Reg. No.															

Question Paper Code	13846

M.Tech. - DEGREE EXAMINATIONS, NOV / DEC 2025

Seventh Semester

M.Tech. - Computer Science and Engineering (5 Years Integrated)

20CSEL604 - BIOINFORMATICS

Regulations - 2020

Duration: 3 Hours Max. Marks: 100

PART - A (MCQ) (10 × 1 = 10 Marks)

Answer ALL Questions

		Marks	*K – Level*	*CO*
1.	Which software is widely used for visualizing protein structures? (a) Excel (b) PyMOL (c) Photoshop (d) Word	*1*	*K1*	*CO1*
2.	The PDB file format is primarily used for storing: (a) DNA sequences (b) Protein 3D structures (c) Phylogenetic trees (d) RNA transcripts	*1*	*K1*	*CO1*
3.	Which international organization provides free access to structural and functional protein data? (a) WHO (b) EMBL-EBI (c) NASA (d) UNESCO	*1*	*K1*	*CO2*
4.	GenBank is a (a) Protein classification tool (b) Database of genetic sequences (c) Tool for graph plotting (d) Lab-based DNA scanner	*1*	*K1*	*CO2*
5.	An example of uncurated biological data is: (a) UniProtKB/Swiss-Prot (b) GenBank submissions (c) RefSeq curated database (d) PDB with annotations	*1*	*K1*	*CO3*
6.	The primary goal of data mining in bioinformatics is to: (a) Predict protein tertiary structure only (b) Extract hidden patterns and relationships in biological data (c) Store sequences in FASTA format (d) Develop sequencing machines	*1*	*K1*	*CO3*
7.	In 2D DNA walk models, nucleotides are usually mapped to: (a) Four distinct directions in a plane (b) A single linear path (c) A random number (d) Protein motifs	*1*	*K1*	*CO4*
8.	Visualization techniques in bioinformatics mainly help in: (a) Storing protein sequences in databases (b) Interpreting complex biological data patterns (c) Speeding up PCR reactions (d) Reducing DNA replication errors	*1*	*K1*	*CO4*
9.	In bioinformatics, a motif usually refers to: (a) A repeated DNA or protein sequence pattern (b) A full-length gene sequence (c) A 3D protein structure (d) A ribosome binding site	*1*	*K1*	*CO5*
10.	Normalization in microarray analysis is done to: (a) Remove noise and make datasets comparable (b) Detect new motifs (c) Edit the genome (d) Amplify RNA	*1*	*K1*	*CO6*

PART - B (12 × 2 = 24 Marks)

Answer ALL Questions

11.	What is meant by biological data annotation?	*2*	*K1*	*CO1*
12.	Differentiate between structural and functional bioinformatics.	*2*	*K2*	*CO1*
13.	List two roles of internet-based bioinformatics tools in modern research.	*2*	*K1*	*CO2*
14.	Recall the benefits of sharing biological data online for genome projects and other research initiatives.	*2*	*K1*	*CO2*

15.	Write short notes on data curation vs. data cleaning in bioinformatics.	*2*	*K2*	*CO3*
16.	List two examples of clinical/biomedical data sources.	*2*	*K1*	*CO3*
17.	Why is motif discovery important in transcription factor binding studies?	*2*	*K1*	*CO4*
18.	How do machine learning methods assist in DNA data analysis?	*2*	*K1*	*CO4*
19.	Compare 1D and 2D DNA walk models in sequence representation.	*2*	*K2*	*CO5*
20.	Classify gene regulation in controlling cellular functions.	*2*	*K2*	*CO5*
21.	Why is spot extraction important in microarray data analysis?	*2*	*K2*	*CO6*
22.	Identify the significance of a cost matrix in scientific data management systems.	*2*	*K2*	*CO6*

PART - C (6 × 11 = 66 Marks)

Answer ALL Questions

23.	a)	(i) Discuss the process and significance of biological data integration.	*5*	*K2*	*CO1*
		(ii) Elaborate the formatting and data annotation analysis quality in bioinformatics.	*6*	*K2*	*CO1*
		OR			
	b)	Explain the evolution and existing research areas in bioinformatics.	*11*	*K2*	*CO1*
24.	a)	Describe the role of structural bioinformatics in managing and integrating heterogeneous biological datasets.	*11*	*K2*	*CO2*
		OR			
	b)	Explain the function of online search engines in retrieving biological information from bioinformatics databases.	*11*	*K2*	*CO2*
25.	a)	Illustrate the process of data curation and data cleaning in bioinformatics databases. Why are they crucial for accurate biological interpretations?	*11*	*K2*	*CO3*
		OR			
	b)	Demonstrate biomedical data analysis methods. Explain how clinical data, genomic data, and proteomic data are integrated for translational research.	*11*	*K2*	*CO3*
26.	a)	Describe multiple sequence alignment (MSA) techniques. Explain how MSA helps in motif detection, comparative genomics, and protein functional analysis.	*11*	*K2*	*CO4*
		OR			
	b)	Explain machine learning and neural network architectures to demonstrate their use in DNA and protein data analysis, and illustrate with suitable bioinformatics applications.	*11*	*K2*	*CO4*
27.	a)	Elaborate the significance of protein and amino acid motifs in predicting protein structure and function.	*11*	*K2*	*CO5*
		OR			
	b)	Discuss how visualization and fractal analysis together contribute to the interpretation of complex genomic data.	*11*	*K2*	*CO5*
28.	a)	Build cost matrices and tradeoffs in scientific data management, highlighting their role in evaluation models.	*11*	*K3*	*CO6*
		OR			
	b)	Develop the application of gene network analysis in understanding regulatory pathways with a suitable bioinformatics example.	*11*	*K3*	*CO6*