M. SC. BIOINFORMATICS SEM.-III (2013 COURSE)
(CHOICE BASED CREDIT SYSTEMS) : WINTER - 2017

SUBJECT: BIOLOGICAL DATA MINING

Day: Monday **Time:** 02.00 PM TO 05.00 PM
Date: 30/10/2017 **W-2017-1017** **Max. Marks: 60**

N.B:

1) **Q. No. 1and Q. NO. 5 are COMPULSORY.** Out of the remaining questions attempt **ANY TWO** from each sections.
2) Answer to both the sections should be solved in **SEPARATE** answer books.
3) Figures to the right indicate **FULL** marks.

SECTION-I

Q.1 Answer in brief: **(10)**

a) What are biological data types?
b) Diagrammatically explain flat file artifacts.
c) What are errors occurs in 3D structural data?
d) How micro array data gets authenticated?
e) What characteristics should be checked for sequence data?

Q.2 Write short notes on: **(ANY TWO)** **(10)**

a) Simulated annealing
b) Newton-Raphson method
c) Steepest Descent Method

Q.3 Answer the following: **(ANY TWO)** **(10)**

a) Differentiate between supervised and unsupervised genetic algorithm?
b) Explain with example utilization of genetic algorithms in bioinformatics.
c) Describe the drawbacks of genetic algorithms.

Q.4 Give a detailed account on *Ab initio* methods for structure prediction. **(10)**

OR

Explain in detail DNA-array data analysis pipeline. How do you differentiate between significant and non-significant data with the help of it.

SECTION-II

Q.5 Define the following with their applications. **(10)**

a) Dot plot b) GAP penalties c) MSA
d) Smith Waterman algorithm e) Cox- Jaynes Axiomes

Q.6 Answer the following: **(ANY TWO)** **(10)**

a) Explain in brief Needleman-Wunch algorithm.
b) Enlist and elaborate the homology modeling steps. Why we use it?
c) Describe any two structure validation methods.

Q.7 Write short notes on: **(ANY TWO)** **(10)**

a) Hidden Markov Model
b) Neural Network
c) Bayesian modeling

Q.8 What is Ant cottony optimization? Explain its algorithm in detail. How this techniques could be utilized for biological data? **(10)**

OR

What is fuzzy logic system? Explain with example its applications in clustering and classification of proteins.

* * * * *