

AJAY RAUT

Flat no:103, Shriram Park

Bhalekar Wasti, Warje Pune-411052

+91 9762550540 | rautajay0701@gmail.com

BIG DATA ENGINEER

A Professional with 3 years of extensive experience in Big Data Technologies like Spark ,Hadoop,Hive,Sqoop and ETL Processing as well.

Professional Summary:

IT professional with 3 years of experience focused on Big Data development,Talend, ETL, Cloud Development.

- Working as a Big Data developer at Ellicium Ltd, a company focused on providing consulting services in Information Management.
- Good working experience in Spark, Hive, Hadoop, Sqoop, Talend.
- Good knowledge of Spark Architecture, Spark Core, SQL.
- Worked extensively on databases like Oracle and MySQL.
- Working experience in ETL. Created source to target mappings, lookup, filter, joins, switch case, etc.
- Extensive experience in extraction, transformation and loading of data directly from different heterogeneous source systems like flat files, Oracle.
- Proven ability to work independently with business users. Worked with Business users and the IT staff for US based clients.

Skills Profile:

Database:	MySQL, Oracle
ETL Tools	Talend, Sqoop
Big Data	Spark, Hive, Hadoop, Sqoop, HDFS, PySpark, Cloudera's Distribution of Hadoop
Programming Languages	Java, Python
Automation Tool	UIPath
Cloud	AWS

PROFESSIONAL EXPERIENCE

Organization: Ellicium Solutions Pvt Ltd	Client: US Cellular
Duration: Oct 2019 - May 2020	Role: Big Data and Spark Developer
Project : USCC Datahub	
Technology Stack: Spark, AWS, Windows, Talend 7.0, SQL Developer, Hive, Oracle	

U.S. Cellular, is a mobile network operator which owns and operates the fifth-largest wireless telecommunications network in the United States, serving 5 million customers in 426 markets in 23 U.S. states. Ellicium was responsible for creating a Data Lake on AWS and migrating USCC's data from Oracle to Hive and validating the data in target system. My responsibilities in this project include developing Talend jobs for migrating data from Oracle to Hive environment based on client's requirement.

Roles and Responsibilities:

- Data Profiling of the source data systems
- Understand the requirements from the client through BRD.
- Creating Spark applications to get data from source systems, apply business rules and transform them.
- Loading large datasets into Hive tables using Spark Application.
- Writing spark applications to ingest and perform complex operations on data. Logging errors encountered in Spark, analysing errors and fixing them
- Also worked on Talend including Talend Administration Centre, Talend Studio, Standard jobs and Big Data jobs
- Using Talend, create joblets used to pick up dynamic parameters from configuration tables. These included all the custom values that would have to be changed between environments.
- Writing complex SQL Queries S2T mappings for Each entity. Writing DDL to create hive table for optimized performance Performing Unit testing and integration testing.
- Data Validation after Data Ingestion.
- Creating High Level Design Document for each entity.
- Query optimization and ETL fine-tuning.

Organization: Ellicium Solutions Pvt Ltd	Client: KBR
Duration: Aug 2019 - Oct 2019	Role: Big Data and Spark Developer
Project : Impact of NOTAM on flight delay	
Technology Stack: Spark, Java, Eclipse, Hive,Hadoop	

KBR partners with government and industry clients to and comprehensive solutions for a wide range of industries and companies all over the world. With a full portfolio of services, proprietary technologies and expertise ranging from, to mission critical services for aerospace and defense, to large scale energy and chemicals sites and to meeting the physical and digital infrastructure needs of governments and organizations across the globe.

Roles and Responsibilities:

- Data Profiling and requirement analysis.
- Preparing data from large datasets to predict flight delays due to NOTAM.
- Loading large datasets into hive tables using Spark Application.
- Writing Spark applications using Java to ingest and perform complex operations on data.
- Predicting impact of NOTAM on flight delay.

Organization: EC Mobility Pvt. Ltd.	Client: CMORE
Duration: Sept 2017 - Aug 2019	Role: Data Engineer
Project : CMORE's Data Ingestion	
Technology Stack: Spark, Hadoop, Sqoop, Oracle, Hive, SQL	

CMORE has Lines of Businesses like Sales, Risks, Payment, Revenues etc. across the world divided country wise over multiple source servers. It has a large amount of data in various data sources like Relational Databases, Flat Files. In our project, we are ingesting and processing data and providing information to them for analytic purposes.

Roles and Responsibilities:

- Imported required tables from RDBMS (Oracle, SQLServer) to HDFS using Sqoop. Writing DDL to create hive table for optimized performance
- Gathering required information for Data Ingestion.
- Written Hive Queries for Data Analysis to meet the business requirements Data Validation after Data Ingestion.
- Involved in creating Hive tables, loading with data, and writing hive queries.
- Used Hive to analyse the partitioned and bucketed data and compute various metrics for reporting.
- Used Hive optimization techniques during joins and best practices in writing hive scripts using HiveQL.
- Importing data into HDFS and Hive using Sqoop.
- Designed both Managed and External Hive Tables and Defined static and dynamic partitions as per requirement for optimized performance on production data sets.

EDUCATION

- **BE** in Computer Science and Engineering(64.88%) from MSS's CET Jalna in 2017.
- **Diploma** in Computer Engg.(67.03%) from Govt. Polytechnic Jalna in 2014.
- **SSC** (78.55%) from Jivanrao Pare Vidhyalaya Jalna in 2011.

PERSONAL DETAILS

Full Name	: Ajay Gorakhnath Raut
Date of Birth	: 7th Jan 1996
Marital Status	: Single
Nationality	: Indian
Languages Known	: English, Hindi, Marathi
Permanent Address	: H.No 860, 60 ft road, Chandanzira Jalna - 431203