Last Updated on March 13, 2025

Mehedi Hasan Bijoy

mhb6434@gmail.com

linkedin.com/in/mehedi-bijoy | GitHub | Google Scholar | mehedihasanbijoy.github.io

PROFESSIONAL SUMMARY

Specialized in Large Language Models (LLMs), Automatic Speech Recognition (ASR), and MLOps. Skilled at designing, developing, and deploying end-to-end machine learning pipelines. Authored several peer-reviewed papers on small language models (SMLs), optimization techniques, and knowledge distillation, achieving recognition in top-tier journals and conferences. Highly self-motivated, detail-oriented, and committed to continuous learning.

KEY SKILLS

- Natural Language Processing
- Large Language Models
- Retrieval Augmented Generation
- Prompt Engineering & Fine-Tuning
- Automatic Speech Recognition
- Languages: Python, C, C++, JAVA
- Libraries: PyTorch, HuggingFace
- Frameworks: LangChain. Ollama
- MLOps: Docker, Kubernetes, MLflow
- Cloud Computing: AWS
- Web Dev. Tools: Flask, FastAPI
- Databases: SQL, VectorDB
- Data Visualization: Grafana, Tableau
- Web Scraping: Selenium, BS4
- Developer Tools: Git, DVC, Linux, Latex

EXPERIENCE

Mar 24 — May 25 **Research Assistant**

Automatic Speech Recognition Research Group, Aalto University, Finland

Topic: Knowledge Distillation for Large Audio-Language Model Compression

- Designed a novel multi-teacher accent-aware knowledge distillation method for accented English speech recognition, improving out-of-distribution generalizability (Ongoing Master's Thesis).
- Fine-tuned several large audio-language models, including Wav2Vec2.0 and Whisper.
- Developed a multilingual speech emotion recognition system (Under Review at INTERSPEECH).
- Conducted model interpretability analysis to enhance explainability.
- Co-authored 1 peer-reviewed publication in ACMM/MuSe 2024 on Multimodal AI.

Jan 23 — Oct 23 **Lecturer**

Department of Computer Science and Engineering, Bangladesh University of Business and Technology

- Delivered lectures for 5 courses: Pattern Recognition, Data Structures, OOP, SPL, and Discrete Math.
- Taught approximately 400 undergraduate computer science students.
- Designed lecture materials, courseworks, and lab exercises.
- Prepared exam questions, assessed scripts, provided feedback, and graded performance.
- Mentored and supervised students in problem-solving and competitive programming.

Mar 22 — Jun 23 **Research Assistant**

Institute for Advanced Research, United International University, Bangladesh

Topic: Development of Bangla Spell and Grammar Checker

- Proposed a transformer-based detector-purificator-corrector framework for Bangla spelling error correction, improving accuracy by ~10%.
- Built a deep learning-based Bangla grammar checker, achieving a ~20% performance improvement.
- Developed 2 large-scale corpora, significantly reducing Bangla's status as a low-resource language.
- Co-authored 3 peer-reviewed publications: 2 Q1 journal articles and 1 Tier-1 conference paper.

Feb 22 — Jan 23 **Lab Instructor**

Department of Electrical & Computer Engineering, North South University, Bangladesh

- Conducted hands-on programming labs on C and C++, teaching undergraduate students.
- Assisted students in debugging and optimizing their code for efficient problem-solving.
- Prepared lab exam questions, evaluated submissions, provided feedback, and graded performance.

EDUCATION

Master of Science (M.Sc.) in Computer, Communication, and Information Sciences, Aalto University — Sep 2023 – May 2025

- CGPA: 4.71* / 5.00
- Major: Speech and Language Technology
- Minor: Machine Learning, Data Science, and Artificial Intelligence
- Award & Scholarship: (I) Finland Scholarship

Bachelor of Science (B.Sc.) in Computer Science & Engineering, North South University — May 2017 – Sep 2021

- CGPA: 3.81 / 4.00
- Specialization: Artificial Intelligence
- Award & Scholarship: (I) Summa Cum Laude. (II) Merit-Based Tuition Fee Waiver Grant.

TRAINING & CERTIFICATION

- **Complete MLOps Bootcamp** — Udemy — Ongoing
- **Software Engineering with Large Language Models** — FITech — Ongoing
- **MLOps Foundations** — iNeuron — March 2025
- **IELTS Academic** (Overall Score: 7 | Speaking Score: 8) — International Development Program (IDP) — October 2022

ACADEMIC SERVICES

Tutor, Aalto University, Courses: Speech Recognition (Autumn 24) and Statistical Natural Language Processing (Spring 25).
Reviewer, Workshop on Bangla Language Processing at EMNLP 2023.
Teaching Assistant, North South University, Course: Probability & Statistics (Fall 20, Spring 21, Summer 21).

SELECTED PROJECTS

- **MLOps / End-to-End Diabetes Detection** [GitHub repository / DagsHub repository]
  Developed a complete MLOps pipeline for diabetes detection, integrating data versioning (DVC) and experiment tracking (MLflow). Automated data ingestion, preprocessing, model training, and deployment.

- **LLMOps / End-to-End RAG-Powered PDF Query System** [GitHub repository]
  Implemented a Retrieval-Augmented Generation (RAG) system leveraging Llama3-70B-instruct for semantic document search and question-answering.

SELECTED PUBLICATIONS (* Indicates Shared First Author)

Mehedi Hasan Bijoy, et al. "**A transformer-based spelling error correction framework for Bangla and resource scarce Indic languages**." Computer Speech & Language 89 (2025): 101703. https://doi.org/10.1016/j.csl.2024.101703

Mehedi Hasan Bijoy*, et al. "**Panini: a transformer-based grammatical error correction method for Bangla**." Neural Computing and Applications 36.7 (2024): 3463-3477. https://doi.org/10.1007/s00521-023-09211-7

REFERENCES (Available Upon Request)

Dr. [redacted], *Professor* in Speech and Language Processing at Aalto University, Finland.
Email: [redacted]

Dr. [redacted], *Associate Professor*, Dept. of Electrical and Computer Engineering, North South University, Bangladesh.
Email: [redacted]

Dr. [redacted], *Professor*, Department of Computer Science and Engineering, BRAC University, Bangladesh.
Email: [redacted]