Prashant Dnyandeo Bharad

Phone no: +91-7972808186 / 9503125596

Email id : prashant.bharad1997@gmail.com

Address : A/P Chikhli, Dist. Buldana, Maharashtra

Data engineer experienced in working as a team across the different stages of the data pipeline. Well acquainted with ETL process, data pipeline, data management and processing systems. Strong desire to move forward, face new challenges and expand my skill set.

Career Summary

- Total **3.8** years of experience in different domains in IT industry which also includes experience in Hadoop Ecosystem, Spark API's, Cloud Services(AWS), Python, SQL, PaaS such as Data bricks.
- Hands on experience in various big data application phases like data processing, analytics.
- Have experience in building optimized big data pipeline using S3, PySpark and Databricks on AWS cloud.
- In-depth understanding of Spark Architecture including SparkCore, SparkSql, and Data frames.
- Have experience writing data extraction logic in pyspark.
- Experience in manipulating analyzing large datasets and finding pattern insight within unstructured and structured data using pyspark in databricks.
- Experience in working on file compression techniques as well as various file formats like CSV, JSON, ORC, Parquet, AVRO.
- In-depth knowledge of Hadoop architecture and its components like YARN, HDFS, Name Node, Data Node, Job Tracker, Application Master, Resource Manager, Task Tracker and Map Reduce.
- Hands on experience on the transformation glue job to move data from the S3 cleansed to the redshift using lambda and cloudwatch.
- Worked on several AWS cloud technologies such as EC2, S3, Glue, Lambda, Redshift, Cloudwatch and have a good understanding of AWS VPC, Athena, EMR, RDS, IAM.
- Very good understanding of Partitions, bucketing concepts in Hive and both Managed and External tables in Hive to optimize performance.
- Understanding of transferring data from RDBMS to HDFS and Hive tables using Sqoop.
- Hands on development on several tools like Putty, WinSCP etc.
- Well aware of various essential commands in linux.
- Experience on working on an Agile/Scrum Methodology.
- Strong problem solving & technical skills coupled with clear decision-making.

Technical Skill Set

- **Big Data Ecosystems**: Pyspark, Apache Hadoop, Apache Hive, Apache Spark, Sqoop, HDFS, YARN, Databricks on AWS Cloud, ETL, Data Pipelines
- **Cloud Technologies**: AWS (EC2, S3, Glue, Lambda, Redshift, Cloudwatch)
- **Language**: Python, SQL
- **Database** : MySQL, Oracle
- **Operating Systems**: Windows 10/11, Linux Ubuntu/CentOS
- **IDE** : PyCharm, Jupyter Notebook
- **Ticketing Tool**: JIRA

Educational Qualification

- Passout in 2015, Masters in Electronics engineering from Walchand College of Engineering, Sangli (An Autonomous Institute) with 7.0 CGPA.

Work Experience

Project #1

Organization: GMS SOLUTIONS PVT. LTD, BENGALURU

Project Name: Retail Store Data Analysis   **Duration: Feb 2022 to till date**
Role: Data Engineer

Description:
Retail stores generates various types of data such as events, sales, users, products at their POS (Point of sale) & get pushed into the big data pipeline for cleaning, processing and Business analysis. Basically, we have created an End-to-End data pipeline for Retail stores data using Amazon Kinesis, AWS S3, Pyspark, AWS Databricks and PowerBi.

Roles & Responsibilities:

- Daily standup call on start of the day.
- In every sprint we have assigned tasks, that we are completing as per timeline.
- In team, we collaborate and start building the logic.
- We have to filter data and perform multiple operations in order to get meaningful insights that will help the client in decision making and analysis using pyspark in Databricks.
- Optimize code using pyspark for better performance.
- Developing code and testing it on sample data.
- Have responsibility to Write the data extraction logic for the application.
- Weekly call with client for their requirements and further enhancement.

Environment: Pyspark, SQL, Python, Databricks on AWS, Amazon Kinesis, PowerBi

Project #2

Organization: GMS SOLUTIONS PVT. LTD, BENGALURU

Project Name:

Data Warehouse Design for an E-commerce Platform   **Duration: Aug 2020 to Dec 2021**
Role: Data Engineer

Description: The client uses on premises OMS (order management system). It leads to facing the slow performance of large analytical data, so to avoid this we have to prepare a data lake and data warehouse for unifying the data and to get meaningful information from it.

Roles and 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.
➢ Job monitoring.
➢ Work on the transformation glue job to move data from the s3 cleansed to the redshift dimensional and fact tables and job scheduling.
➢ Optimize code using pyspark for better performance.
➢ Creating a DDL script for creating tables in redshift.
➢ Work with the teammates to understand and support them in importing data from source to S3 by using AWS DMS, also work with the platform guy for the orchestration purpose.

Environment: Pyspark, SQL, Python, AWS Glue, AWS Lambda, Amazon Redshift, AWS Cloudwatch.

Personal Information

Name: Prashant Dnyandeo Bharad
Languages Known: English, Hindi, Marathi
Address : Mahavir Nagar, A/P Chikhli, Tq. Chikhli, Dist. Buldana, Maharashtra

Declaration:
I hereby declare that all the information mentioned above is true to the best of my knowledge.

Prashant Dnyandeo Bharad