

TANMAY GUPTA

+91 892-004-1951 | tanmaygupta2198@gmail.com | [LinkedIn](#) | [LeetCode](#)

EXPERIENCE (4 YEARS)

Data Engineer

Dec 2021 – June 2024

ACCENTURE

Gurugram, Haryana

- Developed a **real-time data processing pipeline** for IVR call flow data on Amazon **Kinesis Data Streams**, utilised **Glue** with **PySpark** transformations based on Mean Opinion Score (MOS). This solution reduced over **10,000 quality-related incidents per month**.
- Designed and implemented a **high-performance batch processing audio encoding pipeline** using AWS **Lambda** and **Lambda Layers**, triggered by **S3 events**. Integrated the **FFmpeg** library and applied the **mu-law algorithm**, reducing processing time by 1.2 seconds per prompt, significantly improving call handling throughput.
- Conducted comprehensive analysis of **historical call data stored in Amazon Redshift** to evaluate IVR **automated testing** tools, such as Bespoken and Cyara from AWS Marketplace. Delivered insights through **Power BI**, aiding clients in decision-making with detailed reporting.
- Worked on **refining** existing **ETL pipelines** by **reorganizing database schema** and implementing **SQL and PySpark best practices**, including caching, partitioning, and broadcast joins, resulting in enhanced efficiency and accelerated data transformation processes.
- Managed AWS infrastructure deployments using **Azure DevOps** and **CloudFormation**, enhancing resource provisioning efficiency and streamlining **CI/CD** pipelines for smoother deployment processes.

Technology Associate

Oct 2020 – Dec 2021

AVIZVA

Gurugram, Haryana

- Led **in-house product development**, spearheaded the creation of **NLP-based conversational AI products**, demonstrating strong skills in functional programming and advanced problem-solving.
- Integrated **RASA engine server**, implemented as an independent component within a **Django** application, establishing it as the **NLP core to enhance conversational capabilities**.
- Developed and deployed an **automated Scala Play Framework project setup system**, improving setup efficiency by several hours per project and **reducing setup-related errors** through advanced automation and functional programming techniques.

SKILLS

- ETL & DATA ORCHESTRATION:** Python, Lambda, Glue, Step Functions, PySpark, Pandas, Airflow, Kafka
- DATA WAREHOUSING:** Redshift (Star and Snowflake Schema)
- CLOUD & DATABASES:** AWS (S3, DynamoDB), SQL, MySQL, PostgreSQL, MongoDB
- CI/CD & VERSION CONTROL:** Git, Azure DevOps, Github, Gitlab, CodeCommit, CloudFormation
- DATA VISUALIZATION:** Power BI, QuickSight, Tableau

PROFESSIONAL CERTIFICATION

Amazon Web Services Training and Certification

AWS

AWS Certified Cloud Practitioner (CLF-C01)

Dec 2022 – Dec 2025

EDUCATION

Jamia Millia Islamia University

New Delhi

Bachelor of Technology, Computer Engineering (1st Division)

2016 – 2020

Represented the University in the prestigious **ICPC Asia Region Contest, 2018**