Pradeep Kr

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -2 EXAMINATION- 2016

B.Tech (IT), 6th Semester

COURSE CODE: 10B22CI622 MAX. MARKS: 25

COURSE NAME: Data Mining

COURSE CREDITS:4 MAX. TIME: 1Hr 30 Min

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

---

1. How can we further improve the efficiency of Apriori-based mining? **[3 Marks]**

2. What can we do to secure the privacy of individuals while collecting and mining data? **[3 Marks]**

3. Consider the Transactional Data for an all electronics branch as shown below:

TID	*List of Items_IDs*
T100	I_1, I_2, I_5
T200	I_2, I_4
T300	I_2, I_3
T400	I_1, I_2, I_4
T500	I_1, , I_3
T600	I_2, I_3
T700	I_1, , I_3
T800	I_1, I_2, I_3, I_5
T900	I_1, I_2, I_3

Mine the frequent itemsets using vertical data format and consider the minimum support as two for this problem. **[6 Marks]**

4. Outline the two-two example of data mining for each one of them as mentioned below:
   a. Retail Industry b. Telecommunication Industry c. Biological Data Analysis **[3 Marks]**

5. Consider a database has five transactions. Let *min sup* = 60% and *min con f* = 80%.

TID	Items_bought
T100	{ M,O,N,K,E,Y}
T200	{D,O,N,K,E,Y}
T300	{M,A,K,E}
T400	{M,U,C,K,Y}
T500	{C,O,O,K,I,E}

Find all frequent itemsets using Apriori and FP-growth, respectively. Compare the efficiency of the two mining processes. **[10 Marks]**