Souvik Ghosh

 47/11 Kallyanpur, Dhaka-1216   ✉ souvik7701@gmail.com   ☎ 01701627647   🔗 souvikghosh.netlify.app

in souvik-ghosh-b16b9a308   ⬡ Sheldor7701

About me

I am Souvik Ghosh, a motivated and detail-oriented software developer and researcher, currently pursuing an M.Sc. in Computer Science and Engineering at BUET. I am passionate about software development, machine learning, and problem-solving. I am eager to leverage my skills in research and development while contributing effectively to academia and industry.

Education

MSc. **Bangladesh University of Engineering and Technology**, Computer Science   Feb 2025 – Present
- Current Graduate Student
- Research areas: Bioinformatics, Machine Learning, Deep Learning

BSc. **Bangladesh University of Engineering and Technology**, Computer Science   Feb 2020 – Feb 2025
- CGPA: 3.83/4.0
- **Relevant Coursework:** Compilers, Operating Systems, Computer Networks, Database, DSA, Software Engineering, Information System Design

Experience

BRAC University, Lecturer, CSE Department   Jun 2025 – Present
- Teaching undergraduate courses in CSE and mentoring student projects.
- Continuing research in bioinformatics and deep learning applications.

Apurba Technologies, Full Stack Developer   Mar 2025 – May 2025
- Developed full-stack applications with ReactJS, Node.js, and PostgreSQL.
- Built scalable REST APIs and optimized CI/CD pipelines with Docker and GitHub Actions.

Publications

ResLysEmbed: A ResNet-Based Framework for Succinylated Lysine Residue Prediction Using Sequence and Language Model Embeddings   2025

Souvik Ghosh, Md Muhaiminul Islam Nafi, Mohammad Saifur Rahman

Bioinformatics Advances, vbaf198 (2025) ⬈

Research Experience

Generic Post-Translational Modification (PTM) Site Detection   Ongoing

Supervisor: Dr. Swakkhar Shatabda, Professor, BRAC University
- Developing a general framework for PTM site prediction using protein language model embeddings combined with structural data.
- Modeling proteins as graphs where each amino acid is a node initialized with PLM-based embeddings.
- Exploring Graph Neural Networks (GNNs) and related architectures for effective sequence–structure integration.

References available upon request.

- Aiming for a unified approach that can be adapted to multiple PTM types instead of task-specific models.

Alignment-Free Phylogenetic Tree Construction using Syncmers and TF-IDF Ongoing

Supervisor: Dr. Atif Hasan Rahman, Professor, BUET CSE
- Investigating a novel alignment-free method for phylogenetic inference by replacing traditional k-mers with **syncmers**.
- Introducing **TF-IDF scoring**, inspired by natural language processing, instead of raw frequency or binary presence/absence for sequence representation.
- Constructing syncmer TF-IDF matrices to capture discriminative sequence features.
- Exploring machine learning approaches to train on these representations for accurate phylogenetic tree reconstruction.

Undergraduate Thesis, CSE, BUET Published 2025

Title: ResLysEmbed: A ResNet-Based Framework for Succinylated Lysine Residue Prediction Using Sequence and Language Model Embeddings
Supervisor: Dr. Mohammad Saifur Rahman
- Published in Bioinformatics Advances (2025).
- Achieved state-of-the-art performance in succinylation site prediction, surpassing existing benchmarks.

Deepfake Detection for IEEE SP Cup 2024 Submitted

Supervisor: Dr. Mohammad Saifur Rahman
- Developed hybrid deep learning models for generalizable deepfake detection.
- Explored generative approaches, including diffusion models, GANs, and VAEs for dataset augmentation.
- Evaluated advanced backbone architectures, such as **MaxViT, ConvNeXt, and EfficientNet**, with specialized loss functions like contrastive loss and focal loss for improved accuracy.

Projects

Bits Unplugged | Interactive CS Learning Platform Github ↗ YouTube ↗ Website ↗
- Co-developed an innovative platform focused on enhancing problem-solving skills without coding, emphasizing strategy over syntax.
- Implemented features such as drag-and-drop interactive problem solving, real-time contests, personalized recommendations, and analytics dashboards.
- **Technologies:** Docker, GitHub Actions, Tailwind CSS, Sequelize, PostgreSQL, Supabase, ReactJS, NodeJS, ExpressJS, Render.com

Multimodal Breast Cancer Prognosis Prediction | Machine Learning Project Github ↗
- Designed a multimodal framework using mRNA expression, copy number alteration, and clinical data from the TCGA-BRCA dataset.
- Transitioned from MLP-based models to self-attention and cross-attention mechanisms for improved performance.
- Focused on contrastive learning techniques to handle feature embeddings without classification during training.

Awards & Achievements

Research Grant from RISE BUET

Received a research grant from BUET Research and Innovation Centre for Science and Engineering for an ongoing research project.

Deepfake Detection for IEEE SP Cup 2024

Selected as a finalist in the top 3 teams; the final competition will occur in April 2025.