**Haris Mahmood**

**Professional Summary**

I am a Hadoop Engineer with experience in Java-related technologies, specifically Servlets and JDBC. Used front-end technologies to make full-stack applications. Using the knowledge gained from Java, I have implemented MapReduce into one of the projects for Hadoop.

* Utilized **HTML5**, **CSS**, **Bootstrap**, **JavaScript**, **Java**, and **Hadoop** technologies.
* Constructed webpages that included the use of **HTML5**, **CSS**, **Bootstrap**, and **JavaScript**.
* Used **JavaScript** to fetch the objects from the **Servlet** to display on the webpage.
* Implemented **Servlets**for communication between client-server.
* Incorporated **JDBC** to include the database in web apps.
* Built the database using **Sqoop** to add the data into **MySQL**.
* Included the use of **Sqoop**, and **Hadoop Distributed File System ( HDFS )**to store and organize files.
* Stored **Sqoop** jobs for later use, when having to migrate other data files in the future.
* Used **Hive** to incorporate **MapReduce** through SQL queries.
* Made Java programs that ran along with **YARN** to run **MapReduce**.
* Built a workflow.xml file for **Oozie** workflows.
* Incorporated the use of **EMR** for clusters in the cloud.

**Certifications**

**Education**

Bachelors in Biology from East Carolina University (No degree)

Engineering from Wake Technical Community College (No degree)

**Technical Skill Summary**

|  |  |
| --- | --- |
| Web Technologies | JavaScript, Bootstrap, Servlets, AJAX, HTML5, CSS |
| Database | JDBC, Hive, Hue, OracleSQL, MySQL, SQL |
| Hadoop Development | HDFS, Oozie, YARN, EMR, Hue, Hive, Sqoop, Spark |
| Advanced Java | MapReduce, Servlets, Design Patterns, JDBC, JUnit, MRUnit |
| Server-Side Technologies | MySQL, Apache, Tomcat, Java, Servlets, JDBC |

**Professional Experience**

**Revature May-2019 to Present**

**Biforce**

Revature is in need to improve its business decisions. Based of data on an existing OLTP system, the requirement is to utilize all relevant tools to increase the value of the business by providing new metrics, reaffirming existing ones, and automate the process of the Business Intelligence lifecycle of the training department and company as a whole, with the use of efficient algorithms and tools available within the Hadoop ecosystem, in a physical, and in a cloud cluster.

**Responsibilities:**

* Incorporated Sqoop to import and export files into and out of HDFS.
* Included HDFS to store and organize files.
* Used a RDS to retrieve data.
* Made the schema for the data to be migrated from Caliber 1 to Caliber 2.
* Used Machine Learning to perform a specific task without specific instruction.
* Made a presentation to explain the process of the migration.
* Applied RedShift for Data migration.
* Utilized S3 for using objects that would be needed for different parts of the process.
* Included the concepts of Agile and Scrum to stay on track and keep track of progress throughout the project.

**Environment:**

MapReduce, HDFS, Sqoop, Zookeeper, Cloudera, Oozie, Spark, EMR, Java, Hive, Git, JUnit, EC2, RDS, SQL, Python

**Gender Statistics Analysis - Advanced**

The Gender Statistics Analysis project aims to identify factors regarding gender and its effects on those in higher education. With a heavy focus on geographical location, this project strives to identify special programs geared towards improving the general welfare of women across the globe via investigation and analysis of specific metrics outlined in a semi-structured realistic data set.

**Responsibilities:**

* Transformed project data from local machine into HDFS.
* Hosted source code on GitHub for version control.
* Identified relevant data for each susiness inquiry via data exploration.
* Implemented MapReduce algorithms for each business inquiry.
* Tested functionality of each job with MRUnit.
* Managed project build and needed dependencies via Maven.
* Pinpointed programmatic issues through Log4j.
* Developed and debugged source code in Eclipses.

**Environment:**

HDFS, Sqoop, Hive, Oozie, MapReduce, Git, Cloudera, Unix, Java

**Expense Reimbursement System - Java**

The Expense Reimbursement System (ERS) will manage the process of reimbursing employees for expenses incurred while on company time. All employees in the company can login and submit requests for reimbursement and view their past tickets and pending requests. Finance managers can log in and view all reimbursement requests and past history for all employees in the company. Finance managers are authorized to approve and deny requests for expense reimbursement.

**Responsibilities:**

* Used OracleSQL to create the database.
* Incorporated JDBC to implement the database.
* Used servlets to make requests between front-end and back-end.
* Incorporated JavaScript to display client information on the webpage.
* Included sessions to keep the same information for the client logged in to the website.
* Utilized Bootstrap for a responsive webpage.
* Built a web.xml in order to have the servlets communicate.
* Included the RDS from Amazon Web Services.

**Environment:**

Servlets, Java, JavaScript, HTML, CSS, JDBC, SQL, AJAX, Bootstrap, RDS, Tomcat, Git, Maven