

DATA AND TEXT MINING

DOCUMENTATION RESEARCH AND TRAINING CENTRE, ISI-BC

MID TERM EXAMINATION (2023)

Total Marks: 40

Time: 90 mins.

ANSWER ALL

1. Explain the major issues in Data Mining operations. [4]
2. Find the L^1 , L^2 and L^∞ norm of a pattern $F = [-4, 10, -7, 9, 5]$. Also find the unit vector of F . [3+2]
3. Name TWO significances of Eigen values and Eigen vectors in a database of 10 samples and 5 features for each sample. [2]
4. Describe the two differences in the database management and data mining operations [2]
5. What is the process of Knowledge Discovery in Databases? Describe each of its operation with diagram. [4]
6.
 - Describe the conditional and decision attributes of a data base with examples. [2]
 - What are symmetric and asymmetric binary attributes? Explain with examples. [2]
7. Plot the locations of MEAN, MEDIAN and MODE in a symmetric, positively skewed and negatively skewed data set. [3]
8. Differentiate between feature selection and feature extraction operations in a decision making process. [2]
9. Find the mean and variance vector of [2+2]

$XX =$

4	5	3
3	2	3
3	2	4
1	0	7
2	3	4
5	7	8
7	9	4
3	7	6
2	1	1
1	0	8

10. What is interquartile range and find it for the following vector.
 - $X = [12, 5, 22, 30, 7, 36, 14, 42, 15, 53, 25]$ [1+3]
11. Differentiate between Histogram and Box And Whisker Plot of a database [4]
12. Differentiate the process of decision-making using classification and clustering operations. [2]