



INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR
End-Autumn Semester Examination 2022-23

Date of Examination _____ Session: (FN/AN) Duration: 3hrs Full Marks: 40

Subject No.: CS60050 Subject Name: Machine Learning

Department/ Center/ School: Computer Science and Engineering

Special Instructions (if any): Write answers to all questions of a Section together. Answer **ALL** questions from Section A and **any five** from Section B. Do not answer more than 5 questions in Section B. If there are more questions answered in Section B, answers of the first five questions, in order of their appearances in the answer sheet, will be evaluated.

Section A (10 × 1)

QUESTION 1: The output of a neural network is [1, 2, 7, 4, 5]. Find the softmax probability for 7.

QUESTION 2: Assume, you want to cluster 9 observations into 4 clusters using K-Means clustering algorithm. After first iteration clusters, C1, C2, C3, C4 has following observations:

C1: (2,2), (1,4), (9,9)

C2: (5,5), (-1,4)

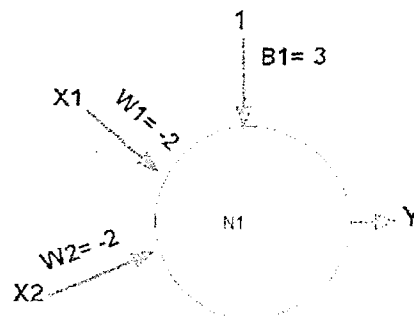
C3: (4,4), (2,-2)

C4: (6,6), (4,0)

What will be the cluster assignment for the point (4,4) in the next iteration?

QUESTION 3: Consider a multilayer feed forward network which takes a 6-D feature vector as input and produces a 3-D output vector. The network has one hidden layer with number of neurons 10. Please note that the input vector is the input to the first hidden layer. What is the dimension of the parametric space over which the optimization process would be carried out to train the network?

QUESTION 4: A neural Network given below takes two binary inputs $X_1, X_2 \in \{0, 1\}$ and the activation function for each neuron is the binary threshold function ($g(a) = 1$ if $a > 0$; 0 otherwise). What logical function does it compute?



QUESTION 5: Calculate the impurity of the following data using Gini Impurity Index.

Attribute 1	Attribute 2	Label
Red	3	Apple
Green	4	Apple
Yellow	1	Lemon
Yellow	2	Banana

Table 1:

QUESTION 6:

Given eigen values of covariance matrix formed using 4 data points as (0.5, 0.25) . What will be the total variance of the data points?

Type: MCQ

QUESTION 7:

Which of the following is an example of a sequential ensemble model?

- a) Bootstrapping
- b) AdaBoost
- c) Random forest
- d) All of the above

QUESTION 8: Consider a naive Bayes classifier with 3 boolean input variables, X1, X2 and X3, and one boolean output, Y. How many independent parameters need to be estimated to train such a naive Bayes classifier?

Type: MCQ

QUESTION 9:

Generally, in which of the following situations the ensemble method works better?

- a) If the individual base models have high correlation among predictions.
- b) If the individual base models have less correlation among predictions.
- c) There is no dependence of correlation on the ensemble output.

QUESTION 10:

Suppose there are 7 models with each having accuracy of 90%. What is the accuracy of their ensemble using majority voting?

Section B (5 × 6)

QUESTION 11:

What is the minimum number of layers needed to implement XOR gate using ANN? Justify your answer and give the implementation. 1+3+2

QUESTION 12:

In the clustering algorithm, Cluster Affinity Search Technique (CAST), how do you define a feature point is close to a cluster. How it is used to grow a cluster from a seed point? 2+4

QUESTION 13:

You are given a dataset [0, 1, 1, 2, 3, 4, 4, 4, 5, 6, 6]. Consider the following kernel,

$$k(d) = 1 - |d| ; \text{ if } |d| \leq 1$$

$$k(d) = 0; \text{ otherwise}$$

where d is the distance of sample x_i to x divided by a parameter called, the bandwidth. Compute the kernel density estimate for $x = 3$, when the bandwidth is 3. 6

QUESTION 14:

Consider the problem of diagnosing Dengue in a patient, given the following binary-valued (i.e., true or false) attributes:

- Fever = patient has fever
- Rashes = patient has rashes
- Nausea = patient has nausea
- Ache = patient has joint ache

Patient	Rashes	Fever	Nausea	Ache	Dengue
P1	T	T	F	F	F
P2	T	T	T	T	T
P3	F	F	T	F	F
P4	F	T	T	F	T

Consider the training set S given below. In each of the four training examples, true or false values for each of the five attributes are given, as well as the correct classification for each example. Compute the information gain, $\text{Gain}(S, A)$, for each attribute A with respect to the training set S. Which attribute would the ID3 algorithm choose to use for the root of the tree? 4+2

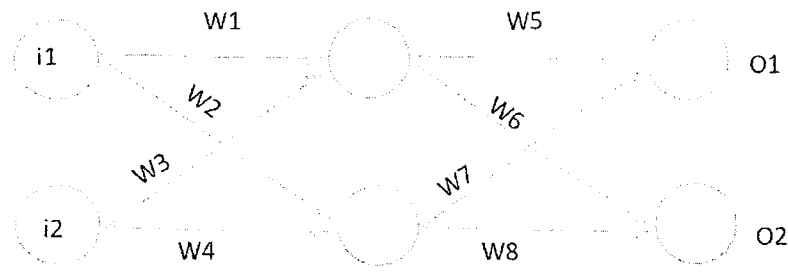
QUESTION 15:

a) Two boxes containing candies are placed on a table. The boxes are labelled B1 and B2. Box B1 contains 10 cinnamon candies and 45 ginger candies. Box B2 contains 5 cinnamon candies and 7 pepper candies. The boxes are arranged so that the probability of selecting box B1 is $\frac{3}{5}$ and the probability of selecting box B2 is $\frac{2}{5}$. Suresh is blindfolded and asked to select a candy. He will win a colour TV if he selects a cinnamon candy. If he wins a colour TV, what is the probability that the marble was from the first box? 4

b) In the Naive Bayes algorithm, suppose that prior for class 1 is greater than class 2, would the decision boundary shift towards the region R1 (region for deciding class 1) or towards region R2 (region for deciding class 2)? Justify. 2

QUESTION 16:

Consider the following neural network. The activation function used after the hidden and output layer is sigmoid. The error is given by $E = \frac{1}{2}(\text{target} - \text{output})^2$. $w_1 = 0.15$, $w_2 = 0.20$, $w_3 = 0.25$, $w_4 = 0.3$, $w_5 = 0.40$, $w_6 = 0.45$, $w_7 = 0.50$, $w_8 = 0.55$.



a) Calculate the o_1 and o_2 after the forward pass. Clearly show the outputs after hidden layer and output layer. What is the total error if target1 and target2 is 0 and 1, respectively?

b) Update w_5 , and w_1 using back propagation. Consider the learning rate as 0.2.

QUESTION 17:

Support vector machines learn a decision boundary leading to the largest margin from both classes. You are training SVM on a tiny dataset with 4 points, all having integer coordinates, as shown in the figure given in the last page. This dataset consists of two examples with class label -1 (denoted with plus), and two examples with class label +1 (denoted with triangles). Find the weight vector w and bias b . 4+2

