Einführung in die

Künstliche Intelligenz

WS12/13 - Prof. Dr. J. Fürnkranz, Prof. Dr. U. Brefeld

TECHNISCHE UNIVERSITÄT DARMSTADT

Beispiellösung für das 6. Übungsblatt (29.01.2013)

Aufgabe 1 Neuronale Netze

a)

$$\begin{aligned} in_h &= W_{x,h} \cdot x + W_{y,h} \cdot y = 0.3 \cdot 1 + 0.7 \cdot (-1) = -0.4 \\ in_i &= W_{x,i} \cdot x + W_{y,i} \cdot y = -0.1 \cdot 1 + (-0.4) \cdot (-1) = 0.3 \\ in_j &= W_{x,j} \cdot x + W_{y,j} \cdot y = 0.2 \cdot 1 + (-0.6) \cdot (-1) = 0.8 \end{aligned}$$

Die angegebene Aktivierungsfunktion $g(x) = x$ gibt die Aktivierungswerte unverändert weiter, d.h. $out_x = in_x$.

$$\begin{aligned} in_a &= W_{h,a} \cdot out_h + W_{i,a} \cdot out_i + W_{j,a} \cdot out_j = (-0.3) \cdot (-0.4) + (-0.8) \cdot 0.3 + (-0.4) \cdot 0.8 = -0.44 \\ in_b &= W_{h,b} \cdot out_h + W_{i,b} \cdot out_i + W_{j,b} \cdot out_j = 0.2 \cdot (-0.4) + 0.2 \cdot 0.3 + 0.6 \cdot 0.8 = 0.46 \end{aligned}$$

Die Ausgabewerte bleiben wiederum unverändert.

b)

1. Berechnen Sie die Fehlerterme Δ_a und Δ_b

   $g'(x) = 1$
   $\Delta_a = Err_a \cdot g'(in_a) = (-0.2 - (-0.44)) \cdot 1 = 0.24$
   $\Delta_b = Err_b \cdot g'(in_b) = (0.9 - 0.46) \cdot 1 = 0.44$

2. Berechnen Sie die Fehlerrate Δ_h

   $\Delta_h = W_{h,a} \cdot \Delta_a \cdot g'(in_h) + W_{h,b} \cdot \Delta_b \cdot g'(in_h) = (-0.3) \cdot 0.24 + 0.2 \cdot 0.44 = -0.072 + 0.088 = 0.016$

3. Berechnen Sie die Gewichtsänderung für das Gewicht $W_{h,a}$

   $W_{h,a} \leftarrow W_{h,a} + \alpha \cdot \Delta_a \cdot out_h = -0.3 + 0.5 \cdot 0.24 \cdot (-0.4) = -0.348$

c) Mit der Identität als Aktivierungsfunktion stellt jedes Neuron eine Linearkombination ihrer Eingaben dar. Da Linearkombinationen von Linearkombinationen wiederum Linearkombinationen sind, können im ersten Fall nur lineare Funktionen gelernt werden. Ist jede beliebige Aktivierungsfunktion erlaubt, können alle stetigen Funktionen gelernt werden.

Aufgabe 2 Lineare Modelle

Die Strecke zwischen den beiden Klassenmittelpunkten ist gegeben durch $\bar{x}_{+1} - \bar{x}_{-1}$. Da die Hyperebene orthogonal zur Strecke sein soll können wir direkt $w = \bar{x}_{+1} - \bar{x}_{-1}$ setzen (w zeigt in Richtung der positiven Klasse). Das lineare Modell hat jetzt die Form

$$f(x) = (\bar{x}_{+1} - \bar{x}_{-1})x + b$$

Die Hyperebene soll die Strecke $\bar{x}_{+1} - \bar{x}_{-1}$ in der Mitte schneiden. Das bedeutet, dass $(\bar{x}_{+1} + \bar{x}_{-1})/2$ auf der Ebene liegen muss. Einsetzen in die Gleichung ergibt

$$f\left((\bar{x}_{+1} + \bar{x}_{-1})/2\right) = (\bar{x}_{+1} - \bar{x}_{-1})(\bar{x}_{+1} + \bar{x}_{-1})/2 + b \stackrel{!}{=} 0$$

Umformen ergibt

$$b = -(\bar{x}_{+1} - \bar{x}_{-1})(\bar{x}_{+1} + \bar{x}_{-1})/2 = -\frac{1}{2}(\bar{x}_{+1}^2 - \bar{x}_{-1}^2)$$

und die Entscheidungsfunktion ist gegeben durch

$$f(x) = (\bar{x}_{+1} - \bar{x}_{-1})x - \frac{1}{2}(\bar{x}_{+1}^2 - \bar{x}_{-1}^2).$$