

VINAY SHARMA

Ballia, U.P. | +91 7379771194 | vinayiet435@gmail.com | LinkedIn/vinayiet | GitHub/vinayiet

Aspiring Data Engineer — Skilled in Azure, AWS, ETL Pipelines, Data Science

EDUCATION

Institute of Engineering & Technology

B.E. in Computer Science & Engineering — CGPA: 7.14/10

St. Michael's School

Intermediate (PCM) — 71%

Jul 2024

Ballia, U.P.

Jul 2019

Ballia, U.P.

EXPERIENCE

GLA University

Technical Trainer (Python, SQL)

Oct 2024 – Present

Mathura, India

- Instructed 270+ students on Python programming, SQL database management, and fundamentals of data engineering.
- Delivered hands-on sessions on building ETL pipelines and cloud integration basics (Azure).

Blackcoffer

Data Scientist Associate

Sep 2024 – Dec 2024

New Delhi, India

- Built serverless ETL pipelines for structured data extraction using AWS Lambda and S3 triggers.
- Deployed containerized applications using Docker on AWS for efficient data processing.
- Designed ML models for classification tasks and optimized them using feature engineering and hyperparameter tuning.

Betaknot AI

Data Science & Engineering Intern

May 2024 – Sep 2024

New Delhi, India

- Managed and validated Azure DevOps pipelines for large-scale data ingestion into Azure SQL.
- Designed and implemented ETL workflows for bulk and incremental loads using PySpark and SQL.
- Improved data quality by developing automated SQL-based validation scripts, reducing error rates by 30%.

PROJECTS

Data Pipeline for Real-Time Weather Data

Python, Azure Functions, SQL, Power BI

- Built an end-to-end ETL pipeline to collect, transform, and load weather data into Azure SQL Database.
- Visualized real-time weather trends through interactive dashboards built in Power BI.

Flight Price Prediction System

Python, Machine Learning, Flask

- Developed ML models to predict flight prices based on features like airline, source, and destination.
- Deployed the model using Flask and created a user-friendly web interface.

Automation DataSync Tool

Python, Regex, SQL

- Automated data extraction from complex PDFs using regex and synced structured data into SQL databases.

Product Recommendation System

Python, Web Scraping, OpenAI APIs

- Built a recommendation engine using web-scraped e-commerce data, fine-tuned using OpenAI embeddings.

Sensor Fault Detection System

Python, Scikit-learn, Flask

- Developed machine learning models to detect anomalies in IoT sensor data streams and deployed as an API.

TECHNICAL SKILLS

- **Data Engineering:** ETL Pipelines, Azure Data Factory, Azure SQL, AWS Lambda, DevOps Pipelines
- **Cloud Platforms:** Azure, AWS
- **Languages:** Python, SQL, JavaScript
- **Databases:** Azure SQL, MySQL, MongoDB, Snowflake
- **Data Science:** Scikit-learn, Pandas, NumPy, Deep Learning
- **Frameworks:** Flask, FastAPI
- **Tools:** Docker, Git, Power BI, Looker Studio, Regex

CERTIFICATIONS

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