F

03000CS467122301

Pages: 2

Reg No.:_____ Name:_____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

B.Tech Degree S7 (S, FE) / S7 (PT) (S, FE) Examination December 2023 (2015 Scheme)

Course Code: CS467

Course Name: MACHINE LEARNING

Max. Marks: 100 Duration: 3 Hours

PART A

Answer all questions, each carries 4 marks. Marks

1		What is meant by "learning" in the context of machine learning?	(4)
2		List out different types of data.	(4)
3		Explain the methods to learn multiple classes for K class classification problem.	(4)
4		Define the following terms i) sensitivity ii) specificity iii)Accuracy and Error rate	(4)
5		Explain Receiver Operating Characteristic (ROC) and special points in ROC curve.	(4)
6		Distinguish between classification and regression trees.	(4)
7		Why do we prefer to combine many learners together?	(4)
8		Describe the term optimal separating line in SVM.	(4)
9		Explain the algorithm for density based clustering (DBSCAN).	(4)
10		Elucidate the features of Agglomerative clustering	(4)

PART B

Answer any two full questions, each carries 9 marks.

11	a)	Explain supervised learning with an example	(6)
	b)	Comment on input representation	(3)
12	a)	In the context of classification problems explain with examples the following: (i) Hypothesis (ii) Hypothesis space.	(6)
	b)	Distinguish between over fitting and under fitting.	(3)
13	a)	Explain feature selection and feature extraction method for dimensionality reduction.	(5)
	b)	What is Bootstrapping method for evaluating classifier performance?	(4)

PART C

Answer any two full questions,each carries 9 marks.

14	a)	Three factories A, B, C of an electric bulb manufacturing company produce respectively 35%. 35% and 30% of the total output. Approximately 1.5%, 1% and 2% of the bulbs produced by these factories are known to be defective. If a randomly selected bulb manufactured by the company was found to be defective, what is the probability that the bulb was manufactures in factory A?	(6)

b) Write the Naïve Bayes algorithm. (3)

15 a) Calculate the information Gain of the attribute Outlook from the table given below (6)

Day	outlook	temperature	humidity	wind	playtennis
D1	sunny	hot	high	weak	no
D2	sunny	hot	high	strong	no
D3	overcast	hot	high	weak	yes
D4	rain	mild	high	weak	yes
D5	rain	cool	normal	weak	yes
D6	rain	cool	normal	strong	no
D7	overcast	cool	normal	strong	yes
D8	sunny	mild	high	weak	no
D9	sunny	cool	normal	weak	yes
D10	rain	mild	normal	weak	yes
D11	sunny	mild	normal	strong	yes
D12	overcast	mild	high	strong	yes
D13	overcast	hot	normal	weak	yes
D14	rain	mild	high	strong	no

b) Figure out the issues involved in decision tree learning. (3)

16 a) Suppose a computer program for recognizing dogs in photographs identifies eight dogs in a picture containing 12 dogs and some cats. Of the eight dogs identified, five actually are dogs while the rest are cats. Compute the precision and recall of the computer program. (4)

b) Define the term entropy. Calculate the entropy of the following data. (5)

Name	Gives birth	Aquatic animal	Aerial animal	Has legs	Class label
human	yes	no	no	yes	mammal
bat	yes	no	yes	yes	bird
cat	yes	no	no	yes	mammal
shark	yes	yes	no	no	fish

PART D

Answer any two full questions, each carries 12 marks.

17 a) What do you mean by a base learner? How are they chosen? (6)

b) Explain the working of random forest algorithm. (6)

18 a) What are the key principles behind Support Vector Machines (SVMs), and how does the algorithm work to classify data points in a given feature space? (6)

b) Differentiate between Bagging and Booting techniques . (6)

19 a) Illustrate K-means clustering algorithm with an example . (6)

b) Describe the working of Complete linkage clustering and single linkage clustering . (6)

*****