

Uddipan Goswami

sohan.goswami02@gmail.com • +91 9123770250 • linkedin.com/in/uddipangoswami

SUMMARY

Results-oriented Software and Data Engineer with expertise in Python, FastAPI, ReactJS, and PostgreSQL. Experienced in developing and testing data-driven applications, integrating APIs, and supporting product design and deployment workflows. Strong foundation in Azure cloudservices, automation, and scalable system architecture.

WORK EXPERIENCE

Mindsprint

Data Engineer

Chennai, India

July 2024 – Present

- Ingested data from S/4HANA, PostgreSQL, and Oracle into Snowflake staging using ADF, improving pipeline stability and reducing manual intervention by 30%.
- Built DBT transformation models (staging, incremental, curated layers), enhancing data quality and accelerating transformation runtimes by 20–25%.
- Automated external data ingestion by extracting JSON/CSV from URLs using Python and loading it into Snowflake internal/external stages.
- Implemented Row-Level Security (RLS) with roles and secure views, ensuring 100% compliant access control for sensitive datasets.

PROJECTS

Airbnb Data Engineering Pipeline

2025

Azure Data Factory • CosmosDB Change-Feed • ADLS Gen2 • Synapse Analytics

- Built ELT pipeline using **Azure Data Factory** to ingest raw Airbnb CSVs hourly into **ADLS Gen2**.
- Processed Cosmos DB change-feed via ADF for filtering, dedupe, and staging upserts.
- Applied SCD-1 upserts to Synapse using ADF to maintain clean dimension tables

Retail Sales Streaming Pipeline

2025

Real-time POS data ingestion optimization for sales and inventory

- Real-time POS data ingestion & optimization for sales and inventory
- Configured Snowpipe to auto-load batch-partitioned POS CSVs from ADLS Gen2 via Event Grid into Snowflake.
- Implemented batching (100–250MB) with date-based partitioning, automated retrying of failed loads, and manual refresh capability for backfilling.

Gait Recognition System

2024

Developed a biometric authentication model with 93.75% accuracy.

- Designed a real-time gait recognition system using OpenCV and Random Forest.
- Conducted testing under different conditions to improve model robustness.
- Implemented feature extraction techniques to enhance classification accuracy.

AI Experience Classification using Few-Shot Learning

2024

Built a binary classifier for AI-related LinkedIn experiences using few-shot learning.

- Implemented Prototypical Networks, RoBERTa, and MAML to minimize labeled data needs.
- Created a dataset of 2,500 LinkedIn profiles and improved accuracy through transfer learning.
- Developed an automated data labeling pipeline to reduce manual effort.

TECHNICAL SKILLS

- **Programming:** Python(PySpark,pandas,NumPy,Matplotlib),
- **Databases:** MS SQL Server, PostgreSQL,HANA
- **Cloud and Tools:** Microsoft Azure,Snowflake,Azure Data Factory,DBT

EDUCATION

- **Kalinga Institute of Industrial Technology**, B.Tech in Computer Science and Engineering
CGPA: 8.92

2020-2024

- **Exchange Program:** CY Tech Paris University

2024