I'm Sanju (he/him), a recent graduate from the Department of Computer Science at Bangladesh University of Engineering and Technology (BUET). Currently, I am an adjunct lecturer at the Department of CSE, BUET, and have prior experience as a GenAI software engineer at ResTech AI, a US-based company. My research focuses on Natural Language Processing, Information Retrieval, and Human-AI Interaction, with a special interest in low-resource and cross-lingual language processing. My career goal is to pursue a PhD in these areas, starting in Fall 2025.

I am currently working as a Research Assistant under Dr. Md. Mostofa Akbar in the Department of Computer Science and Engineering (CSE) at BUET. Our project focuses on developing a web-based platform aimed at automating the evaluation of medical exams. This involves utilizing advanced computer vision models and optical character recognition (OCR) technology to accurately extract answers from handwritten or printed exam scripts. The integration of these technologies aims to streamline and enhance the efficiency and accuracy of the medical exam grading process.

My undergraduate thesis project was under [Dr. Rifat Shahriyar](https://scholar.google.com/citations?user=p-w4hOUAAAAJ&hl=en&oi=ao). We created an *Effective Retrieval-Augmented Generation pipeline for Open Domain Question Answering in Bengali*. Currently, our manuscript (joint first-authored) is under review at COLING’25.

Outside of my academic pursuits, I have a deep passion for travel. In my free time, I enjoy either watching movies or playing the flute, which helps me relax and explore different creative outlets

Undergraduate thesis project under Dr. Rifat Shahriyar and *Abhik Bhattacharjee* (Research Assistant, Department of CSE, BUET). In this work, I co-developed the first-ever Bengali Retrieval-Augmented Generation (RAG) pipeline. This work involved benchmarking two Bengali open-domain question-answering datasets—SQuAD BN and BanglaRQA—using six state-of-the-art embedding models and three retrieval methods, covering both sparse and dense approaches. We evaluated the performance of three large language models (LLMs) with and without the RAG pipeline, analyzing their factual response capabilities on Bengali-specific data in comparison to global data.

We presented a multilingual Theory of Mind (ToM) dataset translated from a bilingual version, encompassing seven major languages, and also developed a culturally nuanced dataset. We evaluated six state-of-the-art large language models (LLMs) on both datasets to assess how cultural relevance impacts their social reasoning abilities. Our analysis focused on examining the variations in LLM performance across different languages and cultural contexts, highlighting the influence of linguistic and cultural diversity on the models' understanding of social reasoning.