**SOMASEKHAR. K**

**E-mail:** [**somasekarak@gmail.com**](mailto:somasekarak@gmail.com) **Mobile:+91-9908660003**

**Experience Summary:**

* Having 4 years of overall IT experience in Application Development in **SQL, Big Data Hadoop & Spark.**
* 2.3 years of extensive working experience on Hadoop framework and its components like**: HDFS, SQOOP, HIVE, MAPREDUCE, FLUME,SCALA & PYTHON.**
* Good knowledge on designing and implementing ETL process to load data from various data sources to HDFS using Sqoop, performing transformation logic using Hive and integration with BI tools for **visualization/reporting**.
* Good working knowledge with **Apache Sqoop, Kafka and Hive.**
* Very good understanding of **Partitions, Bucketing** concepts in **Hive** and designed both **Managed and External** tables in Hive.
* Performed Importing and exporting data into **HDFS** and **Hive** using **Sqoop**.
* Involved in writing the **Hive** queries.
* Knowledge in HBase, MapReduce, Spark and Python.
* Experience in importing and exporting the data using **Sqoop** from **HDFS** to Relational Database systems.
* Experience in handling different file formats like RC, ORC, Parquet and Sequence file, Text file formats.
* Good understanding of Hadoop architecture and hands-on experience with Hadoop components such as Job Tracker, Task Tracker, Name Node, Secondary Name Node, Data Node, Map Reduce concepts and **YARN** architecture which includes Node manager, Resource manager and App Master.
* Having good exposer on **Spark (RDD**) and its working functionality.
* Strong Knowledge on Spark architecture and its integrations like **Spark SQL**, **Data Frames** and **Datasets**.
* Hands on experience on Memory Components **Apache Spark Core, Spark SQL**.
* Having knowledge on Scheduling all Hadoop/ Pig/ **Hive**/ **Sqoop** actions.
* Experience with Cloud computing Technique like **AWS (S3).**
* Strong knowledge in writing complex queries for oracle.
* Hands on experience with **Agile methodology**.
* Excellent problem solving skills with a strong technical background and result oriented team player with excellent communication skills.

**Technical Skills:**

**Big Data : Hive, Sqoop, HDFS, Flume, Spark (pyspark), HBase, MapReduce.**

**Database : Oracle 10g, MySQL.**

**Other Tools : WinSCP, Putty, Eclipse.**

**Operating Systems : Windows, Linux.**

**Languages : SQL, HiveQL, Scala and basics of python.**

**Hadoop Distributions : Cloudera 5.13.0**

**Scheduling Tools : Crontab & basics of Oozie.**

**Work Experience:**

* Working as Big data developer with **Cargill** in Bangalore, Since Nov 2017.
* Worked as Software Engineer with **Accenture** in Bangalore, From Jan 2016 to Nov 2017.

**Education:**

* Bachelor of Technology from **JNTUA University.**

**Project 3:**

**Project Title :** **Cargill Billing System**

**Client : Cargill , Minneapolis**

**Role : Hadoop and Spark Developer**

**Environment : Hadoop, Hive, Sqoop, Scala, HDFS, Apache Spark, Spark SQL**

**Project Duration: Aug-2018 to till date**

**Description:**

Payment solution aims to simplify and connect the entire payments world. Creating a dynamic, evolving solution and providing tools for pro-active collaboration amongst all parties. While disrupting and liberating us from today’s static and restrictive payments technology and business models. The data ingestion happens from RDMS to Hadoop. Once the ingestion happens, perform the analytical queries on top of virtual tables of Spark Data Frame. Based on the aggregated queries, the system is completely transparent to a client. We analyze the incoming data processing through a series of programmed jobs and deliver the desired output and present the data into the portal so that it could be accessed by different teams for various analysis and sales purpose. The Payment platform provides an on-line portal for merchants to search and report on their transactions as well as providing transaction handling capabilities to manage exceptions such as refunds and chargeback.

**Responsibilities:**

* Interacting with clients in gathering and understanding the requirements.
* Reading data from File System into a **Spark RDD.**
* Loading files to HDFS and writing hive queries to process required data.
* Writing the **Transformations** and **Actions** for processing the data.
* Writing the Spark SQL write the SQL queries for incoming data.
* Created hive table with schema and loaded the data using **Spark SQL** and **data frames**.
* Developed Hive external tables to store the processed results in a tabular format.
* Worked in migrating **HiveQL** into **Impala** to minimize query response time.
* Experience in developing **HiveQL** scripts for Data Analysis and **ETL** purposes.
* Partitions of Hive tables, Creating Hive tables to store the processed results in a tabular format.
* Writing the script files for processing data and loading to **HDFS**.
* Moved all log/text files generated by various products into HDFS location.
* Created External Hive Table on top of parsed data.
* Import the data from different sources like **HDFS** into **Spark RDD.**
* Experienced in implementing **Spark RDD transformations**, **actions** to implement business analysis .
* Coordinating the team and resolving the issues of the team technically as well as functionally.

**Project 2:**

**Project Title : CBSM**

**Client : Cargill, Minneapolis**

**Role : Hadoop Developer**

**Environment : Hive, Sqoop, MapReduce and HDFS**

**Project Duration : Nov -2017 to July-2018**

**Description:**

Data from various sources and provided in log file formats. The data is cleansed and extracted using Map Reduce and transformed using Hive and the analyzed data is provided to the digital campaign team. The user data is used to send targeted ads on the promotional items of the Cargill Products.

**Responsibilities:**

* Interacting with clients in gathering and understanding the requirement.
* Importing data from RDBMS to **HIVE**, **HDFS** and Exporting data from **HIVE**, **HDFS** to **RDBMS** by using **Sqoop**.
* Developed worked with **SQOOP** incremental loads to populate **HIVE** external tables.
* Creating HIVE external tables to store the results on top of **HDFS**.
* Involved in loading data from Linux file system to **HDFS**.
* Creating **HIVE** tables with Partitions and **Buckets**.
* Analyzed transactions using Pig scripts and Hive to generate reports for end users.
* Involved in resolving the JIRA issues with in SLAs.

**Project 1 :**

**Project Title : VCE**

**Client : Dell Technologies**

**Role : Technical support engineer**

**Environment : Oracle, SQL, LINUX**

**Tools** : **SQL Developer, TOAD, Putty**

**Project Duration: Mar-2016 to July-2017**

**Description:**

Closed Loop Feedback is a system that gathers the customer’s complaints which are logged through either Service Center/Call Center and performs surveys on those customer’s complaints by calling back to them  in order to get their feedback on their issue resolution and asking net promoter score(NPS) to the customer on their services provided in resolving their issues. Based on this NPS calculation and the feedback received from the customers the service providers can generate the reports on the customer feedback received and can concentrate on enhancing or improving their services on which they are lacking. This system will gives a clear picture to the service providers on which service areas they are depreciating and on which area they have to concentrate more so that a customer will be more satisfied in order to increase their sales.

**Responsibilities:**

* Resolving the tickets based on SLA.
* Handling Production tickets through Requestee and JIRA.
* Generating Reports Daily, Weekly And Monthly.
* Daily jobs include monitoring the application, data flow, solving daily issues.
* Monthly billing activities and reporting activities
* Responsible for providing the support in L1/L2 level depending on the priority of the issues to meet client’s SLA.
* Interacting with customers while resolving the incidents and tasks.
* Responsible for raising tickets in Requestee.
* Give the Weekly and Monthly wise SLA report to the Manager.