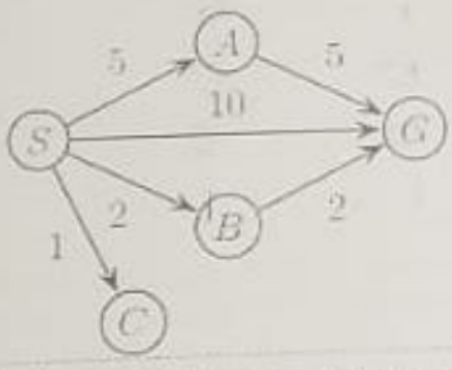
PES University, Bengaluru
(Established under Karnataka Act No. 16 of 2013)

UE18CS303

DECEMBER 2020: END SEMESTER ASSESSMENT B Tech CSE 5th SEMESTER

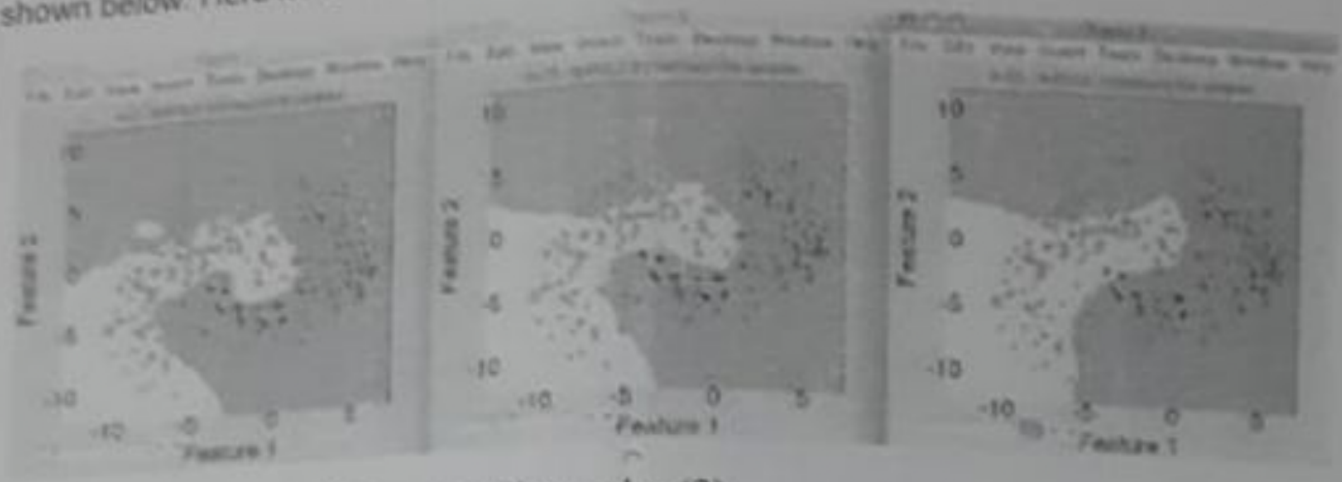
UE18CS303 – MACHINE INTELLIGENCE

Time: 3 Hrs	Answer All Questions	Max Marks: 100

1.	a)	When is a concept complete? (1) When is a concept consistent? (1) Using the completeness and consistency definitions, define a version space (2) When is a version space set convex?(1)	5
	b)	For each of the following activities, give a PEAS description of the task environment and characterize it in terms of the properties A. Exploring the subsurface oceans of Titan. (1) B. Shopping for used AI books on the Internet. (1) C. Playing a tennis match. (1) D. Practicing tennis against a wall. (1) E. Performing a high jump.(1) Example: Playing Soccer Partially observable, stochastic, sequential, dynamic, continuous, multi-agent.	5
	c)	Consider the following search problem:  Node / h: A 3; B 2; C 100; G 0 S is the start State and G is the goal – perform BFS,DFS, UCS, A*, Greedy Search and return the Nodes visited: (1 mark for each search algorithm)	5
	d	Consider the following data set: price / maintenance / capacity / airbag / profitable: low / low / 2 / no / yes low / med / 4 / yes / no low / low / 4 / no / yes low / high / 4 / no / no med / med / 4 / no / no med / med / 4 / yes / yes med / high / 2 / yes / no med / high / 5 / no / yes high / med / 4 / yes / yes high / high / 2 / yes / no high / high / 5 / yes / yes Considering 'profitable' as the binary values attribute, we are trying to predict, which of the attributes would you select as the root in a decision tree with multi-way splits using the cross-entropy impurity measure? Calculations for each attribute carries 1 mark and choice of root node carries 1 mark	5

2. a) Comment on KNN in terms of it being a parametric/non parametric and a Lazy/Eager algorithm? (2) — 5

Suppose you are given the following images(1 represents the left image, 2 represents the middle and 3 represents the right). Now your task is to find out the value of k in k-nn in each of the images shown below. Here k1 is for 1st, k2 is for 2nd and k3 is for 3rd figure.



Arrange K1, K2 and K3 in ascending order (2)

You have been given the following 2 statements. Find out which of these options is/are true in case of k-NN?

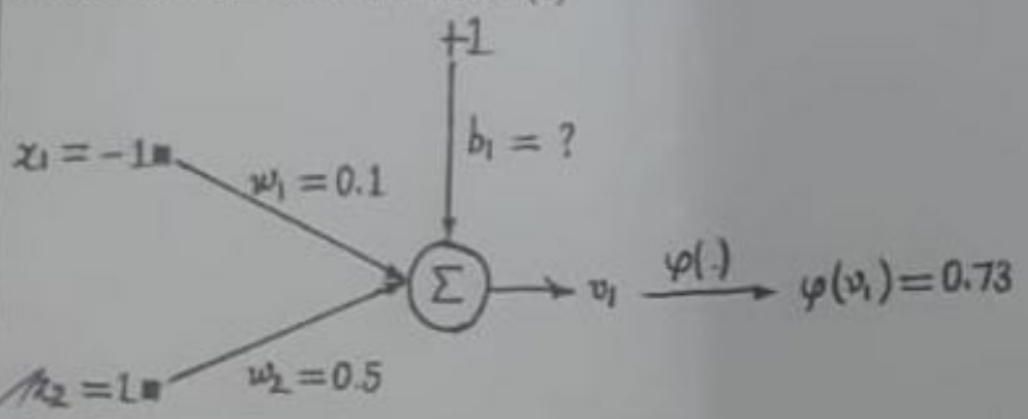
1. In case of very large value of k, we may include points from other classes into the neighborhood.
2. In case of too small value of k, the algorithm is very sensitive to noise.

State as True/False (1)

b) Given the XOR truth table as below draw a neural network without backpropagation with weights initialized for predicting the correct values. (3) — 5

Inputs		Output
A	B	X
0	0	0
0	1	1
1	0	1
1	1	0

Consider a 2 input neuron with logistic activation function and a slope parameter say a = 2. Let the inputs be [-1,1] and weights are [0.1, 0.5] respectively. The output of the neuron is 0.73. The value of the Bias would be what value ? (2)



c) State the primal formulation of SVM (2)
State the dual formulation of SVM (2)

d Compute w, b, alpha's Show all calculations for alphas for a SVM classifier
Suppose we are given the following positively labeled data points in R^2

Suppose we are given the following positively labeled data points in R^2:

$$\left\{\begin{pmatrix}3\\1\end{pmatrix}, \begin{pmatrix}3\\-1\end{pmatrix}, \begin{pmatrix}6\\1\end{pmatrix}, \begin{pmatrix}6\\-1\end{pmatrix}\right\}$$

And the following negatively labelled data points in R^2

$$\left\{\begin{pmatrix}1\\0\end{pmatrix}\cdot\begin{pmatrix}0\\1\end{pmatrix}\cdot\begin{pmatrix}0\\-1\end{pmatrix}\begin{pmatrix}-1\\0\end{pmatrix}\right\}$$

3. a) State the following algorithms in terms of initialization, induction/recursion and termination criteria of the hidden markov model in terms of alpha, beta, gamma and di-gamma probabilities. Please define your symbols for each and draw a diagram for the induction process(1) — 10

1. Forward Algorithm (3)
2. Backward Algorithm (3)
3. Veterbi Algorithm (3)

b) State the 3 advantages of Naïve-Bayes classifier. (3)
We want to classify a student into two classes "hirable (H)" and "non-hirable (N)" based on her percentage marks m in an exam. We found that probability P (H) = 0.8, and P (N) = 0.2. We also found that probability P (m=70|H) = 0.90, and P (m=70|N) = 0.30. The class to which Bayes classifier will classify a student who has obtained 70% marks is? (2) — 5

c) The following are the attributes of a car stolen database. — 5
Attributes are Color , Type , Origin, and the subject, stolen can be either yes or no.
The data set given is as follows

Example No.	Color	Type	Origin	Stolen?
1	Red	Sports	Domestic	Yes
2	Red	Sports	Domestic	No
3	Red	Sports	Domestic	Yes
4	Yellow	Sports	Domestic	No
5	Yellow	Sports	Imported	Yes
6	Yellow	SUV	Imported	No
7	Yellow	SUV	Imported	Yes
8	Yellow	SUV	Domestic	No
9	Red	SUV	Imported	No
10	Red	Sports	Imported	Yes

Classify a Red Domestic SUV as being stolen or Not stolen using Baye's Classifier.

4. a) The database of transactions in a book mart is as follows: — 10
Let the min_support = 27%

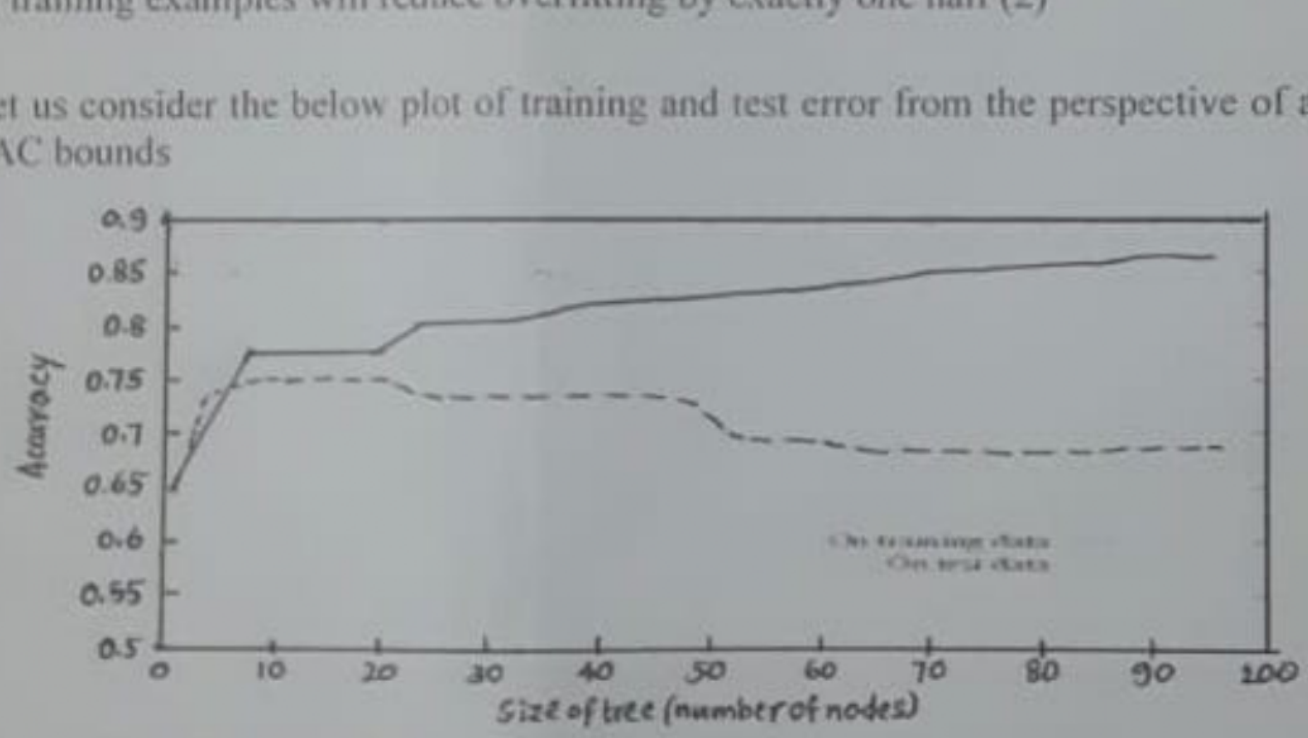
Trans_ID	Items
101	Book, Pen, Eraser
102	Pen, Pencil
103	Notebook, Book, Pen, Eraser
104	Book, Pen
105	Book, Notebook, Eraser

(Use alphabets B for Book, P for Pen, E for Eraser, PN for Pencil and N for Notebook)
1. Find all frequent sets using Apriori algorithm.
2. Write the FP tree, conditional pattern base and conditional FP-Tree

b) Suppose that the data mining tasks is to cluster the following 8 points (with (x,y) representing location) into 3 clusters: A1(2,10), A2(2,5), A3(8,4), B1(5,8), B2(7,5), B3(6,4),C1(1,2), C2(4,9).Plot a rough sketch of these 8 points. The distance function is Euclidean distance. Suppose initially we assign A1, B1 and C1 as the center of each cluster, respectively. Use K-means algorithm to show only the three cluster centers after the first round of execution — 5

c) Differentiate Single Linkage and Complete Linkage Techniques — 3

d) Define with an example unsupervised Learning? — 2

5.	a)	In roulette wheel selection the survival probability is proportional the Rank of the chromosome. Is this statement True/False(1) In Genetic algorithm the cross-over probability and mutation probability can be either High or Low. Should both the probabilities be High or Low or which should be High and which should be low?(1) In travelling saleman problem what is the type encoding and crossover adopted(2) Which GA operation is the most computationally expensive operation(1)	5
	b)	In a travelling salesman problem if there are 10 cities and the direction between cities is not important 1. How many genes do we use in a chromosome?(1) 2. How many genes will be there in the alphabet of the algorithm?(1) The Fitness Function in Genetic Algorithms is - method to measure how fit a candidate solution is in solving the problem. - The objective function for the optimization problem being solved. - a substitute to approximate the survival abilities of individuals in nature Which of the following statements are True (3)	5
	c)	Assume we used 200 training examples to produce the above decision tree plot. If we wish to reduce the overfitting to half of what we observe there, how many training examples would you suggest we use? Justify your answer in terms of the agnostic PAC bound, in no more than two sentences.(2) Give a one sentence explanation of why you are not certain that your recommended number of training examples will reduce overfitting by exactly one half (2) Let us consider the below plot of training and test error from the perspective of agnostic PAC bounds  Consider the agnostic PAC bound $$m \geq \frac{1}{2\epsilon^2}(\ln |H| + \ln(1/\delta))$$ where e is defined to be the difference between errortrue(h) and errortrain(h) for any hypothesis h output by the learner. State in one carefully worded sentence what the above PAC bound guarantees about the two curves in our decision tree plot above(2)	6
	d)	State which of the following are True/False PSO carries out both local and global search (1) PSO is faster and less computationally expensive then GA (1) PSO generally solves minimization problems (1) In PSO cognitive part of the equation requires the global best (1)	4