

D DHRUVA KUMAR

Email: dhruvakumar.308@gmail.com

Mobile: +91-9701513241

Professional Summary:

- 4.5 years of overall experience in IT industry with Experience in areas – Requirement gathering, Analysis, Development and Maintenance and client interactions.
- 3 years of experience in **Big Data** Technologies like **HDFS, HIVE, SPARK, SCALA, SQL, Unix Shell Scripting & Aws**
- Full life cycle experience of big data projects in data ingestion, transforming data and loading back to visualization modules.
- Effectively involved in unit testing, bug tracking, reviewing and analysing result.
- Involved in all the activities of Software Life Cycle.
- Project lifecycle experience Waterfall model and Agile.
- Involved into data analytics based on dashboard.
- Experience working in both team and individual environments.

Experience Summary:

- Working as a **Big Data Engineer** in **Compusol Software PVT Ltd** from Mar 2018 to till date.

Educational Qualification:

- B. Tech from **GIET** in 2015.

Technical Skills:

Big data Technologies	: Spark ,HDFS, Hive
Programming Language	: Scala.
Cloud Platform	: Aws.
Database	: MySQL, Oracle
Other Tools	: IntelliJ IDE
version control Tools	: Git.

Work Experience:

Project-2:

Project Title : Provider Directory Modernization

Client : Globant

Environment : Spark, Scala,Hive

Description:

Comparison of EPIM golden records to the NDB records to be as dashboard for showcase of the mismatch – number of records, type of data elements, etc., The database is intended to used by business teams – EPIM, IT teams and UAT teams. The dashboard will ease the identification and review of the gaps in the matching and information would be used by the connected team to rectify the errors/mismatched records.

Responsibilities:

- Analyse and identify the resources (NDB, EPIM, NICE) to extract information for selected data elements.
- Reading the Json file generated by the EPIM for the input data elements.
- Converting the JSON file data from EPIM into data frames, so the records can be easily compared at data element level
- Comparing the EPIM and NDB records and loading the result data to ES.
- Creation of the Step functions for Spinning up the cluster (EMR), Execution of Multiple Spark Jobs and Terminating the Cluster.

Project-1:

Project Title : User Data Management Platform

Client : The Hindu

Environment : Spark, Scala, Hadoop, Hive & Unix.

Description:

UDMP - contain information related to users and articles read by users. With captured users data is used to recommend them latest updated articles as taking interests of users.so engaging the users in Hindu site and to subscribe to the new plans.

Responsibilities:

- Involved into push our target tables into MySQL database.
- Involved into data analytics based on dashboard.
- Developed Spark Scala scripts for ad hoc analysis.
- Analyzing the source systems data before loading into HDFS.
- Involved in data loading and validating the data fields as part of data loading.
- Involved in creating Hive tables, loading with data and writing hive queries.