
*Answer any FIVE Questions
All Questions Carry Equal Marks*

- | | | |
|------|---|-----|
| 1. a | Define KDD. Explain what are the Data Mining tasks. | 5M |
| b | Given a similarity measure with values in the interval $[0,1]$ describe two ways to transform this similarity value into a dissimilarity value in the interval $[0,\infty]$. | 7M |
| 2. a | Explain about evaluating the performance of a classifier. | 6M |
| b | How to build a decision tree? Explain. | 6M |
| 3. a | Explain frequent item set generation in the Apriori algorithm. | 6M |
| b | What are the methods for expressing attribute test conditions? Explain. | 6M |
| 4. a | Discuss in detail about agglomerative hierarchical clustering. | 7 M |
| b | Write the basic K-means algorithm and explain. | 5 M |
| 5. a | Explain the web terminology related to web data mining. | 6M |
| b | Explain about web Structure Mining. | 6M |
| 6. a | What is data preprocessing? Explain in detail. | 8M |
| b | What are different types of clustering? Explain. | 4M |
| 7. a | Describe strengths and weaknesses of K-means. | 6M |
| b | Explain the web Data mining architecture. | 6M |
| 8. a | What are summary statistics? Explain. | 6M |
| b | What are OLAP operations in the multidimensional data model? Explain. | 6M |
