Course Code	Course Name	Credits
26BY613	ADVANCED BIOINFORMATICS AND ARTIFICIAL INTELLIGENCE IN BIOLOGICAL RESEARCH	04

Course Objectives

- Explain advanced concepts and scope of bioinformatics in biological research.
- Retrieve, analyze, and interpret biological data from major sequence and structural databases.
- Analyze biological information generated from genomics, proteomics, and transcriptomics studies.
- Evaluate the role of artificial intelligence and machine learning in genomic analysis, protein structure prediction, and plant research.

Learning Outcomes

Upon successful completion of this course it is intended that a student will be able to:

- Identify and retrieve biological data from nucleotide, protein, and structural databases.
- Apply sequence alignment methods and substitution matrices for comparative sequence analysis.
- Interpret large-scale biological data from genomics and proteomics studies.
- Explain the application of artificial intelligence and machine learning in bioinformatics and biological research.

Unit 1 - Introduction to Bioinformatics (12 Hrs.)

Scope and significance of bioinformatics in biological sciences. Biological data types and biological data formats (FASTA, ASN.1, PDB, mmCIF). Biological data mining and biological data analysis. Introduction to bioinformatics software and tools.

Unit 2 – Biological Databases (12 Hrs.)

Biological data organization and database search tools. NCBI resources and database retrieval systems. Nucleic acid sequence databases – GenBank, EMBL and DDBJ. Gene expression databases and genome databases.

Unit 3 - Protein and Structural Databases (12 Hrs.)

Protein sequence databases – SWISS-PROT and UniProt. Structural databases – Protein Data Bank (PDB). Functional and pathway databases – KEGG, SCOP and CATH. Bibliographic databases such as PubMed and OMIM.

Unit 4 – Sequence Analysis and Phylogenetics (12 Hrs.)

Sequence comparison and pairwise sequence alignment. Global and local alignment algorithms. Sequence search tools such as BLAST and FASTA. Amino acid substitution matrices – PAM and BLOSUM. Multiple sequence alignment using ClustalW. Phylogenetic analysis using MEGA and PHYLIP.

Unit 5 – Advanced Bioinformatics and Artificial Intelligence (12 Hrs.)

Omics approaches – genomics, transcriptomics, and proteomics. Gene prediction and genome annotation. Protein structure prediction and molecular modelling including recent AI-based approaches. Integration and analysis of large-scale biological data (big biological data). Introduction to artificial intelligence and machine learning in bioinformatics. Applications of AI in genomic data analysis, plant breeding, crop improvement, and drug discovery.

Reference Books:

1. Lesk, A. M. (2019). *Introduction to Bioinformatics*, 5th Edition. Oxford University Press.
2. Mount, D. W. (2017). *Bioinformatics: Sequence and Genome Analysis*, 2nd Edition. Cold Spring Harbor Laboratory Press.
3. Baxevanis, A. D., & Ouellette, B. F. F. (2020). *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*. Wiley.
4. Pevsner, J. (2015). *Bioinformatics and Functional Genomics*, 3rd Edition. Wiley-Blackwell.
5. Xiong, J. (2021). *Essential Bioinformatics*. Cambridge University Press.
6. Zou, J. et al. (2019). *A Primer on Deep Learning in Genomics*. Nature Genetics.