

SECTION-B

2. What is data preprocessing. Explain major tasks in data preprocessing.
3. Explain the process of ranking of web pages.
4. Explain k-means clustering method by taking a suitable example.
5. How to we evaluate classifier's accuracy? Explain different techniques for accuracy estimation
6. Explain the process of web data mining in detail.

SECTION-C

7. Consider the training examples shown in below table for classification problem.

Gender	Car Type	Shirt Size	Class
M	Family	Small	C0
M	Sports	Medium	C0
M	Sports	Medium	C0
M	Sports	Large	C0
M	Sports	Extra Large	C0
M	Sports	Extra Large	C0
F	Sports	Small	C0
F	Sports	Small	C0
F	Sports	Medium	C0
F	Luxury	Large	C0
M	Family	Large	C1
M	Family	Extra Large	C1
M	Family	Medium	C1
M	Luxury	Extra Large	C1
F	Luxury	Small	C1
F	Luxury	Small	C1
F	Luxury	Small	C1
F	Luxury	Medium	C1
F	Luxury	Medium	C1
F	Luxury	Medium	C1
F	Luxury	Large	C1

Apply the split algorithm based on information theory to perform classification of the given dataset. Indicate all intermediate steps properly.

8. Write short notes on :

a) Hierarchy in the web

b) Web structure mining

9. Explain OLAP server architecture. Differentiate ROLAP, MOLAP and HOLAP.

NOTE : Disclosure of Identity by writing Mobile No. or Making of passing request on any page of Answer Sheet will lead to UMC against the Student.