Started on	Thursday, 8 April 2021, 9:01 AM
State	Finished
Completed on	Thursday, 8 April 2021, 10:37 AM
Time taken	1 hour 36 mins
Overdue	51 mins 23 secs
Grade	Not yet graded

Question **1**

Complete

Marked out of 1.00

Consider a convolution layer. The input consists of 8 feature maps of size 20*20. The output consists of 10 feature maps of size 5*5. The convolution is done with a stride of 2 and zero padding, so the output features are of size 10*10.

(a) Determine the number of weights in this convolution layer.

(b) What if we replaced it with a fully connected layer, with the same input and output dimensions. Determine the number of weights in that case.

You can just leave the answer as a product of numbers. Showing steps can get you partial marks (even if your final solution is incorrect).

input = 20*20*8

output = 5*5*10

stride, S = 2

padding, P = 0

let kernel = x*x*10 , 10 as there are 10 channels in output, I assume 3d convulution

formula for ouput size is

((W - K + 2P)/S) + 1

W = 20

(20-x+2(0))/2 + 1 = 5

x = 11

so kernel = 11*11*10

a) total weights = 11*11*10

b) input size = 20*20*8 = X let

output size = 5**5*10 = Y let

Question **2**

Partially correct

Mark 0.50 out of 1.00

What is generally true about sigmoid and ReLU activation functions.

- [ ] a. Both functions have a monotonic first derivative
- [ ] b. Compared to sigmoid, the ReLU is more computationally expensive
- [ ] c. Both activation functions are monotonically non decreasing
- [x] d. The sigmoid derivative s'(x) is quadratic in s(x) ✔

Your answer is partially correct.

You have correctly selected 1.
The correct answers are:
Both activation functions are monotonically non decreasing,

The sigmoid derivative s'(x) is quadratic in s(x)

Question **3**

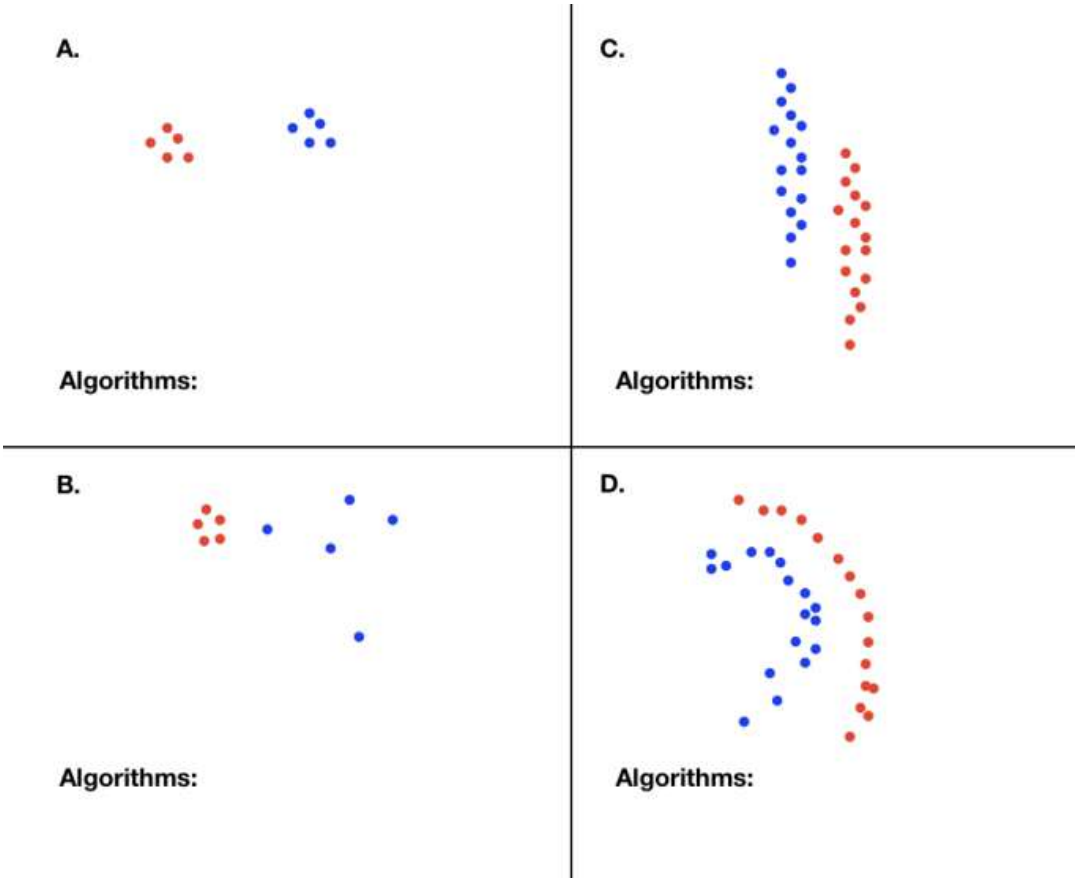
Complete

Marked out of 1.00

We study a problem of two-class classification. We look at three different algorithms (a) K-means (b) Single Link agglomerative clustering and (c) Gaussian Mixture Models (GMM).

Which of the three algorithms can lead to the clusters assignments shown below (red color represents the first cluster and the blue one represents the second cluster). If the cluster configuration cannot be obtained by any of the above three algorithms, mention it in the answer.

Your answer should be in four parts, Case A, Case B, Case C, and Case D. In front of each case, mention the algorithms which can lead to the given result. In case of GMM, the algorithm with higher soft probability is considered the assigned class.



No upload required. Type your solution below:

CASE A - kmeans

CASE B - Single Link agglomerative clustering

CASE C - Gaussian Mixture Models (GMM)

CASE D - NONe

Question **4**

Incorrect

Mark 0.00 out of 1.00

In kernel Support Vector Machines, there exists a one to one mapping ϕ for every kernel k.

Select one:

- True ✖
- False

The correct answer is 'False'.

Question **5**

Not answered

Marked out of 1.00

You are given the following observations

Sound	Texture	Colour	Class
woof	Smooth	White	Dog
purr	Furry	White	Dog
woof	Smooth	Black	Dog
woof	Smooth	Black	Dog
purr	Furry	White	Cat
purr	Smooth	Black	Cat
woof	Furry	Black	Cat
purr	Furry	White	Cat

Given that you observe a white animal with a smooth texture making a woof sound. Apply Naive Bayes classifier to decide if the observed animal is a Cat or a Dog. More specifically, compute the following two values:

P(class=Dog | Sound= woof, Texture = smooth, Color = white)

P(class=Cat | Sound= woof, Texture = smooth, Color = white)

Any normalization constants can be ignored (no need to compute the denominators, which do not affect the decision). Type the solution down below, clearly showing all the steps. Please note there is no option for uploading a handwritten solution. You have to type your solution below.

Question **6**

Correct

Mark 1.00 out of 1.00

A Multi-Layer network (MLP) with linear activation functions is equivalent to a single-layer perception. Assuming that the two use the same loss function and have the same number of inputs.

Select one:

- True ✔
- False

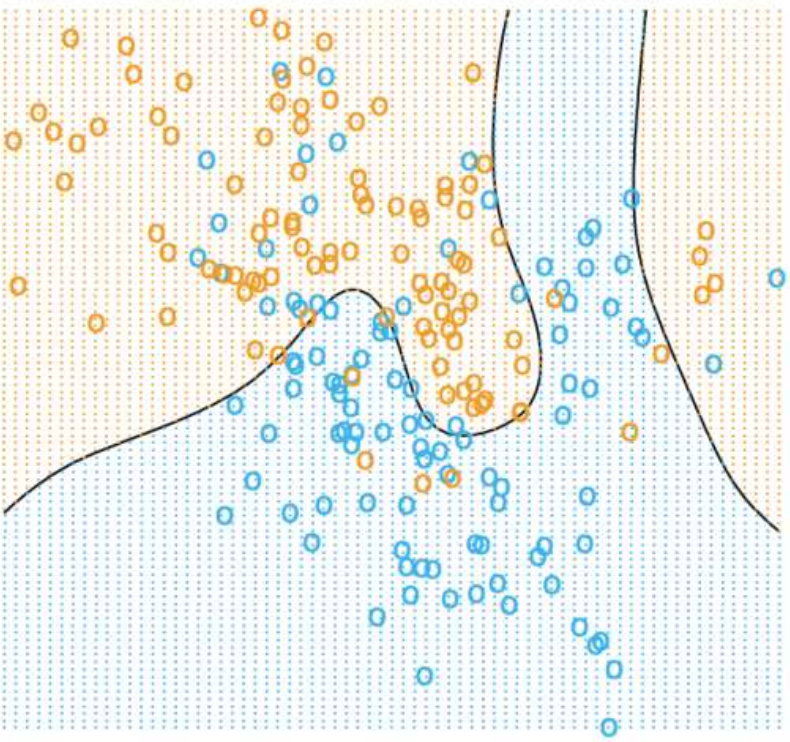
The correct answer is 'True'.

Question **7**

Incorrect

Mark 0.00 out of 1.00

Please observe the decision boundary obtained by a classifier in a two-class classification problem. Separating orange and the blue points.



Which classifier do you think might have resulted in the above boundary?

- a. 1-Nearest Neighbour (1 NN) ✖
- b. Linear SVM
- c. Logisitic Regression
- d. None of these

Your answer is incorrect.

The correct answer is:
None of these

Question **8**

Not answered

Marked out of 1.00

$$z = w_0 + w_1 x + w_2 x^2 + w_3 x^3$$

$$y = 1 + e^x$$

$$L = \frac{1}{2} (\log y - \log t)^2$$

- Determine the backprop derivation for $\frac{\partial L}{\partial w_2}$.
- x is the input, y is your prediction and t is the ground truth.
- L is the loss function.
- You need to show each step of your work.

◄ Quiz1

Jump to...