LOYOLA COLLEGE (AUTONOMOUS), CHENNAI – 600 034

M.C.A. DEGREE EXAMINATION – **COMPUTER APPLICATIONS**

FIFTH SEMESTER – **NOVEMBER 2019**

16/17PCA5MC03 – DATA MINING

Date: 31-10-2019 Dept. No. Max. : 100 Marks

Time: 09:00-12:00

Part-A

Answer ALL Questions **(10 * 2= 20)**

1. Differentiate data mining from data warehousing.
2. What do you understand about knowledge discovery?
3. Mention the various tasks to be accomplished as part of data pre-processing
4. Write the strategies for data reduction.
5. What is Bayesian theorem?
6. Compare clustering and classification.
7. What is meant by web usage mining?
8. Define text mining.
9. Differentiate between star and snowflake schema.
10. Write down some applications of data mining.

Part – B

Answer ALL Questions **(5 * 8= 40)**

11. a) Describe the architecture of typical data mining system with neat Sketch.

(or)

b) List out the primitives for specifying a data mining task.

12. a) What is back propagation? Explain classification by back propagation.

(or)

b) What is classification? Explain if then rule based classification with suitable example.

13. a) Explain hierarchical methods of clustering.

(or)

b) Discuss the different types of clustering methods.

14. a) What is web mining? Discuss the various web mining techniques.

(or)

b) Explain the mining of spatial databases.

15. a) Describe OLAP operations in multidimensional data model.

(or)

b) Diagrammatically illustrate and discuss the three tier data warehousing architecture.

Part – C

Answer any TWO Questions **(2 * 20= 40)**

16. a) Describe the data mining functionality and examine. What kinds of patterns can be mined?

b)

RID	Age	Income	Student	Credit_Rating	Class:buys_computer
1	Youth	High	No	Fair	No
2	Youth	High	No	Excellent	No
3	Middle_aged	High	No	Fair	Yes
4	Senior	Medium	No	Fair	Yes
5	Senior	Low	Yes	Fair	Yes
6	Senior	Low	Yes	Excellent	No
7	Middle_aged	Low	Yes	Excellent	Yes
8	Youth	Medium	No	Fair	No
9	Youth	Low	Yes	Fair	Yes
10	Senior	Medium	Yes	Fair	Yes
11	Youth	Medium	Yes	Excellent	Yes
12	Middle_aged	Medium	No	Excellent	Yes
13	Middl_aged	High	Yes	Fair	Yes
14	Senior	Medium	No	Excellent	No

Predict the class label of a tuple using naïve Bayesian classification for the given training data. The tuple to classify is $\boldsymbol{X}$ = (*age = youth, income = medium, student = yes, credit rating = fair*).

17. a) Explain the various methods for detecting outliers.

b) Discuss the mining of text data mining.

18. a) Discuss in detail of applications of data mining for biomedical and DNA data analysis and telecommunication industry.

b) Discuss the social impacts of data mining systems.

****