

## Question Bank DWDM

### Unit-1

1. Explain about different schemas in dataware houses.
2. Write difference between OLAP & OLTP
3. Explain about dataware house Architecture.
4. What are the OLAP operations performed in Data Cube?
5. Write short notes about OLAP, ROLAP, MOLAP, and HOLAP

### Unit-2

1. Explain KDD process with neat diagram.
2. Explain about data mining architecture or Integration of Data mining system with a Database.
3. Write application of Data Mining.
4. What are Major issues in Data Mining
5. Write steps in Data warehouse implementation.

### Unit-3

1. Define and describe the use of mean, mode, median, standard deviation in the representation of data objects.
2. Explain data processing steps with example.
3. Write short notes about Measuring Data Similarity and Dissimilarity.
4. Write short notes about data cleaning and Data Transformation.
5. What is Metadata? Explain different types of Meta data.

### Unit-4

1. What is cluster analysis?. Discuss categorization of clustering methods.
2. Explain with relevant example, how association rules can be mined from large datasets.
3. Explain Apriori Algorithm and solve the below problem assume  $\text{min\_sup}=2$ .

Transaction ID	Items Bought
T100	I1,I2,I5
T200	I2,I4
T300	I2,I3
T400	I1,I2,I4
T500	I1,I3
T600	I2,I3
T700	I1,I3
T800	I1,I2,I3,I5
T900	I1,I3,I3

4. What is association rule mining? Define five metrics used for measuring classifier performance.
5. Explain about Market-Basket Analysis? With suitable example.
6. Explain about K-Mean and K-medoids Algorithm with example.

### Unit-5

1. What is classification? Explains two steps in classification.
2. Solve the below problem using Decision Tree Induction and Draw the Classification Tree.

Age	Income	Student	Credit Rating	Buy Computer
Youth	High	No	Fair	No
Youth	High	No	Excellent	No
Middle_age	High	No	Fair	Yes
Senior	Medium	No	Fair	Yes
Senior	Low	Yes	Fair	Yes

Senior	Low	Yes	Excellent	No
Middle_age	Low	Yes	Excellent	Yes
Youth	Medium	No	Fair	No
Youth	Low	Yes	Fair	Yes
Senior	Medium	Yes	Fair	Yes
Youth	Medium	Yes	Excellent	Yes
Middle_age	Medium	No	Excellent	Yes
Middle_age	High	Yes	Fair	Yes
Senior	Medium	No	Excellent	No

3. Explain Bayesian classification algorithm and analyze the probability for person buy computer or not above problem.
4. Explain about Rule Based Classification and Design extract the list of Rules from the Problem.
5. Describe about the Prediction and Linear Regression, non-linear regression.

#### **UNIT-6**

1. Explain about The Business Pressures-Responses- Support Model with neat Diagram.
2. Differentiate Transaction Processing versus Analytic Processing
3. What is Business Intelligence and Discuss the A Framework for Business Intelligence.
4. Write about Major Tools and Techniques of Business Intelligence.
5. Write short notes about Intelligence Creation and Use and BI Governance and Successful BI Implementation.

### Assignment Questions

1. Explain with relevant example how association rules can be mined from large data sets
2. What is cluster analysis? Discuss categorization of clustering methods. Explain k-mean clustering method
3. Discuss the challenges in creating and maintaining a data warehouse
4. Explain Data Cleaning and Data Transformation concepts
5. Explain different types of attributes with example and basic descriptors as, mean, mode and median
6. Explain what are the practical application of data mining
7. Given the transactions below

100	{J, B}
200	{M, Ch, Br, S}
300	{C, Br}
400	{M, Br, S, P, A, T}
500	{Ch, S, Bf}
600	{J, Br, K}

- i) What is the support and confidence for following rules?  
 Rule1 : M  $\Rightarrow$  B    Rule2 : Br  $\Rightarrow$  M
  - ii) If the minimum support is set to 50%, how many large itemsets will be formed?
8. Explain rule induction method using sequential covering algorithm. Explain the steps of the rule generation in brief for a rule based classifier.
  9. What data dissimilarity? Given two objects represented by the tuples (22,1,42,10) and (20, 0,36,8).
    - i) Compute the Euclidean distance between the two objects
    - ii) Compute Manhattan distance between the two objects
    - iii) Compute the Minkowski distance between the two objects, using  $q = 3$
  10. Write short notes on
    - i) Use of dashboards
    - ii) Classification and Prediction
  11. Write a note on
    - iii) Transaction processing Vs. Analytic processing
    - iv) Major tools and techniques of BI