PES University, Bengaluru

(Established under Karnataka Act No. 16 of 2013)

UE15CS353/UE17CS303

DECEMBER 2020: END SEMESTER ASSESSMENT (ESA) B TECH
SEMESTER 5th

UE15CS353/UE17CS303 -Machine Learning

Time: 3 Hrs	Answer All Questions	Max Marks: 100

1 a) Differentiate Classification Versus Prediction with an Example. **6**

1 b) **10**

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

Given above dataset do the following

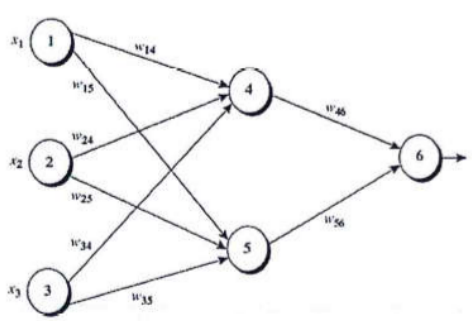
1. Find Information Gain for all the attributes.
2. Construct Decision Tree for the above Dataset.

1 c) Mention the drawbacks of ID3 Algorithm(Decision Trees). **4**

2 a) Perfom KNN Classification on following dataset and predict class for x(Attribute1=3, Atribute2=7) where k=3. **7**

Attriubte1	Attribute2	class
7	7	False
7	4	False
3	4	True
1	4	True

2 b) **10**



Initial Input, Weight, and Bias Values

x_1	x_2	x_3	w_{14}	w_{15}	w_{24}	w_{25}	w_{34}	w_{35}	w_{46}	w_{56}	θ_4	θ_5	θ_6
1	0	1	0.2	−0.3	0.4	0.1	−0.5	0.2	−0.3	−0.2	−0.4	0.2	0.1

Given above Artificial Neural Network with Initial Input, Weight, Bias Values. Calculate the error at each node i.e node 6,node 5,node 4.**(Only error at each node to be calculated)**

2 c) List 3 Activation functions with their equations. **3**

3	a)		12

age	Income	Student	Credit rating	buys_computer
<=30	high	n	f	n
<=30	high	n	e	n
31...40	high	n	f	y
>40	medium	n	f	y
>40	low	y	f	y
>40	low	y	e	n
31...40	low	y	e	y
<=30	medium	n	f	n
<=30	low	y	f	y
>40	medium	y	f	y
<=30	medium0	y	e	y
31...40	medium	n	e	y
31...40	high	y	f	y
>40	medium	n	e	n

Given attributes are age, income, student, credit rating, buys_computer in the above table.**(f-fair,e-excellent, n-no,y-yes)**
Given Class attribute is:

C1:buys_computer = 'yes'
C2:buys_computer = 'no'

Compute P(X|Ci) for each class. Check whether the sample given below belongs to yes or no class using Bayesian Classification Technique.

X = (age <=30, Income = medium, Student = yes Credit_rating = Fair)

	b)	Define Boosting and Bagging briefly? How they can improve the performance?	8
4	a)	A database has 5 transactions. Let min sup = 60% and min confidence = 80%. TID items bought T100 {M, O, N, K, E, Y} T200 {D, O, N, K, E, Y } T300 {M, A, K, E} T400 {M, U, C, K, Y} T500 {C, O, O, K, I ,E} Find all frequent item sets using Apriori and FP-growth, respectively. Compare the efficiency of the two mining processes.	12
	b)	Consider the following data set consisting of the scores of two variables on each of seven subjects in the range 1 to 7.	8

Subject	A	B
1	1.0	1.0
2	1.5	2.0
3	3.0	4.0
4	5.0	7.0
5	3.5	5.0
6	4.5	5.0
7	3.5	4.5

This data set is to be grouped into two clusters. As a first step 1 and 4 define the initial cluster means. Using K – means algorithm generate the initial and final clusters.

5	a)	List all the steps briefly in PCA.	5
	b)	Explain Fitness function, selection methods, crossover, mutation in Genteic Algoirthms	8
	c)	Can SVD be used for rectangular matrices? List some of the applications of SVD.	7