

SOHAM PATRA

Data Engineer - 3 YoE | +91 9546722131 | sohampatra1234@gmail.com

TECHNICAL SKILLS

- **Programming Languages:** Python, PySpark, PL/SQL, Shell Scripting
- **Databases:** MySQL, OracleDB, MongoDB
- **Data Warehouses:** Hive, Snowflake, AWS Redshift
- **Message Queue:** Confluent Kafka
- **Distributed Computation Frameworks:** Apache Hadoop, Apache Spark, Apache Kafka
- **Dashboarding Tools:** Power Bi
- **Workflow Management:** Airflow
- **AWS Services:** S3, Lambda, Cloudwatch, Redshift, Athena, Appflow, Glue, SNS, SQS
- **Build Tools & Containers:** Docker
- **Control Systems and Documentation:** Git, Jira, Confluence,

WORK EXPERIENCE

Data Engineer – Cognizant

Sept 2023 – Present

- Developed end-to-end **data pipelines** for large amount of financial transactions with optimized data ingestion which improved performance by reducing the runtime by 20% and scheduled the ETL process to achieve robust data delivery.
- Designed and implemented a comprehensive **data model** in Oracle database for a new management company, leading to a 10% improvement in data retrieval time.
- Communicated with key stakeholders and Data Analyst to understand requirements and modify existing **data transformation** process based on the insights, resulted in reducing time for data analysis.
- Collaborated with the testing team for data integrity testing to evaluate data based on quality, accuracy and consistency which helped to get improve the analysis based on the data.

Jr Software Engineer-Cognizant

Aug 2021 – Aug 2023

- Developed automation scripts using Ansible to automate Unix jobs for data refresh which helped saving hours of effort to run scripts manually with 80% reduction in efforts.
- Conceptualized and implemented custom ETL process utilizing Oracle SQL and PL/SQL for processing financial transactions resulting in clean and usable data for financial analysis.
- Contributed in A/B testing for ETL pipeline to optimize it for faster query execution which improved the performance by 15% and also reducing cost

EDUCATION

B Tech -Electronics Engg

KIIT UNIVERSITY, BBSR

Jul 2017 – May 2021, CGPA 8.0/10

PERSONAL PROJECT

• Chest Cancer Classification:

The objective of this project is to detect chest cancer

The model can detect if the person is suffering from chest cancer or not by seeing the x ray images for chest

Tech stack for model building : tensorflow, CNN, mlFLOW, DVC, FLASK API

EDUCATION & CERTIFICATIONS

- **Complete Data Science Boot camp 2024** On Udemy By 365 datascience
- **Generative Ai Fundamentals** By Cognizant
- **Azure AZ: 900** By Microsoft