

Tejas Agrawal

9172390187 | tejasagrawal36@gmail.com | linkedin.com/tejas | github.com/tejas

SUMMARY

Data Engineer with hands-on experience in building scalable data pipelines using PySpark, Kafka, and Airflow. Skilled in Python, SQL, AWS, and Databricks. Built ETL pipelines processing millions of records daily. Portfolio: tejasagrawal.com

TECHNICAL SKILLS

Programming Languages: DSA, DSA-C++, DSA-Python, Python, SQL, PySpark, Scala, Shell Scripting, C++, JavaScript

Data Engineering: Azure Databricks, Apache Spark, Delta Lake, Unity Catalog, PySpark, Apache Airflow, ETL Pipelines, Data Modeling (Star/Snowflake), Data Warehousing

Cloud Technologies: Azure (ADF, ADB, VNET, Virtual Machines, EntraID, Data Lake), AWS (S3, EC2, VPC, IAM, Lambda, RDS)

Databases: PostgreSQL, MySQL, MS SQL Server, MongoDB, Redshift

Tools & Platforms: Databricks, GitHub, Docker, Linux, GitHub Actions

Organizational Tools: Confluence (Documentation) , JIRA and Service Now (as a Ticketing Tools), Follows ITIL Guidelines.

Other: CI/CD, Logging & Monitoring, Documentation, Agile Methodology

PROJECT

Disha Academy

Apr 2025 – May 2025

Data Engineering Project

Databricks, Apache Spark, PySpark, SCALA, SQL, Data Pipeline

- Consolidated data from 110+ educational institutes into a unified Azure Databricks platform for analytics and reporting.
- Designed scalable ETL pipelines in Databricks using Apache Spark to clean, transform, and validate large datasets.
- Built a Medallion Architecture (Bronze, Silver, Gold) on ADLS Gen2 for structured and governed data processing.
- Governed data with Unity Catalog, implementing fine-grained access controls, lineage tracking, and schema management.
- Optimized Spark jobs using partitioning, Z-order clustering, and caching, improving performance by 40
- Orchestrated complex workflows with Databricks Jobs API and monitored pipelines using SQL dashboards and Azure Monitor.
- Ensured data quality with Delta Live Tables expectations and automated anomaly detection across ETL layers.
- Implemented rolling sum aggregations over time-based windows using PySpark Window functions to enable trend analysis and cumulative reporting on large datasets.
- Built PySpark Window functions for ranking, rolling sums, and lag/lead analyses on partitioned datasets.

Tejas Agrawal - iCode

Jan 2025 – Feb

MERN

NodeJS, ReactJS, SQL, Docker, RESTFUL-API

- Launched a revenue-generating web app with 2K+ users, built end-to-end systems, and showcased projects at tejasagrawal.com.
- Developed a professional portfolio at tejasagrawal.com to showcase achievements and projects.

TejasAI

March 2025

MERN stack Project

ReactJS, NodeJS, google ai studio, RealtimeDB, Dotenv, RESTFUL-API

- Designed a clean, modern UI for showcasing skills, successful SaaS projects, and achievements at tejai.com.
- Developed a professional Generating AI tejai.com to showcase achievements and projects.

EyeNex

Jun 2024

Fluid Eye Detection System

Flask, Deeplearning, Tailwind-CSS, JinjaTemplating

- Smart eye detection using SVM.
- Highlights Achieved 89.6% accuracy in ML model performance.

EXPERIENCE

Design Lyric - Internship

DEC 2024 – MAY 2025

Data Engineer (Remote)

Pune, India

Built and optimized scalable ETL pipelines using PySpark, SQL, and Databricks Notebooks to process large volumes of log and analytics data, integrating multiple structured and unstructured data sources into a centralized Azure Data Lake Storage Gen2. Created Azure Databricks Jobs and Delta Live Tables to automate and monitor end-to-end data workflows with email alerting, logging, and error handling. Collaborated with the backend team to expose clean datasets through REST APIs, **FAST-API**: for product analytics dashboards, enhancing real-time data access. Leveraged Unity Catalog for centralized data governance, ensuring secure, controlled access to critical datasets, and optimized performance with Delta Lake for ACID transactions, schema enforcement, and fast, scalable querying on large data sets.

EDUCATION

Government Polytechnic, Nagpur

Aug 2017 – May 2021

Polytechnic in Information Technology

RTMNU University

Aug 2021 – 2025

B.Tech in Computer Science Engineering

PERSONAL DETAILS

Language:

English, Hindi, Marathi

Address

*At Post **Akola** Maharashtra INDIA pin: 444103*