AUTUMN END SEMESTER EXAMINATION-2023
5th Semester B.Tech (DE-II)

MACHINE LEARNING

CS 3035

(For 2022 (L.E), 2021 & Previous Admitted Batches)

Time: 3 Hours   Full Marks: 50

Answer any SIX questions.
Question paper consists of four SECTIONS i.e. A, B, C and D.
Section A is compulsory.
Attempt minimum one question each from Sections B, C, D.
The figures in the margin indicate full marks.
Candidates are required to give their answers in their own words as far as practicable
and <u>all parts of a question should be answered at one place only</u>.

SECTION-A

1.   Answer the following questions.   [1 × 10]

(a)   Illustrate the basic difference between Machine Learning and Traditional Programming.

(b)   What are the issues with gradient descent algorithm?

(c)   Distinguish the similarities and dissimilarities between k-Means and k-Nearest Neighbors.

(d)   Explain the difference between likelihood and probability.

(e)   Why log function is used as cost function in logistic regression?

(f)   Define Gini Impurity. What is the range of Gini Impurity?

(g)   Explain what happens, when there is no clear separating hyperplane in SVM.

(h)   What are the two types of pruning used in Random Forests?

(i)   Why KKT conditions are useful?

(j)   List some advantages of using Multilayer perceptron over a Single-layer Perceptron.

SECTION-B

2. (a) Compare and contrast among Supervised, Unsupervised and Semi-Supervised learning with one example of application. How Reinforcement learning is different from others? [4]

(b) Perform KNN Classification on following dataset and predict the class for (height=170 and weight=57) with K=5 using Euclidean distance. [4]

Height(CM)	Weight(KG)	Class
167	51	Underweight
182	62	Normal
176	69	Normal
172	65	Normal
173	64	Normal
174	56	Underweight
169	58	Normal
173	57	Normal
170	55	Normal
170	57	?

3. (a) What do you mean by regularization? Explain the difference between L1 and L2 regularization in linear regression, and when should each be used. [4]

(b) Describe the terms with example: accuracy, sensitivity, specificity, F1-score, and ROC-AUC. Also, find the accuracy, precision, recall and F1-score for the following confusion matrix: [4]

	Actual Positive	Actual Negative
Predicted Positive	[TP]=150	[FP]=10
Predicted Negative	[FN]=20	[TN]=120

SECTION-C

4. (a) Derive the gradient ascent training rule for weight update equation of Logistic regression by using maximum likelihood function. [4]

(b) Consider the following data set and predict the class of new instance X= (Slow, Rarely, No) using Naive Bayes classification algorithm. [4]

Sl. No	Swim	Fly	Crawl	Class
1	Fast	No	No	Fish
2	Fast	No	Yes	Animal
3	Slow	No	No	Animal
4	Fast	No	No	Animal
5	No	Short	No	Bird
6	No	Short	No	Bird
7	No	Rarely	No	Animal
8	Slow	No	Yes	Animal
9	Slow	No	No	Fish
10	Slow	No	Yes	Fish
11	No	Long	No	Bird
12	Fast	No	No	Bird

5. (a) List out the link functions? Also discuss some common link functions and their uses. [4]

(b) Evaluate the Euclidean distance, Manhattan distance, Minkowshi distance and Cosine distance for the following two points. [4]

P1: (1, 0, 2, 5, 3) P2: (2,1, 0, 3, -1)

6. (a) Consider the following data set. [4]

Color	Size	Act	Age	Inflated
YELLOW	SMALL	DIP	ADULT	F
YELLOW	LARGE	STRETCH	ADULT	T
YELLOW	LARGE	STRETCH	CHILD	F
YELLOW	LARGE	DIP	ADULT	F
YELLOW	LARGE	DIP	CHILD	F
PURPLE	SMALL	STRETCH	ADULT	T
PURPLE	SMALL	STRETCH	ADULT	T
PURPLE	SMALL	STRETCH	CHILD	F
PURPLE	SMALL	DIP	ADULT	F
PURPLE	SMALL	DIP	CHILD	F

Calculate the information gain of each attribute. State which attribute should be used as the first root node based on the information gain parameter.

(b) Compare the difference between Bagging and Boosting algorithms. Explain Adaboost algorithm with a neat diagram. [4]

SECTION-D

7. (a) Explain the Principal Component Analysis (PCA) and reduce the following dataset step-by-step from 2 dimensions to 1. [4]

Feature	Example 1	Example 2	Example 3	Example 4
x	2	1	0	-1
y	4	3	1	0.5

(b) Discuss the Mercers conditions for a valid kernel function. Name at least four valid kernel functions used in SVM. Write down the expression for RBF and Polynomial Kernel. [4]

8. (a) Use McCulloch-Pitts Neuron to implement AND function (take binary data representation). [4]

X1	X2	Y
1	1	1
1	0	0
0	1	0
0	0	0

(b) Draw an ANN architecture for 3 inputs, 2 hidden layer each with 2 neurons and one output. Derive the estimated output using sigmoid transfer function. [4]

*****