

# HARSHALI PARKHI

## Data Engineer

### My Contact

✉ parkhiharshali24@gmail.com

☎ 7823836793

📍 Pune

### TECHNOLOGY SKILLS

- **Programming Language:**
  - Python
- **Big Data Ecosystem:**
  - Spark, Pyspark, HADOOP, HDFS, HIVE
- **IDE:**
  - PyCharm, Jupiter notebook
- **Cloud Technologies:**
  - AWS (S3, GLUE, EC2 & EMR, Cluster, RedShift, Athena, Sagemaker), Snowflake
- **Database:**
  - SQL, MySQL, Oracle
- **Other:**
  - GIThub, Agile, Jira

### Education Background

- **Bachelor of Computer Application (BCA)**  
**G.H.Raisoni Collage of Technology, Nagpur**  
**Completed in 2020**

### Languages

- English
- Hindi
- Marati
- Gujarati

### About Me

Experienced IT professional with a strong background in development, PySpark, and AWS seeking a challenging role in a dynamic organization. Skilled in data extraction, building optimized big data pipelines, and working with various AWS services. Proficient in handling structured and unstructured data from multiple sources. Capable of creating data models and maintaining applications to meet project requirements.

### Professional Experience

**DGCT INFOSOFT PVT LTD**

*JUNE 2020 - AUG 2023*

**Payroll company Name - Aaxeleron Technology**

**Client - Cognizant**

*AUG 2023*

- 3+ years of diverse IT experience in development, PySpark, and AWS.
  - Expertise in data extraction and building efficient big data pipelines
  - Proficient in AWS services like S3, EC2, Glue, EMR, Redshift, Athena, and SageMaker.
  - Strong experience in Hive, MySQL, Oracle, and Snowflake databases

### Projects

#### 1. Data migrate & data analytics

**(Domain : Banking)**

- Identified and analyzed new data sources.
- Developed custom data pipelines using AWS Sagemaker.
- Automated data processing pipelines with AWS services.
- Collaborated with teammates to import data from source to S3 using AWS Sagemaker.

#### 2. Data Analytics

**(Domain : Finance)**

- Designed and implemented data pipelines for ETL of financial data.
- Collaborated with data scientists and analysts for data availability.
- Implemented data governance and security measures for sensitive financial information.
- Optimized data processing and storage systems.