

BIRLA INSTITUTE OF TECHNOLOGY, MESRA, RANCHI
(END SEMESTER EXAMINATION)

CLASS: BTECH
BRANCH: AI & ML

SEMESTER : IV/ADD
SESSION : SP/2025

SUBJECT: AI203 MATHEMATICS FOR DATA SCIENCE

TIME: 3 Hours

FULL MARKS: 50

INSTRUCTIONS:

1. The question paper contains 5 questions each of 10 marks and total 50 marks.
 2. Attempt all questions.
 3. The missing data, if any, may be assumed suitably.
 4. Before attempting the question paper, be sure that you have got the correct question paper.
 5. Tables/Data hand book/Graph paper etc. to be supplied to the candidates in the examination hall.
-

- | | | CO | BL | |
|--------|--|---------|----|---|
| Q.1(a) | In a group of five friends – Rohan, Raj, Seema, Rahul, and Meera – it was customary to share lunch together each day for a period of 7 days. Each member contributed ₹50 per day towards the shared lunch. However, two irregularities occurred during the week: <ul style="list-style-type: none">● Meera developed a food allergy and discontinued lunch after Day 2.● Raj went on a family tour and missed lunch on Days 3, 4, and 5. (i) Create a graph to represent the situation stating number of vertex and edge in the graph.
(ii) Is your graph weighted graph? If yes, say what weights represents in your graph.
(iii) Calculate the total amount of money spent on lunch over the 7-day period, stating the individual contribution of each. | [2+1+2] | 1 | 5 |
| Q.1(b) | Find Eigen values and Eigen vectors of given matrix $A = \begin{bmatrix} -5 & 2 \\ -7 & 4 \end{bmatrix}$ | [2+3] | 3 | 3 |
| Q.2(a) | A student noticed an unusual pattern in the behavior of a professor at their university. Through extended observation and informal questioning, the student discovered the following behavioral tendencies: <ul style="list-style-type: none">● On hot days, the professor is more likely to bring a drink, doing so with a probability of 0.7, and not bringing a drink with a probability of 0.3.● On cold days, the pattern is reversed – the professor brings a drink only 20% of the time, and comes without a drink 80% of the time. The student also gathered information about how the weather tends to change from one day to the next in that region: <ul style="list-style-type: none">● If one day is hot, the following day is also likely to be hot with a 75% chance, but there is a 25% chance that it becomes cold.● If one day is cold, then it tends to remain cold the next day with a 70% chance, and there is a 30% chance that it becomes hot. At the time the student began recording, there was a 55% chance that the day was hot, and a 45% chance that it was cold.

(i) Clearly identify all the components of HMM model like hidden states, observed states, initial probabilities, size of emission and transition matrix.
(ii) Find emission and transition probabilities.
(iii) Draw HMM diagram to represent the problem. | [2+1+2] | 2 | 5 |
| (b) | Over the course of four days, the student recorded the professor's behavior as follows: No Drink, No Drink, No Drink, Drink(NNND). Now, several months later, the student wants to reconstruct the most likely sequence of weather conditions for those four days.
Give stepwise calculations to find most probable weather sequence using viterbi algorithm.
Give most probable sequence with the final probability of sequence. | [4+1] | 2 | 3 |

PTO

- Q.3(a) For given data find within class and between class scatter [3+2] 3 4
 $X1=(x1,x2)=\{(4,1),(2,4),(2,3),(3,6),(4,4)\}$
 $X2=(x1,x2)=\{(9,10),(6,8),(9,5),(8,7),(10,8)\}$
 Where $X1$ and $X2$ represents the data points corresponding to two classes present in data.
- Q.3(b) Compare PCA, SVD and LDA in terms of type of method, data type, optimization parameter, applications and outcome(show graphically wherever applicable). [0.5+1+2+0.5+1] 3 4
- Q.4(a) For given data [2+1.5+1.5] 4 3

x	4	5	7	10	11	13
$f(x)$	48	100	294	900	1210	2029

Use Newton's Divided Difference Interpolation Formula and

- (i) Show calculations stepwise.
 (ii) Find the interpolation polynomial.
 (iii) Estimate the values of the function at $f(8)$, $f(9)$, and $f(15)$.

- Q.4(b) Given below are the data relating to the sales of a product in a district. [2+3] 4 4

Year	2000	2001	2002	2003	2004	2005	2006
Prod. of Sugarcane	40	45	46	42	47	50	46

Fit a straight line trend by the method of least squares and tabulate the trend values. Take year 2003 as initial year.

- Q.5(a) Let's say that 80% of all business startups in the IT industry report that they generate a profit in their first year. If a sample of 10 new IT business startups is selected, Identify model parameters, distribution and find the probability that exactly seven will generate a profit in their first year. [1+1+2] 2 2
- Q.5(b) State true/false with 2-3 lines/graphical justification: [2+2+2] 5 4
 (i) $d(x^T Ax)/dx=2Ax$ where A is a symmetric matrix in $R^{n \times n}$ space and x is vector in R^n space.
 (ii) In Gradient Descent, a large learning rate can cause the algorithm to overshoot the minimum and fail to converge.
 (iii) Optimal shortest path in travelling salesman problem can be calculated using gradient decent.