

ASIF KHAN

8051527967 | asif31866@gmail.com |

DATA ENGINEER

EDUCATION

JRE GROUP OF INSTITUTION

Btech. ECE

GRADE : 65%

Gr Noida, UP

Aug. 2012 – May 2016

TECHNICAL SKILLS

Big data platform : Hadoop, Spark
Hadoop framework : Hive, Sqoop, Hbase, Mysql , pyspark, Python
Platform : Windows(8/10), Linux(Ubuntu, cent OS),
Cloud infra : AWS(EC2,S3,IAM,EMR,Glue,Redshift), AZURE(ADF,BLOB)

EXPERIENCE

RF Engineer

Jan 2017-March 2018

AIRTEL (Full Time)

- Conducting stationary and mobility drive test for all technologies (2G/3G/4G/Volte) using standard testing tools
- Responsible to trouble shoot the reasons for poor network performance and provide suitable optimization suggestions.
- Managed a process re-engineering project to improve and consolidate end-to-end service processes; restructured communication flow among 10 departments and cut down paperwork by 75%.

Network Associate Engineer

G-SAT 11 ISRO (SATELLITE COMMUNICATION) (Full Time)

April 2019- May 2021

- Led a series of integration initiatives for satellite systems with three departments, establishing protocols that facilitated smoother transitions and improved compatibility, now adopted as best practices across the organization.
- Orchestrated the integration of satellite systems into existing networks by collaborating with five cross-functional teams, resulting in a 40% increase in data transmission efficiency and reduced downtime for end users.
- Managed the installation and maintenance of satellite ground equipment, ensuring compliance with industry standards.
- Analyzed and optimized network performance to enhance data transmission efficiency.
- Created detailed reports on system performance and recommended improvements.

DATA ENGINEER PROJECTS

Data integration & Filtering | Hadoop,Spark,Hive,Mysql

Nov-2022 - May2023

- Engineered and optimized ETL processes utilizing Apache NiFi to efficiently ingest and transform 500M+ records daily into Hadoop clusters, improving data accuracy by 25% and reducing processing time by 40%
- Involved in Hadoop-based data processing pipelines, resulting in a 20% improvement in data ingestion efficiency
- Collaborated with data scientists to deploy machine learning models on Spark clusters, enhancing real-time analytics capabilities.
- Designed and maintained Hive data warehouses, ensuring optimal query performance and data accessibility for business analysts.
- Implemented Sqoop jobs to transfer data between Hadoop and relational databases, reducing data transfer times by 25%.
- Managed HBase clusters for storing and retrieving large-scale semi-structured data, achieving 99.9% uptime and data integrity.
- Worked closely with business analysts to define data requirements and transform raw data into meaningful insights.
- Created interactive dashboards using Python libraries; streamlined reporting efforts for three core business units, reducing report generation time by 40 hours monthly while ensuring accurate visual representation of complex dataset
- Conducted data profiling and cleansing activities to maintain data accuracy and quality. Assisted in the design and implementation of data governance policies and procedures.

Adidas Sales Data Analysis: Databricks,Pyspark,SQL

- The objective of this project is to analyze the adidas sales database for the year 2020 and 2021 and identify to help improve sales performance and optimize business strategies.