Paper Id: **113224**   Roll No. ☐☐☐☐☐☐☐☐☐☐☐

B TECH
(SEM VI) THEORY EXAMINATION 2018-19
DATA WAREHOUSING AND DATA MINING

Time: 3 Hours   *Total Marks: 70*

Note: **1.** Attempt all Sections.

SECTION A

1. Attempt *all* questions in brief.   2 x 7 = 14

a.	Define Data Mining?
b.	What is classification?
c.	Explain data generalization.
d.	Why is data preprocessing required?
e.	Define data compression?
f.	Explain Association rules.
g.	What is backup and recovery?

SECTION B

2. Attempt any *three* of the following: 7 x 3 = 21

a.	Explain data cleaning and phases of data cleaning.
b.	What are the various requirements of clustering in data mining?
c.	How does a Snowflake schema differ from a STAR schema?
d.	Define the terms-- Aggregation, Historical Information and Query Facility.
e.	What is data warehousing with its advantages and disadvantages?

SECTION C

3. Attempt any *one* part of the following: 7 x 1 = 7

(a)	Explain data integration, transformation and loading.
(b)	Define data reduction? Write down the various strategies for data reduction.

4. Attempt any *one* part of the following: 7 x 1 = 7

(a)	Find out the mean, variance and standard deviation for the height of animals 555mm,450mm,165mm,410mm and 300mm.
(b)	Explain attribute relevance analysis with example.

5. Attempt any *one* part of the following: 7 x 1 = 7

(a)	How CURE is different from CHAMELEON.
(b)	Explain Density-based and Grid-based clustering methods.

6. Attempt any *one* part of the following: 7 x 1 = 7

(a)	Write down the difference between OLTP and OLAP.
(b)	Explain various architectures of data warehouse.

7. Attempt any *one* part of the following: 7 x 1 = 7

(a)	Explain web mining and types of web mining with example.
(b)	What are the essential differences between the ROLAP and MOLAP?