Hadoop Developer Hadoop Developer Hadoop Developer   Having 3+ years of overall IT experience, diversified exposure in Software Process Engineering, developing, building enterprise applications using Java, SQL, Big Data Technologies in various domains like Financial sector and Telecom.   Strong working experience, on different Big Data Platforms like Cloudera (CDH), Hortonworks (HDP) and MapR Distribution.   Strong working experience with Ingestion, Storage, Querying, Processing and Analysis of Big data Integration and Talend Integration.   Worked on developing SPARK Applications using Scala and Java, Spark Core, Spark Data Frames, Spark SQL and Spark Streaming API's for Fast processing of Data.   Experience in developing Map-Reduce programs and custom UDF's for data processing using Java and Scala.   Hands on experience working on different Hadoop components like Big Data Talend, Kafka, Spark, Hive, HBase, Sqoop, Cassandra and Zookeeper and used tHDFSInput, Thdfs Output, TPigLoad, tPigFilterRow and tsqoopImport and export.   Involved in Creating Hive Tables and load processed data into tables and fine-tuned performance using different optimization techniques.   Worked on importing and exporting data using Apache Sqoop from RDBMS to Hadoop Platform and vice versa.   Knowledge on troubleshooting failures in spark applications and fine-tuning for better performance.   Hands on experience working with Dimensional Modelling(Star Schema and Snowflake Schema)   Experience in using Zookeeper for coordinating the distributed applications.   Good experience with NoSQL databases like HBase and Cassandra.   Experience in Object Oriented Analysis, Design (OOAD) and development of software using UML Methodology, good knowledge of J2EE design patterns and Core Java design patterns.   Expertise in various JAVA/J2EE technologies such as JSP 2.0, Servlets 2.x, Struts 1.2/2.0, Hibernate 2.0/3.0 ORM, Spring 2.0/3.0, JDBC.   Very good experience in complete project life cycle (design, development, testing and implementation) of Client Server and Web applications.   Possess excellent communication and analytical skills along with a can-do attitude.   Experience in developing Web Applications with various Open Source frameworks: Spring, Hibernate 2.0/3.0 ORM Frameworks.   Proficient in Core Java, J2EE, JDBC, Servlets, JSP, Exception Handling, Multithreading Concepts.   Experience in Testing and documenting software for client applications.   Good experience in using Data Modelling techniques to find the results based

on SQL and PL/SQL queries.   Worked and learned a great deal from Amazon Webservices (AWS) Cloud Services like EC2, S3 and EMR.   Experience working with different databases, such as Oracle, SQL Server and MySQL, writing stored procedures, functions, joins, and triggers for different Data Models.   Experience in using version control tools like GITHUB, BitBucket share the code snippet among the team members.   Experience in working in Agile project management.   Strong motivational skills, familiarity with various technologies, ability to learn quickly, dealing with people, commitment to work and believes in hard-working. Work Experience Hadoop Developer Cisco September 2018 to March 2019 Project Name: Marketing - IT.   Project description: The primary goal of this project is to develop its enterprise data hub for storing, accessing and analyzing Marketting data. At Cisco , company stores large volumes of customer data from different sources like web pages, transactions , and SFDC. This will be core to its overall information delivery strategy and allows team comprehensive view and study of customer behaviour and target the customer based on their intrests.   Responsibilities: Responsible for ingesting large volumes of data into Hadoop Data Lake Pipeline on daily basis   Read data from different source like Oracle and Hive using Apache Spark and process consumed data using Spark and load the processed to HIVE and downstream applications for different purposes.   Developed Apache spark application using Java to perform different kind of validations and standardization on fields based on certain validations rules on incoming Data.   Designed and Implemented ETL Process to load data from Source to Target Tables.   Once data processed loaded data into different Hive Tables based on the derived fields.   Used Java based Spark RDD's, Data Frames and Datasets Spark- SQL faster processing of ETL workloads.   For performance optimization implemented different optimization techniques optimizations like Partition, Bucketing and Map Side joins for fast query processing.   Excellent working Knowledge in Spark Core, Spark SQL, Spark Streaming.   Implemented Spark using Scala and Spark SQL for faster testing and processing of data.   Involved in converting Hive/SQL queries into Spark transformations using Spark Data Frames and Scala.   Design data ingestion and integration process using SQOOP, Shell Scripts & Pig, with Hive   Optimizing existing algorithms in Hadoop using Spark Context, Spark-SQL, Data Frames and Pair RDD's   Worked on Error

Handling techniques and tuning ETL flow for better performance   Generated Java APIs for retrieval and analysis on No-SQL database such as HBase.   Worked on troubleshooting and Performance optimization on Spark Application to improve the over-all processing time and make more error tolerant for the pipeline.   Collaborated with the infrastructure, network, database, application and BA teams to ensure data quality and availability.   Used NOSQL HBase for storing batching information.   Used Rally tracking tool for assigning and defect management.   Worked in Agile development environment and actively involved in daily Scrum and other design related meetings.   Used BitBucket to share the code snippet among the team members.   Environment: Spark 2.11, Oracle 11.2.0.3, Apache Solar 6.6.5, Datastax Cassandra 3.0.0, Hbase1.1.8, Talend, Bitbucket, Rally. Hadoop Developer T- Mobile - Atlanta, GA June 2017 to August 2018 Project Name: Continuous Service Providing.   T- Mobile is one of the top Telecommunication company in US providing cell phone services. It's the second Biggest mobile network operator serving clients with high quality mobile networks solutions.   For Continuous Service Providing and detecting the towers, cables and routers before getting failed. T- mobile stores large volumes of customer usage data which provides company an accurate, data- driven pictures of user experience and the areas experienced by outage. Mainly useful to estimate and setup repairs to cell towers based on how it impacts customer experience. By predicting and identifying the tower behavior, which tower needs maintenance and fixing the repairs before they become inactive which helped T-Mobile in improving customer experience and saved costs.   Responsibilities: Responsible for ingesting large volumes of data into Hadoop Data Lake Pipeline on daily basis.   Developed Spark programs to Transform and Analyze the data.   Build data pipeline using Kafka and developed Kafka producers to stream data from external source to Kafka Topics.   Developed Kafka Consumer API's in Scala for consuming data from Kafka topics and process consumed data using Spark and load the processed streams to HIVE for analytics.   Developed Spark applications using Scala for performing data cleansing, event enrichment, data aggregation and data preparation needed for reporting teams to consume.   Experienced in handling large datasets using Spark in Memory capabilities, using Broadcast variables in Spark, efficient Joins, transformations and other capabilities.   Used Scala

based Spark RDD's and Spark- SQL for faster testing and processing of ETL workloads and loaded the results into HIVE.   Designed and Developed end-to-end process from various source systems to staging area from staging to Data Marts.   Transformed RDD's into Data Frames using CaseClass and toDF( ) also published the end reports to HDFS.   Involved in creating HIVE tables, imported processed data into tables and done analytics as per business developments.   For performance optimization implemented partitions, bucketing and compression techniques and Hive on Tez for fast query processing.   Worked on troubleshooting and Performance optimization on Spark Application to improve the over-all processing time and make more error tolerant for the pipeline.   Used Zookeeper to provide coordination services to the cluster.   Used MySQL for storing customer data, moved data to HDFS when needed and run Spark jobs for analyzing customer behavior.   Worked on NOSQL database HBase for Further analysis   Run Sqoop jobs to migrate data from MySQL Database to Hadoop File System.   Generated daily final report's on processed data using Tableau.   Used JIRA tracking tool for assigning and defect management. Worked in Agile development environment and actively involved in daily Scrum and other design related meetings.   Environment: Hortonworks (HDP), Hadoop, HDFS, Hive, Spark, Scala, Kafka, Sqoop, Oozie, Zookeeper, MySQL, HBase, Jira, Tableau. Java Developer Mahindra Convivia Technologies - Bangalore, KARNATAKA, IN May 2014 to November 2015 Bangalore - India May 2014- Nov 2015   Project Name: Integrated Messaging (IM) Role: Java Developer   Mahindra Convivia is a value-added service provider for mobile operations and sells solutions in Consumer value management, Business and Financial Solutions. It provides mobile apps and a variety of voice, SMS and Internet Services.   Developed application which provides all the VAS services like SMS, MMS, Data using Integrated Messaging Platform which enables the user to configure services based on their requirements. Application also provides insight of the traffic and load on the application with statistics.   Responsibilities: Involved in all phases of project development - requirements gathering, analysis, design, development, coding and testing.   Developed the application using Spring MVC and involved in setting up the Spring Bean Profiling. Developed the data layer for the applications using Spring Hibernate ORM and developed various business logic

and services using HQL.   Used Hibernate named queries concept to retrieve data from the database and integrate with Spring MVC to interact with back end persistence system (SQL Server).   Used Maven script for building and deploying the application.   Strong experience in SQL and knowledge of MySQL Server database   Wrote SQL queries to fetch the statistics from the database and displayed the statistics in the application UI.   Experience in ETL software development.   Involved in code review and documentation review of technical artifacts.   Testing of the product: Unit Testing, Regression Testing, and Integration Testing.   Used Tortoise SVN for source code versioning and code repository.   Used Jenkins for Continuous Integration Builds and Deployments (CI/CD)   Used JIRA to track bugs.   Environment: Java, Hibernate, Spring MVC, MySQL, Maven 3.2, JIRA, Jenkins, Agile, Eclipse IDE. Additional Information Technical Skills: BigData Technologies HDFS, Spark, Hive, MapReduce, Sqoop, Oozie, HBase and Streaming Kafka-Connect   Processing Spark, Spark Streaming, Spark SQL, Hive,   Big Data Platforms Cloudera, Hortonworks, MapR   Databases/ RDBMS MySQL, SQL/PL-SQL, MS-SQL Server 2005, Oracle 9i/10g/11g   Language Java, Scala, PL/SQL   ETL Tools Kafka, Sqoop, ETL ,Talend Data Integration, Big Data Integration, Informatica Power Center   Software Life Cycles SDLC, Agile models   Cloud Platforms Amazon EC2, S3, EMR.   Bug Tracking Tools Jira, Rally

Name: Deborah Gilbert

Email: whitelaura@example.net

Phone: (256)269-0618