

Khondoker Ittehadul Islam

FINAL-YEAR EM-LCT MASTERS STUDENT

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

Education

University of Groningen, Saarland University

DOUBLE DEGREE MASTER'S IN NATURAL LANGUAGE PROCESSING

Groningen & Saarbrücken

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

B.Sc. (ENGG.) IN COMPUTER SCIENCE AND ENGINEERING

Sylhet, Bangladesh

2017 - 2021

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

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. **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: BLP workshop @ IJCNLP-AACL 2025 [🔗](#) and on-going.

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

- Digitalized the administrative workflow process of Bangladesh Transport Ministry.
- Incorporated Theory-of-Computation (ToC) concepts to automate the chain-of-decisions.
- Production: <https://www.gems.gov.bd/> [🔗](#)

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

- Evaluated the performance of neural network 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 [🔗](#)

Student Projects

OCR on Historical Documents

2025

University of Groningen

Groningen, Netherlands

- 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.pdf> [🔗](#)

Fallacy Detection On Presidential Speech

2025

University of Groningen

Groningen, Netherlands

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

Fake News Detection on Low-Resourced Language

2022

Fordham University

Remote

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

Skills

Technical Stack Python, Java, R, PyTorch, HF's Transformers, vLMs, Praat, Git, Slurm

Interpretable ToolKit Seaborn, t-SNE, inSeq, ContextCite, LogitLens

Research Stack Paper Reviewing, Academic Writing, Grant Writing, Project Management

Hackathons

2025 Edge AI contest [🔗](#), Edge Impulse

Integrated a graph-based RAG system providing farm management guidance from academic sources.

Awards

2025 Best NLP Student Project Award, @GroNLP, University of Groningen

Competed with 7 other student teams to produce impactful research.

2025 Google & Zendesk Scholarship, LxMLS

Received expenses to participate in Lisbon ML Summer School.

2024 NL Scholarship, Nuffic, Dutch Ministry of Education, Culture, and Science

Awarded stipend to support education.

2024 Consortium Scholarship, Erasmus Mundus LCT Program

Received reduced annual program fee.

2023 Outstanding Reviewer, BLP Workshop @ EMNLP 2023

Assisted program chairs and authors through constructive feedback.

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 course structural insights 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

2025 Erasmus Mundus LCT Summer School 2025, DFKI & Saarland University

Workshops: Low-Resource Machine Translation, Interpreting Transformer Internals with Sparse Auto-encoders

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