



# RISHABH CHAUDHA

Data Engineer

 chaudha.rishabh@gmail.com

 8074 282 973

 Hyderabad, Telangana

 linkedin.com/in/rishabh-chaudha-48957150

## SKILL

**Big Data Ecosystem:** Spark, Hadoop, Spark Streaming, Kafka, MapReduce, HDFS, Hive, Sqoop

**Cloud:** AWS

**Language:** Python, Scala, JAVA

**IDE :** Eclipse, IntelliJ

**Tools:** SyncSort DMX Express

**Schedulers :** Autosys, Oozie

**Databases :** Teradata , SQL

## EDUCATION

Bachelor of Engineering / Computer Science  
RGPV, Bhopal 2012 - 2016

## ACHIEVEMENTS

Cognizant Certified  
Data Scientist 2020

## PROFESSIONAL SUMMARY

Data Engineer with 4+ years of experience in analysis, design, development, customization, and Implementation of Big Data Applications. Have Worked on a number of strategically important technology projects in various business domains. Currently helping in building Streaming pipelines for building AIOPS Datalake.

## EXPERIENCE

### Data Engineer

*Cognizant Technology Solutions / October 2019 – Present*

- Developed Streaming Pipelines to ingest the data into datalake from various monitoring sources.
- Identify relevant data sources to mine for client business needs and collect large structured/unstructured datasets and variables.
- Involved in Migration of legacy systems to distributed system environment, reducing processing time by 50%.
- Helped in movement of RDBMS systems to cloud infrastructures.
- Converted Oracle procedures into spark sql.

### Hadoop Developer

*Tata Consultancy Services / August 2016 – September 2019*

- Involved in creating an enterprise level consolidated environment of deposit data that presents a robust history of deposit behaviour.
- Development of ETL application as per client requirement using Spark, MapReduce, Hive, DMX and Unix script.
- SME for creating the CDC & Snapshot process for deposit data.
- Handled a team of 6 members and successfully delivered a project.
- Knowledge Transfer to new joiners at Hyderabad and Chennai locations for relevant business application logic.

## Projects

### AIOPS – Data Lake

Ingestion of heterogeneous data from many different sources about all components of the IT environment — networks, applications, infrastructure, cloud instances, storage and more to hive and Building Anomaly models to predict the outages.

### MLF - Merrill Lynch Foreign

Creating an enterprise level consolidated environment of deposit data related with Foreign Accounts. Coding the CDC, Snapshot process, generation of surrogate key to uniquely identify each foreign deposit account and processing the data to Downstreams and Teradata for further analysis.

### Ranger

Profiling the consumption and invoice details data received from source systems with smart meter and IDR usage and aggregating it on brands, plans & contracts. Comparison of the data generated with the actual invoices generated by the vendors.