DHAKSHANA V

DATA ENGINEER

Extremely motivated to constantly develop my skills and grow professionally. I am confident in my ability to come up with interesting ideas for unforgettable Challenging campaigns.

EDUCATION

- **Bachelor Degree in Computer science and Engineering** (Full time course)
  2013-2017
- 10th & 12th Above 80%

- Participated in the LiveLife Education Programme conducted by LiveLife Education at Avinashilingam Institute for Home Science and Higher Education for Women,
- Animation and Android App Development based Implant Training conducted by UNIQ SOLUTIONS

TECH SKILLS

- Hadoop ecosystem and frameworks - **HDFS, Yarn, Hive, Sqoop**
- Framework - **Apache Spark**
- SQL-based technologies - **MySQL, Oracle DB**
- Programming Language - **Pyspark**
- Operating system - **UNIX, LINUX, MS Windows**
- Trained in **BIGDATA** at IBM, Bangalore.
- Cloud Data Mining Tools **IBM SPSS Modeler**
- **Kafka** ETL & ELT Pipe line analysis

CONTACT

+91 78 0689 9599

dhakshanashiva@gmail.com

@ Dhakshu_Shivan

257/1 Pallakadu 2nd Street, Chinnathirupathy - Post Salem - Tamilnadu - 636007

WORK EXPERIENCE

IBM India Pvt Ltd

DATA ENGINEER (Full Time)

March 30 2019 - Till date

Project/Client - VODAFONE IDEA Pvt Ltd.

- **Aggregation** Project to provide ways to analyze and systematical extraction, transformation, KPI extractions, load and store large and complex data in adequate time and optimizing the complexities deals with traditional data-processing applications.

- **Legal consolidation** Application used to get requests from the legal users to track particular MSISDN's and IMEI's activities for mentioned number of days. For generating reports for the request, application used to receive large amount of data's on daily basis from many source systems in different formats like .CSV, .TXT, .DEL. Validating and Loading the required data into the source tables also taking care by legal app.

CONTRIBUTION IN AGGREGATION

- Implemented the complete Pyspark framework with real-time processing.
- Used Spark 2.1 API to stream data from various sources
- Developed Pyspark framework using Spark SQL and Data Frames for aggregation.
- Worked with Sqoop to ingest and retrieve data from various Databases.
- Created Schema in Hive with performance optimization using Bucketing and Partitioning.
- Written Hive queries to transform data for further downstream processing
- Cleared SIT and UAT for aggregations.
- Involved in end to end Deployment process.

CONTRIBUTION IN LEGAL CONSOLIDATION

- Created Extraction Module to fetch data from Hadoop servers and loading into Hive using Pyspark.
- Created Control_file for moving all the user's request from Legal_app to Oracle database.
- Contributed in the development of Retrival_application.