Roll No.						



# PRESIDENCY UNIVERSITY

## **BENGALURU**

### **End - Term Examinations - MAY 2025**

School: SOCSE	Program: CSE				
Course Code: CSE3069	Course Name: Introduction to bioinformatics				
Semester: IV	Max Marks: 100	Weightage: 50%			

CO - Levels	CO1	CO2	CO3	CO4	CO5
Marks	26	26	24	24	-

#### **Instructions:**

- (i) Read all questions carefully and answer accordingly.
- (ii) Do not write anything on the question paper other than roll number.

#### Part A

## Answer ALL the Questions. Each question carries 2marks.

10Q x 2M=20M

1.	Each of the 20 amino acids has a common core structure, outline the representation.	2 Marks	L1	CO1
2.	Define the importance of bioinformatics.	2 Marks	L1	CO1
3.	Translate reverse complement of: TACGGCTAAATTGCCATGCTAT.	2 Marks	L2	CO1
4.	Expand DDBJ and BLAST	2 Marks	L1	CO2
5.	Explain secondary database with an example	2 Marks	L2	CO2
6.	Briefly explain SAM file format	2 Marks	L2	CO2
7.	Briefly explain progressive alignment method (PAM).	2 Marks	L2	CO3
8.	Define match, mismatch and gaps in sequence alignment.	2 Marks	L1	CO3
9.	Differentiate sequence motif and structure motif.	2 Marks	L2	CO4
10.	Define phylogenetic.	2 Marks	L1	CO4

## Part B

Description			Answer the Questions.	Total Marks 80M			
b. Explain the four major biological macromolecules in detail.   10 Marks   L2   CO	11.	a.	With a neat diagram, Explain the basic unit of life that is	10 Marks	L2	CO1	
12. a. Differentiate between DNA and RNA with their structure.  10 Marks			responsible for all life's processes.				
12.   a.   Differentiate between DNA and RNA with their structure.   10 Marks   L2   CO		b.	Explain the four major biological macromolecules in detail.	10 Marks	L2	CO1	
b. Explain various types of omics in detail.  10 Marks		1	Or				
13. a. Explain two substitution matrix with an example.  b. Demonstrate the reconstruction of substring for the following sequence: 'AGTC', 'GTCA', 'TCAG', 'CAGT', 'AGTT', 'GTTA', 'TAGC', TAGC', 'AGCT', 'GCAT', 'CTAT', TATG', 'ATGC', 'TGCA', 'GCAT'.  14. a. With suitable examples, explain local and global alignment of sequences.  b. Illustrate approximate pattern matching for DNA sequence SSS: "TGCATCGATCGTACGTAGCTAGCATGTAGCATGTAGCATGTAGCATGTAGCATGTAGCATGTAGCATGTAGCATGTAGCATGTAGCATGTAGCATGTAGCATGTAGCATGTAGCATGTAGCATGTAGCATGTAGCATGTAGGTAG	12.	a.	Differentiate between DNA and RNA with their structure.	10 Marks	L2	CO1	
b. Demonstrate the reconstruction of substring for the following sequence: 'AGTC', 'GTCA', 'TCAG', 'CAGT', 'AGTT', 'GTTA', 'TTAG', 'TAGC', 'AGCT', 'GCAT'.  Or  14. a. With suitable examples, explain local and global alignment of sequences.  b. Illustrate approximate pattern matching for DNA sequence SSS: "TGCATCGATCGTACGTAGCTAGCTAGCATGTAGCTAGCTA		b.	Explain various types of omics in detail.	10 Marks	L2	CO1	
sequence: 'AGTC', 'GTCA', 'TCAG', 'CAGTT, 'AGTTA', 'TTAG', 'TAGC', 'TAGC', 'GCAT'.  14. a. With suitable examples, explain local and global alignment of sequences.  b. Illustrate approximate pattern matching for DNA sequence SSS: "TGCATCGATCGTACGTAGCTAGCTAGCTAGCTAGCTAGCT	13.	a.	Explain two substitution matrix with an example.	10 Marks	L2	<b>CO2</b>	
TTAG', 'TAGC', 'AGCT', 'GCTA', 'CTAT', 'TATG', 'ATGC', 'TGCA', 'GCAT'.  14. a. With suitable examples, explain local and global alignment of sequences.  b. Illustrate approximate pattern matching for DNA sequence SSS: "TGCATCGATCGTACGTAGCTAGCAT" Pattern P: "CGTAC" with maximum mismatches d: 3  15. a. Calculate the alignment score of following alignment: S1: AGCTTTCGAACCTTACGATCGATCGATCGATCGAGCTAGCT		b.	Demonstrate the reconstruction of substring for the following	10 Marks	L3	CO2	
GCAT'.   Or   I			sequence: 'AGTC', 'GTCA', 'TCAG', 'CAGT', 'AGTT', 'GTTA',				
14. a. With suitable examples, explain local and global alignment of sequences.  b. Illustrate approximate pattern matching for DNA sequence SSS: "TGCATCGATCGTAGCTAGCTAGCAT" Pattern P: "CGTAC" with maximum mismatches d: 3  15. a. Calculate the alignment score of following alignment: S1: AGCTTTCGAACCTTAGGATCGACTGGAGCTAGCTTAGGC S2: AGTACGATTAGGACGAAGGGATTACGATCTAGGT Scoring scheme: Match = +3, Mismatch = -2, Gap = -3  b. Apply the Needleman-Wunsch Algorithm to the following sequences and present the global alignment: Seq1 = GCATGCG Seq2 = GATTACA Match: +1, Mismatch: -1, Gap: -1  16. a. Illustrate 3 components of sequence alignment with suitable example.  b. Apply the Smith-Waterman Algorithm to the following sequences and present the local alignment: Seq1 = ACCGTGA Seq2 = GTGAATA Match: +1, Mismatch: -1, Gap: -1  17. a. Solve by identifying the GATTACA motif in the following sequences using GIBBS algorithm. ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 CO-			'TTAG', 'TAGC', 'AGCT', 'GCTA', 'CTAT', 'TATG', 'ATGC', 'TGCA',				
14. a. With suitable examples, explain local and global alignment of sequences.  b. Illustrate approximate pattern matching for DNA sequence SSS: "TGCATCGATCGTACGTAGCTAGCATC" Pattern P: "CGTAC" with maximum mismatches d: 3  15. a. Calculate the alignment score of following alignment: S1: AGCTTCGAACCTTAGGATCGACTGGACCTGAGCTAGCTAG			'GCAT'.				
Sequences.   Sequences.   Sequence   Seque			Or				
b. Illustrate approximate pattern matching for DNA sequence SSS: "TGCATCGATCGTAGCTAGCTAGCTAGCAT" Pattern P: "CGTAC" with maximum mismatches d: 3  15. a. Calculate the alignment score of following alignment: S1: AGCTTTCGAACCTTACGATCGACTGGAGCTAGCTTAGC S2: AGTACGATAGGACGAAGGGATTACGATCTAGGT Scoring scheme: Match = +3, Mismatch = -2, Gap = -3  b. Apply the Needleman-Wunsch Algorithm to the following sequences and present the global alignment: Seq1 = GCATGCG Seq2 = GATTACA Match: +1, Mismatch: -1, Gap: -1  16. a. Illustrate 3 components of sequence alignment with suitable example.  b. Apply the Smith-Waterman Algorithm to the following sequences and present the local alignment: Seq1 = ACCGTGA Seq2 = GTGAATA Match: +1, Mismatch: -1, Gap: -1  17. a. Solve by identifying the GATTACA motif in the following sequences using GIBBS algorithm. ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 COMMARKS	14.	a.	With suitable examples, explain local and global alignment of	10 Marks	L2	<b>CO2</b>	
### TGCATCGATCGTAGCTAGCAT" Pattern P: " CGTAC" with maximum mismatches d: 3  15. a. Calculate the alignment score of following alignment:     \$1: AGCTTTCGAACCTTACGATCGACTGGAGCTAGCTTAGC     \$2: AGTACGATAGGACGAAGGGATTACGATCTAGGT     \$5: Scoring scheme : Match = +3, Mismatch = -2, Gap = -3      b. Apply the Needleman-Wunsch Algorithm to the following sequences and present the global alignment:     \$5: Seq1 = GCATGCG     \$6: Seq2 = GATTACA     Match: +1, Mismatch: -1, Gap: -1  16. a. Illustrate 3 components of sequence alignment with suitable example.  10 Marks			sequences.				
15. a. Calculate the alignment score of following alignment: S1: AGCTTTCGAACCTTACGATCGACTGGAGCTAGCTTAGC S2: AGTACGATAGGACGAAGGGATTACGATCTAGGT Scoring scheme : Match = +3, Mismatch = -2, Gap = -3  b. Apply the Needleman-Wunsch Algorithm to the following sequences and present the global alignment: Seq1 = GCATGCG Seq2 = GATTACA Match: +1, Mismatch: -1, Gap: -1  Tor  16. a. Illustrate 3 components of sequence alignment with suitable example.  b. Apply the Smith-Waterman Algorithm to the following sequences and present the local alignment: Seq1 = ACCGTGA Seq2 = GTGAATA Match: +1, Mismatch: -1, Gap: -1  17. a. Solve by identifying the GATTACA motif in the following sequences using GIBBS algorithm. ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 CO-		b.	Illustrate approximate pattern matching for DNA sequence SSS:	10 Marks	L3	<b>CO2</b>	
15. a. Calculate the alignment score of following alignment: S1: AGCTTTCGAACCTTACGATCGACTGAGCTTAGC S2: AGTACGATAGGACGAAGGGATTACGATCTAGGT Scoring scheme: Match = +3, Mismatch = -2, Gap = -3  b. Apply the Needleman-Wunsch Algorithm to the following sequences and present the global alignment: Seq1 = GCATGCG Seq2 = GATTACA Match: +1, Mismatch: -1, Gap: -1  Tor  16. a. Illustrate 3 components of sequence alignment with suitable example. b. Apply the Smith-Waterman Algorithm to the following sequences and present the local alignment: Seq1 = ACCGTGA Seq2 = GTGAATA Match: +1, Mismatch: -1, Gap: -1  17. a. Solve by identifying the GATTACA motif in the following sequences using GIBBS algorithm. ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 CO-			" TGCATCGATCGTACGTAGCTAGCAT" Pattern P: " CGTAC" with				
S1: AGCTTTCGAACCTTACGATCGACTGGAGCTAGCTTAGC S2: AGTACGATAGGACGAAGGGATTACGATCTAGGT Scoring scheme : Match = +3, Mismatch = -2, Gap = -3  b. Apply the Needleman-Wunsch Algorithm to the following sequences and present the global alignment: Seq1 = GCATGCG Seq2 = GATTACA Match: +1, Mismatch: -1, Gap: -1  Tor  16. a. Illustrate 3 components of sequence alignment with suitable example.  b. Apply the Smith-Waterman Algorithm to the following sequences and present the local alignment: Seq1 = ACCGTGA Seq2 = GTGAATA Match: +1, Mismatch: -1, Gap: -1  17. a. Solve by identifying the GATTACA motif in the following sequences using GIBBS algorithm. ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an  10 Marks L3 CO			maximum mismatches d: 3				
S1: AGCTTTCGAACCTTACGATCGACTGGAGCTAGCTTAGC S2: AGTACGATAGGACGAAGGGATTACGATCTAGGT Scoring scheme : Match = +3, Mismatch = -2, Gap = -3  b. Apply the Needleman-Wunsch Algorithm to the following sequences and present the global alignment: Seq1 = GCATGCG Seq2 = GATTACA Match: +1, Mismatch: -1, Gap: -1  Tor  16. a. Illustrate 3 components of sequence alignment with suitable example.  b. Apply the Smith-Waterman Algorithm to the following sequences and present the local alignment: Seq1 = ACCGTGA Seq2 = GTGAATA Match: +1, Mismatch: -1, Gap: -1  17. a. Solve by identifying the GATTACA motif in the following sequences using GIBBS algorithm. ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an  10 Marks L3 CO	15.	a.	Calculate the alignment score of following alignment:	10 Marks	L3	<b>CO3</b>	
S2: AGTACGATAGGACGAAGGGATTACGATCTAGGT Scoring scheme : Match = +3, Mismatch = -2, Gap = -3  b. Apply the Needleman-Wunsch Algorithm to the following sequences and present the global alignment: Seq1 = GCATGCG Seq2 = GATTACA Match: +1, Mismatch: -1, Gap: -1  Tor  16. a. Illustrate 3 components of sequence alignment with suitable example.  b. Apply the Smith-Waterman Algorithm to the following sequences and present the local alignment: Seq1 = ACCGTGA Seq2 = GTGAATA Match: +1, Mismatch: -1, Gap: -1  17. a. Solve by identifying the GATTACA motif in the following sequences using GIBBS algorithm. ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an  10 Marks L3 CO							
b. Apply the Needleman-Wunsch Algorithm to the following sequences and present the global alignment: Seq1 = GCATGCG Seq2 = GATTACA Match: +1, Mismatch: -1, Gap: -1  16. a. Illustrate 3 components of sequence alignment with suitable example.  b. Apply the Smith-Waterman Algorithm to the following sequences and present the local alignment: Seq1 = ACCGTGA Seq2 = GTGAATA Match: +1, Mismatch: -1, Gap: -1  17. a. Solve by identifying the GATTACA motif in the following sequences using GIBBS algorithm. ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 CO-							
b. Apply the Needleman-Wunsch Algorithm to the following sequences and present the global alignment:  Seq1 = GCATGCG Seq2 = GATTACA Match: +1, Mismatch: -1, Gap: -1  Tor  16. a. Illustrate 3 components of sequence alignment with suitable example.  b. Apply the Smith-Waterman Algorithm to the following sequences and present the local alignment:  Seq1 = ACCGTGA Seq2 = GTGAATA Match: +1, Mismatch: -1, Gap: -1  17. a. Solve by identifying the GATTACA motif in the following sequences using GIBBS algorithm.  ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 CO-							
sequences and present the global alignment: Seq1 = GCATGCG Seq2 = GATTACA Match: +1, Mismatch: -1, Gap: -1  Or  16. a. Illustrate 3 components of sequence alignment with suitable example.  b. Apply the Smith-Waterman Algorithm to the following sequences and present the local alignment: Seq1 = ACCGTGA Seq2 = GTGAATA Match: +1, Mismatch: -1, Gap: -1  17. a. Solve by identifying the GATTACA motif in the following sequences using GIBBS algorithm. ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 CO-		b.		10 Marks	L3	CO3	
Seq1 = GCATGCG Seq2 = GATTACA Match: +1, Mismatch: -1, Gap: -1  16. a. Illustrate 3 components of sequence alignment with suitable example.  b. Apply the Smith-Waterman Algorithm to the following sequences and present the local alignment: Seq1 = ACCGTGA Seq2 = GTGAATA Match: +1, Mismatch: -1, Gap: -1  17. a. Solve by identifying the GATTACA motif in the following sequences using GIBBS algorithm. ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 CO-							
Seq2 = GATTACA Match: +1, Mismatch: -1, Gap: -1  Or  16. a. Illustrate 3 components of sequence alignment with suitable example.  b. Apply the Smith-Waterman Algorithm to the following sequences and present the local alignment: Seq1 = ACCGTGA Seq2 = GTGAATA Match: +1, Mismatch: -1, Gap: -1  17. a. Solve by identifying the GATTACA motif in the following sequences using GIBBS algorithm. ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 CO-							
16. a. Illustrate 3 components of sequence alignment with suitable example.  b. Apply the Smith-Waterman Algorithm to the following sequences and present the local alignment: Seq1 = ACCGTGA Seq2 = GTGAATA Match: +1, Mismatch: -1, Gap: -1  17. a. Solve by identifying the GATTACA motif in the following sequences using GIBBS algorithm. ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 CO-							
16. a. Illustrate 3 components of sequence alignment with suitable example.  b. Apply the Smith-Waterman Algorithm to the following sequences and present the local alignment: Seq1 = ACCGTGA Seq2 = GTGAATA Match: +1, Mismatch: -1, Gap: -1  17. a. Solve by identifying the GATTACA motif in the following sequences using GIBBS algorithm. ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 CO-			Match: +1, Mismatch: -1, Gap: -1				
b. Apply the Smith-Waterman Algorithm to the following sequences and present the local alignment:  Seq1 = ACCGTGA Seq2 = GTGAATA Match: +1, Mismatch: -1, Gap: -1  17. a. Solve by identifying the GATTACA motif in the following sequences using GIBBS algorithm.  ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 CO-		I	Or				
b. Apply the Smith-Waterman Algorithm to the following sequences and present the local alignment:  Seq1 = ACCGTGA Seq2 = GTGAATA Match: +1, Mismatch: -1, Gap: -1  17. a. Solve by identifying the GATTACA motif in the following sequences using GIBBS algorithm.  ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 COA	16.	a.	Illustrate 3 components of sequence alignment with suitable	10 Marks	L3	CO3	
sequences and present the local alignment: Seq1 = ACCGTGA Seq2 = GTGAATA Match: +1, Mismatch: -1, Gap: -1  17. a. Solve by identifying the GATTACA motif in the following sequences using GIBBS algorithm. ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 COA			example.				
Seq1 = ACCGTGA Seq2 = GTGAATA Match: +1, Mismatch: -1, Gap: -1  17. a. Solve by identifying the GATTACA motif in the following sequences using GIBBS algorithm. ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 CO4		b.	Apply the Smith-Waterman Algorithm to the following	10 Marks	L3	<b>CO3</b>	
Seq2 = GTGAATA Match: +1, Mismatch: -1, Gap: -1  17. a. Solve by identifying the GATTACA motif in the following sequences using GIBBS algorithm. ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 COA			sequences and present the local alignment:				
Match: +1, Mismatch: -1, Gap: -1  17. a. Solve by identifying the GATTACA motif in the following sequences using GIBBS algorithm. ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 CO4			Seq1 = ACCGTGA				
17. a. Solve by identifying the GATTACA motif in the following sequences using GIBBS algorithm.  ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 COA			Seq2 = GTGAATA				
sequences using GIBBS algorithm. ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 CO4			Match: +1, Mismatch: -1, Gap: -1				
sequences using GIBBS algorithm. ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 CO4	17	а	Solve by identifying the GATTACA motif in the following	10 Marks	I.3	<b>CO4</b>	
ACCATGACAG GAGTATACCT CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 CO4	1/1	<b>.</b>		IO PIUI NO	20	301	
CATGCTTACT (random selected sequence) CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 CO4							
CGGAATGCAT  b. Interpret the information provided by a sequence logo with an 10 Marks L3 CO4							
b. Interpret the information provided by a sequence logo with an 10 Marks L3 CO4							
example and its relevance in bioinformatics.		b.		10 Marks	L3	<b>CO4</b>	
			example and its relevance in bioinformatics.				

	Or									
18.	a.	Illustrate the significance of motifs in the study of genomic	10 Marks	L3	CO4					
		sequences.								
	b.	Apply the UPGMA algorithm to construct the phylogenetic tree, and interpret the evolutionary relationships between the sequences. Discuss the implications of these relationships in the context of their genetic similarities and differences.  1. GGTCAGTCAA  2. CCTAGGTTGA  3. TGGCATACTG  4. ATGCGTAGCT  5. GATTACAAGT	10 Marks	L3	CO4					