Faysal Al Mahmud

Mymensingh, BD | (+880) 1756-282696 | faysalalmahmud78@gmail.com
faysalalmahmud | faysalalmahmud | faysalalmahmud.github.io

Objective

Machine Learning Engineer proficient in building end-to-end NLP and Computer Vision pipelines. Expert in extracting actionable insights from large-scale unstructured datasets using Python & SQL. Proven track record of delivering scalable, data-driven solutions and technical documentation within collaborative, high-volume environments.

Experience

- **MasterCourse** *Jun 2025 - Nov 2025*
  Trainee Data Scientist Remote
  - Built full-cycle pipelines from **Web Scraping** to training **NLP and Computer Vision** models.
  - Developed and optimized architectures using **PyTorch, fastai, and Transformers**.
  - Integrated models into production via **Flask/Gradio APIs** and created **Tableau/Streamlit** dashboards.

Projects

- **AI Patent CPC Code Classifier: A Deep Learning NLP Pipeline for Multi-Label Classifier** *Oct 2025*
  Tools: Selenium, PyTorch, fastai, blurr, Transformers, Pandas, DistilBERT, Scikit-learn, ONNX, Gradio, Flask, Render ()
  - Scraped and processed **39,681 patent abstracts** from Google Patents using **Selenium** and **Pandas**.
  - Fine-tuned a **DistilBERT** model (using **fastai/blurr**) for multi-label classification, achieving **85.6% accuracy** and **85.5% F1-score**.
  - Optimized for production by compressing the model with **ONNX** and deploying via **Flask API** on Render.

- **Surgical Equipment Recognizer: A Deep Learning Identification Tool** *Sep 2025*
  Tools: fastai, PyTorch, DenseNet121, Gradio, DuckDuckGo API, Hugging Face, GitHub Pages ()
  - Built an end-to-end computer vision pipeline, scraping 5,000+ images via DuckDuckGo API and training a **DenseNet121** model with fastai.
  - Achieved **86% accuracy** classifying 20 types of surgical instruments after evaluating multiple CNN architectures.
  - Deployed an interactive **Gradio** app on **Hugging Face Spaces**, integrated with a GitHub Pages front-end for a live demonstration.

- **Analysis of Healthcare Professional in BD: A Comprehensive Data Analysis Project** *Aug 2025*
  Tools: Python, Selenium, Pandas, Tableau ()
  - Scraped and processed data of **6,000+ doctors** to study specialization, experience, and regional coverage.
  - Built interactive **Tableau dashboard** with maps, charts, and specialty filters.
  - **85% of doctors** concentrated in **Dhaka**, rural divisions underserved;
  - **Immunologists avg. 30 yrs** experience vs. **Psychologists 8 yrs**.

Technical Skills

- **AI & Machine Learning:** PyTorch, Scikit-learn, fastai, Transformers (Hugging Face), blurr, ONNX, Computer Vision (DenseNet121), NLP (DistilBERT)
- **Languages & Databases:** Python (Flask), SQL (MySQL), C++, Java
- **Data Engineering & Analysis:** Pandas, NumPy, Selenium, BeautifulSoup, Tableau, Matplotlib, Seaborn
- **Deployment & MLOps:** Gradio, Streamlit, Render, Vercel, Git, GitHub, Jupyter/Colab

Education

- **Brahmaputra International University** *Jan 2022 - Dec 2025*
  BSC in Computer Science and Engineering Jamalpur, Bangladesh
  - CGPA: 3.83/4.00

Achievements

- **Honorable Mention** *06 - 07 December 2024*
  ICPC Asia Dhaka Regional Contest 2024