Leela Kumar
Hadoop developer

Mobile No: +91-7207750595
Email ID: sreeram.leelakumar001@gmail.com

EXPERIENCE SUMMARY

- I am Senior Engineer with Mind Tree with a total of 4.2 years of IT experience.
- 2 . 6 years of exclusive experience in Hadoop and its environment components like HDFS, Spark Scala, Hive, Sqoop, Zookeeper, PySpark, Python, AWS.
- Good working knowledge with Apache Spark and Hadoop Ecosystems.
- Knowledge on SQL.
- work Experience in Python (Pandas).
- Good knowledge in Cloudera Hadoop Distribution.
- Knowledge of job workflow management and coordinating tools like Oozie and Airflow.
- Worked in a strong Agile Methodology team to achieve high throughput.
- Experience in Relational database (RDBMS), Individual capability to manage the project modules and to write Complicated Ad-hoc queries using SQL.
- Experience in writing custom UDF's in Hive and custom Input Formats in Map Reduce.
- Good Knowledge in writing Spark Applications in Python and PySpark.
- Replaced existing MR jobs and Hive scripts with Spark SQL, Spark data transformations for efficient data processing.
- Good working knowledge in Scala and Python in Hadoop.

SKILLS

HDFS

SPARK

SQOOP

HIVE

AWS

PYTHON

SCALA

AIRFLOW

EDUCATIONAL QUALIFICATION

B.Commerce (Computers Applications) with 73% from **Vikrama Simhapuri University,** Andhra Pradesh in 2017.

WORK EXPERIENCE:

Currently working as Hadoop Developer in **Mind Tree.**

PROJECT SUMMARY

Project #1:

- ➢ **Project Name** : Human Activity Patterns from Smart Home **Big Data** for Healthcare Applications.
- ➢ **Environment** : Linux, Cloudera.
- ➢ **Technologies** : Sqoop, Hive, PySpark, Spark SQL and Oozie.
- ➢ **Role** : Hadoop Developer.
- ➢ **Duration** : Aug-2021 to tilldate.

Description:

Converting the data stored in HDFS into Spark RDD Performing operations on data in Hadoop using Spark Context, Spark SQL, Data Frames and Pair RDD Experience in handling large datasets using Partitions, Spark In-memory capabilities, Broadcasts in Spark. Written Spark SQL code for huge structured the purpose of publishing this in this project, we will process the metadata datasets coming with the thousands of records of patients from various hospitals.

Roles and Responsibilities:

- Developed Spark SQL queries for processing HIVE tables by using hive context.
- Experienced in implementing Spark RDD transformations, actions to implement business analysis
- Developed a script in Scala to read all the Parquet Tables in a Database and parse them as Jsonfiles, another script to parse them as structured tables in Hive.
- Implemented more business requirements using spark core, Spark SQL, Scala.
- Involved in creating data frames using SparkSQL
- Developed Oozie workflows for scheduling and maintaining several batch jobs to run automatically depending on Business requirements.

Project #2:

- ➢ **Project Name** : NetGen.
- ➢ **Environment** : Linux, Cloudera.
- ➢ **Technologies** : Sqoop, Hive, Spark, Scala, AWS, Airflow.
- ➢ **Role** : Hadoop Developer.
- ➢ **Duration** : Feb 2020 to June 2021.

Project Description:

NetGen is one of the biggest health care client 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 #3:

Client: Nestle

Role: ETL Developer

Skills: MS Excel, My SQL, SQL

Duration : Sep-2018 to Dec-2019.

Description:

Nestle S.A. is a Swiss transnational food and drink company headquartered in Vevey, Vaud, Switzerland. It has been the largest food company in the world, measured by revenues and other metrics. Nestles products include baby food, medical food, bottled water, breakfast cereals, coffee and tea, confectionery, dairy products, ice cream, frozen food, pet foods, and snacks.

Roles and Responsibilities:

- Involved in understanding the client requirements and identifying the sources.
- Well versed with complex SQL Queries.
- Good experience in customizing and generating reports from the data Collected from different data-sources (retail stores) across the globe as per business needs.
- Mentor junior team members in the SQL Queries, Report Generation and data load monitoring.