b) Differentiate between value-based and policy-based reinforcement learning methods 3

8. a) What is the Bellman Equation? Write the Bellman Expectation Equation for the value function. 3

b) What is Q-Learning? How Does Q-Learning Work? 7

******

B.Tech Even Semester Examination, May, 2025

Computer Science & Engineering

(6th Semester)

Course No: CSE-603(21)

(Machine Learning)

Full Marks: 50

Pass Marks: 25

Time: 2 hours

Note:

1. Attempt any five questions.
2. Begin each answer in a new page
3. Answer parts of a question at a place
4. Assume reasonable data wherever required
5. The figure in the right margin indicates full marks for the question
6. All the mathematical symbols and abbreviations have their usual meanings.

1 a) What's the trade-off between bias and variance? 3

b) When should you use classification over regression? 3

c) What are the basic assumptions to be made for linear regression? 4

2. a) What is underfitting andoverfitting in Machine Learning and how do we prevent it? 4

b) Given the dataset below to learn a decision tree which predicts if people pass (Yes or No), based on their previous GPA and whether or not they studied.

GPA	Studied	Passed
L	F	F
L	T	T
M	F	F
M	T	T
H	F	T
H	T	T

What is the entropy H(Passed)?

What is the entropy H(Passed | GPA)?

What is the entropy H(Passed | Studied)? 6

3. a) Explain the Gradient Descent algorithm. Write down the update formula in each iteration. Explain the effects of high and low learning rates. 4+2

b) Below is a diagram if a single artificial neuron (unit):

Figure 1: Single unit with three inputs.

The node has three inputs x = (x1, x2, x3) that receive only binary signals (either 0 or 1). How many different input patterns this node can receive? What if the node had four inputs? Five? Can you give a formula that computes the number of binary input patterns for a given number of inputs? 2

c) Suppose we have a multilayer perceptron (MLP) model with 17 neurons in the input layer, 25 neurons in the hidden layer and 10 neuron in the output layer. What is the size of the weight matrix between the hidden layer and output layer? 2

4. a) What is the Naive assumption in the Naive Bayes Models? At what stage will you use Bayes rule in this algorithm?

b) What is advantage of using neural networks over linear regression?

c) What is a kernel in SVM? Why do we use kernels in SVM? 424

5. a) Why is feature extraction important? 2

b) Explain the difference between a test set and a validation set? 3

c) Explain the forward feature selection and backword feature selection methods. 5

6. a) Write 2 similarities and 2 differences between Kmeans algorithm and EM algorithm 4

b) How do you choose the number of clusters in clustering algorithms 3

c) What is a dendrogram and how is it used in hierarchical clustering? 3

7. a) Define the components of a Markov Decision Process (MDP). How is it used in reinforcement learning? 7