Reg. No. :

B.E. / B.Tech. DEGREE EXAMINATIONS, NOV/ DEC 2024

Fifth Semester

Artificial Intelligence and Data Science

U20AI502 – ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

(Regulation 2020)

Time: Three Hours   Maximum: 100 Marks

Answer ALL questions

PART – A   (10 x 2 = 20 Marks)

1.   Define omniscient agent. Give example.

2.   Compare informed and uninformed search.

3.   Infer about the term 'fact' in Prolog programming?

4.   State the role of ontology in knowledge representation?

5.   Compare supervised and unsupervised learning.

6.   Inscribe about the purpose of the least squares method in regression?

7.   Justify why slack variables are used in soft-margin SVM.

8.   Summarize the role of the confusion matrix in selecting a classification threshold.

9.   Mention the significance of transition probabilities in an HMM.

10. Identify the impact of dimensionality reduction on model performance.

11. (a) Enumerate the components of a problem? Narrate them in brief. Write the problem formulation for 8 puzzle problem solving agent.   (16)

(OR)

(b) Consider the graph given below and find the total cost incurred using best first search. Assume A as the source node and H as the goal node. The heuristic values for the corresponding nodes are given in the table.   (16)
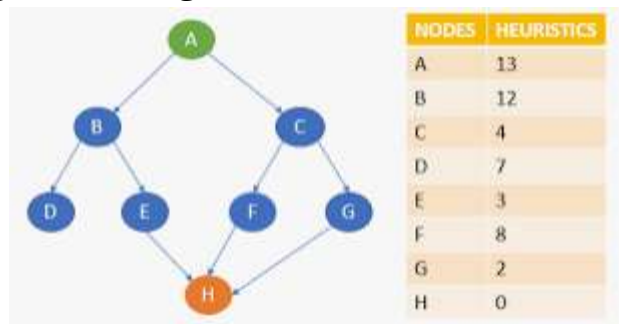
12. (a) Evaluate the efficiency of forward chaining versus backward chaining. Discuss factors that influence the performance of each method, including time complexity and space complexity. Use examples to illustrate your points.   (16)

(OR)

(b) Compare and contrast different artificial intelligence techniques used in intelligent systems, such as machine learning, rule-based systems, and natural language processing. Discuss the strengths and weaknesses of each technique in the context of intelligent system development.   (16)

13. (a) (i) Assume that customer id, income, age, city, interest rates are considered as independent variables. To predict the loan amount that can be paid by the customer, identify which regression model suits this kind of problem. Narrate how you can solve the problem as a step by step approach.   (8)

(ii) Briefly elucidate the concept of Multiple Linear Regression.   (8)

(OR)

(b) Illustrate the concept of logistic regression. Also find the probability of the patient having obesity using logistic regression. The weights of two patients are 120 Kg and 60Kg respectively. Let's assume the parameters b0=-7 and b1=1/12 .   (16)

14. (a) (i) Illustrate the concept of bias, variance, over fitting, under fitting using learning curves. (8)

(ii) In case the models behave differently (different performance) when trained with different dataset derived from the same population, whether the models are said to have low or high variance? Justify. (8)

(OR)

(b) Sketch the process of creating a decision tree. Analyze the advantages and disadvantages of using decision trees for classification and regression tasks. Discuss their interpretability, handling of missing values, and tendency to overfit. Provide examples to support your analysis. (16)

15. (a) (i) Using K-Means Clustering Algorithm, cluster the following points into two clusters C1 and C2. The points are A(2,2), B(1,1), C(3,2), D(3,1) and E(1.5,0.5). Assume A(2,2) and B(1,1) are the centers of the clusters C1 and C2 respectively.(8)

(ii) Illustrate the strength and weakness of k-means in comparison with the k-medoids algorithm. (8)

(OR)

(b) Demonstrate the concept of advanced supervised cleaning with example. (16)