KARIMULLA SHA SHAIK   +91-8919179781
                                            19kshah@gmail.com

Having 5+ years of experience in IT out of which 2+ years into Hadoop and Spark

EXPERTISE INCLUDES:

- ➤ I have 5+ years of experience working in IT operations out of which 2+ years into Hadoop and Spark Ecosystems.
- ➤ 2+ years of Experience in Ingestion, Storage, Querying, Processing and Analysis of Big Data.
- ➤ Hands on experience on Hadoop Ecosystem Development including HDFS, Sqoop, Hive, YARN, NoSQL, MongoDB, Cassandra and HBase.
- ➤ In depth understanding of HDFS architecture.
- ➤ Extensive experience in importing/exporting data from/to RDBMS, the Hadoop Ecosystem using Apache Sqoop.
- ➤ Good experience in analyzing data using Hive QL.
- ➤ Good experience over Hive functionalities by writing custom UFDs.
- ➤ Good Experience in writing SQL queries.
- ➤ Hands on experience in Linux Commands.
- ➤ Performed import from multiple tables using joins from Sqoop to HDFS with various file formats.
- ➤ Performed optimizations in hive, joining tables like Map side join, Bucket join.
- ➤ Experienced with Apache SPARK Scala & Python both programming by improving the performance and optimization of the existing algorithms in Hadoop using Spark Session, Spark Context, Spark SQL, Data Frame and Pair RDD's.
- ➤ Involved in creating Broadcasting & Accumulators in Spark jobs.
- ➤ Experience in handling different file formats like Parquet and Sequence file format, csv files.
- ➤ Sound technical Knowledge on Spark streaming with Python, Kafka, Zookeeper and Nifi.
- ➤ Involved in writing the Spark and Hive APIs to reduce the job execution time.
- ➤ Knowledge on Cassandra.
- ➤ Experience with various IDE's for development of project (InteliJ-idea, PyCharm).
- ➤ Experience in writing programs in PyCharm using different libraries like arcPy, Pandas, numPy etc.
- ➤ Excellent communication, interpersonal, analytical skills and strong ability to perform as part of team
- ➤ Zeal in learning new Technologies.
- ➤ Hard working and enthusiastic.

TECNICAL SKILLS:

- ➤ HDFS, YARN, Sqoop, Hive, NoSQL, MongoDB, Cassandra and HBase.
- ➤ Apache SPARK (Python & Scala)
- ➤ Python, Scala, Kafka, Nifi and Hive
- ➤ InteliJ-idea, PyCharm
- ➤ Cloudera CentOS, Ubuntu 16.4

TECHNICAL EXPERTISE:

Working with SSSBI Pvt.Ltd., Tirupati from 1st April 2017 to till date

PFIZER Inc. American pharmaceutical corporation	
Project Name	PFIZER Inc.
Duration	January 2018 to till Date
Technology	Hadoop, Hive, SQOOP, SPARK, SCALA, MySql Server, Cloudera, CentOS, Ubuntu 16.4
Role	Hadoop and Spark Developer
Responsibilities	➤ Expertise in designing and deployment of Hadoop cluster and different Big Data analytic tools including Hive, Sqoop, Apache Spark, with Cloudera Distribution. ➤ Involved in Data extraction, Transformation and loading from source to target system. ➤ Involved in loading and transforming large sets of structured, semi-structured and unstructured data and analysed them by running Hive queries and Spark Jobs with python. ➤ Used Sqoop to dump data from relational database into HDFS. ➤ Involved in creating Hive tables, loading with data and writing hive queries, joining tables, writing HQL script for batch of queries. ➤ Responsible for developing data pipeline using Sqoop to extract the data from weblogs and Store in HDFS. ➤ Experienced with batch processing of data sources using Apache Spark. ➤ Experienced in implementing Spark RDD transformations, actions to implement business analysis. ➤ Implemented partitioning, dynamic partitions and buckets in HIVE.

MAERSK A/S, transport, logistics and energy sectors.	
Project Name	Maersk
Duration	April 2017 to December 2017
Technology	Hadoop, Spark, Scala, Hive, SQOOP, Linux and MySQL
Role	Hadoop Developer
Responsibilities	➤ Day to Day activity involves importing and exporting data from RDBMS to HDFS using Sqoop. ➤ Writing the script files for processing data and loading to HDFS ➤ Involved in creating the Maven project and configuring the required dependencies. ➤ Created Hive External tables to avoid data lose. ➤ Loaded and transformed large sets of structured, semi-structured and unstructured data in various formats like text, JSON, Parquet ➤ Installing and monitoring of Hadoop and Spark Stacks in Linux Environment

GIS Engineer   1st March, 2014 to 31st March, 2017

Govt. of AP Forest Dept, Hyderabad | AP

Responsibilities:

- ✓ Developed various modules under APFMIS (A.P. Forest Management Information System) http://www.fmis.ap.gov.in/

- ✓ Manage Department Dashboard by showing Planting, Nursery, Seed dibbling etc., activities

- ✓ Finalizing actions in Modules like NCMIS (Neeru-Chettu MIS), PMIS (Plantation MIS), FPMIS (Forest Protection & Fire Management)

- ✓ **WBI Clearance:** Wood Based Industry clearance is ArcGIS based module, we developed some Python tools for this, as per Supreme Court orders clearance for WBI is more than 5KMs crow-fly distance from the nearest RF Boundary. Taken the input location of WBI as Lat-Long (geo-coordinates) and given decision whether it is acceptable or not. Output gives nearest RF Name, Compt.No, distance from the RF Boundary.

Environment: Python programming, ASP.Net, Visual Studio, Java Script, C++, SQL, Server, ArcGIS, ArcMap 10.1, Windows & Linux

Education Details:

- ➢ **M.Tech(CSE)** from JNTU Anantapur with 75.16 %, in 2013.
- ➢ **M.Sc (C.S)** from S.K.University, Anantapur, with 76.97 %, in 2009.
- ➢ **B.Sc(C.S)** from S.K.University, Anantapur, with 62.94 %, in 2004.
- ➢ **Intermediate** from A.P. State Board of Intermediate with 60.30 %, in 2001
- ➢ **SSC** from A.P. State Board with 60.16 %, in 1999

Personal Details:

➢ **Date of Birth**	:	19-08-1984
➢ **Gender**	:	Male
➢ **Nationality**	:	Indian
➢ **Marital Status**	:	Married
➢ **Passport No**	:	L2248092
➢ **Languages Known**	:	English, Hindi, Telugu & Urdu.
➢ **Mobile**	:	+91 8919179781
➢ **e-mail**	:	19kshah@gmail.com

I hereby declare that the information that I have furnished is authentic, and true to the best of my knowledge.

(KARIMULLA SHA SHAIK)