



# PRESIDENCY UNIVERSITY

BENGALURU

Roll No.																			
----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

## Make-up Examinations - December 2025

Date: 29- 12- 2025

Time: 01:00pm - 04:00pm

<b>School:</b> SOCSE	<b>Program:</b> B. Tech		
<b>Course Code :</b> CSE3001	<b>Course Name :</b> Artificial Intelligence and Machine Learning		
<b>Semester:</b> MK	<b>Max Marks:</b> 100	<b>Weightage:</b> 50%	

CO - Levels	C01	C02	C03	C04	C05
<b>Marks</b>	<b>4</b>	<b>29</b>	<b>19</b>	<b>34</b>	<b>14</b>

### Instructions:

- (i) Read all questions carefully and answer accordingly.
- (ii) Do not write anything on the question paper other than roll number.

### Part A

Answer ALL the Questions. Each question carries 2marks.

10Q x 2M=20M

1	What is Artificial Intelligence? List some applications of AI.	2 Marks	L1	C01
2	What is the Turing Test and how is it used in the context of Artificial Intelligence?	2 Marks	L1	C01
3	Give the algorithm for K-means clustering.	2 Marks	L1	C04
4	Name the different supervised feature selection techniques	2 Marks	L1	C02
5	List the different evaluation metrics for regression model.	2 Marks	L1	C02
6	What is Support Vector Machine.	2 Marks	L1	C03
7	What is Ensemble Learning? Write one advantage and disadvantage of ensemble learning	2 Marks	L1	C03
8	Define Clustering. List out the different types of clustering.	2 Marks	L1	C04
9	What is forecasting in the context of machine learning?	2 Marks	L1	C05
10	List out the types of patterns occur in Time Series data?	2 Marks	L1	C05

## Part B

Answer the Questions

Total 80 Marks

<b>11.</b>	<b>a.</b>	Explain the concept of <b>Logistic Regression</b> and how it differs from Linear Regression. Why is the sigmoid function used in Logistic Regression?	<b>10 Marks</b>	<b>L2</b>	<b>C02</b>
------------	-----------	---	---------------------	-----------	------------

or

<b>12.</b>	<b>a.</b>	What is Feature selection technique in machine learning. Explain the different types of supervised feature selection.	<b>10 Marks</b>	<b>L2</b>	<b>C02</b>
------------	-----------	---	---------------------	-----------	------------

<b>13.</b>	<b>a.</b>	What is Time series data. Explain the types of patterns occur in Time Series data?	<b>10 Marks</b>	<b>L2</b>	<b>C05</b>
------------	-----------	--	---------------------	-----------	------------

or

<b>14.</b>	<b>a.</b>	Explain the importance of forecasting in machine learning and discuss the error metrics used in evaluating time series forecasting models.	<b>10 Marks</b>	<b>L2</b>	<b>C05</b>
------------	-----------	--	---------------------	-----------	------------

<b>15.</b>	<b>a.</b>	What is partitioned based clustering. Write the algorithm for K-means clustering and how do you choose the optimal value of K.	<b>10 Marks</b>	<b>L3</b>	<b>C04</b>
------------	-----------	--	---------------------	-----------	------------

Or

<b>16.</b>	<b>a.</b>	What is Hierarchical clustering. List out the differences between agglomerative and divisive clustering with dendrogram.	<b>10 Marks</b>	<b>L1</b>	<b>C04</b>
------------	-----------	--	---------------------	-----------	------------

<b>17.</b>	<b>a.</b>	<p>A dataset contains the following information about whether students will "Play Outside" based on two features: <b>Weather</b> and <b>Homework</b>.</p> <table border="1" style="margin-left: auto; margin-right: auto; border-collapse: collapse;"> <thead> <tr> <th style="padding: 5px;">Weather</th> <th style="padding: 5px;">Homework</th> <th style="padding: 5px;">Play Outside</th> </tr> </thead> <tbody> <tr> <td style="padding: 5px;">Sunny</td> <td style="padding: 5px;">Done</td> <td style="padding: 5px;">Yes</td> </tr> <tr> <td style="padding: 5px;">Sunny</td> <td style="padding: 5px;">Not Done</td> <td style="padding: 5px;">No</td> </tr> <tr> <td style="padding: 5px;">Cloudy</td> <td style="padding: 5px;">Done</td> <td style="padding: 5px;">Yes</td> </tr> <tr> <td style="padding: 5px;">Cloudy</td> <td style="padding: 5px;">Not Done</td> <td style="padding: 5px;">Yes</td> </tr> <tr> <td style="padding: 5px;">Rainy</td> <td style="padding: 5px;">Done</td> <td style="padding: 5px;">No</td> </tr> <tr> <td style="padding: 5px;">Rainy</td> <td style="padding: 5px;">Not Done</td> <td style="padding: 5px;">No</td> </tr> </tbody> </table> <p>(a) Calculate the <b>entropy</b> of the target variable "<b>Play Outside</b>".            (b) Calculate the <b>information gain</b> for both features, "<b>Weather</b>" and "<b>Homework</b>", and determine which feature is the most suitable root node for the decision tree.</p>	Weather	Homework	Play Outside	Sunny	Done	Yes	Sunny	Not Done	No	Cloudy	Done	Yes	Cloudy	Not Done	Yes	Rainy	Done	No	Rainy	Not Done	No	<b>15 Marks</b>	<b>L3</b>	<b>C02</b>
Weather	Homework	Play Outside																								
Sunny	Done	Yes																								
Sunny	Not Done	No																								
Cloudy	Done	Yes																								
Cloudy	Not Done	Yes																								
Rainy	Done	No																								
Rainy	Not Done	No																								

Or

18.

The following dataset contains weather conditions and whether a person plays tennis or not:

Day	Outlook	Temperature	Humidity	Wind	Play Tennis?
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Cool	Normal	Weak	Yes
14	Rain	Mild	High	String	No

a.

1. Calculate the Prior Probabilities for each class (Yes/No for "Play Tennis").

2. For a new instance:

- Outlook = Sunny, Temperature = Cool, Humidity = High, Wind = Strong  
Use the Naive Bayes Classifier to predict whether the person will Play Tennis.

3. Show all the steps, including:

- Likelihood probabilities for each attribute value.
- Posterior probabilities for each class.
- Final prediction based on the class with the highest posterior probability.

15  
Marks

L3 CO2

19.

Consider the following dataset with 4 points:

Point	X	Y
P1	2	3
P2	5	4
P3	3	7
P4	8	8

a.

Use the K-Means clustering algorithm to cluster the points into 2 clusters ( $k = 2$ ). The initial centroids are:

$$C1 = (2, 3)$$

$$C2 = (5, 4)$$

15  
Marks

L3

C03

Or

20.

Consider the following dataset, where we want to classify a new point (6, 6) using the K-Nearest Neighbors (KNN) algorithm with  $k=3$ .

Point	X	Y	Class
P1	2	3	A
P2	5	4	B
P3	9	6	A
P4	4	7	B
P5	8	1	A

a.

1. Calculate the Euclidean distance of the point (6, 6) to all the points in the dataset.

2. Identify the 3 nearest neighbours.

3. Determine the class of the new point using majority voting.

4. Show all calculations.

15  
Marks

L3

C03

21.	a.	Implement Ensemble learning using python code for Bagging and Boosting classifier using scikit-learn, and to evaluate its performance on a dataset. Assume our own dataset.	20 Marks	L3	CO4
-----	----	---	-------------	----	-----

**Or**

22.	a.	<p>Write a Python program to implement the K-Means clustering algorithm for a given dataset. Your solution should include the following:</p> <ol style="list-style-type: none"> <li>1. <b>Load and preprocess the dataset</b> (e.g., handle missing values, normalize features if necessary).</li> <li>2. <b>Implement the K-Means algorithm</b> using a library such as scikit-learn or manually if specified.</li> <li>3. <b>Print Cluster Centres:</b> Find the average income and age</li> <li>4. <b>Visualize the clusters</b> using a 2D or 3D plot for the given dataset.</li> <li>5. <b>Find the optimum number of clusters:</b> Using Elbow method.</li> </ol> <p>Assume our own dataset.</p>	20 Marks	L3	CO4
-----	----	--	-------------	----	-----

**\*\*\*\*\* BEST WISHES \*\*\*\*\***