	b	Consider the data set shown in Table 7b. Using the naive Bayes approach, predict the class label "species" for a test sample X with the following attributes X = {Color=Green, Legs=2, Height=Tall, Smelly=No}					

Table 7b

Sl. No	Color	Legs	Height	Smelly	Species
1	White	3	Short	Yes	M
2	Green	2	Tall	No	M
3	Green	3	Short	Yes	M
4	White	3	Short	Yes	M
5	Green	2	Short	No	H
6	White	2	Tall	No	H
7	White	2	Tall	No	H
8	White	2	Short	Yes	H

		OR					
8	a	What are ensemble learning models? Explain bagging and boosting in detail.	9	1	1		
	b	Describe the characteristics of logistic regression.	7	2	2		
9	a	Distinguish the following types of clustering: i) Hierarchical versus Partitional ii) Exclusive versus Overlapping versus Fuzzy iii) Complete versus Partial	6	2	2		
	b	Describe how empty clusters, outliers, centroid updates are handled in k-means clustering algorithm.	6	1	5		
	c	Write a note on cluster tendency.	4	3	3		
		OR					
10	a	How to determine the correct or natural number of clusters in the dataset? Explain.	4		1		
	b	Explain different classification-oriented measures of cluster validity.	5		2		
	c	What are the strengths and weaknesses of k-means algorithm? Apply k(=2)-means algorithm over the data (185, 72), (170, 56), (168, 60), (179, 68), (182, 72), (188, 77) up to two iterations and show the clusters. Initially choose two objects as initial centroids and use Euclidean distance as distance measure.	7		3		

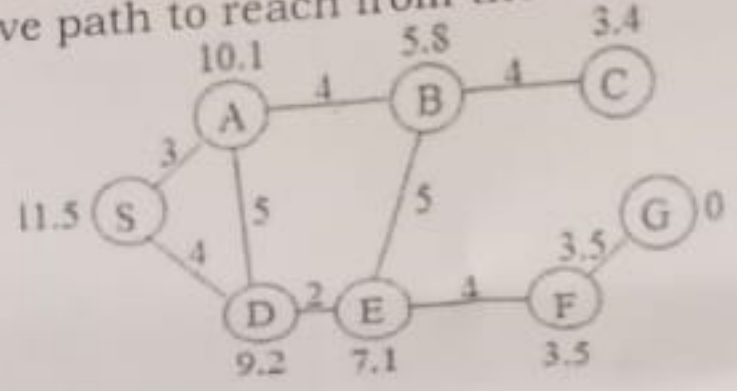
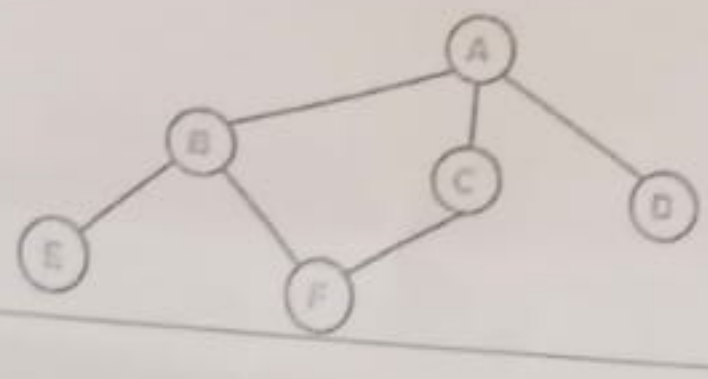
4	a	Apply A* algorithm to the graph shown in Fig 4a to find the cost-effective path to reach from the start state S to goal state G.  Fig 4a	6	3	2
	b	Explain the following local search algorithms: i) Simulated annealing ii) Hill Climbing Search	10	2	5
5	a	Discuss methods for expressing attribute test conditions with examples.	6	1	2
	b	Explain any four characteristics of decision tree classifiers.	6	2	1
	c	Write a note on cross-validation model evaluation technique.	4	2	3
		OR			
6	a	Explain the classification model selection approach to estimate the generalization performance of a model.	8	2	1
	b	Consider the training examples shown in Table 6b for a binary classification problem. i) What is the entropy of this collection of training examples with respect to the class attribute? ii) What are the information gains of a1 and a2 relative to these training examples? iii) For a3, which is a continuous attribute, compute the information gain for every possible split. iv) What is the best split (among a1, a2, and a3) according to the information gain? v) What is the best split (between a1 and a2) according to the Gini index?			

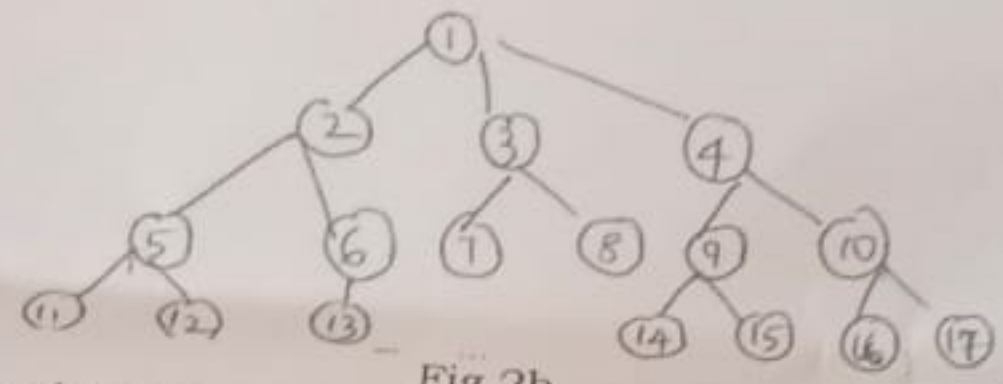
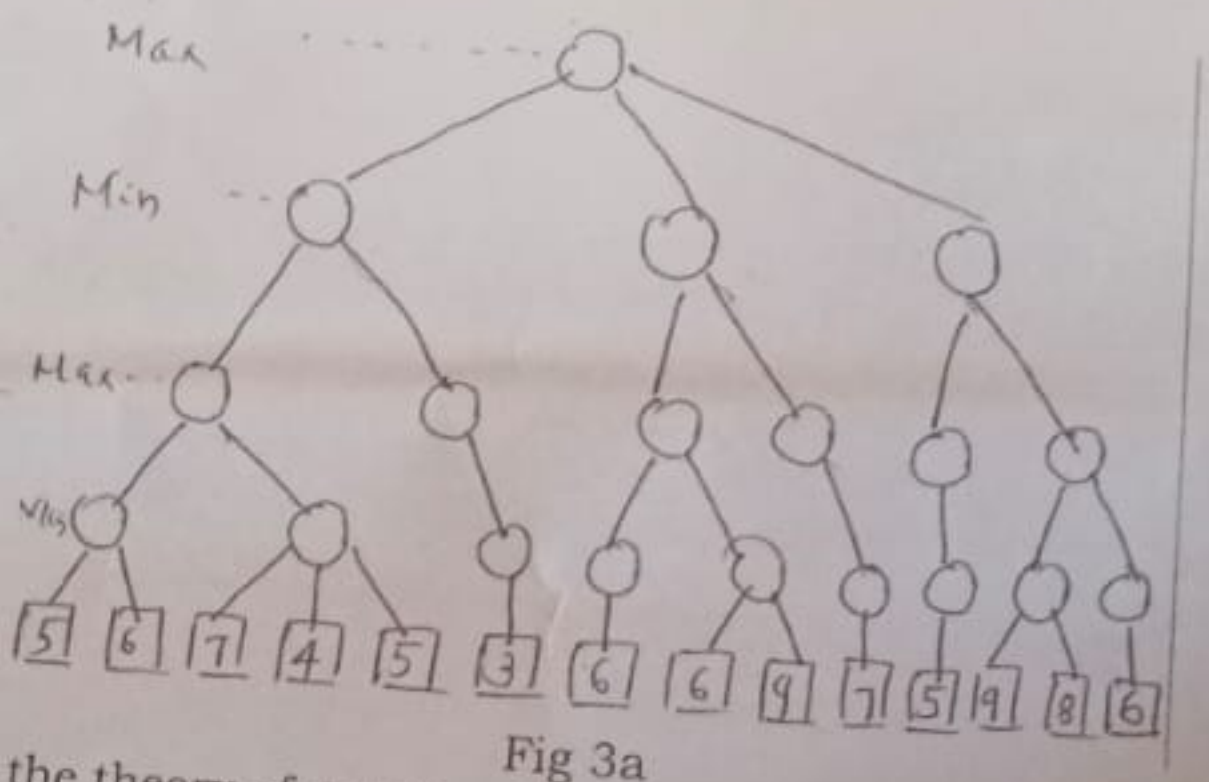
Table 6b

Instance	a1	a2	a3	Target Class
1	T	T	1.0	+
2	T	T	6.0	+
3	T	F	5.0	-
4	F	F	4.0	+
5	F	T	7.0	-
6	F	T	3.0	-
7	F	F	8.0	-
8	T	F	7.0	+
9	F	T	5.0	-

				8	3
7	a	Write the k-nearest neighbor classifier algorithm. Describe its characteristics.	8	2	

1.14	Apply *BFS* method for the following graph shown in fig 1.14 and write order in which vertices are visited.  		2	3	2

PART-B

2	a	Explain the following agent programs i) Model-based reflex agents ii) Goal-based agents	5	2	2
	b	Apply the following uninformed search strategies for the state space shown in Fig 2b. Suppose the goal state is 15, list the order in which nodes will be visited. i) Depth-limited search with depth limit 3 ii) Iterative deepening depth-first search Assume: Root is at depth 0  Fig 2b			
	c	Discuss any three properties of task environments.	5	3	1
3	a	Explain alpha-beta pruning. Apply alpha-beta pruning algorithm to prune the game tree shown in Fig 3a. Assuming child nodes are visited from left to right. Show all final alpha and beta values computed at root, each internal node explored, and at the top of pruned branches. Also show the pruned branches.  Fig 3a	6	4	3
	b	Explain the theory of genetic algorithm with an example.	10 6	3 2	3 1

OR

21AI52

USN | | | | | | | | 0 | 5 | 7 |

RV COLLEGE OF ENGINEERING®

(An Autonomous Institution affiliated to VTU)

V Semester B. E. Examinations April/May -2024

Common to CS /IS / AIML

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Time: 03 Hours *Maximum Marks: 100*

Instructions to candidates:

1. Answer all questions from Part A. Part A questions should be answered in first three pages of the answer book only.
2. Answer FIVE full questions from Part B. In Part B question number 2 is compulsory. Answer any one full question from 3 and 4, 5 and 6, 7 and 8, 9 and 10.

PART-A

		M	BT	CO
1.1	_____ agent programs select actions on the basis of the current percept, ignoring the rest of the percept history.	1	2	2
1.2	Differentiate between episodic agent environment and sequential agent environment.	2	1	1
1.3	List the different ways to evaluate the performance of a search algorithm.	1	1	2
1.4	Iterative deepening algorithm is a combination of _____ and _____ algorithms.	1	2	5
1.5	_____ algorithm is a local search algorithm that moves continuously upward (increasing) until the best solution is attained.	1	1	1
1.6	Differentiate between informed search and uninformed search algorithms.	2	2	3
1.7	How classification model can be used as predictive and descriptive model?	2	2	5
1.8	Entropy value of _____ represents that the data sample is pure and entropy value of _____ represents that the data sample has a 50 – 50 split belonging to two categories.	1	1	2
1.9	What is the time complexity of the K-means clustering algorithm?	1	1	3
1.10	The value of the silhouette coefficient can vary between _____ and _____.	1	3	1
1.11	_____ algorithm combines the predictions made by multiple decision trees, where each tree is generated based on the values of an independent set of random vectors.	1	2	5
1.12	Differentiate between the pre-pruning and post-pruning strategies for decision tree induction.	2	1	2
1.13	Mention two reasons for model over fitting.	2	2	1