DATA AND TEXT MINING

DOCUMENTATION RESEARCH AND TRAINING CENTRE, ISI-BC

MID-TERM EXAMINATION   (2025)

Total Marks: 40   **Time: 90 mins.**

ANSWER Any FIVE from SEC: A

and all from SEC: B

SEC: A

1. Explain the major challenges of the data mining process.   [7]
2. Discuss the differences in the process of data mining and databases with examples.   [7]
3. List ANY six types of attributes used in a data mining task with examples.   [7]
4. Why data pre-processing is so important? Explain any four types of pre-processing approaches with examples.   [3+4]
5. What is Knowledge Discovery in Databases (KDD) refer to? Provide a complete flow process with discussion.   [7]
6. Find the variance-covariance matrix for the following dataset   [7]

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

SEC: B

7. Describe the motivations for standardization and normalization of data sets. Give at least one method of performing these operations   [5]

=====================END of the Question Paper=====================