

6. Übungsblatt

Aufgabe 22 Aktivierungsfunktionen

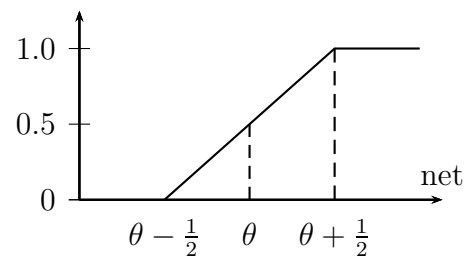
Es ist wichtig, dass mindestens eine Schicht eines mehrschichtigen Perzeptrons nichtlineare Aktivierungsfunktionen benutzt, da sonst seine Fähigkeiten stark eingeschränkt sind: Zeigen Sie, dass ein mehrschichtiges Perzeptron mit beliebig vielen Schichten, in dem die Netzeingabefunktion jedes Neurons die gewichtete Summe seiner Eingänge (abzüglich eines Biaswertes θ) und sowohl die Aktivierungsfunktion als auch die Ausgabefunktion jedes Neurons lineare Funktionen sind, äquivalent ist zu einem zweischichtigen Perzeptron mit den gleichen Eigenschaften!

Hinweis: Die Struktur eines neuronalen Netzes wird manchmal auch, wie in der Vorlesung erwähnt, durch eine Gewichtsmatrix dargestellt.

Aufgabe 23 Gradientenabstieg

Gegeben sei ein zweischichtiges Perzeptron mit n Eingängen und einem Ausgang, in dem das Ausgabeneuron die gewichtete Summe der Eingänge als Netzeingabefunktion, eine semilineare Funktion

$$f_{\text{act}}(\text{net}, \theta) = \begin{cases} 1, & \text{wenn } \text{net} > \theta + \frac{1}{2}, \\ 0, & \text{wenn } \text{net} < \theta - \frac{1}{2}, \\ (\text{net} - \theta) + \frac{1}{2}, & \text{sonst,} \end{cases}$$



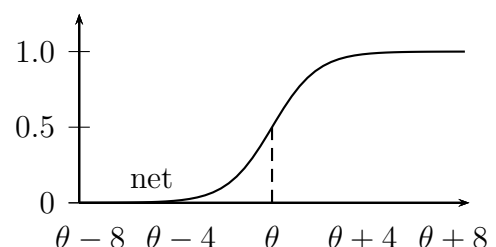
(siehe Zeichnung) als Aktivierungsfunktion und die Identität als Ausgabefunktion besitzt.

Leiten Sie die Änderungsregel für die Gewichte ab, die sich aus einem Ansatz mit Gradientenabstieg ergibt, wenn der Netzfehler als Summe (über alle Muster) der quadrierten Differenzen von gewünschter und tatsächlicher Ausgabe berechnet wird!

Aufgabe 24 Gradientenabstieg

Gegeben sei ein zweischichtiges Perzeptron mit n Eingängen und einem Ausgang, in dem das Ausgabeneuron die gewichtete Summe der Eingänge als Netzeingabefunktion, eine logistische Funktion

$$f_{\text{act}}(\text{net}, \theta) = \frac{1}{1 + e^{-(\text{net} - \theta)}}$$



(siehe Diagramm) als Aktivierungsfunktion und die Identität als Ausgabefunktion besitzt.

Leiten Sie die Änderungsregel für die Gewichte für ein einzelnes Trainingsmuster ab, die sich aus einem Ansatz mit Gradientenabstieg ergibt, wenn der Netzfehler als absoluter Betrag der Differenz von gewünschter und tatsächlicher Ausgabe berechnet wird! Welche Änderungen müsste man (bei mehrschichtigen Perzeptren) an dem Rückpropagationsverfahren vornehmen?

Aufgabe 25 Fehlerrückpropagation

Betrachten Sie ein mehrschichtiges Perzeptron mit 4 Eingängen, 5 Neuronen in der versteckten Schicht und 3 Ausgabeneuronen, das ein 4-dimensionales Dreiklassenproblem lösen soll. Für das Training dieses Netzes soll die Fehlerrückpropagation verwendet werden. Des Weiteren wird angenommen, dass

- die Ausgabeneuronen den *tangens hyperbolicus* und
- die Neuronen der versteckten Schicht die logistische Funktion

als Aktivierungsfunktion benutzen.

- a) Definieren Sie für das Netz sinnvolle Ausgabewerte für die Trainingsdaten für jede der drei Klassen.
- b) Entwickeln Sie die Formel für die Gewichtsänderung der Verbindung des 3. versteckten Neurons zum 2. Ausgabeneuron. Seien Sie dabei so genau wie möglich.
- c) Entwickeln Sie die Formel für die Gewichtsänderung der Verbindung des 1. Eingabeneurons zum 3. versteckten Neuron.