

BACHELOR OF SCIENCE (COMPUTER SCIENCE) (CBCS - 2018 COURSE)
S.Y.B.Sc.(Computer Science) Sem-IV :SUMMER- 2022
SUBJECT : DATA WAREHOUSING & DATA MINING-II

Day : Friday
Date : 15-07-2022

S-20113-2022

Time : 03:00 PM-06:00 PM
Max. Marks : 60

N.B.

- 1) All questions are **COMPULSORY**.
 - 2) Figures to the right indicate **FULL** marks.
 - 3) Draw neat and labelled diagram **WHEREVER** necessary.
-

Q.1 Attempt any **TWO** of the following: **(12)**

- a) Explain K – means algorithm.
- b) Explain association as a data mining method.
- c) Describe the decision tree learning method.

Q.2 Attempt any **TWO** of the following: **(12)**

- a) Explain Naïve Bayesian classification method.
- b) Explain cluster analysis in detail.
- c) Describe OLAP systems with its characteristics.

Q.3 Attempt any **TWO** of the following: **(12)**

- a) Explain data mining method classification with suitable example.
- b) Explain characteristics of data mining.
- c) Describe web mining with its advantages.

Q.4 Attempt any **THREE** of the following: **(12)**

- a) State the features of data warehouse.
- b) List and define the four steps of ETL process.
- c) What is prediction? Give one example for prediction.
- d) Write a note on machine learning.

Q.5 Attempt any **FOUR** of the following: **(12)**

- a) Compare classification with clustering.
- b) Write a short note on regression.
- c) Write a short note on outlier analysis.
- d) Write a short note on data discretization.
- e) What are the different types of data to be mined by data mining ?
- f) Define the following terms:
 - i) Data
 - ii) Support
 - iii) Confidence

* * *