Santosh Walde

E mail ID: santosh.walde23@gmail.com

Contact Number: +91-9325308557

Professional Summary

- Having total **5.9 Years** of IT and Infrastructure experience as **"Software Developer"** complete hands on **Hadoop, Apache Spark** using **scala** Technologies.

- Having good Hands on experience of **Distributed data processing systems** such as Hadoop.

- Having good knowledge of **Cassandra and spark** for data processing and analysis.

- Having good knowledge of Hadoop **HDFS** using Spark and hive.

- Having good knowledge of **Sqoop** and **Flume** for importing/export data into HDFS.

- Having hands on experience of **Spark Framework** (**operation on RDD, Data Frame, Dataset, Spark SQL) using Scala.**

- Having good Hands on experience of **Unix and Shell scripting.**

- Flexibility and adaptability in regards with new Technologies and Environments.

- Fundamentals of Spark using the Dataframe API

- Understanding partitioning of data

- Responsible for developing, enhancing & maintaining applications with Scala and Java.

- Analysing and performance tuning Spark queries e.g. looking at the DAG

- Creating Scala/Spark jobs for data transformation and aggregation.

- Knowledge of Hadoop and its ecosystem of technologies especially Hive

- Contributes to quality assurance by writing unit and functional tests.

- Ensures development happens for all Software Components in accordance with Detailed Software Requirements specification, the functional design and the technical design document.

- Basic knowledge of UNIX and Shell Scripting

- Ability to analyse databases directly using query language tools such as SQL

  ➢ **Non-Technical**
- Use of JIRA / Confluence

- Provides Level 3 support

- Contributes to problem and root cause analysis.

- Good with analytical skills.

- have good understanding of architecture

- having to share information, transfer knowledge and expertise to team members.

- having design and write code in accordance with provided business requirements

- have good to work in a fast paced environment with competing and alternating priorities with constant focus on delivery on sprint.

3. Figmd India Pvt Ltd

(US Healthcare org)   (**Nov_2019 till present**)

Software Developer

Description	• The project is all about the processing clinical and healthcare data of US different registries to generate score of provider and practices based on US health CRA act & law.
Role	Software Developer
Responsibility	• Developed module to Reading data from CDR server and proceed using business logic to to generate data in standards into Cassandra. • Developed module Reading data from Cassandra to process as per business logic to meat goal of client requirement & dumped into postgres and cassandra.to that scorecard will be visible on dashboard. • PRO data processing and sharing analysis to QA delivery team. • Loading intermediate data shared by delivery team to Cassandra for visual to dashboard. • Deployment of generated jar and releasing it every month so that operation team use it for processing it • Data cleaning and filtering and parsing using Spark. • Created DAG using cloud composer to schedule job for automation
Team Size	4
Environment	GCP, Cassandra, Hive, Apache Spark with scala.

2. Globant India Pvt Ltd   **(May_2019 to Sep_2019)**
Software Engineer

• Worked on POC Spark Streaming & Apache Kafka
• Worked on POC using flume where data transferred to HDFS
• Worked on POC pipeline set up
• Dummy data proceed.

1. Wipro Technologies   **(June_2015 to May_2019)**

Project Engineer

Description	The project is all about the Telecom customer and network log processing where the particular section is for helping out the Support team in decision making for customer offers and plan for them and trace-out congestion details inside network environments. We have couple of use cases where we have trace-out existing plan and uses for each customer depending on business logic .
Role	Engineer
Team Size	11

Responsibility	• Requirement Analysis: Analyzing the requirements and communicating with BA team for better understanding of the requirements. • Process to logs file using spark for required outcomes • Transform the business logic to implement SPARK • Design schema & Load data in Hive and MySql • Migrating manual process into automation using Spark • Develop a Spark application with Scala. • Read data into RDD converted to dataframe and dataset to perform further processing • Develop and schedule flume job to transfer data from the local file system to HDFS • Developed ETL pipeline with the help of Apache Sqoop, Apache Hive and Apache Spark.
Environment	HDFS, Hive, flume, sqoop, Apache Spark .

1. **CRM Module**   6 months

Description	The project is all about the customer issue, field Engineer and Service Desk who are handle status of customer problems, process goes through all of the phases once customer raised ticket and individually updated their status for further action to resolve customer problem as earlier.
Role	Developer
Team Size	09
Responsibility	● Requirement gathering : gathering the requirements and communicating with team for better understanding of the requirements and workflow to make it easy way. ● Involved in supporting, analysis, design, coding and testing as per the client requirement. ● Used DAO design pattern in the application design and implemented the Service and Data Layer. ● Implemented Action classes, Action forms. ● Implemented the DAO classed using Hibernate with MySQL Database to handle the data layer.
Environment	JAVA/J2EE, JSP, Hibernate, Spring, Apache tomcat, MySQL, Eclipse.

Technical Profile

Languages **:** Strong hands on C,C++,JAVA,SCALA,PYTHON.

J2EE Technologies: Core Java, Servlets, JSP, JDBC, Hibernate, spring

Web Technologies : HTML, Java Script And CSS.

Web Server **:** Apache Tomcat.

Operating Systems: Windows,UNIX.

Databases **:** MySQL, Oracle, NoSQL(Cassandra)

Framework **:** Hadoop, Apache Spark.

Additional Skills

Beginner: SPARK MLlib & Deep Dive into Spark MLlib.

Educational Qualification

Course	Year of passing	Board/University	Percentage
MCA	2015	Sinhgad Institute(University Of Pune)	66.35%
BCS	2011	Dr. BAMU Aurangabad (Maharashtra)	67.53%
HSC(Sci)	2008	Aurangabad Board (Maharashtra)	71.83%
SSC	2006	Aurangabad Board (Maharashtra)	67.06%

Personal Details

Date of Birth **:** 01/Jul/1990

Marital Status **:** Married

Languages Known: English, Hindi and Marathi.

Current Address : Flat no-402, Tatyaba Complex, Manaji Nagar, Narhe, Pune-411041.