|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Roll No. |  |  |  |  |  |  |  |  |  |  |  |  |



**PRESIDENCY UNIVERSITY**

**Bengaluru**

|  |
| --- |
| **End - Term Examinations – JANUARY 2025** |
| **Date:** 13- 01- 2025 **Time:** 9.30 am – 12.30 pm |

|  |  |  |
| --- | --- | --- |
| **School:** SOCSE | **Program:** B. Tech CSE | |
| **Course Code :** CSE3069 | **Course Name :** Bioinformatics | |
| **Semester**: V | **Max Marks**:100 | **Weightage**: 50% |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **CO - Levels** | **CO1** | **CO2** | **CO3** | **CO4** |
| **Marks** | **26** | **26** | **24** | **24** |

**Instructions:**

1. *Read all questions carefully and answer accordingly.*
2. *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** | Explain the processes of transcription and translation in gene expression. | **2 Marks** | **L1** | **CO1** |
| **2** | List the Omics of Science branch. | **2 Marks** | **L2** | **CO1** |
| **3** | Identify the complement of given sequence.  ATCGTTTTAAACCGGGTTTAAAGCA | **2 Marks** | **L2** | **CO1** |
| **4** | Define any Multiple sequence alignment tool. | **2 Marks** | **L2** | **CO2** |
| **5** | Define Bioinformatics and provide an example of its application. | **2 Marks** | **L1** | **CO2** |
| **6** | Describe Phylogenetic. | **2 Marks** | **L1** | **CO2** |
| **7** | List the primary function of carbohydrates. | **2 Marks** | **L1** | **CO3** |
| **8** | Explain which organization/company maintains the GenBank. | **2 Marks** | **L2** | **CO3** |
| **9** | State the importance of RNA. | **2 Marks** | **L3** | **CO4** |
| **10** | Illustrate k-mers with brief explanation. | **2 Marks** | **L1** | **CO4** |

**Part B**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Answer the Questions Total 80 Marks.** | | | | | |
| **11.** | **a.** | Distinguish the importance of databases in the field of Bioinformatics. [10 Marks] | **20 Marks** | **L2** | **CO1** |
| **b.** | Differentiate between DNA sequencing methods and explain their respective applications. [10 Marks] |
| **Or** | | | | | |
| **12.** | **a.** | What is Biological Database? Explain the GENOMIC databases in detail. [10 Marks] | **20 Marks** | **L2** | **CO1** |
| **b.** | Explain the purpose and usage of CLUSTALW and MSF file formats in multiple sequence alignment . [10 Marks] |
|  |  |  |  |  |  |
| **13.** | **a.** | Outline different types of database retrieval tools and describe their specific roles in storing and retrieving genetic information. Discuss the advantages of using specialized tools over general ones . [10 Marks] | **20 Marks** | **L2** | **CO2** |
| **b.** | What is a gene expression and explain the process of 3d folding of a protein structure . [10 Marks] |
| **Or** | | | | | |
| **14.** | **a.** | Identify the index positions for the Pattern = "ACG" in the text  Text = "ACGTACGCTGACACCGGACTTAAGGACGGTAC"  d = 2 . [10 Marks] | **20 Marks** | **L2** | **CO2** |
| **b.** | Explain substitution matrix example.[10 Marks] |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **15.** | **a.** | Compute frequent k mer in text where k=3 ATCGATCGAATCGGGCTAGCTAGCTAG  CTAATCGATCGGGCTAGC. [10 Marks] | **20 Marks** | **L3** | **CO3** |
| **b.** | Explain what is goodness of the alignment with its importance in real time and find the score of the sequence considering the default values for match and mismatch and gap values  ATCGGGTTCCAAACCGGGATTTACACA  ATCCC\_TCC\_AGGGGGGAA\_TTACACC. . [10 Marks] |
| **Or** | | | | | |
| **16.** | **a.** | Apply the Needleman-Wunsch Algorithm to the following sequences and present the global alignment: . [10 Marks]  Seq1 = TCCAGGTCAGGTCA  Seq2 = CAGAGG  Match: +4 Mismatch: -3 Gap: -1 | **20 Marks** | **L3** | **CO3** |
| **b.** | Examine the role of omics technologies (e.g., genomics, proteomics) in advancing our understanding of biological systems. . [10 Marks] |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **17.** | **a.** | Apply the UPGMA algorithm to create the phylogenetic tree, and analyze the resulting tree to interpret the evolutionary relationships between the sequences. Discuss the implications of these relationships in the context of their genetic similarities and differences. [10 Marks]  1. GATTACAAGT  2. CCTAGGTTGA  3. TGGCATACTG  4. ATGCGTAGCT  5. GGTCAGTCAA | **20 Marks** | **L3** | **CO4** |
| **b.** | Interpret the information provided by a sequence logo and its relevance in bioinformatics.  [10 Marks] |
| **Or** | | | | | |
| **18.** | **a.** | Identify GATTACA motif in the following sequences using GIBBS algorithm. [10 Marks]  ACCATGACAG (random selected sequence)  GA**GTATACCT**  CAT**GCTTACT**  **CGGAATG**CAT | **20 Marks** | **L3** | **CO** |
| **b.** | Apply phylogenetic tree represents evolutionary relationships among species. [10 Marks] |

**\*\*\*\*\* BEST WISHES \*\*\*\*\***