**Balakrishna** Email : sbrbigdata@gmail.com

Phone : +91 9059008245

**PROFESSIONAL SUMMARY :**

* Around 5 years of professional experience with extensive usage of **Big Data(Hadoop)** technologies(around 3 years) and **Java** (around 2 years).
* Good understanding of developing/migrating applications in Big-Data ecosystem and experience in development of applications using Java technologies.
* **Production implementation** experience on multiple projects using **Map Reduce**, **PIG, HIVE, HBase,** Apache **Sqoop, Spark, Oozie.**
* Good Knowledge on **HDFS**, **YARN, Scala, Teradata.**
* Good understanding of developing/migrating **ETL** end to end in to Big-Data ecosystem.
* Experience in developing applications using Spring and Hibernate.
* Experience in using IDE such as Eclipse.
* Hands on experience in developing web applications using Weblogic and Tomcat.
* Sound knowledge of the entire SDLC.
* Experienced in various technologies, domains, project implementations and support projects.
* Having good knowledge in Banking, Media, Telecom& Manufacturing domains.
* Fast paced learner and quick solution provider who can easily work under tight timelines and always deliver to the delight of the client.
* Skillful with dedication to work as independently as well as a team member to solve the end user issues with my analytical and technical knowledge.

**EDUCATION:**

Bachelor of Technology, Information Technology, JNTU-Hyderabad, India, July 2011.

**TECHNICAL EXPERTISE :**

Hadoop Ecosystem : Map Reduce, Pig, Hive, Spark, Oozie, Flume, Zookeeper

Sqoop, Apache Storm, Kafka.

Operating Systems : UNIX, Windows

File Systems : HDFS, UNIX, Windows

DBMS :HBase (No SQL), MySQL, Oracle

Languages : UNIX Scripting,Java,Scala

Tools : SVN, Maven, Eclipse

App & Web Servers : Weblogic, Tomcat

Frameworks : Struts, Spring and Hibernate

**WORK EXPERIENCE :**

**Agile Solutions Pvt Ltd (www.agileinfosolutions.com), Bangalore *October 2012 – Apr 2017***

**Hadoop cum Java Developer**

**Common Format: December 2016-Mar 2017**

**Roles and Responsibilities:**

* Implementing Performance, stabilizing and optmizing existing projects using **Spark** ecosystem.
* Utilization of Spark framework to analyse the project datasets.
* Analysing the performance Spark over HiveQL.

|  |  |
| --- | --- |
| Project 5 : **TeratoHadoop** | **July 2016 – December 2016** |
| Role: Team Member | Team size : 11 |
| Environment:**HBase, Hive, UNIX scripting, Teradata, Spark, Kafka** | |

**Business Use:** Cost of data maintenance is reduced – Centralized Hadoop data pool.

**Roles & Responsibilities :**

* Migration of data repository from Teradata to Hadoop.
* Configurable application to Intake of data from source systems to Hadoop.
* Provision enriched and transformed data to Downstream and Teradata system for business continuity.
* Implemented generic and scalable frame work to create Teradata common format in big data ecosystem.
* An alternative to hive queries - Built Metadata driven map reduce framework to read HBase tables and provision data. Reduction in container usage and performance gain when benchmarked with Hive ORC queries.
* Implementing of 200 CF with more than 2000+ tables and loading to Teradata.

|  |  |
| --- | --- |
| Project 4 :**Data Lake** | **January 2016 – June 2016** |
| Role: Team Member | Team size : 11 |
| Environment:**HBase, Hive, Pig, Core Java, Sqoop,UNIX shell scripting** | |

**Project description:**

Existing DWH system holds the data in form of multiple files and locations including oracle database / mainframe data / xml input files and UNIX system log files. A consolidated HDFS data pool is created for better analysis of business data.

**Business benefits:**

* Consolidated data pool providing more opportunities to transform data to information.
* Enabling the IT system to cope with expanding business and huge flow of data.
* Availability of huge data from past dates for analysis and using business intelligence tools.
* Faster retrieval and transformation of data in to valuable information.
* More accurate analytic results and benefit to business predictions.

**Responsibilities handled:**

* Involved in Data Migration from RDBMS(Oracle) to HDFS for data analytics using MR and Hive
* Hadoop Environment integrated with various data sources i.e., Oracle DB, Mainframe data which are unstructured(logs, error files, XMLs and text files).
* Perform data analysis using Pig and Hive.
* Involved in loading data from UNIX file system to HDFS.

**DataIngestion: November 2015-January 2016**

* Creation of Generic tool to pull incremental data and updates from structured datasource to Hadoop File System.
* Usage of Sqoop jobs to ingest and Hive on Hbase to analyse and make a move to ease the data ingestion mechanism thereby productivity.

|  |  |
| --- | --- |
| Project 3 :**Data Interpretation System** | **February 2015 – November 2015** |
| Role: Team Member | Team size : 11 |
| Environment: **HBase, Hive, Pig, Oozie, SQOOP, Core Java, UNIX Shell scripting** | |

Client’s inventory and prediction modeling was managed through dozens of spreadsheets, requiring significant manual processes and creating a potential for human error.Our solution ingested all inventory data into HDFS which allows for the creation of predictive models that are dynamically refreshed to provide better insight. After using distributed framework and the data process, client is able to develop more accurate models for ideal floor plans and inventory logistics. In addition, the Hadoop-based system allows business users to easily interpret data outputs.

**Responsibilities handled:**

* Developed multiple MapReduce jobs in java for data cleaning and preprocessing.
* Perform data analysis using Hive and Pig.
* Automation script to monitor HDFS and HBase through cronjobs.
* Plan, design, and implement processing massive amounts of marketing information, complete with information enrichment, text analytics, and natural language processing.
* Involved in loading data from UNIX file system to HDFS.
* Prepare multi-cluster test harness to exercise the system for performance and failover.
* Develop high-performance cache, making the site stable and improving its performance.

|  |  |
| --- | --- |
| Project 2 : **Employee Management System** | **February 2014 – December 2014** |
| Role: Team Member | Team size : 11 |
| Environment: **Core Java, JSP, Spring, Hibernate, Oracle, Tomcat, Eclipse** | |

This is a quality assurance tool to evaluate the voice,email and chat transactions in detail so that the organization can be able to focus on customer needs and results are shared with management team for use in compensation, computation, remedial efforts for individual agents. It involved different modules significantly Agent, Client, Management modules.

**Responsibilites Handled :**

* Involved in implementation of the agent module of the application.
* Understanding functional specification as well as technical aspects of the project.
* Interacted with the client to gather reuirements for development and with different departments to understand the requirements and develop the application user friendly and coordinated with developers to delivser the project efficiently.
* Involved in writing queries that perform DML,DDL operation using Oracle Database.
* Participated in project review meetings on a weekly basis.
* Developed Hibernate mapping files, configuration files and POJOs.
* Developed server side components containing business logic using Spring MVC and DAO using Spring ORM functionality.

|  |  |
| --- | --- |
| Project 1 : **RedCabs** | **January 2013 – December 2013** |
| Role: Team Member | Team size : 11 |
| Environment: **Core Java, JSP, Servlets, HTML, JavaScript, Oracle, Tomcat, Eclipse** | |

This project aims to develop a centralized application for cab booking system.This manages details of Customers,Drivers and Cabs. This provides a common interface designed for every user(Customer, Admin and Driver) and consists of four modules. The capabilities included are centralized access to services,a common interface look and feel,profile management and integrated development architecture.

**Responsibilities handled :**

* Developed JSP pages, HTML pages as per the requirement.
* Developed servlet pages for business logic implementation.
* Responsible for developing client side validation script using Javascript.
* Involved in development of tables for the respective module using Oracle.