



Q.P. Code: 24210

[Total Marks 80]

3 Hours)

- i. Q. 1. is Compulsory.
- ii. Attempt any three from the remaining.
- iii. Assume suitable data.

- Q. 1 (a) Explain Regression based classifier (5)
- (b) What is Data Preprocessing? Explain the different methods for the Data Cleansing phase (5)
- (c) What is hierarchical clustering? Explain any two techniques for finding distance between the clusters in hierarchical clustering. (5)
- (d) Define "Data mining", Enumerate five example applications that can benefit by using Data Mining (5)
- Q. 2 (a) Explain clearly working of DBSCAN using appropriate diagrams (10)
- (b) Explain Data mining as a step in KDD. Give the architecture of typical Data Mining system (10)
- Q. 3 (a) Explain BIRCH algorithm with example (10)
- (b) Explain different visualization techniques that can be used in data mining (10)
- Q. 4 (a) Explain Multilevel association rules with suitable examples (10)
- (b) Define classification, issues of classification and explain ID3 classification with example (10)
- Q. 5 (a) Explain Bagging and Boosting of classifiers (10)
- (b) Explain Box plot summary with example (10)
- Q. 6 (a) Explain Business Intelligence Issues (10)
- (b) What is clustering? Explain k-means clustering algorithm. Suppose the data for clustering - {2, 4, 10, 12, 3, 20, 11, 25} Consider $k=2$, cluster the given data using above algorithm. (10)