Question Paper

MANIPAL ACADEMY OF HIGHER EDUCATION

SCHOOL OF INFORMATION SCIENCES (SOIS)
SECOND SEMESTER ME - MEDICAL SOFTWARE
DEGREE EXAMINATION- APRIL
Wednesday, 25 April 2018
Time : 10:00 am to 1:00 pm
Data Mining - Elective 1 [MMS 616.2]

Marks: 100   **Duration: 180 mins.**

Answer all the questions.

1)	Explain the Major issues in Data Mining	(10)
2)	Discuss the different OLAP operations available in the Multidimensional Data Model.	(10)
3)	Why preprocess the data? Discuss the different strategies for data reduction 2+8	(10)
4)	Explain step by step the Decision Tree Induction algorithm	(10)
5)	Predict the class label of an unknown sample using naïve Bayesian classification, given the training data below. The unknown sample is,X= (age="<=30", income="medium", student="no", credit_rating="fair")	(10)

Age	Income	Student	credit_rating	class: buys_computer
<=30	High	No	Fair	No
<=30	High	No	Excellent	No
31...40	High	No	Fair	Yes
>40	Medium	No	Fair	Yes
>40	Low	Yes	Fair	Yes
>40	Low	Yes	Excellent	No
31...40	Low	Yes	Excellent	Yes
<=30	Medium	No	Fair	No
<=30	Low	Yes	Fair	Yes
>40	Medium	Yes	Fair	Yes

Age	Income	Student	Credit_rating	class: buys_computer
<=30	Medium	Yes	Excellent	Yes
31...40	Medium	No	Excellent	Yes
31...40	High	Yes	Fair	Yes
>40	Medium	No	Excellent	No

6) What is Association Rule Mining? Explain the Market Basket Analysis   2+8   (10)

7) Write a short note "Mining Frequent Patterns without Candidate Generation".
Construct the FP tree for the following Database.   (10)

TID	Items bought	(ordered)frequent items
100	{f, a, c, d, g, i, m, p}	{f, c, a, m, p}
200	{a, b, c, f, l, m, o}	{f, c, a, b, m}
300	{b, f, h, j, o, w}	{f, b}
400	{b, c, k, s, p}	{c, b, p}
500	{a, f, c, e, l, p, m, n}	{f, c, a, m, p}

8) What is cluster analysis? Explain in detail the major methods in cluster analysis   2+8   (10)

9) Write a short note on following advanced data mining   5x2   (10)
   a) Web mining   b) Spatial mining

10) Discuss the following data mining application domains   5x2   (10)
   a) Science and Engineering   b) Retail and Telecommunication

-----End-----