

Khondoker Ittehadul Islam

✉ khondoker.islam9@gmail.com | 🏠 khondokerislam.github.io | 💻 github.com/KhondokerIslam | 🔗 linkedin.com/in/-kh/ | 🎓 Google Scholar

Education

University of Groningen

Master of Arts in Language and Communication Technologies (LCT)

Groningen, Netherlands

September 2024 - Present

Shahjalal University of Science and Technology

B.Sc. (Engg.) in Computer Science and Engineering

Sylhet, Bangladesh

Jan 2017 - Apr 2021

- **Thesis:** Examining the generalization capabilities of Multi-lingual Language models in low-resource settings
- CGPA: 3.35/4.00

Publications

- **Khondoker Ittehadul Islam**, Sudipta Kar, Md Saiful Islam, and Mohammad Ruhul Amin, “SentNoB: A Dataset for Analysing Sentiment on Noisy Bangla Texts,” *Findings of EMNLP 2021* (acceptance rate: 11.2%).
- **Khondoker Ittehadul Islam**, Tanvir Hossain Yuvraz, Md Saiful Islam, and Enamul Hassan, “EmoNoBa: A Dataset for Analyzing Fine-Grained Emotions on Noisy Bangla Texts,” *AACL-IJCNLP 2022* (acceptance rate: 26.5%).
- **Khondoker Ittehadul Islam**, Md Saiful Islam, and Md Ruhul Amin, “Sentiment analysis in Bengali via transfer learning using multi-lingual BERT,” *ICCIT 2020*.

Professional Experience

Software Engineer

REVE Systems, E-Governance Solutions

April 2021 - August 2024

- > Optimized **Thymeleaf's caching strategy** to accommodate frequent user role changes, achieving an **80% server response time** reduction for **≈600 users** across **365+ roles**.
- > Collaborated with client, business, and UI/UX team to design and implement several **vehicle servicing workflow** architectures adopted by **64+ regional government offices**.
- > Enhanced operational insights for **11 vehicle management modules**, including inventory and complaint, with **versatile analytics**.

Java Spring Boot Spring Security Hibernate Thymeleaf Gradle

Research Experience

- Analyzing Sentiment on Low-Resource Multi-Dialect Dataset ([pdf](#) / [dataset](#) / [code](#) / [video](#))
 - Examined the **negative results** of **popular pre-trained language models** in low-resource texts
 - Presented a **diversity estimation method** of the corpus to demand generalized classification methods
- Detecting Multi-Label Emotion using LLMs on Multi-Dialect Dataset ([pdf](#) / [dataset](#) / [code](#))
 - Error analyzed the **poor generalization performance** of LLMs in **multi-dialect settings**
 - Formulated a **heuristic technique** to determine the **annotation quality** of 80+ annotators

Skills

Programming Languages Java (Fluent), Python (Fluent), JavaScript (Fluent), C++ (Competent), R (Competent)

Frameworks & Libraries PyTorch, HuggingFace, Scikit, NLTK, Spring Boot, Spring MVC, Spring Cloud, Hibernate

Awards and Scholarships

- 2024 **Consortium Scholarship**, Erasmus Mundus
- 2024 **Holland Scholarship**, Nuffic
- 2023 **Outstanding Reviewer**, BNLN @ EMNLP 2023
- 2021 **Travel Scholarship**, EMNLP 2021
- 2017 **Emerging Robotics Participant**, SUST Robotics Club



Professional Services

2024-present **PC Member**, ReMa Linguistics

2021 **Student Volunteer**, EMNLP 2021

2013-2016 **Junior Reserved Officer's Training Corps (JROTC)**, SSHS

Programming Records

2018	ACM ICPC Dhaka Regional Preliminary 2018 , Team: SUST_RedPanda,  codemarshal.org/preli-18	<i>Rank: 102</i>
2017	ACM ICPC Dhaka Regional Preliminary 2017 , Team: SUST_STS,  codemarshal.org/preli-17 CodeForces , UserName: _kh CodeChef , UserName: ettehad97	<i>Rank: 218</i> <i>Max Rate: 1376</i> <i>Max Rate: 1648</i>

Linguistic Skills

Bangla	Native	
English	Intermediate	<i>IELTS 7</i>
Hindi	Conversational	
Arabic	Reading and Writing	