PRESIDENCY UNIVERSITY
BENGALURU

SCHOOL OF ENGINEERING
END TERM EXAMINATION - DEC 2022

Semester : Semester V - 2020

Course Code : ECE3020

Course Name : Sem V - ECE3020 - Computational Intelligence and Machine Learning

Program : B.Tech. Electronics and Communication Engineering

Date : 16-JAN-2023

Time : 9.30AM - 12.30PM

Max Marks : 100

Weightage : 50%

Instructions:
(i) Read all questions carefully and answer accordingly.
(ii) Question paper consists of 3 parts.
(iii) Scientific and non-programmable calculator are permitted.

PART A

ANSWER ALL THE FIVE QUESTIONSQ   5 X 2 = 10M

1. Machine learning is classified into supervised, unsupervised and reinforcement learning the k-means algorithm is a
   A) Supervised learning algorithm
   B) Unsupervised learning algorithm
   C) Semi-supervised learning algorithm
   D)Weakly supervised learning algorithm

   (CO1) [Knowledge]

2. Machine learning can have many methods of learning ._____is the machine learning algorithm that can be used with unlabeled data.
   A) Regression algorithms
   B) Clustering algorithms
   C) Instance based algorithms
   D) All of the above

   (CO2) [Knowledge]

3. Of the following examples, which would you address using an supervised learning algorithm?
   A) Given email labeled as spam or not spam , learn a spam filter
   B) Given a set of new articles found on web , group them into set of articles about the same story.
   C) Given a database of customer data , automatically discover market segments and group customers into different market segments.
   D) Find the patterns in market basket analysis

   (CO3) [Knowledge]

4. The definition of classifying is categorizing something or someone into a certain group or system based on certain characteristics. The average squared difference between classifier predicted output and actual output.
   A) mean squared error
   B) root mean squared error
   C) mean absolute error
   D) mean relative error

   (CO1) [Knowledge]

5. For a classification task, instead of random weight initializations in a neural network, we set all the weights to zero. Which of the following statements is true?
   A) There will not be any problem and the neural network will train properly
   B) The neural network will train but all the neurons will end up recognizing the same thing
   C) The neural network will not train as there is no net gradient change
   D) None of these

   (CO2) [Knowledge]

PART B

ANSWER ALL THE TWO QUESTIONS   2 X 15 = 30M

6. Probabilistic discriminative models are majorly divided into two types a) generative model b) discriminative models. Consider a data set C={(0,1),(1,1),(1,0),(0,1),(1,0),(1,1),(0,1)},find the relation between generative and discriminant models for the given data set.

   (CO3) [Comprehension]

7. Ant colony optimization (ACO) is an optimization algorithm which employs the probabilistic technique and is used for solving computational problems and finding the optimal path with the help of graphs. An ACO is a population-based Algorithm that can be used to find approximate solutions to difficult optimization problems. In ACO, a set of software agents called artificial ants search for good solutions to a given optimization problem.

   Consider the problem of finding the optimum order in which the numbers from 1 to 9 are arranged so that the cost of order is maximum. Assume that six ants have the cost functions (C1, C2, C3, C4, C5, and C6). Consider the following are the orders selected by the six ants along with the corresponding Cost as given below.

ANT Number	ORDER									COST
ANT1	3	2	4	1	7	5	8	6	9	C1
ANT2	5	8	9	2	7	3	6	4	1	C2
ANT3	5	6	8	2	3	9	7	1	4	C3
ANT4	8	4	9	6	3	1	2	7	5	C4
ANT5	5	6	7	3	4	1	9	2	8	C5
ANT6	4	6	7	8	1	2	5	9	3	C6

   (CO2,CO1) [Comprehension]

ANSWER ALL THE THREE QUESTIONS   **3 X 20 = 60M**

8. Logistic Regression is one of the most popular linear classification models that perform well for binary classification but falls short in the case of multiple classification problems with well-separated classes. While Linear Discriminant Analysis (LDA) handles these quite efficiently. LDA can also be used in data preprocessing to reduce the number of features just as Principle component analysis which reduces the computing cost significantly.

   Consider two data sets as mentioned below, what will be suitable weight vector which will be used to perform classification as well as dimensionality reduction.

   $$X1= \{(4,1), (2,4), (2,3),(3,6),(4,4)\}$$
   $$X2= \{(9,10),(6,8),(9,5),(8,7),(10,8)\}$$

                  (CO3,CO2) [Application]

9. Particle swarm optimization (PSO) is a population-based stochastic optimization algorithm motivated by intelligent collective behavior of some animals such as flocks of birds or schools of fish. Since presented in 1995, it has experienced a multitude of enhancements. As researchers have learned about the technique, they derived new versions aiming to different demands, developed new applications in a host of areas, published theoretical studies of the effects of the various parameters and proposed many variants of the algorithm.

   Consider a scenario that, five particles (Say P1, P2, P3, P4 and P5) are moving around the solution space (Say Z). Each particle moves around the solution space randomly but at the same time attracted by other poles, its past best position (solution) and the best position (solution) of the whole swarm (collection of particles). These poles modify the velocity vector of the particles at each iteration.

   How these swarms (Say P1, P2, P3, P4 and P5) modify their velocity vectors in the form of their position and reaches their destination (Say Z). Form an algorithm with suitable equations

                  (CO2,CO3) [Application]

10. K-means algorithm is an iterative algorithm that tries to partition the dataset into K-pre-defined distinct non-overlapping subgroups (clusters) where each data point belongs to only one group. It tries to make the intra-cluster data points as similar as possible while also keeping the clusters as different as possible. It assigns data points to a cluster such that the sum of the squared distance between the data points and the cluster's centroid is at the minimum

    Consider data sets X and Y as given below. Show the steps of calculation for data points until final clustering is done where no data points are changing clusters.

X	1	2	2	3	4	5
Y	1	1	3	2	3	5

               (CO2,CO3) [Application]

★★★★★