P. G. Diploma in Chemo-informatics
Annual Examination – 2010
Data Sequencing, Mining and Visualization
PGDC – 107

Time allowed : 2½ Hours   Max Marks:70

Section – A

Attempt all questions.   [5x2]
  1. What is Data Mining?
  2. What are the characteristics of a Data Warehouse?
  3. Define Orienteering.
  4. Differentiate between data and information.
  5. What are the basic data types in visualization system?

Section – B

Attempt any **five** questions   [5 x 6 = 30]
  1. Explain the stages/processes identified in data mining and knowledge discovery.
  2. Differentiate between data mining and machine learning.
  3. Explain OLTP. How it is different from data warehousing?
  4. What is clustering? How it is different from classification in data mining?
  5. Explain two softwares used fir chemical data mining.
  6. What do you mean by database search algorithms?
  7. Describe some potential applications of data mining.

Section – C

Attempt any **THREE** questions   [3 x 10 = 30]
  1. What is gene expression? Explain the five technologies involved in high throughout analysis of differential gene expression and purposes for which high throughput gene expression is used.
  2. Describe the various mapping methods used for data visualization and explain geometry viewer with the help of diagram.
  3. What do you mean structured Query Language? Explain with suitable example.
  4. Explain the concept of comparing different models using visualization.
  5. Explain the data warehouse model. What are the advantages of data warehouses over conventional systems?