

DISHA GUPTA

Lucknow | +91 6393738728 | dishaa28118@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

Aspiring Data Engineer with a strong foundation in Python, SQL, and data pipelines, passionate about designing efficient ETL workflows, data lakes, and AI/ML-ready datasets. Experienced in automating data cleaning, transformation, and visualization through modern tools like FastAPI, Streamlit, and Power BI.

TECHNICAL SKILLS

Languages: Python, SQL, Shell (Basic), HTML/CSS

Frameworks & Libraries: FastAPI, Pandas, NumPy, PySpark (Learning), Scikit-Learn

Databases & Tools: PostgreSQL, MySQL, Power BI, Excel, Docker, Git, VS Code

Concepts: ETL Pipelines, Data Lakes, Data Cleaning, API Development, Automation

Soft Skills: Analytical Thinking, Communication, Problem-Solving, Team Collaboration

EXPERIENCE

Virtual Internship Trainee – Tata Data Visualisation (Forage)

Duration: July 2024 – April 2025

- Designed data pipelines for visual reporting using Power BI aligned with business KPIs.
- Performed data cleaning and transformation for clear, accurate dashboards.
- Applied visualization techniques to convert raw datasets into actionable insights supporting business decisions.

PROJECTS

1. Data Cleaning as a Service (Daas_API) — FastAPI | Docker | Python

[GitHub](#)

- Engineered a data preprocessing microservice using **FastAPI** for automated data ingestion, deduplication, missing value imputation, and schema validation.
- Designed modular **ETL** components for data extraction and transformation, ensuring high-quality structured datasets.
- Containerized the solution using **Docker**, improving scalability and reusability.

2. Business Health Dashboard (AI-Driven) — Streamlit | Plotly | Python

[GitHub](#)

- Developed a functional dashboard with integrated **ETL pipeline** for data ingestion, and visualization.
- Implemented automated anomaly detection and PDF report generation (via **FPDF**) to streamline business reporting.
- Enabled real-time insights for decision-makers through an AI-enhanced data interface.

3. Real-Time Object Detection Web App (YOLOv8, Streamlit, OpenCV)

[GitHub](#)

- Developed a **real-time object detection web app** using **YOLOv8**, **Streamlit**, and **OpenCV**, supporting images, videos, and live webcam feeds.
- Integrated multiple **YOLOv8 model variants** for dynamic speed–accuracy comparison.
- Enabled **custom model uploads** and **interactive result downloads** through an intuitive Streamlit interface.

EDUCATION

Bachelor of Technology in Computer Science & Engineering

Shri Ramswaroop Memorial University, Lucknow — Aug 2022–Present

CGPA: 7.54 / 10 (Current)

CERTIFICATIONS

- Power BI – Microsoft (Jul 2024)
- Data Science and Business Analytics Workshop – IIT Mandi (Mar 2025)
- Deloitte Data Analytics Job Simulation – Forage (Apr 2025)
- Learning: ETL Pipeline Development with Apache Airflow (In Progress)