

Aspiring Data Scientist and Machine Learning enthusiast dedicated to continuous learning and professional growth. Passionate about solving complex challenges in generating models in Computer Vision and RAG based system in Natural Language Processing (NLP). Aim to leverage advanced analytical skills and innovative mindset to create impactful solutions that benefit both the company and society as a whole.

## EXPERIENCE

### Junior Machine Learning Engineer

October 2024 — Present

Omdena Inc.

USA

- Collaborating with 50+ ML and Data Engineers from 15+ countries to predict Soil Nutrients from data collected from different regions of Bangladesh
- The project aims to recommend fertilizers based on the predicted micro nutrients
- Working on designing feature engineering pipelines that capture key soil characteristics through EDA and developing AI models to provide nutrient predictions

### TA (Data Science)

June 2024 — Present

MasterCourse BD

Dhaka, Bangladesh

- Assisting the course instructor
- Engaging with students to identify and solve problems

### Machine Learning Engineer

December 2023 — May 2024

Innovative Skills BD

Dhaka, Bangladesh

- Developed a face recognition system with state of the art DL algorithms to achieve high accuracy and adaptability in real time system
- Conducting workshops for DS fundamentals like numpy, pandas as a course instructor

## SKILLS

### Programming Languages & Sql

Python, MySQL, PostgreSQL

### Web Scraping

BeautifulSoup, Selenium, Scrapy

### Data Analysis & Visualization

Numpy, Pandas, Matplotlib, Seaborn, Tableau

### Machine Learning Tools

Tensorflow, Keras, Pytorch, Sklearn, Scipy, OpenCV, Fastai, Huggingface

### Version Controlling Systems

Git, Github, Dagshub

### Operating Systems

Linux (Ubuntu, Mint), Windows 10/11

### Miscellaneous

Jupyter Notebook, Anaconda, Google Colab, VS code, IntelliJ IDEA IDE

### Soft Skill

Leadership, Teamwork, Time management, Communication, Critical thinking  
Quick learner, Problem Solver, Presentation

## PROJECTS

### Benagli Book Recommendation System (Github)

Tools & Skills : Python, Scrapy, Pandas, GoogleTranslator API, Deep translator, Lang detect, Gensim, Fasttext

- Developing an end-to-end **recommendation system** that provides personalized suggestions for **Bengali books**.
- **300K+** book data scraped from the Rokomari website, involving extensive preprocessing to structure and clean the data for effective model training
- The preprocessing pipeline included custom handling of features such as author, title, and genre, along with a **refined search** strategy to improve recommendation accuracy
- Worked on different models (**Gensim, Fasttext**) to examine their ability to generate relevant results based on specific titles, genres, or authors

### Multi-label Text Classifier with Integrated Transcriber and Summarizer(Github) (App)

Tools & Skills: Python, Selenium, Pandas, Fast ai, Hugging Face transformer, Blurr, Flask, Onnx, Git

- A multi-label text classifier that **transcribes** video/audio, **classifies**, **summarizes** and **extracts** important topics
- Can be used for summarizing and extracting topics from videos, audios or long text
- Fine-tuned **distilroberta-base**, achieved **0.42** F1 score (Micro), used **faster-whisper** for transcribing video/audio, compressed a **317 MB** model to **80 MB** using **ONNX**
- Deployed using **Flask, Onrender** and **Hugging Face** spaces

### Image Based Body Measurement Estimation

*Tools & Skills: Python, Pytorch, Pandas, Matplotlib, OpenCV, Dagshub*

- Implementation of Body Measurement model following the paper [Human Body Measurement Estimation with Adversarial Augmentation](#)
- The output model outputs **14** different body measurement from silhouette images of a person

### Amazon Data Science Books Analysis ([Github](#)) ([Tableau Dashboard](#))

*Tools & Skills: Python, Selenium, Geckodriver api, Tableau public, Git*

- A data visualization project on book data scraped from Amazon using python and selenium
- Scraped **1300+** data science books from **Amazon**.
- Preprocessing and data manipulation done with **Python** and visualization report generated with **Tableau**.

### Bangladesh Property Data Analysis ([Github](#)) ([App](#))

*Tools & Skills: Python, Numpy, Pandas, Scikit-learn, Regression algorithms (Linear regression, Decision Tree Regressor, Random Forest regressor)*

- Analyzing and training different regression models on one of the biggest property data of Bangladesh
- Reduced **MAE 5800** to **5600** and serialized the best model using **joblib**
- Built a user-friendly app using **Stremlit**

## EDUCATION

---

### BSc (Engineering) in Electrical & Electronic Engineering

April 2017 — November 2022

*Hajee Mohammad Danesh Science & Technology University*

## ACTIVITIES

---

Into the Computer Vision - Classify, Segment, Detect & Generate

Organized by: Innovative Skills BD (Completed as one of the top performer)