extract actionable insights.
- Evaluated and fine-tuned time-series models to improve demand prediction accuracy and model robustness.

Breast Cancer Classification | IIT Roorkee**March 2022- April 2022**

Worked on Breast cancer data. Analyzed different parameters and applied Logistic regression, K-Nearest neighbor, SVM, Kernel SVM, Naive Bayes, Random Forest classification methods.

CERTIFICATES

- AWS Certified Cloud Practitioner Sept 2025
 - AWS Certified AI Practitioner May 2025
 - AWS Partner: Generative AI on AWS Essentials (Business) July 2024
 - Business and Operations Analytics, NITIE by David Simchi-Levi (MIT Professor) Oct 2022
-

ACHIEVEMENTS

- First Runner-Up for Eureka Awards – Innovative Solution of the Year at the Tech Mahindra Star Awards 2025
- Winner of the Bravo Award for successfully leading Cisco PoC
- Secured AIR-16072 in JEE Advanced 2016
- Received ₹25,000 Meritorious Student Prize from MP Chief minister.