**Rohit Ravva**

# 901-326-1862

[**rravva8@gmail.com**](mailto:rravva8@gmail.com)

[linkedin.com/in/rohit-ravva](http://www.linkedin.com/in/rohit-ravva)

**PROFESSIONAL SUMMARY**

* Over 7 years of professional IT experience with over 4 Years of Big Data experience and deep insight on Hadoop Eco System.
* Experienced Big Data Consultant with a demonstrated history of working in the banking, healthcare and logistics industries. Skilled in BigData - Hadoop, Spark, Python, Oracle, Teradata, PL/SQL and Data warehouse architecture.

**EDUCATION**

* Masters of Science in Managing Information Technology from Sullivan University(3.7/4.0) **Jun 2013 –Sep 2014**
* Masters in Computer Science from Eastern Illinois University (3.7/4.0) **Aug 2009 - Dec 2010**
* Bachelors in Computer Science from Kakatiya University, India (8.0/10) **Aug 2005 - Apr 2009**

**CERTIFICATION**

* Apache Spark Certified Developer

**TECHNICAL SKILLS**

**Big Data Ecosystem**: Hadoop, MapReduce, Pig, Hive, HBase, Flume, Sqoop, Impala, Oozie.

**Spark** **components**: RDD, Spark SQL (Data Frames and Dataset), Spark Streaming.

**Languages**: Python, Core Java, SQL, C.

**Databases**: Oracle, MySQL, SQL Server.

**IDE’S & Command line tools:** Eclipse, IntelliJ, GIT, Autosys, MS Project 2007, Toad and WinSCP.

**Reporting Tools:** SAP Business Objects 3.x,4.0, Tableau, Bex

**Build** **Management** **Tools**: Maven, Ant, Jenkins

**Operating systems**: UNIX, Linux, and Windows Variants

**WORK EXPERIENCE**

**Bank of America - Charlotte, NC Apr 2016 - Present**

**Big Data/Spark Consultant**

**Responsibilities**

* Implementing end-to-end solutions in support of Enterprise Credit/Quantitative Risk Tech products.
* Developed end-to-end data pipeline using Sqoop to store data into HDFS from existing RDBMS system and further processing through Spark.
* Designed and developed spark framework to convert existing Java mapreduce framework to spark.
* Converted existing hive scripts to Spark applications using RDD’s/Dataframes to transform data and loading into HDFS.
* Converted existing sqoop, hive jobs to SparkSQL applications to read data from Oracle using JDBC and write it to hive table.
* Developed sourcing framework to source data from different datasources and performing DI & DQ checks for UI and batch jobs.
* Converted complex oracle and Netezza SQLs into Hive HQLs. Developed ETL using Hive, Oozie, shell scripts and Sqoop.
* Used Hive to analyze the partitioned and bucketed data and compute various metrics for reporting.
* Extensively worked on performance optimization of hive queries by using map-side join, parallel execution, cost based optimization and vectorization
* Developed Shell scripts as part of the process flow for several cases such as wrapper scripts for Hive, Spark, oozie, mapreduce jobs etc.
* Extensively worked on Text, ORC, Avro and Parquet file formats and compression techniques like Snappy, Gzip and Zlib.
* Experience in Job management using Autosys scheduler and Developed job processing scripts using Oozie workflow.
* Worked in agile environment namely Scrum and Kanban methodologies. Actively participated in Monthly deployments.

**AESOP Works - Memphis, TN Feb 2015 – Mar 2016**

**Sr. Hadoop Developer**

**Responsibilities**

* Implemented the data pipeline using Big Data tools including Hive, Pig, Oozie, Sqoop,Map Reduce Programs.
* Used Sqoop to extract and load incremental and non-incremental data from RDBMS systems into Hadoop.
* Responsible for analyzing large data sets and derive customer usage patterns by developing new MapReduce programs.
* Implemented MapReduce programs for pre-processing and cleansing the data in HDFS obtained from heterogeneous data sources to make it suitable for ingestion into Hive schema for analysis.
* Worked on creating combiners, partitions, and distributed cache to improve the performance of MapReduce jobs
* Responsible for debug, optimization of Hive scripts and also implementing DE duplication logic in Hive.
* Experienced in writing Hive validation scripts that are used in validation framework.
* Used Hive to analyze the partitioned and bucketed data and compute various metrics for reporting.
* Create/Modify shell scripts for scheduling various data cleansing scripts and ETL loading process
* Production Rollout Support that includes monitoring the solution post go-live and resolving any issues that are discovered by the client and client services teams.

**Fedex TechConnect - Memphis, TN Feb 2014 – Jan 2015**

**Hadoop Developer**

**Responsibilities:**

* Involved with Data Science team in gathering the requirements and transforming the Business requirements to technical designs.
* Worked on analyzing Hadoop cluster using different big data analytic tools including Pig, Hive and MapReduce.
* Developed Map Reduce programs in Java for parsing the raw data and populating staging tables.
* Written custom mapreduce programs that includes use of custom Data Types, Input Format, Record Reader, Output Format and Record Writer for Map Reduce computations.
* Imported data from various data sources like MySQL and Netezza using Sqoop, SFTP, performed transformations using Hive, Pig and loaded data back into HDFS.
* Created multiple Hive tables, implemented Dynamic Partitioning and Buckets in Hive for efficient data access.
* Involved in creating Hive External tables, also used custom SerDe’s based on the structure of input file so that Hive knows how to load the files to Hive tables.
* Solved performance issues in Hive and Pig scripts with understanding of Joins, Group and Aggregation.
* Exposed Hive tables in impala to minimize query response time to end users.
* Troubleshoot and resolve data quality issues and maintain high level of data accuracy in the data being reported.
* Optimized Mapreduce jobs to use HDFS efficiently by using various compression techniques.

**Deloitte - Nashville, TN Mar 2013 – Dec 2013**

**Hadoop Developer**

**Responsibilities:**

* Responsible for loading the customer’s data and event logs from Oracle database, Teradata into HDFS.
* Involved in creating Hive tables, loading with data and writing Hive queries, which will run internally in MapReduce way.
* Implemented different analytical algorithms using MapReduce programs to apply on top of HDFS data.
* Designed a data warehouse using Hive and created Partitioned tables in Hive.
* Load and transform large data sets of structured, semi structured and unstructured data.
* Worked with various Hadoop file formats, including Text Files, SequenceFile, RCFile.
* Create/Modify shell scripts for scheduling various data cleansing scripts and ETL load process.

# Intuitive Surgical Inc - Sunnyvale, CA Aug 2013 – Dec 2013

**Sr. BI Consultant**

**Responsibilities:**

* Used data visualization software and data modeling tools to analyze and communicate complex data and information clearly and efficiently to business users.
* Assisted in the development of reporting, analysis and visualization of business data, metrics and measures in order to allow staff to align and track against overall business strategies, goals and objectives, drive innovation, insight and decision making.
* Supported existing data models and provide clarifications to business on the existing BI reports or data.
* Using SQL skills to optimize, analyze, and improve database SQL performance on any platform.
* Extensively worked with HP-ALM to update testing status and worked with users regarding the test cases.

**BlackBerry - Atlanta, GA Nov 2011 – Jun 2012 BI Consultant**

**Responsibilities**:

* Gathered requirements from client, analyzed and prepared the requirement specification document.
* Experience in writing PL/SQL and in developing and implementing Stored Procedures, Packages and Triggers.
* Experience on modelling principles, database design and programming, creating E-R diagrams and data relationships to design a database.
* Worked with different sources such as Oracle, SQL and Flat files.

**Fedex TechConnect- Memphis, TN Apr 2011 – Aug 2011**

**BI Consultant**

**Responsibilities:**

* Gathering system requirements from Business Owners and Business Users.
* Interacted with Business Users to understand their specific Business needs for reporting for their daily business
* Created Reports using all the functions like Drill Down, Prompts, Dimensional and Measure variables to show accurate Results.
* Scheduled reports and Implemented dynamic cascading prompts.