

# Mehedi Hasan Bijoy

mhb6434@gmail.com

[LinkedIn](#) • [GitHub](#) • [Google Scholar](#) • [Portfolio](#)

## Experience

- Mar 24 - May 25    Research Assistant**  
[Automatic Speech Recognition Research Group](#), Aalto University  
**Project:** Knowledge Distillation for Large Audio-Language Model Compression
- Multilingual Speech Emotion Recognition
  - Accented English Speech Recognition (ongoing Master's thesis)
- Outcome:** Co-authored peer-reviewed publication in ACMM/MuSe2024.
- Jan 23 - Oct 23    Lecturer**  
Department of CSE, Bangladesh University of Business and Technology  
**Courses:**
- CSE103 - Discrete Mathematics
  - CSE111 & CSE112 - Structured Programming Language and Lab
  - CSE121 & CSE122 - Object Oriented Programming Language and Lab
  - CSE231 & CSE232 - Data Structures and Lab
  - CSE467 - Pattern Recognition
- Mar 22 - Jun 23    Research Assistant**  
Institute for Advanced Research, United International University  
**Project:** Development of Bangla Spell and Grammar Checker  
**Outcomes:** Co-authored peer-reviewed publications in *Computer Speech & Language*, *Neural Computing and Applications*, and *EMNLP/BLP2024*.
- Feb 22 - Jan 23    Lab Instructor**  
Department of Electrical & Computer Engineering, North South University  
**Courses:**
- CSE115L - Programming Language I Lab
  - CSE225L - Data Structures and Algorithms Lab
- Oct 20 - Sep 21    Teaching Assistant**  
Department of Mathematics & Physics, North South University  
**Course:** MAT361 - Probability & Statistics

## Academic Credential

- M.Sc. in Computer, Communication and Information Sciences  
Aalto University, Finland Sep 2023 - May 2025  
CGPA: 4.73\* / 5.00  
Major: **Speech and Language Technology**  
Minor: **Machine Learning, Data Science and Artificial Intelligence**
- B.Sc. in Computer Science & Engineering  
North South University, Bangladesh May 2017 - Sep 2021  
CGPA: 3.81 / 4.00  
Specialization: **Artificial Intelligence**

## Programming Skills

- **Languages:** Python, C, C++, SQL, Java
- **DL Frameworks:** PyTorch, HuggingFace, LangChain
- **ML Libraries:** Scikit-learn, NLTK, Gensim, Pandas, Numpy, Matplotlib, Seaborn
- **Web Development Tools:** FastAPI, Gradio
- **Data Visualization Tool:** Tableau
- **Web Scraping Libraries:** Selenium, BeautifulSoup, Requests, Urllib, Wikipedia-API
- **Developer Tools:** LaTeX, Git, Google Colab, Jupyter Notebook, Eclipse, SSMS, SQL Workbench

## List of Relevant Courses

- Software Engineering with Large Language Models, CS-AJ1040, Aalto University, Finland.
- Large Language Models, ELL881/AIL821, IIT Delhi, India. <Audited>
- Statistical Natural Language Processing, ELEC-E5550, Aalto University, Finland.
- Machine Learning, CS-C3240, Aalto University, Finland.
- Deep Learning with Python, CS-EJ3311, Aalto University, Finland.
- Deep Learning, CS-E4890, Aalto University, Finland.
- Speech Processing, ELEC-E5500, Aalto University, Finland.
- Speech Recognition, ELEC-E5510, Aalto University, Finland.
- Supervised Machine Learning, CS-E4715, Aalto University, Finland.
- Natural Language Processing, CSE495, North South University, Bangladesh.

## Open-Source LLM Projects

- **A web application for general questions and answers.** ([code](#))  
Developed using Mistral-7B-instruct, Gradio and FastAPI.
- **RAG-Powered PDF Query System.** ([code](#))  
Developed using Llama2.

## Awards & Scholarships

- Finland Scholarship for graduate studies at Aalto University, Finland.
- Summa Cum Laude distinction in Bachelor of Science.
- Partial tuition fee waiver grant for undergraduate studies at North South University, Bangladesh.

## Standard Test Scores

- **International English Language Testing System (IELTS)**  
**Overall: 7.0 | Speaking - 8.0**  
**Date of Exam - October 29, 2022**

## Research Interests

Knowledge Distillation, Multilinguality, Large Language Models, Speech & Language Processing

## Publication († indicates equal contributions)

- **Multimodal Humor Detection and Social Perception Prediction**  
ACM Multimedia 2024, MuSe 2024 / [Code](#)  
Authors: Mehedi Hasan Bijoy†, Dejan Porjazovski†, Nhan Phan†, Guangpu Huang†, Tamás Grósz, and Mikko Kurimo
- **A transformer-based spelling error correction framework for Bangla and resource scarce Indic languages**  
Computer Speech & Language / [Paper](#) / [Code](#)  
Authors: Mehedi Hasan Bijoy†, Nahid Hossain†, Salekul Islam, and Swakkhar Shatabda
- **Panini: A Transformer Based Grammatical Error Correction Method for Bangla**  
Neural Computing and Applications / [Paper](#) / [Code](#)  
Authors: Nahid Hossain†, Mehedi Hasan Bijoy†, Salekul Islam, and Swakkhar Shatabda
- **Advancing Bangla Punctuation Restoration by a Monolingual Transformer-based Method and a Large-Scale Corpus**  
EMNLP 2023, BLP 2023 / [Paper](#) / [Code](#) / [Slides](#) / [Oral Presentation](#)  
Authors: Mehedi Hasan Bijoy†, Mir Fatema Afroz Fariat†, Mahbub E Sobhani, Tanzid Ferdoush, and Swakkhar Shatabda
- **Towards Sustainable Agriculture: A Novel Approach for Rice Leaf Disease Detection using CNN and Enhanced Dataset**  
IEEE Access / [Code](#) / [Paper](#)  
Authors: Mehedi Hasan Bijoy, Nirob Hasan, Mithun Biswas, Suvodeep Mazumdar, Andrea Jimenez, Faisal Ahmed, Mirza Rasheduzzaman, and Sifat Momen
- **Image Tagging by Fine-tuning Class Semantics Using Text Data from Web Scraping**  
ICCIT 2021 / [Paper](#) / [Oral Presentation](#)  
Authors: Mehedi Hasan Bijoy†, Nirob Hasan†, Md. Tahrim Faroque Tushar, and Shafin Rahman

## Research Papers (Under Review)

- **Multi-Teacher Language-Aware Knowledge Distillation for Multilingual Speech Emotion Recognition** [[Code](#)]  
Authors: Mehedi Hasan Bijoy, Dejan Porjazovski, Tamás Grósz, and Mikko Kurimo