Md. Abdur Rakib Mollah

Machine Learning Engineer | Problem Solver

+88-01878-099-706 | rakib1703115@gmail.com | github.io/portfolio/ | rakib-mollah | abdurrakibmollah/

Summary

I completed my BSc in Computer Science & Engineering from Rajshahi University of Engineering & Technology. I am currently working as a Machine Learning Engineer at Euclido.Inc.- a Canada-based Ed-tech start-up. I have a keen interest in Data Analysis, Machine Learning, Deep Learning, Computer Vision, and NLP. I have published a few research papers at various conferences.

I have completed my thesis, which involved researching Natural Language Processing (NLP) with a specific focus on text generation in the Bangla language. During this project, I utilized language models like GPT-2 and BERT. This endeavor not only allowed me to make contributions to the field but also expanded my knowledge and expertise in NLP. Overall, with a curious mindset and hands-on experience, I am an enthusiastic learner committed to staying informed about the latest advancements in Artificial Intelligence.

Experience

Machine Learning Engineer-I *Waterloo, Ontario, Canada*

EUCLIDO.INC *March 2024 - Present*

- Designed and built a chatbot that leverages Retrieval-Augmented Generation (RAG) combined with LangGraph for dynamic, context-aware responses.
- Created an AI agent to autonomously manage the ETL process, automating data ingestion, transformation, and loading while significantly reducing manual intervention.
- Implemented a knowledge graph to establish relationships between chatbot sessions, enhancing contextual understanding and providing deeper insights into user interactions.

Research Assistant *Rajshahi, Bangladesh*

DEPARTMENT OF CSE, RUET *Jan 2022 - Jul 2022*

- Worked on a research paper in my 6th semester under the supervision of my respected teachers, it was later published.

Skills

Programming & Development	Python (Sci-Kit Learn, TensorFlow, PyTorch, JAX), C++, Java, JavaScript
Machine Learning & AI	LLMs, Retrieval-Augmented Generation (RAG), Vector Databases (Pinecone, ChromaDB)
Prompt Engineering	Chain of Thought (CoT), Tree of Thought (ToT), Langchain for tool integration
Cloud & ML Deployment	AWS (S3, Lambda, EC2), Azure AI Services, CI/CD Pipelines (Jenkins, GitHub Actions), Docker
Database Management	PostgreSQL, MySQL, MongoDB
DevOps & Software Engineering	Git, Linux, Websockets, RESTful APIs, Flask, FastAPI, Docker Image Management

Research Work

Enhancing Political Stance Detection via Knowledge Graph Retrieval and Chain-of-Thought Reasoning. *ACL Anthology, 2025*

UNDER REVIEW *March 2025*

- Enhanced political stance detection in understudied contexts by combining structured knowledge graphs with chain-of-questions reasoning (GRASP-ChoQ), achieving up to a 40% higher F1 score over zero-shot baseline methods.

Adapting Contextual Embedding to Identify Sentiment of E-commerce Consumer Reviews with Addressing Class Imbalance Issues. *iCACCESS, 2024*

ACCEPTED HERE* *March 2024*

- Optimized sentiment classification in e-commerce reviews by leveraging RoBERTa embeddings and XGBoost with re-weighted loss, achieving 83.84% multiclass and 93.29% binary accuracy.

Detection of Fake News with RoBERTa Based Embedding and Modified Deep Neural Network Architecture. *ICCIT, 2023*

PUBLISHED HERE* *Dec 2023*

- Improved fake news classification by integrating RoBERTa embeddings with a modified deep neural network, achieving higher detection accuracy than existing benchmarks.

Exploring Deep Convolutional Neural Networks: A Grad-CAM Enhanced Comparative Study for Automated COVID-19 Diagnosis from Chest X-Ray Images. *SMART GENCON, 2023*

PUBLISHED HERE* *Dec 2023*

- Achieved 99.13% binary and 97.66% multi-class accuracy in COVID-19 diagnosis by leveraging Grad-CAM visualization and evaluating multiple CNN architectures.

Improving Alzheimer's Disease Diagnosis on Brain MRI Scans with an Ensemble of Deep Learning Models. *ICAIIHI, 2023*

ACCEPTED HERE* *Dec 2023*

- Developed an ensemble model for early Alzheimer's detection, surpassing benchmark accuracies with 99.96% and 98.90% results on ADNI datasets.

An Ensemble Approach for Identification of Distracted Driver by Implementing Transfer Learned Deep CNN Architectures. *ICCIT, 2022*

PUBLISHED HERE* *Dec 2022*

- Enhanced driver distraction detection by ensembling MobileNet and DenseNet121, leading to improved classification accuracy on dashcam images.

Generation of Bangla Literature Texts by Fine-tuning Generative Pre-trained Transformer.

B.SC THESIS*

- Developed a GPT-2-based Bangla text generator that replicates Satyajit Ray's writing style, evaluating performance through Masked vs. Causal Language Modelling.

Education

Rajshahi University of Engineering and Technology (RUET) *Rajshahi, Bangladesh*

B.SC IN COMPUTER SCIENCE AND ENGINEERING *Jan 2018 - Sep 2023*

- *CGPA: 3.02/4*

Projects

DetectVideoShotLength Developed a Python package for detecting scene cuts in videos using histogram and SSIM analysis. The tool identifies frame changes, calculates shot durations, and visualizes scene length distributions. Available here.*

Bangla Text Generation using GPT-2 Fine-tuned GPT-2 to generate Bangla text in the style of Satyajit Ray's writing. The notebook is available here.*

Bangtex I created bangtex — a Python package that automates the conversion of native Bangla into bangtex transliteration for easier compilation in pdfLatex compiler. More information here.*

Line Follower Robot using Arduino UNO Designed and programmed an autonomous line-following robot as part of a peripherals and interfacing project. The project is uploaded here.*

Live Weather Forecast with Django Implemented a Django-based weather application that fetches real-time weather data using a public API. The complete project is uploaded here.*

Certifications

April 2023 **NLP**, Natural Language Processing with Classification and Vector Spaces. *Coursera*

May 2023 **Machine Learning**, Machine Learning with Python *Freecodecamp*