## COURSE CONTENT

| BIOINFORMATICS | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **II Semester: CSE** | | | | | | | | |
| **Course Code** | **Category** | **Hours / Week** | | | **Credits** | **Maximum Marks** | | |
| | | **L** | **T** | **P** | **C** | **CIA** | **SEE** | **Total** |
| **2425824** | **Foundation** | 3 | 0 | 0 | 3 | 40 | 60 | 100 |
| **Contact Classes: 45** | **Tutorial Classes: Nil** | **Practical Classes: Nil** | | | | **Total Classes: 45** | | |
| **Prerequisites:** Basic understanding of molecular and cellular biology. | | | | | | | | |

**Course Overview:**

This course is an interdisciplinary field that combines biology, computer science, mathematics, and statistics to store, analyze, interpret, and visualize biological data such as DNA, RNA, and proteins. It is essential in genomics, proteomics, personalized medicine, and many areas of life sciences research.

**Course Objectives:**

1. To understand the fundamental concepts of bioinformatics and the biological basis of sequence analysis.
2. To study the Central Dogma, Bio-XML, and XML technologies used in biological data representation and exchange.
3. To learn Perl and Bioperl programming techniques for processing and analyzing biological sequence data.
4. To understand database technologies, architectures, and management systems used in bioinformatics applications.
5. To analyze sequence alignment algorithms, phylogenetic analysis methods, and protein structure prediction techniques.

**Course Outcomes: After Completion of the Course, Students should be able to**

1. Analyze how the central dogma of molecular biology relates to data organization and knowledge representation in bioinformatics.
2. Write Perl programs to process DNA and protein sequences, including transcription and reverse complement calculation.
3. Implement biological data storage using flat file, relational, and object-oriented databases
4. Evaluate the accuracy and performance of sequence analysis tools and alignment techniques.
5. Evaluate clustering methods and tools for protein visualization and structure prediction in evolutionary studies.

**UNIT - I: The Central Dogma & XML (Bio XML) for Bioinformatics:** Watson's definition, information flow, from data to knowledge, Convergence, the organization of DNA, the organization of Proteins, Introduction, Differences between HTML and XML, fundamentals of XML, fundamentals of XML namespaces. Introduction to DTDs, Document type Declarations, declaring elements, declaring attributes, working with entities XML Schemas, Essential Concepts, working with simple types, working with complex types, Basic namespaces issues.

**UNIT - II: Perl (Bioperl) for Bioinformatics:** Representing sequence data, program to store a DNA sequence, concatenating DNA fragments, Transcription, Calculating the reverse complement in Perl, Proteins, files, reading proteins in files, Arrays, Flow control, finding motifs, counting Nucleotides, exploding strings into arrays, operating on strings, writing to files, subroutines and bugs.

**UNIT - III: Databases:** Flat file, Relational, object-oriented databases, object Relational and Hypertext, Data life cycle, Database Technology, Database Architecture, Database Management Systems and Interfaces.

**UNIT - IV**: **Sequence Alignment Algorithms:** Biological motivations of sequence analysis, the models for sequence analysis and their biological motivation, global alignment, local alignment, End free-space alignment and gap penalty, Sequence Analysis tools and techniques.

**UNIT - V: Phylogenetic Analysis:** Introduction, methods of Phylogenetic analysis, distance methods, the neighbor- Joining (NJ) method, The Fitch/ Margoliash method, character-based methods, other methods, Tree evaluation and problems in phylogenetic analysis, Clustering, Protein structure visualization and Protein structure prediction.

## TEXT BOOKS:


1. S.C. Rastogi, N. Mendiratta, "Bioinformatics Methods and Applications", CBS publications, 2004
2. James D. Tisdall, "Beginning Perl for Bioinformatics" O'Reilly media, 1st Edition, 2001.


## REFERENCE BOOKS:


1. D.R. Westhead, J.H. Parish, "Bioinformatics" Viva books private limited, New Delhi (2003)
2. Att Wood, "Bioinformatics" Pearson Education, 2004
3. Bryan Bergeron, M.D, "Bioinformatics Computing" Pearson Education, 2003.


## ELECTRONIC RESOURCES:

1. https://www.coursera.org/specializations/bioinformatics
2. https://www.coursera.org/learn/bioinformatics
3. https://www.coursera.org/learn/bits-introduction-to-bioinformatics
4. https://www.edx.org/learn/bioinformatics

## MATERIALS ONLINE:

1. Course template
2. Tutorial question bank
3. Tech talk and Concept Video topics
4. Open-ended experiments
5. Definitions and terminology
6. Assignments
7. Model question paper – I
8. Model question paper – II
9. Lecture notes

10. E-Learning Readiness Videos (ELRV)