

Md. Ayon Mia

☎ +8801635690439 | ✉ mdayonrahman100@gmail.com | 🌐 Ayon128 | 🎓 Google Scholar | 🌐 ayon128.github.io



EDUCATION

- **Chittagong University of Engineering & Technology** Feb 2019 - June 2024
Bachelor of Science Chattogram, Bangladesh
 - GPA: 3.83 (Honors)/4.00 (7th position)
 - Coursework: OOP, DBMS, OS, Computer Network, Computer Architecture, System Design and Analysis, Compiler Design, AI, ML, etc.
 - **Thesis:** *Code-Mixed Romanized Bengali to English Translation with Sequence-to-Sequence Deep Learning Model*
- **Rajuk Uttara Model College** 2018
Higher School Certificate Dhaka, Bangladesh
 - Grade: 5.00/5.00
- **Uttara High School & College** 2016
Secondary School Certificate Dhaka, Bangladesh
 - Grade: 5.00/5.00

RESEARCH INTEREST

Natural Language Processing (NLP), large Language Models (LLMs), Multimodal Learning, Machine Learning (ML), Deep Learning (DL) and Low Resource Language

EXPERIENCE

- **Dhaka International University**  Jul 2024 - Present
Lecturer at Dept. of Computer Science and Engineering Dhaka, Bangladesh
 - Conducting undergraduate level courses in Computer Science and Engineering
 - Mentoring students in academic and research activities
- **Spectrum Engineering Consortium Ltd**  Sep 2023 - Oct 2023
Software Engineer Intern Dhaka, Bangladesh
 - Developed an algorithm for vehicle scheduling
 - Created backend server with proper database schemes
 - Achieved good output from the system
 - Successfully implemented UI using MERN stack

PUBLICATIONS

C=CONFERENCE, J=JOURNAL, W=WORKSHOP

- [W.5] **KCRL@DravidianLangTech 2025: Multi-View Feature Fusion with XLM-R for Tamil Political Sentiment Analysis**
M.A. Mia, F. Haq, M.T.A. Shawon, G.S.M. Mursalin, M.I. Khan
In *The Fifth Workshop on Speech and Language Technologies for Dravidian Languages* at NAACL 2025
- [W.4] **KCRL@DravidianLangTech 2025: Multi-Pooling Feature Fusion with XLM-RoBERTa for Malayalam Fake News Detection and Classification**
F. Haq, M.T.A. Shawon, M.A. Mia, G.S.M. Mursalin, M.I. Khan
In *The Fifth Workshop on Speech and Language Technologies for Dravidian Languages* at NAACL 2025
- [W.3] **Golden_Duck at #SMM4H 2024: A Transformer-based Approach to Social Media Text Classification**
M.A. Mia, M. Yahan, H. Murad, M.I. Khan
In the 9th Social Media Mining for Health Applications (SMM4H) Workshop at ACL 2024
- [W.2] **EmptyMind at BLP-2023 Task 1: A Transformer-based Hierarchical-BERT Model for Bangla Violence-Inciting Text Detection**
U Das, K Fatema, MA Mia, M Yahan, MS Mowla, MDF Ullah, A Sarker, H Murad
In Bangla Language Processing(BLP) Workshop at EMNLP 2023
- [W.1] **EmptyMind at BLP-2023 Task 2: Sentiment Analysis of Bangla Social Media Posts using Transformer-Based Models**
K Fatema, U Das, MA Mia, MS Mowla, M Yahan, M.F. Ullah, A Sarker, H Murad
In Bangla Language Processing(BLP) Workshop at EMNLP 2023
- [C.1] **Enhanced Semantic Relatedness Prediction using XLM-RoBERTa and CNNs with K-Fold Cross-Validation**
M.A. Mia, F.S. Tamim, Z.S. Taheri
Accepted at ICDMIS. In press (Springer) [\[Preprint\]](#)

- [J] **BanGRev: A novel Bengali tech gadget review corpus utilizing triple-pooling feature fusion with explainable AI**
M.A. Mia, A. Rahman, G.S.M. Mursalin, M.I. Khan, I.H. Sarker [\[Preprint\]](#)
- [J] **Code-Mixed Language Translation: An Integrated Framework with Enhanced Preprocessing and N-gram Optimization for Cross-Lingual Communication**
M.A. Mia, M. Fahim, G.S.M. Mursalin, M.J.I. Basher, H. Murad, M.R. Islam, M.I. Khan, I.H. Sarker [\[Preprint\]](#)
- [C] **Optimizing Eye Disease Classification with Vision Transformers: An Analysis of Bangladeshi Fundus Images**
A. Rahman, M. Hasan, M.A. Mia, M.I. Khan [\[Preprint\]](#)
- [W] **Transformer-Based Approaches for Multilabel Classification and Span Detection of Manipulation in Ukrainian Social Media**
A.R.Dhruba, M.M.Alam, k.S. Sayma, M.A. Mia, G.S.M. Mursalin, M.I. Khan [\[Preprint\]](#)












PROJECTS

- **DocuMind-RAG – Intelligent PDF Assistant** Apr 2025 - May 2025
Tools: Python, Streamlit, LangChain, LLMs, FAISS, Sentence-Transformers [\[G\]](#)
 - Built a conversational PDF assistant using Retrieval-Augmented Generation (RAG) for natural language querying
 - Integrated LLMa for answer generation and FAISS for efficient semantic search over embedded document chunks
 - Developed a clean, interactive UI with Streamlit for real-time PDF analysis and question answering
- **CUET Hall Service: Hall management system** Sep 2023 - Jan 2024
Tools: Laravel, MySQL, Bootstrap, CSS, PHP etc. [\[G\]](#)
 - Architected three-panel system (Admin, Student, Service Provider) using Laravel for streamlined hall service management
 - Implemented comprehensive admin panel with user management and request approval workflow features
 - Developed student portal for service requests with priority levels and real-time status tracking notifications
 - Created service provider interface with task management and automated report generation capabilities
- **T20 Score Predictor: A cricket score prediction system** Oct 2022 - Jan 2023
Tools: Python, NumPy, Pandas, Matplotlib, XGBoost, Flask, HTML/CSS etc. [\[G\]](#)
 - Built interactive web interface using Flask and implemented XGBoost regression model for cricket score predictions
 - Designed feature engineering pipeline and created user-friendly interface for real-time score forecasting
- **Fruit And Vegetable Recognition: A CNN-based classification system** Oct 2022 - Jan 2023
Tools: Python, TensorFlow, VGG-19 etc. [\[G\]](#)
 - Implemented fine-tuned VGG-19 architecture achieving 94% classification accuracy across 136 classes
 - Developed robust image preprocessing pipeline and optimized model parameters for improved performance
- **Electricity Billing System: Utility management application** Jan 2021 - Feb 2021
Tools: Java, Java Swing etc. [\[G\]](#)
 - Developed comprehensive billing calculation system
 - Created automated billing generation based on consumption data

SKILLS

- **Programming Languages:** Python, C, Java, C++
- **Libraries & Frameworks:** NumPy, Pandas, Matplotlib, Seaborn, Scikit-Learn, Tensorflow, Pytorch, Flask
- **Version Control:** Git
- **Database Systems:** MySQL
- **Problem Solving:** Solved 800+ problems (Codeforces, HackerRank, AtCoder, etc.)
- **Tools & Software:** MS Word, PowerPoint, Microsoft Excel

HONORS AND AWARDS

- **Best paper award** Dec 2023
Association for Computational Linguistics 
 - Title: EmptyMind at BLP-2023 Task 1: A Transformer-based Hierarchical-BERT Model for Bangla Violence-Inciting Text Detection
- **Technical Scholarship [Funded my undergraduate studies]** 2019 - 2024
Chittagong University of Engineering and Technology 
- **1st at FakeDetect-Malayalam, detecting fake news in Malayalam news articles across five categories** Nov 2024
DravidianLangTech-2025, NAACL 2025 
- **2nd at Political Multiclass Sentiment Analysis of Tamil X (Twitter) Comments** Nov 2024
DravidianLangTech-2025, NAACL 2025 
- **6th at Fake News Detection in Dravidian Languages: Classify social media posts into original or fake** Nov 2024
DravidianLangTech-2025, NAACL 2025 
- **5th in Marathi Subtask at SemEval 2024 Task 1: Semantic Textual Relatedness (STR)** Sep 2023
The 18th International Workshop on Semantic Evaluation, NAACL 2024 
- **7th in Telugu Subtask at SemEval 2024 Task 1: Semantic Textual Relatedness (STR)** Sep 2023
The 18th International Workshop on Semantic Evaluation, NAACL 2024 
- **9th at BLP Shared Task 1: Violence Inciting Text Detection (VITD)** Aug 2023
Bangla Language Processing(BLP) Workshop at EMNLP 2023 
- **7th at BLP Shared Task 2: Sentiment Analysis of Bangla Social Media Posts** Aug 2023
Bangla Language Processing(BLP) Workshop at EMNLP 2023 
- **1st at Sentiment Analysis of Code-Mixed Text CLBLP 2023 (Task 2)** Feb 2023
International Workshop on Computational Linguistics and Bangla Language Processing (CLBLP 2023) 
- **6th at Multimodal Sentiment Analysis CLBLP 2023 (Task 1)** Feb 2023
International Workshop on Computational Linguistics and Bangla Language Processing (CLBLP 2023) 

VOLUNTEER EXPERIENCE

- **Research Mentor** 2024 - Present
Khan Computational Research Lab (KCRL)
 - Mentor undergraduate students in thesis research and academic writing
 - Guide research methodologies and technical implementations
- **Adviser** 2025 - Present
DIU Computer Programming Club (CPC)
 - Guide students in organizing programming competitions and tech events
 - Mentor club members in developing their programming and problem-solving skills
- **Adviser** 2024 - Present
DIU Film & Photography Club
 - Guide students in organizing photography exhibitions and film festivals
 - Mentor club members in developing their creative and technical skills
- **Financial Secretary** 2022 - 2023
RAJUK Ex- Students' Association (RESA), CUET
 - Managed financial operations and budget allocation for association activities
 - Maintained transparent financial records and coordinated fundraising initiatives

CERTIFICATIONS

- **AI Agents Fundamentals** by Hugging Face March 2025
- **The Reasoning Course - Building AI models that reason** by Hugging Face March 2025
- **Python Matplotlib** by Great Learning Academy Nov 2023
- **Data Visualization** by kaggle Nov 2023
- **Pandas** by kaggle Nov 2023
- **Neural Networks and Deep Learning** by DeepLearning.AI Oct 2023

ADDITIONAL INFORMATION

Languages: Bengali (Native), English (Professional Working Proficiency)

Interests: Exploring new things, Reading research papers & blogs, Listening songs, Watching Movies