**Srivani K**

**Big Data Consultant**

**SUMMARY:**

* 7 years of experience in Designing, Developing Web applications and Big data applications using **Java, J2ee, Oracle and Cloudera Hadoop based Big data technologies**
* Experience in writing Map Reduce programs for both **Hive and pig in java.**
* Experience in data load management, importing and exporting data using **Sqoop and FLUME**.
* Experience in creating Hive Internal/External tables and views using shared metastore, writing scripts in HiveQL, data transformation & file processing using Pig Latin Scripts.
* Knowledge on installation, configuration, support, maintenance of Cloudera’s Hadoop platform along with CDH4&5 clusters.
* Excellent understanding of Hadoop architecture and various components such as HDFS, Job Tracker, Task Tracker, Name node, Data node and MapReduce.
* Trained in Cloudera based Hadoop Distribution.
* Good Working knowledge of Core & Advanced Java, J2EE, Oracle, JavaScript, Struts, spring, Hibernate.
* Experience in application programming using Servlets, EJB’s.
* Designed and develop web based UI application using HTML, CSS, and JSP.
* Well versed in MVC (Model View Controller) architecture using spring, JSF and also implementing JSTL (JSP standard tag library), custom tag development and tiles.
* Good Exposure to Web/application servers such as Apache Tomcat, WebLogic.
* Experience in developing database to create its objects like tables, views, functions, triggers and stored procedures packages using PL/SQL in Oracle.
* Strong analytical and logical ability to work independently or in a team.

**EDUCATION:**

* **Master of Science (MS):** Computer Science, New Jersey Institution of Technology, New Jersey, USA- 2009.
* **Bachelor of Engineering (BE):** Computer Science, India.
* **Certifications:** Oracle Certified Associate – Oracle Database 10g Administrator.

**TECHNICAL SKILLS:**

|  |  |
| --- | --- |
| Big Data Technologies | Hadoop, HDFS, Hive, Pig, HBase, Sqoop, Flume, ZooKeeper, Cloudera CDH4, CDH5, AWS, HiveQL, Pig Latin. |
| Java/J2EE Technologies | JSF, Struts, Servlets, JSP,EJB, Junit and JDBC |
| Programming Languages | C, C++, Java , SQL, PL/SQL, HTML,XML |
| Web Development | HTML5, DHTML, XHTML, CSS, Java Script, AJAX |
| Frameworks | Struts, Hibernate, Spring, JSTL |
| Application/Web Servers | Apache Tomcat, IBM Web Sphere. |
| Databases | Oracle 10g/11g, SQL Server, IBM DB2. |
| Methodologies/Design Patterns | OOAD, OOP, UML , MVC |
| IDEs | Eclipse, Net beans |
| Tools | MSOffice, Microsoft Visio, XML Spy, TOAD |
| Operating Systems | Windows 98/2000/NT/XP, UNIX and LINUX |

**PROFESSIONAL EXPERIENCE:**

**Vaqya Somerset, NJ**

**Hadoop Developer June 2014 to till date**

Vaqya is a voice-based social networking platform that provides social networking opportunities for situations when typing and reading are impossible or inconvenient: like driving, jogging or washing dishes. It allows users to share 30-second voice messages, which are then combined into a smooth uninterrupted stream and enables users to comment in their own voice which carry the true intent and emotion. The followers can choose to follow specific "Squares" for example; a famous personality (Sports, Music, Movie star or a politician etc...) or event/trending news based squares and leave their comments in their own voice if they choose to.

**Responsibilities:**

* Imported logs from web servers with Flume to ingest the data into HDFS.
* Implemented custom interceptors for flume to filter data and defined channel selectors to multiplex the data into different sinks.
* Retrieved data from HDFS into relational databases with Sqoop.
* Parsed cleansed and mined useful and meaningful data in HDFS using Map-Reduce for further analysis
* Fine tuning hive jobs for optimized performance
* Implemented UDFS, UDAFS, UDTFS in java for hive to process the data that can't be performed using Hive inbuilt functions.
* Designed and implemented PIG UDFS for evaluation, filtering, loading and storing of data.
* Developed workflow in Oozie to automate the tasks of loading the data into HDFS and pre-processing with Pig.

**Environment:** Hadoop, Hive, Pig, Sqoop, Oracle10g, HDFS, Oozio, Flume.

**GHP HealthCare Edison, NJ**

**Hadoop Developer Dec 2013-June 2014**

GHP is a healthcare services provider supporting hospitals and clinics and process data of patients and physicians. The project is to analyze and validatethe data from different sources like,hospitals and Physician offices, and make different data sources in HL 7 format. The processed data can be used for further processing by the upstream systems for generating the MIS reports for Business Analysis and decision making.

**Responsibilities:**

* Gathered all the business requirements from the business Partners.
* Wrote Map reduce jobs using various input and output format. Also design custom format as per the business requirements.
* Importing and exporting data from HDFS from Oracle database and vice versa using SQOOP.
* Experience working on processing unstructured data using Pig and Hive.
* Wrote PIG UDFs to perform data cleansing and transforming for ETL activities.
* Wrote HIVE UDF, UDAF, UDTF for Data analysis and Hive table loads.
* Responsible for analyzing and cleansing raw data by performing hive queries and running pig scripts on data.
* Monitored Hadoop cluster job performance and performed capacity planning and managed nodes on Hadoop cluster.
* Involved in converting the data to HL7 Format.
* Developed workflow on Oozie for automating the job flows.

**Environment:** Hadoop, Pig, Sqoop, Oozie, Oracle 10g, Linux OS, MS project

**Tech Democracy** (Client: Citibank, NY) **Edison, NJ**

### Behavior Analysis using Naïve Bayes Algorithm Apr 2013 – Nov 2013

### Java / Hadoop developer

### Description: Naïve Bayes Algorithm is a simple probabilistic classifier based on applying Bayes' Theorem with strong independence assumptions. The application is built to help analyze the data based on the user’s requirement. The result of the application provides the user with web based interface to the probability of the car buyer intention.

**Responsibilities:**

* Involved in Design and Development of technical specification documents using Hadoop technology on Amazon cloud computing (EC, S3).
* Developed an integration wrapper to push computations from local server to the Amazon cloud.
* Developed MapReduce programs to parse the raw data using Clusters.
* Deployed the application in Amazon Web Services (EC2).
* Implemented the Naïve Bayes algorithm to analyze the data according to the user’s requirements.
* Implemented the user’s front end using JSP, HTML, JavaScript, and CSS.

**Environment:** Core Java, JSP, Hadoop, Amazon EC, JavaScript, HTML, CSS, Amazon S3, Eclipse.

**Tech Democracy** (Client: CVS Health, RI ) **Edison, NJ**

**Master Health Apr 2012– Mar2013**

### Java/J2EE Developer

**Description:** A comprehensive package called MASTER HEALTH is concerned with billing module. This module takes care of preparing bills for outpatients and inpatients and produces receipts on payment. And is developed to manage and assist the hospital industry. This integrated system takes care of all front-end and back-end operations that are usually performed in the hospital. MASTER HEALTH is an on-line, integrated, configurable and modular Hospital information system. It works in a multi user environment with graphical user interface using Oracle as backend. Hence doctors and non-technical users can use this system without much training.

**Responsibilities:**

* Extensively involved in gathering requirements, preparing requirements documents.
* Implemented Struts framework.
* Generated User interface Templates using XML/XSL.
* Developed the Java Code using Eclipse as IDE.
* Implemented Action classes, which contain the logic of selecting and using the right Data Access Object (DAO) classes to create and execute, host transaction(s) applicable to current user operation. The request handlers are also responsible for readying up the data received from the host (via DAOs) to meet presentation requirements.
* Used Log4j for logging and debugging.

**Environment:** Java/J2EE, Oracle, Struts, JSP, XML, Eclipse and Apache Tomcat.

**Actavis High hills, NJ**

**Hospital Information Resource Planning System Sep 2010 –Apr2012**

### Java/J2EE Developer

**Description:** HIRPS (Hospital Information Resource Planning System) is the software designed and developed for the automation of hospitals, managing right from registration, admission, diagnosis & treatment to discharge of patients. It manages the activities of every department in a hospital with high-level data security.

**Responsibilities:**

* Involved in Design and development of the application.
* Worked with struts validation and mappings
* Developed JSP pages to meet the requirements.
* Worked with MVC Design patters with Struts.
* Participated in work flow design and framework design.
* Generated User interface Template using JSP.
* Implemented Action classes and Data Access Object (DAO) classes.

**Environment:** Java/J2EE, Struts, JSP, XML, Eclipse, Apache Tomcat.

### Heitman Bethesda, MD

### Real Estate Application Jun 2009 – Aug 2010

### Database developer

### Description: This is a web based real estate application designed on Java technology. The website lets user find a new home, search for listing to buy, sell , rent a house, apartment, flat and search for listing for plot, condo or home. In addition, the application has functionality to register user, upload images and receive comments from the users.

### Responsibilities:

* Involved in all phases of the application development.
* Wrote procedures to call various stored procedures from database.
* Designed the Database.
* Designed/developed tables, views, materialized view, stored procedures, packages and functions.
* Developed various backend application programs, such as Triggers using SQL and PL/SQL language for the management for decision making.
* Wrote the validation XML files and properties file needed for validating user inputs.
* Developed Presentation layer using JSP, HTML and XML.
* Performed unit testing, system testing and user acceptance test.

**Environment**: JSP, Eclipse, Oracle database, PL/SQL, JavaScript, HTML, CSS.