<u>AUFGABE 1</u>: Klassifizierungsverfahren: ID3 vs. Perceptron

Betrachten Sie das folgende Klassifizierungsproblem:

Objekt	Attributwerte X1	X2	Klasse
A	2	3	+
B	3	2	–
C	3	3	+
D	4	3	+
E	4	1	–
F	4	2	+
G	1	1	–
H	1	3	–
I	1	2	
J	2	2	
K	2	1	
L	3	1	

1) Bestimmen Sie den ID3-Entscheidungsbaum für die Klassifizierung der Objekte A bis H. (Berechnung der Entropieabnahme für den ersten Entscheidungsschritt !).

2) Zeigen Sie (graphisch) dass das Problem linear separabel ist, d.h. auch mit Hilfe eines Perceptrons gelöst werden kann.

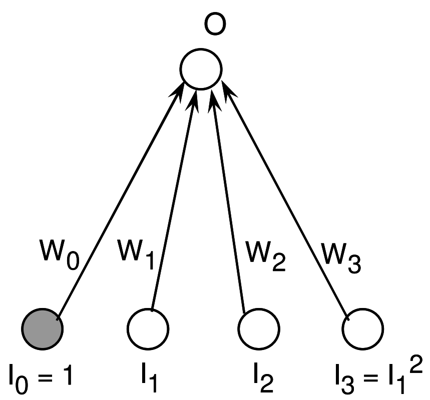
3) Vergleichen und diskutieren Sie die Klassifizierungen der Objekte I bis L durch ein Perceptron mit denjenigen des ID3-Entscheidungsbaums.

AUFGABE 2: Perceptron mit erweitertem Inputraum

Betrachten Sie die folgenden 4 Lernbeispiele *(D = gewünschter Output)* :

Beispiel	**I_1**	**I_2**	**D**
a	– 1	0	+ 1
b	0	0	– 1
c	0	1	– 1
d	1	1	+ 1

1) Zeigen Sie, dass dieses nicht linear separable Klassifizierungsproblem durch ein Perceptron mit erweitertem Inputraum gelöst werden kann:



O = sign(……..)

$$\text{sign}(x) = \begin{cases} 1 & \text{falls } x \geq 0 \\ -1 & \text{falls } x < 0 \end{cases}$$

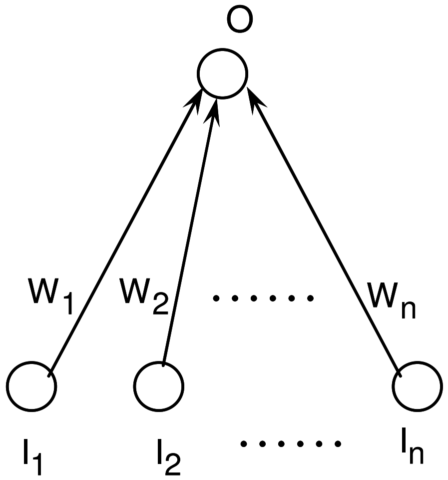
2) Wählen Sie einen konkreten Gewichtsvektor (W_0, $W_{1,}$ W_2, W_3), der das Problem löst, und bestimmen Sie im (I_1, I_2)-Raum die Trennkurven zwischen den Gebieten mit O = +1 bzw. O = –1.

   Stellen Sie diese Trennkurven auch graphisch dar.

<u>Hinweis</u>: *Sie können das Problem vereinfachen, indem Sie sich auf Lösungen mit $W_1 = 0$ beschränken!*

AUFGABE 3: Gradienten-Lernverfahren

Betrachten Sie ein Perceptron mit sigmoidem Output,



$$O = \frac{1-e^{-V}}{1+e^{-V}}$$

$$V = \sum_{i=1}^{n} W_i\, I_i$$

und die Fehlerfunktion

$$F = -(1 + D)\, \ell n(1+O) - (1 - D)\, \ell n(1 - O) \quad ,$$

(D = gewünschter Output).

1) Bestimmen Sie analytisch die Gewichtsänderungen ΔW_i , die sich aus der Minimierung von F mit Hilfe eines Gradientenverfahrens ergeben.

   [Hinweis: $\frac{\partial O}{\partial V} = \frac{1}{2}\left(1 - O^2\right)$.]

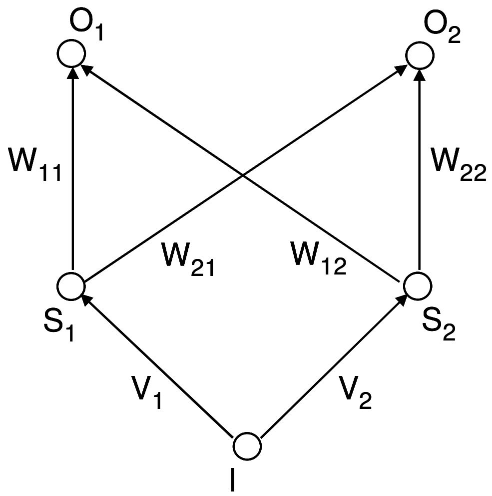
2) Zu Beginn seien alle $W_i = 0$ (d.h. O = 0 für jeden Input).

   Betrachten Sie nun ein Lernbeispiel mit D = 1, bei dem alle Inputs I_i entweder +1 oder –1 sind.

   Wie gross ist der Output O für ein solches Lernbeispiel nach einem Lernschritt mit $\eta = \frac{2}{n}$?

<u>AUFGABE 4</u>: Error-Backpropagation

Gegeben sei das folgende Feedforward-Netzwerk,



$$O_1 = f\,(W_{11}\,S_1 + W_{12}\,S_2)$$

$$O_2 = f\,(W_{21}\,S_1 + W_{22}\,S_2)$$

$$S_1 = f\,(V_1\,I)$$

$$S_2 = f\,(V_2\,I)$$

$$f(x) = \frac{1}{1+e^{-x}}$$

Berechnen Sie die Änderung von V_1 beim Backpropagation-Lernverfahren mit folgender Fehlerfunktion:

$$F = \frac{1}{2}(D_1 - O_1)^2 + \frac{1}{2}(D_2 - O_2)^2.$$

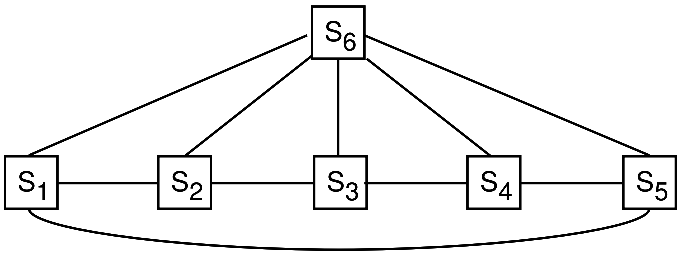
(D_1, D_2 = gewünschte Outputs)

Stellen Sie ΔV_1 explizit als Funktion von O_1, D_1, O_2, D_2, S_1, I und den W_{ij} dar!

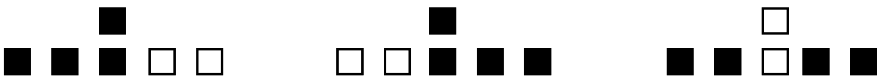
<u>Hinweis</u>: $\frac{df}{dx} = f(x)\,[1 - f(x)]$

<u>AUFGABE 5</u>: Hopfield-Netzwerk

Betrachten Sie das folgende (nicht vollständig verknüpfte) Hopfield-Netzwerk,



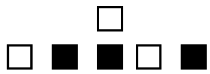
und benutzen Sie die Hebb'sche Lernregel, um in diesem Netzwerk die drei Muster A, B und C zu speichern:



A B C

(wobei: ■ : $S_i = 1$, □ : $S_i = -1$)

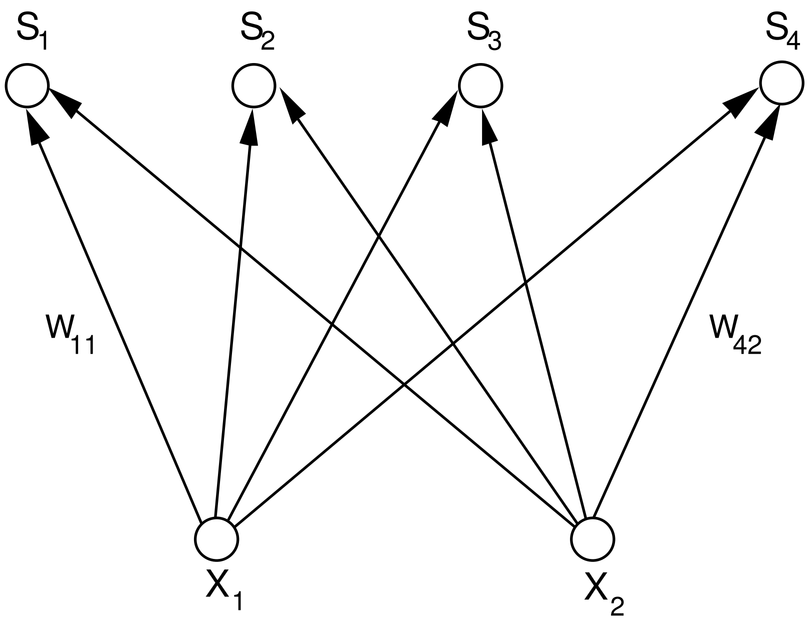
1) Welche Werte ergeben sich für die Gewichte der 10 Verbindungen?

2) Gegen welches Muster konvergiert das Anfangsmuster  wenn Sie die S_i der Reihe nach (S_1, S_2, ...) aufdatieren?

3) Vergleichen Sie die Energie des Anfangsmusters mit derjenigen des stabilen Endmusters.

AUFGABE 6: Winner-Take-All Netzwerk, Kompetitives Lernen

Betrachten Sie ein Netzwerk mit 2 Inputs und 4 Winner-Take-All Output-Neuronen:



$W_{11} = 0.2$	$W_{21} = 0.8$	$W_{31} = 0.6$	$W_{41} = 0.4$
$W_{12} = 0.2$	$W_{22} = 0.2$	$W_{32} = 0.8$	$W_{42} = 0.8$

1) Bestimmen Sie die Winner-Gebiete der 4 Outputneuronen im Inputraum ($0 < X_1, X_2 < 1$).

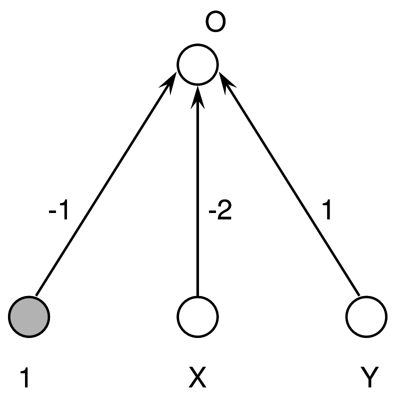
2) Präsentieren Sie dann die beiden Lernbeispiele ($X_1 = 0.4$, $X_2 = 0.5$) und ($X_1 = 0.9$, $X_2 = 0.8$) je einmal, und ändern Sie die Gewichte gemäss dem einfachen kompetitiven Lernverfahren [d.h. nur die Gewichte des Winner-Neurons werden geändert] mit $\eta = 2/3$.

3) Bestimmen Sie die neuen Winner-Gebiete im Inputraum.

<u>Hinweis</u>: Alle Aufgaben lassen sich rein graphisch lösen!

AUFGABE 7: Stochastische Neuronen, Reinforcement Learning

1) Für welche Werte von X und Y ist beim folgenden stochastischen Neuron (Standard-Definition) die Wahrscheinlichkeit, dass O den Wert +1 annimmt, grösser als 0.5 ?



2) Wie werden in einem stochastischen Feedforward-Netzwerk die Gewichtsänderungen ΔW_{ij} berechnet (A_{R-P} – Lernverfahren ohne Bestrafung)?

3) Welche Vorteile hat Reinforcement Learning (z.B. das A_{R-P} – Lernverfahren) im Vergleich mit Gradienten-Lernverfahren (z.B. Error-Backpropagation)?

AUFGABE 8: Lernverfahren und Lernverhalten

1) Warum und wie kann man Hopfield-Netzwerke (bzw. Hopfield-Tank-Netzwerke) auch dazu benutzen, gewisse Typen von Optimierungsproblemen zu lösen?

2) Welche Probleme können beim Error-Backpropagation-Lernverfahren auftreten?

   Was sind die Ursachen für diese Probleme?

   Beschreiben Sie mögliche Massnahmen gegen diese Probleme.

3) Beschreiben Sie die wichtigsten Schritte eines einfachen stochastischen Lernverfahrens (z.B. “Iterative Improvement“).

4) Zur Lösung eines Problems mit 10 Inputs und einem Output möchten Sie ein Multilayer-Perceptron verwenden, das 2 Schichten mit je 4 Hidden Units enthält (aufeinander folgende Schichten sind vollständig verknüpft!).

   Wie viele Lernbeispiele sollten mindestens zur Verfügung stehen, damit das Sinn macht?