Khondoker Ittehadul Islam

ERASMUS MUNDUS LCT MASTERS STUDENT

+49 163 3705859 | khondoker.islam9@gmail.com | khondokerislam.github.io | KhondokerIslam | -kh

Education

University of Groningen, Universität des Saarlandes — *Netherlands & Germany*

MASTER'S IN NATURAL LANGUAGE PROCESSING, LANGUAGE SCIENCE & TECHNOLOGY — *2024 - 2026 (expected)*

- Courses: **Advanced NLP** (9.2/10), **Deep Learning Practicals** (8.0/10).
- CGPA: **7.7/10**

Shahjalal University of Science and Technology — *Sylhet, Bangladesh*

B.Sc. (ENGG.) IN COMPUTER SCIENCE AND ENGINEERING — *2017 - 2021*

- Thesis: Multi-lingual BERT in Low-resource Language Tasks.
- CGPA: 3.35/4.00

Publications

- **Khondoker Ittehadul Islam**, Md Saiful Islam, and Md Ruhul Amin; Sentiment analysis in Bengali via transfer learning using multi-lingual BERT. **ICCIT 2020**.
- **Khondoker Ittehadul Islam**, Sudipta Kar, Md Saiful Islam, Mohammad Ruhul Amin; SentNoB: A Dataset for Analysing Sentiment on Noisy Bangla Texts. **Findings of EMNLP 2021**.
- **Khondoker Ittehadul Islam**, Tanvir Hossain Yuvraz, Md Saiful Islam, Enamul Hassan; EmoNoBa: A Dataset for Analyzing Fine-Grained Emotions on Noisy Bangla Texts. **AACL-IJCNLP 2022**.
- **Khondoker Ittehadul Islam**, Gabriele Sarti; Reveal-Bangla: A Dataset for Cross-Lingual Multi-Step Reasoning Evaluation. ***Under Review*** at BLP workshop @ IJCNLP-AACL 2025.

Professional Experience

2024 - 2025 **Research Assistant**, InCLow Lab, University of Groningen

- *Analyzed LLM reasoning across languages, generation strategies, and parameter scales.*
- *Applied NLI, RAG, translation, and attribution techniques in a cross-lingual evaluation.*
- *Identified higher internal reasoning attribution for low-resource languages.*
- *Publications: Submitted at BLP workshop @ IJCNLP-AACL 2025 and on-going.*

2025 - 2025 **Research Scholar**, University of Groningen

- *Improved linguistic capabilities of vLMs for OCR tasks.*
- *Composited image to augment ancient documents.*
- *Reduced image resolution through page-separation preprocessing.*
- *Heuristically pseudo-labeled multi-class unlabeled documents.*
- *Report: https://arxiv.org/pdf/2508.10356*

2025 - 2025 **Research Scholar**, University of Groningen

- *Leveraged emotional cues in debate speech for fallacy detection.*
- *Mapped tonal emotion values into textual representations to improve comprehension of audio features.*
- *Report: https://arxiv.org/pdf/2509.11127*

2022 - 2022 **Research Assistant**, Fordham University

- *Created a Bangla fake news dataset by collecting source articles.*
- *Performed manual annotation with explanatory justifications.*

2019 - 2021 **Research Scholar**, SUST NLP Research Group, Shahjalal University of Science & Technology

- *Evaluated the performance of encoder-only models on low-resourced Bangla language.*
- *Developed a corpus-level generalization metric.*
- *Designed an algorithm to assess annotation quality of 80+ annotators.*
- Publications: *Findings of EMNLP 2021*, *AACL-IJCNLP 2022*

2021 - 2024 **Software Engineer**, REVE Systems

- **Project Details**: *Automating the administrative workflow process of Bangladesh Transport Ministry.*
- **Notable Implementations**: *Developed administrative benchmark reports using DSA.*
- **Production Site**: `https://www.gems.gov.bd/`

Skills

`Technical Stack`	Python, Java, R, PyTorch, HF's Transformers, vLMs, Git, Slurm
`Interpretable ToolKit`	Seaborn, t-SNE, inSeq, ContextCite, LogitLens
`Research Stack`	Paper Reviewing, Academic Writing, Grant Writing, Project Management
`Development Stack`	Spring Boot, Hibernate, React, Redis, CI/CD, RESTful APIs, Distributed System

Hackathons

2024 **3rd Turku Technical Hackathon** ⬀, RESPONSE
Participated to automate PV energy minimization at the residential scale.

Awards

2025 **Best NLP Student Project Award**, @GroNLP, University of Groningen
Competed with 7 student teams in conducting and producing quality research.

2025 **LxMLS Scholarship**, Google & Zendesk
Received expenses to participate in Lisbon Summer School.

2025 **Erasmus+ Scholarship**, Erasmus Mundus
Stipend to support students in their traineeship.

2024 **NL Scholarship**, Nuffic, Dutch Ministry of Education, Culture, and Science
Grant to support researchers in Master's program.

2024 **Consortium Scholarship**, Erasmus Mundus LCT Program
Received reduced annual program fee for Joint Master's Degree students.

2023 **Outstanding Reviewer**, BLP Workshop @ EMNLP 2023
Provided feedback to assist program chairs and to help authors in improving their work.

2021 **Student Travel Scholarship**, EMNLP 2021
Received expenses to attend the conference.

Academia Services

2024-2025 **Programme Committee Member**, ReMa Linguistics, CLCG, University of Groningen
Represented student cohort and provided insights on course structure to the Dutch Education Ministry.

Complementary Education

2025 **15th LxMLS Summer School**, Instituto Superior Técnico (IST)
Courses: Learning with Linear Models, Introduction to Sequence Models, Pretrained Transformer Models, Causality

2020 **Fundamental of Reinforcement Learning**, Alberta Machine Intelligence Institute via Coursera (University of Alberta)
An online non-credit course authorized by University of Alberta, Alberta Machine Intelligence Institute and offered through Coursera. CREDENTIAL ID: VYBJSWD5UKCH