SASIDHAR DULIPUDI
Email: sasidhardulipudi@gmail.com
Mobile: +91 9908438890   **Hadoop Consultant.**

Experience Summary

I am having 6.2 years of work experience on Big Data Technologies and Analytics at below organizations.

Tata Consultancy Services : June 2015 – June 2018

Capgemini   : July 2018 – Present

Technical Skills

Apache's Open Source Frameworks   : Hadoop, Spark, Impala, Hive, Kafka, Hive

Programming Languages   : Python, Java.

Databases   : MySQL, MariaDB, MongoDB.

Version Control Systems   : Git, Bitbucket

Continuous Integration   : Team Foundation Server, Jenkins

Machine Learning   : Neural Networks and Linear regression

Qualifications

Qualification	Group	Percentage/CGPA
Bachelor of Technology(JNTUK)	Computer Science Engineering	8.14
Standard XII	M.P.C	90.9
SSC	NA	86.7

Project 1

Name   :   **Customer Financial Attribution (CFA)**

Skills used   :   Apache Spark, Apache hive, MariaDB

Description   :   CFA deals with the Lump sum revenue generated by individual source systems like Debit card and Credit card annual fees, ATM income, Insurance Premium etc. The lumpsum amount is attributed back to respective departments based on their contributions.

Project 2

Name	:	**Enterprise Performance Management System - Fund Transfer Pricing**
Skills used	:	Apache Spark, Apache hive, MariaDB
Description	:	FTP individually measures how much each source of funding is contributing to overall profitability of the firm, which then used by the bank as means of outlining the areas of strength and weakness within the funding source basing on the operational needs and the function of loan and deposit divisions.

Project 3

Name	:	**Advanced Communication Surveillance**
Skills used	:	Apache Spark, Apache Hive, MySQL, Apache Tika
Description	:	To improve the performance, an existing Communication Surveillance project, which was developed using R, was migrated to Spark. The analysis was performed on five communication sources that includes Chat and Mail. Project includes efficient extraction of huge zip files using custom API, Parsing the XML, Converting the attachments to text using Apache Tika, filtering the active traders, Lexicon matching, Alert generation for business users.

Project 4

Name	:	**Interactions**
Skills used	:	Spark Streaming, Apache Kafka, Apache Hive, MongoDB, Apache Oozie.
Description	:	Interactions project provides 360-degree view to users by processing Interactions data from various sources (IVR, WebActivity, Salesforce.com). Each source has both Stream and Batch data that has to be processed.

POC

Name	:	**Twitter Sentiment Analysis**
Skills used	:	Apache Spark, Apache Flume, Apache Hive.
Description	:	Twitter sentiment analysis was developed using Spark. Tweets were collected based on keywords using Apache Flume, and using Hive to store the data. Spark was used to process the data and find the sentiment based on the predefined lexicons and distance between them.