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



 **PRESIDENCY UNIVERSITY**

  **Bengaluru**

|  |
| --- |
| **End - Term Examinations – JANUARY 2025** |
| **Date:** 16 – 01- 2025 **Time:** 09:30 am – 12:30 pm |

|  |  |
| --- | --- |
| **School:** SOE | **Program:** B. Tech (ECE) |
| **Course Code :** ECE3020 | **Course Name :** Computational Intelligence and Machine Learning |
| **Semester**: V | **Max Marks**:100 | **Weightage**: 50% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **CO - Levels** | **CO1** | **CO2** | **CO3** | **CO4** | **CO5** |
| **Marks** | **32** | **41** | **27** | **NA** | **NA** |

**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. Each question carries 2marks. 10Q x 2M=20M** |
| **1** | Write short note dimensionality reduction? | **2 Marks** | **L1** | **CO1** |
| **2** | Differentiate between hierarchical and non-hierarchical clustering.  | **2 Marks** | **L1** | **CO1** |
| **3** | Explain the role of a labeled datasets in supervised machine learning. | **2 Marks** | **L1** | **CO1** |
| **4** | What is the role of personal and global best in ACO? | **2 Marks** | **L1** | **CO1** |
| **5** | Write short note on confusion matrix. | **2 Marks** | **L1** | **CO3** |
| **6** | What is an R squared error, and why is it important in regression algorithms? | **2 Marks** | **L2** | **CO2** |
| **7** | What is the indication made by Recall and Precision in confusion matrix? | **2 Marks** | **L1** | **CO2** |
| **8** | Justify the need of testing a model in machine learning. | **2 Marks** | **L1** | **CO2** |
| **9** | What is the role of activation function in neural network? | **2 Marks** | **L1** | **CO1** |
| **10** | Differentiate between Ant Colony Optimization (ACO) and Particle Swarm Optimization (PSO) machine learning algorithms (Any four points). | **2 Marks** | **L1** | **CO1** |

**Part B**

|  |
| --- |
| **Answer the Questions Total 80 Marks** |
| **11.** | **a.** | Write a brief note on validation and testing of a machine learning algorithm with example of it. | **10****Marks** | **L2** | **CO2** |
| **Or** |
| **12.** | **a.** | A researcher wants to predict the price of a house (in $1000s) based on its size (in square feet). The following data is collected:

| **House Size (X) (sq. ft.)** | **Price (Y) ($1000s)** |
| --- | --- |
| 1000 | 150 |
| 1500 | 200 |
| 2000 | 250 |
| 2500 | 300 |
| 3000 | 350 |

Compute the regression line coefficients and Predict the price of a house with a size of 1800 square feet. Comment on the relationship between test marks and study hours. | **10****Marks** | **L3** | **CO2** |
|  |  |  |  |  |  |
| **13.** | **a.** | How would you design a machine learning model for predicting a student’s performance in a final exam using multiple input features like study hours, attendance, and previous test scores? Discuss the steps, challenges, and ethical considerations. | **10****Marks** | **L3** | **CO3** |
| **Or** |
| **14.** | **a.** | The following dataset contains the features of cars based on engine size and fuel efficiency:

| **Engine Size (X1) (cc)** | **Fuel Efficiency (X2) (km/l)** | **Car Type (Y)** |
| --- | --- | --- |
| 1500 | 20 | Sedan |
| 2000 | 15 | SUV |
| 1800 | 18 | Sedan |
| 2500 | 12 | SUV |
| 1600 | 19 | Sedan |

Use the KNN algorithm with K=3 to predict the car type of a car with an engine size of 1900 cc and fuel efficiency of 16 km/l. Show all calculations clearly, including the distance computation and the decision-making process. | **10****Marks** | **L3** | **CO3** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **15.** | **a.** | Supervised learning models can often fail to generalize to unseen data. What strategies would you adopt to improve generalization, and how would you test whether your strategies are successful?  | **10****Marks** | **L2** | **CO2** |
| **Or** |
| **16.** | **a.** | Imagine you have trained a highly accurate supervised learning model for customer churn prediction. However, the model is complex and not interpretable. How would you balance accuracy with interpretability, and why might interpretability matter in this case? | **10****Marks** | **L3** | **CO2** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **17.** | **a.** | Explain why the simple perceptron is referred to as a linear classifier. Use the Ex-OR problem to illustrate your justification, including an appropriate diagram. | **15****Marks** | **L2** | **CO2** |
| **Or** |
| **18.** | **a.** | An electronics company uses a machine learning model to predict whether a manufacturing component is Defective or Non-Defective based on sensor readings during the production process. The actual labels (True Class) and the predicted labels (Predicted Class) for 10 components are as follows:

| **True Class** | **Predicted Class** |
| --- | --- |
| Defective | Defective |
| Defective | Non-Defective |
| Non-Defective | Non-Defective |
| Defective | Defective |
| Non-Defective | Non-Defective |
| Defective | Defective |
| Non-Defective | Non-Defective |
| Non-Defective | Defective |
| Defective | Non-Defective |
| Non-Defective | Non-Defective |

1. Construct the confusion matrix for the above data based on the predicted and actual labels.
2. Calculate the following performance metrics using the confusion matrix:
	* Accuracy
	* Precision
	* Recall
	* F1-Score
3. Interpret the results and evaluate how well the classification model identifies defective components.
 | **15****Marks** | **L3** | **CO2** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **19.** | **a.** | A dataset of customer shopping patterns in retail stores across different regions is provided for training a model to predict the most profitable areas for opening new retail stores. The goal is to identify locations that will attract more customers without negatively impacting the sales of existing stores. In your opinion, which supervised machine learning model would provide the best possible prediction? Justify your choice with a brief explanation. | **15****Marks** | **L2** | **CO3** |
| **Or** |
| **20.** | **a.** | Write a brief note on hierarchical clustering algorithm with suitable real life example. | **15****Marks** | **L3** | **CO3** |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **21.** | **a.** | A company wants to segment its customer base for targeted marketing using the K-means clustering algorithm. They have collected data on customer age and annual income (in thousands) as shown below:

| **Customer** | **Age (years)** | **Annual Income (k$)** |
| --- | --- | --- |
| 1 | 25 | 30 |
| 2 | 35 | 60 |
| 3 | 45 | 90 |
| 4 | 50 | 100 |
| 5 | 30 | 40 |
| 6 | 40 | 80 |
| 7 | 60 | 120 |
| 8 | 55 | 110 |
| 9 | 32 | 45 |
| 10 | 28 | 35 |

**Apply the K-means clustering algorithm** on this dataset to cluster the individuals into **3 groups** (K=3). **Evaluate the final clusters** and provide an interpretation of the results. Discuss which individuals are grouped together and why. | **20****Marks** | **L3** | **CO1** |
| **Or** |
| **22.** | **a.** | A data scientist is assigned to develop a machine learning model to predict employee attrition based on a dataset containing employee demographics, performance metrics, and job satisfaction levels. Write a brief overview of the process for selecting the appropriate model, along with the strategies for training, testing, and evaluating the model. | **20****Marks** | **L3** | **CO1** |

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