

No. of Printed Pages : 5

RCSE–002

**Ph. D. (Course Work) COMPUTER
SCIENCE (PHDCS)**

Term-End Examination

December, 2024

RCSE-002 : MACHINE LEARNING

Time : 3 Hours

Maximum Marks : 100

Weightage : 50%

Note : (i) *Question No. 1 is compulsory.*

(ii) *Answer any **three** questions from the rest.*

1. (a) What is Machine Learning ? Explain life cycle of machine learning. What do you mean by training and testing in machine learning ? Explain how to prepare data for machine learning. 10

- (b) What is labelled data ? How is it different from unlabelled data ? Explain how hypothesis testing can be performed for a large set of data (large sample size). 10
- (c) Explain Find-S algorithm to find a maximally specific hypothesis with a suitable example. 10
- (d) What is Inductive Bias in concept learning ? Explain biased hypothesis space and an unbiased learner with the help of an suitable example. 10
2. (a) Explain with a suitable example how Gini index and entropy play a major role in the selection of attributes for partitioning the training samples in decision tree induction.

- (b) For the following given set of training examples, draw decision tree using CART classification algorithm. Show intermediate steps : 12

Outlook	Temp.	Humidity	Wind	Play Football
Sunny	Hot	High	Weak	No
Sunny	Hot	High	Strong	No
Overcast	Hot	High	Weak	Yes
Rain	Mild	High	Weak	Yes
Rain	Cool	Normal	Weak	Yes
Rain	Cool	Normal	Strong	No
Overcast	Cool	Normal	Strong	Yes
Sunny	Mild	High	Weak	No
Sunny	Cool	Normal	Weak	Yes
Rain	Mild	Normal	Weak	Yes
Sunny	Mild	Normal	Strong	Yes
Overcast	Mild	High	Strong	Yes
Overcast	Hot	Normal	Weak	Yes
Rain	Mild	High	Strong	No

3. (a) Explain K-means algorithm with the help of a suitable example. 10
- (b) Based on the following training examples, identify the class for $X = \{\text{yes, no, yes, no}\}$, using Naïve-Bayes' classifier : 10

Name	Give Birth	Can fly	Live in water	Have legs	Class
Human	Yes	No	No	Yes	Mammals
Python	No	No	No	No	Non-mammals
Salmon	No	No	Yes	No	Non-mammals
Whale	Yes	No	Yes	No	Mammals
Frog	No	No	Sometimes	Yes	Non-mammals
Komodo	No	No	No	Yes	Non-mammals
Bat	Yes	Yes	No	Yes	Mammals
Pigeon	No	Yes	No	Yes	Non-mammals
Cat	Yes	No	No	Yes	Mammals
Leopard	Yes	No	Yes	No	Non-mammal
Shark					

Turtle	No	No	Sometimes	Yes	Non-mammal
Penguin	No	No	Sometimes	Yes	Non-mammal
Porcupine	Yes	No	No	Yes	Mammals
Eel	No	No	Yes	No	Non-mammals
Dolphin	Yes	No	Yes	No	Mammals

4. (a) Explain how to find difference in error of two hypotheses. How will z -test help in this operation ? 10
- (b) What is PAC-learnable concept class ? Also, explain PAC learnability with example. 10
5. Write short notes on the following : $4 \times 5 = 20$
- (a) Vapnik-Chervonenki's dimension
 - (b) True error of hypothesis
 - (c) Supervised and unsupervised learning
 - (d) Confidence interval and confidence level

× × × × × × ×