GURU NANAK INSTITUTE OF TECHNOLOGY

An Autonomous Institute under MAKAUT

2022

MACHINE LEARNING

MCA20-E305G

TIME ALLOTTED: 3Hours **FULL MARKS:70**

The figures in the margin indicate full marks.
Candidates are required to give their answers in their own words as far as practicable

GROUP – A
(Multiple Choice Type Questions)

Answer any ***ten*** from the following, choosing the correct alternative of each question: **10×1=10**

			Marks	CO No
1.	i)	Machine Learning is a Subset of --------- a. Deep Learning b. Artificial Intelligence c. Data Learning d. None of the above	1	CO3
	ii)	Application of Machine learning is _____ a. Email filtering b. Sentimental analysis c. Face recognition d. All of the above	1	CO3
	iii)	The categories in which Machine learning approaches can be traditionally categorized as a. Supervised learning b. Unsupervised learning c. Reinforcement learning d. All of the above	1	CO3
	iv)	____ is the machine learning algorithms that can be used with unlabeled data. a. Regression algorithms b. Clustering algorithms c. Classification algorithms d. All of the above	1	CO3
	v)	Which among the following algorithms are used in Machine learning? a. Naïve Bayes b. Support Vector Machines c. K-Nearest Neighbors d. All of the above	1	CO3
	vi)	_____ is a disadvantage of decision trees? a. Robust to outliers b. Prone to be over fitting c. Both A & B d. None of the above	1	CO3

vii) Which of the following distance metric cannot be used in k-NN? 1 CO3
   a. Manhattan
   b. Euclidean
   c. Mahalanobis
   d. All can be used

viii) The effectiveness of an SVM depends upon: 1 CO3
   a. Selection of Kernel
   b. Kernel Parameters
   c. Soft Margin Parameter C
   d. All of the above

ix) A process by which we estimate the value of dependent variable on the basis of one or more independent variables is called: 1 CO3
   a. Correlation
   b. Regression
   c. Residual
   d. Slope

x) The method of least squares dictates that we choose a regression line where the sum of the square of deviations of the points from the lie is: 1 CO3
   a. Maximum
   b. Minimum
   c. Zero
   d. Positive

xi) Which of the following is not a type of supervised learning? 1 CO3
   a. Classification
   b. Regression
   c. Clustering
   d. None of the above

GROUP – B
(Short Answer Type Questions)
(Answer any *three* of the following) **3 x 5 = 15**

		Marks	CO No
2.	Define the terms Hypothesis space and Version space. Illustrate with an example.	5	CO2
3.	What are training and test data?	5	CO1
4.	Write the Bayes Theorem and write the significance of the theorem.	5	CO4
5.	Write the differences between Linear Regression and Logistic Regression.	5	CO5
6.	Define Supervised and Unsupervised learning methods with example.	5	CO2

GROUP – C
(Long Answer Type Questions)
(Answer any *three* of the following) **3 x 15 = 45**

		Marks	CO No
7 a.	What is Regression?	5	CO2
b.	What are the different forms of regression?	5	CO1
c.	Explain the Difference Between Classification and Regression?	5	CO1
8 a.	What are the applications of Machine Learning?	5	CO4
b.	Discuss the steps for building a Decision Tree.	10	CO4
9 a.	Define covariance of two vectors.	3	CO3
b.	What are the causes of failure of Decision tree?	5	CO3
c.	What are Boosting and Bootstrapping? Explain.	7	CO2
10 a.	What is ensemble modeling?	5	CO5
b.	Classify ensemble models.	5	CO2
c.	Differentiate sequential and parallel models in context to ensemble models.	5	CO5
11.	Write Short note: (Any three)	3x5=15	
a.	Linear regression	5	CO4
b.	Support Vector Machine	5	CO4
c.	KNN	5	CO4
d.	Artificial Neural Networks	5	CO4
e.	Data Mining	5	CO4