

ATUL MAURAY

E-707, Amrapali Princely Estate, Sec.- 76, Noida

Email: **atulkod@gmail.com**

Ph: **+91-9455288557**

Profile Summary:

- **4.5 years** of extensive experience in **GCP, Hadoop Ecosystem** and **Big Data**.
- Hands on experience in **PubSub, Apache Beam, Dataflow, Dataproc, Big Query, Big table, Cloud Compute Engine, Memstore(Redis), Cloud SQL(Postgresql)** and **Google Cloud Storage**.
- Hands on experience in **Apache Airflow, Rest API, Python Scripting** and **golang**.
- Hands on experience in **Spark-APIs(RDD, Dataframe, Dataset), Spark-SQL, Spark Streaming, Scala, Java, Hive, Sqoop, Kafka** and **HBase**.
- Hands on experience in creating **Big query/Hive** tables and working on them using **QL** and also have written **Big queries/Hive Queries** for data analysis to meet business requirements.
- Proficient in analyzing and translating business requirements to functional/technical requirements.
- Excellent analytical, problem solving, communication and interpersonal skills with ability to interact with individuals and can work as part of a team as well as independently.

Career Goal:

To obtain a growth oriented job in the IT industry involving **Senior Big-Data/Data Engineer** where the emphasis is on creative and quality work, dedication and ability of problem solving.

Academic Background:

- MCA from **National Institute of Technology(NIT) Jamshedpur** with 7.64 CGPA in 2016.
- BCA from Punjab Technical University with 75.69% in 2012.

Technical Skills:

- **GCP Services:** PubSub, Apache Beam, Dataflow, Dataproc, Big Query, Big table, Cloud Compute Engine, Memstore(Redis), Cloud SQL(Postgresql) and Cloud Storage.
- **Hadoop Ecosystem:** HDFS, Hive, Sqoop, Kafka, Hbase.
- **Spark Ecosystem :** Spark-Core, Spark-SQL, Spark Streaming.
- **Hadoop Platform:** Hortonworks Data Platform, Cloudera Distributed Environment.
- **Tools:** Eclipse, IntelliJ, Pycharm, Putty, WinSCP, BitBucket, github, SourceTree, Jira.
- **Programming Languages & DSA:** Java, Scala, Python, Golang, Data Structure and Algorithms.
- **Databases:** NoSql(**HBase and Big table**), Postgresql, MySQL

Employment summary:

- **Data Engineer** at **Tokopedia** India Private Limited, Noida (December 2019 to till date)
- **Technical Consultant** at **Iris Software** Private Limited, Noida (from May 2019 to till December 2019)
- **Software Engineer** at **InfoPro Learning** India Private Limited, Noida (From October 2018 to May 2019)
- **Big Data Engineer** at **Maxonic** India Private Limited, Bangalore (From July 2016 to September 2018)

Project Experience:

Customer Data Platform and Campaigns manager Dec 2019 - till date

Description: It manages and runs the campaigns, which help to reach targeted users with discounts, offers, gift cards etc. Customer data platforms help to read streaming data and ingest data into big query and gcp.

Responsibilities:

- Creating multiple dataflow and airflow pipelines to process and ingest live data in bigquery and processing batch data through airflow.
- Reading data from pubsub and process and store into bigquery and gcp by using dataflow and dataproc.
- Fetching newly created campaigns and providing users list to marketing team by using airflow and cached required data in redis.
- Creating endpoints in golang, which help others to get required campaigns data.
- Handled both Structured and semistructured data.
- Use pub sub and kafka and to load data from multiple sources.

Risk and Compliance Solutions October 2018 - May 2019

Description: It allows banks to manage risk and compliance according to the USA anti money laundering rules. Here we generate risk numbers based on person, account and business attribute like attribute, watchlist, network, advertisement , local, media etc.

Responsibilities:

- Implementing new features in scala and java with their unit test cases.
- Write spark jobs by using spark-RDD, spark-dataframe, spark-dataset, spark streaming and spark streaming API in scala.
- Write code in Java or Scala to fetch data from hdfs, hbase and hive APIs.
- Handled both Structured and semistructured data.
- Use Kafka to load data from multiple sources.

Monitor Patient Vital Signs

July 2016 - September 2018

Description: Here, We monitor the patient's vital signs such as temperature, pulse, heartbeat, blood pressure etc and send alerts to the doctors, nurses and care providers so they know instantly changes in patient condition.

Responsibilities:

- Involves creating hive tables, loading data and writing spark sql hive queries.
- Handle both structured and semi-structured data and process those data by hive and mapreduce.
- Involved in creating and optimizing hive tables to store data.
- Use Flume and sqoop to ingest data into hdfs.

I hereby declare that all the information given is correct to the best of my knowledge.

Place: Noida

Atul Mauray