

**Q1 1**

20 Points

Design a two-input perceptron that implements the Boolean function

- (a) AND
- (b) OR
- (c) XOR
- (d)  $\neg A \wedge B$

▼ Q1.jpg

Download

**Q1**

**(A) And gate**

$$y = \begin{cases} 1 & \text{if } u \geq 2 \\ 0 & \text{otherwise} \end{cases}$$

Diagram: A perceptron with two inputs  $x_1$  and  $x_2$ , weights  $w_1$  and  $w_2$ , and bias  $Q_1$ . The output is  $y$ .

Calculations:

- $u = 1 \cdot 0 + 1 \cdot 0 = 0 \quad (0 < 2) \quad y = 0$
- $u = 1 \cdot 1 + 1 \cdot 0 = 1 \quad (1 < 2) \quad y = 0$
- $u = 1 \cdot 0 + 1 \cdot 1 = 1 \quad (1 < 2) \quad y = 0$
- $u = 1 \cdot 1 + 1 \cdot 1 = 2 \quad (2 = 2) \quad y = 1$

**(B) Or gate**

$$y = \begin{cases} 1 & \text{if } u \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

Calculations:

- $u = 2 \cdot 0 + 2 \cdot 0 = 0 \quad (0 < 1) \quad y = 0$
- $u = 2 \cdot 1 + 2 \cdot 0 = 2 \quad (2 > 1) \quad y = 1$
- $u = 2 \cdot 0 + 2 \cdot 1 = 2 \quad (2 > 1) \quad y = 1$
- $u = 2 \cdot 1 + 2 \cdot 1 = 4 \quad (4 > 1) \quad y = 1$

**(C) We can't design XOR with 2 input perceptron. XOR needs hidden layer**

**(D)  $\neg A \wedge B$**

Calculations:

- $u = 2 \cdot 1 + 2 \cdot 0 = 2$
- $u = 2 \cdot 1 + 2 \cdot 1 = 4$
- $u = 2 \cdot 0 + 2 \cdot 0 = 0$
- $u = 2 \cdot 0 + 2 \cdot 1 = 2$

Truth Table:

A	B	$\neg A \wedge B$
0	0	0
0	1	1
1	0	0
1	1	0

Output function:

$$y = \begin{cases} 1 & \text{if } u \geq 4 \\ 0 & \text{otherwise} \end{cases}$$
**Q2 2**

25 Points

A Naive Bayes classifier has to decide whether the document “cat dog” is news about the class a or class b.

document	class
cat dog	a
bird cat	b
cat bird	a
bird dog	b

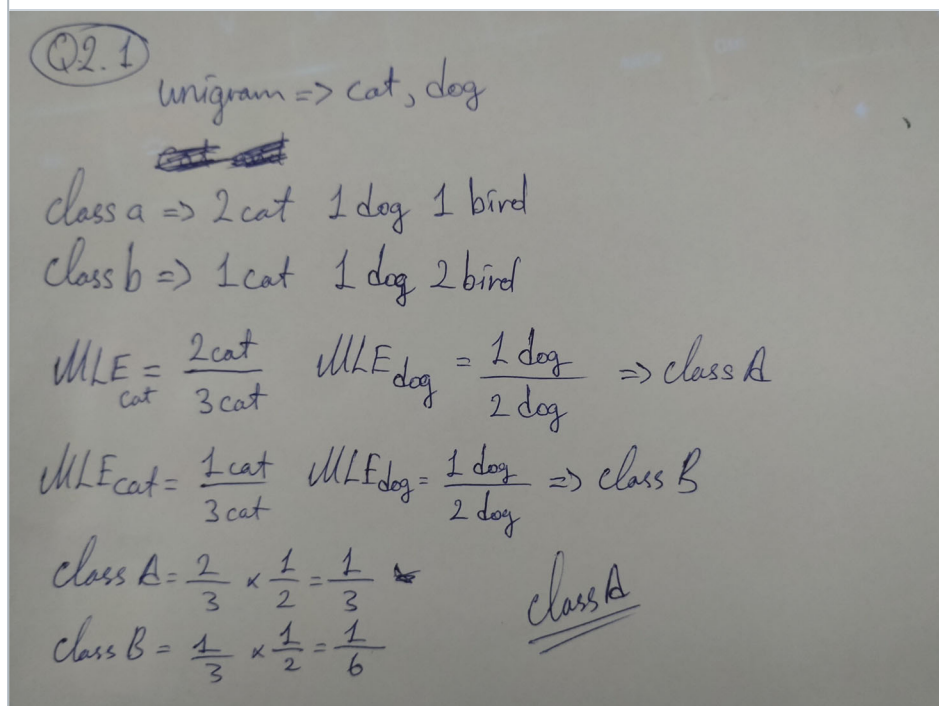
## Q2.1 a

15 Points

Estimate the probabilities that are relevant for this decision from the following document collection using Maximum Likelihood estimation (without smoothing).

▼ Q21.jpg

Download



## Q2.2 b

5 Points

Based on the estimated probabilities, which class does the classifier predict? Explain. (Show that you have understood the Naive Bayes classification rule.)

▼ Q22.jpg

Download

Q2.2 unigram  $\Rightarrow$  cat, dog

Priors:  
 $P(a) = \frac{1}{2}$   
 $P(b) = \frac{1}{2}$

$\hat{P}(w|c) = \frac{\text{count}(w,c)+1}{\text{count}(c)+1V}$

class A will predict.

$\frac{1}{2} \cdot \frac{3}{7} \cdot \frac{2}{7} = \frac{3}{49}$

$\frac{1}{2} \cdot \frac{2}{7} \cdot \frac{2}{7} = \frac{2}{49} \Rightarrow$  class B

class A higher than class B

$P(\text{cat}/a) = \frac{2+1}{4+3} = \frac{3}{7}$   
 $P(\text{cat}/b) = \frac{1+1}{4+3} = \frac{2}{7}$   
 $P(\text{dog}/a) = \frac{1+1}{4+3} = \frac{2}{7}$   
 $P(\text{dog}/b) = \frac{1+1}{4+3} = \frac{2}{7}$

$\frac{1}{2} \cdot \frac{3}{7} \cdot \frac{2}{7} = \frac{3}{49}$   
 $\frac{1}{2} \cdot \frac{2}{7} \cdot \frac{2}{7} = \frac{2}{49}$

Prior A  $P(\text{cat}/A)$   $P(\text{dog}/A)$

### Q2.3 c

5 Points

Practical implementations of a Naive Bayes classifier often use log probabilities. Explain why?

Since the multiplied numbers are fractional numbers, log is used to avoid underflowing.

### Q3 3

25 Points

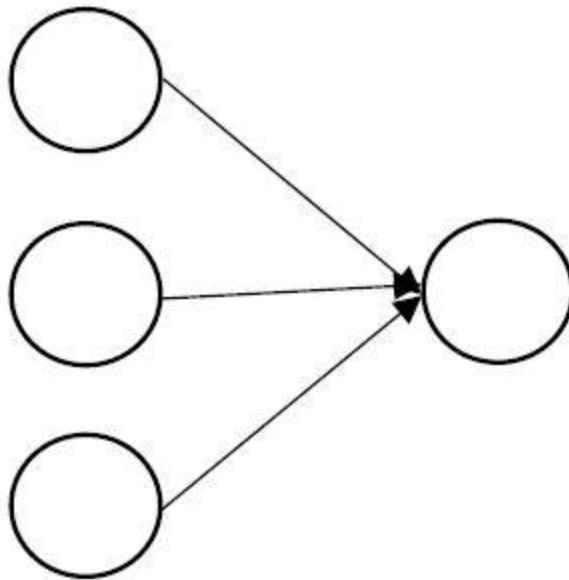
A neural network with three input neurons and 1 output neuron is given below.

Suppose that the output neuron uses an identity activation function, i.e.,  $z = f(z)$  where  $z$  is the total input to the neuron. Let  $d$  be the desired output and  $d=4$ , and let  $E = -d \log z - (1-d) \log (1-z)$  be the cross-entropy error. For the inputs (1,2,1) and weights (1,3,1), perform a forward pass to compute the output of the network and then apply backpropagation with gradient descent for one iteration.

Write down the weights after the forward pass and also after the backward pass separately. Learning rate  $\mu = 10$ .

Hint: The derivative of the identity function is simply 1. The derivative of the logarithm is given below:

$$\frac{\partial \log u}{\partial u} = \frac{1}{u} \qquad \frac{\partial \log(1-u)}{\partial u} = \frac{-1}{1-u}$$



### Q3.1 1

5 Points

Forward pass

### Q3.2 2

20 Points

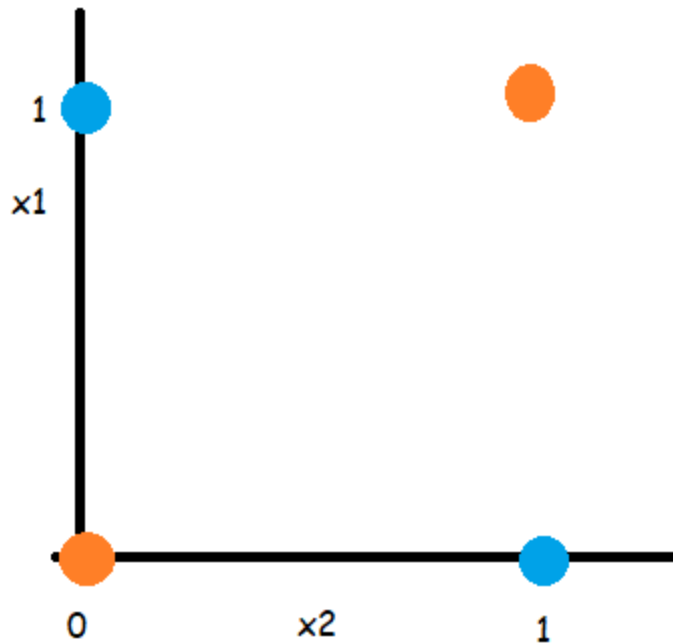
Backward pass (backpropagation)

 No files uploaded

### Q4

30 Points

A dataset with two attributes ( $x_1$ ,  $x_2$ ) and two classes ( $y=1$  (blue) and  $y=-1$  (orange)) are given below.

**Q4.1 1**

10 Points

To classify the given dataset ( $Score(x) = w_0 + w_1x_1 + w_2x_2$ ), can we use logistic regression with the given features? (Yes/No) Why? If Yes, show which weights we need to use for classification?

First)  $x_1 = 1, x_2 = 0 \Rightarrow w_1$  needs to be  $> 0.5$   
 Second)  $x_1 = 0, x_2 = 1 \Rightarrow w_2$  needs to be  $> 0.5$   
 LAST)  $x_1 = 1, x_2 = 1 \Rightarrow w_1 + w_2$  needs to be  $< 0.5 \Rightarrow$   
 impossible due to previous two cases

**Q4.2 2**

20 Points

If we add an extra feature to the model ( $x_1x_2$ ), is it possible to classify the datasets? (Yes/No) Why? If Yes, show which weights we need to use for classification?

No, model need extra layer

Quiz3

GRADED

STUDENT  
MEHMET TAHA USTA

TOTAL POINTS  
25 / 100 pts

QUESTION 1

1

20 / 20 pts

QUESTION 2

2

5 / 25 pts

2.1 a

0 / 15 pts

2.2 b

0 / 5 pts

2.3 c

5 / 5 pts

QUESTION 3

3

0 / 25 pts

3.1 1

0 / 5 pts

3.2 2

0 / 20 pts

QUESTION 4

(no title)

0 / 30 pts

4.1 1

0 / 10 pts

4.2 2

0 / 20 pts