Neural networks 2018 Fall
paper-based closed room test

PPCU-ITK

8th October, 2018

Name: ____________________

NEPTUN ID: ____________________

Time of practice session: ____________________

Question	Points	Student's points
1	20	
2	10	
3	15	
4	20	
Total	65	

Instructions:

1. This examination contains 6 pages, including this page.
2. You have **ninety (90) minutes** to complete the examination. As a courtesy to your classmates, we ask that you not leave during the last fifteen minutes.
3. You may not use any external resources, including lecture notes, books, other students or other engineers.
4. You may use a calculator. You may not share a calculator with anyone.
5. Please sign the below statement.

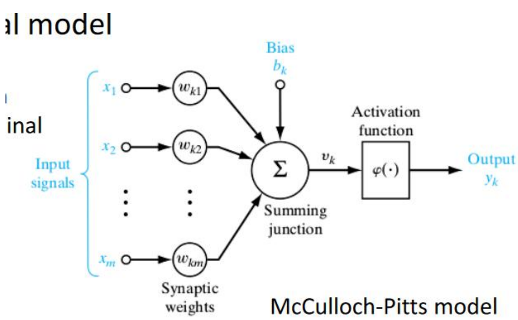
I hereby certify that I will neither give nor receive unpermitted aid on this examination.

Signature: ____

Question 1: The artificial neuron

Overall: [20 pts]

(a) [5 pts] Draw the McCulloch-Pitts model of an n-dimensional artificial neuron and label its components!



(b) [6 pts] What is the output equation of an artificial neuron?

- Output equation:

$$y_k = \varphi\left(\sum_{i=1}^{m} w_{ki}x_i + b_k\right)$$

x_i : input vector (*i: 1….m*)
w_{ki} : weight coefficient vector of neuron k
b_k : bias value of neuron k
o_k : output value of neuron k

Bias

(c) [2 pts] What is the bias term? How is it usually included?

- Bias can be included as:
  - $w_0=b$
  - $x_0=1$

$$y_k = \varphi\left(\sum_{i=0}^{m} w_{ki}x_i\right) = \varphi(\mathbf{w}^T\mathbf{x})$$

In
sig

helps to reach assymetry, and helps the model to fit the data. hyperplane usually goes through the origo. if we want to change it we add the bias.

(d) [7 pts] Define the term activation function. Also, give at least two examples, complete with name, simple graph and mathematical formula.

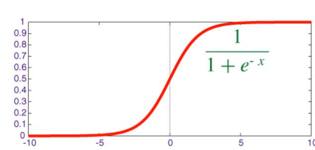
- Activation function: φ(.)
  - Always a nonlinear function
  - Typically it clamps the output (introduces boundarie:
  - Monotonic increasing function
  - Differentiable
    - Important from theoretical point of view
  - Or at least continuous (except in simplified cases)
    - Sophisticated training algorithms require continuity

- **Sigmoid** (or logistic) function is a widely activation function

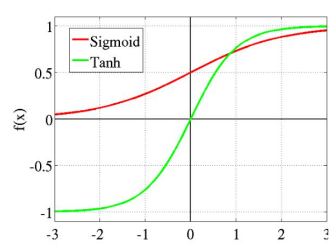
$$y = \varphi(u) = \frac{1}{1+e^{-\lambda u}}$$

- where

$$u = \sum_{i=0}^{m} w_i x_i = \mathbf{w}^T\mathbf{x}$$



- Bipolar activation function: tanh
- Continuously differentiable
- Monotonic
- Useful, when bipolar output is expected
- Hard approximations:
  - Piece-wise
  - Step-wise



Question 2: Set separation by hyperplane

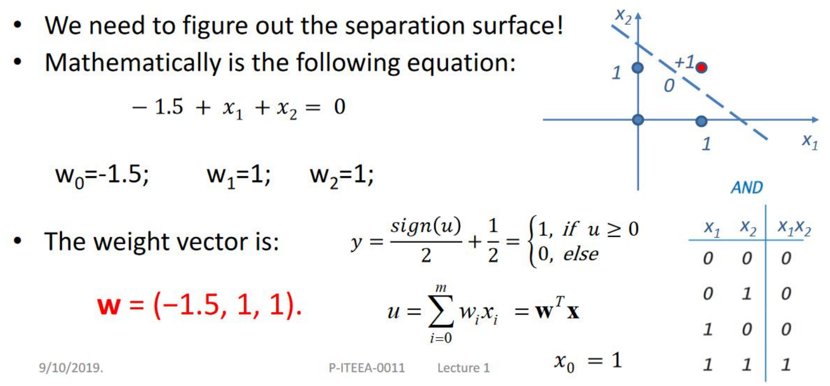
Overall: [10 pts]

(a) [4 pts] When can we say that a boolean function is linearly separable?

two sets are linearly separable precisely when their respective convex hulls are disjoint (colloquially, do not overlap). egyrészt. másrészt kb akkor, ha ábrázolod őket az euklideszi síkba és szépen ketté tudod választani a két halmazt egy egyenessel.

(b) [6 pts] Show that the ~~XNOR~~ (negated XOR) function is linearly separable, while the XOR function is not. You may do so in a graphical way as well.

Ehelyett AND volt mert a XNOR kurvára nem szeparábilis



Question 3: The learning algorithm

Overall: [15 pts]

(a) [5 pts] Write down the 5 steps of the perceptron learning algorithm (introduced by Rosenblatt). Explain the steps in greater detail.

1. *Initialization.*

   Set **w**(0)=**0** or **w**(0)=**rand**

2. *Activation.*

   Select a $\mathbf{x}_k \rightarrow \mathrm{d}_k$ pair

3. *Computation of actual response*

   $$y(k) = sign\left(w^T(k)x(k)\right)$$

4. *Adaptation of the weight vector*

   $$\mathbf{w}(k+1) = \Psi\left(\mathbf{x}(k), \mathbf{w}(k), d(k), y(k)\right)$$

5. *Continuation*

   Until all responses of the perceptron are OK

(b) [10 pts] Apply the above learning algorithm to learn the AND logical function with a single perceptron, with the following intial conditions:

- **all** the weights are initialized to 0
- the learning rate is $^1/_2$

Please show your work.

hát ezt nem tudta megcsinálni a jalal konzin én legalábbis egy büdös szót nem értettem belőle. de egyébként ehhez hasonlóan kell: week2 (perceptron training) 17-21. dia
http://users.itk.ppke.hu/~ekacs/anyagok/felev5/NeurHalok/2019%20Lectures/Week_2_-_perceptron_training.pdf

Question 4: Back-propagation

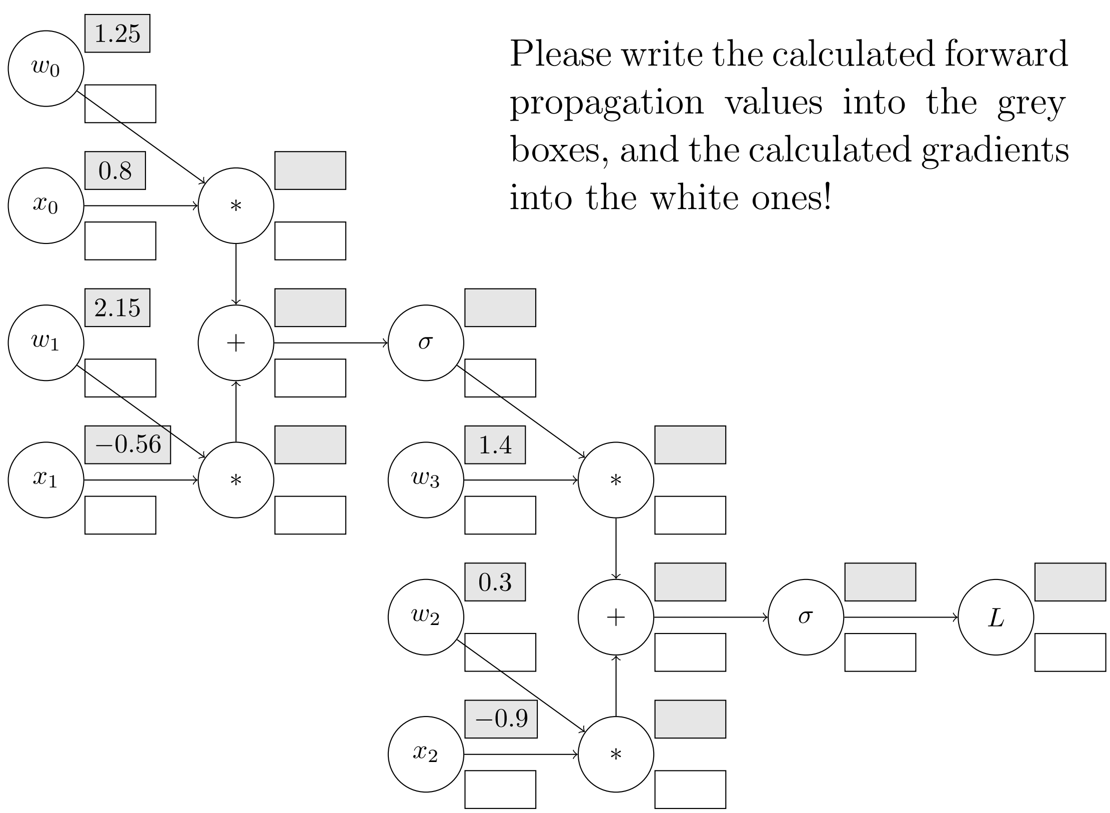
Overall: [20 pts]

(a) [2 pts] Write down the most common use case for the back-propagation algorithm!

¯_(ツ)_/¯

(b) [18 pts] Calculate the gradient for each node of the computational graph below! The σ node represents the logistic sigmoid function. Please show your thinking (next page)!
Here is some help with derivation :)

$$f(x) = e^x \rightarrow \frac{df}{dx} = e^x$$
$$f(x) = ax \rightarrow \frac{df}{dx} = a$$
$$f(x) = \frac{1}{x} \rightarrow \frac{df}{dx} = \frac{-1}{x^2}$$
$$f(x) = c + x \rightarrow \frac{df}{dx} = 1$$
$$\sigma(x) = \frac{1}{1+e^{-x}} \rightarrow \frac{d\sigma}{dx} = \sigma(x) \cdot (1 - \sigma(x)$$



Please write the calculated forward propagation values into the grey boxes, and the calculated gradients into the white ones!

You have reached the end of the exam. If you happen to be done **more than 15 minutes before** the end of the allotted time, please fill in the time you hand in your test, and if you'd like, you may leave the lecture hall quietly. (If you hand in your test at the end of the exam with everybody else, you may leave this field blank.)

Test handed in at: ____________