

# Akanksha Gupta

## Big Data Engineer

Having 3.8+ years Experience in data analysis, data processing and Development in Big Data technologies.

gargakanksha49@gmail.com

9997664034

Saharanpur, India

## WORK EXPERIENCE

### Data Engineer

#### Blackstraw AI

12/2019 - Present

Chennai, India

##### Achievements/Tasks

- Developing and maintaining Spark-Scala code for ETL from Azure BLOB to Azure Synapse Analytics using Databricks.
- Configured the Pipeline for ingesting data from Azure Blob to Azure Synapse Analytics using spark.
- Monitoring and optimizing performance of spark and azure synapse analytics.
- Works on POC for replacement of Snowflake over Azure Synapse Analytic.

### Software Developer

#### Avancer Software Solution Pvt. Ltd

06/2018 - 06/2019

Chennai, India

##### Achievements/Tasks

- Developed Spark core and Spark SQL scripts using Scala for faster data processing.
- Involved in converting Hive/SQL queries into Spark transformation using Spark RDD's.
- Working on Hadoop 1.x/2.x cluster architecture and Monitoring the cluster.
- Working with Hadoop Clusters using Hortonwork(HDP) distribution.

### Software Engineer

#### iSHA Technology Solution Pvt. Ltd.

06/2017 - 06/2018

Noida, India

##### Achievements/Tasks

- Involved in creating Hive tables, loading with data and writing Hive queries which will run internally in map Reduce way.
- Used Hive to analyze the partitioned and bucketed data and compute various metrics for reporting.
- Involve in Requirement Analysis, Design, and Development.

## EDUCATION

### Bachelor of engineering

#### Meerut Institute of Engineering & Technology

07/2012 - 06/2016

Meerut, India

### XII (Senior Secondary Exam)

#### S.V.M Inter College (U.P. Board)

07/2010 - 06/2011

Saharanpur, India

## SKILLS

Hadoop

Apache Spark

Hdfs

Linux

SQL

Hortonwork(HDP)

Microsoft Azure

Hive

Snowflake

Python

Azure Databricks

Azure Data factory

Sqoop

Mapreduce

## PROJECTS

### GCCD(Global Connect Client Delivery) (12/2019 - Present)

- Description:** In GCCD project(Nielsen Client) we analyze the market data with respect to product, period and fact values for many countries.
- Technical Skills Used:** Azure Data Bricks, Apache Spark, Azure Synapse, Snowflake, Azure Blob.
- Role:** Big Data Engineer
- Roles and Responsibilities:**
  - Identified and analyzed the client requirements by interacting with Business Analyst.
  - Developing and maintaining Spark-Scala code for ETL from Azure BLOB to Azure Synapse Analytics using Databricks.
  - Configured the Pipeline for ingesting data from Azure Blob to Azure Synapse Analytics using spark.
  - Monitoring and optimizing performance of spark and azure synapse analytics.
  - Works on POC for replacement of Snowflake over Azure Synapse Analytic.

### AVA GPS TRACKER (06/2018 - 06/2019)

- Description:** Ava gps tracker is an application for the purpose of handling the GPS vehicles the various scenarios. This application have 6 modules which are vehicle management, tracking, reports, role management, alert & Notification.
- Technical Skills Used:** Apache Spark, Hortonwork(HDP), Hive, MongoDB, Core Java.
- Role:** Big Data Engineer
- Roles and Responsibilities:**
  - Developed Spark core and Spark SQL scripts using Scala for faster data processing.
  - Involved in converting Hive/SQL queries into Spark transformation using Spark RDD's.
  - Working on Hadoop 1.x/2.x cluster architecture and Monitoring the cluster.
  - Working with Hadoop Clusters using Hortonwork(HDP) distribution.

### Life Insurance organization (06/2017 - 06/2018)

- Life insurance organization, which has number of insurance products, once the products specification, design and is available for sale, it is presented to the customers through agents and online.

## LANGUAGES

English

Professional Working Proficiency

Hindi

Full Professional Proficiency

## INTERESTS

Badminton

Cricket