

Dipayan Sarkar

Department of Bioinformatics, University of North Bengal
Siliguri - 734013, West Bengal, India.

☎ +91 9735490241 • ✉ dipayanskr@gmail.com • 🌐 dipayanskr.in
🐙 github.com/dipayanskr • 🏠 Google Scholar • 🆔 0000-0002-0765-5689
✉ rs_dipayans@nbu.ac.in • 🤖 huggingface.co/dipayanskr

Research Summary

PhD researcher in bioinformatics and computational biology developing deep learning methods for protein sequence modeling. Work spans sequence-based protein-protein interaction prediction using attention-based architectures with transfer learning from pretrained protein language models, generative modeling of protein binders through autoregressive and diffusion-based approaches. Also builds web platforms that make trained models usable for inference and interaction-network visualization.

Education

University of North Bengal

Ph.D. in Bioinformatics

Supervisor: Dr. Chiranjib Sarkar

Focus: Framework Development for Predicting Protein-Protein Interaction Networks Using Deep Learning Algorithms on Protein Sequences.

India

Expected March 2027

University of North Bengal

M.Sc. in Botany, 77.25% - First Class

Specialization: Biochemistry

Dissertation: "A review on the responses of aquatic macrophytes to phenolic compounds with special reference to their tolerance strategies and future implications."

Supervisor: Dr. Swarnendu Roy, DRS Department of Botany.

India

2021

Ananda Chandra College, University of North Bengal

B.Sc. (Hons.) in Botany, 63.50% - First Class

India

2019

Research Experience

University of North Bengal

PhD Research Scholar

Department of Bioinformatics

Siliguri, India

2022–Present

- Deep learning architectures for sequence-based protein-protein interaction prediction using Attention-based hybrid architecture with transfer learning.
- Sequence based generative modeling of protein binder generation using autoregressive and diffusion-based approaches.
- Development of web-based platforms for model inference and protein interaction network visualization

Publications

2025: Sarkar, D., & Sarkar, C. AttnSeq-PPI: Enhancing protein-protein interaction network prediction using transfer learning-driven hybrid attention. *Biochimica et Biophysica Acta (BBA)—Proteins and Proteomics*, 1874(1), 141102. doi:10.1016/j.bbapap.2025.141102

2026: Sarkar, D., & Sarkar, C. ARACoFusion: Uncertainty-aware calibrated deep learning for protein-protein interaction network prediction in *Arabidopsis thaliana*. *bioRxiv* preprint. doi:10.64898/2026.05.22.727120

2026: Sarkar, D., & Sarkar, C. MoE-Bind: Guiding De Novo Protein Binder Generation with Sparse Experts. *bioRxiv* preprint. doi:10.64898/2026.06.13.732043

2026: Sarkar, D., Bardhan, K., & Sarkar, C. Deep-Interact Studio: An Interactive Deep Learning Model Building Platform for Biomolecular Interaction Prediction. *bioRxiv* preprint. doi:10.64898/2026.07.02.736034

Research Software Developed

Web Servers.....

Deep-Interact Studio: Interactive web platform for building, training, and comparing custom deep learning models for protein–protein, drug–target, RNA–protein, and protein–DNA interaction prediction, with integrated interpretability.

deepinteract.compbiosysnbu.in

AttnSeq-PPI: Deep learning-based protein–protein interaction prediction platform.

attnseqppi.compbiosysnbu.in

Packages.....

EGRNi: R package for ensemble gene regulatory network inference, combining multiple inference methods to improve reconstruction accuracy of gene regulatory networks.

Sarkar, C., **Sarkar, D.**, Parsad, R., & Mishra, D. *CRAN*, v0.1.6

cran.r-project.org/package=EGRNi

Teaching & Mentoring

2022-2024: M.Sc. Bioinformatics (CBCS) - University of North Bengal. Independent theory and laboratory instruction in three papers (9 credits), as assigned by the course in-charge:

- *Basic Computer Applications* (Core, theory and practical, 4 cr; Sem. I, 2022-23 & 2023-24)
- *Computer Programming for Bioinformatics* (Core, theory and practical, 4 cr; Sem. II, 2022-23)
- *Basic Mathematics and Statistics for Biology* (DSE practical, 1 cr; Sem. I, 2022-23 & 2023-24)

2025: Dissertation mentoring - mentoring of an M.Sc. Bioinformatics dissertation student (Central University of South Bihar) on *Deep Learning-Based Model for Prediction of Targeted Sequencing Depth Using Probe Sequences*; submitted May 2025.

Talks & Presentations

2025: A Webtool for Transfer Learning Based Protein-Protein Interaction Network Prediction Using Hybrid Attention - Oral presentation, *Anusandhan 2025* (SARANSH, Wiley-sponsored), Bose Institute, Kolkata, India, 7 March 2025.

Awards

2021: CSIR-UGC NET with JRF, Life Sciences - June 2021

All India Rank 216 | Junior Research Fellowship, CSIR-HRDG, Government of India

2022: Qualified GATE (Graduate Aptitude Test in Engineering) - Life Sciences (XL)

Grants & Computing Resources

2025: IndiaAI Compute Initiative - Awarded GPU cloud compute (NVIDIA L4 GPU) under the IndiaAI Mission for the project *Deep learning-based framework for protein-protein interaction network prediction* (Project ID: P1-S2025070964, Sep 2025 - Aug 2026).

2025: Google TRC (TPU Research Cloud) Award - Awarded access to GCP - TPUs (v4, v5e, v6e) for machine learning research.

Skills

Programming: Python (PyTorch), R, C

Deep Learning: Transformers, attention mechanisms, transfer learning, pre-training and finetuning.

Generative AI: Autoregressive generative models for de novo protein and protein binder design.

Scalable Training: GPU-based training, parallel training on HPC environments

Bioinformatics: Biological sequence analysis, protein-protein interaction and biomolecular network prediction.

Deployment: Nginx, Docker, Flask, FastAPI, Linux server deployment, model-backed web applications.