



April 2021

2019/2012/PGR/PGA

JUSTICE BASHEER AHMED SAYEED COLLEGE FOR WOMEN (*Autonomous*),
CHENNAI – 600 018

IV SEMESTER (**Regular & Arrear**) – Applicable to candidates admitted in the
year 2019 & 2012

II M.Sc. Computer Science/ II M.C.A

Data Mining – MGIVC9(6)/MHIVC12(6)

TIME: 90 Minutes

MAXIMUM MARKS : 50

The first ten minutes should be used for reading the question paper only. The students should not begin to answer the questions in the first ten minutes.

Section-A

Answer all the questions.

15 Marks

I. Fill in the blanks

(5X1=5 Marks)

1. KDD stands for _____.
2. _____ refers how much the data are targeted by users.
3. OLAM stands for _____.
4. A _____ of class – labelled tuples is used to estimate cost complexity.
5. Clustering can also be used for _____.

II. Choose the correct answer

(5X1=5 Marks)

6. To remove noise and inconsistent data is called
 - a) Data Integration
 - b) Data Cleaning
 - c) Data Mining
 - d) Data Pre process
7. Data mining often requires
 - a) Data cleaning
 - b) Data Pre process
 - c) Data Integration
 - d) Data Transformation
8. Apriori is a seminal algorithm proposed by
 - a) Agarwal
 - b) Srikant
 - c) Agarwal & Srikant
 - d) Apriori
9. In Decision tree, the top most node in a tree is the
 - a) Root node
 - b) Head node
 - c) Top node
 - d) List node
10. The method works by grouping data objects into a hierarchy.
 - a) Hierarchical clustering methods
 - b) Tree based method
 - c) Knowledge based method
 - d) Analytical method



April 2021

2019/2012/PGR/PGA

JUSTICE BASHEER AHMED SAYEED COLLEGE FOR WOMEN (*Autonomous*),
CHENNAI – 600 018
IV SEMESTER (**Regular & Arrear**) – Applicable to candidates admitted in the
year 2019 & 2012

II M.Sc. Computer Science/ II M.C.A

Data Mining – MGIVC9(6)/MHIVC12(6)

III. Answer the following in One or Two Sentences

(5X1=5 Marks)

11. Define Data mining.
12. Identify any four fields of data quality.
13. Mention the use of multi-dimensional data mining in cube space.
14. Define Classification.
15. What is Clustering?

Section – B

(3 x 5 = 15 Marks)

Answer any three questions All questions carry equal marks.(Each answer should not exceed 150 words)

16. Explain the concept of data warehouse.
17. Discuss an overview of data reduction strategies.
18. Explain about association rule mining.
19. Analyse the concept of back propagation.
20. Illustrate an overview of basic clustering methods.
21. Analyse the concept of web mining.
22. Evaluate the concept of Naïve Bayesian classification method.
23. Explain briefly about types of Outliers.

Section – C

(2 x 10 = 20 Marks)

Answer any two questions. All questions carry equal marks.(Each answer should not exceed 300 words)

24. Discuss various major issues in Data mining.
25. Analyse the major tasks in Data pre-processing.
26. Describe in detail about Apriori algorithm.
27. Classify various typical requirements of clustering in data mining.
28. Elaborate the concept of Spatial and Temporal mining.

----- All the Best -----



April 2021

2019/2012/PGR/PGA

JUSTICE BASHEER AHMED SAYEED COLLEGE FOR WOMEN (*Autonomous*),
CHENNAI – 600 018

IV SEMESTER (**Regular & Arrear**) – Applicable to candidates admitted in the
year 2019 & 2012

II M.Sc. Computer Science/ II M.C.A

Data Mining – MGIVC9(6)/MHIVC12(6)
