

Professional Summary

4+ years of **IT** experience coupled with 2+ years in **Big Data** Technologies.

Exposure in design and development of solutions for Big Data using the Hadoop eco system technologies (**HDFS, Sqoop, Hive, Apache Spark**).

Experience in creation of on-demand **EC2** instance with Custom step execution using **Amazon Web Services(AWS)**

Developed Spark Applications using spark tools like **Spark-Core** and **Spark-Sql**. Having hands-on experience on **complex data processing**

Extensively worked on Spark using Scala (**Scala-Spark**) and Spark using Python (**Pyspark**)

Capable of processing large sets of **Structured, Semi-structured data**. Experience on **Joins** and **Aggregate** functions on Data frames.

Handled multiple reads and writes involving **S3 buckets**

Loaded and transformed large sets of semi structured data like **JSON, Avro, Parquet, XML, ORC**.

Involved in creating **Hive tables**, data loading and writing **hive queries**.

- Currently Working as Hadoop Developer in **Netfinity Technologies India Pvt Ltd**.

Skills

- **HADOOP**
- **SQOOP**
- **HIVE**
- **SCALA-SPARK**
- **PYSPARK**
- **MYSQL**
- **AWS**

Work History

Project 1:

Project Title	NextGen
Period	08/2020 to Till Date
Technical Stack	Spark/Scala, Sqoop, Hive, Hadoop, AWS-S3.

Project Description:

NextGen is one of the biggest health care clients in US. This project is all about converting different input structure of data into a common structure and then storing it into HDFS to perform data analytics on top of it.

Roles and Responsibilities:

- Involve in technical design approaches.
- Performance optimization based on Scala-spark.
- Responsible for Master data management using Scala-spark.
- Responsible for running spark jobs with huge volumes of data.
- Writing test cases for Spark.

Project 2:

Project Title	EkStep Learning Platform.
Period	06/2019 - 06/2020
Technical Stack	Spark/Scala, Cloudera Distribution.

Project Description:

Ekstep is non profitable organization for giving proper education to children. platform for the children. Any person can post the content on this platform and it is free for everyone.

Roles and Responsibilities:

- Analytics on top of events generated while operating on genie app with scala and Spark.
- Writing test cases for Spark.

Education:

M.B.A from SVU, Tirupati

B.Com Computers from SVU, Tirupati .

Project 3:

Project Title : Web Analytics

Period : 06/2018 to 05/2019

Project Description :

In a banking website, identify the stage up to which a customer has reached in a particular funnel (like new account opening, loan application). This can be used to draw conclusions like how many customers are qualifying for the eligibility stage, proceeding till the last stage and completing the process. Based upon this, the client can also decide whether to campaign for the customer about certain plans or policies on the related schemes which were browsed by him/her.

Roles and Responsibilities:

- Find the most influential features among all features.
- Find the similarity measures between reached customers and dropped at middle stages using correlation, cosine similarity etc.
- Summarize statistics about web pages using D3.