



JK Lakshmipat University, Jaipur
Institute of Engineering and Technology
Machine Learning (CS1138)
End Term Examination, B.Tech., IV Semester

Roll No. : _____

Date: April 22, 2024

Duration: 2.5 hours

MM: 30

There are 5 questions in this paper with sub-parts. You have to attempt ALL questions. Make appropriate assumptions wherever required. Scientific calculators are allowed.

1. Multiple choice questions. Provide proper justification for your choice(s). There will be partial marking only when correct choice(s) and their justification is provided.

(1x6 - COs 2,3,5)

- (a) Which of the following layer(s) have learnable parameters (weights) in a Neural Network?
- A. Batch Norm
 - B. ReLU Activation
 - C. Softmax
 - D. Fully Connected
- (b) Which option(s) is/are True w.r.t. PCA:
- A. The objective function for PCA maximizes the length of the orthogonal projections of the data points to the new axes.
 - B. The objective function for PCA minimizes the length of the orthogonal projections of the data points to the new axes.
 - C. PCA maximizes the variance in the projected space.
 - D. PCA minimizes the variance in the projected space.
- (c) Which of the following objective function(s) is/are convex:
- A. Logistic function
 - B. Distortion function
 - C. Entropy function
 - D. Contrastive Loss
- (d) Which of the following statement(s) is/are True w.r.t. gradient descent:
- A. Decreasing the learning rate can make it diverge.
 - B. It always converges to the global optima.
 - C. Stochastic gradient descent is slower than mini-batch gradient descent.
 - D. Increasing the learning rate can make it diverge.
- (e) For a dataset S, the objective function for a Decision Tree classifier maximizes:
- A. Entropy
 - B. Gini Index
 - C. Information Gain
 - D. It minimizes all of the above functions, instead of maximizing them.

- (f) K-NN implementation can be made more efficient (w.r.t. execution time complexity) using which of the following:
- Stacks
 - KD-Tree
 - Priority Queue
 - Circular Linked List

2. (a) Perform a hierarchical clustering of the one-dimensional set of points $\{1, 4, 9, 16, 25, 36, 49, 64, 81\}$ assuming clusters are represented by their centroid, and at each step the clusters with the closest centroids are merged. Draw the dendrogram illustrating the merging steps.

(4+2 - COs 2,4,5)

- (b) What is the computational complexity for hierarchical clustering with N points each of D dimension. Justify.

3. Answer the following:

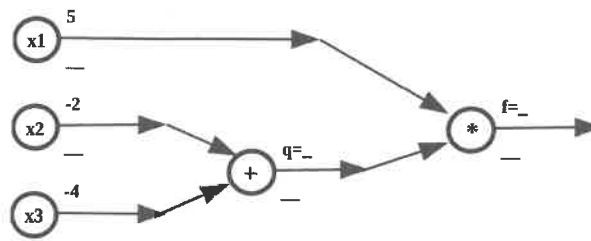
(3+3 - COs 3,4)

- (a) Given the computational graph for the following function and inputs. Show the forward propagation values, and the backprop gradients (marked by $_$) and corresponding derivative equations for each intermediate step.

$$f = x_1 * (x_2 + x_3)$$

[* being the multiplication operator].

where $x_1 = 5$, $x_2 = -2$, $x_3 = -4$



- (b) Given the following confusion matrix for a model, find its accuracy and the balanced accuracy values.

		Predictions		
		Class 0	Class 1	Class 2
Ground Truth	Class 0	4	1	0
	Class 1	0	23	4
	Class 2	10	3	40

4. A dataset is given below. Calculate the Gini index for the attributes Emotion, and Temp. Consider Temp > 100 as a threshold value. Find the Weighted Gini split and the subsequent choice of split attribute. Show the rows IDs that are associated with the child nodes.

(6 - COs 2,4,5)

RowID	Emotion	Temp	StayHome
R1	Sick	98	N
R2	Sick	105	Y
R3	NotSick	98	Y
R4	NotSick	98	N
R5	NotSick	101	Y
R6	Sick	101	N
R7	NotSick	99.5	N
R8	NotSick	102	Y

5. Case Study: Defect Detection

(2+2+2 - COs 1,5)

The manufacturing unit of JK Plastic has 10 machines that are dedicated for cutting fixed length plastic sheets from plastic rolls at the rate of 1 sheet per second per machine. The defects like holes, weft, weaving, burns, etc. may creep into the earlier stages of plastic processing that may appear at any place in the rolls. A major loss to the company occurs when these defects are found in the fixed cut-length plastic sheets, as these full cut-length pieces need to be discarded. At present, the wastage to the company is around 5% i.e., 5% of the cut pieces are wasted due to these defects. The management of the company decides to install a machine learning and computer vision system for each of the 10 machines, such that the machine will halt in real time before the fixed length piece is cut. This will help reduce their wastage as only the defected area may be discarded which will be usually less than the complete cut-length. The only drawback is that removing this defect after the machine has halted, is done manually by a human operator and it takes around 10 - 20 secs to restart the machine. You are given the task of deploying this ML - CV system and you have managed to train a model (which is obviously not perfect) that detects the defects in real-time and halts the machine via an interrupt signal.

- You decide to evaluate the deployed model for a machine. With respect to this application, explain what are True/False Positives, and True/False Negatives for your evaluation.
- A certain run of the machine during testing, gives the following statistics. In 1 hour the machine was halted 250 times, number of cut-lengths without defect were 2000, number of actual defects that occurred were 200, and the model correctly detected 153 defects. Find the precision, recall and F1-score for the currently deployed model.
- Explain which evaluation metric needs to be maximized in order to increase the productivity of the manufacturing unit and for reducing the amount of wastage.

