

OMKAR SHENDE

omkarshendeofficial@gmail.com | +91 - 9323115264 | [LINKEDIN](#)

SKILLS

Languages: Python, SQL, PySpark, Bash and PowerShell

Technologies & Tools: Spark, PostgreSQL, SQL Server, Pandas, Databricks, AWS, Lambda, S3, ECR, API Gateway, Storage Gateway, Docker, Redshift, DynamoDB, EFS, Jira, Git, GitHub, ChatGPT, Bard, GeminiAI pro

EXPERIENCE

Associate Data Engineer

Jan 2023 - Present

Circulant Solutions

- Accelerated **90%** of **SQL** query generation by integrating **ChatGPT** with the database using **Python** and **OpenAI API**.
- Automated a daily **Power BI** dashboard refresh flow, reducing manual effort by **98%** by using **PySpark** on **Databricks**.
- Slashed data load time by **60%** with a Machine Learning pipeline using **Docker**, **S3**, **Lambda**, **ECR** and **API Gateway**.
- Boosted patient selection **accuracy** by **90%** with an ML model's feature selection, model training and prediction.
- Developed a **Serverless SFTP** workflow, achieving a **20% reduction** in **expenses** using **AWS S3**.
- Worked on an E2E **cloud** project using **AWS** services like **S3** and **EFS** that helped save **EC2** instance **cost** by **15%**.
- Streamlined error handling **time** by **25%** while enhancing client data security through a **Databricks ETL** stream by integrating **AWS DynamoDB**, **Redshift**, and **Microsoft Power Automate**.
- Reduced claims data ingestion **time** by **75%** through **Databricks** from **Azure Blob Storage**.
- Developed integration pipeline of **AWS Storage Gateway** for storage in replacement to **EFS** ensuring file access saving **20%** of the cost.
- Conceived a backend project to automate folder permissions using **PowerShell** reducing the **manual time** by **98%**.
- Led a comparative POC evaluating **Gemini AI pro** against **GPT-4** achieving **10% time** savings, **100% cost reduction** and developing a tailored prompt for **98%** accuracy on compound subqueries.

Intern

July 2022 - Dec 2022

Circulant Solutions

- Automated basic **ETL**, reducing **75%** of manual data loading ensuring efficient clinical data ingestion.
- Optimized data flow using Databricks and **AWS Lambda** by automating edge cases, saving **90%** of **manpower**.
- Constructed user-defined unit tests improving reliability of the existing pipeline and incurring **0.75% profit**.
- Proposed a **POC** to automate folder permissions, showcasing **Python**, **PowerShell** and **Task Scheduler** integration to **reduce manpower** by **98%**.

Solutions Design Graduate Engineer

Aug 2021 - Feb 2022

Avvashya CCI Logistics Ltd

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

ACHIEVEMENTS

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

EDUCATION

Dwarkadas Jivanlal Sanghvi College of Engineering, Mumbai University, India

- Bachelor of Engineering | Mechanical Engineering | CGPA - **9.22**

2017 - 2021

Maharashtra State Board, India

- HSC - **83.69%**
- SSC - **93.6%**

June 2017

June 2015

CERTIFICATIONS

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

[Certification Link](#)

[Certification Link](#)

[Certification Link](#)