

Shayekh Bin Islam

[@shayekhbinislam](#) [Google Scholar](#) [in @shayekhbinislam](#) [@ shayekh.bin.islam@gmail.com](mailto:shayekh.bin.islam@gmail.com) [+880-1740595591](tel:+880-1740595591)

Education

Bangladesh University of Engineering and Technology (BUET)

March'18 - May'23

B.Sc. (Engg.) (Honours), Computer Science and Engineering

CGPA: 3.78/4.00

Relevant Coursework: Machine Learning, Artificial Intelligence, Linear Algebra, Statistics, Calculus, Numerical Methods

Publications & Preprints

*=Equal Contribution

- [6] **Open-RAG: Enhanced Retrieval-Augmented Reasoning with Open-Source LLMs** [[Preprint](#)] [[Website](#)]
[Shayekh Bin Islam](#)^{*}, Md Asib Rahman^{*}, K S M Tozammel Hossain, Enamul Hoque, Shafiq Joty, Md Rizwan Parvez
[Accepted in EMNLP 2024 (Findings)]
- [5] **M-RewardBench: Evaluating Reward Models in Multilingual Settings** [[Preprint](#)] [[Website](#)]
Srishti Gureja^{*}, Lester James V. Miranda^{*}, [Shayekh Bin Islam](#)^{*}, Rishabh Maheshwary^{*}, Drishti Sharma, Gusti Winata, Nathan Lambert, Sebastian Ruder, Sara Hooker, Marzieh Fadaee
[Under review in ARR towards ACL 2025 (Score: Avg 3.67, Metareview 4.0/5.0)]
- [4] **MM-Eval: A Multilingual Meta-Evaluation Benchmark for LLM-as-a-Judge and Reward Models** [[Preprint](#)]
Guijin Son, Dongkeun Yoon, Juyoung Suk, ..., [Shayekh Bin Islam](#), ..., Seungone Kim et al
[Submitted in NAACL 2025]
- [3] **INCLUDE: Evaluating Multilingual Language Understanding with Regional Knowledge** [[Preprint](#)]
Angelika Romanou, Negar Foroutan, ..., [Shayekh Bin Islam](#), ..., Marzieh Fadaee, Sara Hooker, Antoine Bosselut et al
[Accepted in ICLR 2025 (Spotlight)]
- [2] **Improving Offline Meta-Reinforcement Learning With Ensemble-Diversified Actor-Critic** [[Manuscript](#)]
[Shayekh Bin Islam](#), Haque Ishfaq, M Rahman
[Submitted to AISTATS 2024]
- [1] **bbOCR: An Open-source Multi-Domain OCR Pipeline for Bengali Documents** [[Preprint](#)] [[Website](#)]
Imam Mohammad Zulkarnain^{*}, [Shayekh Bin Islam](#)^{*}, ..., Md Nazmuddoha Ansary, Asif Sushmit, Farig Sadeque et al
[Submitted to EAAI-AAAI 2024]

Research Experience

Cohere For AI & U of Copenhagen | Evaluating Multimodal LLMs Across Language Boundaries Nov'24 - Now
Advisors: [Marzieh Fadaee](#), [Sara Hooker](#), [Desmond Elliot](#) | Keywords: Multilinguality, VLM Evaluation [[Code](#)]

> Benchmarking the reasoning capability of large vision-language models across 18 languages. [Towards NeurIPS'25]

Sakina.AI, QCRI & Mila | Sakina: An LLM-based Mental-Health Support System July'24 - Now
Advisor: [Ines Arous](#), [Safa Messaoud](#) | Keywords: Patient Simulation, Multi-Agent LLM [US Patent Filing in Progress]

> Developing high-quality patient simulation and synthetic datasets towards CBT and psychoanalysis to assist therapists.

> Adapting existing language models into Arabic culture to build a robust multi-agent diagnosis system.

University of Michigan & Fatima Fellowship | Learning from Instance Gradients Aug'23 - Dec'23
Advisor: [Mohamed El Banani](#) | Keywords: Optimization, Byzantine ML, Data Pruning

> Developing a gradient-based optimization algorithm that selectively back-propagates based on the quality of the gradient signal of each instance in a mini-batch.

> Unifying the ideas of Byzantine ML and Data Pruning methods at the mini-batch level.

Mila | Improving Offline Meta-Reinforcement Learning With Ensemble-Diversified Actor-Critic Mar'22 - Dec'22
Advisor: [H Ishfaq](#), [M Rahman](#) | Keywords: Offline RL, Meta-RL, Uncertainty Estimation [[Manuscript](#)]

> Proposing uncertainty-aware learning in offline meta-RL setup based on ensemble diversity.

Dreamfold.AI & Mila | Foundation Models and Sampling Methods for Protein Design Oct'23 - Jan'24
Advisor: [R Islam](#) | Keywords: RL, Protein Foundation Models, Multi-modality

> Developing RL environments and reward models for protein structure and sequence design. [[Code](#)]

- Bengali.AI & Apsis S. | bbOCR: An Open-source Multi-Domain OCR Pipeline for Bengali Documents** June'23 - Dec'23
Advisor: Farig Sadeque | Keywords: Document OCR, Geometric & Illum. Correction
- > Developing bbOCR: an open-source scalable document OCR system that can reconstruct Bengali documents into a structured searchable digitized format.
 - > Taking geometric distortions and illumination variation into account. [[🌐 Preprint](#)]
- Brac U. | A Scalable Bitcoin Blockchain Network by Dynamic Block Size Adjustment using RL** June'23 - Dec'23
Thesis Co-Supervisor | *Advisor: MGR Alam* | Keywords: Offline RL, Inverse RL
- > Proposing a dynamic block size adjustment system that maximizes the miner's revenue based on the condition of the blockchain network; and exploring online, offline, and inverse RL algorithms in this domain.
- BUET | Bangla Grammatical Error Detection Leveraging Transformer-based Token Classification** Feb'23 - Aug'23
Undergrad Project | *Advisor: Dr. Mohammed Eunus Ali* | Keywords: NLP, GED, GEC, ELECTRA, XLM-RoBERTa
- > Developing models for detecting grammar, spelling, and punctuation errors in Bangla text using transformer-based architectures, ensemble learning, and rule-based methods. [[🌐 Report](#)] [[🌐 Code](#)]
 - > Winning solution for Bhashabhrom: Bangla Grammatical Error Detection Challenge

Professional Experience

- Cohere For AI** [[🌐 Site](#)] June'24 -
Community Researcher
- > Working on multilingual preference alignment for generative, classifier, and DPO reward models.
 - > Developing large-scale multilingual and multimodal reasoning benchmarks.
- Qatar Computing Research Institute (QCRI), HBKU** [[🌐 Site](#)] Jan'24 -
Research Assistant (Remote)
- > Proposing an open-source RAG framework capable of complex reasoning tasks.
 - > Developing efficient Mixture-of-Experts RAG models from dense models.
- Computer Science and Engineering, Brac University** [[🌐 Site](#)] June'23 - May'24
Visiting Lecturer
- > Instructing Artificial Intelligence, Robotics, Software Engineering, and Data Communications courses.
 - > Designing Basic ROS, Kinematics, and PID Control assignments using Husky, UR5, and Hector Drone simulators.
 - > Supervising undergrads working on NLP, Reinforcement Learning, and Robotics toward theses and projects.
- Data Science and Engineering Research Laboratory (DataLab), BUET** Jan'20 - June'20
Undergraduate Researcher
- > Curating datasets for a health-care chatbot in Bengali language using machine learning and NLP.

Academic Service

- Reviewer** NeurIP'25, ICML'25, ACL ARR Dec'24, COLING'25, AISTATS'25, ICLR'25, NeurIPS'24, {Sys2-Reasoning, MATH-AI, SafeGenAI, BehavioralML}@NeurIPS'24, {LLM-Cognition, AutoRL}@ICML'24, {AI4ED, NuCLear}@AAAI'24, ACL-BEA'23
- Member** Association for Computational Linguistics (ACL) 2020–, ACL SIGEDU 2023–
- In-Person Attendance** EMNLP'24 (Miami, FL)
- Virtual Volunteer** NeurIPS'21, ICLR'21, ICML'21, ACL'21, NAACL'21, ACL'22, ICLR'22, ICML'22, ICLR'23
- Co-founder** Robotics and Intelligent Systems Research Group (RIS), Brac University 2023
- Co-founder** BUETSEC: BUET Cyber Security Club 2022
- Organizer** Intra-BUET Capture the Flag (CTF) Competition and Workshop 2023

Other Projects

Expedition Aya 2024 | Cohere for AI | Mentors: Marzieh Fadaee, Sara Hooker July'24 - Aug'24
Won **Silver Prize & Most Promising Project** | Keywords: Knowledge Distillation, Data Filtering, Multi-modality, Perplexity
> **Maya**: Multimodal Aya – curating a new multilingual image-text dataset and finetuning Aya/SigLIP models.
> Identifying mislabeled instances in Aya datasets (instruction and red-teaming datasets).

First NLP Hackathon, Bangladesh | Complex Named Entity Recognition in Bengali Texts Feb'23
Second Runner-Up solution | Keywords: NLP, NER, ELECTRA, XLM-RoBERTa
> Identifications and classification of named entities in Bengali language texts. The project is designed to be used in various applications such as text analytics and information retrieval systems. [[Report](#)] [[Code](#)]

Udacity | People Counter App Mar'20 - June'20
Intel® Edge AI for IoT Developers **Scholarship Project** | Keywords: SSD, Computer Vision, IoT, Edge AI
> Detection of people in a designated area, providing the number of people in the frame, average duration of people in the frame, and total count.
> Leverages deep learning frameworks like Tensorflow, PyTorch, and Caffe combined with Intel's OpenVINO toolkit to impact a real-world computer vision application with an accuracy of about 75% using the SSD512 model.

Technical Skills

Languages Python, Java, C/C++, Shell, JavaScript, HTML/CSS, Matlab, Intel 8086 Assembly Language, MIPS Assembly

ML Tools PyTorch, Hugging Face, Spacy, Tensorflow, OpenCV, NLTK, Pandas, Matplotlib, Scikit-Learn, OpenAI Gym

Development Tools Slurm, Latex, Git, Docker, VS Code, Visual Studio, IntelliJ Idea, Kubernetes, Postman

Talks and Presentations

“Evaluating Reward Models in Multilingual Settings”

> Expedition Aya Lightning Talk, Cohere For AI Community Meetup [[Slides](#)] Aug 2024 (Online, Cohere For AI)

“Introduction to Deep Learning with Keras”

> Keras Community Day Bangladesh 2023 [[Slides](#)] Oct 2023 (Brac University, Bangladesh)

“Bangla Grammatical Error Detection Leveraging Transformer-based Token Classification”

> Bhashabhrom: Bangla Grammatical Error Detection Challenge 2023 [[Slides](#)] Mar 2023 (Dept. of EEE, BUET)

Honours and Awards

Dean's List Scholarship, Bangladesh University of Engineering and Technology (BUET) 2018-19

Education Board Scholarship, Government of Bangladesh 2010-2023 (\$800 in Total)

Cohere For AI's Expedition Aya Hackathon 2024 | Silver Prize Project: Multilingual RewardBench \$4000 API Credits

Cohere For AI's Expedition Aya Hackathon 2024 | Most Promising Project: Multimodal Aya \$1000 API Credits

Fatima Fellowship 2023 | Research Collaboration, Mentoring, and Compute Support Sponsored by Hugging Face

The Trojan Detection Challenge 2023 (LLM Edition) | Compute Grant Received \$500

Bhashabhrom: Bengali Grammatical Error Detection Challenge 2023 | Champion Phase I and II, Awarded \$750

Intel® Edge AI Scholarship 2020 | 750 Seats Full Free Intel® Edge AI for IoT Developers Nanodegree Program

First NLP Hackathon BD, 2023 | Second Runner-UP Awarded \$200

Association for Computational Linguistics (ACL) | Membership 2020-21

Simulation Humanoid Robot Wrestling Competition, ICRA 2023 | Participant to the 1/8 finals

Bengali.AI Speech Recognition 2023 | 67/744 Teams (Top 10%) Kaggle Bronze Medal

National Science Olympiad, Bangladesh | Divisional Champion, Rajshahi 2016