Code No: RT32052   **R13**   SET - 1

III B. Tech II Semester Supplementary Examinations, November -2018
DATA WARE HOUSING AND MINING
(**Common to** Computer Science Engineering and Information Technology)

Time: 3 hours   Max. Marks: 70

Note: 1. Question Paper consists of two parts (**Part-A** and **Part-B)**
2. Answering the question in **Part-A** is compulsory
3. Answer any **THREE** Questions from **Part-B**
*****

PART –A

1	a)	What is data mining? Give an example.	[3M]
	b)	Why concept hierarchies are useful in data mining.	[4M]
	c)	What is data integration and why it is necessary?	[4M]
	d)	What is classification? Explain briefly.	[4M]
	e)	What are the time and space complexities of K-means clustering algorithm?	[4M]
	f)	What is Apriori principle? Explain briefly.	[3M]

PART –B

2	a)	Explain how the evolution of database technology led to data mining.	[8M]
	b)	Describe any five advanced data base systems and applications.	[8M]
3	a)	What is data preprocessing? Why it is necessary?	[5M]
	b)	Explain different data cleaning methods?	[6M]
	c)	What is attribute subset selection? What are different methods used for this?	[5M]
4	a)	Compare and contrast OLAP and OLTP.	[8M]
	b)	What are different schemas for design of a data ware house? Explain with neat sketches.	[8M]
5	a)	Explain Algorithm for decision tree induction with suitable classification example.	[8M]
	b)	Specify the reasons for model overfitting and explain the methods to solve this problem.	[8M]
6	a)	What is association rule Mining problem? Explain Aprori algorithm for finding frequent item sets with example.	[8M]
	b)	What is the difference between mining frequent item sets with candidate generation and without candidate generation? Explain.	[8M]
7	a)	Write about Min, Max, and Average links used in clusterings.	[8M]
	b)	Explain K-means clustering algorithm with its additional issues.	[8M]

*****