Total number of printed pages: **UG/07/UECE714B**

2022

Machine Learning

Full Marks : 100

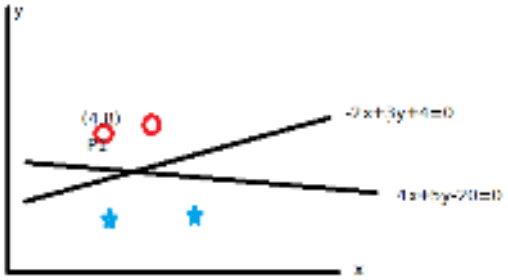
Time : Three hours

The figures in the margin indicate full marks for the questions.

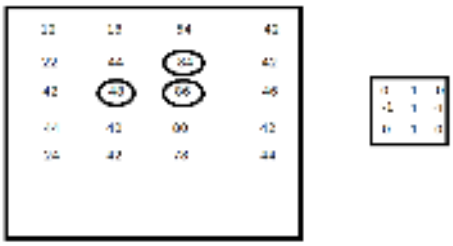
Answer Q-1 ***any four*** *questions from the rest.*

1.	A)	Select the correct alternatives	1x10=10
	a)	What is unsupervised learning i) The features of a group are not explicitly started ii) The number of a group may not be known iii) Both i) and ii) iv) None of the above	1
	b)	If the output of a supervised model is a categorical output then it is a i) Regression ii) Classifier iii) Prediction iv) None of the above	1
	c)	When we have two or more input variables and an output variable then the cost function is represented by a i) Convex curve ii) Concave curve iii) Contour plot iv) None of the above	1
	d)	Which of the following is a measure of the spread of data? i) Mean ii) Variance iii) Covariance	

		iv) Median	
	e)	Which of the following is a widely used and effective machine learning algorithm based on the idea of bagging? i) Random forest ii) Regression iii) Stochastic regression iv) K-means clustering	
	f)	Which of the following is a widely used and effective machine learning algorithm based on the idea of bagging? I) Factor analysis ii) Decision trees are robust to outliers iii) Decision trees are prone to be overfit iv) None of the above	
	g)	**Weighted sums in ANNs are referred to as _____?** I) Activation ii) Support iii) Link iv) Output	
	h)	**Weighted sums in ANNs are referred to as _____?** I) Recurrent Neural Networking ii) Convolutional Neural Networking iii) Auto encoder iv) Deconvolutional Neural Network	2
	i)	**Which of the following boolean operation can be represented by a single perceptron layer?** A) X1 **AND** X2 B) X1 **OR** X2 C) X1 **NOR** X2 D) X1 **XOR** X2 **i) A only** **ii) A, B only** **iii) A, B, and C only** **iv) Only D.**	

	j)	The ***sigmoid*** function defined as, (i) $\frac{1}{\exp(t)+\exp(-t)}$ (ii) t exp (-t) (iii) $\frac{1}{1+\exp(t)}$ (iv) $\frac{1}{1+\exp(-t)}$	
B		Answer all question briefly	2x5
	i)	A point (1,2) undergoes a nonlinear mapping through 2 radial basis functions $\varphi(X) = e^{(-r^2/2)}$ where r is the distance from the cluster center. Assume the center points of the clusters are (0,0) and (4,6).	
	ii)	Consider a single perceptron with sign activation function. The perceptron is represented by weight vector $[0.4\ -0.3\ 0.1]^t$ and a bias $\theta = 0$. If the input vector to the perceptron is X = [0.2 0.6 0.5] then the output of the perceptron is : 	
	iii)	There are two clusters C1 and C2 having C1=[(1,2), (2,4), (4,4),(1,6)] and C2=[(4,6),(8,8),(10,10),(10,8)]. Find the cluster centers of C1 and C2. State in which class an unknown point (5,5) will belong?	
	iv)		

		Find the margins of the point P1 for the above two decision boundaries in the figure.	
	v)	Match the following network and their applications A. Convolutional Neural Network — I) Speech recognition B. Auto encoder — ii) Automatic feature extraction C. Decision Tree — iii) Bagging D. Random forest — iv) Categorical data classification	
2.	a)	Define the following terminologies i) Activation function ii) Perceptron	3x2
	b)	A Multilayer perceptron is classifying handwritten numeral images of 1-10 from MNIST dataset. 100 features was extracted from each image. A multilayer neural network having with two hidden layers of number of neurons [120,30] is used. i) Draw the network with its symbolic diagram. ii) Calculate number of parameters to be optimized including bias in all layers except the input layer. Iii) Suggest a suitable optimization method which can be used for this classification. Justify your reason. iv) suggest a cost function to be used to optimize during training. Write its formula.	2x4
	c)	Discuss how radial basis function resolves XOR problem.	6
3	a)	Compare Multiplayer perceptron and(MLP) Convolutional neural network (CNN) in classification problem. Discuss advantages, disadvantages, limitations of both these classifiers.	6
	b)	State the names of different layers in a CNN and state their functions with examples.	8

	c)	The following image A was convolved with a kernal B. Find out the results for the rounded points. 	6
4.		Table 9.1 (Continued) S. No. | Age | Income | Student | Credit | Buy 4 | >40 | Medium | No | Fair | Yes 5 | >40 | Low | Yes | Fair | Yes 6 | >40 | Low | Yes | Excellent | No 7 | 31–40 | Low | Yes | Excellent | Yes 8 | <30 | Medium | No | Fair | No 9 | <30 | Low | Yes | Fair | Yes 10 | >40 | Medium | Yes | Fair | Yes 11 | <30 | Medium | Yes | Excellent | Yes 12 | 31–40 | Medium | No | Excellent | Yes 13 | 31–40 | High | Yes | Fair | Yes 14 | >40 | Medium | No | Excellent | No a) In respect of the decision tree problem, compute entropy of the dataset. b) State how many attributes are there in the table. b) Find out the gain for all the attributes. c) Calculate split info S(D), and Gain ratio G (S, A="Income"), G (S, A="Credit") Hints: $\text{Split Info}_A(D) = -\sum_{j=1}^{v} \frac{\lvert D_j \rvert}{\lvert D \rvert} \times \log_2\left(\frac{\lvert D_j \rvert}{\lvert D \rvert}\right)$ $\text{Gain}(S, A) = \text{Entropy}(S) - \sum_{v \in \text{values}(A)} \frac{\lvert S_v \rvert}{\lvert S \rvert} \text{Entropy}(S_v)$ d) Based on the Gain, which attribute should be picked up as the next attribute.	2+1+ 4x4 1
5.	a)	Cluster the following eight points (with (x, y) representing locations) into three clusters: A1(2, 10), A2(2, 5), A3(8, 4), A4(5, 8)	10

Initial cluster centres are: A1(2, 10), A4(5, 8) and A7(1, 2).

The distance function between two points a = (x1, y1) and b = (x2, y2) is defined as-

P(a, b) = |x2 – x1| + |y2 – y1|

Use K-Means Algorithm to find the three cluster centers after the second iteration.

Fill the following table for 2 iterations using K-Means clustering

Given Points	Distance from center (2, 10) of Cluster-01	Distance from center (5, 8) of Cluster-02	Distance from center (1, 2) of Cluster-03	Point belongs to Cluster
A1(2, 10)				
A2(2, 5)				
A3(8, 4)				
A4(5, 8)				

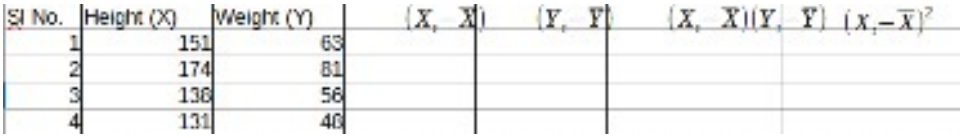
b) State the difference between K-means clustering and K-nearest-Neighbour algorithm for machine learning. Write the algorithm of K-Nearest neighbourhood algorithm — 10

6 .

a) State the steps of Linear regression with necessary mathematical derivation. Assume $(X_i, Y_i), i = 1,2,3,\ldots,N$ are the given input-output pairs. Find the expression of unknown parameters (b_0) and b_1 in $Y_i = b_0 + b_1 X_i$. Show that — 12

$$b_1 = \frac{\sum_{i=1}^{N} (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{N} (X_i - \bar{X})^2}$$

and $b_0 = \bar{Y} - b_1 \bar{X}$

	b)	Fill up the following table and find out the values of b_0 and b_1 . 	8
7.		Write short notes (any two)	10x2
	a)	Support vector machine	
	b)	Auto encoder	
	c)	Stochastic Regression	
	d)	Backpropagation learning.	

Sl No.	Height (X)	Weight (Y)	$(X_i - \bar{X})$	$(Y_i - \bar{Y})$	$(X_i - \bar{X})(Y_i - \bar{Y})$	$(x_i - \bar{x})^2$
1	151	63				
2	174	81				
3	138	56				
4	131	40				