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



 **PRESIDENCY UNIVERSITY**

  **Bengaluru**

|  |
| --- |
| **End - Term Examinations – JANUARY 2025** |
| **Date:** 03-01-2025 **Time:** 09:30 am – 12:30 pm |

|  |  |
| --- | --- |
| **School:** SOIS | **Program:** MCA |
| **Course Code :** CSA4009 | **Course Name :** Artificial Intelligence and Machine Learning |
| **Semester**: III | **Max Marks**: 100 | **Weightage**:50% |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **CO - Levels** | **CO1** | **CO2** | **CO3** | **CO4** |
| **Marks** | **24** | **24** | **28** | **24** |

**Instructions:**

1. *Read all questions carefully and answer accordingly.*
2. *Do not write anything on the question paper other than roll number.*

**Part A**

|  |
| --- |
| **Answer ALL the Questions. 10 x 2 Marks=20 Marks** |
| **1** | Define Artificial Intelligence with example. | 2 Marks | L1 | CO1 |
| **2** | List out the Technological drivers of modern AI. | 2 Marks | L1 | CO1 |
| **3** | Define Machine Learning with example. | 2 Marks | L1 | CO2 |
| **4** | List out Difference between Supervised and Unsupervised Machine Learning. | 2 Marks | L1 | CO2 |
| **5** | List out Difference between Classification and Regression. | 2 Marks | L1 | CO3 |
| **6** | Define Information gain, Entropy and write the appropriate formula. | 2 Marks | L1 | CO3 |
| **7** | List out the different types of loss function. | 2 Marks | L1 | CO3 |
| **8** | What is Elbow method. | 2 Marks | L1 | CO3 |
| **9** | What is a dendrogram in hierarchical clustering. | 2 Marks | L1 | CO4 |
| **10** | Write the accuracy measures used in forecasting. | 2 Marks | L1 | CO4 |

**Part B**

|  |
| --- |
| **Answer the Questions Total 80 Marks.** |
| **11.** | **a.** | List and explain future impact of AI in different sector. | **10 Marks** | **L2** | **CO1** |
|  | **b.** | Explain the use of cross-industry standard process for data mining in data processing. | **10 marks** | **L2** | **CO1** |
| **Or** |
| **12.** | **a.** | Explain rise and fall of the expert system with example. | **10 Marks** | **L2** | **CO1** |
|  | **b.** | Discuss different types of data and V’s in Big data with suitable example. | **10 marks** | **L2** | **CO1** |
|  |  |  |  |  |  |
| **13.** | **a.** | Explain different types of machine learning Techniques with suitable example. | **10 Marks** | **L2** | **CO2** |
|  | **b.** | Explain Feature Selection and Feature Engineering Techniques. | **10 Marks** | **L2** | **CO2** |
| **Or** |
| **14.** | **a.** | Briefly explain data preprocessing in machine learning. | **10 Marks** | **L2** | **CO2** |
|  | **b.** | The Stanford University Business School is assessing whether this year's applicants have GMAT verbal scores significantly higher than the target score of 210. The standard deviation of GMAT verbal scores for all applicants is 8.5. A sample of 42 applicants had an average verbal GMAT score of 212.79. At a significance level of α = 0.05, apply z-test to determine if this year's mean score is statistically greater than the target mean of 210? | **10 Marks** | **L3** | **CO2** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **15.** | **a.** | What is Classification, Explain Decision Tree Algorithm in Details. | **08 Marks** | **L2** | **CO3** |
|  | **b.** | Define confusion matrix and for the given confusion matrix below, calculate the following metrics:1.Calculate the accuracy of the model.2.Calculate the precision for the "Spam" class.3. Calculate the recall for the "Spam" class.4.Calculate the F1-score for the "Spam" class.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Actual/Predicted | Predicted Spam |

|  |
| --- |
|  |

|  |
| --- |
| Predicted Not Spam |

 |
| Actual Spam | 40 | 10 |
| Actual Not Spam | 4 | 45 |

 | **12 Marks** | **L3** | **CO3** |
| **Or** |
| **16.** | **a.** | Explain Bagging and Boosting Algorithm along with suitable Example. | **08 Marks** | **L2** | **CO3** |
|  | **b.** | Given the dataset shown in the table with Brightness and Saturation values, the color class is either Red or Blue. Using the KNN algorithm, predict the color (Red or Blue) for the missing class by assuming K =3. Apply Euclidean distance formula. | **12 Marks** | **L3** | **CO3** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **17.** | **a.** | Define Time series data, Explain components of time series data with example. | **08 Marks** | **L2** | **CO4** |
|  | **b.** | Apply K means Clustering algorithm to divide the following data into two clusters (K=2). C1=(1,1) ,C2=(5,5)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  **X1** | **1** | **2** | **2** | **3** | **4** | **5** |
|  **Y1** | **1** | **1** | **3** | **2** | **3** | **5** |

 | **12 marks** | **L3** | **CO4** |
| **Or** |
| **18.** | **a.** | Explain Briefly:1.MA2.AR3.ARMA4.ARIMA |  **08 Marks** | **L2** | **CO4** |
|  | **b.** | Consider the following data points with pairwise distances:

| **Points** | **A** | **B** | **C** | **D** | **E** |
| --- | --- | --- | --- | --- | --- |
| A | 0 | 2 | 6 | 10 | 9 |
| B | 2 | 0 | 5 | 9 | 8 |
| C | 6 | 5 | 0 | 4 | 5 |
| D | 10 | 9 | 4 | 0 | 3 |
| E | 9 | 8 | 5 | 3 | 0 |

1. Perform **hierarchical clustering** using **single linkage**.
2. Draw the **dendrogram** to represent the clusters.
 | **12 marks** | **L3** | **CO4** |





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