HARSHIT SAXENA

Data Engineer: AWS Certified Developer

Email – harshitprog05@gmail.com   Contact- 08224046433   LinkedIn-https://www.linkedin.com/in/harshit-saxena-a083b3176

SUMMARY

3 years IT experienced and result-oriented Apache Spark & Scala developer possessing a proven track record in software development and data pipelines management using Apache Spark, Kafka, Slick, Scala, Play Framework, HDFS, Vertica, MySQL and Apache Ignite. Certified AWS developer and proficient in processing structured and unstructured data & developing data pipelines using Apache Spark to analyze huge data set, identify pattern & gain valuable insight. Capability in accomplishing product life cycle management in agile methodology. Having skills of Spark cluster management using Kubernetes and capacity of planning for end-to-end data management and performance optimization. Experienced in all phases of product development from planning and requirement gathering to implementation and testing and also having an experience as an business Intelligence automation engineer.

WORK EXPERIENCE

PROJECT: *Lumenore Data Layer (Product)*   *[Oct'18- Present]*
Organization: Netlink Software Group America Inc.

An ETL Data Processing web based tool which has the capability to fetch data from multiple sources to the data warehouse, in this tool we are developing graphs and Jobs which consist of multiple transformations. This is tool is working as an ETL data pipeline for processing the data from 180+ data sources.

ROLE   : *Spark/Scala Developer*
TECHNOLOGIES: *Scala, Spark, Data frames, play framework, Rest API's, Kubernetes, Vertica, SBT, Spark SQL, RDD and MYSQL*

- Leading, Managing, Planning and Designing the whole architecture of the product and involved in development of various types of transformation to aid a developer/user to perform data manipulation, data cleansing activities.
- Prepared an organized data to provide proficient and valuable insights to customer.
- Planned and designed the implementation and integration of data pipeline.
- Performed a POC by setting up spark cluster on EC2 in AWS.
- Overcame the challenges of storing & processing the structure, streaming and Unstructured data with the help of Spark, Kafka and Scala.
- Configured various data connectors to fetch data from multiple sources including S3.
- Also involved in setting up a spark-cluster using kubernetes resource manager.

MANAGEMENT ROLE:

- Leading the team of 7 members to ingest the data from structured and unstructured source into Data warehouse.
- Effort estimation and planning of task.

OBJECTIVE

Highly analytical individual with strong apt for learning and collaborative skills. Looking to obtain a Data Scientist/Data Engineer position, to utilize prior data scientist experience. Offering expertise in Scala, Spark, HDFS, Kafka, Apache Ignite, Kubernetes, Rest API's and analyzing large complex data set using a variety of tools.

KEY TECHNICAL SKILLS

SCALA

APACHE SPARK

BIG DATA ANALYTICS

BUSINESS INTELLIGENCE

HDFS

KAFKA

PLAY FRAMEWORK

KUBERNETES

REST API

RDBMS

DATA PROCESSING

DATA PIPELINES

SOFT SKILLS

PROBLEM SOLVING

CRITICAL THINKING

CREATIVE THINKING

INTERPERSONAL SKILLS

PROJECT: *Self Service (Product)*   *[June'19 - Present]*
Organization: Netlink Software Group America Inc.

Business Intelligence tool to provide data analytics that enable business users to access and work on their data even if they do not have a background knowledge of statistical analysis and BI.
ROLE   : *Lead Architect and Developer*
TECHNOLOGIES: *Spark, Scala, Apache ignite , Kafka , ORC, HDFS, Vertica, MYSQL, Play framework, SBT, Mongo DB, RDD and Spark SQL*

- Developed various data connectors and provide connectivity.
- Performed multiple data cleansing activities.
- Responsible for maintaining and developing the architecture of this data pipeline.
- Developed a governance module and also defined the data model for this module.
- Responsible for loading customer data in proficient manner to the database.

PROJECT: *SXM TAP*   *[June'17-Jan'18]*
Organization: Netlink Software Group America Inc.

SXM training and Audit Plus provides detail insights of trainer performance and reporting in the dashboard with respect to different key performance index
ROLE   : *Automation Engineer*
TECHNOLOGIES: Selenium Web driver, HTML, Java Script, MySQL, Java

- Responsible for analyzing the data.
- Developed a user automation module to perform unit testing on various dashboards.
- Developed automation scripts with the help of selenium web driver.
- Knowledge about waterfall methodology.

ACADEMICS
Bachelors of Engineering (CSE) (73.60%) Shri Ram institute of Information Technology.
[2013-2017]   Gwalior, M.P
Class 12TH (70.6%) *M.P BOARD 2013*
Class 10TH (73.1%) *M.P BOARD 2011*

ACHIEVEMNTS & CERTICATIONS
AWS Certified Developer (9HTXZ3Y2C1B1QF9M)
Vertica Technical Certification
CMMI Level 3 Process Owner
Rising Star for Q-2
Appreciation certification as an innovator in Q-4 2019

PERSONAL INFORMATION
DOB – Sept 09th, 1996
S/O – Mr. Alok Saxena & Mrs. Yojna Saxena
Permanent Address – DD Nagar Gwalior
Marital Status – Unmarried