JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

END SEMESTER EXAMINATION-2016

B.Tech (BI) VI Semester

COURSE CODE: 16B11BI612

MAX. MARKS: 35

COURSE NAME: Data Warehousing and Mining for Bioinformatics

COURSE CREDITS: 04

MAX. TIME: 2 HRS

Note: All questions are compulsory. Carrying of mobile phone during examinations will be treated as case of unfair means.

1. (a) How do we identify whether an association rule is strong? Explain on the basis of four different measures. (4)
(b) Why is it useful to identify closed frequent items and maximal frequent itemsets? Explain lucidly with examples. (4)
(c) Highlight the applications of association rule mining in bioinformatics. (2)
(d) Discuss the significance of Apriori algorithm and how is it used to mine association rules. What are its limitations and how are these solved? (5)

2. You are provided with the hormone level values for a cohort of individuals suffering from a particular problem. 200, 300,400, 600, 1000
Use these methods to normalize the following group of data: (a) min-max normalization by setting min=0 and max=1 (b) z-score normalization (c) normalization by decimal scaling (6)

3. Distinguish between: (6)
(a) Symmetric binary similarity and asymmetric binary similarity
(b) Worlds within worlds and Tree maps
(c) Scatter plot and scatter plot matrix

4. Explain what is concept hierarchy generation. Discuss the different types of operations which are performed on data cubes for generating concept hierarchies. (4)

5. Why and how do we perform wavelet transformation and Principal components analysis (PCA) on data prior to generating any machine learning based models? Which types of bioinformatics datasets can be transformed using these techniques and why? (4)