JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -2 EXAMINATION- 2025

B.Sc.(Hons.)- V Semester

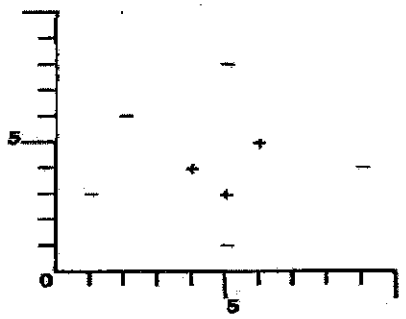
COURSE CODE (CREDITS): 18B1WCI634 (2) MAX. MARKS: 25

COURSE NAME: MACHINE LEARNING

COURSE INSTRUCTORS: SKP MAX. TIME: 1 Hour 30 Min

Note: *(a) All questions are compulsory.*
(b) Calculator is allowed.

QN	Question	CO	Marks
Q1	*ID3 searches for just one consistent hypothesis, whereas the CANDIDATE-ELIMINATION algorithm finds all consistent hypotheses. Consider the correspondence between these two learning algorithms.* **Example** ***Sky*** ***AirTemp*** ***Humidity*** ***Wind*** ***Water*** ***Forecast*** ***EnjoySport*** 1 Sunny Warm Normal Strong Warm Same Yes 2 Sunny Warm High Strong Warm Same Yes 3 Rainy Cold High Strong Warm Change No 4 Sunny Warm High Strong Cool Change Yes (a) Show the decision tree that would be learned by ID3 assuming it is given the four training examples for the ***Enjoy Sport?*** Target concept on the above dataset. (b) What is the relationship between the learned decision tree and the version space that is learned from these same examples? Is the learned tree equivalent to one of the members of the version space? (c) Add the following training example, and compute the new decision tree. This time, show the value of the information gain for each candidate attribute at each step in growing the tree. Sky | AirTemp | Humidity | Wind | Water | Forecast | EnjoySport Sunny | Warm | Normal | Weak | Warm | Same | No	1,2	2+1+2
Q2	Give decision trees to represent the following boolean functions: a) $A \wedge \neg B$ b) A XOR B	2	1+2
Q3.	*Consider the instance space consisting of integer points in the x, y plane and the set of hypotheses H consisting of rectangles. More precisely, hypotheses are of the form $a \leq x \leq b, c \leq y \leq d$, where a, b, c, and d can be any integers. Also, Consider the version space with respect to the set of positive (+) and negative (-) training examples shown below.*	2	1+1+2+2



(a) What is the S boundary of the version space in this case? Write out the hypotheses and draw them in on the diagram.
(b) What is the G boundary of this version space? Write out the hypotheses and draw them in.
(c) Suppose the learner may now suggest a new x, y instance and ask the trainer for its classification. Suggest a query guaranteed to reduce the size of the version space, regardless of how the trainer classifies it. Suggest one that will not.
(d) Now assume you are a teacher, attempting to teach a particular target concept (e.g., $3 \leq x \leq 5, 2 \leq y \leq 9$. What is the smallest number of training examples you can provide so that the CANDIDATE-ELIMINATION algorithm will perfectly learn the target concept?

Q4. Consider the given Dataset:

S.N.	Weather	Temperature	Humidity	Wind	PlayTennis
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes

a) Using **C4.5**, calculate Gain Ratio for each attribute and determine the root node.
b) Using **CART**, calculate Gini Index for all attributes (consider binary splits for continuous attributes) and determine the root node.
c) Are the root nodes the same for C4.5, and CART? Justify.

Q5. Design a two-input perceptron that implements the boolean function $A \wedge \neg B$.

Q4.		2	3+3+2
Q5.		3	3

*** Best of luck ***