

## DATA ENGINEER

### KIRAN DEEPAK BHALEKAR

Phone +91 8329837920

Email: [kiranbhalekar43@gmail.com](mailto:kiranbhalekar43@gmail.com)

LinkedIn: <http://www.linkedin.com/in/kiran-bhalekar-a95584166>

## SUMMARY

- ◆ Overall 3+ years of experience in IT industry including technical proficiency in big data environment with extensive expertise in development on Hadoop ecosystem and Cloud Data Engineering(AWS).
- ◆ Hands on experience with Amazon EC2, Amazon S3, Amazon RDS, IAM, Cloudwatch, Lambda, EMR, Sagemaker, Step Function, Glue and other services of AWS family.
- ◆ Hands on experience with big data core components and Eco System including Data Ingestion and Data Processing (Pyspark, Hive, SQOOP, HDFS)
- ◆ Have experience writing data extraction logic in SQL.
- ◆ Perform tasks such as writing scripts, SQL queries, Python cases.
- ◆ Experience in using Hadoop distribution like Cloudera.
- ◆ Develop batch processing, integration solution and process structured and Semi-structured data.
- ◆ Experience in transferring data from RDBMS to HDFS and Hive tables using Sqoop.
- ◆ Experience in creating tables, partitioning, bucketing, loading in Hive.
- ◆ Gather and process raw data at scale and analyse processed data.
- ◆ Resolve JIRA tickets raised by Client.
- ◆ Very Good understanding of Partitioning, Bucketing concepts in Hive and designed both Internal and External tables in Hive to optimize performance.
- ◆ Strong problem solving & technical skills coupled with clear decision-making.
- ◆ Hands-on Hadoop eco-system. Good exposures in Requirement Analysis, Customization and user acceptance.

## Education

- ◆ Bachelor of engineering in Computer Science And Technology, MSS college of engineering and technology, Jalna.

## DATA ENGINEER

### Technical Skills Set

- ♦ **Big Data Technologies:** Hadoop, Mapreduce, HDFS, Sqoop, Hive, HBase, Kafka, Zookeeper, YARN, Apache Spark.
- ♦ **Programming:** Python, SQL, PySpark.
- ♦ **Cloud Technology:** AWS.
- ♦ **Operating Systems:** Linux(Ubuntu), Windows 7/8/10.
- ♦ **Visualisation Tool:** Power BI.

### PROFESSIONAL EXPERIENCE:

**Organization:** Advent Informatics

**Project Name:** Sentiments Analysis

**Oct 2021 – Present**

**Role:** Data Engineer

---

#### Description:

Sentiment Analysis takes the daily responsibilities which include data cleaning and monitoring the data pipeline after deployment. We get the data from our client on s3 bucket then lambda is triggered. It further executes the step function. The step function starts the sagemaker instance where our job is cleaned and passed to data scientist and our final output is also given on s3 bucket.

#### Responsibilities:

- ♦ New Development as per requirements.
- ♦ Enhancement of the already developed applications.
- ♦ Support if there are any issues or job failures.

#### Environment:

AWS Services used: S3, Lambda, Cloudwatch, Step functions, Glue, Sagemaker, ECR.

## DATA ENGINEER

**Project Name: Medicinal Warehouse**

**March2020 –Sept2021**

**Role: Data Engineer**

---

### **Description:**

In this project we receives various feeds of products in the CSV format from upstream into the AWS, from S3 we import this data perform cleaning and after validating this cleaned data we store it in AWS RDS in different Schema using Glue PySpark job. Again extract the cleaned data from RDS and apply the business logics as per the client requirements and load this final data to another S3 bucket for frontend applications of SAP using Glue PySpark.

### **Responsibilities:**

- ◆ Attaining a daily scrum call with team members and discussing the task on a particular sprint.
- ◆ In every sprint we have assigned tasks that we are completing as per timeline.
- ◆ Weekly call with client for their requirements and further enhancement.
- ◆ Analysing PROD tickets – Issues can be data related or code related.
- ◆ Job monitoring
- ◆ Import data from S3 raw bucket and do the cleaning and masking and again
- ◆ move to s3 cleanse bucket by using glue.
- ◆ Worked on the transformation glue job to move data from the s3 cleanse to the redshift dimensional and fact tables.
- ◆ Validating the records once the transformation glue job completed the loading of data into redshift.
- ◆ Work with the teammates to understand and support them in importing data from source to S3 by using AWS DMS.

### **Environment :**

AWS S3, AWS Glue, AWS RDS, SAP, JIRA

### **Personal Details:**

Name : Kiran Deepak Bhalekar

Gender: Female

Date of Birth: 15 -July-1996

Nationality: Indian

### **Declaration:**

I hereby declare that the information furnished above is true to the best of my knowledge

Date: / / 2022

Kiran Deepak Bhalekar