

# TABASSUM HOSSAIN

Kolkata, India

9062873120/ 9836624083 | [tabassum.hossain2025@gmail.com](mailto:tabassum.hossain2025@gmail.com) | <https://www.linkedin.com/in/tabassum-hossain-5b6a36186/>

## Summary

Data Engineer with hands-on experience designing and building data pipelines, ETL workflows, and analytical solutions using Databricks, Delta Lake, PySpark, SQL, and Azure Cloud services. Proficient in developing scalable data ingestion, transformation, and processing systems, with a strong focus on reliability, automation, and performance optimization.

## Skills

**Programming:** Python, SQL

**Technologies:** PySpark, Databricks, Delta Lake, Azure Data Factory, Azure Data Lake, Data Warehouse, Hive, HDFS, Data Visualization, Numpy, Pandas, Matplotlib, Machine learning, NLP, LLM.

**Tools:** Databricks, Jupyter Notebook, Visual Studio Code

**Professional Skills:** Communication, Documentation, Problem-Solving, Collaboration

**Language Skills:** English, Bengali, Hindi

## Work Experience

### Freelance

#### Data Engineer

**Jan 2025 – Present**

- Designing and optimizing ETL workflows and analytics pipelines using Python, SQL, and Spark on cloud environments.
- Developing scalable data ingestion and transformation processes improving data accessibility and performance.

#### AlphaDnA (formerly Gnosis lab)

#### Data Engineer

**Oct 2023- Dec 2024**

- Collaborated on a legal NLP project developing a translation platform for English-to-Bengali case files.
- Utilized Python, Pandas, and NLP tools to preprocess and manage 10,000+ text documents efficiently.
- Worked with Azure Databricks and cloud storage to enhance data processing and deployment workflows.
- Supported LLM-based automation and data visualization for improved document analysis accuracy.

### Codevidhya

#### Computer Science Lecturer

**July 2020- Sept 2023**

- Delivered interactive lessons on Python programming and logical problem-solving for 100+ students.
- Mentored learners and created structured course material to enhance student engagement and performance.

## Certifications

1. The Complete Python Bootcamp from Zero to Hero in Python – Udemy
2. Data Engineering Bootcamp with GCP and Azure cloud- Udemy
3. The Ultimate Big Data Master's Program with Cloud Focus – Trendytech (Ongoing)

## Education

### Jadavpur University

**Aug 2017 – Jul 2019**

M.Tech in Biomedical Engineering,  
Kolkata, India

**GPA: 8.11**

### MAKAUT

**Aug 2013- Jul 2017**

B.Tech in Instrumentation & Control  
Engineering,  
Kolkata, India

**GPA: 8.66**

## Publications

Tabassum Hossain, Arnab Rakshit, & Amit Konar. (2020, January). Brain-Computer Interface based User Authentication System for Personal Device Security. In 2020 International Conference on Computer, Electrical & Communication Engineering (ICCECE) (pp. 1-6). IEEE. doi:10.1109/ICCECE48148.2020.9223069