

Post Graduate Diploma in Chemoinformatics
Annual Examinations – 2006

Paper: PGDC – 107
Data sequencing, Mining and Visualization

Time allowed: Three hours

Maximum Marks: 80

SECTION – I

Q1. Attempt all the objective type questions given below. All questions carry equal marks. 1X20

- i) Logistic regression is a generalization of _____.
- ii) Modeling is simply the act _____.
- iii) A pattern can be defined as given a set of facts _____ some measure of certainty.
- (iv) CART stands for _____.
- (v) One of the best ways to summarize data is to _____.
- (vi) The concept behind clustering is to _____ a large set of data into smaller groups based on a similarity metric.
- (vii) Drug discovery and development is in itself a search and _____ problem.
- (viii) The use of data warehouse is to _____.
- (ix) Common forms of interaction are _____.
- (x) Data visualization, data is transformed into _____.
- (xi) Decision tree are a way of representing a series of rules that lead to a _____.
- (xii) A _____ starts with an input layer, where each node corresponds to a predictor variable.
- (xiii) Data mining does not replace _____ technique.
- (xiv) Define Data?
- (xv) Define Information?
- (xvi) What is Knowledge?
- (xvii) The scientific field concerned with the creation of intelligent behavior in a ~~modern~~ *machine* is called _____.
- (xviii) The use of the data ware house is to _____.
- (xix) ECA stands for _____.
- (xx) Statistics is a branch of Mathematics concerning the _____.

SECTION – II

Attempt any five questions. All questions carry equal marks.

6 X 5

- Q2. Explain scalability in Data Mining.
- Q3. What technological Infrastructure is required for Data Mining?
- Q4. Describe common techniques of Data Visualization.

- Q5. Explain working of decision Tree.
- Q6. Describe Data Warehousing.
- Q7. Differentiate hierarchical and non-hierarchical clustering.
- Q8. Explain the logistics of personalized medicine.

SECTION – III

Attempt any three questions. All questions carry equal marks.

10 X 3

- Q10. Describe the application of data mining in Chemo informatics.
- Q11. What do you understand by Molecular Data Mining?
- Q12. Describe the applications of Data mining software.
- Q13. What do you understand by the term Neural Networks? Neural Net makes a prediction. Discuss.
- Q14. Explain Data Warehouse Design for Pharmaceutical Drug Discovery Research.
- Q15. Explain Data Visualization and write down the characteristics of data Visualization.