Mahmud Un Nabi

Senior Data Engineer turning data complexity into scalable business-impacting insights

Dhaka, Bangladesh | +880 155 246 53 31 | LinkedIn | Github
Born 13/11/1998 | munmridul@gmail.com
I am a Bangladeshi citizen. I am open for relocation or remote roles. My notice period is 60 days.

TECHNICAL SKILLS

Python (6 years of project experience) | SQL (5) | ELT pipelines (5) | Event Driven Architecture Modelling (5) | Dimensional Data Modelling (5) | Snowflake (4) | PostgreSQL (3) | MySQL (2) | MongoDB (2) | DBT (3) | PySpark (2) | Fivetran (3) | Airbyte (3) | Kafka (2) | Apache Airflow (4) | Argo Workflow (3) | AWS (S3, EKS, IAM, Redshift, Application Load Balancer) | Docker (5) | Kubernetes (4) | Terraform (2) | GitHub Actions (3) | FastAPI (3) | Pydantic (3) | Test Driven Development (5) | Pytest (5) | MLflow (2) | GraphRAG (1) | Prompt Engineering (1) | Google ADK (1) | Apache Iceberg (1) | Open Table Format (1)

WORK EXPERIENCE

06/2022 – today

Senior Data Engineer
Optimizely, Dhaka, Bangladesh (A US-based global enterprise, empowering businesses to deliver optimized digital experiences at scale using AI-driven experimentation, personalization, and marketing orchestration)

08/2025 - today

Project: Modern Data Stack Migration (AWS RDS Postgres to Snowflake) - A migration project to upgrade Optimizely's data warehouse to a modern cloud system (Snowflake), reducing cost and data reliability across the Optimizely organization.

- Migrating Airflow ELT pipelines to Argo Workflows on Kubernetes with dbt, enabling retries and backfills to achieve a 99.7% pipeline success rate.
- Deployed Airbyte on AWS EKS to replace Fivetran, cutting cloud costs by 50%.
- Designed Medallion architecture-based DBT pipelines enabling analytics across 30+ mission-critical business dashboards.
- Implemented Point-In-Time tables in DBT by leveraging DBT macros, eliminating ambiguity across different but relevant data by 100 %.
- Implemented Terraform-based Snowflake RBAC for 2,000+ tables, enabling data security and governance with auditable access control.

Technologies used: Python | SQL | ELT Pipelines | Snowflake | Postgresql | DBT | AWS (EKS, S3) | Apache Airflow | Airbyte | Fivetran | Argo Workflow | Docker | Kubernetes | Github Actions | Terraform | Dimensional Data Modelling

03/2024 - 08/2025

Project: DWBrainstormer (AI agent) - AI-powered data query engine for general knowledge discovery in Optimizely's data.

- Deployed a knowledge-graph-based AI agent for semantic data discovery, serving 40% of monthly recurring adhoc data requests.
- Built an Actor-Critique AI agent (using Google ADK) for automated Snowflake SQL generation, reducing manual query effort by 35–40%.
- Implemented agent-driven, schema-aware data quality tests with DBT, Argo, and GitHub Actions, reducing production incidents by 30%.

Technologies used: Python | FastAPI | DBT | AWS EKS | AWS S3 | AWS Application Load Balancer | GraphRAG | Google ADK | Prompt Engineering

07/2023 - 03/2024

Project: DataSync (Snowflake to Salesforce Reverse ETL) - Seamless integration of Salesforce and Snowflake, syncing customer insights from Snowflake to Salesforce to power smarter targeting, along with higher converting campaigns.

- Engineered a high-throughput Python-based Snowflake-Salesforce CRM integration, reliably ingesting up to 30 million rows per day.
- Refactored the legacy pipeline to incremental, idempotent processing, achieving atleast 50% reduction in cloud run costs.

Technologies used: Python | SQL | Snowflake | Pytest | OOP | SOLID | Argo Workflow | Docker | Kubernetes | Test Driven Development | Github Actions

06/2022 - 07/2023

Project: ML Pipeline Infrastructure - Scalable ML pipeline infrastructure, reducing deployment time and accelerating Optimizely's ML-driven customer churn prediction analytics.

- Optimized ML retraining by offloading pandas-based data transformations to Snowflake SQL, cutting runtime 60%
- Scaled churn prediction analytics 5x by orchestrating 20+ ensemble ML models with memory-efficient, fault-tolerant Argo pipelines on AWS EKS.

Technologies used: Python | AWS | Kubernetes | MLflow | Scikit-learn

01/2021 – 06/2022

Software Engineer
Shohoz Limited, Dhaka, Bangladesh (One of the largest online ticketing platforms in Bangladesh)
Project: PySpark-based ETL- Implemented PySpark data ingestion jobs (up to 30 GB/day from MySQL and MongoDB), optimizing joins and storage to deliver analytics that cut go-to-market strategy review time by 3 days.
Technologies used: Python | PySpark | Open Table Format | Apache Iceberg | MySQL | MongoDB

EDUCATION

01/2017 – 11/2021

Bachelor of Science, Computer Science and Engineering
Independent University Bangladesh, Dhaka, Bangladesh

CERTIFICATIONS

Modelling Data Warehouse with DataVault 2.0 | Neural Networks and Deep Learning | Improving Deep Neural Networks | Convolutional Neural Networks

Technical Articles

Test Driven Development in Data Engineering using DBT on Postgres

LANGUAGE SKILLS

English (Fluent) | Bengali (Native) | German (A1 - Learning)

HOBBIES

Playing First-person shooting game | Car enthusiast