CA5T3H   **BIO-INFORMATICS**   **Credits : 4**

Lecture Hours : 4 periods / week   Internal assessment : 30   Marks

Semester and Examination: 70   Marks

Course Description:

Students are expected to know the fundamentals of Engineering in Medicine and biology, which is emerging as an interesting field. Students are expected to use The Internet extensively to understand the subject.

Course Objective:

• Learn to create Biological database of available sequences
• Learn how to search for similar sequences
• Learn the concept of available computerized biological databases & Software tools
• Learn how to visualize the architecture of a biological input data
• Learn the need for sequence analysis
• Understanding of tools for Identification, Prediction, Visualization and Pattern Matching from biological sequences and structure pf proteins
• Can classify the structure of protein with the help of available software tools
• Learn how to develop the tools for pattern matching from available sequences

UNIT I:

Introduction: Definitions, Sequencing, Biological sequence/structure, Genome Projects, Pattern recognition an prediction, Folding problem, Sequence Analysis, Homology and Analogy.

UNIT II:

Protein Information Resources: Biological databases, Primary sequence databases, Protein Sequence databases, Secondary databases, Protein pattern databases, and Structure classification databases.

UNIT III:

Genome Information Resources: DNA sequence databases, specialized genomic resources.

UNIT IV:

DNA Sequence analysis: Importance of DNA analysis, Gene structure and DNA sequences, Features of DNA sequence analysis, EST (Expressed

Sequence Tag) searches, Gene hunting, Profile of a cell, EST analysis, Effects of EST data on DNA databases.

UNIT V:

Pair wise alignment techniques: Database searching, Alphabets and complexity, Algorithm and programs, Comparing two sequences, sub-sequences, Identity and similarity, The Dotplot, Local and global similarity, different alignment techniques, Dynamic Programming, Pair wise database searching.

UNIT VI:

Multiple sequence alignment: Definition and Goal, The consensus, computational complexity, Manual methods, Simultaneous methods, Progressive methods, Databases of Multiple alignments and searching

UNIT VII:

Databases: Secondary database searching Importance and need of secondary database searches, secondary database structure and building a sequence search protocol

UNIT VIII:

Analysis packages: Analysis package structure, commercial databases, commercial software, comprehensive packages, packages specializing in DNA analysis, Intranet Packages, Internet Packages.

Learning Resources :

Text Books:

1. Introduction to Bioinformatics,T K Attwood & D J Parry-Smith, Essex: Addison WesleyLongman, 1999.

Reference Books:

1. Bioinformatics For Dummies Jean-Michel Claveriw, Cerdric Notredame, WILEY DreamTech India Pvt. Ltd, 2/e 2006.
2. Sequence Analysis in A Nutshell, Scott Markel &Darryl Leon, O'REILLY, 1/e, 2003.