Name : Yatish Kumar Sharma

Module lead, Mindtree Technologies

Noida

Email-sandhya.yatish@gmail.com

Contact– 9319331358

- Around 5 years and 8 month of IT experience, including 3 years and 7 month of experience in dealing with Apache Hadoop Ecosystem like HDFS, Map Reduce, Hive 2.1, Sqoop 1.99.7, Spark 2.1.0, Python 3.7 and Big Data Analytics and 2 years of experience in Development Core Java, JavaScript,XML JQuery and Wordpress.
- Excellent understanding of Hadoop Architecture and underlying Hadoop framework including Storage Management.
- Experience to write the optimize code in pyspark for structure data processing.
- Involved in creating POCs to ingest and process structure data using Spark 2.1.0 and HDFS.
- Experience in analyzing data using HiveQL 2.1.
- Experience in working with Spark Data Frames and RDD.
- Experience in importing and exporting data using Sqoop1.99.7 from HDFS to RDBMS and vice-versa.
- Certificate for "spark developer" and "Accessing Haddop data using hive" from IBM (CONGITIVE CLASS.ai)
- Expert in writing complicated SQL Queries and database analysis for good performance.
- Experience in complete project life cycle (design, development, testing and implementation) of Client Server and Web applications.
- Expert in Building, Deploying and Maintaining the Applications.
- Experienced in preparing and executing Unit Test Plan and Unit Test Cases after software development.
- Excellent analytical, Interpersonal and Communication skills, fast learner, hardworking and good team player.

Work Experience
Project01 - CEM - Nokia
Company - Mindtree Technologies
March 2020 -June 2020

Customer experience management product is developed by nokia and it is used for customer experience and quality experience for telecom users and it has data from different sources and hive and spark as technologies.
Responsibilities:

- Involve in requirement gathering and client presentation
- wrote optimized hive query
- performed join on stream data
- Worked on CQI (Customer quality indexing) for MTN

- Raised Bug in JIRA if found any.

  Environment:

  HDFS, CDH 6.1, MY SQL , Scrum, JIRA , Hive

Project 2- H&M

Company-HCL Technologies

Feb 2018 to March 2020

This project is related to one of leading name in e-commerce domain. H&M generate millions data on daily basis from e-commerce website and its physical store. to optimize the produced data, we usespark, python ,sqoop, spark SQL and power BI . processed results are displayed on sales portal or shared as report to respective stake holder

Responsibilities:

- Bulk importing of data from data source to HDFS and transform data in flexible ways by using sqoop .
- Optimize the existing or newly created Sqoop query
- Create the Spark session and Read the data from the HDFS By Spark API.
- Used Spark   2.1 over Cloudera Hadoop YARN 2.5.2 for creating Data Frame
- Implemented Batch processing of data sources using   Apache Spark 2.1 Cleaning and validating
- Create spark sql UDF according and apply on data frame according to requirement
- Write Sql query according to requirement.
- Save the result in HDFS.
- Spark for improving the performance and optimization of the existing algorithms in Hadoop using Spark Session Data Frame, Pair RDD's, Spark, vectorization , file format.

  Environment:
  HDFS, , Sqoop, CDH 6.1, MY SQL , Scrum, JIRA , Spark2.1

Education and medical payment system
OCT 2016 to Jan 2018
Big Data Engineer , Company–HCL Technologies

Project: Medical & education payment system, this program is run by federal government. through this system drug and device companies make the payment to physician and hospital and education institute. Payment regarding the travel, research, gifts , speaking fees and meal.

Responsibilities:
- Collaborated with the Internal/Client BA's in understanding the requirement and architect a data flow system.
- Used sqoop to Inject the data from My SQL to HDFS.
- Wrote the hive query according to the requirement.

- Optimize Sqoop command to faster the injection.
- Used partitioning, bucketing and Vectorization for Hive query optimization.
- Raised Bug in JIRA if found any.
- Removal of duplicate value form Hive table .

Environment:
>   CDH 5.8.3, HDFS 2.7.3, Hive , Sqoop

Project 4- ICICI Lombard General Insurance Company – HCL Technologies
Nov 2014 to Sep 2016

Reserve Valuation project is mainly for calculating reserve amounts for Incurred but Not Approved (IBNA) and Long Term Disability (LTDV) products for future needs and analysis purpose. Claims, premium, reserve feeds come from different group mainframe machines. These feeds will be processed and loaded into according databases. This data goes through the various phases, process steps and process types. Business partners define rules for data manipulation and put some adjustments for data corrections. Reserve calculations and allocation will be done based on reserve category and experience levels. After all processes are done, output feeds will be created and loaded into one single group reserve database (GRDB).

Responsibilities:
- from XML and a business process layer to allow rapid development Worked with customers togather requirements and develop use cases, project and test and deployment of web services Designed a flexible and highly dynamic framework that incorporates robust
- aggregating back-end service invocations.
- Developed the UI pages using HTML, DHTML, JavaScript, AJAX, JQUERY, JSP and tag libraries etc.
- Produced as well as consumed the WSDL based web services using WSDL, SOAP, JAX-WS, AXIS, Apache CXF, JAXB etc.
- Developed web services using Restful technology to support JSON and XML using JAX- RS.
- Worked as a Data Modeler for the team.
- Built Custom business validations using validation framework.
- Created new connections through application coding for better access to DB2 database and involved in writing SQL & PL/SQL – Stored procedures, functions, sequences, triggers, cursors, object types etc.
- Involved in writing developer test cases.

Environment:
Java 1.5, J2EE, HTML, DHTML, JavaScript, CSS, AJAX, JQUERY, Servlets, JSP, SQL

Additional Information
TECHNICAL SKILLS

Scripting Language\ Operating Systems -

Language	Version
Python	3.7
Window	10

Git Shell	2.12.0

Web Development\ Hadoop Ecosystem-

Hadoop 2.7.x, Spark 2.1.0, MapReduce , Apache 2.1.1, Sqoop 1.99.7, Flume 1.7.0, Yarn

IDE Application\ Database:-

MySql 5.x, Notepad++ 7.3.2,Pycharm, Jupyter

Other Tool:- Git 2.12.0, JIRA 6.4

Reporting tool:- Power BI

Certification:-

1. "Spark developer "certificate from IBM (CONGITIVE CLASS.ai)

2. "Accessing Hadoop data Using Hive" certificate from IBM (CONGITIVE CLASS.ai).

Academic Credentials:-

Examination	Discipline	School/College	Board/University	Year	Aggregate
B.TECH	ECE	Govt Eng. College Jhalawar	RTU ,Kota	2014	61 %
12	PCM	Maa Bharti S.S. School	RBSE	2010	64.38%
10th	ALL	S.S. School ,CKB	RBSE	2008	77.83%

Personal Dossier

Date Of Birth: 20-07 -1993

Father's Name: Mr. Rambabu Sharma

Gender: Male

Languages Known: Hindi , English

I hereby declare that all the details in the application are true, complete and correct to the best of my knowledge and belief.

Yatish KUMAR SHARMA