Neha Pradeep Patil

Contact: +91 7776946489 | Email: nehapatil12110@gmail.com

OBJECTIVE

To work in an organization where I can contribute my knowledge and growth to enhance my working capabilities, professional skills, and business efficiencies to serve my organization in the best possible way with sheer determination and commitment to realize the organizational objectives.

PROFESSIONAL SUMMARY

- Overall 2.5 years of Experience in IT industry as data engineer including Hadoop Ecosystem, Pyspark, Cloud Services (AWS), SQL database such as ORACLE , NoSQL Databases such as HBase.
- Hands on experience with Hadoop Eco System including Data Ingestion and Data Processing (Pyspark, Hive, HDFS, HBase)
- Have hands on experience Spark Architecture, SparkSql.
- Designed Both Managed and External tables in hive for optimization performance
- Uploaded and processed terabytes of data from various structured sources into HDFS (AWS Cloud).
- Performed tasks in transferring data from RDBMS to HDFS and Hive tables using Spark API's.
- Having experience on using optimization techniques in Hive, Spark if required.
- Knowledge on Amazon Web Services such as RDS, EMR, S3, IAM.
- Analytical mind and ability to learn-unlearn-relearn concepts.

TECHNICAL SKILLS

- Big Data Ecosystems: Hadoop, SQL, Pyspark, Hive, HBase, Yarn.
- Cloud Technologies: AWS (EC2, RDS, S3, EMR, IAM).
- Language: Python.
- Database: MYSQL,Oracle.
- Project Management Tool: Agile & JIRA**.**

PROFESSIONAL EXPERIENCE

- Jun 2019 – Oct 2022 - NexGen Cloud Services

PROJECT DETAILS:

Project #1
Project Title: Financial Intelligence Unit.
Environment: Spark, HDFS, Hive, Phoenix,Hbase, MsSql.   **Role: Data Engineer**

Description: Financial intelligence unit is the gathering of information about the financial affairs of entities of interest, to understand their nature and capabilities, and predict their intentions. Generally the term applies in the context of law enforcement and related activities.

Roles and Responsibilities:
- Export and import batch data and delta data into HDFS, HBASE and Hive using PySpark.
- Engage with BA to understand the requirements clearly.
- Load, transform large sets of structure data.
- Involve in loading data from AWS RDS (PostgreSQL) to HBASE And Hive Environment i.e., HDFS using Apache Spark.
- Involved in creating tables, loading data into hive tables and writing Spark queries in Pycharm.
- Involved in processing data and do operations on it e.g., doing validation of data, deduplication and other operations as per the client requirement.
- Resolve JIRA tickets raised by Client.

Project #2

Project Title: Data Migration and Data Lake.

Environment: RDS,AWS S3,HDFS.

Description: The purpose of this project is to clean and validate the data of different category of products sent by upstream using Pyspark and also store the validated data into RDS . Apply extract logic on RDS to get some business insights and storing the output in AWS S3.

Roles and Responsibilities:
- Understand the client requirement for developing the project.
- Imported required data from source to target (RDS) location using Spark.
- Writing required logic to load as well as extract necessary data and writing queries that will   run internally.
- Used API to connect to fetch files to load the necessary data from source to target destination.
- Import data from S3 raw bucket and do the cleaning and masking and again move to S3   cleanse bucket.
- Extracting CSV or XML files according to customer's requirement at desired S3 location.
- Worked with various HDFS file formats like Parquet, RC, ORC.
- Written logic to extract the data according to customers requirement

EDUCATION

Bachelor's degree in civil Engineering from Ndmvp's KBTCOE Nashik, Savitribai Phule University, Maharashtra. (Aug'15-Jul'18)

Personal Details:

Name: Neha Pradeep Patil

Address: Plot no 24,Dekhu Road Gulmohar Colony No.2, Amalner Dist-Jalgaon Maharashtra.

Date of Birth: 01-02-1997
Gender :- Female
Languages known: English, Hindi, Marathi.