SUMANTH AKULA

Contact No: +91-9581865517

Email Id: aakulasumanth@outlook.com

PROFESSIONAL SUMMARY:

- Around 4+ years of overall experience in program development which includes hands on experience in **Big Data Ecosystem** related Technologies.

- 3.8 years of comprehensive experience as a **Big Data Developer**.

- Hands on Experience in **SPARK** using **SCALA**.

- Experience in using **Yarn, Hive, Sqoop, Impala, HBase, Pig** and **Cloudera (CDH5)**.

- Experience in working with **MapReduce** programs using Apache Hadoop for working with Big Data.

- Experience in importing and exporting data using **Sqoop** from **HDFS** to Relational Database Systems and vice-versa.

- Extending **Hive** and **Impala** core functionality.

- Experience in analyzing data using **HiveQL**, **Pig** Latin and custom **Spark** programs in Scala.

- Worked on NoSQL databases like **HBase.**

- Experience in job workflow scheduling and monitoring tools like **Oozie, RLM**

EMPLOYMENT:

- Currently Working as Senior Data Analyst for IBM.

TECHNICAL ENVIRONMENT/SKILLS:

- **Languages:** Core Java, Scala

- **Big Data Tools:** Hadoop, HDFS, MapReduce, Hive, Pig, HBase, Impala, Spark, Sqoop, Oozie, Zookeeper, Kafka, Cloudera CDH5, HiveQL, Pig Latin, Hortonworks

- **RDBMS:** Oracle, MySQL, SQL Server

- **No SQL:** HBase

- **Scripting & Query Languages:** Shell Scripting, SQL

- **Environment:** Unix, Linux, Windows

- **Scheduling Tools:** Oozie, RLM, Autosys

WORK EXPERIENCE:
Client: CITI BANK AML & DDAR Project (April 2021-Present)
Company: L&T infotech
Role: Senior Data Analyst/Business Analysis

Description:

Analyzing the data based on the retention factor and generating the reports for the data deletion purpose using different process frameworks and data quality checks.

Environment: Hadoop, Map-Reduce, Spark, Scala, Hive, Sqoop, RLM, Autosys, Bit bucket, Git, Jenkins, Oracle, SQL.

Responsibilities

- Preparing the business logic based on the guidelines of respective region regulatory Authorities.
- Loading the data using Sqoop into the framework of Hive meta store tables and views.
- Preparing the build by pushing the code to Git using the Jenkins.
- Deploying the code using the RLM and verify the Linux servers.
- Checking the data loaded with various conditions across different layers of the framework.
- Adding required logic to the framework using spark Scala.

PRODUCTS AND PLATFORMS (JAN2019-APRIL 2021)
Client: IBM/Cloudera
Company: HCL Technologies
Software Developer
Description

Obtaining the Cases from the various clients. Re-solving the tickets based on the priority.

Environment: Hadoop, Map-Reduce, Spark, Hive, Sqoop, Impala, Oozie, Oracle

Responsibilities

- Checking for the cause for the issues and their reasons to resolve
- Analyzing the requirements and preparing for best solutions.
- Involved in importing data from RDBMS to HDFS using Sqoop.
- Using Impala to perform data Validation on the data ingested using Sqoop.
- Data stored on HDFS and processed by using various tools in Hadoop.
- Running of the Spark submit jobs and monitor using Oozie for the log details.
- Writing Scala code for the process to be done as client requirement.
- Experienced in Performing tuning of Spark Applications for setting right Batch Interval time, correct level of Parallelism and memory tuning.

DATA LAKE CREATION(July2018-DEC2018)
Client: Versced/EMIDS   Payroll: Intelcubes incorporated
Hadoop/Spark Developer
Description

Creating a Datalake for all the Healthcare data of the client from different sources to a single general format, so that any Data Scientist can easily query the dataset and get the desired possible combination of the

required details. And the data will be stored in Hdfs browser with the relevant partition as parquet files. Also, the resultant data is stored as tables in impala integrated with hive.

Environment: Hadoop, Map-Reduce, Spark, Hive, Sqoop, Impala, Oozie, Oracle,

Sql server management studio, Scala, CDH5, Shell scripting.

Responsibilities

- Analyzing the requirements and preparing for best solutions.
- Involved in importing data from RDBMS to HDFS using Sqoop.
- Using Impala to perform data Validation on the data ingested using Sqoop.
- Data stored on HDFS and processed by using various tools in Hadoop.
- Running of the Spark submit jobs and monitor using Oozie for the log details.
- Writing Scala code for the process to be done as client requirement.
- Experienced in Performing tuning of Spark Applications for setting right Batch Interval time, correct level of Parallelism and memory tuning.
- Loaded the data into Spark RDD and do in memory data computation to generate the output response.
- Optimizing of existing algorithms in Hadoop using Spark Context, Spark-SQL, Data Frames and Pair RDD's.
- Performed advanced procedures like text analytics and processing, using the in-memory computing capacities of Spark using Scala.
- Experienced in handling large datasets using partitions, Spark in Memory capabilities, Broadcasts in Spark, Effective & efficient Joins, Transformation and other during ingestion process itself.

EDUCATION:

Bachelor of Engineering   JUNE 2010 TO MAY 2014

     R.M.K. College of Engineering and Technology

     Anna University, Chennai

     CGPA: 7.0

12th Std. (Board of Intermediate Education, A.P)   JUNE 2008 TO APRIL 2010

     Narayana jr college Nellore AP

     Secured 88.1%

10th Std. (Secondary School Certificate, A.P)   JUNE 2007 TO APRIL 2008

Ravindrabarath em school, Nellore AP

Graduated 2008

Secured 83.3%.

DECLARATION:

 The above information is true to the best of my knowledge and belief.

Date:

Place: Bangalore   Name : Sumanth A