+91 8124884686 | danveer4686@gmail.com | https://danveer4686.github.io/portfolio

Danveer Sharma

Sr Data Engineer || 7 years of exp || Spark, Scala, GCP, Big Data

CAREER SUMMARY

Aug 2019 – Present

SENIOR DATA ENGINEER – Star Tv Network (3 years)
PROJECT: DATA DRIVEN APPLICATIONS for Star sales team
OneView and MINT are internal products of Star which make organization's revenue and viewership KPIs available for various teams. I am **leading backend team** in building **scalable datalake** solutions as well as **graphql APIs** Apps to support MINT and OneView.
RESPONSIBILITIES:

- Design and development of Datalake-App which involves brewing **spark-scala** with **GCP** and **AWS** services (BigQuery, GCS, Cloud function, Postgres) to bolster star-datalake.
- Development of Middleware-App which makes best use of **scala's functional paradigm** in creation of flexible graphql-apis to make datalake available for front-end Apps.
- Job scheduling using **Airflow** Dags in **Python.**
- Cloud functions for etl pipeline automation.
- **Kubernetes** deployment for scalable applications.
- Product **performance tuning** using different file formats (**Avro, ORC**), **resource allocation**, **joins, actions, transformations**, **Spark encoders and Spark listener.**

Sep 2018 – Aug 2019

SENIOR SOFTWARE DEVELOPER – Sterlite tech ltd (1 year)
PROJECT: PRODUCT DEVELOPMENT for telecom clients
INTELLZA is an under development proprietary product of Sterlite. Product involves brewing spark-scala for data ingestion (inputs from **HDFS, Kafka, Hive, RDBMS**), data refinery (various transformations like **filter, merge, correlation, enrichment, lookups and aggregator**) and data store (**HDFS, Kafka, Hive, RDBMS, Hbase**) components. All components are utilized in creation of data pipelines architecture based on json format. Data pipelines are scheduled using **oozie workflow and coordinator** for batch and streaming jobs.
RESPONSIBILITIES:

- Development of various data refinery components like **filter, merge, correlation, enrichment, lookups and aggregator.**
- Development of data storage components for storing detailed data in **Hive, HDFS, Hbase, Kafka** and summarized data in **RDBMS.**
- Job scheduling for batch and streaming using **Oozie workflow** and coordinator.
- Solving **use cases for telecom clients** using data pipelines, providing insights on **Tableau/Grafana** visualisation dashboard and Using **ELK stack** for logs aggregation and visualisation.
- Developed **Unix shell scripts** for end users to use the product in easier manner.

July 2015 – Sep 2018

BIG DATA ENGINEER – Infosys Ltd. (3 years)
PROJECT: SPARK SCALA DEVELOPMENT for a German Bank client
Modernisation of VES (Validation Enrichment & Services) rules is a decentralised **FDW** project of banking client. Project involves BRE (Business Rules Extraction) from **Informatica ETL mappings** and development of VES rules in **Spark, Scala**. Enriched data then is loaded into Hive tables. Functionality of spark scala has been tested by validating **Hive tables** against **Oracle** tables of staging area.

SKILLS

Languages	**SPARK API (RDD, DF AND DS)**	★★★★★★★☆☆☆
	SCALA 2.13.X	★★★★★★☆☆☆☆
	SQL & HIVE QL	★★★★★★★☆☆☆
	JAVA CORE	★★★★★★☆☆☆☆
	UNIX SHELL SCRIPTING	★★★★☆☆☆☆☆☆
Framework	**HADOOP 2.X**	★★★★★★☆☆☆★
	HDFS, YARN, MAPREDUCE	★★★★★★☆☆☆☆
Database/DWH	**BIGQUERY**	★★★★★★★☆☆☆
	POSTGRES	★★★★★★★☆☆☆
	HBASE, KAFKA	★★★★☆☆☆☆☆☆
Cloud Services	**GCS, S3**	★★★★★☆☆☆☆☆
	CLOUD FUNCTION	★★★★★☆☆☆☆☆
	APACHE AIRFLOW	★★★★★☆☆☆☆☆
	CLOUD BUILD	★★★★★☆☆☆☆☆
	DOCKER & KUBERNETES	★★★★★☆☆☆☆☆

EDUCATION

2011 - 2015	**B.TECH MECHANICAL** – Vellore Institute of Technology, Vellore (8.04)
2010	**HIGHER SECONDARY EDU.** – Govt. Sr. Sec. School, Bayana, Rajasthan (81.23)
2008	**SECONDARY EDU.** – Adarsh Vidhya Mandir, Bayana, Rajasthan (84.33)

PERSONAL DOSSIER

- **DATE OF BIRTH** :15-AUG-1992
- **GENDER** :MALE
- **LANGUAGES KNOWN** :ENGLISH, HINDI
- **PLACE** :MUMBAI(MH), INDIA