Course title: Advanced Bioinformatics and Computational Biology			
Course code: BBP 176	**No. of credits:** 2	**L-T-P:** 22-8-0	**Learning hours:** 30
Pre-requisite course code and title (if any): Science graduate			
Department: Department of Biotechnology			
Course coordinator:		**Course instructor:**	
Contact details:			
Course type: Core		**Course offered in:** Semester 1	

Course description:

This course is designed to introduce students of various academic backgrounds to the interdisciplinary knowledge that underlies Bioinformatics and Computational Biology. The students will get acquainted with fundamentals of computers, operating systems and basic concepts in computing and networking. The students will be trained to use computational tools and approaches to extract information from different types of bioinformatics data (gene, protein, disease, etc.) and to analyse them in their area of research work.

Course objectives:

1. To educate the interdisciplinary nature of advances in bioinformatics and computational biology.
2. To provide basic understanding of how biological data is stored and retrieved from various biological databases.
3. To develop an understanding of algorithms of sequence alignment (pair-wise and multiple) and scoring algorithms.

Course contents

Module	Topic	L	T	P
1	**Introduction to Information Technology**			
1.1	**Concepts in information processing and peripherals** Information technology, an overview of current IT applications, Difference between data and information, Information system, Programming languages, Input and output devices, port introduction	2		
1.2	**Internet technology and World Wide Web** Data compression, entropy of information, networking in computers, operating system, Intranet and internet, file transfer protocols, world wide web, internet requirements, Internet-a global network, host& terminal, TCP/IP, common protocols used in internet, web browsers, internet addresses, domain names, basic concepts of HTML, web search engines, electronic mail.	3	2	
1.3	**Information search and data retrieval** Introduction, tools for web search, data retrieval tools, data mining of biological databases, biological databases, difference between primary, secondary and tertiary databases, types of databases. Probabilistic information retrieval, language models for information retrieval, managing	6	2	

	bioinformatics tools, command line sequence extraction and analysis.			
2	**Sequence Analysis and alignment**			
2.1	**Sequence analysis** Sequence analysis, various file formats for bio-molecular sequences: Genbank, FASTA, GCG, MSF, NBRF-PIR etc. Basic concepts of sequence similarity, identity and homology, definitions of homologues, orthologues, paralogues. Scoring matrices: basic concept of a scoring matrix, PAM and BLOSUM series. Applications of substitution matrices in protein sequence alignment and evolution. Sequence-based Database Searches: what are sequence-based database searches, BLAST and FASTA algorithms, various versions of basic BLAST and FASTA, filtering and gapped BLAST.	6	2	
2.2	**Pairwise and Multiple sequence alignments** Basic concepts of sequence alignment, Needleman & Wunsch, Smith & Waterman algorithms for pairwise alignments, Progressive and hierarchical algorithms for MSA. Use of pairwise alignments and Multiple sequence alignment for analysis of Nucleic acid and protein sequences and interpretation of results	5	2	
	Total	**22**	**8**	**0**

Evaluation criteria:

1. Minor test- (Modules 1) 40%
2. Major test (end semester)- (Modules 1 and 2) 60%

Learning outcomes:

1. An understanding of information processing and data retrieval (Minor test 1)
2. Familiarity with the sequence analysis of nucleic acids and proteins using bioinformatics software (Minor test 2)
3. An understanding of sequence alignment and its use in biological research (Minor test 1, Minor test 2 and Major test)

Pedagogical Approach:

1. Classroom lectures and tutorials
2. Case studies and discussion with examples from original research articles

Skill Set:

1. Data analysis: bioinformatics methods for analysis of sequence data

Employabilty:

1. Bioinformatics software development companies.
2. Academic Institutions and Industries involved in R&D.

Materials:

Suggested Readings

1. David W Mount, Bioinformatics: Sequence And Genome Analysis, 2nd Edition, Cold Spring Harbor Press
2. Durbin et al (2007) Biological Sequence Analysis: Probabilistic models of protein and Nucleic acids .Cambridge University Press.
3. Korf Ian, Yandell Mark, Bedell Joseph. BLAST: an essential guide to the basic local alignment search tool. Shroff Publishers and Distributors Pvt. Ltd., 2003. ISBN: 8173665125.
4. Teresa Attwood, Parry-Smith David J. Introduction to Bioinformatics. Publisher: Pearson Education (Singapore) Pte.Ltd. 2001. ISBN:8178085070.

Case studies

Websites:

http://www.ncbi.nlm.nih.gov/

www.embl.org

www.ddbj.nig.ac.jp

www.uniprot.org

www.rcsb.org/

Additional information (if any):

Student responsibilities:

1. Class attendance.
2. Study of reading materials as specified by course instructor
3. Self-study

Course reviewers:

1. Dr. Dinesh Gupta, ICGEB, New Delhi
2. Prof. B.N. Mishra, Institute of Engineering & Technology, Lucknow