**RV COLLEGE OF ENGINEERING®**

**(An Autonomous Institution affiliated to VTU)**

**IV Semester B. E. Sept – 2024 Examinations**

**Department of Artificial Intelligence and Machine Learning**

**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 ALL full questions from part B.

|  |  |  |
| --- | --- | --- |
| **S. No** | **PART A** | **Marks** |
| 1.1 | Identify the various capabilities that a computer has to possess to pass Turing test | 01 |
| 1.2 | Define Rational Agent (RA) and discuss the usage of RA in Healthcare applications | 02 |
| 1.3 | For each of the following agents, develop a PEAS description of the task  environment:   1. Robot soccer player 2. Internet book-shopping agent | 02 |
| 1.4 | For the real time “AIRLINE” travel problem, identify the five components involved in problem formulation in-terms of problem-solving agents. | 01 |
| 1.5 | Consider the following state space shown in Fig.1.5. Suppose the goal state is 11. List the order in which nodes will be visited for breadth-first search and depth-limited search with limit 3.    Fig.1.5: State space | 02 |
| 1.6 | Analyze and identify issues in following machine learning scenarios and propose solution(s) for them:   1. A student dataset has an employee age value as 230. 2. An object recognition model training set error is 95% and test set error is 15%. | 02 |
| 1.7 | The average squared difference between classifier predicted output and actual output is called as\_\_\_\_\_\_\_\_\_ | 01 |
| 1.8 | **In a decision tree algorithm, how can you reduce the level of entropy from the root node to the leaf node?** | 01 |
| 1.9 | **In the decision tree, the measure of the degree of probability of a particular variable being wrongly classified when it is randomly chosen is called \_\_\_\_\_.** | 01 |
| 1.10 | While predicting the spam email using a classification model following data are recorded.  i. Correct prediction – 0 spam, 75 non-spam  ii. Incorrect prediction- 30 spam, 700 non-spam  Calculate Accuracy and sensitivity. | 02 |
| 1.11 | What are some of the hyperparameters of the random forest regressor that help prevent overfitting? | 02 |
| 1.12 | There are a number of choices for the proximity function, centroid, and objective function that can be used in the basic K-means algorithm and that are guaranteed to converge. What are they? | 02 |
| 1.13 | How do you handle missing or corrupted data in a dataset? | 01 |
|  | **PART B** |  |
| 2a. | Explain how AI works in the following areas:   1. Google Search Engine 2. Voice Assistants | 04 |
| 2b. | Justify with reasons how utility-based agents is different from model- based reflex agents with appropriate diagrams. | 06 |
| 2c. | Apply the Depth first search algorithm for the following graph shown in Fig. 2c.    Fig. 2c   1. Display the output visiting each node, and 2. Develop the DFS spanning tree. | 06 |
| 3a. | Consider the following graph shown in Fig. 3a.    Fig. 3a  Find the most cost-effective path to reach from start state A to final state J using A-star Algorithm. | 06 |
| 3b. | For the following two-ply game tree shown in Fig. 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.    Fig. 3b | 10 |
|  | **OR** |  |
| 4a. | Given an initial state of 8-puzzle problem and final state to be reached-    Fig. 4a  Find the most cost-effective path to reach the final state from initial state using A-star Algorithm for the Fig. 4a. | 10 |
| 4b. | With neat sketches discuss the following classical search algorithms:   1. Hill climbing 2. Simulated Annealing | 06 |
|  |  |  |
| 5a. | Identify the key differences between predictive and descriptive machine learning modeling techniques in terms of approach, applications, algorithms used, evaluation measures. | 04 |
| 5b. | Illustrate the bias-variance trade-off issue in machine learning. Identify the techniques used to avoid high bias and high variance scenarios. | 06 |
| 5c. | Analyze and identify issues in the following machine learning scenarios and propose solutions (s) for them   1. A student dataset has an employee age value of 230. 2. An object recognition model training set error is 95%, and the test set error is 15%. 3. Malignant tumour classification model training set error is 25%, and test set error is 23%. | 06 |
|  | **OR** |  |
| 6a. | For instance, the following table informs about decision making factors to play tennis at outside for previous 14 days:   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | Day | Outlook | Temp | Humidity | Wind | Decision | | 1 | Sunny | Hot | High | Weak | No | | 2 | Sunny | Hot | High | Strong | No | | 3 | Overcast | Hot | High | Weak | Yes | | 4 | Rain | Mild | High | Weak | Yes | | 5 | Rain | Cool | Normal | Weak | Yes | | 6 | Rain | Cool | Normal | Strong | No | | 7 | Overcast | Cool | Normal | Strong | Yes | | 8 | Sunny | Mild | High | Weak | No | | 9 | Sunny | Cool | Normal | Weak | Yes | | 10 | Rain | Mild | Normal | Weak | Yes | | 11 | Sunny | Mild | Normal | Strong | Yes | | 12 | Overcast | Mild | High | Strong | Yes | | 13 | Overcast | Hot | Normal | Weak | Yes | | 14 | Rain | Mild | High | Strong | No | | 1. Calculate the IG (information gain) for each attribute (feature) 2. Choose a feature for each node to split 3. Build Final Tree | | | | | | | 10 |
| 6b. | The following data are recorded while predicting the spam email using a classification model.   1. Correct prediction – 0 spam, 75 non-spam 2. Incorrect prediction- 30 spam, 700 non-spam   Calculate the Accuracy, sensitivity, precision, and F-measure of the model. Identify the problem with the above classification model. | 06 |
| 7a. | With relevant mathematical equations discuss Naïve Bayes Classification. Consider the following data in table 7a. How would a Naive Bayes classifier used to classify the sample X (Cow, Medium, Black)? Write step-by-step approach to solution.  Table 7a. Dataset  IMG_256 | 10 |
| 7b. | For the given dataset shown in Table 7b, apply Logistic Regression.  Table 7b.   |  |  | | --- | --- | | Hours Study | Pass (1)/Fail (0) | | 29 | 0 | | 15 | 0 | | 33 | 1 | | 28 | 1 | | 39 | 1 |  1. Assume the model suggested by the optimizer for odds of posing the course is: log(odds) = -64+2\*hours. 2. Given the class, calculate the probability of pass percentage for the student who studies for 33 hours. 3. Compute the minimum number of hours of preparation student should do to pass the course with the probability of > 95%. | 06 |
|  | **OR** |  |
| 8a. | Reflect on working of KNN algorithm. Consider the data provided in Table 8a, using KNN algorithm with value of k = 3, find the weight for data: Height= 5.5 and Age= 38. Write step-by-step approach to solution.  Table 8a dataset | 10 |
| 8b. | With neat sketches, explain boosting and bagging ensemble learning techniques. | 06 |
| 9a. | Distinguish between the following:   1. Hierarchical versus Partitional 2. Exclusive versus Overlapping versus Fuzzy 3. Complete versus Partial | 09 |
| 9b. | Illustrate and explain Bisecting K-means algorithm. | 07 |
|  | **OR** |  |
| 10 a. | Given fifteen points in the Cartesian coordinate system and the requirement to form 3 clusters (K=3), if 3 centroids are randomly selected from the data—specifically the points A2 (2, 6), A7 (5, 10), and A15 (6, 11) respectively, how would the process of clustering proceed over 3 iterations?   |  |  |  |  | | --- | --- | --- | --- | | Point | Coordinates | Point | Coordinates | | A1 | (2,10) | A8 | (4,9) | | A2 | (2,6) | A9 | (10,12) | | A3 | (11,11) | A10 | (7,5) | | A4 | (6,9) | A11 | (9,11) | | A5 | (6,4) | A12 | (4,6) | | A6 | (1,2) | A13 | (3,10) | | A7 | (5,10) | A14 | (3,8) | |  |  | A15 | (6,11) | | 10 |
| 10 b. | How can cohesion and separation metrics be used to evaluate the quality of clusters in an unsupervised learning context? | 06 |

Signature of Scrutinizer: Signature of Chairman

Name: Name: