



SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - I
FALL SEMESTER 2026-2027

Programme Name & Branch : B.Tech (BDS)
Course Code and Course Name : BCSE334L Predictive Analytics
Faculty Name(s) : Prof. J Anuradha, Prof. Anbarasi M
Class Number(s) : VL2026270102454, VL2026270102453
Date of Examination : 10.08.2026
Exam Duration : 90 minutes

Maximum Marks:

General instruction(s):

- Answer All Questions
 - M - Max mark; CO - Course Outcome; BL - Blooms Taxonomy Level (1 - Remember, 2 - Understand, 3 - Apply, 4 - Analyse, 5 - Evaluate, 6 - Create)
 - Course Outcomes: (Type the CO statements covered in this question paper. Use number as per the syllabus copy)
1. Understand the importance of predictive analytics and processing of data for analysis.
 2. Describe different types of predictive models.
 3. Apply regression and classification model on applications for decision making and evaluate the performance.

Q. No	Question	M	C
1.	Identify the type of data analytics and justify. i. A retail store manager reviews a weekly sales report generated by the POS system. The report shows that umbrella sales increased by 45% during the last week. ii. The HR department generates an annual dashboard displaying employee turnover rates across departments. The visual chart highlights that the engineering department lost 12% of its staff. iii. A streaming service displays a summary page to a user showing their total watch time. The page states they spent 120 hours watching comedy movies last month. iv. A school principal opens a spreadsheet containing the final exam grades of all students. The sheet calculates that the average score for the math exam was 78%. v. An e-commerce website compiles a list of the top ten most visited product pages. The marketing team uses this static list to see past website traffic trends.	10	1
2.	a) A small business records the daily revenue (in dollars) over a week, yielding the following 7 data points: (10, 14, 18, 25, 50, 120, 400). Show that this dataset is highly positively skewed. Apply Logarithmic transformation technique to reduce the skewness. b) For the given data perform discretization (bin width = 3), Max-Min normalization and handling missing value. X: 34, 25, 36, 42, --, 66, 83, 23, 65, 92	10	
3.	i) Identify the type of outliers for the given scenario and justify. a. A credit card fraud detection system notices a single transaction of 15,000 dollars from an account that normally spends between 20 and 50 dollars per transaction.	10	2



SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - I
FALL SEMESTER 2026-2027

	b. A smart home thermostat records a temperature reading of 32 degrees Celsius at 2:00 AM during a peak winter night in a region where winter nights always drop below 5 degrees Celsius. c. An ECG monitor records heart rate values that are all individually within the normal range of 60 to 100 beats per minute. However, the values stay exactly at 80 beats per minute for an extended period of 2 hours without any of the minor fluctuations typically seen in human heart rates. ii) An online retailer tracks customer behavior using two metrics: Number of Items Purchased (x) and Total Checkout Time in seconds (y). Normal shoppers take time to review items, while automated bot accounts checkout instantly with massive bulk orders. Detect the outlier if any using KNN with k=3. User1: (2,120), user2: (3, 130), User3: (1,110), User4: (2,115), User5: (40,5)																							
4.	a) Differentiate between Model based and Data Driven Models with suitable example. b) Brief on Propensity, cluster and collaborative predictive models with an example for each models.	10	3																					
5.	A streaming platform uses collaborative filtering to recommend movies to its users. Ratings are given on a scale of 1 to 5. A blank cell indicates that the user has not yet watched or rated that movie. <table border="1"><thead><tr><th>User</th><th>Inception</th><th>Matrix</th><th>Interstellar</th></tr></thead><tbody><tr><td>U1</td><td>5</td><td>4</td><td>4</td></tr><tr><td>U2</td><td>?</td><td>2</td><td>2</td></tr><tr><td>U3</td><td>4</td><td>5</td><td>4</td></tr><tr><td>U4</td><td>2</td><td>3</td><td>?</td></tr></tbody></table> i. Calculate the adjusted rating matrix. ii. Calculate item similarity matrix. iii. Predict the missing Rating.	User	Inception	Matrix	Interstellar	U1	5	4	4	U2	?	2	2	U3	4	5	4	U4	2	3	?	10	3	3
User	Inception	Matrix	Interstellar																					
U1	5	4	4																					
U2	?	2	2																					
U3	4	5	4																					
U4	2	3	?																					
