

# DEEPALI VISHNU JADHAV

---

Aurangabad, Maharashtra

+91 9284664427

[deepalijdc22@gmail.com](mailto:deepalijdc22@gmail.com)

## CAREER OBJECTIVE

To obtain a challenging position in an organization that requires highly motivated and creative people. I wish to build a promising career with a growth-oriented organization which promotes a learning atmosphere and continuous skill development.

## PROFESSIONAL SUMMARY

- Over 3+ years of IT experience, with around 2+ years into Big Data Technology.
- Have Experience in building optimized big data pipeline using HDFS, Hive, HBase, PySpark on cloud premises.
- Have a hands-on experience on Amazon Web Services mainly S3, EC2, EMR, RDS.
- Have experience with Data Flow, Data Pipeline and Workflow Management Tool.
- Experience in processing large amounts of structured and unstructured data.
- Perform tasks such as writing scripts and SQL queries.
- Develop Batch processing and process Structured and Unstructured Data.
- Good experience on Spark Architecture including Spark-Core, Spark-SQL.
- Experience in using Hadoop distribution like Cloudera.
- Experience in transferring data from RDBMS to HDFS and Hive tables using Hive and Spark API's.
- Experience in creating tables, partitioning, bucketing, loading in Hive.
- Gather and process raw data at scale and analyze processed data.
- Analytical mind and ability to learn-relearn concepts.
- Ability to work effectively with associates at all levels within the organization.

## TECHNOLOGY PROFICIENCY

- Hadoop/ Bigdata: HDFS, MapReduce, Hive, HBASE, Zookeeper
- Cluster Configuration: AWS
- Distributions: Cloudera
- Programming Languages: C, Python, JavaScript, C#, .net
- Databases: SQL, HBase, Oracle
- IDE: Jupyter notebook, PyCharm, Microsoft Visual Studio 2008/2012
- Scripting: HTML, CSS, JavaScript, XML
- Tools: Putty
- Operating System: Windows, UNIX/LINUX.

## TECH SKILLS:

**PRIMARY SKILL :** Spark, Pyspark, Hadoop, Oracle, Python, MS-SQL, Apache Hadoop, HDFS MR, Apache Hive, Sqoop, Big Data, AWS S3, EC2, EMR, RDS, HBase, Zookeeper, PysparkSql, Git, Github, JIRA

**SECONDARY SKILL:** Apache Kafka, Apache NIFI, Sqoop, Power BI, Unix/Shell Scripting, Linux, AWS Glue, Lambda, Redshift, Databricks – basics.

## EXPERIENCE

**Company Name:** Advent Informatics, Pune

**Designation:** Data Engineer.

**Duration:** July 2019 – Present

**1. Project Name:** Data Analytics for Retail Domain

**Client:** Europe Location

**Duration:** October 2021 – Present

**Description:** Taken the data from clients Database inject into the s3 using Kafka ingestion tool. Same for the Delta data. All the required data then Move into our storage unit. Process the data, add some business logic on top of it, clean it and then send it back to s3 for storage. The reporting team pulls data from s3. Creates a visualization, generate report at the end required data move into client s3 location.

### **Roles and Responsibilities:**

- Involved in Collecting Business Requirements from Business Users, Translate into Technical Design.
- Work closely with the business and analytics team in gathering the system requirements.
- Define all possible Test Cases along with Test Data.
- Import batch data and delta into s3, using PySpark.
- We have to filter data and perform multiple operations in order to get meaningful insights that will.
- Implement Data Validation, Quality Checks.
- Reconcile the imported or transformed data as per business requirement.

**Environments:** Databricks, Spark SQL, PySpark, Aws (s3,RDS), GitHub, AWS.

**2.Project Name :** Enterprise Data warehouse Analysis

**Duration:** February 2020 - August 2021

**Description:** The Purpose of The Project is to do the ingestion of the Telecom data that is present on cloud storage as well as Clients Database to Data pipeline which was modelled as per the client. It is Cloud to HBase ingestion Project. Data Migration Utility to migrate existing system into Data hub which was built on top of Hadoop and uses spark for operations.

**Roles and Responsibilities:**

- Understand the upstream source nature and work with business cases.
- Connect That Created Postgre instance using workbench.
- Initial Data load using Initial\_Dataload.sql Check table structure and sample data.
- Initial Data load using Initial\_Dataload.sql Check table structure and sample data  
Shared responsible for administration of Hadoop, Hive.
- Copy all dependencie on EMR.
- Written Spark job to Process data.
- Import Ref data from RDBMS to Hbase.
- Create Hive tables deployment to transformed data using hive DDL table.
- To process delta, Use below script Copy data into Hbase at/ data location.
- Customer can extract report from s3 and can use for their further use like visualization etc. And extra
- To analyze and solve problems and put hot fixes on production.

**Environments:** Spark, Hbase, Phoenix, Hive, RDS, S3, AWS, postgresql.

**EDUCATIONAL QUALIFICATION**

Year of passing	Examination/Degree	Institute/University	% Score
2011-2015	Bachelor of Engineering (CSE)	Dr. BAMU University	60%
2008-2011	Computer Engineering (Diploma)	Government Polytechnic, Jalna(MSBTE)	61%
2007-2008	X (High School)	Godavari Public School, Aurangabad	87.53%

**PERSONAL SKILLS**

- Eager to learn new things as well as good learner.
- Strong analytical, design skills and problem-solving abilities.
- Good management skills.
- Leadership skills.

I here by declare that the information given here with is correct to my knowledge and I will responsible for any discrepancy.

**JADHAV DEEPALI VISHNU**