1STELLA MARIS COLLEGE (AUTONOMOUS) CHENNAI –600 086
(For candidates admitted from the academic year 2011 – 12 & thereafter)

SUBJECT CODE: 11BI/PC/DM44

M. Sc. DEGREE EXAMINATION, APRIL 2016
BIOINFORMATICS
FOURTH SEMESTER

COURSE : CORE
PAPER : DATA MINING AND MACHINE LEARNING
TIME : 3 HOURS **MAX. MARKS: 100**

SECTION – A

ANSWER ALL QUESTIONS **(20 X 1=20)**

Fill in the Blanks:

1. Genetic Algorithm is based on feature_____.
2. _____ an essential process where intelligent methods are applied in order to extract data patterns.
3. The_____ is responsible for fetching the relevant data, based on the user's data mining request.
4. OLAP stands for _____ _____ _____ _____.
   a. Online Assessment Protocol.
   b. Online Analytical processing
   c. On Large databases Association program.
   d. On Largescale Assessment protocol.
5. Types of learning technique in Neural Networks are _____ & _____.
6. The following step involved in genetic algorithm is _____.
   a. Outlier analysis
   b. Crossing – over
   c. Clustering
   d. SOM
7. The KDD process in datamining stands for_____ _____ _____.
   a. Kyoto discovery in datamining.
   b. Knowledge Discovery in Databases.
   c. K – means databases
   d. K – means data in clustering databases.
8. Neural Networks are complex_____with many parameters.
9. The network that involves backward links from output to the input and hidden layers is called as_____.
10. A Bayesian network provides a _____ description of the domain.

Answer in a line or two:

11. Transactional Databases.
12. Hierarchial Clustering.
13. Binning methods.

..2

14. A feed forward network and a recurrent network.
15. Normalization.
16. Meta data
17. Data warehouse
18. ARM
19. SOM
20. Perceptron.

SECTION – B

ANSWER ANY FOUR QUESTIONS. EACH ANSWER SHOULD NOT EXCEED 500 WORDS. ALL QUESTIONS CARRY EQUAL MARKS. DRAW DIAGRAMS WHEREVER NECESSARY (4 X 10 = 40)

21. Association Analysis.
22. Data Mining functionalities.
23. Cross over Techniques.
24. Apriori Algorithm.
25. Support Vector Machine Technique.
26. Learning Algorithm.
27. Outlier Analysis.

SECTION – C

ANSWER ANY TWO QUESTIONS. EACH ANSWER SHOULD NOT EXCEED 1200 WORDS. ALL QUESTIONS CARRY EQUAL MARKS. DRAW DIAGRAMS WHEREVER NECESSARY (2 X 20 = 40)

28. Explain Genetic Algorithm, its features and methodology.
29. Describe the major clustering methods.
30. Classification and Prediction classification in Data Processing.
31. Outline the major research challenges of data mining in one specific application domain, such as stream/sensor data analysis, spatiotemporal data analysis, or bioinformatics.

****