



# P.E.S. College of Engineering, Mandya - 571 401

(An Autonomous Institution affiliated to VTU, Belagavi)

**Fifth Semester, B.E. - Computer Science and Engineering**

**Semester End Examination; Feb. - 2021**

**Data Mining and Warehousing**

Time: 3 hrs

Max. Marks: 100

### Course Outcomes

The Students will be able to:

CO1: Analyze different data models used in data warehouse.

CO2: Apply different preprocessing techniques for different attributes.

CO3: Determine frequent item set using association rules.

CO4: Apply different classification techniques to classify the given data set.

CO5: Analyze different clustering techniques.

**Note:** I) PART - A is compulsory. Two marks for each question.

II) PART - B: Answer any **Two** sub questions (from a, b, c) for Maximum of 18 marks from each unit.

Q. No.	Questions	Marks	BLs	COs	POs
<b>I : PART - A</b>		<b>10</b>			
I a.	When slice and dice operations are used?	2	L2	CO1	PO1
b.	Write two examples to illustrate the issues of data mining.	2	L2	CO2	PO1
c.	Give an example when rule based classifier and Bayesian classifiers are applied for classification.	2	L2	CO3	PO2
d.	Distinguish between maximal frequent item set and minimal item set.	2	L2	CO4	PO1
e.	How does DBSCAN quantify the neighborhood of an object?	2	L2	CO5	PO2
<b>II : PART - B</b>		<b>90</b>			
<b>UNIT - I</b>		<b>18</b>			
1 a.	Distinguish between OLTP and OLAP with two examples for each.	9	L4	CO1	PO1
b.	Construct a snow-flake scheme for a data warehouse which consists of three dimensions time, doctor and patient and the two measures 'count' and 'charge', where 'charge' is the fee that a doctor charges a patient for a visit. Assume suitable attributes for each table.	9	L4	CO1	PO1
c.	Illustrate with an example how data generalization is done?	9	L4	CO1	PO1
<b>UNIT - II</b>		<b>18</b>			
2 a.	Illustrate with an example, how data mining is crucial to the success of a business. What data mining functions does this business need?	9	L4	CO2	PO3
b.	Explain the importance of data preprocessing before constructing a data warehouse.	9	L4	CO2	PO1
c.	What is data reduction? Explain any one data reduction technique.	9	L4	CO2	PO2
<b>UNIT - III</b>		<b>18</b>			
3 a.	Write an algorithm for decision tree induction. Explain with an example.	9	L4	CO3	PO3
b.	Explain with an example how and when Bayesian classifier is used for classification?	9	L3	CO3	PO2
c.	Write and explain basic sequential covering algorithm.	9	L2	CO3	PO1

**UNIT - IV****18**

- 4 a. Explain with an example how FP-Tree is constructed. 9 L2 CO4 PO2
- b. Define the terms:
- i) Frequent item sets 9 L2 CO4 PO1
  - ii) Support and Confidence
  - iii) Association Rules
- c. Generate all frequent item sets for the transaction given in Table. 4(c) using Apriori principle. Assume minimum support count = 2.

Transaction Id	Items
100	I1; I3; I4
200	I2; I3; I5
30	I1; I2; I3; I5
40	I2; I5
50	I1; I2; I3; I4

9 L5 CO4 PO3

**UNIT - V****18**

- 5 a. Illustrate with example how K-means clustering algorithm is used? 9 L4 CO5 PO3
- b. Mention the difference between clustering and classification. Illustrate with an example when clustering is applied? 9 L3 CO5 PO3
- c. Distinguish between Hierarchical clustering and partitioning clustering techniques. 9 L4 CO5 PO2

\* \* \*