



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

Roll No. : _____

Date: May 06, 2025

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). No marks without proper justification. (1x4 + 2 - COs 2,3,5)
 - (a) In hierarchical clustering, which statement(s) is/are True?
 - A. Upper bound on computational complexity is $O(N^3)$, assuming constant dimensions.
 - B. Clustroids represent the average of points belonging to a cluster.
 - C. Minkowski distance is the only valid distance measure.
 - D. Priority Queues can improve the execution time complexity.
 - (b) In LASSO regression, if the regularization parameter $\lambda = 0$, then which of the following statement(s) is/are True?
 - A. This LASSO model can be used for feature selection.
 - B. Large coefficients are penalized.
 - C. The loss function is as same as the ridge regression loss function.
 - D. The loss function is as same as the ordinary least square loss function.
 - (c) Which of the following is a true statement about gradient descent (GD)?
 - A. GD can be expensive to run on datasets with a large number of samples.
 - B. Decreasing the learning rate, keeping all other hyperparameters fixed, guarantees that the error of our estimated parameters will decrease.
 - C. When training, we should not update the bias (aka offset, or intercept) term using GD.
 - D. An advantage of GD over SGD is that GD requires only a single update step to converge.
 - (d) Given the following graph, for what (minimal) value of k (in kNN) will the query point "?" be negative. Enter "0" if this is impossible. Ties are broken at random - try to avoid them.



- (e) How does Random Forest reduce variance error as compared to Decision Trees? Mention its design choices.

2. (a) Write the steps of K-Means algorithm along with its optimization objective. Discuss its time complexity and how can we make it more efficient?
- (b) What is the effect on kNN algorithm, if the dimensions increase? Show mathematically.

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

3. Answer the following:

(2+1+3 - COs 3,4)

- (a) Draw the computational graph for the following function and compute the forward pass and backward pass values for inputs $a = 2, b = 1$.
 $f = (a + b) * (b + 1)$ [* being the multiplication operator].

- (b) Consider a neural network with 8 layers trained on a dataset of 800 samples with a batch size of 10. How many forward passes through the entire network are needed to train this model for 5 epochs?

- (c) What are activation functions and why are they necessary for neural networks? Discuss their mathematical implications. Mention any 3 activation functions.

4. A dataset is given below to find if John likes a type of food appealing or not based on its temperature, taste and size. Construct the full decision tree for this dataset using ID3 algorithm. Show entropy and information gain values for all the splits. (6 - COs 2,4,5)

Appealing	Temperature	Taste	Size
No	Hot	Salty	Small
No	Cold	Sweet	Large
No	Cold	Sweet	Large
Yes	Cold	Sour	Small
Yes	Hot	Sour	Small
No	Hot	Salty	Large
Yes	Hot	Sour	Large
Yes	Cold	Sweet	Small
Yes	Cold	Sweet	Small
No	Hot	Salty	Large

5. Answer the following questions:

(4+2 - COs 1,5)

- (a) What are the steps for reducing an n -dimensional data to a k -dimensional data using PCA, where $k < n$? How can we choose the value of k ?
- (b) What is the Crowding problem in t-SNE? Explain.