Darshan Shinde

Data Engineer

Profile:

Intend to build a career with leading corporate of IT Industries with committed and dedicated people, who will help me to explore myself fully and realize my potential and willing to work as a key player in challenging and creative environment. I don't have affinity towards any particular technology, but have got good exposure to Spark, Big Data, Cloud Services technology etc. that can give me a jump start in my career and help achieve objectives for the organization where I'll work.

Education:

- Mechanical Engineering, L R Tiwari college of Engineering, Mumbai (February 2017)
- Diploma in Mechanical Engineering, M H Saboo Siddik Polytechnic, Mumbai (June 2013)
- SSC, Balmohan Vidyamandir , Mumbai (June 2009)

Employment History:

Data Engineer,
BITWISE GLOBAL, PUNE (JULY 2019 – PRESENT)

- I have 3 years of experience in IT Industry.
- Experience of 2 years in relevant of Big Data Domain and Spark API
- Experience in data Transformation, Spark-SQL and transferring processed data from S3 to Database.
- Have experience writing data extraction logic in SQL.
- Developed PySpark programs and created the data frames and worked on data cleaning & transformations.
- Knowledge of Amazon Web Services such as [AWS S3,AWS Lambda ,AWS Glue, AWS Redshift, AWS RDS, IAM]
- Having Knowledge of creating table, partitioning, bucketing and loading in hive.

Project:

1. Data Analysis
ROLE: BIG DATA DEVELOPER

- Responsible for development of new lambda applications and glue jobs.
- Involved in monitoring and handling Glue jobs.
- Writing required logic to load as well as extract necessary data and writing queries that will run internally.
- Import data from S3 raw bucket and do the cleaning and masking and again move to S3 cleanse bucket by using glue.
- Extracting CSV or XML files according to customer's requirement at desired S3 location.
- Written logic to extract the data according to customer's requirement.

2. E-Commerce Data Lake:

ROLE: BIG DATA DEVELOPER

- Roles and Responsibility:
- Understand the upstream source nature and work with business cases.
- Data for each customers of client is organized by in S3 bucket in particular format so that to it would be easy for further process. Firstly data is stored into the S3 raw in parquet format and after that it is used for further processing in glue jobs.
- AWS Glue is use for validation, masking and various transformation of job. So AWS glue cleanse the data and store into S3 processed bucket for further transformation.
- Further Glue job will take data from S3 processed and do the various transformations.
- After transformation on processed data in AWS glue push the data into Redshift Data-warehouse.
- Basically AWS Redshift is used to store the processed data for analytical needs.

Personal Details:

Name: Darshan Haresh Shinde

Address: Mumbai, India

Contact no.: +91 7977012925

Email: darshanshinde4930@gmail.com

Key Skills:

- **Programming Language:**   Python
- **BigData Technology:**   Apache Hadoop, HIVE
- **Framework:**   Spark, PySpark, Spark-SQL
- **Dev.IDE:**   Pycharm, Jupyter notebook
- **Other Tools:**   WinSCP, Putty, SQL Workbench, MS Office
- **Operating System:**   Windows and Linux (Ubuntu)
- **Good knowledge of:**   MySQL, Oracle
- **Cloud Technologies:**   AWS (S3, IAM, Glue, RDS, EMR)

Declaration:

The above given information is true to the best of my knowledge.

Date: Place: