

OMKAR SHENDE

+91-9323115264

omkarshendeofficial@gmail.com

omkarshende1

TylerDurden786

TECHNICAL SKILLS

Languages: Python, SQL, PowerShell, Shell Scripting

Databases: SQL Server, PostgreSQL, MySQL, SQLite

Frameworks: Spark, Pandas, Pytest, Unittest, Scikit, Git, Docker, Selenium

Technologies: GitHub, Azure Databricks, AWS Lambda, Glue, S3, ECR, EC2, Postman

EXPERIENCE

Circulant Solutions

July 2022 – Present

Associate Data Engineer

- Improved **90%** of SQL query generation by integrating **OpenAI API** with the database using **Python**.
- Built a bioinformatics NGS workflow, ensuring validated ingestion of big data into **S3** and its smooth processing through **AWS EFS** for report generation and output archival, cutting **Databricks** cost by **15%**.
- Developed a streamlined pipeline that integrates shell triggers for SFTP, **Databricks** flow monitoring, **DynamoDB** metadata validation and view creation in **RedShift** reducing error handling time by **80%**.
- Automated a daily **Power BI** dashboard refresh flow, reducing manual effort by **98%** by using **PySpark**.
- Reduced claims data ingestion time by **75%** through **PySpark** on **Databricks** from **Azure Blob Storage**.
- Automated a backend project of folder permissions by **PowerShell** reducing manual time by **98%**.
- Slashed data load time by **60%** with an **ML** pipeline using **Docker**, **S3**, **Lambda** and **ECR**. Boosted patient selection accuracy by **90%** with an **ML** model's feature selection, model training and prediction.
- Automated a basic **ETL**, reducing **75%** of manual data loading ensuring efficient clinical data ingestion.
- Proposed a POC to automate provisioning CRUD tasks for fast creation, updation and deletion of the users.

Cognizant

April 2022 – June 2022

Programmer Analyst Trainee

- Completed an intensive training in Python, SQL and Data Warehousing enhancing data analysis skills.

Avvashya CCI Logistics Ltd

September 2021 – February 2022

Solutions Design Graduate Engineer

- Optimized cost data solutions by tracking productivity and managing procurement data.

PROJECTS

Real-Time Road Journey Pipeline

Tech Stack: Kafka, Zookeeper, PySpark, Docker, AWS Redshift, Glue

- Designed and implemented a real-time data streaming pipeline to monitor road journeys, contributing to enhanced road safety and accident prevention.

LinkedIn Skills Scrapper

Tech Stack: Selenium, Python

- Devised a Python script to iterate over the LinkedIn pages and search for the most relevant skills of a specified designation.

ACHIEVEMENTS

- Received **Spot Award** for the development project that featured rigorous unit testing.
- Achieved Client Appreciation for ETL automation by reducing the manual verification of data inconsistencies.

EDUCATION

Indian Institute of Technology Madras

January 2024 – Present

Diploma in Data Science



Mumbai University, Dwarkadas.J.Sanghvi College of Engineering

2017 – 2021

Bachelors of Engineering in Mechanical Engineering

CGPA - 9.22/10

CERTIFICATIONS

- Lakehouse Fundamentals - Databricks 
- Fundamentals of Generative AI - Databricks 
- Machine Learning with Apache Spark - Coursera 