

# TAFIQUE HOSSAIN KHAN

+91-8926148518

tafiquehossain2003@gmail.com

GitHub Profile

in LinkedIn Profile

## PROFILE SUMMARY

Data Science student with hands-on experience in building machine learning projects using various libraries and MLOps tools. Proficient in machine learning and deep learning algorithms, with practical knowledge of end-to-end model deployment. Skilled in applying data science techniques to solve real-world problems effectively.

## EDUCATION

**Gandhi Institute of Technological Advancement, Bhubaneswar, India** **2025**  
*Bachelor of Technology CSE With Specialization in Data Science - CGPA - 8.91*

**Vyasagar Public School, Odisha, India** **2021**  
*XII - Secondary Education - Percentage - 64%*

**Vyasagar Public School, Odisha, India** **2019**  
*X - Primary Education - Percentage - 76.6%*

## Technical Skills

- **Languages:** Python, Java, MySQL
- **Concepts:** Machine Learning Algorithms, Statistical Analysis, ETL Process, Feature Engineering, Model Evaluation, Data Wrangling, Data Cleaning, Mathematics for Machine Learning & Deep Learning
- **Tools:** Docker, Git, DVC, MLflow, Apache Airflow, DAGsHub, Power BI
- **Libraries:** pandas, NumPy, scikit-learn (sklearn), Matplotlib, Seaborn, Plotly, NLTK, SpaCy, Streamlit

## PROJECTS

### Food Delivery Time Prediction

- Employed feature extraction techniques to enhance model accuracy, utilizing DVC for data versioning and integrating MLflow and DagsHub for streamlined experiment tracking and model registration.
- Developed an automated training pipeline with Apache Airflow and containerized the application using Docker, leveraging the registered MLflow model for accurate delivery time predictions.

### Restaurant Sentiment Analysis & Recommendation System

- Developed a web app that analyzes 1000+ reviews using NLP and sentiment analysis for personalized restaurant recommendations.
- I used cosine similarity for the content-based recommendation system and a filtering recommendation system based on cuisine and ratings.
- Validated sentiment analysis results through hypothesis testing, creating a user-friendly interface with Streamlit.

### Health Prediction Web App

- Developed a Health Condition Prediction web app for diabetes and heart disease risk assessments, using machine learning for early health insights.
- Implemented symptom-based predictions and handled imbalanced datasets with techniques like SMOTE to enhance prediction accuracy.
- Built with Python and Streamlit, creating an interactive and user-friendly experience for personalized health assessments.

## CERTIFICATIONS

### Data Science Masters

Platform: pw skills

Certificate ID: 4675c2db-3652-4d84-a0f7-7381f75b7d79

Completion Time: February 2024

View Certificate