GÖTEBORGS UNIVERSITET

STUDENT

0026-XKF

TENTAMEN

TIG122 Tentamen

Kurskod	--
Bedömningsform	DO
Starttid	13.05.2024 12:00
Sluttid	13.05.2024 15:00
Bedömningsfrist	--
PDF skapad	20.01.2025 15:41
Skapad av	Catarina Elg

General information

Welcome to TIG122_VT24 -- Artificial Intelligence -- 2nd examination session:

- Standard duration of the examination: **3 hours.**
- Number of questions: **18.**
- Number of available points: **20.** Each correctly answered question earns you one point, except for two questions, revealingly titled "Convolutional and Pooling layer" and "Vanishing gradients", which may earn you two points each.
- You will have no access to external props: no books, slides, internet, calculators, and phones.
- From 13:30, the course coordinator will be present in the exam room to answer questions about the exam.
- Question types:
  - **Multiple choice (one correct answer): 5 questions**
  - **Multiple choice (multiple possible correct answers): 2 question** --> ATTENTION: selecting the wrong answer(s) zeroes the gain from that question!
  - **Free text: 5 questions** --> Show your knowledge/understanding of the issue at hand in a relatively short form.
  - **Basic calculations: 6 questions** --> No complex calculations! In case you have to report your calculations, consider utilizing the character * to represent the multiplication operators (i.e. both for the basic multiplication and the "dot product" between matrices).

Grading of the digital examination

(information cloned from "Canvas TIG122 VT24 > Syllabus")

The Bonus points will be added to the examination score, after being transformed by the "ceiling" function, i.e. the Bonus points will be transformed into the least integer number greater than or equal to their value (e.g. 0.5 Bonus points are transformed into 1).

Total score = Examination score + ceil(Bonus points)

- Total score ≥ 14 --> VG
- 8 ≤ Total score <14 --> G
- Total score < 8 --> U

If you have received a negative result on the examination of March 18th, the bonus points will be computed again only if they allow an upgrade from U to G (i.e. they will not be considered to upgrade from G to VG). If the examination on May 13th is your first examination for the course, the bonus points will be in any case added to your DISA result.

Warm wishes!

1 Turing test

Discuss the structure of the Turing test.
Is that a good way to tell if a computer program is "intelligent"? Why/why not?
Fill in your answer here

The Turing Test is a way of testing a program's intelligence by how well it can act human. A program passes the test when the human it's chatting with thinks they are chatting with another human.

I think it's a good measure of social intelligence, but it doesn't say anything on how intelligent the program is in other aspects. As per the Chinese Room, one can reason if it does truly understand the terms it is using since it hasn't experienced them in other ways than by words and terms.

Ord: 90

Besvarad.

2 Decision trees

Which statement is correct?
Select one alternative:

- Bootstrapping starts by removing repeated data points.
- Random forests combine bootstrapping and limited tree size. ✓
- The output of hierarchical clustering depends on the starting point.

Rätt. 1 av 1 poäng.

3 Regression trees

Which statement is correct?

Select one alternative:

- When constructing regression trees, the goal is to minimize the Gini impurity.
- When constructing regression trees, the goal is to minimize the sum of the squared residuals. ✓
- When constructing regression trees, the goal is to minimize the number of Gini dimensions.

Rätt. 1 av 1 poäng.

4 ReLU activation

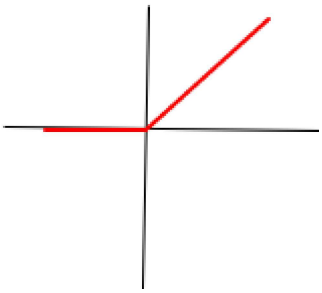
Consider a unit with a ReLU activation function, input x = (2,3), weights w = (−1, 2), and threshold θ = 2.
Calculate its activation a. (report your calculations)

Fill in your answer here

X*W = 2*(-1) + 3*2 - 1*(-2) = -2 + 6 - 2 = 2
(I'm adding the threshold as a weight w_0 with a constant input x_0=1, thus treating the threshold as 0)

Since the sum is 2, the activation a = 2.

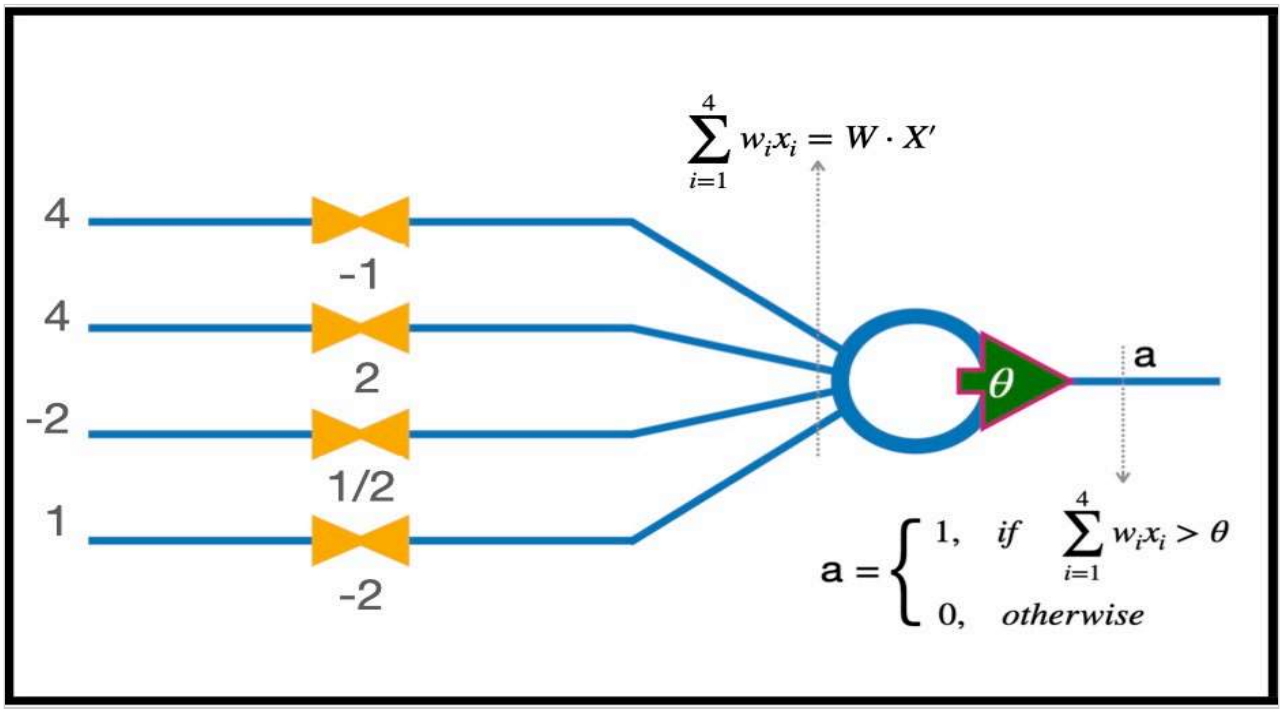
ReLU



Ord: 45

Besvarad.

5 Step activation with θ



Consider the unit represented in the figure, with its input and weight distribution, step function activation, and a threshold θ = 2.
Calculate the weighted sum and its corresponding activation. (report your calculations)

Fill in your answer here

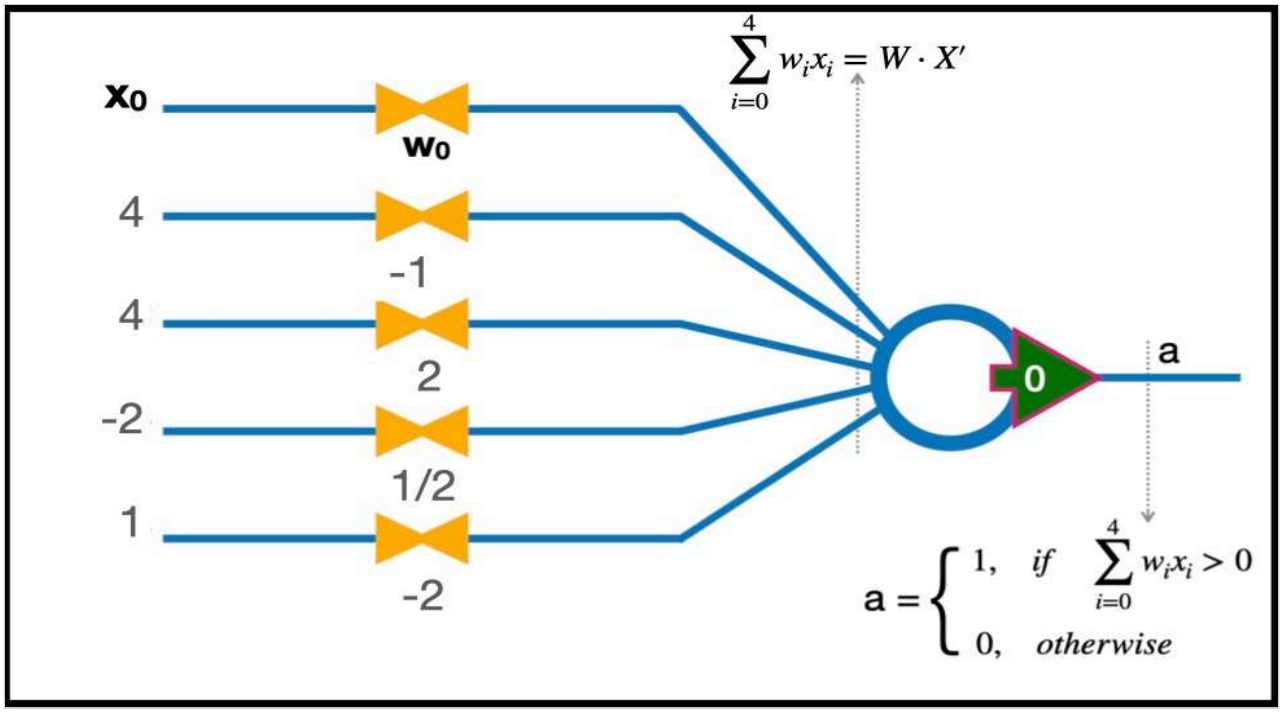
Sum
X * W = 4*(-1) + 4*2 + (-2)*1/2 + 1*(-2) = -4 + 8 - 1 - 2 = 8 - 7 = 1

1 < 2 => **a = 0**
(sum is less than threshold, thus the activation is 0)

Ord: 43

Besvarad.

6 Step activation with extra input



Consider the step function unit from the previous question, with its input and weight distribution, and a threshold θ = 2.
Complete the equivalent description of the unit by indicating the values for x_0 and w_0 (see figure above). Calculate the weighted sum and its corresponding activation. (report your calculations)

Fill in your answer here

x_0 = 0
w_0 = -2

Sum
X * W = 4*(-1) + 4*2 + (-2)*1/2 + 1*(-2) + 1*(-2) = -4 + 8 - 1 - 2 - 2 = 8 - 9 = -1

-1 < 0 => **a = 0**

Ord: 43

Besvarad.

7 Activation function

A unit (artificial neuron) generates an activation value of -1. Which of the following activation functions are compatible with that fact?

Select one or more alternatives:

- [x] Sign
- [ ] Tanh
- [ ] ReLU
- [ ] Sigmoid

Delvis rätt. 0.5 av 1 poäng.

8 Matrix layer representation

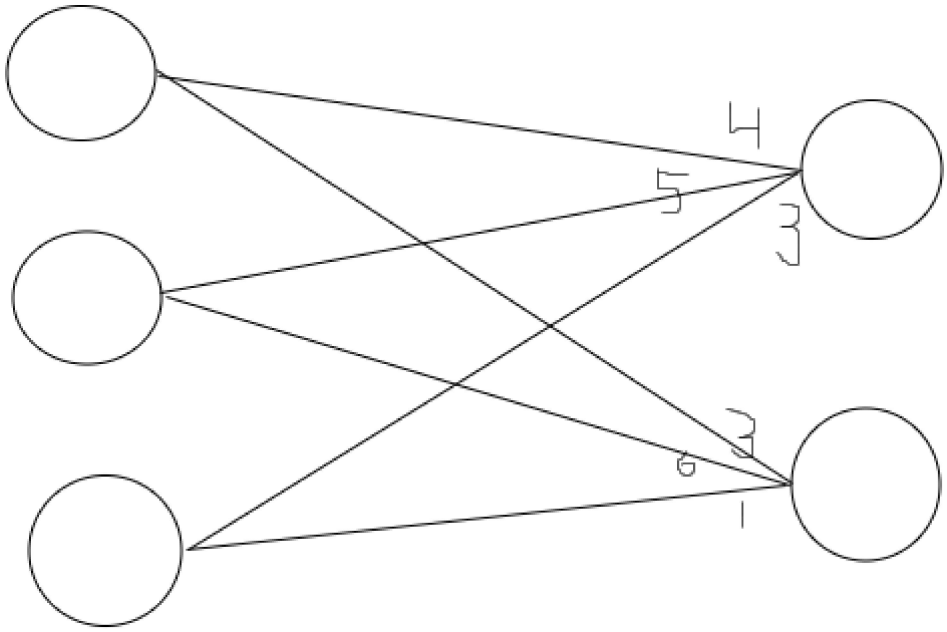
$$\begin{pmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \end{pmatrix} = \begin{pmatrix} 4 & 5 & 2 \\ 3 & 6 & 1 \end{pmatrix}$$

The matrix of weights displayed above represents the first hidden layer of a fully connected multiple layer network.
Specify the number of inputs, the number of units in the hidden layer, and the weight value of each connection between input and the second unit of the hidden layer.

Fill in your answer here

Inputs: 3
Units in the hidden layer: 2

Weights of second unit in the hidden layer = 3, 6, 1



Ord: 20

Besvarad.

9 Layout 1

Consider a fully connected network, with three input units, a first hidden layer constituted of four units, a second hidden layer of three units, and a two unit output layer.
Calculate the total number of weights and the total number of biases.

Fill in your answer here

Biases = 4 + 3 + 2 = 9
Weights = 3*4 + 4*3 + 3*2 = 12 + 12 + 6 = 30

Ord: 24

Besvarad.

10 Layout 2

```
model = Sequential()
model.add(Dense(10, activation='sigmoid', input_shape=(100,)))
model.add(Dense(10, activation='softmax'))
```

Consider the code snippet in the figure above.
Briefly describe the network layout. For each layer, specify the number of weights and biases.

Fill in your answer here

100 input units, 10 in a first hidden layer, 10 output units

Biases = 10 + 10 = 20
Weights = 100 * 10 + 10 * 10 = 1000 + 100 = 1100

Ord: 34

Besvarad.

11 Backpropagation 1

Is the Tanh function a suitable activation function for training by backpropagation? Why?

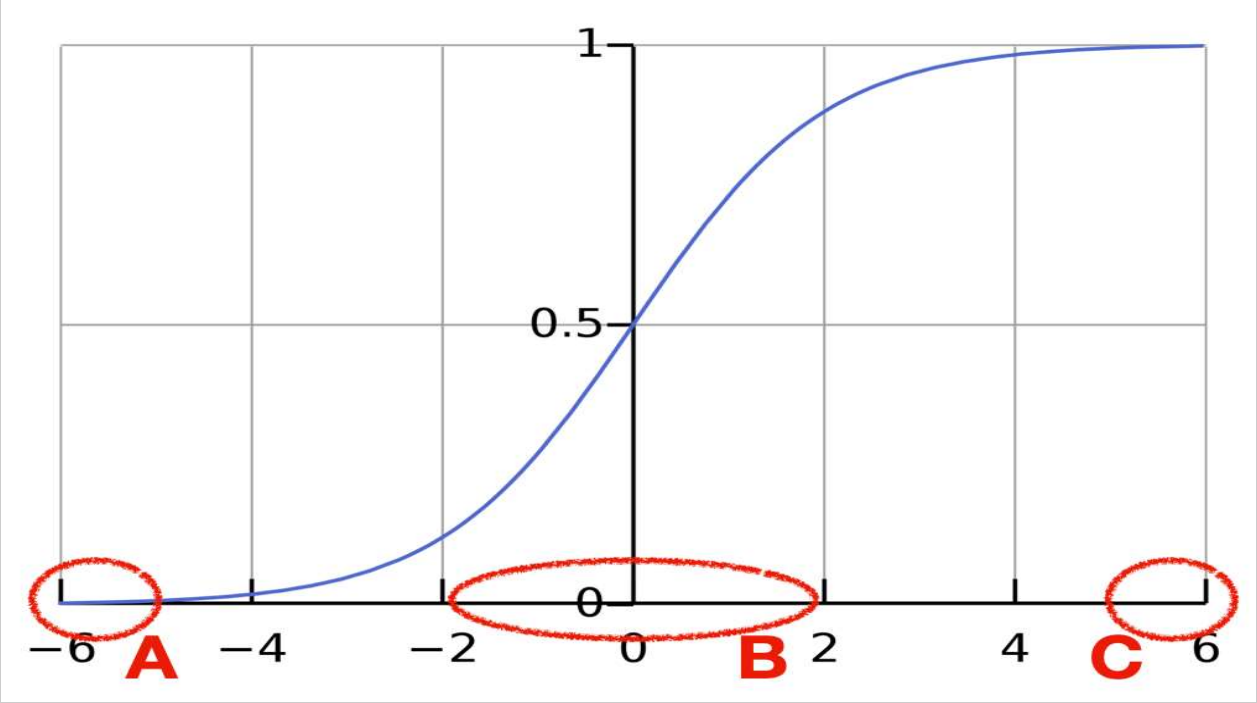
Fill in your answer here

Yes! Because as opposed to something like the sign function it has a curve, like the sigmoid function. Thus it has gradients, which are use in backpropagation to determine which way to adjust the parameter.

Ord: 35

Besvarad.

12 Backpropagation 2



Refer to the figure. During training by backpropagation, what is the ideal working region for the argument of the sigmoid activation function and why?

Select one alternative:

- A - Because of the lower value of the function.
- B - Because of the maximum value of the function's slope. ✓
- C - Because of the limited value of the function's slope.
- C - Because of the higher value of the function.
- B - Because of the intermediate value of the function.

Rätt. 1 av 1 poäng.

13 Learning rate

What is the effect of the learning rate η?
Select one alternative:

- To determine the size of the mini-batch.
- To control the level of update of the weights during learning. ✓
- To determine the level of dropout.
- To control the level of activation of the trained unit.

Rätt. 1 av 1 poäng.

14 Batches

With reference to the use of mini-batches in backpropagation, which statements are correct?
Select one or more alternatives:

- [x] The sequence of training samples is different for each epoch. ✓
- [ ] Considering each epoch, the weights are updated along the direction of the steepest gradient.
- [x] During each epoch, each item in the training set is shown to the network once and only once. ✓
- [ ] Splitting the training data into smaller batches can result in increased memory usage by the system.

Rätt. 1 av 1 poäng.

15 Number of neurons

Imagine you have trained a network for image classification with 64 units in a given hidden layer. Reducing the number of units in that hidden layer to 32 does not result in any reduction in the performance. Based on that observation, what is the most reasonable action to take next?

Select one alternative:

- Test the performance after halving the value of the learning rate.
- Test the performance for 128 units. ✗
- Test the performance for 16 units. ✓

Fel. 0 av 1 poäng.

16 Sat

Name the technique that can prevent saturation, particularly during the initial epochs of training a neural network using backpropagation.

Fill in your answer here

Limiting the amounts of epochs.

Ord: 5

Besvarad.

17 Vanishing gradients

Is it effective to use Glorot initialization and batch normalization simultaneously? Explain why or why not.

Fill in your answer here

Glorot initialization is a way to initialize the weights so that the activation is normalized (in sigmoid function). Batch normalization is applied after each layer during propagation. So I don't see why they would be incompatible, seing as they are used at different points of time in the process.

Ord: 49

Besvarad.

18 Convolutional and pooling layers

Consider an 6x6x3 RGB image (i.e. an 6x6 image with three color channels) with no padding. Apply one 3x3 kernel with a stride of 1. What is the size and depth of the resulting activation map?

Then, apply onto such an activation map a 2x2 max pooling layer with a stride of 2. What is the size and depth of the final map?

Fill in your answer here

The final map is **2x2x1.**

The 3x3 (actually 3x3x3 if I understand it correctly) kernel:
6x6x3 -> 4x4x1

And then the max pooling layer (stride 2):
4x4x1 -> 2x2x1

Ord: 29

Besvarad.