Gouri Bhosale

📍 Pune   ✉ gouribhosale1996@gmail.com   📞 +91-8999844316   🔗 Website-GouriPortfolio   in gouribhosale

⌗ gouri-github

Profile Summary

I am a **Senior Data Engineer** with expertise in building **automated data pipelines** and workflows using Python, PySpark, and Snowflake. I specialize in creating scalable solutions for data integration, processing, and analytics in cloud environments like **Databricks** and **AWS**. In my recent projects, I developed automated scripts, scheduled and managed them to run seamlessly, ensuring efficient data delivery and reliability. With over 2.5 years of experience, I focus on optimizing **ETL processes**, maintaining data quality, and supporting business needs through robust engineering solutions.

Technologies/Skills

- **Cloud Computing:** AWS S3, Redshift, RDS, IAM, EC2, VPC
- **Data Engineering:** PySpark,Databricks, ETL, Data Pipelines
- **Automation and DevOps:** Terraform,CI/CD,GitHub Redwood RMJ
- **Databases:** SQL,PostgreSQL,SQLite
- **Web Technologies:** HTML5,CSS3,JavaScript
- **Development Tools:** Jupyter Notebook,databricks,snowflake

Education

Master of Computer Application [MCA]   2018-2021
- Chhatrapati Shahu Institute of Business Education and Research,Kolhapur.

Bachelor of Computer Application [BCA]   2015-2018
- V.P. Institute of Management Studies and Research,Sangli

Experience and projects

Capgemini, Devops engg   Capgemini,Pune
                                                                                                     June 2022 current
Project 1: Client: Sage Therapeutics

Title: Commercial and medical datahub/Market Access datahub

- Facilitated data integration and loading into Snowflake via a 3-layer architecture (Source to Core, Core to Conformed, Conformed to Analytics Ready).
- Maintained and improved existing automated data pipelines created with Python and PySpark in Databricks.
- Implemented automated scripts to ensure efficient data flow and integration from various sources (CSV,txt,table) to Snowflake.
- Designed and implemented data models (ERD, DFD) for scalable data solutions.
- Performing deployments on Test and Production environments, ensuring smooth operations across teams.
- Worked closely with cross-functional teams to design automation for seamless **data integration**.

Project 2:

Title: Data Integration and Query Automation with AWS Redshift and PySpark.

- Use Redshift data warehouse as the source and S3 as the target.
- Upload CSV files into an S3 bucket and load data from the S3 bucket into the Red-

shift data warehouse using S3 as the data source.

- Establish a connection between PySpark and Redshift using the data available in the Redshift warehouse and store the data into a PySpark DataFrame.
- Perform PySpark DataFrame queries to generate the required reports.

Project 3:

Title: Development of web application .

- Developed a web application to manage employee information, including details, experience, qualifications, and aspirations.
- Created front-end pages using HTML5, CSS3, and JavaScript and connected them with a SQLite backend.
- Gathered project requirements, aligned with business needs, and created detailed documentation.
- Designed data models (ERD, DFD) and developed forms for login, registration, and profile management with built-in validation.
- Conducted manual testing to ensure the application was bug-free and provided an excellent user experience.

Hobbies

- Playing badminton
- Carrom
- Design- website design,video creation,logo creation,UIUX

Address

Address: Menlo Homes Hinjewadi phase 2 Plot no 912,B wing

Pune