

Krishna Barkule

BIG DATA Engineer

Accomplish 2.6 years of experience in Big Data ecosystem for designing and executing solutions to complex business problems. Having the capability to explore, learn and understand newer business domains and technology. Problem-solving capability peered with good communication skills.

Experience

Big Data Engineer

[Apr 2020 – Present]

Codevizer Technologies Pvt. Ltd

1) Project Name – Financial Intelligence Big Data Pipeline

- The Purpose of The Project is to ingest the customers' full load data that is present on RDBMS storage to a Data Hub modeled as per an LDM. It's Cloud to Hive ingestion Project in Real Time. Data Migration Utility to migrate existing system into Data hub which was built on top of Hadoop.

Description & Responsibilities:

- Import Data into the Hbase from various Relational Databases (Oracle, SQL) using Spark
- Write Hive DDL to Create Hive Table for Optimize Query Performance.
- Ingest Flat Files like Delimited, Fixed Length, etc. into Hive Warehouse.
- Raise JIRA for Infrastructure, Platform issues.
- Implement Data Validation, Quality Checks, and Profiling.
- Involved in importing data from various RDBMS into HDFS using Pyspark, including Incremental Load to populate Hive External Table and vice-versa.
- Involved in importing data from various RDBMS into Hive Tables, including Queries using Pyspark.
- Designed both Managed and External Hive Tables and Defined static and dynamic partitions as per requirement for optimized performance on production datasets.

Primary Skills:

- Hadoop, Yarn, Map Reduce, Hive, Sqoop, and Hbase
- MySQL
- Pyspark, Python
- AWS, EMR, EC2, Role, RDS, S3.

Secondary Skills:

- Aws Glue, Aws Ogie
- Power Bi

Educational Details:

BACHELOR OF COMMERCE

Shhiv Chhatrapti College ,N-3 , cidco ,Aurangabad .

- Graduated with 53.23 % .

HSC COMPUTER TECH.

Vasantrao Naik college, Aurangabad

- Passed with 65.15% .

SSC MARATHI

M.I.T High School ,N-3 ,Cidco Aurangabad .

- Passed with 56.20% .

Languages:

Marathi, Hindi, English

- Worked with various File Formats like Text File, SequenceFile, ORC Files, Avro Files, and various Compression Formats like Snappy, and bzip2.
- Written Hive Queries for Data Analysis to meet the business requirements.
- Create Shell-script/Spark jobs for extracting data from RDBMS and file sources and load them into the data warehouse.
- Create Pyspark/Spark jobs for data transformation and aggregation.
- Produce unit tests for Spark transformations and helper methods.

The environment used – PostgreSQL 12.5-R1, 2. EMR(5.33) , Spark(2.4.7) ,Python(3.1.2) ,Hadoop(2.10.1) ,Hive(2.3.7) ,Hbase(1.4) ,Phoenix(phoenix-4.14.3) ,S3 ,Linux ,Workbench(Build 127) .

Automotive Mfrs. Pvt. Ltd

[Aug 2016 – Jan 2020]

Worked as **Back office Executive** in the automobile sector for 2.5 years as the Service Marketing And Customer Relationship Manager. The design offers According to the customer relation base. Analyze the data and Distribute data in Team track teamwork after accomplishing making a report and creating an insight report. Analyze the business flow.

Skills Summary

- Excellent understanding of **Hadoop architecture** and different components of the **Hadoop ecosystem**.
- Processing of **CSV and JSON-type** data structure.
- Developed **Pyspark script** for data analysis to meet the business requirement.
- Importing and exporting data from different databases like **MySQL into HDFS/ S3 and Hive using Pyspark**.
- Developed hive queries for **partitioning and bucketing** for the hive tables for better performance
- Designed both **Managed and External tables**
- Experienced in building **Spark applications, RDD, Dataframe, and Datasets also submitting Spark job**
- Good communication, collaboration, and team interactive skills with

Address and Contact Details

New Hanuman Nagar,
Aurangabad

Phone – 9822327692

E-Mail –

krishnabarkule@gmail.com

Declaration – All information in this resume is true and correct to the best of my knowledge and beliefs.