SRN

PES UNIVERSITY, BENGALURU-560085
(Estd. Under Karnataka Act 10 of 2013)

UE16CS353

MAY 2019: END-SEMESTER ASSESSMENT (ESA) B.TECH. VI SEMESTER

UE16CS353 –MACHINE LEARNING

Time: 3 Hrs | Answer All Questions. | Max Marks: 100

1.

a) Define Inductive bias. What is the inductive bias of the Candidate-Elimination algorithm? **[2]**

b) Clearly define overfitting. Explain two ways to overcome overfitting in Decision Trees. **[1+6]**

c) Briefly explain the following ways of selecting chromosomes in a typical Genetic Algorithm. **[6]**

a) Roulette wheel
b) Ranking
c) Tournament selection

d) Consider the following dataset. Using Shannon's Entropy, find out which attribute appears as the root node when the ID3 algorithm is applied. What is the max gain at the root? Show calculations clearly. **[5]**

BP	Sugar levels	Haemoglobin	Risk?
High	Normal	Low	No
High	High	Normal	Yes
Very High	Normal	Low	Yes
High	High	Normal	No
Very High	High	Low	Yes

2.

a) Explain three methods of avoiding local minima in an Artificial Neural Network. **[6]**

b) A custom activation unit is defined as $f(x) = \log_e(1 + e^x)$. Using this activation unit, show that the weight update rule (Using back-propagation for training) for an output neuron in an ANN is given by: **[6]**

$\Delta w_{ji} = \eta*(t_d - o_d) * \sigma(net_j) * x_{ji}$ (with usual conventions)

where σ is the sigmoid function and net_j is the output of the j^{th} neuron measured BEFORE the activation unit.

Assume Stochastic Gradient Descent is used and the Error term used is the Square Error.

c) For the following data points, use the SVM technique to find the marginal and the optimal hyperplanes. Hence prove that the dot product of any point on the optimal hyperplane with any point on the line **y = x**, is 0. **[6+2]**

	Attr1	Attr2	target
Instance x1	-2	0	-1
Instance x2	0	2	+1

SRN

3.	a)	Consider the following 1-D data. Decision stumps are used for this data as part of an Adaboost setup.	5+2

X	8	2	9	3	7	4	6	5	1
target	+	-	-	-	+	-	+	-	+

i) What are the starting weights of the instances?
ii) What is the weight of the first learner?
iii) What are the weights of the instances for the second learner?
iv) What is the weight of the second learner?

If a new instance is to be classified now, one of the learners is useless. Which one? Why?

NOTE: You can eyeball the data for the decision-stump splits.

b) Explain the terms "Naïve" and "Bayes" in the Naïve Bayes Classifier. For the tabular data shown below, apply the Naïve Bayes Classifier and give the classification for a person who has a runny nose, has mild headache, does not suffer from fever and has chills. — 5

Headache	Fever	Chills	Runny nose	Flu?
Strong	N	N	Y	N
Mild	Y	Y	Y	Y
Mild	Y	Y	N	N
No	N	Y	Y	Y
Strong	Y	Y	N	Y
Mild	Y	N	Y	Y
No	N	N	N	N
Strong	Y	N	Y	Y

c) The mood of a human being is modeled as a first-order Hidden Markov process. The moods are <Happy, Unhappy>. The possible observations are <Smile, Frown>. Calculate the likelihood of the observation sequence **<Frown, Smile>**. Show that the likelihood is the same if you use any of the following methods. — 2+3 +3

i) Brute Force
ii) Forward Rule only
iii) Forward and Backward rule together (with t = 1).

π = { 0.6, 0.4 }

A =	Happy	Unhappy
Happy	[0.7	0.3]
Unhappy	[0.4	0.6]

B =	Smile	Frown
Happy	[0.8	0.2]
Unhappy	[0.3	0.7]

4.	a)	Clearly explain how the DBSCAN algorithm is able to form clusters of arbitrary shapes and also separate out noisy data.	8

SRN

	b)	Consider the following transactions. i) Banana, ii) Apple, Carrot, Banana iii) Banana, Apple iv) Banana, Apple, Orange v) Carrot, Banana vi) Apple, Banana vii) Carrot, Orange, Apple • Using the FP-Growth algorithm, build the FP-Tree for the above transaction set. • Hence find all the frequent itemsets ending with 'Apple'. Assume a support threshold of 25%	5+2
	c)	For the following data, apply the k-means clustering algorithm and form 2 clusters. (0,0), (0,1), (1,0), (1,1), (4,4), (4,5), (5,4), (5,5). Initial centroids are (0,0) and (0,1). Which cluster has the lower mean square distance? Note: Show calculations clearly. Use Euclidean distance measure	4+1
5	a)	Briefly explain how SVD can be used for Data Compression. Perform SVD on the following matrix and express it as a product of 3 matrices. $A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix}$	3+5
	b)	With a schematic diagram clearly explain the main components of a Convolutional Neural Network.	2+4
	c)	Explain the architecture of a Generative Adversarial Network. What is the main challenge when implementing GANs?	4+2