MUSTERERKENNUNG, 1. 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 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 Merkmalsraum 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?

Aufgabe 5. Der „Fischer Klassifikator“ [Fischer, 1936] ist ein Verfahren zur Klassifikation von Mustern $x \in \mathbb{R}^n$ in $|K|$ Klassen $k \in K$. Als Parameter dienen $|K|$ Vektoren $w^k \in \mathbb{R}^n$, $k = 1, \ldots, |K|$. Mit Hilfe dieser wird der Raum $\mathbb{R}^n$ in $|K|$ Gebiete $X^k \subset \mathbb{R}^n$ wie folgt zerlegt:

$$X^k = \{x \in \mathbb{R}^n \mid \langle x, w^k \rangle > \langle x, w^i \rangle \quad \forall i \neq k\}.$$

Die Lernaufgabe besteht in der Konstruktion eines solchen Klassifikators anhand gegebener Trainingsdaten $(x^1, k^1), \ldots, (x^l, k^l)$. Zeigen Sie, dass diese Aufgabe durch eine verallgemeinerte Variante des Perzeptron Algorithmus gelöst werden kann.
Hinweis: In jedem Lernschritt werden ein Muster (x, k) (aus den Trainingsdaten) und ein i bestimmt, für die $\langle x, w^k \rangle > \langle x, w^i \rangle$ nicht erfüllt ist. Die Vektoren w^k und w^i werden korrigiert. Wie?