

S.No. 32

17PCS08

(For the candidates admitted from 2017-2018 onwards)

M.Sc. DEGREE EXAMINATION, APRIL/MAY 2018.

Second Semester

Computer Science

DATA MINING TECHNIQUES

Time : Three hours

Maximum : 75 marks

PART A — (5 × 5 = 25 marks)

Answer ALL questions.

1. (a) Explain the need for human direction of data mining.

Or

- (b) Explain the graphical methods for identifying outliers.

2. (a) Explain the validation of principal components.

Or

- (b) Write a brief note on factor rotation.

3. (a) Write a short note on distance function.

Or

(b) Explain the concept of decision trees.

4. (a) Explain complete-linkage clustering.

Or

(b) Explain self organizing maps.

5. (a) Explain the terms — support, confidence, frequent itemset, and the Apriori property.

Or

(b) Explain the business understanding phase of the direct-mail marketing problem.

PART B — (5 × 10 = 50 marks)

Answer ALL questions.

6. (a) Discuss about the common data mining tasks.

Or

(b) Explain (i) Data cleaning (ii) Measures of centre and spread.

7. (a) Explain principal components analysis. How is this applied to the Houses dataset?

Or

(b) Explain user-defined composites with an example.

8. (a) Explain the k-nearest neighbor algorithm in detail.

Or

(b) Discuss about classification and regression trees.

9. (a) Explain Kohonen networks with an example.

Or

(b) Explain pseudo-F statistic with an example.

10. (a) Discuss about affinity analysis and market basket analysis.

Or

(b) Explain the data preparation phase of the direct-mail marketing problem.