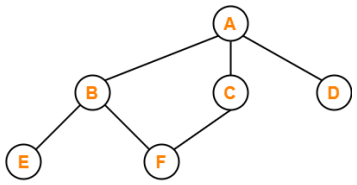
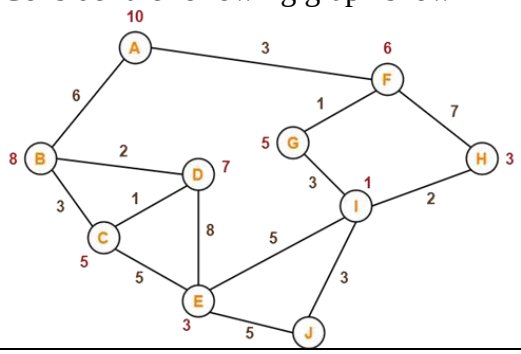
UG

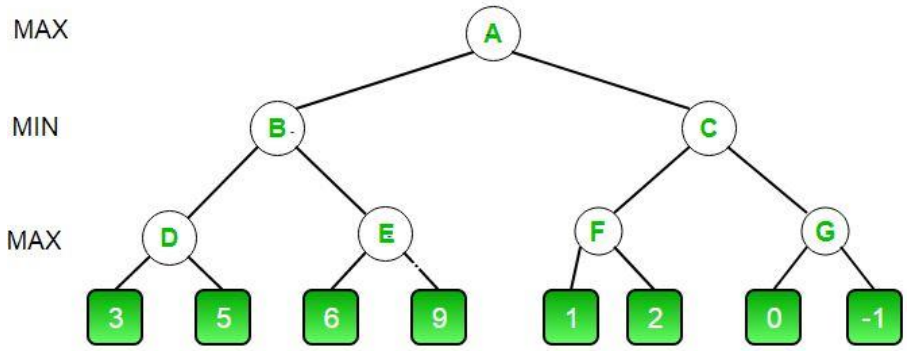
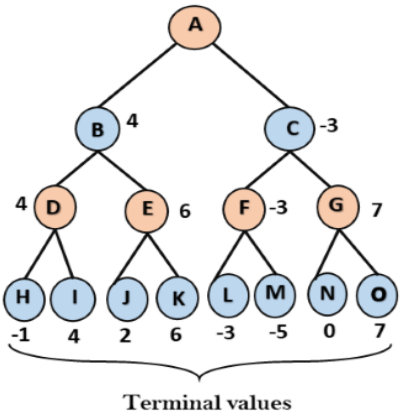
RV College of Engineering®, Bengaluru-560 059
Autonomous Institution affiliated to Visvesvaraya Technological University

Manuscript of Question Paper

COURSE CODE:21AI52	SEM: V
COURSE TITLE : Artificial Intelligence and Machine Learning	
Duration of Paper: 03 Hrs	

Instructions to Candidates:
1. Answer all questions from Part A
2. Any 5 Full questions from Part B choosing one from each side. (Question No.2 is compulsory)

Question No	PART A	Marks
1.1	State two advantages of DFS?	02
1.2	How 8 queens problem can be viewed as a Constraint Satisfaction Problem?	02
1.3	Define Hill Climbing problem?	02
1.4	Define two parameters in Alpha Beta Pruning?	02
1.5	State any two advantages of Decision Tree?	02
1.6	State two reasons for model over fitting?	02
1.7	State two methods of estimating the probability of data when it has continuous attributes?	02
1.8	State two necessary condition for an ensemble classifier?	02
1.9	State any two data characteristics that affect the proximity in unsupervised learning?	02
1.10	State two limitations of K-means algorithm?	02
	PART B	
2(a)	Differentiate between model based reflex agents and Goal based reflex agents with an example?	**6**
2(b)	Apply BFS method for the following graph? Give its advantages and disadvantages? 	**10**
3(a)	Consider the following graph shown in figure 3.a: 	**6**

	Find the most cost-effective path to reach from start state A to final state J using A* Algorithm.	
3(b)	For the following two-ply game tree shown in figure 3b., the terminal nodes show the utility values computed by the utility function. Use the Min-Max algorithm to compute the utility values for other nodes in the given game tree.  	**10**
	OR	
4(a)	Discuss how gaming can be considered as a search strategy?	06
4(b)	Apply min-max algorithm for the given graph?  	10
5(a)	Consider the given dataset given below. Write a pseudo code and Illustrate the Decision Tree algorithm to build a model using Information gain as a metric to split the attributes?	10

Day	Outlook	Temp	Humidity	Wind	Play Volleyball
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	Rain	Mild	High	Strong	No
D13	Overcast	Hot	Normal	Weak	Yes
D14	Overcast	Mild	High	Strong	Yes

5(b)	Summarize the splitting of attributes based on continuous Attributes?	06
	OR	
6(a)	Consider the given dataset given below. Write a pseudo code and Illustrate the	06

	Decision Tree algorithm to build a model using Gini as a metric to split the attributes? 	Age	City	Income Group	Car
---	---	---	---
Medium	Rural	Normal	Yes
Medium	Rural	Normal	No
Eleder	Metro	High	Yes
Medium	Urban	High	Yes
Young	Metro	High	No
Eleder	Metro	Normal	Yes
Young	Urban	High	No
Young	Rural	Normal	Yes		
6(b)	Discuss the characteristics of decision tree?	06
7(a)	Given a dataset, train a K-nearest neighbors (KNN) model using the KNN algorithm with K=3 and Euclidean distance. Additionally, provide a pseudo code for the training process. After training the model, predict its response for a specific test example? **Test Example: User ID - 15621045, Gender – Female, Age – 20, Purchased-?** 	User ID	Gender	Age	EstimatedSalary	Purchased
---	---	---	---	---
15624510	Male	19	19000	0
15810944	Male	35	20000	0
15668575	Female	26	43000	0
15603246	Female	27	57000	0
15804002	Male	19	76000	0
15728773	Male	27	58000	0
15598044	Female	27	84000	0
15694829	Female	32	150000	1
15600575	Male	25	33000	0
15727311	Female	35	65000	0
15570769	Female	26	80000	0
15606274	Female	26	52000	0
15746139	Male	20	86000	0
15704987	Male	32	18000	0
15628972	Male	18	82000	0
15697686	Male	29	80000	0
15733883	Male	47	25000	1
15617482	Male	45	26000	1
15704583	Male	46	28000	1
15621083	Female	48	29000	1
15649487	Male	45	22000	1
15736760	Female	47	49000	1		10
7(b)	Discuss different methods to improve KNN efficiency?	6
	OR	
8(a)	Consider the data set shown below: 	Instance	A	B	C	Class
---	---	---	---	---
1	0	0	1	−
2	1	0	1	+
3	0	1	0	−
4	1	0	0	−
5	1	0	1	+
6	0	0	1	+
7	1	1	0	−
8	0	0	0	−
9	0	1	0	+
10	1	1	1	+	 (a) Estimate the conditional probabilities for P(A = 1|+), P(B = 1|+), P(C = 1|+), P(A = 1|−), P(B = 1|−), and P(C = 1|−) (b) Use the conditional probabilities in part (a) to predict the class label for a test sample (A = 1, B = 1, C = 1) using the naive Bayes approach.	10

<table>
<tr><td></td><td>(c) Compare P(A = 1), P(B = 1), and P(A=1, B = 1). State the relationships between A and B.</td><td></td></tr>
<tr><td>8(b)</td><td>Discuss the different ways to constructing the Ensemble Classifiers?</td><td>06</td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td>9(a)</td><td>Discuss the role of clustering in Machine Learning? Discuss the different types of clusters in Machine Learning</td><td>10</td></tr>
<tr><td>9(b)</td><td>Demonstrate the K-Means algorithm for the following data considering K =3
<table>
<tr><td>Point</td><td>Coordinates</td></tr>
<tr><td>A1</td><td>(2,10)</td></tr>
<tr><td>A2</td><td>(2,6)</td></tr>
<tr><td>A3</td><td>(11,11)</td></tr>
<tr><td>A4</td><td>(6,9)</td></tr>
<tr><td>A5</td><td>(6,4)</td></tr>
<tr><td>A6</td><td>(1,2)</td></tr>
<tr><td>A7</td><td>(5,10)</td></tr>
<tr><td>A8</td><td>(4,9)</td></tr>
<tr><td>A9</td><td>(10,12)</td></tr>
<tr><td>A10</td><td>(7,5)</td></tr>
<tr><td>A11</td><td>(9,11)</td></tr>
<tr><td>A12</td><td>(4,6)</td></tr>
<tr><td>A13</td><td>(3,10)</td></tr>
<tr><td>A14</td><td>(3,8)</td></tr>
<tr><td>A15</td><td>(6,11)</td></tr>
</table></td><td>6</td></tr>
<tr><td></td><td>OR</td><td></td></tr>
<tr><td>10(a)</td><td>Discuss the need for Bisecting K –means and write the pseudo code for the same?</td><td>10</td></tr>
<tr><td>10(b)</td><td>Discuss the different measures used in Measuring the clustering validity?</td><td>6</td></tr>
</table>