

V Semester B.C.A. Examination, February/March 2024
(NEP Scheme)

(Freshers)

COMPUTER SCIENCE
Data Mining (Elective – I)

Time : 2½ Hours

Max. Marks : 60

Instruction : Answer **any 4** questions from **each** Sections.

SECTION – A

I. Answer **any 4** questions. **Each** carries 2 marks.

(4×2=8)

- | | |
|--|---|
| 1) Define KDD and Data mining. | 2 |
| 2) What is market basket analysis ? | 2 |
| 3) What is correlation ? | 2 |
| 4) Explain similarity measures. | 2 |
| 5) Differentiate between bottom-up and top-down strategy in hierarchical clustering. | 2 |
| 6) Define support and confidence in Association rule mining. | 2 |

SECTION – B

II. Answer **any 4** questions. **Each** carries 5 marks.

(4×5=20)

- | | |
|---|---|
| 7) With an example, explain where data mining is crucial to the success of a business. What data mining functionalities does this business need ? | 5 |
| 8) Explain Data mining process in detail. | 5 |
| 9) State Bayes Theorem. Explain Bayesian classification. | 5 |
| 10) Discuss data mining issues in detail. | 5 |
| 11) Write sampling algorithm. | 5 |
| 12) What do you understand by outliers ? Explain with an example. | 5 |



SECTION – C

III. Answer **any 4** questions. **Each** question carries **8** marks.

(4×8=32)

13) Discuss the tasks of data mining with an examples.

8

14) Explain KNN classification in detail with an example.

8

15) a) Using the data given below, draw OC curves assuming that the output 2 column is the correct classification and output 1 is what is seen.

Draw 3 curves.

5

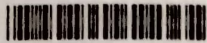
b) Construct a confusion matrix assuming output is the correct assignment and output 1 is actually made.

3

Name	Gender	Height	Output 1	Output 2
Kristina	F	1.6m	Short	Medium
Jim	M	2 m	Tall	Medium
Maggie	F	1.9m	Medium	Tall
Martha	F	1.88m	Medium	Tall
Stephanie	F	1.7m	Short	Medium
Bob	M	1.85m	Medium	Medium
Kathy	F	1.6m	Short	Medium
Dave	M	1.7m	Short	Medium
Worth	M	2.2m	Tall	Tall
Steven	M	2.1m	Tall	Tall
Debbie	F	1.8m	Medium	Medium
Todd	M	1.95m	Medium	Medium
Kim	F	1.9m	Medium	Tall
Amy	F	1.8m	Medium	Medium
Wynette	F	1.75m	Medium	Medium

16) Explain Algometric algorithm with an example.

8



- 17) For the following data, construct a decision tree and explain the terms Root node, Decision node, leaf node, sub pruning, parent node and child node.

8

Day	Weather	Temperature	Humidity	Wind	Play
1	Sunny	Hot	High	Weak	No
2	Cloudy	Hot	High	Weak	Yes
3	Sunny	Mild	Normal	Strong	Yes
4	Cloudy	Mild	High	Strong	Yes
5	Rainy	Mild	High	Strong	No
6	Rainy	Cool	Normal	Strong	No
7	Rainy	Mild	High	Weak	Yes
8	Sunny	Hot	High	Strong	No
9	Cloudy	Hot	Normal	Weak	Yes
10	Rainy	Mild	High	Strong	No

- 18) Explain Apriori Algorithm.

8
