

## EDUCATION

**University of Dhaka** - Institute of Information Technology, Bachelors of Science in Software Engineering(BSSE) **3.93/4.00**  
Accomplishment: Secured the 70th position out of **86,000+** students in the admission examination for University of Dhaka.

## RESEARCH GRANTS

**Advanced Research Fund** | Samsung **2024**  
During my academic internship in Samsung R&D Institute, Bangladesh, I submitted a proposal for the Advanced Research Fund (ARF). After many verification and discussions, it was finally approved by headquarters. A grant of 3 man-months for 1 year was approved for my team against the project proposal, which is equivalent to ~3000000 Taka. Further details are classified. .

## AWARDS

- Team "Software Project Lab" received "**Best innovator team**" for 1st quarter during my academic internship in Samsung. Jul, 2024
- Secured the 1st Runners-Up position in Blockchain Segment, ITVerse2.0, Bangladesh. Nov, 2023
- Secured the 2nd Runners-Up position in Notre Dame Inter-University Ideathon Contest, Bangladesh. Oct, 2023
- Secured 1st Runners-up position in Deep Learning Sprint, BUET, Bangladesh. Jun, 2023
- **Champion** of Datathon, Coding Object Private Limited, India. Dec, 2020
- Received "Rising Star" award in Intra-sport Competition, IIT, University of, Bangladesh Jan, 2020
- 8 Merit Position Awards, 6+ Music Awards and 2 Perfect Attendance Award in School

## SCHOLARSHIP

- Received **Mitsubishi UFJ Foundation Scholarship** Program FY 2024, Japan Sep,2024
- Received Govt. Scholarship in Higher Secondary School Certification Examination, Bangladesh 2015-2017
- Received Govt. Scholarship in Junior School Certification Examination, Bangladesh. 2019-2023
- Received 5 Scholarships from High School due to academic excellence 2010-2017

## PUBLICATIONS

**Where Journalism Silenced Voices: Exploring Discrimination in the Representation of Indigenous Communities in Bangladesh** Oct, 2024  
On-Review in CSCW (CORE A) Conference '25

Using both quantitative and qualitative methods, this research analyzes **2.4M** Bangla newspaper articles to find strong support for discriminatory representation and its impacts. We employed a total of 8 methods to find strong support for our hypothesis, including **LLM** and **CorEx**

**Evaluating Human Comprehension of Labeled Transition System Designs** *Rejected* in ICSE Design Workshop '25 Oct, 2024  
Developed metrics to assess human comprehension of formal design. Applied findings in a design repair tool, reducing comprehension time by 39%, highlighting the value of human-centered metrics for formal design interpretation.

**GraphRAG Pipeline on Industry Data**, Extended Abstract, *Accepted* in SBPA, Samsung R&D (Specifics are classified) May 2024

**Bengali Document Layout Analysis -- A YOLOV8 Based Ensembling Approach**, Arxiv Oct 2023  
Experimentation results on Bangla Document Layout Analysis on the BadLaD dataset in DLSprint'23 where we placed 2nd.

## EXPERIENCE

**Samsung R&D Institute, Bangladesh (SRBD)** | Engineer **2024-25**

- o Currently working in collaboration with Oxford Semantics for Ontology based Knowledge Engineering process.
- o Built a RAG pipeline on industry data that uses Knowledge-Graph to augment its performance. Since industry data with its vocabularies and project-specific meanings create a low-resource regularized domain, I trained custom vector embeddings to improve RAG results.
- o Submitted 2 SBPA extended abstracts and 1 Advanced Research Fund (ARF) proposal during my brief 6 months academic internship.
- o HQ Approved 1 SBPA extended abstract. I also secured the ARF. Contributed to R&D in tasks involving Smart Home, Fuzz and NLP.

**NSF Funded Project "ACOSUS"** | State University of New York, Old Westbury, USA **Mar 2024 - Oct 2024**

It's a joint project between 5 universities from the USA, namely Northeastern Illinois University, University of Houston - Victoria, California State Polytechnic University, Humboldt, State University of New York, Old Westbury and University of Texas at El Paso. I work on the NLP technical part where I collected Reddit data and analyzed them to analyze how students in community college are affected by social media information.

**Softanbees Technologies Limited**, Kolkata, India | Data science Intern **Nov 2021 – Dec 2021**

- o Creating models like Early Prediction of Birth Weight Based on Maternal Factors.
- o Constructing documentation for research and technical content.

- o Founding member of the research group Bangla AI Research, Technology and Application (BARTA) under Dr. Ahmedul Kabir's supervision. With our efforts, the group now has grown to 13+ members, with external collaborators from Canada, USA and industry collaborators (SoftBD Ltd).
- o Researched topics in NLP, including Bayesian language models, normalization, word embeddings, Bangla Sign Language, spelling correction, topic modeling, information retrieval, stemmers, POS tagging, and more. Long experience with LLM based research tasks. Hosted 24+ meeting
- o Established 7+ NLP and HCI projects under Dr. Kabir's supervision to onboard new undergraduates. Managed discussions with the Center for Disability Development, Bangladesh Deaf School and Chakma indigenous group for our survey tasks.
- o Trained Syntax aware Transformer, Sign Language Recognition, Sign Mimicking, News Frame classifier, Topic Model, Word Embedding, RAG, Fine-Tuned BERT, Non-negative Matrix Factorization, SVD, Small Language Model (SLM), Spelling Correction, GRU, RNN, LSTM.
- o Organized dissemination sessions to promote BARTA's research philosophy, leading to a marked increase in undergrad research engagement.
- o Supervision: Oversee junior undergraduates in Bangla Sign Language, Chakma Language, and Knowledge Graph projects, for which I created initial prototypes to help them get started and currently guide them in writing papers.

## PROJECTS

### Automated-Static-Control-Flow-Testing | Python, Angular, FastAPI

### Program Analysis

- o Built a static program analysis tool to detect data flow bugs in C programs. Parsing C code to generate static control flow and data flow graphs from scratch and then hoisting variables to test using heuristics.
- o I used it in my c assignments from the first semester and identified 5 data flow bugs.
- o Built a simple static program analysis tool that constructs AST and then identifies common code smell patterns from the AST.

### Assembly Line Management System | Versatile Garments Ltd | NodeJS, TypeScript, Angular, FastAPI, MySQL

### Machine Learning

- o Conducting experimentation to find a NMF-based approach at detecting congestion in the assembly line.
- o Performing Software Requirements Specification after communicating with Versatile Garments Ltd to design our system.

### Video Feature Based Sign Mimicking Evaluation System | Python, Angular, TypeScript, FastAPI, Keras

### NLP

- o A platform where teachers can upload their videos in sign language, students can view them and take quizzes.
- o Multiple experimentation starting from data creation to testing. Additionally, deployment in AWS Cloud Server.

### BNLP Knowledge Graph | Org Mode, Emacs Lisp, Firebase

### HCI

- o The project is aimed at creating a hierarchical graph for BNLP, with subject nodes representing key topics and paper nodes representing papers related to those topics. We used tools readily available to use for nothing to fast-prototype it.

### Bangla Sign Language Learning Platform | Collaborative Project | Angular, TypeScript, FastAPI, MySQL

### HCI

- o A platform where teachers can upload their videos in sign language, students can view them and take quizzes. Sign mimicking model.
- o Multiple experimentation starting from data creation to testing. Additionally, deployment in AWS Cloud Server. BdSL recognition model.

### Dhaka University Bus Tracking System | Collaborative Project | Azure, React, JavaScript, LeafletJS, MySQL

### Deployed

- o Under the guidance of Dr. Naushin Nowar, leading a team of six towards creating a fast and scalable bus tracking system.
- o Designed a distributed, anonymous implementation of Bus Tracking System. along with their business model.

### Bangla Text to Gloss Translation | Python

### NLP

- o Annotated 2000+ synthetic data via LLM and generated 120K+ gloss data from rulesets - the only gloss dataset in Bangla.
- o Trained/fine tuned GRU, RNN, LSTM, BART, mBERT, seq-to-seq with attention, syntax aware transformer to find the best model.

### NLP Library | C++, Gprof, gdb, Cmake

### NLP

- o Implemented language model and a wide variety of smoothing and segmentation techniques to preprocess and post-process the data.
- o Showed an application of the NLP library with a noisy channel model for spelling correction tasks and optimized existing research work.

## DEVELOPMENT SKILLSET

<b>Languages:</b>	Python, JS, Solidity, E-Lisp, TS, Java, R, C, C++, SQL, Bash, Php, Rust, NuSMV, LTSA, TLA+, Alloy, Coq
<b>Libraries and Frameworks:</b>	Angular, NodeJS, OpenCV, Hardhat, Scikit-Learn, Actix Framework(Rust), BNLTK, Solana, Metaplex, MEAN, spaCy, OpenCV, SciKit, BNLTK, Keras, PyTorch, Panda, Numpy, Matplotlib, LangChain, Ollama
<b>Development Tools:</b>	Git, Docker, Azure Cloud, Linux/UNIX, Emacs, Jupyter Notebook, Emacs, Wireshark, nmap
<b>Others:</b>	Software Requirements Specification, Database Administration, Linux System Administration

## CERTIFICATES

2023 “KolpoKoushol”, Workshop by Universal Machine Inc, Silicon Valley, California, US.

2020 “Industrial Data Science in Python”, Coding Objects Private Limited, Kolkata, India.