Institut für Mess- und Regelungstechnik
mit Maschinenlaboratorium
Karlsruher Institut für Technologie
(KIT)
Prof. Dr.-Ing. C. Stiller

Solutions for exam
"Machine Vision"
February 22, 2021

Question 1 (6 points)

Let $f,g : \mathbb{R} \to \mathbb{R}$ be functions and $\alpha,\beta \in \mathbb{R}$. Prove the following property of the Fourier transform:

$$\mathcal{F}\{\alpha \cdot f(t) + \beta \cdot g(t)\} = \alpha \cdot \mathcal{F}\{f(t)\} + \beta \cdot \mathcal{F}\{g(t)\}$$

Note, that $\mathcal{F}\{h\}$ denotes the Fourier transform of a function h.

Solution

$$\begin{aligned}
\mathcal{F}\{\alpha \cdot f(t) + \beta \cdot g(t)\} &= \int_{-\infty}^{\infty} (\alpha \cdot f(t) + \beta \cdot g(t)) \cdot e^{-2\pi ikt} dt \\
&= \alpha \cdot \int_{-\infty}^{\infty} f(t) \cdot e^{-2\pi ikt} dt + \beta \cdot \int_{-\infty}^{\infty} g(t) \cdot e^{-2\pi ikt} dt \\
&= \alpha \cdot \mathcal{F}\{f(t)\} + \beta \cdot \mathcal{F}\{g(t)\}
\end{aligned}$$

Question 2 (2+2 points)

Assume the two scenarios (a) and (b) described below. For each scenario name two possible techniques to improve the image quality. If you name more than two techniques, only the first two count for the answer.

(a) A photographer is standing at the coast and wants to make a picture of an island two kilometers away. However, on the pictures the island appears too small so that the details of the island cannot be recognized. What can the photographer do to get more details on the pictures?

(b) In a machine that should sort letters according to their destination a camera makes pictures of the letters while they are passing by the camera with high speed. The distance between the letters and the camera is always the same and it is previously known. The problem is that the images are too dark. You are allowed to change the overall setup of the machine, to change the control parameters of the camera, and to change the lens of the camera. You are not allowed to replace the camera by another one. Mind that improving the image brightness should not lead to blurry images!

Solution

(a) Possible answers are:

- replace the lens by a lens with larger focal length/increase the focal lens in case of a zoom lens
- use an imager with larger resolution
- rent a boat, approach the island and reduce the distance between the island and the position from which the photo was taken

(b) Possible answers are:

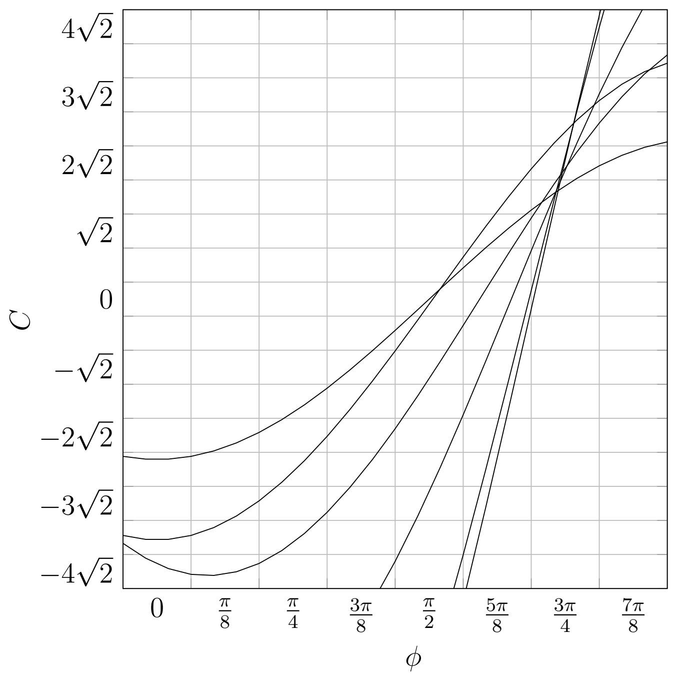
- increase the aperture (since the distance to the letters is known and fixed, the decrease in depth of field does not matter)
- add light sources to the machine
- increase the gain of the camera or post process the images by multiplying all gray values by a factor > 1 (not preferable solution but okay as answer)

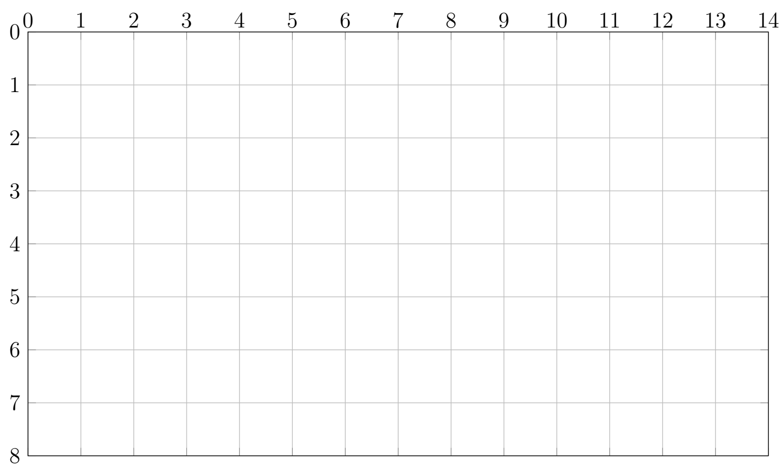
Question 3 (6 points)

Assume that six edge pixels have been extracted from a camera image. The Hough transform has been applied to each of the six pixels. The illustration below shows the Hough space with the resulting curves. The cells of the accumulator array are shown as gray boxes. Find the maximum in the accumulator array and mark it. Then, transform the maximum back into a line equations in normal form. Draw that line in the image below. Mind that the origin of the image is in its left upper corner, that an angle of 0 points to the right, and an angle of $\frac{\pi}{2}$ points downwards.

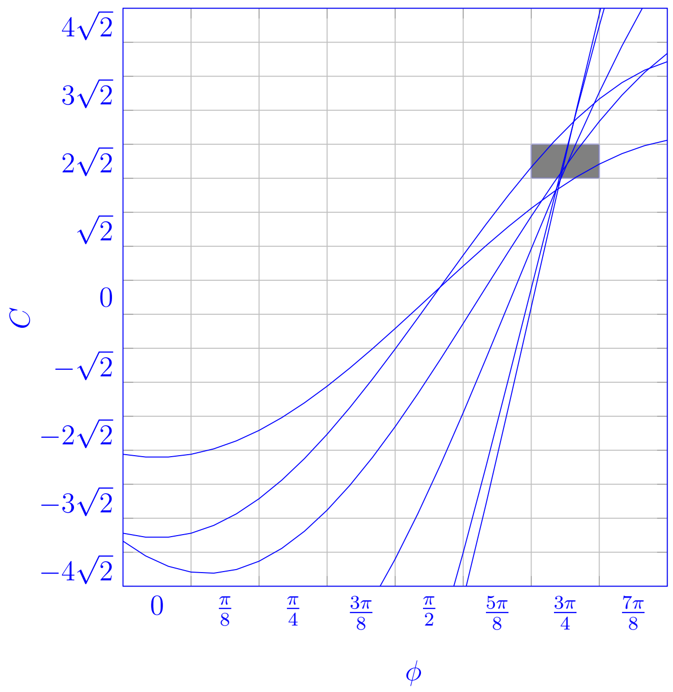
Remark:

θ	0	$\frac{\pi}{8}$	$\frac{\pi}{4}$	$\frac{3\pi}{8}$	$\frac{\pi}{2}$	$\frac{5\pi}{8}$	$\frac{3\pi}{4}$	$\frac{7\pi}{8}$
$\sin(\theta)$	0	$\frac{\sqrt{2-\sqrt{2}}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2+\sqrt{2}}}{2}$	1	$\frac{\sqrt{2+\sqrt{2}}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2-\sqrt{2}}}{2}$
$\cos(\theta)$	1	$\frac{\sqrt{2+\sqrt{2}}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2-\sqrt{2}}}{2}$	0	$-\frac{\sqrt{2-\sqrt{2}}}{2}$	$-\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{2+\sqrt{2}}}{2}$





Solution



The line equation is

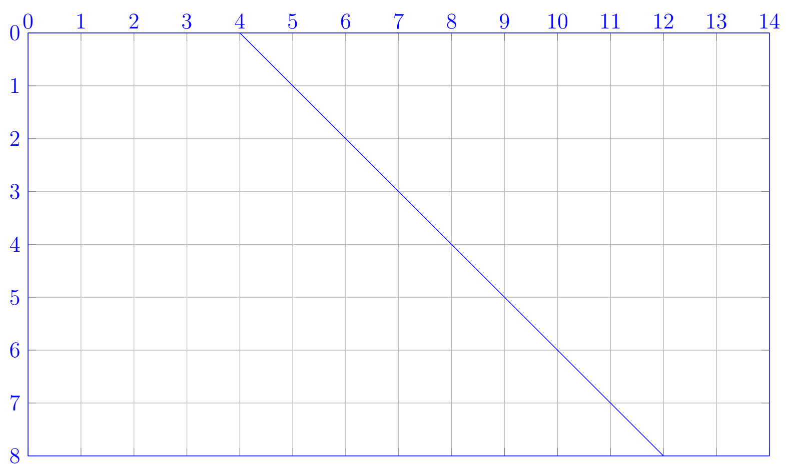
$$u \cdot \cos \frac{3\pi}{4} + v \cdot \sin \frac{3\pi}{4} + 2\sqrt{2} = 0$$

$$v(u) = -u \cdot \cot \frac{3\pi}{4} - \frac{2\sqrt{2}}{\sin \frac{3\pi}{4}}$$

$$v(u) = u - 4$$

$$\vec{n} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad c = \frac{4}{\sqrt{2}}$$

$$\langle \vec{n}, \vec{x} \rangle + c = \langle \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \vec{x} \rangle + \frac{4}{\sqrt{2}} = 0$$



Question 4 (4 points)

Assume we want to compare different colors which we represent in the HSV color space. Discuss briefly whether or not the Euclidean distance between two HSV triplets is a reasonable metric to measure the similarity of colors. The Euclidean distance between the two HSV triplets (h_1, s_1, v_1) and (h_2, s_2, v_2) is defined as

$$d_E((h_1, s_1, v_1), (h_2, s_2, v_2)) = \sqrt{(h_1 - h_2)^2 + (s_1 - s_2)^2 + (v_1 - v_2)^2}$$

Solution

The hue value of a color is represented in a degrees. Therefore, a hue value of $359°$ is very similar to a hue value of $0°$. However, the Euclidean distance between those hue value is quite large. Therefore, the Euclidean distance is not a reasonable metric to measure color similarity in the HSV color space.

(other explanations are also possible)

Question 5 (8 points)

In this task we want to apply the k-means algorithm to segment a graylevel image into foreground and background by choosing $k = 2$. The algorithm is initialized with two prototype graylevels. Then, all pixels are assigned to the two segments and the new prototype graylevels are recalculated. These two steps are repeated until convergence.
The code below shows a Matlab implementation of the k-means algorithm with $k = 2$. The function takes as inputs a graylevel image I and two prototype graylevels inside the vector c. It returns the fully segmented image J in which each pixel takes the graylevel of the respective prototype.
Unfortunately, the code contains four logical errors. Find those errors and replace them by proper code.
You get one point for finding each error and one point for fixing it. Adding errors to the code or marking a proper part as erroneous will lead to minus points. The whole task will be awarded with at least zero points.

```
function [ J ] = k_means( I, c )
   while ( true )
     I_seg = ( I + c( 1 )) .^ 2 < ( I - c( 2 )) .^ 2;
     c_update( 1 ) = sum( sum( I .* I_seg )) / sum( sum( I_seg ));
     c_update( 2 ) = sum( sum( I * ( 1 - I_seg ))) / sum( sum( 1 - I_seg ));

     if ( c = c_update )
       J = c( 1 ) * I_seg + c( 2 ) * ( 1 - I_seg );
     else
       c = c_update;
     end
   end
end
```

Solution

```
function [ J ] = k_means2( I, c )
   while ( true )
     I_seg = ( I - c( 1 )) .^ 2 < ( I - c( 2 )) .^ 2;
     c_update( 1 ) = sum( sum( I .* I_seg )) / sum( sum( I_seg ));
     c_update( 2 ) = sum( sum( I .* ( 1 - I_seg ))) / sum( sum( 1 - I_seg
));

     if ( c == c_update )
       J = c( 1 ) * I_seg + c( 2 ) * ( 1 - I_seg );
       break;
     else
       c = c_update;
     end
   end
end
```

Question 6 (7 points)

Assume a camera with thin lens which is described by its matrix of intrisic parameters

$$A = \begin{pmatrix} 500 & 0 & 400 \\ 0 & 500 & 300 \\ 0 & 0 & 1 \end{pmatrix}$$

We know that the optical focal length of the lens is $f_{lens} = 4.5$ mm and that the spacing of two neighboring pixels on the imager is 0.01 mm. Calculate the distance in front of the camera at which objects are mapped in a perfectly sharp way to the image plane.

Remark: Mind that the focal length of the lens f_{lens} is different from the focal length of the camera f_{camera}. Use millimeters as the unit for the focal lengths f_{lens} and f_{camera} and pixel per millimeter as the unit for the scaling factors α and β.

Solution

From A we obtain that $\alpha' = \beta' = 500$ pixels. Since $\alpha' = \alpha \cdot f_{camera}$ and $\alpha = 100\ \frac{pixels}{mm}$ we obtain $f_{camera} = \frac{500\ pixels}{100\ \frac{pixels}{mm}} = 5\ mm$. f_{camera} models the distance between the lens and the image plane. Together with the lens equation we get the distance z as

$$\frac{1}{f_{lens}} = \frac{1}{f_{camera}} + \frac{1}{z} \Rightarrow z = \left(\frac{1}{f_{lens}} - \frac{1}{f_{camera}}\right)^{-1} = 45\ mm$$

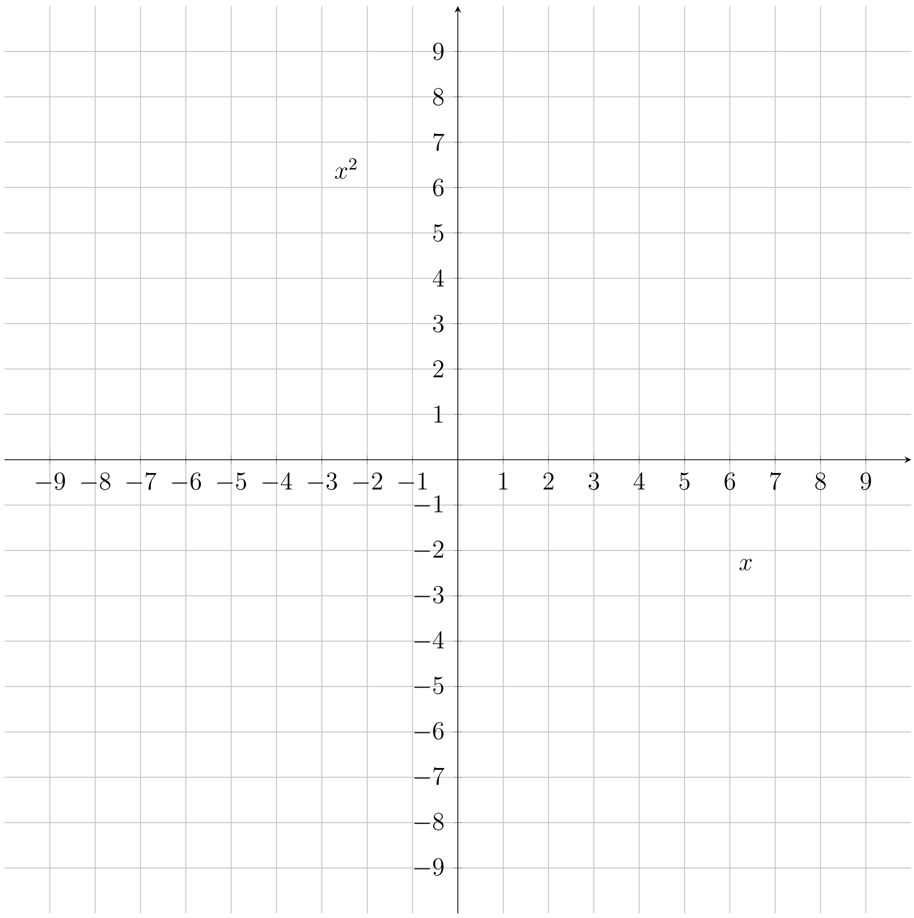
Question 7 (5 + 2 points)

Assume a binary classification problem with the six one-dimensional patterns shown in the table below.

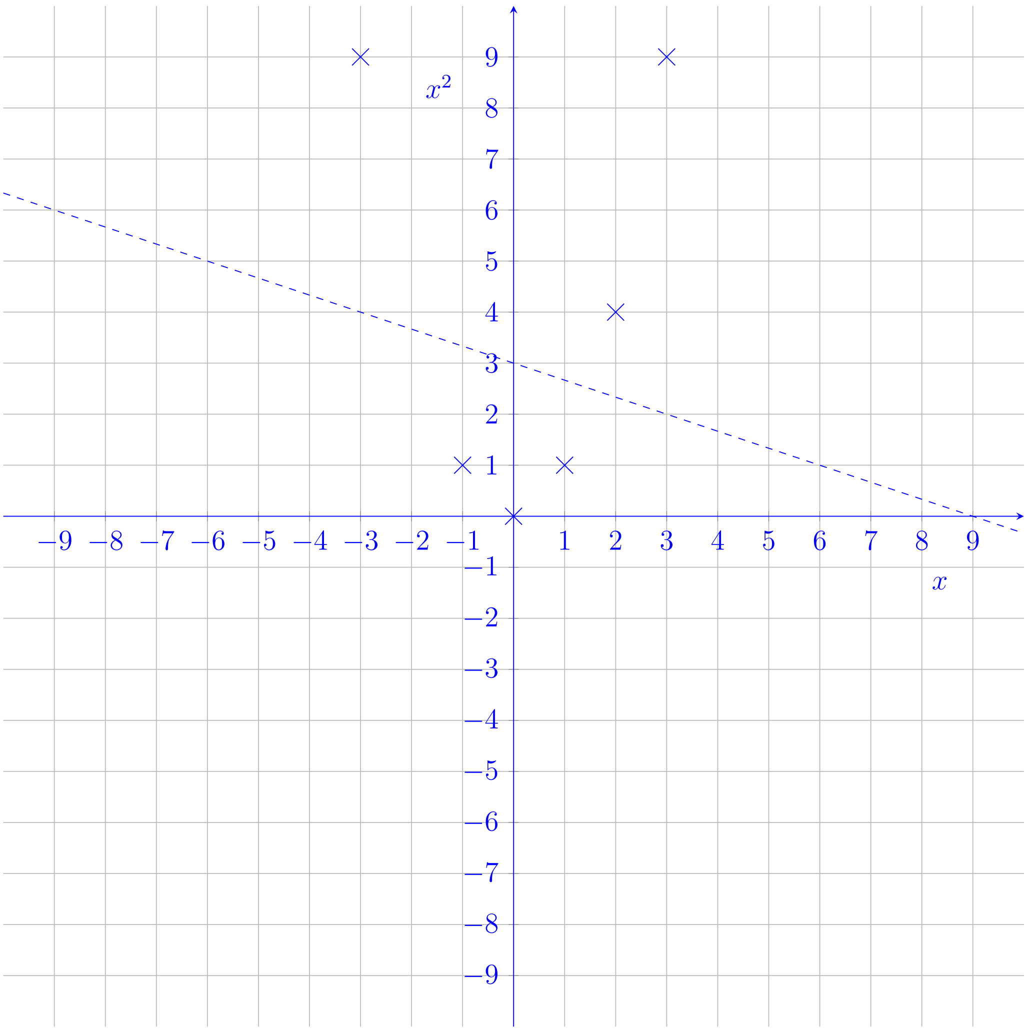
(a) Apply the transform $\Phi_1(x) = \begin{pmatrix} x \\ x^2 \end{pmatrix}$ to each pattern and add the result in the diagram below. Are the patterns linearly separable after the non-linear transform Φ_1? Justify your answer briefly.

(b) Check whether the problem is linearly separable using the non-linear transformation $\Phi_2(x) = \begin{pmatrix} x^3 \\ x^2 \end{pmatrix}$ rather than Φ_1. Justify your answer briefly.

pattern No.	class label $d^{(i)}$	pattern $x^{(i)}$
1	$+1$	-3
2	-1	-1
3	-1	0
4	-1	1
5	$+1$	2
6	$+1$	3



Solution



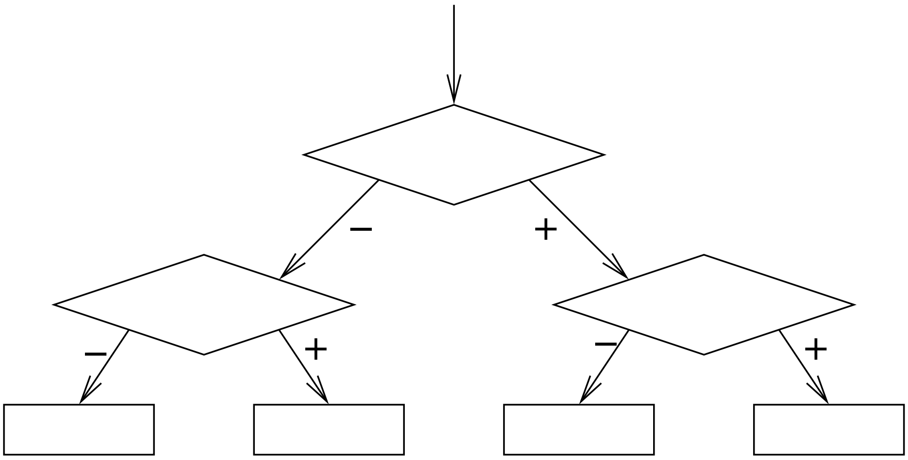
(a) The patterns in feature space are linear separable e.g. by the classifier in the graph.

(b) Still, the patterns are separable with the same line as in (a)

Question 8 (10 points)

For a binary classification problem we want to train a decision tree with three internal nodes and four leaf nodes as depicted below. We have 16 training examples $(\vec{x}^{(i)}, d^{(i)})$ and can choose among five possible binary classifiers $c_1, \dots, c_5$. The true class label $d^{(i)}$ of each training example $\vec{x}^{(i)}$ and the output of each of the five classifiers are listed in the table below.

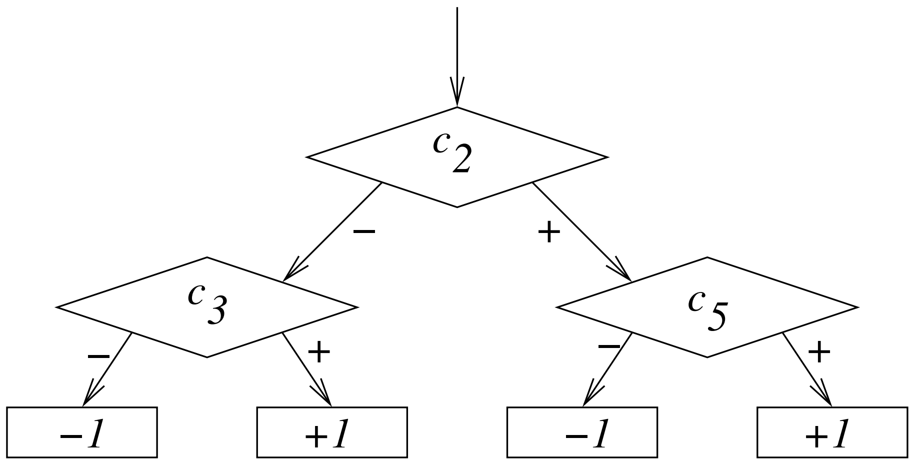
Select the best classifier for each internal node and provide the class labels to the leaf nodes according to the training algorithm introduced in the lecture. You may enter your solution in the illustration of the decision tree.



example	label $d^{(i)}$	output of classifiers c_1	c_2	c_3	c_4	c_5
$\vec{x}^{(1)}$	$+1$	$+1$	-1	$+1$	-1	$+1$
$\vec{x}^{(2)}$	$+1$	$+1$	-1	$+1$	-1	-1
$\vec{x}^{(3)}$	$+1$	$+1$	$+1$	-1	-1	$+1$
$\vec{x}^{(4)}$	$+1$	$+1$	$+1$	-1	-1	$+1$
$\vec{x}^{(5)}$	$+1$	-1	$+1$	$+1$	-1	$+1$
$\vec{x}^{(6)}$	$+1$	-1	$+1$	-1	$+1$	$+1$
$\vec{x}^{(7)}$	$+1$	-1	$+1$	$+1$	$+1$	$+1$
$\vec{x}^{(8)}$	$+1$	-1	$+1$	$+1$	$+1$	$+1$
$\vec{x}^{(9)}$	-1	$+1$	-1	-1	$+1$	$+1$
$\vec{x}^{(10)}$	-1	$+1$	-1	-1	$+1$	$+1$
$\vec{x}^{(11)}$	-1	$+1$	-1	-1	$+1$	-1
$\vec{x}^{(12)}$	-1	$+1$	-1	-1	-1	-1
$\vec{x}^{(13)}$	-1	-1	$+1$	$+1$	-1	$+1$
$\vec{x}^{(14)}$	-1	-1	$+1$	$+1$	$+1$	-1
$\vec{x}^{(15)}$	-1	-1	-1	-1	$+1$	$+1$
$\vec{x}^{(16)}$	-1	-1	-1	-1	$+1$	$+1$

Solution

For the root node we select the classifier with the best performance on all examples. If we count the errors we find that c_1 makes 8 errors, c_2 makes 4 errors, c_3 makes 5 errors, c_4 makes 11 errors, c_5 makes 6 errors. Hence we select c_2. The examples 1, 2, 9–12, 15, 16 go to the left branch, examples 3–8, 13, 14 to the right branch. Classifier c_3 is selected for the internal node on the left branch since it partitions all examples on the left branch correctly. The two leaf nodes in the left branch are labeled with -1 and +1. On the right branch we don't find a classifier that is perfect for all examples, however we find that c_5 only makes one error for example 13. The two leaf nodes are labeled with -1 and +1. The result is shown in the figure below.



Question 9 (8 points)

Create a multi layer perceptron (MLP) that implements the function

$$x \mapsto \begin{cases} 0 & \text{if } x \leq -4 \\ x+4 & \text{if } -4 < x \leq 0 \\ 2 \cdot x + 4 & \text{if } 0 < x \end{cases}$$

The MLP should be composed out of at most four perceptrons. For each perceptron provide the activation function and the weights. Don't forget the bias weight. Draw the structure of the MLP.

Solution

There are several solutions possible. The smallest network uses two perceptrons. We present a solution with two perceptrons and one with three perceptrons.

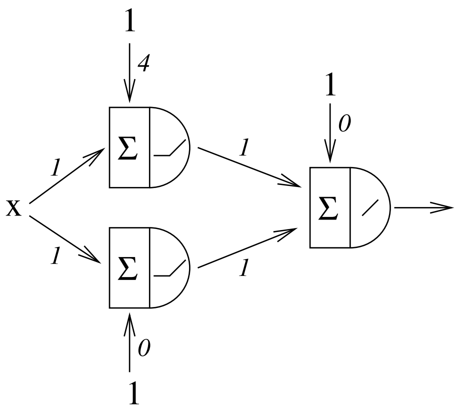
- Solution with 3 perceptrons: The left part ($x \leq 0$) can be modeled with a single perceptron with ReLU activation function and weight 1 for the input from x and bias weight 4. The second hinge can be generated by a second perceptron that takes x as input, uses the ReLU activation function and has a weight of 1 for the input from x and bias weight 0. The output of these two perceptrons are added in a third perceptron, with both input weights being equal to 1 and the bias weight being equal to 0. The third perceptron uses the linear activation function (ReLU would also be possible).
- Solution with 2 perceptrons: A solution with just two perceptrons with ReLU activation function can be build with a structure as shown in the plot below on the right side. This MLP calculates the function

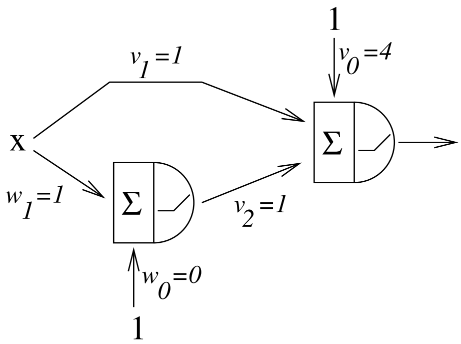
  $$x \mapsto \max\{0, v_0 + v_1 x + v_2 \max\{0, w_0 + w_1 x\}\}$$

  If we assume $v_2 > 0$ we can simplify the term and obtain

  $$x \mapsto \max\{0, v_0 + v_1 x, (v_0 + v_2 w_0) + (v_1 + v_2 w_1)x\}$$

  The first entry 0 can be used to model the target function for values of $x \leq 4$. The second entry $v_0 + v_1 x$ can be used to model the interval $-4 < x \leq 0$ if we choose $v_0 = 4$ and $v_1 = 1$. The third entry $(v_0 + v_2 w_0) + (v_1 + v_2 w_1)x$ can be used to model the function for $0 < x$. We can choose $v_2 > 0$ arbitrarily, e.g. $v_2 = 1$. Then, $w_0 = 0$ and $w_1 = 1$ solve the task.





Gesamtpunkte: 60