Md. Ayon Mia

📞 +8801635690439 | ✉ mdayon.cse@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

[J.1] **Enhanced Semantic Relatedness Prediction using XLM-RoBERTa and CNNs with K-Fold Cross-Validation**
<u>M.A. Mia</u>, F.S. Tamim, Z.S. Taheri, M.A Talukder
The Journal of Engineering (JoE), Wiley and The Institution of Engineering and Technology, 2025

[C.7] **BANMIME: Misogyny Detection with Metaphor Explanation on Bangla Memes**
<u>M.A. Mia</u>, A.M.R. Mazumder, K.S. Sayma, M. Fahim, M.T.H. Fuad, A.M.M. Rahman, M.I. Khan
*Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2025**)*

[C.6] **BANHATEME : Understanding Hate in Bangla Memes through Detection, Categorization, and Target Profiling**
<u>M.A. Mia</u>, M. Fahim
Second Bangla Language Processing Workshop (BLP 2025), co-located with ICJNLP-AACL 2025 **Accepted[Preprint]**

[C.5] **Prompting vs Fine-Tuning for Bengali Violence Detection in Low-Resource Settings**
<u>M.A. Mia</u>, A.Hasan, A. Rahman, G.S.M. Mursalin, M.I. Khan
28th International Conference on Computer and Information Technology (ICCIT 2025) **Accepted [Preprint]**

[C.4] **SETBoost: An Interpretable Machine Learning Approach for Predicting Software Employee Turnover Tendency**
A. Rahman, M.I. Hossain, <u>M.A. Mia</u>, M.I. Khan
28th International Conference on Computer and Information Technology (ICCIT 2025) **Accepted [Preprint]**

[C.3] **Vision Transformers for Multi-Class Eye Disease Classification: Enhancing Early Detection in Resource-Constrained Healthcare Settings**
A. Rahman, M. Hasan, <u>M.A. Mia</u>, M.I. Khan
3rd International Conference on Big Data, IoT and Machine Learning (BIM 2025) **Accepted** [Preprint]

[C.2] **Fraud Detection with DistilBERT: A Transformer-Based Approach to Behavioral Banking Sequences**
M.M. Hasan, A.A. Mahin, S. Chakraborty, M. Afrose, <u>M.A. Mia</u>, M.A. Based
3rd International Conference on Big Data, IoT and Machine Learning (BIM 2025) **Accepted** [Preprint]

[C.1] **Explainable Self-Attentive Transformer Model for Bangla Mental Health Disorder Detection**
A.S.H. Hridoy, N.S. Shohan, M.S. Alif, N.M.A. Mahin, <u>M.A. Mia</u>
International Conference on Intelligent Data Analysis and Applications (IDAA 2025) **Accepted** [Preprint]

[W.8] **KCRL@DravidianLangTech 2025: Multi-View Feature Fusion with XLM-R for Tamil Political Sentiment Analysis**
<u>M.A. Mia</u>, 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.7] **KCRL@DravidianLangTech 2025: Multi-Pooling Feature Fusion with XLM-RoBERTa for Malayalam Fake News Detection and Classification**
F. Haq, M.T.A. Shawon, <u>M.A. Mia</u>, G.S.M. Mursalin, M.I. Khan
In The Fifth Workshop on Speech and Language Technologies for Dravidian Languages* at NAACL 2025

[W.6] **PhantomTroupe@CASE 2025: Multimodal Hate Speech Detection in Text-Embedded Memes using Instruction-Tuned LLMs**
Amin, M.A. Horaira, M.T.A. Shawon, <u>M.A. Mia</u>, M.I. Khan
The 8th Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text |
Accepted [Preprint]

[W.5] **ID4Fusion@CASE 2025: A Multimodal Approach to Hate Speech Detection in Text-Embedded Memes Using ensemble Transformer based approach**
T.B. Rashfi, M.T.A. Shawon, <u>M.A. Mia</u>, M.I. Khan
The 8th Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text |
Accepted [Preprint]

[W.4] **CUET_KCRL at CheckThat! 2025: EnsembleNet with RoBERTa-Large for Subjectivity Detection in News Articles**
M.T.A. Shawon, F. Haq, <u>M.A. Mia</u>, G.S.M. Mursalin, M.I. Khan
CLEF 2025 Conference and Labs of the Evaluation Forum

[W.3] **Golden_Duck at #SMM4H 2024: A Transformer-based Approach to Social Media Text Classification**
<u>M.A. Mia</u>, 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, <u>MA Mia</u>, M Yahan, MS Mowla, M.F. 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, <u>MA Mia</u>, MS Mowla, M Yahan, M.F. Ullah, A Sarker, H Murad
In Bangla Language Processing(BLP) Workshop at EMNLP 2023

UNDER REVIEW PAPERS   C=CONFERENCE, J=JOURNAL, W=WORKSHOP

[J] **Code-Mixed Language Translation: An Integrated Framework with Enhanced Preprocessing and N-gram Optimization for Cross-Lingual Communication**
<u>M.A. Mia</u>, M. Fahim, G.S.M. Mursalin, M.J.I. Basher, H. Murad, M.R. Islam, M.I. Khan, I.H. Sarker
Engineering Applications of Artificial Intelligence

[J] **BanGRev: A novel Bengali tech gadget review corpus utilizing triple-pooling feature fusion with explainable AI**
<u>M.A. Mia</u>, A. Rahman, G.S.M. Mursalin, M.I. Khan, I.H. Sarker
New Generation Computing

[J] **Addressing Rating-Review Discrepancy in a Novel Dataset: An Explainable Pharmaceutical Product Evaluation using Multi-Stream Attention Transformer**
A. Rahman, <u>M.A. Mia</u>, M.I. Khan, I.H. Sarker
PLOS ONE

[J] **BanCyB : A Hybrid Transformer-LSTM Framework for Explainable Multi-Label Cyberbullying Detection in Low-Resource Setting**
M.M. Hasan, M.S. Arefin, <u>M.A. Mia</u>, P.K. Dhar
Discover Applied Sciences (First review completed)

PROJECTS

- **DocuMind-RAG – Intelligent PDF Assistant** *Apr 2025 - May 2025*
  Tools: Python, Streamlit, LangChain, LLMs, FAISS,Sentence-Transformers [🔗]
  - 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. [🔗]
  - 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. [🔗]
  - 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. [🔗]
  - 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. [🔗]
  - 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*

  ◦ 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

• **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

• **Fundamental of LLMs** by Hugging Face *May 2025*
• **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