|  |  |
| --- | --- |
| Full Name | Tejaswini Nallagatla |
| Contact Number | 330-245-4908 |
| Current location | Chicago IL |
| Email Address | tejaswinin218@gmail.com |
| Skype ID | tejaswinin218@gmail.com |

**Professional Summary**

* Having around 7+ years of Professional experience in IT Industry, involved in Developing, Implementing and maintenance of various web based applications **using Java, J2EE and Big Data Ecosystems** experience on Windows and Linux environments.
* Over more than 2 years of work experience on Big Data Analytics with hands on experience on writing Spark and Map Reduce jobs on Hadoop Ecosystem including Hive, Pig, Sqoop and Flume.
* Excellent knowledge on Hadoop Architecture and ecosystems such **as HDFS, Job Tracker, Task Tracker, Name Node, Data Node** and Map Reduce programming paradigm.
* Knowledge on installing, configuring and using Hadoop ecosystem and components **like Hadoop Map Reduce, HDFS, HBase, Oozie, Hive, Sqoop, Pig, Zookeeper** and **Flume.**
* Strong understanding of Hadoop internals, different compressions like**AVRO, JSON,** and different file formats.
* Experience in importing and exporting data using Sqoop from HDFS to Relational Database Systems and vice-versa.
* Extending Hive and Pig core functionality by writing custom UDFs.
* Proficiency in Spark using Scala for loading data from the local file systems like HDFS, Amazon S3, Relational and NoSQL databases using Spark SQL, Cassandra and Import data into RDD and Ingesting data from a range of sources using Spark Streaming.
* Extensive experienced in using different Spark modules like Spark transformations, Mlib, Graphx, Streaming andSparkQL.
* Developed **Apache Spark** jobs using **Scala** in test environment for faster data processing and used **Spark SQL** for querying.
* Analyzed large amounts of data sets using **Pig scripts** and **Hive scripts.**
* Exploring with various modules of Spark and working with Data Frames, RDD and Spark.
* Performed **map-side joins** on RDD.
* Experience in ETL operations on Hive to Spark.
* Knowledge in usingNiFito automate the data movement between different Hadoop systems.
* Performed visualizations according to business requirements using visualization tools like Tableau.
* Designed and developed Tableau dashboards, installed and configured Tableau Server on enterprise wide deployments.
* Worked with Core Java and J2EE technologies such as **Servlets, JSP, EJB, JMS, JDBC**, **Threads, Multi-Threading, Collections** and **Exception handling**
* Experienced in developing applications using **Model-View-Controller** (MVC) Architecture and Spring framework.
* Experience in developing and consuming Web Services using **REST, SOAP, XSD, XML, UDDI, JSON** and **WSDL**.
* Experience in **Deploying** web application using application servers **WebLogic, Apache Tomcat, WebSphere** and **JBOSS.**
* Used Version Control tools like GIT, CVS, SVN and Clear Case.
* Good Experience on SDLC (Software Development Life cycle).
* Experienced in coding **SQL, PL/SQL, Procedures/Functions, Triggers and Packages** on database (RDBMS) packages like Oracle.
* Experienced in web development using **HTML**/**HTML5, DHTML, XHTML, CSS/CSS3, JavaScript, Angular JS, Node JS** technologies.

**EDUCATION:**

Bachelors of Technology in Computer Science, JNTU, 2009.

**Technical Skill Set:**

|  |  |
| --- | --- |
| Hadoop/Big data technologies | Spark, Hive, HBase, Sqoop, Pig, MapReduce, YARN, flume, Oozie, Zoo Keeper |
| Java & J2EE Technologies | JDK 1.6, 1.7, JDBC, EJB, Servlets, JSP, JSTL, JSF, JMS, JNDI, JAF, JTA, JPA, JCA, JAAS, JAXR, JAXP, RMI, Multi-Threading, Collections, Generics, Serialization & Exception Handling |
| Programming Languages | Java, C, C++, Python, Scala |
| Amazon Web Services(AWS) | Elastic Map Reduce(EMR), Amazon EC2, Amazon S3 |
| Web Servers | Apache Tomcat, Apache HTTP Server, IBM Http Server |
| Web technologies | XML, HTML, CSS, Java script |
| Testing | JUnit, JProfiler, JMeter, Mockito |
| IDE Tools | Eclipse 4.0, Net Beans 8.1, IntelliJ IDEA IDE, Scala Build Tool (SBT), PL/SQL Developer |
| Databases | Oracle 9i/10g/11g, PLSQL, MySQL, Cassandra |
| Platforms | Unix/ Linux, Windows |
| SDLC Methodologies | Agile Methodology, Waterfall and iterative. |
| Other Tools: | Splunk, Putty, WinSCP, DataLake, Talend, Tableau, GitHub, Apache NiFi |

**PROFESSIONAL EXPERIENCE**

**ROLE: Hadoop/BigData Engineer November 2016 - Current**

**Client: AIG – Wilton, CT**

**Responsibilities:**

* Installed and configured Hadoop **MapReduce, HDFS**, Developed multiple **MapReduce jobs** in java for data cleaning and preprocessing.
* Experience in Hadoop distributed file system **Cloudera**.
* Experienced in managing and reviewing **Hadoop log files.**
* Great familiarity with **Hive joins**& used **HQL** for querying the databases eventually leading to complex **Hive UDFs**.
* Responsible for implementing POC's to migrate iterative map reduce programs into Spark transformations using **Spark** and **Scala**.
* Used **Scala** to write the code for all the use cases in **Spark**and **Spark SQL.**
* Expertise in implementing Spark and Scala application using higher order functions for both batch and interactive analysis requirement.
* Implemented **SPARK batch jobs**.
* Analysis of Problem Domain to find out the optimum solution on Big Data technologies.
* Delivering couple of initiatives on big data analytics and large data processing using Hadoop M/R, HBase, and Hive, Sqoop, HBase.
* Developed and executed shell scripts to automate the jobs.
* Wrote complex **Hive queries** and **UDF**s.
* Worked with Spark core, **Spark Streaming** and **spark SQL** modules of Spark.
* Worked on reading multiple data formats on **HDFS** using **PySpark**.
* Involved in converting Hive/SQL queries into Spark transformations using **Spark RDDs, Python** and **Scala.**
* Exploring with Spark various modules of Spark and working **with Data Frames, RDD** and**Spark Context.**
* Developed a data pipeline using Spark and Hive to ingest, transform and analyzing data.
* Performed **map-side joins** on RDD.
* Performed visualizations per business requirements using custom visualization tool.

**Environment**: Unix/Linux, Cloudera, Apache Spark 1.6.1, Scala 2.11.8(Dynamic Build), Hive, HDFS, YARN, Sqoop, HBase, Oozie, Jenkins, Oracle 11/10g, DB2, MySQL, Shell Scripting, Eclipse, ETL Tool (Tableau), PL/SQL, Java, JSP, JDBC

**ROLE: Hadoop/BigData Engineer February 2015 – October 2016**

**Client: Nike – Irving, TX**

**Responsibilities:**

* Evaluated business requirements and prepared detailed specifications that follow project guidelines required to develop written programs.
* Responsible for building scalable distributed data solutions using **Hadoop.**
* Involved in the loading of structured and unstructured data into **HDFS**.
* Loaded data from **MySQL** to **HDFS** on regular basis using **Sqoop Import/Export**.
* Analysed large amounts of data sets to determine optimal way to aggregate and report on it.
* Involved in requirement and design phase to implement Streaming **Lambda Architecture** to use real time streaming using **Spark** and **Kafka**.
* Exported analyzed data to relational databases using SQOOP for visualization to generate reports for the BI Team. Writing Pig Latin scripts to process the data and written UDF in java and python for both Hive and Pig.
* Developed **Spark-SQL** statements for processing data. Worked on **SPARK** engine creating batch jobs with incremental load through **STORM, KAFKA, SPLUNK, FLUME, HDFS/S3, KINESIS, Sockets** etc.
* Installed and configured Hadoop using **Cloudera Distribution**.
* Used Hive Queries in**Spark-SQL** for analysis and processing the data.
* Developed **Junit tests** for testing MapReduce and performed testing using small sample data.
* Created MapReduce jobs which were used in processing survey data and log data stored in HDFS.
* Used **Pig Scripts** for data cleaning and data pre-processing.
* Used MapReduce jobs and pig scripts.
* Experienced in fixing errors by using debug mode of **Talend.**
* Experienced in managing and reviewing the **Hadoop log files**.
* Migrated **ETL** jobs to **Pig** scripts do Transformations, even joins and some pre-aggregations before storing the data onto HDFS.
* Implemented external tables to store the processed results using **HIVE**.
* Involved in developing **Hive UDFs** and reused in some other requirements.
* Worked on performing Join operations.
* Developed the **Sqoop** scripts to export data from **HDFS** and **MySQL** Database.
* Written **HIVE Queries** for analysing data in **Hive warehouse** using **Hive Query Language** (HQL).
* Designed and implemented a data analytics engine based on Scala to provide trend analysis, regression analysis and **machine learning predictions** as web services for survey data.
* Expertise in programming using **Scala,** built Scala prototype for the application requirement and focused on types of **functional Scala**.
* Involved in converting **Hive/SQL** queries into **Spark** transformations using **Spark RDD** and **Scala.**

**Environment**: Cloudera, Apache Spark 1.2.1, Scala, Hive, HDFS, YARN, Sqoop, HBase, Kafka, Flume, DB2, MySQL, Eclipse, ETL Tool (Informatica), PL/SQL, Java, Python, JSP, JDBC

**ROLE: Hadoop Developer September 2013 – January 2015**

**Client: First Data- Austin, TX**

**Responsibilities**:

* Cloudera Hadoop installation and configuration of multiple nodes using **Cloudera Manager** and **CDH 4.X/5. X.**
* Designed documents and estimated efforts for the project.
* Developed Map Reduce Programs using **MRv1** and **MRv2** (**YARN**).
* Responsible for processing unstructured data using Pig and Hive.
* Developed **Pig Latin scripts** for extracting data.
* Used Pig for data loading, filtering and storing the data.
* Developed **HIVE queries** for the analysts.
* Extensive experience in Spark Streaming (version 1.5.2) through core Spark API running Python, Scala, Java & Python Scripts to transform raw data from several data sources into forming baseline data.
* Developed UDFs in Scala, Java & Python as and when necessary to use in PIG and HIVE queries
* Developed Java code to stream the Packet tracer data into Hive using rest full services.
* Worked on migrating data from **Mongo DB** to Hadoop.
* Worked on integrating **SFDC** with Hadoop.
* Extracted the data from **MySQL** into **HDFS** using **Sqoop**.
* Involved in running Hadoop jobs for processing millions of records of text data for batch and online   
  processes by using Tuned/Modified SQL.
* Designed and published workbooks and dashboards using **Tableau Dashboard**/Server 6.X/7.X

**Environment**: Cloudera, Hadoop(HDFS), Map Reduce, Spark, Hive, Java, Scala, JDK, UNIX Shell Scripting, MySQL, Eclipse, Tableau 8.X/9. X.

**ROLE: Java/J2EE Developer April 2012 – August 2013**

**Client: Bank Of America – Charlotte, NC**

**Responsibilities:**

* Design and review **class diagrams** and **state diagrams**
* Followed **n-tier** architecture to develop the web application.
* Done requirement analysis in design phase
* Extensively used design patterns like **Singleton, Prototype, Factory.**
* Used Hybrid **waterfall** methodology to develop the project.
* Worked with **SQL** to update/modify the database.
* Design and implemented Enterprise Integration Application solutions to the clients.
* Designed the front end using **JSP, JavaScript, Angular JS, HTML** and **CSS**.
* Experience working with **Hibernat**e Template of Spring Framework and Hibernate Interceptors.
* Redesigned all **hibernate entity** classes from **XML** files and used **Hibernate annotations**.
* Designed **REST API** for effective, low cost application integration.
* Supported the application in the production phase and fixed any bugs during that phase.
* Developed **Web services** using **REST** based and **SOAP** calls.
* Used **Camel ESB** for integrating with different platforms.
* Developed experience with the **Hibernate Query/Criteria Language** to execute the database operations.
* Applied **MVC** design patterns using JavaBeans and conducted simultaneous queries and retrievals using **Java multi-threading** techniques.
* Involved in **code reviews** with peers.
* Involved in testing using **JUNIT** for Unit Testing.

**Environment:** Groovy, Grails2.1.0, SpringSourceTool, Collections, GORM, JavaScript, HTML, MySQL, hibernate, Ajax, JQuery, Tomcat6.0, Agile.

**ROLE: Java/J2EE Developer October 2010 – February 2012**

**Client: Idea labs Pvt.Ltd - India**

**Responsibilities:**

* Involved in Analysis, Design, Development, Unit Testing and Load Testing of the Application.
* Performed Code Reviews and responsible for Design, Code and Test signoff.
* Assisting the team in development, clarifying on design issues and fixing the issues.
* Involved in designing test plans, test cases and overall Unit and Integration testing of system.
* Development of the logic for the Business tier using Session Beans (Stateful and Stateless).
* Built data-driven Web applications with server side J2EE technologies like Servlets/JSP. Generated dynamic Web pages with Java Server Pages (JSP).
* Used Struts **MVC Framework, CSS, DHTML, XHTML and HTML** for developing UI Screens.
* Utilized **PL/SQL** for stored procedures.
* Developed **Eclipse** Rich Client Platform framework for creating *Java* client-side applications.
* Developed **Hibernate** Mapping files and Domain objects.
* The GUI was designed on the base of **MVC** design-patterns and **Swing APIs** are used extensively.
* Developed GUI using, Swing for Customer entry form and result form.
* Developed **JUnit** Test Cases for Service Layer and **DAO**s.
* Implemented Oracle Queries and **Triggers**using **SQL Server, SQL** and **PL/SQL**.

**Environment:** Java, J2EE, Analysis, Design, Development, Unit Testing, Servlets, JSP, Struts, MVC, JavaScript, CSS, HTML, XHTML, DHTML, Hibernate, PL/SQL, Eclipse, SQL, SQL Server.

**ROLE: Associate Java/J2EE Developer January 2010 – September 2010**

**Client: SattvaQ IT Solutions - India**

**Responsibilities:**

* Involved in all phases of **Software Development Lifecycle** (SDLC) of the project.
* Involved in Design and development of UI using **JSP, HTML** and **JavaScrip**t.
* Hands on **Informatica** for ETL process. Involved in Analysis, Design, Development, Unit Testing and Load Testing of the Application.
* Built data-driven Web applications with server side J2EE technologies like **Servlets/JSP**. Generated dynamic Web pages with Java Server Pages (JSP).
* Developed front-end screens using **JSP, HTML** and **CSS**.
* Developed server-side code using **Struts** and **Servlets**.
* Developed core java classes for exceptions, utility classes, business delegate, and test cases.
* Developed SQL queries using MySQL and established connectivity.
* Worked with **Eclipse** using **Maven** plugin for Eclipse IDE.
* Wrote **Client-Side validations** using **JavaScript**.
* Extensively used **JQuery** for developing interactive web pages.
* Developed the user interface presentation screens using HTML, XML, and CSS.
* Developed the **Shell scripts** to trigger the **Java Batch job**, Sending summary email for the batch job status and processing summary.
* Co-ordinate with the QA lead for development of test plan, test cases, test code and actual testing responsible for defects allocation and those defects are resolved.
* Application was developed in **Eclipse IDE** and was deployed on **Tomcat server**.
* Involved in Agile scrum methodology.
* Supported for bug fixes and functionality change.

**Environment:** Java/J2EE, Oracle 10g, SQL, PL/SQL, JSP, Hibernate, WebLogic 8.0, HTML, AJAX, Java Script, JDBC, XML, UML, JUnit, Eclipse.