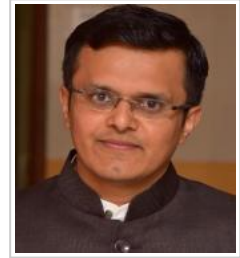


Hitesh Sharma  

[Immediate Joiner] Senior Data Engineer seeking roles in Big Data Development, Agile Methodology, Currently working on Different Big Data component like Apache Spark ,Hadoop,Scala ,Hive, Apache Airflow, PostgreSQL, Python,Amazon S3

Current Designation: Senior Data Engineer

Total Experience: 8 Year(s) 0 Month(s)

Current Company: gNAP Service Pvt. Ltd.

Notice Period: 1 Month

Current Location: Noida

Highest Degree: MCA [Computers]

Pref. Location: Delhi / NCR,Pune,Noida

Functional Area: IT Software - Application Programming / Maintenance

Role: Software Developer

Industry: IT-Software/Software Services

Marital Status: Single/unmarried

Key Skills: **Apache Spark , SCALA, Hive, Apache Airflow, Postgresql, Amazon S3 , Python**Verified : **Phone Number | Email - id**Last Active: **Jan-Mar 2021**Last Modified: **Jan-Mar 2021**

Summary

Over all 8 years of experience in Design and Development of Software Applications including 4 year of experience as Spark developer using Big Data Technology like Spark Ecosystems.

Proven Expertise in performing analytics on Big Data using Spark, Postgres, Elastic, Hive.

Understanding of spark architecture including Spark Core, Spark-SQL, Data Frame.

Expertise Spark-SQL with various data source like Parquet, CSV.

Exposure to Agile software development.

Microsoft Certified ASP.Net 3.5 Application Developer

Hands on experience in all phases of Software Development Life Cycle (SDLC).

Have working experience in Scala, C#, .NET, Visual Studio 2015, ASP.Net 4.5

Work Experience

gNAP Service Pvt. Ltd. as Senior Data Engineer

May 2020 to Till Date

Client Name Nokia Communication Pvt. Ltd.

Project Name CSDP

Project Detail Cell site degradation Predication (CSDP) is for UAE. In CSDP we just predicate that how the network will work in future or at any specific event based on cell site information in CSDP we are covering multiple network technologies like 2g, 3g, 4g and, we are running our jobs for offline and online mode. We received the cell site information from Datahub team for each technology like 2g, 3g, 4g. Once we received different input file then we perform some transformation and action on these file and store out put into S3 and Postgres for DS and UI team so that they can predicate and make dashboard for better understanding. We are using airflow for job automation for which we are using python for creating our dags.

Role Involved with multiple team like data science and UI team. We Processed large set of data coming through Datahub team in various format on daily basis. After performing some transformation and action on data we store it into S3 and Postgres so that data science team, UI development team can you use it for future network prediction and data visualization purpose.

Platform Scala, Python, Spark Core, Spark SQL, Postgres, Airflow, Kubernetes, Gitlab

R1 RCM Global Pvt. Ltd. as Senior Software Engineer

Apr 2015 to May 2020

Product Name RCM Reporting Tool

Project Detail Initially we received data via HL7 in various format. Previously there was SSIS package which process the data and store it into RDBMS. After that reporting team and other product team use that data. But now we are converting SSIS implementation into spark application using Scala for fast processing. After data analysis, we store it into various data store system as per requirement. Once we process data, the data science team and other teams work on that data.

Role Involved with data analysis and data science team. Processed large set of data coming via HL7 in various format on daily basis. Based on the requirement of reporting team performs various Spark-SQL operation on large data set. After analyzing data store into HBase so that data science team, tableau team and web development team can you use it for patient denial and data visualization purpose. Using Scala for writing spark application.

Platform Apache Spark ,Scala, Spark Core, Spark SQL, Hive

Pyramid IT Consulting Solutions as Software Engineer
Mar 2014 to Apr 2015

Project Title DMS

Project Details This project is for Think gate, it's a US based company. Document Management System (DMS) is the product, which is used by Think gate's client. In DMS there are two teams, DI team and UI team, I was in UI team. Think gate clients send different file to think gate or some time they put these file on think gate server. So, In UI part we manage their client information through UI and DI team maintain file data as per client rules.

Role Development

Platform MVC, Entity framework, LINQ, jQuery, SQL 2008, Kendo Control

Education

UG: **B.Sc (p.c.m.)** from **Dr. Bhim Rao Ambedkar University (BRAU), Bihar** in **2006**

PG: **MCA (Computers)** from **G.K.V.HARIDWAR** in **2007**

IT Skills

Skill Name	Version	Last Used	Experience
Spark Core, Spark-SQL, Hadoop, HDFS		2020	4 Year(s)
HIVE		2020	4 Year(s)
VM Virtual Box		2020	3 Year(s)
AWS Config and S3		2020	2 Year(s)
Asp.net 4.5, Ajax, HTML		2016	4 Year(s)
MVC 5.0, Entity framework		2016	4 Year(s)
Web API		2016	4 Year(s)
Scala, C#, VB.Net		2020	4 Year(s)
Python		2020	1 Year(s)
Pycharm, IntelliJ		2020	3 Year(s)
Microsoft Visual Studio 2017/19		2016	4 Year(s)
IIS		2016	4 Year(s)
TFS, SVN, Bit Bucket, Git Labe, JIRA		2020	7 Year(s)

Languages Known

Language	Proficiency	Read	Write	Speak
HINDI	Expert			
ENGLISH	Proficient			

Projects

Project Title: CSDP(Cell site degradation Predication)

Client: Nokia Communication

Nature of Employment: Full Time

Project Location: Noida

Role: Sr. Programmer

Duration: May 2020 - Till Date

Onsite / Offsite: Offsite

Team Size: 4

Skill Used: Spark , SCALA, Airflow ,Amazon S3 , Postgresql, Pytohn, Kubernetes

Role Description: Role Involved with multiple team like data science and UI team. We Processed large set of data coming through Datahub team in various format on daily basis. After performing some transformation and action on data we store it into S3 and Postgres .

Project Details: Project Detail Cell site degradation Predication (CSDP) is for UAE. In CSDP we just predicate that how the network will work in future or at any specific event based on cell site information in CSDP we are covering multiple network technologies like 2g, 3g, 4g and, we are running our jobs for offline and online mode. We received the cell site information from Datahub team for each technology like 2g, 3g, 4g. Once we received different input file then we perform some transformation and action on these file and store out put into S3 and Postgres for DS and UI team so that they can predicate and make dashboard for better understanding. We are using airflow for job automation for which we are using python for creating our dags.

Project Title: R1 Reporting Tool

Client: US Hospital

Nature of Employment: Full Time

Project Location: Noida

Role: Sr. Programmer

Duration: Aug 2016 - May 2020

Onsite / Offsite: Offsite

Team Size: 6

Skill Used: Apache Spark ,Scala, Spark Core, Spark SQL, Hive

Role Description: Working as developer

Project Details: roject Detail Initially we received data via HL7 in various format. Previously there was SSIS package which process the data and store it into RDBMS. After that reporting team and other product team use that data. But now we are converting SSIS implementation into spark application using Scala for fast processing. After data analysis, we store it into various data store system as per requirement. Once we process data, the data science team and other teams work on that data.

Role Involved with data analysis and data science team. Processed large set of data coming via HL7 in various format on daily basis. Based on the requirement of reporting team performs various Spark-SQL operation on large data set. After analyzing data store into HBase so that data science team, tableau team and web development team can you use it for patient denial and data visualization purpose. Using Scala for writing spark application.

Affirmative Action

Category: General

Physically Challenged: No

Work Authorization

Job Type: Permanent

Employment Status: Full time