Mahedi Hasan **Rasel**

Machine Learning Engineer, Open Sourcerer

mahedi.hrasel@gmail.com mahedi GitHub LinkedIn +8801400488939

EXPERIENCE

Business Automation Ltd. | Machine Learning Engineer — Dec 2023 - Present

- Digitized handwritten prescriptions with OCR model, leveraging Generative LLM models for output formation and Layout-based document digitalization using LayoutLM and Donut model.
- Implemented Bangla Law Consultancy chatbot with interaction-based instruction dataset using Llamma-3, featuring RAGsystem and effective tokenization method.
- Extracted Land Records information using OCR, Bangla Document OSD functioning, and Name Entity Recognition employing the Mistral Model with 4-bit quantization for effective memory allocation.
- Executed BIDA and EBS Employee Behavioral Log Text Summarization enhancing efficacy in task-based accomplish ment utilizing Mistral, Pegasus, and ML Models.
- Contributed to EBS Inshight-db database QnA using ChatGPT API and TAPAS model.

InflexionPoint Technologies BD Ltd. | Artificial Intelligence Engineer (intern) — April 2023 - Nov 2023

- Excelled in various data-related tasks, including:
  - Data analysis
  - Feature engineering
  - Data preprocessing
  - Model creation
- Skilled in SQL for effective data handling.

SKILLS

Programming Languages

• Python, • C • C++ • Bash Scripting

Machine Learning Frameworks

• PyTorch • Tensorflow • HuggingFace Transformers • Scikit-Learn • GenAI(RAG, PromptEngineering, Fine Tuning)
• Modules: NLTK, Spacy, Pandas, OpenCV, Pillow, Numpy, Matplotlib, Seaborn

Database

• MySQL • POSTGRESQL • SQLite • MongoDB

Web framework

• Flask • FastAPI • Django

DevOps MLOps

• AWS(VPC, EC2, Fargate, S3, Lambda) • Docker-compose • CICD(jenkins, github Action) • GNU/Linux, • Terraform , • Ansible
• Kubernetes • Sonarqube
• Amazon SageMaker • Vertex AI • Vast.ai(for ai model training)

NOTABLE PROJECTS

Name Entity Recognition (NER) with MISTRAL, BERT, and FLAN T5 | Python, Docker

- Name Entity Recognition (NER) system implemented using MISTRAL, BERT, and FLAN T5 models. These models can identify and classify entities such as persons, organizations, and locations within text data. Mistral, and Flan T5 models are fine-tuned on Cyber Security dataset for classifying the class of texts as well.

Doctor's Handwritten Text Extraction from Medical Prescriptions | Python, Jupyter Notebook

- This project aims to automate the extraction of handwritten text from medical prescriptions. It employs an algorithm to segment handwritten medicine names from prescription images, saving each segmented line as an individual image. These images are then annotated and compiled into a CSV file containing the image filenames and their corresponding annotated text. A TrOCR model is trained using this annotated data to recognize handwritten text. Upon inference, new prescription images are segmented, and the segmented images are passed to the TrOCR model for text extraction. The extracted text is then organized according to specific requirements using the Gemma-7b large language model.(Private repo)

Resume Information Extractor from PDF | Python, Jupyter Notebook, Flask , Docker

- This project is a web application that extracts key information (name, email, phone, university, and degree) from resumes in PDF format using OCR and named entity recognition. The application utilizes Flask for the web interface, PyTesseract for text extraction from images, and the GLiNER model for identifying and extracting relevant information from the extracted text.

Intelligent Question Answering and Code Generation Chatbot for Tabular Data | Python, Flask , Docker

- The goal of this project is to build an intelligent system that enables users to ask questions about tabular data in natural language and automatically generates code snippets to answer those questions. Enable the system to connect with and interact with different datasets stored in structured formats like CSV, Excel. Depending on the dataset, the model can visualize user queries.

Automated Passport Number Tracking Using Image Verification Identification | Python , Docker

- Employing the VGGFace deep learning model in tandem with MTCNN (Multi-task Cascaded Convolutional Networks) , this project pioneers precise face detection and recognition within image verification systems. Beyond merely tracking passport numbers, it integrates an automated model training mechanism triggered at specified intervals upon receiving new images, guaranteeing ongoing refinement and peak performance.

AI-Driven Legal Consultant Chatbot | Llama3, Unsloth, xformers

- Developed an AI-powered legal assistance chatbot for the Bangladeshi legal system, providing instant, accurate responses to common legal queries in Bengali.

EDUCATION

BANGLADESH UNIVERSITY OF BUSINESS AND TECHNOLOGY, Dhaka | BSc in CSE. 2019-2023

CGPA: 3.38/4

Courses - Object Oriented Programming, Data Structures Algorithms, Database Systems Operating Systems, Artificial Intelligence and Expert system, Distributed Database Systems, Machine learning, Neural Network and Fuzzy systems, Cyber Security and Digital forensic, Data Mining, Pattern Recognition

ACHIEVEMENTS/AWARDS

- Top 10 Innovator Award from Digital Bangladesh Mela
- Digital Bangladesh Mela | Innovator
- Contestant Certification from ICPC 2021

PUBLICATIONS

Journal Articles

- A Stacked Ensemble Machine Learning Approach for the Prediction of Diabetes
  Authors: Khondokar Oliullah; Mahedi Hasan Rasel; Md. Manzurul Islam; Md. Reazul Islam; Md. Anwar Hussen Wadud; Md. Whaiduzzaman
  Journal of Diabetes and metabolic Disorders, 603-617, 2024.
  [Link to Paper]

Conference Papers

- Traffic Congestion Detection Using Machine Learning and Deep Learning
  Authors: Mahedi Hasan Rasel; Hitesh Chakraborty; Shadia Jahan Mumu; Md Mahmudul Hasan; Sadia Afrin
  [Link to Paper]

Articles

- Understanding Optical Character Recognition (OCR) [Link]
- How to Fine-Tune an Embedding Model with Sentence Transformers [Link]
- Building an Integrated Data Pipeline for a Retail Shop [Link]

REFERENCES

Ms. Sadah Anjum Shanto
Assistant Professor Department of CSE, BUBT
Email: sshanto@bubt.edu.bd
Relationship: Former Advisor and Mentor