Professional Summary

13+ years of total experience in the field of Information Technology

Having 5 years of experience in Big Data and Analytics, Hadoop,PySpark, Sqoop, Hive, Oozie, HBase and Kafka.

Good experience in Hive partitioning, bucketing and perform different types of joins on Hive tables and implementing Hive serdes.

Good knowledge on the various file formats and compression techniques

Having experience on Spark streaming and SQL.

Having good knowledge on AWS Lambda Services.

Experienced in Python Programming,Pandas,Numpy and Scikit-learn libraries

Involved in all phases of the SDLC including requirement analysis, design, coding, testing, implementation and maintenance.Working on projects based on Agile methodology

Having good knowledge on Machine Learning and EDA

Professional Tenure:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sr No. | Name of organization | Joining Date | Reliving Date | Last designation |
| 1 | Cognizant | June 2016 | - | Senior Associate |
| 2 | Wipro Technologies | Oct 2010 | June 2016 | Technical Lead |
| 3 | Steria(Xansa) | July 2007 | Oct 2010 | Software Engineer |

Technical Skills

* **Distribution System**: Couldera
* **Programming languages:** Python, R Programming,Cobol,ANSI SQL
* **Scripting**:Shell Scripting and PySpark,Python(Pandas,Numpy & Scikit learn)
* **Data Scheduling**: Oozie and Zena
* **Data Extraction**:Spark, Hive, HBase, Kafka and Sqoop
* **Test Framework**: PyUnit
* **Cloud :** AWS
* **Tools**:Putty, FileZilla, Github and SQL Developer
* **DataBase** : Microsoft SQL Server, DB2, MySQL
* I**DE** PyCharm,Jupyter Notebook and IBM Data Studio

Certifications:

* Full Stack Data Science Certified from Jigsaw Academy(JNFDS1001333IL)
* AWS –Solution Architect Certified Associate(Validation Number G26ZBF9CJFR1QZ9G)
* DP-200 Implementing an Azure Data Solution(Microsoft Certifcation id 990541924)
* CCA175-Hadoop and Spark Developer Certified

Soft Skills

Pro-active and Self confident, People Mangement , Leadership, Excellent Communication skills,Assertive, Quick learner,Zeal to learn new technologies/Process,Team Player.

Professional work experience

Jan 2019 – Till Date **GPD Data Ingestion and Claims Integration**

Client **Health Care Service Corporation, USA**

Environment **Big Data- Spark 2.1, Python, Sqoop, Hive, HBase, Kafka and Unix Shell Scripting**

* Currently guiding a Data Engineering team of 13 and working as an individual developer for creating Data Pipelines.
* Created a tool for automating testing for Big Data ETL, which compares the Source vs Target based on mapping metadata, resulted in saving 60% manual effort.
* Created automation tools for Data Ingestion Framework, which reduced the development effort by 50%.
* Involved in code review and bug fixing for improving the performance.
* Designed and developed Kafka to Kafka processing (Ingestion layer to the Integration layer)
* Performed performance tuning and achieved scalability.
* Analysis and Requirement Gathering from the Business Analyst.
* Working on JIRA items assigned and attending daily stand up, grooming and retrospective calls.
* Coordinating and helping the team technically to meet the deadlines.
* Ability to work individually as well as in the Group.

June 2016-Dec 2018 **Enrollment,** **Provider and Authorization and Referal system**

Client **Wisconsin Physician Services,USA**

Environment **Big Data- Spark 1.6, Sqoop, Hive,Python and Unix Shell Scripting**

* Developed Spark programs for the merging of source data into desired output format based on business logic using Spark SQL Lib and Hive HQL queries.
* Worked on data migration from RDBMS to HDFS using Sqoop.
* Data pipeline creation for the ETL processes using Oozie workflows.
* Processing the data in HDFS, analyzing the data, and involved in writing transformation to process data using Spark SQL.
* Improved the performance and optimization of the query by using Partitioning and bucketing in Hive.
* Preparing the Unit / System test plans and executing the Unit / System testing.
* Worked with using different kind of compression techniques to save data.
* Working on JIRA items assigned and attending daily stand up calls.
* Coordinating and helping the team technically to meet the deadlines.

Oct 2010- June 2016 **Credit AMS Development**

Client **Standard Bank of South Africa, Johannesburg**

Environment **Python,Unix Scripting,COBOL, JCL, VSAM, CICS, MQ, DB2, expeditor,** **JAVA Web services, Web sphere MQ,IBM Data Studio**

* Involves in requirements discussion, architectural design meeting in order to evaluate the impact and effort estimation and translate into design documentation.
* Responsible to perform impact analysis on the projects by gathering requirements.
* Prepare detailed design in accordance with functional specification requirements.
* Involves in all the phases of development projects - Functional discussion, Tech Spec preparation, Design document preparation, Coding, Unit testing, defect fix and post production support activities.
* To identify all credit applications that is impacted CIF decommission.
* Delivers all change work packages on time as planned in the project activity schedule.
* Responsible to fix daily batch job abends in production for the development projects involved and to ensure client business will not effect on the next day
* Worked as onsite coordinator for the credit development project in SBSA and assign the work to Wipro Offshore resources.
* Responsible to review the code changes, Test cases & results delivered by offshore.

July 2007- Oct 2010 **Tesco Payroll system for UK/ROI**

Client **Tesco Corporation, U.K**

Environment **COBOL, JCL, VSAM,DB2**

* I was involved in fixing issues in day-to-day job abends and code fixes being raised by payroll users, and supporting the system in 24 \* 7 environments.

Education Background

* M.C.A –Computer Applications from Govt institute Of Technology–Vellore-Anna University