GOURI BHOSALE

DATA ENGINEER

Mumbai, Maharashtra | 8999844316 | gouribhosale1996@gmail.com
LinkedIn | Website | GitHub

SUMMARY

Results-oriented Data Engineer with 3+ years of experience in building scalable data solutions using PySpark, AWS Services, and Databricks. Proven ability to deliver high-quality software while managing end-to-end data workflows, including ingestion, transformation and deployment. Seeking to apply my skills in Big Data and Cloud Technologies to drive impactful and efficient data-driven solutions.

TECHNICAL SKILLS

- PySpark, Python, SQL
- Azure Databricks, AWS
- ETL/ELT Development, CICD
- Data Pipeline Optimization
- GitHub, Snowflake

- Incident Management, Root Cause Analysis
- Redwood RMJ (Orchestration tool)
- Build comprehensive data lineage in Databricks

PROFESSIONAL EXPERIENCE

Data Engineer, Capgemini   **June 2022 - Present**

- Designed and optimized scalable PySpark data pipelines, increasing processing performance by 30% for high-volume data workflows and critical business use cases.
- Integrated multiple data sources, including SQL databases, AWS S3, and Snowflake Data Warehouse, supporting dynamic data loading for end-to-end data processing.
- Expertise in AWS cloud infrastructure including VPC, Load Balancing, EC2, IAM, and S3,ensuring scalable and secure solutions.
- Collaborated with cross-functional teams to optimize data workflows, driving improvements in data quality, and increasing operational efficiency.
- Established and implemented GitHub version control best practices, enhancing collaboration and code traceability; guided teams on efficient branching strategies.
- Built and managed centralized data repositories in Databricks for processing and accessing market access and medical data for analytics.
- Implemented end-to-end CI/CD pipeline for seamless data integration from various sources to target layers, streamlining deployment and release cycles.
- Utilized data analysis techniques to identify data trends, patterns, and correlations, providing actionable business insights and improving decision-making.
- Created and reviewed technical design documentation/SOP Documents to ensure accuracy and consistency in reporting solutions and data models.
- Proficient in Pyspark, to build custom scripts and automate data workflows.
- Created and maintained automated ETL pipelines to ensure efficient, reliable data flow and integration across multiple source systems into Snowflake.
- Led weekly Test and Production deployments, ensuring smooth operations across teams

EDUCATION

CSIBER, Kolhapur   **Aug 2018 - 2021**

Master of Computer Applications [MCA]

ADDITIONAL INFORMATION

- **Languages Known :** English, Hindi, Marathi
- **Certifications:** AWS Cloud Practitioner, AWS Database Specialty, Azure Data Fundamentals
- **Awards/Activities**: Achieved first prize in Entry Level Capgemini Training program, recognized as a "Rising Star" in 2025 for outstanding performance and contributions for (I&D).