

2020

COMPUTER SCIENCE — HONOURS

Paper : DSE-A-2

(Data Mining and Its Application)

Full Marks : 50

The figures in the margin indicate full marks.

*Candidates are required to give their answers in their own words
as far as practicable.*

Answer **question no. 1** and **any four** questions from the rest.

- | | | |
|-----------|--|-------|
| 1. | Answer any five questions : | 2×5 |
| | (a) What do you mean by Clustering? | |
| | (b) Explain different data mining tasks. | |
| | (c) Define : Regression. | |
| | (d) What do you mean by outlier detection? | |
| | (e) How can you check the efficiency of a classifier model? | |
| | (f) What do you mean by predictive modeling? | |
| | (g) Differentiate between OLTP and OLAP. | |
| | (h) What is a ROC graph? | |
| 2. | (a) Explain the various stages in knowledge discovery process. | |
| | (b) Describe Classification as a two-step process. | 5+5 |
| 3. | (a) Why is Data Cleaning required? | |
| | (b) What is a data cube? | |
| | (c) Explain the concept of Principal Component Analysis. | 3+2+5 |
| 4. | (a) Discuss the need of human intervention in the data mining process. | |
| | (b) How concept hierarchies are used for data discretization? | |
| | (c) What do you mean by Data scrubbing tools? | 4+4+2 |
| 5. | (a) Distinguish between Unsupervised and Supervised Learning. | |
| | (b) Briefly outline the major steps of decision tree classification. | 5+5 |

Please Turn Over

6. (a) What is fact table?
(b) Discuss about K-Means algorithm.
(c) Write on different applications of K-means algorithm. 2+5+3
7. (a) What is the usability of data compression?
(b) What is Neural Network? Explain activation function.
(c) What are the different methods of feature selection? 4+4+2
8. (a) What are the four major features of Data Warehouse?
(b) Differentiate between star schema and snowflake schema.
(c) Explain over fitting. 4+4+2
-