



Ganesh Bollapalli

SENIOR DATA ENGINEER

✉ bollapalliganesh@gmail.com

📍 Gachibowli, Hyderabad, India

☎ +91 9502539635

Profile

Senior BigData Developer with over 4 years of experience in Cloudera Hadoop technologies. Experienced in working with Spark, Scala, Python, Hive, Sqoop, Kafka, Airflow, Unix shell scripting, Teradata, Informatica, and Autosys. Able to work independently and as a part of a team.

Academic Qualification

Bachelor's in Electronis and Communication, RVR & JC College of Engineering, Acharya Nagarjuna University, Guntur, Andhra Pradesh

June 2016 — September 2020

Graduated with CGPA: 9.0

Employment History

Senior Data Engineer, LTIMindtree, Hyderabad

December 2023 — Present

Projects :

HSBC Project : Wholesale Data Refinery, Creating Core Data Assets for Credit and Lending Domain

Major Responsibilities :

- Created and Enhanced core Data assets based on business requirements
- Performing Tuning and Optimization techniques on existing DataFactory code
- Monitor and Manage scheduled airflow dags
- Ensure consistency and standardization in MDAs and GDAs for optimal performance

Big Data Developer, TATA CONSULTANCY SERVICES, Hyderabad

November 2020 — December 2023

Projects :

Bank Of America : AML Customer Profile Application, Decommission Teradata and used Hive as the Data-warehouse

Major Responsibilities :

- Converted Bteq shell scripts to Hadoop environment based on business requirement
- Developed spark utilities which loads data into hive from Teradata Tables or Flat Files
- Developed Informatica workflows in workflow Designer
- Worked on Performance tuning for Hive queries and spark jobs
- Created Autosys Jils for the Developed Jobs and scheduled and monitored them on SLA basis

Technical Skills

Apache Hadoop

Apache Spark

Scala

Python

Hive

Sqoop

AirFlow

Apache Kafka

Unix Shell Scripting

Teradata

Informatica

Intellij Idea

Autosys

Git Version Control

Azure Cloud

Links

[LinkedIn](#)

[GitHUB](#)

[Portfolio](#)

Soft Skills

Adaptability

Communication

Time Management

Team-Work

Leadership

Languages

English

Telugu

Hindi

Personal Projects

Spark Scala

Tool :- Clodera Data Platform

- Analysis of All Time IPL Data Using Spark Scala
- <https://github.com/Ganesh-Bollapalli/IPLData-Analysis-Using-Spark-Scala>

PySpark

Tool :- Databricks Community Edition

- Analysis of Flight System Data using pySpark
- <https://github.com/Ganesh-Bollapalli/Flight-System-Analysis-using-pySpark>

Hadoop RealTime Data Pipeline

Tool :- Amazon EC2 Instance

- RealTime Data Pipeline using Kafka, Spark Streaming and Cassandra
- <https://github.com/Ganesh-Bollapalli/RealTime-DataPipeline-Using-Kafka-and-SparkStreaming>

SQL

Tools :- MySQL WorkBench 8.0, Jupyter Notebook

- Analysis of UK Accident and World Population Data using sql
- <https://github.com/Ganesh-Bollapalli/SQL-Projects>

Python

Tool :- Spyder

- Files Handling in a folder using Python Regex
- <https://github.com/Ganesh-Bollapalli/File-Handling-Using-Python-Regex>

Professional Achievements

- Received Star of the Month Award For Crucial Contribution in the Project.
- Received On the Spot award for delivering the Project within the deadline with support from the team.
- Received Performer Of the Month Award For Contributing to Customer Business Value By involving in Various Proof of Concepts.

Certifications

Microsoft Azure Fundamentals (AZ-900)

Microsoft Artificial Intelligence Fundamentals (AI-100)

Apache Spark with Scala and Python, Udemy.

Apache Kafka Fundamentals with Hands-on, Udemy.