MUSTERERKENNUNG
1. SEMINAR – NEURONALE MODELLE

Aufgabe 1. Gegeben sei eine Stichprobe $(x^1,\dots,x^l)$ 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. Man betrachte die quadratischen Entscheidungsregeln $\mathbb{R}^n \to \{1,2\}$:

$$f(x) = \begin{cases} 1 & \text{wenn} \quad x^T \cdot A \cdot x + \langle x,b\rangle + c < 0 \\ 2 & \text{sonst,} \end{cases}$$

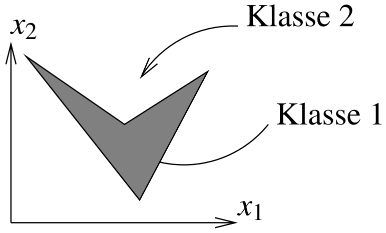
mit einer $n \times n$ Matrix A, einem Vektor $b \in \mathbb{R}^n$ und einer Konstante c. Zeigen Sie, dass man eine solche Entscheidungsregel mit Hilfe des Perzeptron Algorithmus anlernen kann.
Hinweis: Transformieren Sie dazu den Merkmalsraum in einen geeignet gewählten höherdimensionalen Raum, in dem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

Aufgabe 3. 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 $w^k \in \mathbb{R}^n$, $k = 1,\dots,|K|$. Mit Hilfe dieser wird der Raum $\mathbb{R}^n$ in $|K|$ konvexe Kegel $X^k \subset \mathbb{R}^n$ 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 $(x^1,k^1),\dots,(x^l,k^l)$. 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?

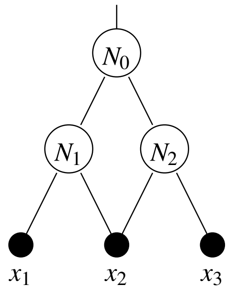
Aufgabe 4. Ein Feed-Forward Netz erhält als Inputs zwei reellwertige Signale x_1 und x_2. Dieses Netz soll alle diejenigen Muster der ersten Klasse zuordnen, die innerhalb des unten skizzierten Gebietes im $\mathbb{R}^2$ liegen. Alle anderen Muster sind der zweiten Klasse zuzuordnen.



Wieviel Schichten werden benötigt um einen solchen Klassifikator mit binären Schwellwertneuronen zu realisieren? Wie ist das Netz zu organisieren (d.h. wie sind die Neuronen in diesem Netz mit einander verbunden, welche Funktionen realisieren sie, welche Bedeutungen haben die Schichten des Netzes etc.)?

Aufgabe 5.

a) Für binäre Muster (x_1, x_2, x_3), $x_i \in \{0, 1\}$ soll ein Klassifikator realisiert werden, der die Muster $(0,0,0)$, $(0,0,1)$, $(0,1,1)$, $(1,1,1)$ der ersten Klasse – und alle anderen Muster der zweiten Klasse zuordnet. Dafür soll das abgebildete zweischichtige Feed-Forward Netz verwendet werden.

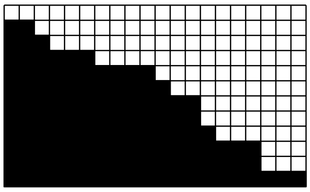


Wie muss man die Gewichte und Schwellen der 3 Neuronen wählen, damit das Netz den geforderten Klassifikator realisiert?

b) Verallgemeinern Sie die Aufgabe für den Fall, dass der Klassifikator für binäre Muster $(x_1, x_2, \ldots, x_n)$, $x_i \in \{0, 1\}$ realisiert werden soll. Zu der ersten Klasse gehören die Muster, für die $x_{i+1} \geq x_i$ für alle i gilt. Der zweiten Klasse werden alle anderen Muster zugeordnet.

Aufgabe 6. Betrachten Sie ein einfaches Hopfield-Netz (ohne externe Inputs) mit zwei binären Neuronen und symmetrischen Gewichten. Zeigen Sie, dass sich im Fall der parallelen Dynamik ein oszillierendes Verhalten einstellen kann.

Aufgabe 7. Man konstruiere ein Hopfield Netz, das die folgende Bildklasse modelliert. Die Bilder sind schwarzweiß. Die Grenze zwischen dem schwarzen bzw. dem weißen Gebiet (die Gebiete können jeweils leer sein) ist dabei eine monoton fallende Kurve. Ein Beispiel eines solchen Bildes ist in der Abbildung unten dargestellt.



Das Netz zur Modellierung diese Klasse soll folgende Eigenschaften haben. Die Neuronen sind zu je zwei in Kolumnen zusammengefasst und entsprechen den Pixelwerten schwarz/weiß. Gewichte verbinden Neuronen einer Kolumne als auch Neuronen benachbarter Kolumnen.

a) Wie ist das intrakolumnare Gewicht zu wählen, damit pro Kolumne höchstens ein Neuron aktiv ist?

b) Wie sind die Gewichte zwischen den Kolumnen und der Schwellwert zu wählen, damit folgendes gilt. Alle Bilder aus der oben beschriebenen Klasse entsprechen Netzkonfigurationen maximaler Energie. Alle anderen Bilder entsprechen Netzkonfigurationen mit kleineren Werten der Energie.

Aufgabe 8. Man betrachte ein Hopfield-Netz mit kolumnarer Struktur ohne externe Inputs. Das Netz ist so organisiert, dass pro Kolumne genau ein Neuron aktiv ist (siehe die vorige Aufgabe). Die Neuronen einer Kolumne seien nummeriert. Sei y eine Konfiguration des Netzes und $n(y)$ die Anzahl der „Cracks“ – die Anzahl der Paare benachbarter Kolumnen, in denen die aktiven Neuronen unterschiedliche Nummern haben. Konstruieren Sie das Netz derart, dass seine Energie proportional zur negativen Anzahl der Cracks ist, d.h. $E(y) \sim -n(y)$.