NATIONAL INSTITUTE OF TECHNOLOGY, KURUKSHETRA

THEORY EXAMINATION

Question Paper

Month and year of the Examination: **Dec-2022**

Programme: **B.Tech.** Semester: - **7th Semester**

Subject: - **Machine Learning** Course No: - **ITPC-41**

Number of Questions to be Attempted: **5** Maximum Marks: - **50**

Total No. of Questions: **5** Time Allowed: - **3 Hours**

Total No. of Pages used: **2**

Note: - There will be internal choice in Question no. 2

Ques 1 (a)	What is Decision Tree? Explain its advantages and disadvantages.	(4)
Ques 1 (b)	What is a kernel in SVM? Why do we use kernels in SVM?	(4)
Ques 1 (c)	Write a short note on Unsupervised Learning with example.	(2)
Ques 2 (a)	Find linear regression equation for the following two sets of data: x: 2, 4, 6, 8 y: 3, 7, 5, 10	(3)
Ques 2 (b)	What is the difference between logistic regression and SVM (Support Vector machine)?	(3)
Ques 2 (c)	Define Cross Validation. Explain k-fold cross validation. What do you mean by **k** in k-fold cross validation?	(4)
	OR	
Ques 2 (a)	Define Bayes Rule. Explain Naïve Bayes' Classifier and its application.	(3)
Ques 2 (b)	Given all the previous doctor have seen (below table are their symptoms and diagnosis). Does doctor believe that a patient with the given data sample {chills= Y, runny nose= N, headache= Mild, Fever= Y} has the flu? (Use Find S Algorithm).	(3)
Ques 2 (c)	Write a short on note Linear Discriminant Analysis (LDA). Differentiate between two class LDA and three class LDA.	(4)

chills	runny nose	headache	fever	flu?
Y	N	Mild	Y	N
Y	Y	No	N	Y
Y	N	Strong	Y	Y
N	Y	Mild	Y	Y
N	N	No	N	N
N	Y	Strong	Y	Y
N	Y	Strong	N	N
Y	Y	Mild	Y	Y

Ques 3 (a)	Use K Means clustering to cluster the following data into two groups. Assume cluster centroid are **m1**=2 and **m2**=4. The distance function used is Euclidean distance. {2, 4, 10, 12, 3, 20, 30, 11, 25}.	(4)
Ques 3 (b)	Write a Short note on Artificial Neural Network. Also describe the similarity with biological neural network.	(4)
Ques 3 (c)	Write a short note on Chi-Square test with example.	(2)
Ques 4 (a)	What do you mean by Convolutional Neural Network? Why do we prefer Convolutional Neural Network for image data as input?	(5)
Ques 4 (b)	Draw decision tree for the following dataset. Explain each step in detail.	(5)
Ques 5 (a)	Discuss the Random forest model in detail. What are the features of random forest?	(4)
Ques 5 (b)	Define Clustering. How it is differed from Classification.	(3)
Ques 5 (c)	Classify whether a special paper {X1=3, X2= 7} is good or bad using KNN. Let us assume K=3.	(3)

Instance	Classification	a1	a2
1	+	T	T
2	+	T	T
3	-	T	F
4	+	F	F
5	-	F	T
6	-	F	T

X1: acid durability	X2: strength (kg/sq. fit)
7	7
7	4
3	4
1	4