**Work Experience:**

* Total 4 Year and 5 months experience in IT Industry worked with various domain Technology, Insurance and Banking.
* 3 Years and 6 months working as a Programmer Analyst (Feb 2014 – 30 Aug 2017) in Cognizant Technology Solutions, Chennai.
* 11 Months in Capgemini as a Associate Consultant (4th Sep 2017 – Till Date) , Chennai
* Expertise in Big Data Ecosystem and Spark and its components having around 3 years of experience.
* Expertise in Data Warehousing and Business Intelligence and Analytical Skills
* Excellent working experience of Pentaho ETL, Pentaho BI Suite, Oracle, DB2, My SQL, R Programming, Python for Data Analysis.

**Technical Skills:**

|  |  |
| --- | --- |
| **Scripting Language** | ANSI SQL, UNIX Shell Scripting, Scala , |
| **Database** | DB2 , Oracle , TeraData, My SQL |
| **Big Data (Hadoop)** | Hadoop 2.7.3, Hive 1.2, Map Reduce, Pig, Sqoop, Spark 2.0 , Hortonworks, Spark-SQL, Spark Data Frames |
| **ETL Tool** | Pentaho Data Integration 5.4.0 (PDI) , Informatica |
| **Reporting Tool** | Pentaho Report Designer |
| **Analytical Skills** | Python for Data Analysis , R Programming, Basics of ML , Data Structure and Algorithms |

**Professional Summary:**

* Expert in developing Data Analytics/Business Intelligence solutions using Pentaho, Hadoop Ecosystem- Sqoop, Hive, PIG, Flume, SQL,SPARK-Spark SQL, Spark Streaming.
* Expert in working on Hadoop Ecosystem and Data Warehousing ETL.
* Excellent knowledge of Data-warehousing concepts, ETL and Data Quality.
* Expertise in using databases Oracle, MySQL to bring data In/Out for developing Analysis and Mining Solutions.
* Knowledge of the Big Data, Hadoop Concepts, Map-reduce and HDFS, Data Warehousing Concepts.
* Working experience in Development and Design Roles.
* As part of assignments, involved in Requirement Analysis, Application development and Maintenance, Functional Studies, Quality Reviews and testing.
* Knowledge of Data Analysis using and Dealing with different types of Data in Big Data.
* Industry experience in Logistics, Insurance with Functional knowledge in daily operation of these industries.

**Certifications:**

* **Cognizant Certified Professional (CCP) : Working with Hadoop v2.0**

Big data & Hadoop basics, May 2016

* **Cognizant Certified Professional (CCP) : Data Warehousing**

Basic of Data Warehousing, Oct 2014

* **Cognizant Certified Professional (CCP) : Pentaho**

Pentaho Practitioner, Jan 2016

**Project Details:**

* *Base Ingestion/Extraction Framework*

Scope of this project is to perform data ingestion from RDBMS sources to HDFS. File and data ingestion from RDBMS is accomplished using Sqoop and ingested in HDFS. Using Big Data the emphasis is to enhance the output and make it available at the earliest to the analytics teams for their requirements.

Spark Data Frames are being designed to increase the batch jobs performance.

Roles and Responsibilities:

*Big Data Components:* Hadoop 2.7.3, Sqoop, Hive 1.2, Map Reduce, UNIX Shell Scripting, Spark-SQL, RDBMS, Scala.

* Perform data ingestion using Sqoop and Spark (Using CSV and JSON files).
* Sqoop the RDBMS data.
* Creating Data Frames using input source (CSV and JSON files) and loading it into HDFS (Hive tables) using Spark SQL.
* Create mapping in HQL to generate reports and files to feed downstream applications.
* Working on RDD’s and Data Frames.
* Validate and match reports and exports generated in Hadoop & Spark in sync with TD/Oracle/SQL Server exports
* Deployment and support all post production activities
* Development includes creating tables, partitioning, bucketing of table in Hive.
* Experience with Queries Performance Tuning.
* Provide solution/workaround for all the issues logged in Production after go live.
* Manage and review Hadoop log files.
* *Commercial Data Lake*

Migration of current Application’s from various sources to HDFS and creating A **Commercial Data Lake** for data analysis and report generation, Landing Zone- Where the data get just sqooped from various Sources either Full Load or Incremental Load, Staging Zone- Zone in which the change Data Capture or append Operations takes place, Standardization Zone- Applying the various transformations as per business requirement performs various joins on various tables and creates view and tables as per frequently access data and various analysis perspectives

Curated Zone- Final Zone where final data is stored and available for various analysis purposes.

Roles and Responsibilities:

*Big Data Components:* Hadoop Ecosystem- Sqoop, Hive, Pig, Map Reduce, UNIX Shell Scripting, R

* Implementation of SCD transformations into Hadoop and HIVE.
* Migration of RDBMS table logics to hive.
* Developed queries required to interact with HIVE tables.
* Implementation of Data Warehousing Logic in Hadoop with Sqoop and HIVE
* Analysis of Data Using R and Creating Graphs.
* Taken regular sessions with ELT team members to enhance technical and business knowledge on ETL & Hadoop.
* Successfully delivered a project working from documentation to proposed output.
* *Insight Compass (Financial Transformation Wave I)*

To carter the ETL and BI (Reporting) requirements of the clients and deployment of the same in various servers and maintenance of the same (servers).

Roles and Responsibilities:

*Component Used:* Pentaho BI Suite (Pentaho Data Integration, Pentaho Report Designer)

* Understanding the ETL Specification Documents and CR’s for transformation requirements.
* Integration of the reports with Job codes and other JavaScript files.
* Worked in Automations of ETL codes.
* Created ETL code using the transformations based on the business requirement.
* Have an excellent development experience in Business Intelligence using Pentaho Report Designer.
* Developing SQL queries both for ETL and Reporting requirements.
* Making Reports at various hierarchal levels.
* Implementing Reporting concepts as per the requirement.
* Involved in UAT support as part of delivery and post-production activities.

**Academic Qualifications:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Degree/Education** | **Institute/College/School** | **Institute/College/School** | **Passing Year** | **Marks Obtained** |
| B.Tech (Computer Science) | Guru Nanak Institute of Engineering and Technology Nagpur | RTM Nagpur University, Maharashtra | 2013 | 67% |
| Higher Secondary | Kendriya Vidyalaya VSN Nagpur Maharashtra | CBSE, New Delhi | 2009 | 71% |
| Secondary | Kendriya Vidyalaya Hasimara West Bengal | CBSE, New Delhi | 2007 | 82% |

**Date of Birth:** 3rdSeptember 1991

**Declaration:** I hereby certify that all the information provided above is true to the best of my knowledge.

(Gaurav Sharma)