IFM 5.6 Künstliche Intelligenz, 2017/18

Blatt 7: Perzeptron, MLP (14 Punkte)

Carsten Gips, FH Bielefeld

Praktikum: 19.12.2017

ILIAS: 19.12.2017

1 Perzeptron und Lernalgorithmus (4 Punkte)

a) Implementieren Sie den Perzeptron-Lernalgorithmus (vgl. VL 08). Ihr Programm soll eine Tabelle mit den Spalten

w_1, w_2, $-\theta$, x_1, x_2, k, $\sum_{i=1}^{n} w_i x_i - \theta$, vorhergesagte Klasse, Änderung? (=,+,-)

ausgeben, die eine Zeile je Lernschritt enthält.

Lernen Sie die boolesche Funktion `OR`. Geben Sie die vollständige Lerntabelle an.

Anders als in der Vorlesung besprochen, sollen die Gewichte w_1 und w_2 sowie die Schwelle θ jeweils mit dem Wert 0 initialisiert werden. Die Lernschrittweite η sei 0.5. Nutzen Sie die sign-Funktion als Aktivierungsfunktion:

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

b) Zeichnen Sie eine geometrische Interpretation des Perzeptrons nach dem Lernen (Trennebene, Gewichtsvektor, Trainingsbeispiele).

Thema: Verständnis Perzeptron und Ablauf Perzeptron-Lernalgorithmus

2 Perzeptron und Entscheidungsgrenze (2 Punkte)

a) Betrachten Sie das durch den folgenden Gewichtsvektor gegebene Perzeptron: $(w_0, w_1, w_2)^T = (2, 1, 1)^T$ (mit der sign-Funktion als Transferfunktion).

Zeichnen Sie die Trennebene und markieren Sie den Bereich, der mit +1 klassifiziert wird.

b) Welche der folgenden Perzeptrons haben die selbe Trennebene? Welche weisen exakt die gleiche Klassifikation auf?

- $(w_0, w_1, w_2)^T = (1, 0.5, 0.5)^T$
- $(w_0, w_1, w_2)^T = (200, 100, 100)^T$
- $(w_0, w_1, w_2)^T = (\sqrt{2}, \sqrt{1}, \sqrt{1})^T$
- $(w_0, w_1, w_2)^T = (-2, -1, -1)^T$

Thema: Verständnis Interpretation Perzeptron (Trennebene)

3 Perzeptron-Lernregel (2 Punkte)

Zeigen Sie, dass durch die Anwendung der Perzeptron-Lernregel die Trennebene tatsächlich in die richtige Richtung verschoben/verdreht wird, d.h. dass für den neuen Gewichtsvektor $\mathbf{w}_{neu}$ gilt:

- Falls der Eingabevektor $\mathbf{x}$ die Klasse 1 hat, dann ist $\langle \mathbf{w}_{neu}, \mathbf{x} \rangle \geq \langle \mathbf{w}_{alt}, \mathbf{x} \rangle$.
- Anderenfalls gilt $\langle \mathbf{w}_{neu}, \mathbf{x} \rangle \leq \langle \mathbf{w}_{alt}, \mathbf{x} \rangle$.

Dabei sind $\mathbf{w}$ und $\mathbf{x}$ die erweiterten Gewichts- und Eingabevektoren.

Thema: Verständnis Perzeptron-Lernregel

4 Lineares MLP (2 Punkte)

Gegeben sei ein MLP mit linearen Transferfunktionen, d.h. für jedes Neuron berechnet sich der Output durch die gewichtete Summe der Inputs: $y = g(\langle \mathbf{w}, \mathbf{x} \rangle)$, wobei $g(a) = a$ gilt, also $y = \langle \mathbf{w}, \mathbf{x} \rangle$.

Zeigen Sie, dass dieses Netz durch eine einzige Schicht mit linearen Neuronen ersetzt werden kann. Betrachten Sie dazu ein zweilagiges Netz bestehend aus einer Ausgabe- und einer versteckten Schicht.

Thema: Verständnis Bedeutung nichtlinearer Transferfunktionen

5 XOR (2 Punkte)

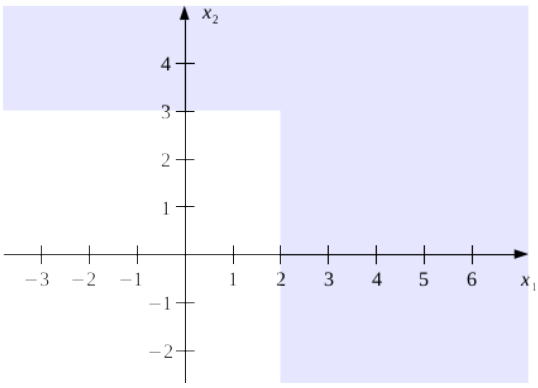
Konstruieren[1] Sie ein zweischichtiges MLP (bestehend aus einer Ausgabe- und einer versteckten Schicht), welches ein `XOR` implementiert. Als Transferfunktion soll wieder die sign-Funktion genutzt werden.

Zeichnen Sie die Architektur Ihres Netzes auf und geben Sie die Gewichte an und weisen Sie nach, dass Ihr Netz tatsächlich ein `XOR` implementiert.

Thema: Funktionsweise Perzeptron, Konstruktion von Perzeptron-Netzen

6 Perzeptron-Netze (2 Punkte)

Konstruieren Sie nun ein Netz mit drei Perzeptrons (Transferfunktion wieder die sign-Funktion), welches für zwei Eingabevariablen x_1 und x_2 die in der folgenden Abbildung blau-grau dargestellten Bereiche mit $+1$ klassifiziert.



Thema: Funktionsweise Perzeptron, Konstruktion von Perzeptron-Netzen

[1] Konstruieren, nicht trainieren/lernen!