KESHAV SAINI



**Phone**: (+91) 9990742008

**E-mail**: keshavkrsaini@gmail.com



**Career Summary**

Cloudera Spark and Hadoop Certified Developer with 3.5 years of experience in Spark, Sqoop, Kafka, Hive, Oozie, SQL Server and SSRS*.*Skilled data analyst with experience in migration, integration and analysis of data in big data environments.



**PROFESSIONAL EXPERIENCE and RESPONSIBILTIES**

**Yes Bank Ltd. Jan 2017 - Present**

**Big Data Developer:**

Analysis of transaction, non-transaction and relationship data. Transfer the traditional RDBMS to Hadoop clusterware and implement analytical tools to provide reports to the bank. The responsibilities are detailed below -:

* Performing analytics using Big Data technologies like Hadoop, Spark.
* Responsible for storing and maintaining data in Sql and No-sql databases like Oracle, Hbase.
* Implementation of Hadoop cluster ware.
* Manipulating and storing data on huge cluster ware using Spark , Spark SQL and Mapreduce.
* Creating scripts for high level data extraction using Pig and Hive scripts.
* Usage of Sqoop, Flume for transportation of data between relational database and No-sql databases.
* Using Spark Machine learning for Predictive analysis.
* Using Kafka and spark streaming for real time data transfer and processing.

**Tools** : Hadoop 2.7.1, Spark Framework, Python, Core Java, Pig, Hive, Hbase, Sqoop, Flume, Mapreduce, Machine Learning, Kafka, Spark streaming, Spark SQL, Oracle and Clusterware Adminstration.

**Toluna India Pvt. Ltd August 2014– Jan 2017**

**Database and Enterprise Architect :**Responsible for the architecture and design of the data storage tier for data warehousing and analysis. The responsibilities are detailed below -:

* Perform analytics on Big data using Apache Spark and Mapreduce.
* Maintenance of Sql and No-Sql databases like Mysql, Sql Server, Hbase.
* Using ETL technologies for data extaction.
* Creating scripts for high level data extraction using Pig and Hive scripts.
* Usage of Sqoop, Flume for transportation of data between relational database and No-sql databases.
* Using visualization tools to develop reports for bussiness

**Tools**: Hadoop 2.7.1, Apache Spark, Hive, Pig, Sqoop, Mysql, Sql Server, Tableau, SSRS.

**eFleet Systems Pvt. Ltd Jun 2013– August 2014**

**Software Developer:**

Responsible for developing electronic management information system for supply chain management(eMist). Activities performed include data analysis using SQL and Visualization using SSRS.



**Technical Skill Sets**

* **Big Data Platforms**: Cloudera, Open Source
* **Big Data Glossary**: Sqoop, Hive, Oozie, Spark,Kafka,Hbase.
* **Languages**: Python, Hql and SQL
* **Visualization** : Tableau, SSRS
* **Databases**: MySQL, Sql Server
* **Platforms:** Linux, Windows



**Certifications**

* **Cloudera Certified Spark and Hadoop Developer** (**Lic. No: 100-016-594**).
* **Edureka Certified Hadoop Developer.**



**EDUCATIONAL CREDENTIALS**

|  |
| --- |
|  |
| B. Tech (CSE) MDU 2009-2013 Agg. 63% |
|  |
| Intermediate (PCM) CBSE 2009 Agg. 82% |
|  |
| High School CBSE 2007 Agg. 74% |

**DECLARATION**

I hereby declare that the above-mentioned information is correct up to the best of my knowledge and belief.

**Place**: Gurgaon(India)

**Date** : **(KESHAV SAINI)**