

SAHIL KHAN

Toronto, ON, Canada

sahil82764@gmail.com | [Portfolio](#) | [LinkedIn](#) | [GitHub](#)

PROFESSIONAL SUMMARY

Microsoft Certified Fabric Data Engineer Associate and a Master's graduate in Data Science, specializing in building scalable, automated ETL/ELT solutions across the Azure ecosystem. My expertise lies in Python, PySpark, and SQL, and I focus on architecting robust, CI/CD-integrated pipelines to transform raw data into governed, high-availability datasets. I thrive on accelerating business value through reliable automation.

EDUCATION

Master's - Data Science	09/2024-08/2025
Memorial University of Newfoundland	St. John's, NL, CA
Postgraduate Certificate - Artificial Intelligence	09/2023-08/2024
Durham College	Oshawa, ON, CA
Bachelor of Technology - Electronics and Communication	08/2017-07/2021
Guru Tegh Bahadur Institute of Technology	New Delhi, India

SKILLS

Programming and Databases: Python, SQL, PL/SQL, R, KQL, Snowflake, NoSQL/MongoDB
Data Engineering: Spark, PySpark, Databricks, ETL/ELT Design, Airflow, dbt, Delta Lake, Unity Catalog
Azure Cloud: Fabric, OneLake, Data Factory, Synapse Analytics, Event Hubs, CosmosDB, ADLS Gen2, Functions, LogicApps
Analytics & Visualization: Power BI, Tableau; Statistical Analysis
Testing & Quality Assurance: Postman, ALM, Test Script Development, Functional Testing
Project Management & DevOps: Ms Excel, JIRA, ALM, Confluence, Git, CI/CD, Docker, Kubernetes
AI & MLOps: Azure AI Services, Azure Machine Learning, Model training/validation, ML Libraries, API design/integration

WORK EXPERIENCE

DATA ENGINEER	04/2022 – 08/2023
WOBBLRR	New Delhi, INDIA
<ul style="list-style-type: none">Orchestrated end-to-end data migrations, building robust Fabric Pipelines to ingest data from over 8 diverse on-premise and cloud sources into Microsoft OneLake.Engineered scalable data transformations using PySpark notebooks, implementing a Medallion architecture to refine raw data into a governed Gold layer within the Fabric Lakehouse, reducing manual reporting effort by 75%.Designed and implemented efficient batch and incremental loading patterns to process over 2GB of daily telemetry and transactional data, optimizing Delta Table storage in OneLake for high-performance analytics.Drove a 60% improvement in Power BI dashboard latency by creating curated Gold-layer Delta Tables and enabling Power BI's DirectLake mode to deliver near real-time analytics.Developed parameterized data pipelines to create a reusable ingestion framework, accelerating the onboarding of new data modules while integrating with Git for CI/CD across development and production environments.Enabled self-service analytics for over 25 business users by exposing key business KPIs through the SQL analytics endpoint of the Lakehouse, providing a secure and queryable layer for internal applications.	
SOFTWARE ENGINEER	07/2021-03/2022
DXC TECHNOLOGY	Bangalore, INDIA
<ul style="list-style-type: none">Analyzed business and functional requirements across 3 banking modules, developing and executing 150+ test scripts covering positive, negative, and edge cases with 95% test coverage.Managed 200+ manual test cases according to user stories, release plans, and sprint schedules using Jira, contributing to 4 successful sprint deliveries with zero critical defects in production.Conducted comprehensive functional testing of banking applications using Postman, identifying and tracking 85+ software defects to closure using ALM and Jira, achieving 98% defect resolution rate.	

ADDITIONAL EXPERIENCE

WALMART ASSOCIATE[PT]

10/2024 – PRESENT

WALMART CANADA

Mount Pearl, NL

- Unloaded and sorted merchandise, organizing products for efficient inventory processing and store restocking.
- Maintained cold chain standards by ensuring proper temperatures for refrigerated and frozen goods.
- Kept shelves fully stocked by moving and organizing merchandise from receiving to the sales floor.

RETAIL SALES ASSOCIATE[PT]

10/2023 – 08/2024

MOKSHA CANNABIS

North York, ON

- Provided detailed product information, met sales targets, and created a positive store environment.
- Managed inventory, ensured product availability, and complied with all age restriction regulations.
- Processed POS/cash transactions, fulfilled online orders, and was responsible for opening and closing duties.

CAMPUS WALK STUDENT[PT]

09/2023 – 12/2023

DURHAM COLLEGE

Oshawa, ON

- Ensured the safety of individuals by offering assistance and escorting them to their destinations.
- Followed Security Officer directives and safety protocols.
- Collaborated with a team partner and maintained a professional appearance as per program standards.

PROJECTS

Fake News Detection | [Github](#) | Built a web application with a scalable data pipeline, fine-tuned transformer models (BERT) via transfer learning in PyTorch for contextual text classification, performed model evaluation on accuracy/latency/memory trade-offs, and introduced an explainable AI approach with real-time fact-checking for media verification.

Vehicle Maintenance Prediction | [Github](#) | Developed a predictive vehicle maintenance solution using supervised machine learning with EDA and feature engineering, deploying a tuned decision tree via a Streamlit app to deliver proactive maintenance insights and actionable recommendations.

PowerBI Reports | [Github](#) | Sample dashboards showcasing financial, sales, and telemetry data.

CERTIFICATIONS

- **Microsoft DP-700:** Fabric Data Engineer Associate
- **Microsoft DP-900:** Azure Data Fundamentals
- **Microsoft AI-900:** Azure AI Fundamentals
- **Databricks Certified:** Data Engineer Associate
- **Google Certified:** Data Analytics
- **IELTS General Training:** Band 7, CEFR C1