

THE KAVERY ENGINEERING COLLEGE*(An Autonomous Institution)***M.E. DEGREE END SEMESTER THEORY EXAMINATIONS, APRIL-MAY 2025****Second Semester****Computer Science and Engineering****CP4252/PCST2423 - Machine Learning****Regulations – 2021/2024****Duration : 3 Hrs****Maximum : 100 Marks****PART – A (ANSWER ALL QUESTIONS) (Marks 10*2=20)**

Q.No	Questions	BLT	CO	Marks
1.	In machine learning perspective, what is Bayesian decision theory?	L1	1	2
2.	Define the term overfitting. Give one solution for it.	L2	1	2
3.	How reinforcement learning is different from unsupervised learning?	L4	2	2
4.	Outline the philosophy of radial basis function(RBF). What does a RBF do?	L2	2	2
5.	Justify the necessity of dimensionality reduction in the context of machine learning.	L4	3	2
6.	Differentiate model based and model free learning.	L4	3	2
7.	What is hard and soft clustering in machine learning?	L1	4	2
8.	Suppose the data sample S contains N = 40 examples and the hypothesis H commits 10 errors over the examples, what will be the sample error?	L3	4	2
9.	What is perceptron? When does the perceptron fails to converge?	L2	5	2
10.	What is the difference between RNN and LSTM?	L4	5	2

PART – B (ANSWER ALL QUESTIONS) (Marks 5*16 = 80)

Q.No	Questions	BLT	CO	Marks
11.	a With appropriate examples, explain the various types of machine learning approaches and discuss the steps in designing the learning system.	L2	1	16
	Or			
	b State and explain the statistical models which are used for machine learning.	L2	1	16
12.	a Explain the principal component analysis(PCA) algorithm that reduces dataset dimensionality with retention of information. Explain how PCA works.	L3	2	16
	Or			
	b i Illustrate how to calculate the distance between the test data and the training data for kNN.	L3	2	8
	i Elaborate the random forest algorithm in detail.	L3	2	8

13.	a	Write K-Means algorithm: Suppose that the data mining task is to cluster the following eight points (with (x, y) representing location) into three clusters. A1(2, 10), A2(2, 5), A3(8, 4), B1(5, 8), B2(7, 5), B3(6, 4), C1(1, 2), C2(4, 9). The distance function is Euclidean distance. Suppose initially we assign A1, B1, and C1 as the center of each cluster, respectively. Use the k-means algorithm to show only i) The three cluster centers after the first round of execution ii) The final three clusters	L4	3	16
		Or			
	b i	How the distance between clusters is measured in hierarchical clustering? Explain the use of this measure in making decision on when to stop the iteration.	L4	3	8
	i	How to recompute the cluster centroids in the k-means algorithm?	L4	3	8
14.	a	State the principle of Hidden Markov Model. Describe the types of HMM with neat sketch.	L3	4	16
		Or			
	b	Given 14 training examples of the target concert play tennis with attributes outlook, temperature, humidity and wind. The frequency of play tennis = 9 Frequency of not play tennis = 5 Conditional probabilities are given as $P(\text{outlook} = \text{rainy} \text{Play} = \text{Yes}) = 2/9$ $P(\text{temp} = \text{cool} \text{Play} = \text{Yes}) = 3/9$ $P(\text{humidity} = \text{High} \text{Play} = \text{Yes}) = 3/9$ $P(\text{Windy} = \text{true} \text{Play} = \text{Yes}) = 3/9$ $P(\text{Outlook} = \text{rainy} \text{Play} = \text{No}) = 3/5$ $P(\text{temp} = \text{cool} \text{Play} = \text{No}) = 1/5$ $P(\text{humidity} = \text{High} \text{Play} = \text{No}) = 4/5$ $P(\text{Windy} = \text{true} \text{Play} = \text{No}) = 3/5$ Classify the new instance whether play = yes or No (Outlook = sunny, Temp = cool, Humidity = high, wind = strong) using Naive Bayes Classifier.	L3	4	16
15.	a	Explain Multilayer Neural Network with back propagation in detail with gradient descent optimization.	L2	5	16
		Or			
	b	Why Do You Prefer Convolutional Neural Networks over Artificial Neural Networks for image data? Justify with an example.	L2	5	16