

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST-2 EXAMINATION (Oct 2018)

B-Tech (7th SEM)

Dr. Rajinder Sandhu

Course Code: 11B1WCI832

Max. Marks: 25

Course Name: INFORMATION RETERIVAL

AND DATA MINING

Max. Time: 1 Hr 30 Min

Course Credit: 3

Note: All questions are compulsory

Q. No. 1 Why is naïve Bayesian classification called naïve? Briefly outline the major ideas and working of naïve Bayes classification. [3 Marks]
[CO-2]

Q. No. 2 Write an algorithm for k-nearest neighbour classification given k and n, the number of attributes describing each tuple. [3 Marks]
[CO-2]

Q. No. 3 Write an algorithm for FCM with explaining the computation of sum of squared errors, centroid and updating fuzzy pseudo partition. [3 Marks]
[CO-3]

Q. No. 4 Use the distance measure in Table below, perform single link hierarchical clustering. Show your results by drawing a dendrogram. The dendrogram should clearly show the order in which the points are merged. [5 Marks]
[CO-3]

	P1	P2	P3	P4	P5	P6
P1	0.00	0.71	5.66	3.61	4.24	3.20
P2	0.71	0.00	4.95	2.92	3.54	2.50
P3	5.66	4.95	0.00	2.24	1.41	2.50
P4	3.61	2.92	2.24	0.00	1.00	0.50
P5	4.24	3.54	1.41	1.00	0.00	1.12
P6	3.20	2.50	2.50	0.50	1.12	0.00

Q. No. 5 Explain the following terminology in relation to performance evaluation of classification algorithm with example: [5 Marks]
[CO-4]

- Confusion Matrix
- TP, FN, FP, TN
- Accuracy
- Precision
- Recall

Q. No. 6 Write short note on: [6 Marks]
[CO-3]

- Prototype Clustering
- Grid Based Clustering
- Chameleon Clustering