MACHINE LEARNING, 5. SEMINAR – NEURON

Aufgabe 1. Gegeben sei eine Stichprobe $(x^1,\ldots,x^l)$ von Datenpunkten $x^l \in \mathbb{R}^n$. Man finde unter allen Hyperebenen, die durch den Ursprung gehen, diejenige, deren *summarischer* vorzeichenbehafteter Abstand[1] zu allen Datenpunkten maximal ist.
Hinweis: Eine Hyperebene wird durch ihren Normalenvektor w beschrieben (Der Normalenvektor sei normiert, d.h. $\|w\| = 1$).

a) Geben Sie den vorzeichenbehafteten Abstand eines Datenpunktes x von der Hyperebene als Funktion von x und w an. Wie ergibt sich die Zielfunktion der Aufgabe?
b) Welcher Vektor w^* maximiert die Zielfunktion?

Aufgabe 2. Für Muster $x \in \mathbb{R}^n_{++}$ (d.h. $x_i > 0$), die in zwei Klassen zu teilen sind, wird die Menge der Entscheidungsregeln

$$\prod_i x_i^{w_i} \gtrless b$$

betrachtet. Dabei sind der Gewichtsvektor $w \in \mathbb{R}^n$ sowie der Schwellwert b Parameter der Entscheidungsregel. Wie kann man die unbekannten Parameter anhand einer klassifizierten Lernstichprobe $\big((x^l,k^l),\ldots\big)$ mithilfe des Perzeptron Algorithmus lernen?

Aufgabe 3. Man betrachte die quadratischen Entscheidungsregeln $f : \mathbb{R}^n \to \{0,1\}$:

$$f(x) = \begin{cases} 1 & \text{wenn} \quad x^T \cdot A \cdot x + \langle x, b\rangle + c < 0 \\ 0 & \text{sonst,} \end{cases}$$

mit einer $n \times n$ Matrix A, einem Vektor $b \in \mathbb{R}^n$ und einer Konstante c. Zeigen Sie, dass man eine solche Entscheidungsregel mit Hilfe des Perzeptron Algorithmus anlernen kann.
Hinweis: Transformieren Sie dazu den Inputraum $\mathbb{R}^n$ in einen geeignet gewählten höherdimensionalen Raum, in dem die Entscheidungsregel linear wird.

Aufgabe 4. Man betrachte die „gewöhnliche“ Perzeptron Aufgabe, in der die unbekannten Parameter einer Entscheidungsregel $\langle x, w\rangle \gtrless b$ anhand einer klassifizierten Lernstichprobe zu bestimmen sind. Zusätzlich sei verlangt, dass bestimmte Parameter positiv sein müssen, zum Beispiel für eine bestimmte Komponente i_* des Gewichtsvektors $w_{i_*} > 0$ gilt. Wie lässt sich der Perzeptron Algorithmus ergänzen so, dass er nur solche Entscheidungsregel zulässt, die diese Nebenbedingungen erfüllen?

[1] Positiv von einer Seite der Hyperebene und negativ von der anderen Seite.

Aufgabe 5. Betrachten Sie die boolesche Formel

$$y = (x_1 \vee \bar{x}_2) \wedge x_3,$$

die jedem Dreitupel $x = (x_1, x_2, x_3)$ von booleschen Variablen $x_i \in \{0,1\}$ (Input) einen booleschen Wert $y \in \{0,1\}$ (Output) zuordnet. Zeigen Sie, dass sich diese Abbildung mithilfe eines „Netzes“ aus Schwellwertneuronen repräsentieren lässt. Die Neuronen des Netzes realisieren elementare (z.B. nur zweistellige) boolesche Funktionen und sind mit einander so verbunden, dass der Output eines Neurons als Input für ein anderes Neuron dienen kann.

Verallgemeinern Sie die obige Aussage auf beliebige boolesche Formeln, d.h. beliebige Abbildungen $\{0,1\}^n \to \{0,1\}$.