

NITESH RANJAN SINGH

Glasgow, United Kingdom | 3090808S@student.gla.ac.uk | +44 7900 945064 | [LinkedIn](#) | [GitHub](#) | [Leetcode](#) | [Portfolio](#)

Professional Summary

Experienced **Data Engineer** (3+ years) with a strong foundation in **Azure Data Engineering**, ETL pipelines, and big data technologies. Currently pursuing an **MSc in Data Science** at the **University of Glasgow** with a focus on applied industry practices. Proven ability to design and optimize scalable data solutions using Azure Data Factory, Databricks, and PySpark. Certified in **Microsoft Azure (DP-203)** and **Databricks**. Now seeking a Summer Internship/ Graduate Role in Data Engineer/Data Analyst/Data Science role in the UK to apply cloud-native data engineering/science expertise in real-world business contexts.

Key Skills & Tools

- Languages: Python, SQL, PySpark
- Cloud: Azure (Data Factory, ADLS, Databricks, Fabric), Snowflake
- Data and BI Tools: Apache Spark, Kafka, Airflow, Power BI
- ETL & Warehousing: ADF, Data Lakes, Delta Lake, Azure Synapse
- **Certifications:**
- - Microsoft Certified: Azure Data Engineer Associate (DP-203)
- - Databricks Certified Data Engineer Associate
- - IBM Data Science Professional Certificate

Education

- **MSc Data Science – University of Glasgow, UK** **Sep 2025 - Sep 2026**
 - Key Modules: Big Data: Systems, Programming & Management, Introduction to Data Science & Systems, Machine Learning & AI for Data Scientists, Deep Learning, Information Visualization, Research & Professional Skills.
 - Scholarship: **£10,000** Excellence Award.
- **B. Tech in Computer Science Engineering – UPES, India** **June 2018 – May 2022**
 - Final Year Project: **Digi-Attendance** – IOT-based attendance system using facial recognition: Developed an automated attendance system using **Python (OpenCV, face_recognition, PyQt5)** for real-time facial recognition through a camera interface. Designed a GUI to clock in/out employees, record timestamps in CSV files, and calculate total working hours. Implemented live video feed processing and face encoding to ensure accurate, contactless attendance tracking.

Professional Experience

- **Data Engineer – IQVIA, Gurgaon, India** **Dec 2022 – Aug 2025**
 - Built and maintained end-to-end ETL pipelines for clinical and commercial healthcare data using ADF, ADB, and ADLS
 - Led data migration of 1B+ records from on-prem to Databricks using PySpark and ADF, reducing run-time by 35%
 - Automated data quality checks and metadata management workflows using Python and ADF
 - Collaborated with data scientists to productionize ML pipelines with model scoring in Databricks
 - Tools: Azure ADF, Databricks, Snowflake, Power BI, Git, Jira
 - Awards: IQVIA Impact Award FY 2023-24(Ovation) & 2024-25(Spotlight)
- **Analyst – Axtria, Noida, India** **Jan 2022 – Dec 2022**
 - Created SQL-based dashboards and ETL pipelines to support pharma sales analytics
 - Improved performance of reporting processes by restructuring queries and indexes
 - Provided ad-hoc data solutions using Power BI and Python scripts for visualization

Projects

- **AutoPipelineAI** – Developed an LLM-driven autonomous ETL and Data-Ops framework that automates data ingestion, transformation, and insight generation using local open-source models. Integrated Streamlit dashboard, Ollama-based Llama3 querying, and smart data visualization for seamless end-to-end analytics. [AutoPipelineAI](#)
- **Mathematical Model for Linear Data** – Built a regression-based root cause analysis tool on US Crime Data using only NumPy for mathematical modeling, without external ML libraries. Achieved MSE = 0.0452 and $R^2 = 0.9835$, demonstrating high model accuracy through fully manual implementation of linear regression.

Additional Information

- Work Authorization: Post-study work visa (Graduate Route), eligible to work full-time in the UK after graduation
- Languages: English (Fluent), Hindi (Native)