F

10000CS467122001

Pages: 3

Reg No.:_____ Name:_____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

Seventh Semester B.Tech Degree Examination (Regular and Supplementary), December 2020

Course Code: CS467

Course Name: MACHINE LEARNING

Max. Marks: 100 Duration: 3 Hours

PART A

Answer all questions, each carries 4 marks. Marks

1 Define supervised learning and unsupervised learning with example. (4)

2 Differentiate between overfitting and underfitting with suitable diagrams. (4)

3 Write down the forward selection algorithm for implementing the subset selection procedure under dimensionality reduction. (4)

4 Express the confusion matrix table for two class datasets. (4)

5 Calculate the Gini index of the given number of class label with bird=2, fish=2 and reptile=1. (4)

6 Describe any four types of activation functions used in Artificial Neural Networks. (4)

7 Explain the basic problems of Hidden Markov Models. (4)

8 What is the purpose of Kernel functions used in Support Vector Machines? (4)

9 Distinguish between K-means Clustering and Hierarchical Clustering technique. (4)

10 Write any four commonly used equations for measures of distance between data points of numeric data in Hierarchical Clustering. (4)

PART B

Answer any two full questions, each carries 9 marks.

11 a) Distinguish between classification and regression with example. (5)

b) List any four applications of machine learning. (4)

12 a) What is Probably Approximately Correct (PAC) Learning? (4)

b) Illustrate the concept of Vapnik-Chervonenkis (VC) dimension. (5)

13 a) Explain the procedure for performing a Principal Component Analysis (PCA) on a given data set. (5)

b) What is reinforcement learning in machine learning and list any two applications? (4)

PART C

Answer any two full questions, each carries 9 marks.

14 Show the first splitting attribute for decision tree by using ID3 algorithm with the following data set. (9)

Day_No.	View	Temperature	Moisture	Breeze	Play Cricket
1	Bright	Hot	High	Weak	NO
2	Bright	Hot	High	Strong	NO
3	Cloudy	Hot	High	Weak	YES
4	Rain	Mild	High	Weak	YES
5	Rain	Cool	Normal	Weak	YES
6	Rain	Cool	Normal	Strong	NO
7	Cloudy	Cool	Normal	Strong	YES
8	Bright	Mild	High	Weak	NO
9	Bright	Cool	Normal	Weak	YES
10	Rain	Mild	Normal	Weak	YES
11	Bright	Mild	Normal	Strong	YES
12	Cloudy	Mild	High	Strong	YES
13	Cloudy	Hot	Normal	Weak	YES
14	Rain	Mild	High	Strong	NO

15 a) Illustrate Naïve Bayes algorithm for the dataset having *n* features. (5)

b) Explain the Receiver Operating Characteristics (ROC) Space in machine learning. (4)

16 a) Describe various types of regression models based on type of functions. (4)

b) Explain the issues involved in decision tree learning. (5)

PART D

Answer any two full questions, each carries 12 marks.

17 a) Describe the features of soft margin hyperplane and explain how it is computed. (6)

b) Explain the bagging and boosting methods used in learning algorithms. (6)

18 a) Write the algorithm for DIANA (**DI**visive**ANA**lysis) of hierarchical clustering technique. (6)

b) Illustrate the algorithm for K-means clustering of unsupervised learning. Write the disadvantages and applications of K-means clustering. (6)

19 a) Demonstrate an algorithm to find the SVM classifier and describe the mathematical formulation of the SVM problem. (6)

b) Construct the dendrogram using Complete Linkage Method for the given data points. (6)

	a	b	c	d	e
a	0				
b	8	0			
c	2	6	0		
d	5	4	8	0	
e	10	9	1	7	0

****