Postgraduate Seminar Series

Venue: Graduate Seminar Room
Date & Time: June 20, 2026 at 3:00 PM

Speaker Information

A.K.M. Mehedi Hasan (Std No. 0422052025) is a part time M.Sc. student in the department of CSE, BUET. He completed his undergraduate studies from Bangladesh University of Engineering and Technology (BUET) in 2022. His research interest lies in the field(s) of Bioinformatics, Machine Learning, and Deep Learning. He is currently doing his postgraduate thesis under the supervision of Dr. Md. Shamsuzzoha Bayzid. He will be speaking about her ongoing research in this talk.

A TRANSFORMER-BASED MODEL FOR QUARTET INFERENCE FROM MULTIPLE SEQUENCE ALIGNMENTS USING DEEP LEARNING APPROACHES FOR PHYLOGENETIC TREE RECONSTRUCTION

Accurate phylogenetic inference from multiple sequence alignments is essential for understanding evolutionary relationships among species. However, quartet topology inference remains challenging when alignments are short, noisy, or contain insertion-deletion patterns that are often treated as missing data by traditional methods. This study investigates deep learning-based quartet inference from four-taxon multiple sequence alignments and introduces a pair-first axial transformer architecture designed specifically for phylogenetic topology prediction. The proposed model represents all pairwise taxon relationships, applies local site convolution, uses axial attention across taxon-pair and alignment-site dimensions, and incorporates learned site-weighted pooling with a topology-aware classifier. CNN-based models are first developed as neural baselines, followed by the transformer model for both un-gapped and gapped alignments. Comprehensive evaluations on simulated quartet datasets show that the proposed transformer outperforms traditional phylogenetic methods and previous CNN-based approaches across all tested settings. Notably, it achieves strong improvements on gapped alignments, reaching 90.25% accuracy for 500 bp data and 96.67% accuracy for 1000 bp data, demonstrating the usefulness of explicit gap modeling. The trained model is further applied to real biological datasets, including skink, Amborella-water lily, and turtle phylogenomic datasets, by converting model-predicted topology probabilities into weighted quartets and reconstructing species trees using wQFM. Overall, the study shows that deep learning can provide effective and biologically meaningful quartet evidence for downstream phylogenetic reconstruction.