AJAY S WADILE
Data Engineer

✉ **ajswadile123@gmail.com** 📱 **+91 7219767565, +91 8421506712**

📍 **Dondaicha,Dhule,Maharashtra,India** in **linkedin.com/in/ajay-wadile-433703206**

SUMMARY

- Having **2.5** years of experience in design, implement and support different big data applications by using Apache Spark, Hadoop and AWS.
- Hands on experience with Big Data core components and Eco System including Data Ingestion and Data Processing (PySpark, Hive, Sqoop, Hadoop).
- Involve in Loading data from PostgreSQL to HBase and from HBase to Hive using PySpark job.
- Developed PySpark applications using Spark SQL, DataFrame API with optimization techniques if required.
- knowledge of Power BI visualization tool.
- Have experience writing data extraction logic in SQL.
- Capable of processing large sets of structured, unstructured data.
- Knowledge of AWS (Amazon Web Services) (S3, EMR, RDS, EC2, IAM, Athena)
- Experience in using Hadoop Distribution package like Cloudera.
- Experience in transferring data from RDBMS to HDFS, Hive and vice versa using Sqoop.
- Experience in creating tables, Partitioning, Bucketing, Loading in Hive.
- Gather and process Raw data at scale and analyze processed data.
- 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.

Technical skills:

Hadoop Ecosystem	Hadoop, Hive, Sqoop, PySpark
Programming Language	Python, SQL
Cloud Services	S3, EMR, RDS, EC2, IAM, Athena
Operating Systems	Linux, Windows
Relational Databases	MySQL, Oracle 19C, PostgreSQL

Visualization Tool	Power BI
Tools/IDLE	Putty, SQL Workbench, SQL Developer, Jupyter Notebook, PyCharm

Professional experience:

Organization: Zielotech InfoSolutions India llp.   **DOJ: 01 Jun 2020**
Designation: Big Data Developer

1 . Credit Card Data Analysis: Duration: Oct 2021 – Present

Description:

Taken the Historical data from Clients Database Ingest into the HBase using PySpark. Same for the Delta data which is came from Clients S3 location All the required data then move into Hive staging table after that as per the requirement the moving those data in Hive External table at the end the analysis required data move into Clients S3 location.

Responsibilities:

- Import Data into the HBase from Relational Database (PostgreSQL) for Historical data and HDFS for delta data using Spark.
- Write Phoenix DDL to Create HBase Table.
- Write Hive DDL to Create Hive Table for Optimize Query Performance.
- Implement Data Validation.
- Involved in import data from HBase into Hive Managed Tables using Spark which includes Incremental Load and some transformations.
- Involved in import data from Hive Managed Table into Hive External Tables which includes Queries using Spark.
- Designed Both Managed and External Hive Tables and Defined Static and Dynamic partitions as per requirement for optimized performance.
- Worked with File Format like csv.
- Written Hive Queries for Data Analysis to meet the Business requirements.
- Monitor PySpark job if failed.
- Perform deduplication on data to satisfy business requirement.

Environment: Hadoop, Hive, HBase, PySpark, AWS S3, EMR.

2. Data Migration :   Nov-2020 to July-2021

Description:

Data Migration involves in loading data from S3 Raw bucket and Perform transformations and do processing as per Business logic using PySpark Job and provide real time data to client in S3 clean Bucket.

Responsibilities:
- Extracting data from S3 Raw Bucket using PySpark job.
- Writing PySpark job to clean data.
- Apply transformation on data as per business logic using PySpark job.
- Store clean data into another S3 Clean Bucket.
- Involved in working on Spark SQL Code as an alternative approach for Faster data Processing and Better Performance.
- Monitor PySpark job if failed.

Environment: PySpark, AWS S3, EMR.

Education:
- Bachelor of Engineering from North Maharashtra University, Jalgaon in year of 2020 with percentages of 60.19%.
- Completed HSC from Maharashtra State Board ,Nashik in Feb 2016 with percentages of 63.85%.
- Completed SSC from Maharashtra State Board, Nashik in March 2014 with percentages of 81.60%.

Personal Details:

Full Name	: Ajay Suresh Wadile
Date of Birth	: 02th August 1998
Languages Known	: English, Hindi, Marathi.