

RAJAN SAHA RAJU

Mailing Address:
ICT Cell, Shahjalal University of Science and
Technology, Sylhet-3114, Bangladesh

Cell: +8801740886331
Email: rajanrajusust@gmail.com
Website: <https://rajan-sust.github.io>

EDUCATION

B.Sc. in Computer Science & Engineering 2015-2019
Shahjalal University of Science and Technology, Sylhet, Bangladesh
Ranked 4th out of 65 students; CGPA: **3.81/4.00** (With Distinction)
Thesis: *Bangla Text-to-Speech (TTS)*; Advisor: Prof. Mohammad Shahidur Rahman

RESEARCH INTERESTS

Artificial Intelligence, Computational Biology, Spoken Language Processing

RESEARCH PROJECTS

Kingfisher [1], a three-stage hybrid Bangla text normalization framework for Text-to-Speech, combining LLM-based tokenization and semiotic class annotation, lexicon-driven context-aware verbalization, and automated error correction. Achieved 96% overall accuracy (95% CI: 95–97%) across diverse Bangla texts, outperforming Sparrowhawk, the only previously available Bangla normalizer.

VirusTaxo [3], a dedicated virus classifier with ~93% accuracy at the genus level across 682 viral genera; outperformed existing tools in metagenomic benchmarks, accurately identified the taxonomy of the novel virus SARS-CoV-2 even when excluded from the database, and released as an open-source framework accessible to the community.

ProtSEC [7], a training-free protein sequence embedding method in complex space using Fast Fourier Transform (FFT). Unlike protein language models (ESM2, ProtBERT, ProtT5), which require GPUs for inference, ProtSEC generates embeddings on CPU with substantially lower runtime while exceeding their performance on sequence similarity search.

AWARDS AND HONORS

Champion (National Level), AI For Bangla 2.0	2023
1st Runner-up, 10th National Undergraduate Mathematics Olympiad, Sylhet Region	2018
62nd place out of 296 teams (first quartile), ICPC Asia Dhaka Regional Contest	2018
Champion (Dhaka Division), National Creative Talent Hunt in Math and Computer	2014
Govt. scholarship (general) in Secondary School Certificate (SSC) Examination	2013
Govt. scholarship (talent pool) in Junior School Certificate (JSC) Examination	2010

PUBLICATIONS

Peer-reviewed journal papers:

- [1] **Rajan Saha Raju**, Arif Ahmad, Mohammad Shahidur Rahman. “[Kingfisher: A hybrid LLM-augmented Bangla text normalization for enhanced text-to-speech](#)”. *Natural Language Processing Journal* (Feb. 2026).
- [2] Reem Alotaibi, Madini Alassafi, Md. Saiful Islam Bhuiyan, **Rajan Saha Raju**, Md Sadek Ferdous. “[A Reinforcement-Learning-Based Model for Resilient Load Balancing in Hyperledger Fabric](#)”. *Processes* (Nov. 2022).
- [3] **Rajan Saha Raju**, Abdullah Al Nahid, Preonath Chondrow Dev, Rashedul Islam. “[VirusTaxo: Taxonomic classification of viruses from the genome sequence using k-mer enrichment](#)”. *Genomics* (Jun. 2022).
- [4] Rashedul Islam, **Rajan Saha Raju**, Nazia Tasnim, Istiak Hossain Shihab, Maruf Ahmed Bhuiyan, Yusha Araf, Tofazzal Islam. “[Choice of assemblers has a critical impact on de novo assembly of SARS-CoV-2 genome and characterizing variants](#)”. *Briefings in Bioinformatics* (Sep. 2021).

Conference papers:

- [5] Prithwiraj Bhattacharjee, **Rajan Saha Raju**, Arif Ahmad, M. Shahidur Rahman. “[End-to-End Bangla Speech Synthesis](#)”. *International Conference on Science & Contemporary Technologies* (Aug. 2021).
- [6] **Rajan Saha Raju**, Prithwiraj Bhattacharjee, Arif Ahmad, Mohammad Shahidur Rahman. “[A Bangla Text-to-Speech System using Deep Neural Networks](#)”. *International Conference on Bangla Speech and Language Processing* (Sep. 2020).

Preprint:

- [7] **Rajan Saha Raju**, Rashedul Islam. “[ProtSEC: Ultrafast Protein Sequence Embedding in Complex Space Using Fast Fourier Transform](#)”. *bioRxiv* (Oct. 2025).

WORK EXPERIENCE

Assistant Computer Programmer

Jan. 2024 - Present

ICT Cell, Shahjalal University of Science and Technology, Sylhet

- Core technical contributor of the online undergraduate admission portal
- Responsible for the development and maintenance of all major software systems, including the official website, student portal, and result processing system

Assistant Network Engineer

Oct. 2022 - Jan. 2024

ICT Cell, Shahjalal University of Science and Technology, Sylhet

- Administered authoritative DNS and maintained security of public-facing web servers
- Evaluated and deployed fiber-to-the-home (FTTH) connectivity for campus residential buildings, supervising installation to extend wireless network coverage

Bioinformatician

Jul. 2021 - Apr. 2022

Child Health Research Foundation (CHRF), Dhaka

- Maintained the Nextstrain Bangladesh build for real-time tracking of SARS-CoV-2 evolution
- Performed metagenomic data analysis, including quality control, read trimming, host DNA removal, assembly, gene prediction, taxonomic profiling, and functional annotation

Junior Software Engineer (Team: Machine Learning)

Jun. 2020 - Dec. 2020

REVE Chat, Dhaka

- Designed and implemented a web crawler to collect data from diverse sources
- Improved the accuracy of a chatbot by ~5% by implementing entity extraction, intent matching, and contextual analysis

TEACHING EXPERIENCE**Instructor**, Teaching Advanced ICT, B.Ed Program

Jan. 2026 - Jun. 2026

Institute of Education and Research (IER), SUST

- Teaching evaluation (out of 10): mean 9.7 ($\sigma=0.4$)

Programming Mentor

Jan. 2018 - Dec. 2018

Border Guard Public School & College, Sylhet

- Mentored and trained high school students for the National High School Programming Contest

PROFESSIONAL SERVICE**Journal Reviewer**: International Journal of Wavelets, Multiresolution and Information Processing (2026); Journal of Engineering, Wiley (2026).**OPEN SOURCE SOFTWARE****Kingfisher**, a hybrid LLM-augmented text normalizer for Bangla text-to-speech, released with the first public Bangla text normalization dataset.<https://github.com/Rajan-sust/BanglaTextNorm>**VirusTaxo**, a tool for taxonomic classification of viruses from metagenomic sequences.<https://github.com/omics-lab/VirusTaxo>**TECHNICAL SKILLS****Programming Languages**: Python3, R, C++, Java, Go, Bash, PHP**Bioinformatics**: BLAST, HMMER, MMseqs2, fastp, Megahit, SPAdes, samtools, Nextstrain**Libraries**: PyTorch, scikit-learn, pandas, Matplotlib, Tidyverse, SciPy, XGBoost**Databases**: MySQL, MongoDB, Vector DB (Milvus, Qdrant)**Web**: HTML, CSS, JavaScript, FastAPI, Laravel, Apache, Nginx**Networking Tools**: phpIPAM, Oxidized, LibreNMS, Cloudflare**Misc.**: Git, Linux (Ubuntu, CentOS), Docker, LaTeX, Cloud (DigitalOcean, AWS)