

BTech (CS/IT) VI Semester Examination, 2020
CS 317: Artificial Intelligence and Machine Learning

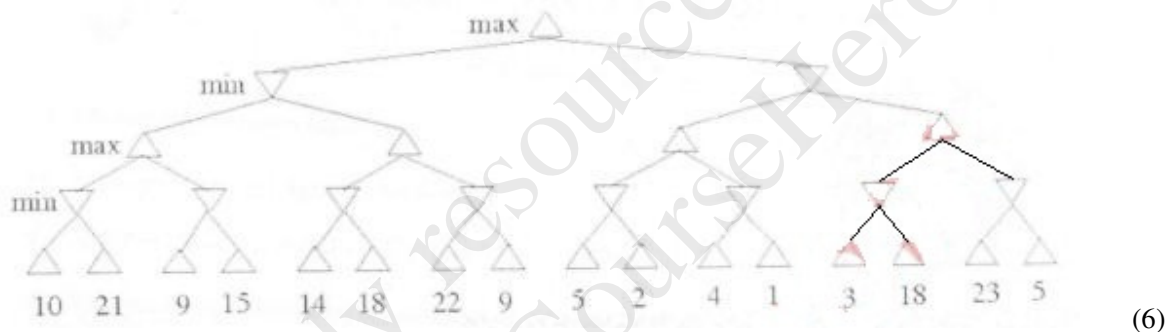
Maximum Marks - 60

*The question paper is divided into three Sections. Students are required to attempt **five** questions in all, selecting **not more than two** questions from each Section.*

All questions carry equal marks.

SECTION A

1. (a) After the initial setbacks, researchers observed that intelligence requires knowledge. What disadvantages does knowledge possess? (6)
(b) What do you understand by informed state space? (3)
(c) State the conditions under which A* is admissible (3)
2. (a) Apply Alpha-Beta pruning on following example considering first node as MAX. (6)



- (b) Why are data reduction techniques applied on a given dataset? How can we handle missing values in a dataset (6)
3. Consider the following statements:
(i) Ram was a man
(ii) Valmiki was a saint
(iii) Ravan was a ruler
(iv) All Hindus were either loyal to Ravan or hated him
(v) Everyone is loyal to someone
(vi) People only try to assassinate rulers they are not loyal to
(vii) All saints are Hindu
(viii) Ram tried to assassinate Ravan
(a) Convert these statements into facts and rules (6)
(b) Using resolution answer the question "Did Ram hated Ravan" (6)

SECTION B

4. (a) For a SunBurn dataset given below, construct a decision tree.

Name	Hair	Height	Weight	Location	Class
------	------	--------	--------	----------	-------

Name	Hair	Height	Weight	Location	Class
Sunita	blonde	average	light	no	yes
anit	blonde	tall	average	yes	no
kavita	brown	short	average	yes	no
sushma	blonde	short	average	no	yes
xavier	red	average	heavy	no	yes
balaji	brown	tall	heavy	no	no
ramesh	brown	average	heavy	no	no
swetha	blonde	short	light	yes	no

In this dataset, Name, Hair, Height, Weight and Location are Dependent Variables (Predictors) and Class is the independent variable (target). (6)

(b) The values of independent variable x and dependent value y are given below:

X	Y
1	0.0319
0	0.8692
1	1.9566
2	3.0343

Find the least square regression line $y=ax+b$. Estimate the value of y when x is 10. (6)

5 (a) Consider the following dataset of stolen vehicles. Colour, Type and Origin are dependent variables and Stolen is an independent variable.

Colour	Type	Origin	Stolen
Red	Sports	Domestic	Yes

Red	Sports	Domestic	No
Red	Sports	Domestic	Yes
Yellow	Sports	Domestic	No
Yellow	Sports	Imported	Yes
Yellow	SUV	Imported	No
Yellow	SUV	Imported	Yes
Yellow	SUV	Domestic	No
Red	SUV	Imported	No
Red	Sports	Imported	Yes

Using Naïve Bayes classifier, classify the new data (Colour=Red, Type=SUV, Origin=Domestic). (6)

- (b) Find out a problem which cannot be solved easily using the architectures familiar to you. Suggest an architecture to solve this. Why do you think that this architecture can solve the problem? Suggest a procedure to derive the learning rule(s). (6)

- 6 (a) Consider the following dataset for loan classification. Age and Loan Amount are dependent variables and Default? is an independent variable.

Age	Loan	Default?
25	40000	N
35	60000	N
45	80000	N
20	20000	N
35	120000	N
52	18000	N
23	95000	Y
40	62000	Y
60	100000	Y
48	220000	Y
33	150000	Y

Using a K-Nearest Neighbour classifier, classify the new data (Age=47 and Loan Amount = Rs 132000). Use Euclidian Distance for computation. Classify the new data for 1 and 5 nearest neighbours. (6)

- (b) Consider the following confusion matrix for 3 variables

		Predicted		
		A'	B'	C'
Actual	A	25	5	2
	B	3	32	4
	C	1	0	15

Calculate Precision, Recall, F-Measure and Accuracy for Class B (6)

SECTION C

7. (a) Enumerate the modalities of Reinforcement Learning for a Game of Chess (6)
 (b) Why do we need to perform cross-validation while training our models? How can it improve the performance of the models? (6)

8. (a) Consider the following data.

2, 11, 8, 13, 10, 2, 13, 5, 4, 12, 6, 10, 11, 6, 7, 13, 1

Using K-Means Clustering, categorize this data into 3 clusters. Assume random initial centroid values and Euclidian distance function. Show the working for each iteration. (6)

(b) Consider the following dataset

X1	X2
12	10
15	14
13	11

Find the distance between X and Y data points where $Y_1=9$ and $Y_2=12$. Compute the distance for Manhattan Distance and Makowski Distance with $q=5$. (6)

9. Consider the following corpus

<s>Banasthali Vidyapith is a University with a difference </s>

<s> This is a fully residential University </s>

<s>Banasthali Vidyapith is in India </s>

(a) Generate a Markov Chain of Order 1 for the corpus (8)

(b) Check the validity of the sentence <s>**Banasthali Vidyapith is a fully residential University in India** </s> using the model developed (4)