

Amol Bhagavanrao Kakde

Add: Plot No.03,Mahalaksmi Colony,
N2 **Aurangabad** 431003.

Mob No: +91 7507547356
Email : amolkakde31@gmail.com

CAREER OBJECTIVE

To work in an organization where I can contribute and enhance my knowledge and growth. To enhance my working capabilities, professional skills, business efficiencies and to serve my organization in best possible way with sheer determination and commitment to realize the organizational objectives.

CAREER SUMMARY

- Overall 3 + years of experience in IT Industry.
- Over 2.5 years of relevant experience in Big Data technology.
- Hands on experience with Big Data core components and Eco System including Data Ingestion and Data Processing (Spark)
- Capable of processing large sets of structured, semi-structured and unstructured data.
- Hands on experience in data analysis library Python Pandas.
- Experience working on an Agile methodology.
- Knowledge on Spark and used Hive to analyze sensor data and cluster users based on their behavior in the events.
- Good understanding and knowledge of Hadoop architecture and Hands on experience with Hadoop components such as Name Node, Data Node and Map Reduce concepts, Spark Execution Concepts and HDFS Framework.
- Hadoop development tools on several tools like Putty, WinSCP etc.

TECHNOLOGY PROFICIENCY

- **Platforms** : Windows, Linux.
- **Programming language** : Python.
- **Dev IDE**: PyCharm, SQL Developer, Jupiter Notebook.
- **Big Data Ecosystem** : Hadoop, HDFS ,Spark, YARN, HBase
- **Distribution Package** : Cloudera
- **Cloud Technologies** : AWS (S3, EC2 & EMR Cluster, Redshift, Athena)
- **Relational database** : MySQL.
- **Another tools and technology** : Jira , GitHub.

ORGANIZATIONAL SCAN

- **Organization:** Zielotech InfoSolutions India.
- **Designation:** Data Engineer
- **Location:** Pune
- **Duration:** From June - 2019 to Till Date

WORK EXPERIENCE

Project #1: Credit Card Data Analysis

Duration: August 2021 to Till Date

Role: Data Engineer

The Purpose of the Project Credit Card industry, in the recent past, has seen a huge surge in volume, stiff competition, various product offerings and various incidences of fraud. Banks and Financial Institutions dealing in credit card require information on the credit card on account of the unique nature of credit card exposures.

Roles & Responsibilities:

- Import Data into the Hbase from Relational Database (Postgre) for Historical data and HDFS for delta data using Spark.
- Write Phoenix DDL to create Hbase Table.
- Write Hive DDL to Create Hive Table for Optimize Query Performance.
- CSV Files like Delimited, Fixed Length, etc. into Hive Warehouse.
- Raise JIRA for Infrastructure, Platform issues.
- Implement Data Validation, Quality Checks, Profiling.
- Involved in import data from Hbase into Hive Managed Tables using Spark which includes Incremental Load and some transformations.
- Involved in import data from Hive Managed Table into Hive External Tables which includes Queries using Spark.
- Designed both Managed and External Hive Tables and Defined static and dynamic partitions as per requirement for optimized performance on production datasets.
- Worked with various File Formats like Text File, ORC Files, Avro Files, Parquet Files and Compression Formats like Snappy.
- Written Hive Queries for Data Analysis to meet the business requirements.
- Create PySpark jobs for importing data into Hbase.
- Create PySpark jobs for data transformation and aggregation.
- Produce unit tests for Spark transformations and helper methods

Project #2: Retail store Data Pipeline

Duration: June-2020 to July-2021

Role: Data Engineer

The purpose of the project is to do the ingestion of the super market data that is present in Mainframe system to Data Warehouse. It's Mainframe to HDFS ingestion project in Realtime. Data migration utility to migrate existing system into MySQL. Later the extraction would be done in PowerBI.

Roles and Responsibilities:

- Retail stores, serving the customers worldwide with a range of products like groceries, home decor, furniture, electronics.
- These stores generate the sales data at their POS (Point of sale) & get pushed into the big data pipeline for cleaning, processing and Business analysis.
- Basically, we have created a End to End data pipeline for Retail stores data using HDFS, Apache Spark, MySQL And PowerBi.
- We have to filter data and perform multiple operations in order to get meaningful insights that will help the client in decision making and analysis.
- Basically, it's A Maven project and we have written logic in Spark Scala using intelliJ IDEA.
- Have responsibility to Write the data extraction logic for the application.
- Have responsibility to write efficient code as per required and requested.

EDUCATIONAL QUALIFICATION

- Done B.E from Dr. BAMU University Aurangabad in June, 2018.
- Done Polytechnic from MSBTE Mumbai, in June 2015.
- Done SSC from Maharashtra State Board, Aurangabad in June 2011.

PERSONAL DETAILS

Date of Birth : 17 June 1995

Languages known : English, Hindi and Marathi

I hereby declare that the above information given is true and correct to the best of my knowledge and belief.

Date:

Place:

(Amol Bhagavanrao Kakde)