



# MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT

(AN AUTONOMOUS INSTITUTION)

(Approved by AICTE, New Delhi & Affiliated to JNTUH, Hyderabad)

Accredited by NAAC with 'A' Grade & Recognized Under Section 2(f) & 12(B) of the UGC act, 1956

## COURSE CONTENT

BIOINFORMATICS								
II Semester: CSE								
Course Code	Category	Hours / Week			Credits	Maximum Marks		
		L	T	P		C	CIA	SEE
2525816	Foundation	3	0	0	3	40	60	100
		Practical Classes: Nil			Total Classes: 45			
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 biological sequence analysis.
2. To develop skills in representing and processing biological data using XML and BioXML techniques.
3. To apply Perl/BioPerl programming for DNA, RNA, and protein sequence analysis.
4. To study database models and management systems used for storing and retrieving biological information.
5. To analyze sequence alignment, phylogenetic methods, and protein structure prediction techniques for biological research.

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

1. Apply XML concepts to represent and manage biological data for knowledge extraction in genomics.
2. Analyze DNA and protein sequences using Perl or BioPerl scripts to automate motif discovery and sequence manipulation in bioinformatics research.
3. Evaluate different database architectures for storing and retrieving biological information.
4. Distinguish sequence alignment algorithms and gap-penalty models to determine accurate techniques for genomic sequence comparisons.
5. Create phylogenetic trees and protein structure visualizations by integrating clustering and prediction methods to study evolutionary relationships and structural biology.

**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)