BTPC2006 BIOINFORMATICS AND BIO-STATISTICS (3-0-0)

Module I
Searching Biological Data From databases: Finding the information stored and its retrieval methods at NCBI, EMBL and DDBJ, Protein Data Bank, CSD, Uniprot, PIR, SwissProt, TrEMBL, SNP, Metabolic pathways databases KEGG, MetaCyc, Domain classification databases: SCOP, CATH, Pfam Retrieving microbial and viral genome information.

Module II
Introduction to Operating systems, Linux commands, File transfer protocols ftp and telnet, Introduction to Bioinformatics and Computational Biology, Biological sequences, Biological databases, Genome specific databases, Data file formats, Data life cycle, Database management system models, Basics of Structured Query Language (SQL).

Module III
Sequence Analysis, Pair wise alignment, Dynamic programming algorithms for computing edit distance, string similarity, shotgun DNA sequencing, end space free alignment. Multiple sequence alignment, Algorithms for Multiple sequence alignment, Generating motifs and profiles, Local and Global alignment, Needleman and Wunsch algorithm, Smith Waterman algorithm, BLAST, PSIBLAST and PHIBLAST algorithms.

Module IV
Statistics: Definition, functions and limitations. Treatment of data: frequency distribution, Graph of Frequency Distribution Descriptive Measures: Averages and Dispersions (Grouped and ungrouped). Probability: Concepts, definition and elementary problems based on definition.

Module V
Inference: Definition-parameter, Statistic sampling distributors, standard error, Test of Hypothesis, type I and Type II errors. Large sample tests: Z tests, small sample tests: t and F tests. Chi-square test: Goodness of fit and Test of independence.

Course Outcomes (COs):
1. Demonstrate advanced knowledge of bioinformatics concepts, algorithms, tools, and methods for analyzing biological data.
2. Apply statistical methods to analyze and interpret biological data effectively, including measures of central tendency and dispersion.
3. Utilize bioinformatics tools for sequence alignment, database searches, and genome analysis.
4. Understand and evaluate various biological databases and their applications in research.
5. Conduct hypothesis testing using statistical methods, including t-tests and chi-square tests, to analyze experimental data.

Program Outcomes (POs):
1. Acquire a comprehensive understanding of bioinformatics and biostatistics principles applicable to real-world biological problems.

2. Prepare for careers in research or industry by developing critical thinking and problem-solving skills in data analysis.
3. Integrate knowledge from interdisciplinary fields such as biology, computer science, and statistics to enhance research capabilities.
4. Foster a commitment to lifelong learning and professional development in the rapidly evolving fields of bioinformatics and biostatistics.
5. Exhibit leadership qualities and teamwork skills necessary for collaborative scientific research and projects.

Books:
1. Mount DW, Bioinformatics: Sequence and Genome Analysis, Spring Harbor Press
2. Arthur Lesk, Introduction to Bioinformatics, Oxford University Press.
3. Baxevanis AS and Ouellette BF, Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins, Wiley International Science.
4. Biostatistics: Rao KS, Himalayan Publishng House