Machine Learning: Exercise Sheet 4

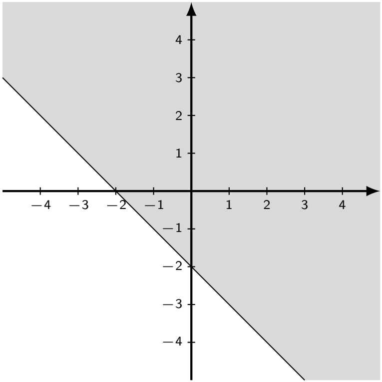
Manuel Blum
AG Maschinelles Lernen und Natürlichsprachliche Systeme
Albert-Ludwigs-Universität Freiburg

mblum@informatik.uni-freiburg.de

Exercise 1: Perceptrons

Given is a perceptron with weight vector $(w_0, w_1, w_2)^T = (2, 1, 1)^T$.

(a) Plot the partition of $\mathbb{R}^2$ that is realized by this perceptron in a diagram and mark the area where the perceptron outputs 1.



$$y = f_{step}\left(w_0 + \langle \vec{w}, \vec{x} \rangle\right)$$

Exercise 1: Perceptrons

Given is a perceptron with weight vector $(w_0, w_1, w_2)^T = (2, 1, 1)^T$.

(b) Which of the perceptrons with the following weight vectors have the same hyperplane and which represent exactly the same classification as the perceptron given above?

	$(w_0, w_1, w_2)^T$	same hyperplane	same classification
(I)	$(1, 0.5, 0.5)^T$	$\times$	$\times$
(II)	$(200, 100, 100)^T$	$\times$	$\times$
(III)	$(\sqrt{2}, \sqrt{1}, \sqrt{1})^T$		
(IV)	$(-2, -1, -1)^T$	$\times$	

Exercise 2: Perceptron Learning

(a) Apply the perceptron learning algorithm for the following pattern set until convergence. Start with weight vector $(w_0, w_1, w_2, w_3)^T = (1, 0, 0, 0)^T$. Apply the patterns in the given order cyclically. For each step of perceptron learning write down the applied pattern, the classification result and the update of the weight vector.

$$(4,3,6)^T \in \mathcal{N}, \quad (2,-2,3)^T \in \mathcal{P}, \quad (1,0,-3)^T \in \mathcal{P}, \quad (4,2,3)^T \in \mathcal{N}$$

pattern	output	classification	update	new weight vector
				$(1,0,0,0)^T$
$(1,4,3,6)^T \in \mathcal{N}$	$f_{step}(1)$	false positive	$-(1,4,3,6)^T$	$(0,-4,-3,-6)^T$
$(1,2,-2,3)^T \in \mathcal{P}$	$f_{step}(-20)$	false negative	$+(1,2,-2,3)^T$	$(1,-2,-5,-3)^T$
$(1,1,0,-3)^T \in \mathcal{P}$	$f_{step}(8)$	true positive	unchanged	unchanged
$(1,4,2,3)^T \in \mathcal{N}$	$f_{step}(-26)$	true negative	unchanged	unchanged
$(1,4,3,6)^T \in \mathcal{N}$	$f_{step}(-40)$	true negative	unchanged	unchanged
$(1,2,-2,3)^T \in \mathcal{P}$	$f_{step}(-2)$	false negative	$+(1,2,-2,3)^T$	$(2,0,-7,0)^T$
$(1,1,0,-3)^T \in \mathcal{P}$	$f_{step}(2)$	true positive	unchanged	unchanged
$(1,4,2,3)^T \in \mathcal{N}$	$f_{step}(-12)$	true negative	unchanged	unchanged
$(1,4,3,6)^T \in \mathcal{N}$	$f_{step}(-19)$	true negative	unchanged	unchanged
$(1,2,-2,3)^T \in \mathcal{P}$	$f_{step}(16)$	true positive	unchanged	unchanged

finished, weight vector $(2,0,-7,0)^T$ classifies all patterns correctly

Exercise 2: Perceptron Learning

(b) Show that the problem given by the following pattern set cannot be solved with a single perceptron. For this purpose, apply the perceptron learning algorithm for the given patterns. Start with weight vector $(w_0, w_1, w_2)^T = (1, 0, 0)^T$.

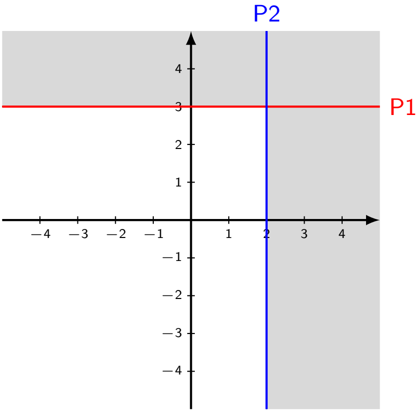
$$(1,1)^T \in \mathcal{P}, \quad (1,0)^T \in \mathcal{N}, \quad (0,0)^T \in \mathcal{P}, \quad (0,1)^T \in \mathcal{N}$$

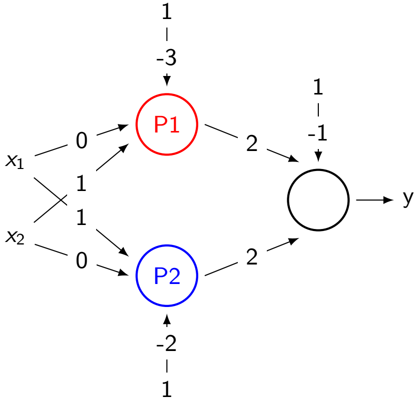
pattern	output	classification	update	new weight vector
				$(1,0,0)^T$
$(1,1,1)^T \in \mathcal{P}$	$f_{step}(1)$	true positive	unchanged	unchanged
$(1,1,0)^T \in \mathcal{N}$	$f_{step}(1)$	false positive	$-(1,1,0)^T$	$(0,-1,0)^T$
$(1,0,0)^T \in \mathcal{P}$	$f_{step}(0)$	true positive	unchanged	unchanged
$(1,0,1)^T \in \mathcal{N}$	$f_{step}(0)$	false positive	$-(1,0,1)^T$	$(-1,-1,-1)^T$
$(1,1,1)^T \in \mathcal{P}$	$f_{step}(-3)$	false negative	$+(1,1,1)^T$	$(0,0,0)^T$
$(1,1,0)^T \in \mathcal{N}$	$f_{step}(0)$	false positive	$-(1,1,0)^T$	$(-1,-1,0)^T$
$(1,0,0)^T \in \mathcal{P}$	$f_{step}(-1)$	false negative	$+(1,0,0)^T$	$(0,-1,0)^T$

finished, weight vector $(0,-1,0)^T$ occurs twice
the problem is not solvable (cycle theorem)

Exercise 3: Perceptron Networks

Develop a perceptron network with two input variables x_1 and x_2 and at most three perceptrons that exactly classifies the marked area (and the boundary). Illustrate the topology (structure) of the network and the weights of its neurons.





Exercise 4: Nonlinear Feature Spaces

The following training patterns are given:

$$(-3) \in \mathcal{N}, (-2) \in \mathcal{N}, (-1) \in \mathcal{P}, (0) \in \mathcal{P}, (1) \in \mathcal{P}, (2) \in \mathcal{N}, (3) \in \mathcal{N}$$

This pattern set is not linearly separable in $\mathbb{R}$. We will use nonlinear features in order to classify the patterns using a single perceptron.

(a) Plot the patterns in input space and mark each positive pattern with output 1 and each negative pattern with output 0.

(b) Use the function $(h_1, h_2)^T = g(x) = (x, x^2)^T$ to build a feature space for the patterns. Plot the patterns in feature space and mark each positive pattern with output 1 and each negative pattern with output 0. Show that the patterns are linearly separable in feature space (give a perceptron that classifies correctly) or derive a contradiction from the data and the model.

(c) Use the function $h = g(x) = x^3$ to build a feature space for the patterns. Plot the patterns in feature space and prove or disprove that the patterns are linearly separable in feature space.

$\rightarrow$ *blackboard*

Exercise 5: Winner-takes-all networks

Consider the pattern set and the initial prototypes in the figure below.

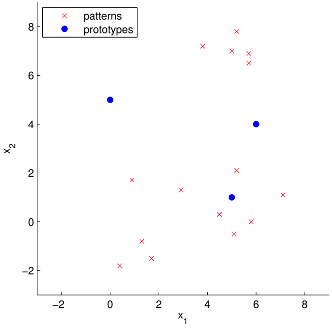
(a) For each input pattern calculate the closest prototype.

(b) Perform one update of the VQ algorithm ($\epsilon = 0.4$) for patterns 1-3 and report the resulting weight vectors of the relevant prototypes.

(c) Using the initial prototypes, perform one iteration of the k-means algorithm. What are the weight vectors of the resulting prototypes? Does the algorithm converge after the first iteration?

#	pattern
1	$(1.7, -1.5)^T$
2	$(0.9, 1.7)^T$
3	$(0.4, -1.8)^T$
4	$(1.3, -0.8)^T$
5	$(2.9, 1.3)^T$
6	$(7.1, 1.1)^T$
7	$(5.1, -0.5)^T$
8	$(5.8, 0.0)^T$
9	$(5.2, 2.1)^T$
10	$(4.5, 0.3)^T$
11	$(5.0, 7.0)^T$
12	$(5.7, 6.9)^T$
13	$(5.2, 7.8)^T$
14	$(3.8, 7.2)^T$
15	$(5.7, 6.5)^T$

#	prototype
1	$(6, 4)^T$
2	$(0, 5)^T$
3	$(5, 1)^T$



Exercise 6: Programming tasks

$\rightarrow$ Demo