Reg.No. \_\_\_\_\_\_\_\_\_\_\_\_\_



**End Semester Examination – Nov / Dec – 2019**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| **Code :** | **14BI2001** | Duration : | **3hrs** |
| **Sub. Name :** | **ANALYTICAL BIOINFORMATICS** | Max. marks : | **100** |

**ANSWER ALL QUESTIONS (5 x 20 = 100 Marks)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Q. No.** | **Sub Div.** | **Questions** | **Course**  **Outcome** | **Marks** |
| 1. | a. | Describe the use of principal component in biological data analysis. | CO1 | 4 |
| b. | Define database. Write a detailed note on major databases containing genomic data with data file format information. | CO2 | 16 |
| (OR) | | | | |
| 2. | a. | Compute the dynamic programming table, alignments and associated sequence identities for the two strings. You have a scoring scheme where  A match gives +8, mismatch gives −5, a gap costs −3  Illustrate the best Local alignment for the same two DNA sequences.  Sequence A: CTTAACTGA  Sequence B: CGGATCATCA | CO2 | 20 |
|  |  |  |  |  |
| 3. | a. | Define scoring matrix. | CO2 | 3 |
|  | b. | Explain briefly the methods of analysis used to find values for a PAM and BLOSUM amino acid similarity matrix. | CO3 | 17 |
| (OR) | | | | |
| 4. | a. | Explain the importance of heuristic algorithms in biological sequence analysis. Also describe in detail the algorithm scheme of BLAST and FASTA. | CO2 | 20 |
|  |  |  |  |  |
| 5. | a. | Differentiate between Parsimony, Distance and Likelihood-based algorithms with limitation. | CO1 | 10 |
|  | b. | What is a neighbour joining tree? How it is different from UPGMA? | CO2 | 10 |
| (OR) | | | | |
| 6. | a. | Describe the format of microarray data. | CO1 | 3 |
|  | b. | Breifly describe the software tool intergarated protein protein interaction (PPI) databases. | CO1 | 17 |
|  |  |  |  |  |
| 7. | a. | Label the online source available databases and tools and methods involved in promoter prediction . | CO1 | 10 |
|  | b. | Explain the principle involved in gene prediction. Justify the difference in methods used for gene prediction in prokaryotes and eukaryotes. | CO1 | 10 |
| (OR) | | | | |
| 8. | a. | Define and differentiate between motif and domain. | CO2 | 6 |
|  | b. | Describe how to build an HMM to identify exons and introns in genome sequences. | CO3 | 7 |
|  | c. | Discuss the Smith–Waterman algorithm. What is the complexity and the relationship with the problem of ﬁnding the longest common subsequences? |  | 7 |
|  | | **Compulsory**: |  |  |
| 9. | a. | Give a brief account on various protein sequence databases and data formats. | CO1 | 10 |
|  | b. | Briefly describe the secondary and tertiary protein structural databases with integrated sotftware solution. | CO1 | 10 |