

Trushant Anup Bhongale

Contact: +91-7350222978

Email Id: tbhongale@gmail.com

Tara Residency, Flat No.03,

Tapowan Road, Pimpri,

Pune – 411017.

Professional Summary

- Working as Senior Big Data Engineer in Capgemini, Pune. Having over 3.2 Years' of IT Experience on Apache Hadoop Big Data Framework, Spark, Kafka, MapReduce, Sqoop, Hive, Pig, Flume, Java, SQL, Shell Scripting.
- Good knowledge of Core Java – OOP, Collection Framework, Generics, Multithreading, Exception Handling, Serialization, Garbage Collection.
- Knowledge of Spark Ecosystem, Transformation and Actions in Scala and Java.
- **Hortonworks** Data Platform Certified Developer on **Hadoop Ecosystems**: <http://bcert.me/sryifkux> (Dec – 2015)
- **Databricks** Certified Developer on **Apache Spark**: <https://drive.google.com/file/d/0ByUcArb9kSqObVA5NWfxSVpScIk/view?usp=sharing> (Sep – 2016)
- Excellent understanding of all aspects of the Software Development Life Cycle and ability to adapt open source technologies.
- My career has focused on to work in challenging environments in the field of Big Data and Analytics; contributed on the Organization's growth and success.

Skills Profile

Technical

Operating System	Windows 7, Ubuntu 14.04, CentOS 6.4
Environment	Apache Hadoop Big Data Framework, MapReduce, Spark, Java
Big Data Ecosystem	Hive, Sqoop, Pig, Flume, Kafka
Database	MySQL 5.0, Oracle 10g, SQL Server, MongoDB
Other	GitHub, SVN, Eclipse, Cloudera Manager, Ambari.

Professional Experience

Capgemini India Pvt. Ltd.

March 2015 – Present

BABAR - Data Migration to Business Big Data Lake

The objective of this project to migrate all the client data to the defined business data lake. The client is financial giant where they have their historical data stored in different sources. We perform UAT on developed code and perform production deployment.

Roles and Responsibility

- Involve with ingesting data received from various relational database providers, on HDFS for analysis and other big data operations.
- Write MapReduce jobs to perform operations like copying data on HDFS and defining job flows load and transform large sets of structured, semi-structured and unstructured data.
- Write MapReduce to ingest XML, EBCDIC, Fixed Length, and Delimited Flat Files into Hive.
- Create Hive tables to import large data sets from various relational databases using Sqoop and export the analyzed data back for visualization and report generation by the BI team.
- Very good understanding of Partitions, Bucketing concepts in Hive and designed both Managed and External tables in Hive to optimize performance.

- Write PIG script to perform ETL operations on large dataset and develop UDFs in Java when necessary to use in Hive and PIG queries.
- Experience in using Sequence files, RCFile, ORCFile, Avro, Parquet file formats.
- Involve in managing and reviewing Hadoop log files and uploading the final results and written queries to analyze them.
- Use SVN for version control, JIRA for internal bug/defect management, MapReduce.
- Write Oozie workflow and coordinator configuration to automate the ingestion process.

Environment: Hadoop (Hortonworks), MapReduce, Hive, Sqoop, Flume, Pig, Java Jdk1.7, Eclipse 3.6, SVN, GitHub

Real Time Weblog Ad-hoc Analytics

The objective of this project to support Capgemini Talent Ad-hoc Analysis.

- Collect weblog data in real time.
- Perform ETL operations in real time and apply schema afterwards.
- Provide result for ad-hoc analysis by executing queries.

Environment: Hadoop (Hortonworks), Kafka, Spark

GlobeIT Solutions Pvt. Ltd.

Dec 2013 – Feb 2015

A System to Filter Unwanted Post from Wall of User

We evaluate an automated system, called Filtered Wall (FW), which is able to filter unwanted post from OSN user's wall. We exploit Machine Learning (ML) text categorization techniques to automatically assign with each short text message a set of categories based on its content. After classification, system provides a powerful rule layer exploiting a flexible language to specify Filtering Rules (FRs), by which users can state what contents, should not be displayed on their walls. Filtering Rules exploit user profiles, user relationships as well as the output of the ML categorization process to state the filtering criteria to be enforced.

Roles and Responsibility

- Developed user interfaces using JSP, HTML and CSS.
- Designed business components using DAO.
- Designed business logic using Pojo, Servlet and Collections.
- Used JavaScript for performing front end validations.

Environment: Java, JSP, Servlet, HTML, CSS, JavaScript, MySQL, Apache Tomcat, Eclipse, SVN.

Certifications

- **Hortonworks** Data Platform Certified Developer on **Hadoop Ecosystems**: <http://bcert.me/sryifkux>
- **Databricks** Certified Developer on **Apache Spark**: <https://drive.google.com/file/d/0ByUcArb9kSqObVA5NWFxSVpScIk/view?usp=sharing>

Honors and Awards

- **Rising Star for Q2 2016** at Capgemini – 2016.
- **Best Intern 2014** at GlobeIT Solutions – 2014.

Education

- University of Pune, Pune, Maharashtra **MCA** – 2014.
- University of Pune, Pune, Maharashtra **BCA** – 2011.

Personal Information

Date of Birth	-	25 th December, 1989
Current Location	-	Pune
Nationality	-	Indian
Passport No	-	N8671043
Pan Card No	-	BETPB7646K
Mobile No	-	+91-7350222978
