Code No :

R20

III B. TECH I SEMESTER REGULAR EXAMINATION - 2022
MACHINE LEARNING
(CSM)

Time : 3 Hours **Max. Marks : 70**

Note : Answer **ONE** question from each unit **(5 × 14 = 70 Marks)**

~~~~~~

		UNIT-I	
1.	a)	Define Machine Learning and Describe the process of estimating f?	[7M]
	b)	Describe the process of assessing Model's accuracy?	[7M]
		(OR)	
2.	a)	Distinguish between supervised learning and Unsupervised learning. Illustrate with an example.	[7M]
	b)	Discuss any four examples of machine learning applications.	[7M]
		UNIT-II	
3.	a)	Consider the training data in the following table where Play is a class attribute. In the table, the Humidity attribute has values "L" (for low) or "H" (for high), Sunny has values "Y" (for yes) or "N" (for no), Wind has values "S" (for strong) or "W" (for weak), and Play has values "Yes" or "No". **Humidity / Sunny / Wind / Play** L / N / S / No H / N / W / Yes H / Y / S / Yes H / N / W / Yes L / Y / S / No What is class label for the following day (Humidity=L, Sunny=N, Wind=W), according to naïve Bayesian classification?	[7M]
	b)	Suppose we collect data for a group of students in a statistics class with variables X1 = hours studied, X2 = undergrad GPA, and Y = receive an A. We fit a logistic regression and produce estimated coefficient, $\hat{\beta}0 = -6$, $\hat{\beta}1 = 0.05$, $\hat{\beta}2 = 1$. (a) Estimate the probability that a student who studies for 40 h and has an undergrad GPA of 3.5 gets an A in the class. (b) How many hours would the student in part (a) need to study to	[7M]
~~~~~~

		have a 50 % chance of getting an A in the class?	
		(OR)	
4.	a)	Compare Linear regression with K-Nearest neighbors.	[7M]
	b)		[7M]

	Pepper	**Ginger**	**Chilly**	**Liked**
A	True	True	True	False
B	True	False	False	True
C	False	True	True	False
D	False	True	False	True
E	True	False	False	True

"Restaurant A" sells burgers with optional flavors: Pepper, Ginger, and Chilly. Every day this week you have tried a burger (A to E) and kept a record of which you liked. **Using Hamming distance, show how the 3NN classifier with majority voting would classify** { **pepper: false, ginger: true, chilly: true**}

		UNIT-III	
5.	a)	Explain bootstrap briefly.	[7M]
	b)	What are the advantages and disadvantages of k-fold cross validation relative to LOOCV?	[7M]
		(OR)	
6.	a)	Explain model selection techniques in detail?	[7M]
	b)	Differentiate ridge regression and lasso regression.	[7M]
		UNIT-IV	
7.	a)	The dataset in the below Table will be used to learn a decision tree for predicting whether a mushroom is edible or not based on its shape, color and odor.	[7M]

Shape	Color	Odor	Edible
C	B	1	Yes
D	B	1	Yes
D	W	1	Yes
D	W	2	Yes
C	B	2	Yes
D	B	2	No
D	G	2	No
C	U	2	No
C	B	3	No
C	W	3	No
D	W	3	No

(a) What is the entropy H(Edible | (Odor = 1 OR Odor = 3))?

(b) Which attribute would the ID3 algorithm choose to use for the root of the tree (no pruning)?

		(c) Draw the full decision tree that would be learned for this data (no pruning).	
	b)	Consider the table below, which contains data about second-hand Car listings. We want to build a decision tree classifier to predict whether it is an acceptable deal to buy a listed second-hand car. We consider the two classes 'acc' and 'unacc'. a) Calculate the GINI index of the root node of a decision tree with the given training data(first seven rows) b) Build the decision tree using the ID3 algorithm. c) Use the decision tree to classify the unknown points in the table.	[7M]

No	Pricing	Condition	Seating	Safety	Acceptability
1	High	Great	Four	High	Acc
2	Low	Good	Four	Low	Acc
3	Low	Great	Two	High	Acc
4	Low	Good	Two	High	Unacc
5	High	Bad	Four	Low	Unacc
6	High	Great	Four	Low	Unacc
7	Low	Good	Two	Low	Unacc
8	High	Good	Four	Low	?
9	Low	Great	Two	Low	?

(OR)

8.	a)	Explain random forest classifier in detail.	[7M]
	b)	Explain Boosting Techniques in detail.	[7M]

UNIT-V

9.	a)	Explain Support Vector Machines Classifier in detail.	[7M]
	b)	Explain maximal margin classifier in detail?	[7M]

(OR)

10.	a)	Explain K – Means Clustering algorithm with a suitable example.	[7M]
	b)	Explain hierarchical clustering in detail.	[7M]

*** * *