

Sahil Khan

Data Engineer

Seeking a challenging role, which will utilize my current skills and also introduces new challenges of the industry and learn new things to develop my skillset.

✉ sahil82764@gmail.com

📍 New Delhi, India

🐙 github.com/sahil82764

☎ +91-8800263260

🌐 linkedin.com/in/sahil-khan-845aa41b3

WORK EXPERIENCE

Data Engineer

Wobblrr

04/2022 - Present

New Delhi, India

Roles and Responsibilities

- Designed and Implemented a data ingestion solution to transfer data from various open APIs to Azure SQL DB and Azure Data Lake.
- Developed and maintained efficient data ingestion solutions on Databricks and Azure to analyze real-time fleet data.
- Used Spark and Databricks to process, analyze and uncover trends in real-time fleet data.
- Designed a data lake architecture to store and process large amounts of structured and unstructured data.
- Orchestrated the designing, testing and debugging of new and existing ETL & reporting processes.
- Improve decision-making and operational efficiency by providing real-time insights and collaborating with cross-functional teams.
- Utilized incremental data ingestion to minimize data duplication and improve data accuracy.
- Leading a team of interns in multiple projects.

Associate Professional Software Engineer

DXC Technology

07/2021 - 04/2022

Bangalore, India

DXC Technology is a Fortune 500 global IT services leader.

Roles and Responsibilities

- Analyzed Business, Functional Requirements and Design Review Documents to develop Test Scripts for positive, negative, edge cases.
- Performed Functional testing of the application to test the system for both functional and business requirements for positive and negative scenarios using ALM or Jira.
- Actively involved in Daily Scrum, Sprint Planning, Sprint Review and Retropective.
- Creating and Executing manual test cases in JIRA for various User Stories based on Release and Sprint Plan.
- Mitigate risks to Business through timely and precise execution of manual test scripts under moderate supervision.
- Documented software application defects and track it to closure.

EDUCATION

B.Tech - Electronics and Communication

Guru Tegh Bahadur Institute of Technology, New Delhi

08/2017 - 07/2021

CGPA: 7.8

TECHNICAL SKILLS

Platform: Windows / Linux

Language: Python, SQL, PySpark and SparkSQL

Framework: Databricks(Azure & AWS), Spark, Delta Lake

Databases: Azure SQL DWH, MongoDB, CosmosDB

Data Integration: Azure Data Factory

Serverless Computing: Azure Functions

Data visualization: Ms PowerBI and Tableau

Machine Learning, MLOps and Docker

PROJECTS

Forest Fire Prediction

- Project for predicting Algerian Forest Fires and Fire Weather Index using machine learning with python.
- In this project, we focused on whether certain weather features could predict forest fires in these regions.
- https://github.com/sahil82764/Forest_Fire_Predictor

Insurance Fraud Detection

- To build a classification methodology to determine whether a customer is placing a fraudulent insurance claim.
- We discussed features regarding person's income and policy details and see whether we could predict the fraudulent claims or not
- https://github.com/sahil82764/Fraud_Detection_Insurance_ML

Power BI Sample Reports/Dashboards

- Financial Reporting
- E-Commerce Sales Reporting
- Fleet Management and many more
- <https://github.com/sahil82764/Power-BI-Reports>

CERTIFICATES

Databricks Certified Data Engineer Associate
(12/2022 - 12/2024)

Microsoft Certified: Data Engineer Associate
(01/2023 - 01/2024)

Microsoft Certified: Azure Data Fundamentals
(11/2021 - Present)

Microsoft Certified: Azure AI Fundamentals
(03/2023 - Present)