

Senior Big Data Developer - Spark Senior Big Data Developer - Spark Senior Big Data Developer - Spark - Huntington Bank Piscataway, NJ 6+ years IT experience with 3 years in Big Data Hadoop development and 3 years in Java/J2EE technologies Sound domain knowledge in the area of banking, insurance, healthcare and catering Certified Cloudera Spark and Hadoop Developer & Oracle Java SE 8 Programmer I & II Hands on Experience in Hadoop ecosystem including HDFS, Spark, Hive, Pig, Sqoop, Impala, Kafka Oozie, Flume, NiFi, HBase, ZooKeeper and MapReduce Expertise in Java, Scala, C++, C and scripting languages like Python Experience in Spark Streaming, Spark SQL in a production environment Hands-on experience on RDD architecture, implementing Spark operations and optimizing transformations Experienced with distributions including Amazon Web Service and Cloudera CHD 5 Worked on building, configuring, monitoring and supporting Cloudera Hadoop CHD 5 Extensive experience in data ingestion technologies, such as Flume, Kafka and Sqoop Utilize Kafka, NiFi and Flume to gain real-time and near-real time streaming data in HDFS from different sources Extensive experience in creating Hive tables and queries using HiveQL Experience in Hive partitions and bucketing to optimize performance Experience in designing time driven and data driven automated workflow using Oozie Experience in NoSQL databases, such as HBase 0.98 and MongoDB 3.2 and ensured faster access to data on HDFS Worked with RDBMS including MySQL 5.5, Oracle 10g and PostgreSQL 9.x Extracted data from log files and push into HDFS using Flume In depth understanding of Hadoop Architecture, workload management, schedulers, scalability and various components, such as HDFS, MapReduce and YARN Good knowledge of Data Mining, Machine Learning and Statistical Modeling algorithms including K-Means, Decision tree, Perceptron, Winnow, Linear regression, SVM, Adaboost, Neural Network and Navie Bayes Experienced in Machine Learning and Data Mining with Python, R and Java Skilled at Data Visualization with Tableau Good knowledge in UNIX shell commands. Hands on experience in MVC architecture and J2EE frameworks like Struts 2, Spring MVC Experience in web development using HTML, CSS, Javascript, JQuery and Hibernate Familiar with Agile methodology standards and Test Driven Development Extensive Experience in Unit Testing with JUnit, MRUnit and Pytest Excellent communication skills.

Successfully working in fast-paced multitasking environment in collaborative team, a self-motivated enthusiastic learner.

TECHNICAL SKILLS Hadoop Eco-systems Web Technologies SOAP, REST, JSP 2.0, JavaScript, Servlet PHP, HTML5 HDFS, MapReduce, HBase, Spark 1.3+, Hive, Pig, Kafka 1.2+, Sqoop, Flume, NiFi, Impala, Oozie, ZooKeeper Operation Systems Programming Languages Linux (CentOS, Ubuntu), Windows, Mac OS Java 6+, Scala 2.10+, Python, C, C++, R, PHP, SQL, JavaScript, Pig Latin Machine Learning algorithms Regression, Perceptron, Naive Bayes, K-means, Relational & NoSQL Databases Decision tree, SVM MySQL, Oracle, PostgreSQL, HBase, MongoDB Work Experience Senior Big Data Developer - Spark Huntington Bank - Indianapolis, IN November 2015 to Present Description: Huntington Bank is a full-service banking provider. Their business involves retail and commercial financial services to individuals, families and business. The purpose of this project is to build a real time solution for credit card fraud detection based on technologies including Hadoop, Spark Streaming and Apache Kafka. Also, we built new processing pipelines over transaction records, user profiles, files, and communication data ranging from emails, instant messages, social media feeds.

Responsibilities: Designed and implemented scalable infrastructure and platform for large amounts of data ingestion, aggregation, integration and analytics in Hadoop, including Spark, Hive, Pig and HBase Loaded large sets of structured, semi-structured, and unstructured data with Sqoop and Flume Wrote Sqoop scripts to import, export and update the data between HDFS and Relational databases Created Flume configure file to collect, aggregate and store the web log and event data. Loaded, transformed and analyzed data using Hive queries (HiveQL) Configured the Spark cluster as well as integrating it with the existing Hadoop cluster Utilized Kafka to capture and process real time and near-real time streaming data Developed Spark programs in Scala to perform data transformation, data streaming and analysis Utilized historical data stored in HDFS and HBase to build machine learning model which can be used to make predictions on live events Worked with analytics team to build statistical model with Spark MLlib Worked with Oozie and Zookeeper to manage job workflow and job coordination in the cluster Performed unit testing using JUnit and PyTest

Environment: Hadoop, HDFS, Spark Streaming, Zookeeper, Oozie, HBase, Hive, Sqoop, Flume,

Kafka, JUnit, PyTest, Scala Big Data Developer - Hadoop The Hanover Insurance Group - Somerset, NJ October 2013 to September 2015 Description: The Hanover Insurance Group is the holding company for several property and casualty insurance. This company mainly focused on home, auto and business insurance, it also offers wide variety of flexibility and claims. The project is to build a fully distributed HDFS and integrate necessary Hadoop tools. Meanwhile, our team needs to support analytic team to build risk management models for insurance product analysis and innovation. Responsibilities: Responsible for building scalable distributed data solutions using Hadoop on Amazon EC2 Installed and configured Hadoop clusters and Hadoop tools for application development including Hive, Pig, Sqoop, Flume, Zookeeper and Oozie Created multiple Hive tables with partitioning and bucketing for efficient data access. Extracted and loaded customer data from databases to HDFS and Hive tables using Sqoop Used Flume to transfer log source files to HDFS Performed data transformations, cleaning and filtering, using Pig and Hive Worked with analytic team to prepare and visualize tables in Tableau for reporting Developed workflow in Oozie to automate the tasks of loading the data into HDFS and pre-processing with Pig Performed unit testing using JUnit and MRUnit Environment: Hadoop, HDFS, YARN, MapReduce, Sqoop, Flume, Hive, Pig, Zookeeper, Oozie, Oracle, JUnit, MRUnit Java Developer JPMorgan Chase - Brooklyn, NY February 2012 to August 2013 Description: JPMorgan Chase is one of the largest banks in the United States. It provides services including markets insurance, investment products and standard banking transactions. The aim of this project is to create a web application that bank phone officers will be able to view credit card information and query different transactions such as address change, credit card summary and card reissue. Responsibilities: Developed the application implementing Spring MVC architecture with Hibernate as ORM framework Developed user interface by using JSP, HTML5, CSS3 and JavaScript Implemented DAO using JDBC for database connectivity to MySQL database Wrote SQL for querying, inserting and managing as required on the database object Implemented user input validations using JavaScript and JQuery Developed test cases and performed unit test using JUnit framework Used Agile methodology for the development of the project Environment: Eclipse, Java, Spring

MVC, Hibernate, JSP, HTML, CSS, JavaScript, MySQL, JUnit Java Developer JD.com - ??? May 2011 to September 2011 Description: JD.com is a Chinese electronic commerce company. It is the largest B2C online retailers in China, and a major competitor to Alibaba TaoBao. The project is to design and implement different modules including product recommendation and some webpage implementation. Responsibilities: Designed and developed of application using Spring MVC framework with Agile methodology Developed JSP and HTML5 pages using CSS and JavaScript as part of the presentation layer Hibernate framework is used in persistence layer for mapping an object-oriented domain model to database Developed database schema and SQL queries for querying, inserting and managing database Implemented various design patterns in the project such as Data Transfer Object, Data Access Object and Singleton Used Maven scripts to fetch, build, and deploy application to development environment Created RESTful web service interface to Java-based runtime engine Used JUnit for functional and unit testing code Environment: Eclipse, Java, Spring MVC, Hibernate, JSP, HTML, CSS, JavaScript, Maven, RESTful, Oracle, JUnit

Front End Developer Chung How Kitchen - Stony Brook, NY December 2010 to April 2011 Descriptions: Chung How Kitchen is a Chinese restaurant in Stony Brook, NY. Chung How Kitchen managed to display the restaurant information to their customers. This project implemented interactive navigation to the website. Responsibilities: Communicated with clients to clearly define project specifications, plans and layouts. Created layouts and wireframes by using Adobe Dreamweaver Developed User-Interface views using HTML, CSS, Bootstrap Implemented fundamental web functions using JavaScript and JQuery Fixed cross browser compatibility issues for Chrome, Firefox, Safari, and IE Implemented dynamic web applications using AJAX and JSON Developed functional prototypes and iterations for testing. Environment: Eclipse, Adobe Dreamweaver, Java, HTML, CSS, Bootstrap, JavaScript, JQuery, AJAX

Education Master of Science in Computer Science Purdue University - West Lafayette, IN Bachelor of Science in Computer Science Stony Brook University - Stony Brook, NY

Name: Isabella Barker

Email: davidsmith@example.com

Phone: 612-448-2627