USN

RV COLLEGE OF ENGINEERING
Autonomous Institution affiliated to VTU
V Semester B.E.
Model Question Paper
DEPARTMENT OF INFORMATION SCIENCE AND ENGINEERING

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING
(2022 SCHEME) Model Question Paper

Time: **03 Hours**   *Maximum Marks:* **100**

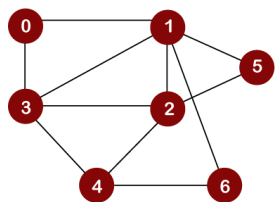### Instructions to candidates:
1. Answer all questions from Part A. Part A questions should be answered in first three pages of the answer book only.
2. Answer FIVE full questions from Part B. In Part B question number 2 is compulsory. Answer any one full question from 3 and 4, 5 and 6, 7 and 8, and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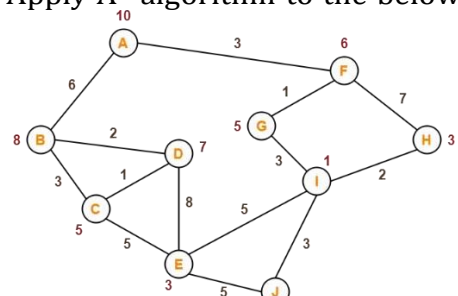
PART-A

1	1.1	What is the primary advantage of depth-first search (DFS) over breadth-first search?	01
	1.2	Differentiate between a rational agent and a non-rational agent.	02
	1.3	A* search considered optimal when the heuristic is admissible. Give reason	01
	1.4	In a two-player game, how does the minimax algorithm ensure that both players play optimally?	02
	1.5	How does simulated annealing differ from hill-climbing when reaching a local optimum?	01
	1.6	In K-means, the objective is to minimize the _____ within each cluster, which is the sum of squared distances between data points and their cluster centroids.	01
	1.7	In Logistic Regression, the parameters are learned by minimizing the _____ function, typically using gradient descent.	01
	1.8	Bagging improves the stability of a classifier by generating multiple datasets through _____ sampling and training a separate model on each dataset.	02
	1.9	In a decision tree, the Gini index of a split is 0.5, and in another split, it is 0.3. Which split is considered better for classification? Give reason	02
	1.10	What are resubstitution errors ?	02
	1.11	When pruning a decision tree, why is it important to balance model complexity and accuracy?	01
	1.12	Ensemble methods combine multiple individual models to create a _____ model that typically outperforms any individual model	01
	1.13	In a Naive Bayes classifier, the likelihood of observing a feature value given a class is calculated using _____ probability.	01

	1.14	A major disadvantage of the Nearest Neighbor Classifier is that it can become computationally expensive as the number of _____in the dataset increases.	01
	1.15	To improve the accuracy of K-means clustering, it may be helpful to apply _____to the dataset before clustering to normalize the data.	01

PART-B

2	a	Explain how a problem-solving agent works. Describe the steps it follows to solve a problem, using an example of a typical problem like route planning	6
	b	Traverse the given graph using BFS and DFS and compare the results	10

3	a	Compare and contrast A* search with Greedy Best-First Search. Use examples to illustrate their differences.	10
	b	Apply A* algorithm to the below given graph and find the optimal path	6
		OR	
4	a	What are local search algorithms? Identify and explain using examples, the key features that distinguish them from classical search methods.	10
	b	Critically evaluate the effectiveness of simulated annealing in finding global optima for complex search spaces.	6

5	a	Construct a decision tree for the dataset given using ID3 algorithm	

Outlook	Temperature	Humidity	Wind	Play Golf
Sunny	Hot	High	Weak	No
Sunny	Hot	High	Strong	No
Overcast	Hot	High	Weak	Yes
Rain	Mild	High	Weak	Yes
Rain	Cool	Normal	Weak	Yes
Rain	Cool	Normal	Strong	No
Overcast	Cool	Normal	Strong	Yes
Sunny	Mild	High	Weak	No
Sunny	Cool	Normal	Weak	Yes
Rain	Mild	Normal	Weak	Yes
Sunny	Mild	Normal	Strong	Yes
Overcast	Mild	High	Strong	Yes
Overcast	Hot	Normal	Weak	Yes
Rain	Mild	High	Strong	No

10

	b	How does overfitting affect the generalizability of a model. Provide examples to illustrate.	06
		OR	
6	a		Predicted: Pass	Predicted: Fail
---	---	---
Actual: Pass	50	10
Actual: Fail	5	35 | For the confusion matrix given above , calculate the accuracy, precision, F1 Score and Recall. Suggest the relevance of each metric	08
	b	Assess the trade-off between model complexity and performance in decision trees, providing examples.	08

7	a	What are the characteristics of nearest neighbor classifiers. How does the value of k influence the bias-variance tradeoff? Provide examples for different values of k.	08
	b	Compare Naive Bayes and k-NN classifiers in terms of their assumptions, computational requirements, and performance in imbalanced datasets.	08
		OR	
8	a	Give the general framework of ensemble learning. How do bagging and boosting differ in terms of how they train base models?	08
	b	Compare logistic regression with decision trees and support vector machines in terms of interpretability, training time, and performance on linearly separable data.	08

9	a	Compare and contrast different types of clustering approaches, such as hierarchical, partitioning, density-based, and grid-based clustering.	08
	b	Provide a comparative analysis of K-means clustering and other clustering approaches like hierarchical clustering and DBSCAN in terms of scalability, accuracy, and applicability.	08
		OR	
10	a	Explain the strengths and weaknesses of K-means clustering. Provide suggestions for improving the algorithm's robustness in real-world applications.	08
	b	Suggest how assessing the Significance of Cluster Validity Measures can be done	08

Signature of Scrutinizer: Signature of Chairman
Name: Name: