

ROHIT GUPTA

Pune, India | +91-9419128023 | rohit.ajaygupta@gmail.com | [linkedin.com/in/rohit-gupta-001/](https://www.linkedin.com/in/rohit-gupta-001/) | github.com/rohitgupta29 |
| <https://leetcode.com/u/user1839g/> | <https://medium.com/@RohitAjaygupta> |
Microsoft Certified Azure Data Engineer

PROFILE SUMMARY

Results-driven Data Engineer with expertise in ETL processes, data modeling, and database management. Proficient in Data Structures, Python, SQL and Apache Spark, ensuring efficient data flow and system optimization. Collaborates with cross-functional teams to deliver scalable solutions aligned with business needs. Adept at staying current with emerging technologies for continuous process improvement.

EXPERIENCE

Cognologix Technologies, Pune, Associate Data Engineer | Sep 2021 - Present

- Designed and developed Data pipeline Solution including Data Modelling, ETL, Aggregations for visualization, Security and logging using Azure Synapse, ADLS Gen2, Entra ID and KQL for mobility domain.
- Designed and implemented scalable data pipeline solution using AWS services such as S3, EMR and RDS for an ed-tech domain.
- Developed and maintained data processing and transformation scripts using Python and SQL as a boiler plate utility using PySpark, postgresSQL, mongoDB and Delta Lake for fasten processes during client engagement for projects.
- Designed and Developed a secure Data Pipeline ecosystem for Mobility Domain using Azure Synapse Analytics.
- Assisted in design and implementation of Data pipeline Orchestration using Airflow for smooth orchestration of batch data.
- Developed a critical connector API using flask for handling nested json data in a data warehouse and operational data store. The major challenge solved was to handle dependency management, null values in nested jsons and using multithreading for faster processing.
- Developed a plug and play tool for data validation of large files using Great Expectations and Flask. The tool generates a report providing insights through table and graph for any file with data discrepancies. It saved us 70% time devoted to Data Validation before trusting data for Ingestion and use for ETL processes.
- Developed a monthly employee insight report generation tool for managers and HR using machine learning and data analytics using scikit-learn, pandas, Matplotlib and Seaborn. Deployed the tool on AWS EC2. It helped voice manager's concerns in work from home scenario.
- Developed an NLP based web application using social media to analyze linguistic markers and generate individual based insights using Flask, Pandas, Matplotlib and Plotly Dashboard. Idea was to help the company HR understand if an employee is facing emotional issues during work from home scenario in covid times.
- Have Written various blogs on behalf of the DataAI department on an internal website for employees sharing experience with different technologies used in different projects. Have also provided many blogs individually.

Prepared for Civil Services (Economy, Indian Constitution, Psychology) June 2014-Jan 2020

Syamantak- School without Walls (mentoring and guidance) **Volunteer** Jan 2014- April 2014

- Mentored special children through skill based learning for better cognitive development and self-reliance.

NTPC Unchahar (Team work) **Engineering trainee** Nov 2012- Aug 2013

- Received Letter of appreciation from AGM Operations for performance in a plant overhauling project.

TECHNICAL SKILLS

Data engineering:	System Design, data warehousing, data modeling
Machine Learning:	Data Structures, Algorithms, Sci-kit Learn, Spacy, textblob,NLTK, LLMs
Languages:	Python (Flask,NumPy, Pandas), Pyspark, SQL, Java (spring boot)
DevOps Services:	Docker, Kubernetes, PostgreSQL, Linux, Shell Scripting
AWS Services:	S3, EC2, IAM, EMR, EKS
Azure Services:	Synapse Analytics, Azure Databricks, ADLS , Streaming Services, Entra Id
Tools & Technologies:	Linux, Airflow, kafka, Plotly Dash, Microsoft Excel, draw.io

CERTIFICATIONS

- Microsoft Certified: Azure Data Engineer Associate | Credential ID: 4F7E535CA55FF2DA
- Advanced SQL ([HackerRank](#)) | Credential ID: 4865cba21b27
- Certified AI Professional (NIELIT, Calicut) | Credential ID 7729
- Preparing for Databricks Certified Data Engineer Professional

EDUCATION

SHRI MATA VAISHNO DEVI UNIVERSITY, J&K

Bachelor of Technology, Mechanical Engineering June 2012