C128572(022)

B. Tech. (Hon's) (Fifth Semester) Examination

Nov.-Dec. 2023

(AICTE Scheme)

(Data Science Engg. Branch)

INTELLIGENT DATA ANALYSIS

Time Allowed : Three hours

Maximum Marks : 100

Minimum Pass Marks : 35

Note : *Attempt all questions. Part (a) from each question is compulsory and answers any* ***two*** *of the remaining (b), (c) and (d).*

Unit-I

1. (a) What is Data Mining? Explain in brief.   4

   (b) What do you mean by data? Explain various types of data used in machine learning.   8

(c) Why central tendency is an important measure in data analysis? Explain all the measure under central tendency calculation.   8

(d) What is data visualization? Explain different techniques used for data visualization.   8

Unit-II

2. (a) What do you mean by pre-processing of data. Explain in brief.   4

(b) How to measure Data Similarity and Dissimilarity. Explain different tools to measure similarity and dissimilarity.   8

(c) What is data transformation? Explain different techniques of data transformation.   8

(d) How artificial neural network (ANN) is different from biological neural network? List different types of ANN model.   8

Unit-III

3. (a) Briefly explain association analysis in data mining.   4

(b) Explain FP growth algorithm in brief. 8

(c) Consider following data sample. Execute k-nearest neighbour algorithm to classify data sample. Classify a new entry where brightness is 20 and saturation is 35. 8

S. No.	1	2	3	4	5	6	7
Brightness	40	50	60	10	70	60	25
Students	20	50	90	25	70	10	80
Class	Red	Blue	Blue	Red	Blue	Red	Blue

(d) Explain Maximal Frequent Itemset and Closed Frequent Itemset. 8

Unit-IV

4. (a) What is cluster? Explain different technique used to make cluster. 4

(b) Consider following data and execute DBSCAN algorithm. Assume = 3.5 and MinPts = 3. Find core, boundary and noise point. 8

Feature	D1	D2	D3	D4	D5	D6	D7	D8
X	5	8	3	4	3	6	6	5
Y	7	4	3	4	7	7	1	5

(c) Discuss issue associated with k-means algorithms. Explain different approach to improve the performance of k-means algorithm. 8

(d) Differentiate between agglomerative and Divisive Hierarchical clustering algorithm. 8

Unit-V

5. (a) What do you mean by anomalies data mining? Describe different types of anomalies. 4

(b) Discuss different types of anomalies detection techniques. 8

(c) Explain classification-based anomalies detection technique with suitable example. 8

(d) What are the different challenges in anomalies detection? Explain in brief. 8