Rohit Kumawat

Data Engineer

Contact No : +91 810-740-9988
LinkedIn : https://www.linkedin.com/in/rohit-kumawat-0233a277

Email: Rhtkumawat55@gmail.com
GitHub:https://github.com/kumawat55

Data Engineer equipped with 3+ years of extensive experience in application development. Employs excellent leadership skills and multi-tasking strengths. Demonstrated ability to improve data operations.

Experience

Employer : Tata Consultancy Services Ltd.
Project #1: High Performance Solution Engineering (Team - HPSE)
Client: Humana Inc.
Role: Application Developer (GCP/Big Data/Spark)
09/2018 – Present   Hyderabad, India

Responsibilities

- Developing StreamSets pipelines to migrate data from origins like (SQL Server, Oracle, PostgreSQL) to GCP destinations (Google cloud bucket, Big Query) or vice-versa.
- Developing auto scalable GCP (Dataflow) pipelines for batch data processing using Apache Beam and python.
- Developing pyspark scripts for extracting data from on-premises sources like (SQL server, Oracle) .
- Developing pyspark scripts for extracting data from different file formats (feather files, text files, csv files, json files, Avro files, parquet etc.) for performing various transformations and loading data to google Cloud storage and cloud SQL tables.
- Developing server less Cloud functions using python to perform specific event-based tasks.
- Developing SQL scripts for managing tables in Cloud SQL and remotely trigger them using python-based cloud functions.
- Responsible for design and development of big data solutions.
- Test cases preparation, reviewing with peers and performing unit testing or System integration testing.
- Understanding the criticality of the customer from Data perspective, along with the responsibilities in updating the Roadblocks, Status, and progress of the assignments of the team in the Scrum/Standup meeting.
- Learning new technologies as part of career aspiration and applying the methods learned in the day to day developing approaches.

Skills

- **Google Cloud Platform**
  - Data Flow
  - Gcloud, gsutil
  - Cloud Storage
  - Cloud Functions
  - Cloud SQL
  - Big Query
  - DataProc
  - PubSub
- **Hadoop Framework**
  - Hive
  - Oozie
  - Map Reduce
  - HDFS
  - Impala
- **Programming Skills**
  - Python3
  - Big Query
  - C/C++, Java
  - Hive QL, SQL
  - Spark/Pyspark
  - Shell-scripting
  - Git
- **Tools**
  - StreamSets
  - Hue™ 4.1.0 - The Hadoop UI
  - Putty, Toad Oracle
  - PyCharm
  - SQL Server Management Studio

Project #2: Digital Data Analytics – EDA1
Client: Humana Inc.
Role: Application Developer (Big Data/Spark)
09/2017 – 08/2018   Hyderabad, India

Responsibilities

- Developing hive QL that works seamlessly on Hadoop clusters.
- Developing Spark/Pyspark code to process unstructured data and storing in HDFS.
- Creating and maintaining the hive tables equivalent to non-traditional sources tables.
- Extracting out derived information by implementing complex queries for Hive and storing results in derived tables, basically ETL operations.
- Building the process how to decode the data and transforms the data.
- Developing Python scripts for loading large amount of data.
- Managing workflows and scheduling oozie jobs with Python and shell-scripting or through web interface.
- Validating the data across the files and databases, end to end along with negative test scenarios to verify the performance and to avoid junk data in the source, Target.
- Test cases preparation and performing unit testing or System integration testing.

Certifications

- **Oracle Cloud Infrastructure 2019 Certified Architect Associate**
  Issued: Jun 2020   Expires: Dec 2021
- **Oracle Cloud Infrastructure 2019 Certified Architect Professional**
  Issued: Jul 2020   Expires: Jan 2022
- **Oracle Cloud Infrastructure 2019 Cloud Operations Certified Associate**
  Issued: Jul 2020   Expires: Jan 2022
- **Red Hat Certified System Administrator (RHCSA EX-200)**
- **Red Hat Certified Engineer (RHCE EX-300)**

Trainings

- **Tata Consultancy Services Ltd., Hyderabad**
  Role: Trainee (Initial Learning Program)
  06/2017 – 09/2017 Hyderabad,India
  Achievements/Tasks
  - o Attended internal training on Java, along with the practical hands-on live projects.
  - o Attended internal training on Hadoop tools and their fundamentals under expert's guidance.

- **Maxgen Technologies Pvt. Ltd. (Pune)**
  Role: **Intern (Java developer)**
  06/2016 – 07/2016   Pune, India
  Achievements/Tasks
  - o Developed java-based modules for live project of library system under expert's guidance. My role include gathering requirements from experts, developing modules by writing code, maintenance troubleshooting and debugging.

Education

- **Bachelor of Technology**
  Grade: 70.97 %
  Computer Science - 2017
  Arya college of Engineering & I.T.
  Jaipur(Rajasthan,India)
- **Senior Secondary Cert.**
  Grade : 73.04 %
  Stream : Mathematics & Science
  Tilak Public School
  Jaipur(Rajasthan,India)

Achievements

- **Inter Account Competition (01/2016 – 09/2016)**
  Basketball, Volleyball, Throw ball winner for HUMANA account in TCS.
- **Smart India Hackathon (04/2017 – 04/2017)**
  Top -8 listed team in Smart India Hackathon (2017) with project River monitoring System in Indian forest ministry.
- **TCS–CodeVita Season 5 (07/2016 – 09/2016)**
  Cleared TCS code vita season 5 (round1 and round2) and got offer from TCS.

Hobbies

- Basket Ball
- Cricket
- Painting
- Snooker
- Watching Scientific-Fictions

Languages

- English
- Hindi
- Rajasthani

Personal Projects

- **Impractical Python - Learning module (03/2019 – Present)**
Integrated python module for various types of algorithms (searching, sorting, graph, trees, math Etc.), Basic idea is to integrate large sets of algorithms at one place and reflecting them on web interface with their run time implementation/working nature for beginners.
- **Analyzing Published Datasets (01/2018 – Present)**

  Analyzing published data sets from "Kaggle" or from different sources and creating standalone applications with specific sets of analyzing questions.
  Link: https://github.com/kumawat55/bigdata
- **River Monitoring System (04/2017 – 04/2017)**

  A physical device for restoration of polluted river water, with sensor like temperature, sensor,Conductivity, pH sensor configured with Arduino.
- **Library module development (06/2016 – 07/2016)**
Java application development using Eclipse (IDE), MySQL database, MVC -model.

Declaration

I hereby solemnly declare that all the statements made above are true to the best of my knowledge and Belief and nothing has been concealed, suppressed or fabricated.

Place: Hyderabad   **Rohit Kumawat**