

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

**NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA**

(An Autonomous Institute Affiliated to AKTU, Lucknow)

B.Tech.

SEM: III - THEORY EXAMINATION (2021 - 2022)

Subject: Bioinformatics

Time: 03:00 Hours

Max. Marks: 100

General Instructions:

1. All questions are compulsory. It comprises of three Sections A, B and C.
  - Section A - Question No- 1 is objective type question carrying 1 mark each & Question No- 2 is very short type questions carrying 2 marks each.
  - Section B - Question No- 3 is Long answer type - I questions carrying 6 marks each.
  - Section C - Question No- 4 to 8 are Long answer type - II questions carrying 10 marks each.
  - No sheet should be left blank. Any written material after a Blank sheet will not be evaluated/checked.

SECTION A

20

1. Attempt all parts:-

- |      |   |   |
|------|---|---|
| 1-a. | Which database contain the 3D structure of protein? (CO1)   | 1 |
|      | <ol style="list-style-type: none"> <li>1. GenBank</li> <li>2. DDBJ</li> <li>3. PDB</li> <li>4. EMBL</li> </ol>  |   |
| 1-b. | PatentScope is governed by which organization? (CO1)  | 1 |
|      | <ol style="list-style-type: none"> <li>1. CIPO</li> <li>2. WIPO</li> <li>3. Ekaswa</li> <li>4. EPIDOS</li> </ol>  |   |
| 1-c. | Which is not multiple sequence alignment tool? (CO2)  | 1 |
|      | <ol style="list-style-type: none"> <li>1. CLUSTAL W</li> <li>2. T-COFFEE</li> <li>3. PHI-BLAST</li> <li>4. Pileup</li> </ol>  |   |
| 1-d. | Which algorithm is followed in global alignment? (CO2)  | 1 |
|      | <ol style="list-style-type: none"> <li>1. Smith-Waterman algorithm</li> <li>2. Needleman-Wunch algorithm</li> <li>3. k-tuple algorithm</li> <li>4. Heuristic algorithm</li> </ol> |   |
| 1-e. | Which is not a phylogenetic tree construction method? (CO3)   | 1 |
|      | <ol style="list-style-type: none"> <li>1. UPGMA</li> <li>2. REBASE</li> <li>3. Neighbor joining</li> <li>4. Maximum Parsimony</li> </ol>  |   |
| 1-f. | Which algorithm or method is used for preparation of evolutionary tree? (CO3)   | 1 |

1. Fitch-Margoliash
  2. Smith-Waterman
  3. Needleman-Wunch
  4. Heuristic
- 1-g. What does CRISPR stand for? (CO4) 1
1. Clustered International Societies of Photosynthesis Research
  2. It is the name of an American candy bar
  3. Clustered Regularly Interspaced Short Palindromic Repeats
  4. Clustered repeat sequences
- 1-h. The enzymatic activity of the Cas9 protein can be compared to: (CO4) 1
1. An anchor
  2. Scissors
  3. Glue
  4. Stapler
- 1-i. Procheck and Verify 3D are molecular docking tools. (CO5) 1
1. TRUE
  2. FALSE
- 1-j. \_\_\_\_\_ is the machine learning algorithms that can be used with labeled data. (CO5) 1
1. Supervised learning algorithms
  2. Unsupervised learning algorithms
  3. Heuristic algorithms
  4. All

2. Attempt all parts:-

- 2-a. What are model organism databases? Name any two. (CO1) 2
- 2-b. What are the common problems of blast tool? (CO2) 2
- 2-c. Differentiate between cladograms and dendrograms? (CO3) 2
- 2-d. What are the various applications of CRISPR? (CO4) 2
- 2-e. Explain classification and regression in terms of machine learning? (CO5) 2

#### SECTION B

30

3. Answer any five of the following:-

- 3-a. Name and explain various *in silico* tools for analysis of restriction digestion of DNA sequences? (CO1) 6
- 3-b. Give the names of tools used for biological data submission in NCBI? What are the major differences between these tools? (CO1) 6
- 3-c. How Entrez tool is integrated with different databases? What is its significance? (CO2) 6
- 3-d. Write down various components of BioMart- Ensembl databases? Explain each component in detail. (CO2) 6
- 3-e. Explain Fitch-Margoliash method of phylogenetic analysis in detail. How Fitch-Margoliash method is used in phylogenetic analysis? (CO3) 6
- 3-f. What do you mean by genome-wide association studies (GWAS)? How GWAS is useful in detecting the human diseases? (CO4) 6
- 3-g. Discuss about decision tree in detail? What are the applications of decision trees in machine learning? (CO5) 6

#### SECTION C

50

4. Answer any one of the following:-

- 4-a. What are the various conditions in which BankIt and Sequin are used for the sequence submission? (CO1) 10
- 4-b. If we want to know the information of any granted patent, where we can get that information? Explain the various search options. (CO1) 10

5. Answer any one of the following:-

- 5-a. Scoring matrices are the basis of sequence alignment. What are the different types of scoring matrices? Explain in detail. (CO2) 10
- 5-b. How global alignment is different from local alignment? Explain with the help of examples. (CO2) 10

6. Answer any one of the following:-

- 6-a. What is the importance of studying the evolution in bioinformatics? How the evolutionary relationship can be established between different species? Explain in detail. (CO3) 10
- 6-b. Write a brief note on UPGMA method of phylogenetic analysis. Explain the steps involved in the process of constructing phylogenetic tree using UPGMA. (CO3) 10

7. Answer any one of the following:-

- 7-a. What are the various advantages of genome editing technologies? Discuss the various methods of genome editing in detail. (CO4) 10
- 7-b. Explain the various types of biomarkers? What is the importance of biomarkers in disease diagnosis? Explain with the help of examples. (CO4) 10

8. Answer any one of the following:-

- 8-a. Describe the support vector machines (SVM) algorithm in detail. How does SVM works? What are the various types of SVM? (CO5) 10
- 8-b. Discuss the significance of molecular modelling in drug designing? Give the details of various methods of molecular modelling in bioinformatics? (CO5) 10