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



**PRESIDENCY UNIVERSITY**

**Bengaluru**

|  |
| --- |
| **End - Term Examinations – JANUARY 2025** |
| Date: 07 – 01- 2025 Time: 09:30 am – 12:30 pm |

|  |  |  |
| --- | --- | --- |
| **School:** SOCSE | **Program:** B.Tech- CSE | |
| **Course Code :** CSE3069 | **Course Name :** Introduction to Bioinformatics | |
| **Semester**: VII | **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** | Define Bioinformatics and its importance. | **2 Marks** | **L1** | **CO1** |
| **2** | State the application areas of Bioinformatics. | **2 Marks** | **L1** | **CO1** |
| **3** | List the family of biological molecules. | **2 Marks** | **L1** | **CO1** |
| **4** | Describe Phylogenetic. | **2 Marks** | **L2** | **CO2** |
| **5** | Describe Genomics. | **2 Marks** | **L2** | **CO2** |
| **6** | List the primary function of carbohydrates. | **2 Marks** | **L1** | **CO2** |
| **7** | Explain Functional Genomics. | **2 Marks** | **L2** | **CO3** |
| **8** | Calculate the complementary strand for the following sequence: GCTACTGTAAACTG. | **2 Marks** | **L3** | **CO3** |
| **9** | Explain which organization/company maintains the GenBank. | **2 Marks** | **L2** | **CO4** |
| **10** | State the importance of RNA. | **2 Marks** | **L1** | **CO4** |

**Part B**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Answer the Questions Total 80 Marks.** | | | | | |
| **11.** | **a.** | Differentiate between DNA AND RNA with neat diagram  [10 Marks] | **20 Marks** | **L2** | **CO1** |
| **b.** | Explain the following. [10 Marks]  i. Alignment  ii. Sequence  iii. Plasma membrane  iv.Genomics  v.Phylogenetic. |
| **Or** | | | | | |
| **12.** | **a.** | Explain how reverse complement sequences are generated and their utility in Bioinformatics. [10 Marks] | **20 Marks** | **L2** | **CO1** |
| **b.** | Explain the principles of structural and functional analysis to explain the role of a specific biological molecule in a cellular process. Provide a detailed example, including how the structure of the molecule relates to its function. [10 Marks] |
|  |  |  |  |  |  |
| **13.** | **a.** | Define Bioinformatics. Explain 3 components of bioinformatics along with applications of bioinformatics. [10 Marks] | **20 Marks** | **L2** | **CO2** |
| **b.** | Explain the processes of transcription and translation in gene expression. [10 Marks] |
| **Or** | | | | | |
| **14.** | **a.** | What is Biological Database? Explain the Database Architecture and types of database in detail. [10 Marks] | **20 Marks** | **L2** | **CO2** |
| **b** | Differentiate between DNA sequencing methods and explain their respective applications. [10 Marks] |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **15.** | **a.** | Illustrate substring reconstruction with proper steps. [10 Marks] | **20 Marks** | **L3** | **CO3** |
| **b.** | Apply the Needleman-Wunsch Algorithm to the following sequences and present the global alignment: [10 Marks]  Seq1 = ATCGGATTCAA  Seq2 = CGGTCA  Match: +3 Mismatch: -2 Gap: -2 |
| **Or** | | | | | |
| **16.** | **a.** | Differentiate FASTA and FASTQ with pros and cons in the file formats with an example. [10 Marks] | **20 Marks** | **L3** | **CO3** |
| **b.** | Compute the index for the following sequence [10 Marks]  Pattern = "ATTG"  Text = "AATGTTGCATTGACGATTGCATTGGGGATTATTGGTTT"  d = 1 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **17.** | **a.** | Identify the GATTACA motif in the following sequences using GIBBS algorithm. [10 Marks]  A**CCATGACA**G  GAGTATACCT (random selected sequence)  CAT**GCTTACT**  C**GGAATGC**AT | **20 Marks** | **L3** | **CO4** |
| **b.** | Plan the steps involved in finding the best match of sequence using Basic Local Alignment Search Tool (BLAST). [10 Marks] |
| **Or** | | | | | |
| **18.** | **a.** | Show construction of a sequence logo for the following sequences and analyze the conservation of nucleotides at each position 1. GATTACAAGT  2. CCTAGGTTGA  3. TGGCATACTG  4. ATGCGTAGCT  5. GGTCAGTCAA  6. ACTGCTTGGA  7. TTGACCGTAG   Analyze the resulting sequence logo to determine the significance of any conserved regions. Discuss how these conserved regions might relate to functional or structural aspects of the sequences. [10 Marks] | **20 Marks** | **L3** | **CO4** |
| **b.** | 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. ATGCTATCTA  2. ATCCTCACTT  3. TCGTCATCAT  4. TATCTATCTA  5. ACTGCTACAT |

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