

**H.T.No.**

--	--	--	--	--	--	--	--	--	--

**Code No: CS1913****GEC-R14****M. Tech II Semester Regular/Suppl. Examinations, July 2017****DATA MINING AND DATAWAREHOUSING****(Computer Science Engineering)****Time: 3 Hours****Max. Marks: 60****Note:** Answer any **FIVE** questions. All Questions carry equal Marks.**5 × 12 = 60M**

1. a) Explain KDD process with a neat diagram? (6M)  
b) What are the various methods of attribute subset selection and explain? (6M)
2. a) What is normalization and explain methods for normalization (6M)  
b) Why is data mining said to be multi-disciplinary field? Explain. (6M)
3. a) Distinguish OLTP and OLAP? (6M)  
b) Discuss various operations on data cube. (6M)
4. For the given data set create a Decision Tree? And explain about the knowledge obtained from it? (12M)

Outlook	Temp(f)	Humidity (%)	Windy	Class
sunny	79	90	True	play
sunny	56	70	False	play
sunny	79	75	True	no play
sunny	60	90	True	no play
overcast	88	88	False	no play
overcast	63	75	True	play

5. a) What is a Decision Tree? What are the advantages and Disadvantages of Decision Tree classifications? (6M)  
b) Explain training Bayesian belief networks? (6M)
6. a) Explain support and confidence in Association rule mining? (6M)  
b) Explain Apriori algorithm? (6M)
7. a) What are the types of data in cluster analysis? (6M)  
b) Explain briefly about DBSCAN? (6M)
8. a) Explain K-Means clustering algorithm? (6M)  
b) Explain Agglomerative hierarchical clustering algorithm? (6M)

**\*\*\*\*\***