Prof. Dr. Claus Weihs, Prof. Dr. Katharina Morik
Julia Schiffner, Marco Stolpe

Dortmund, 9. Juni
Abgabe: bis zum 16.06. um 10:00 Uhr an
marco.stolpe@tu-dortmund.de

Übungen zur Vorlesung
Wissensentdeckung in Datenbanken
Sommersemester 2009
Blatt 8

Aufgabe 8.1 (4 Punkte)

In dieser Aufgabe geht es darum, den Effekt einer Φ-Transformation auf Datenpunkten zu untersuchen. Gegeben sind die Datenpunkte aus der rechts stehenden Tabelle. Transformieren Sie die Daten mit den nachfolgenden Funktionen Φ_i und geben Sie die transformierten Werte in einer Tabelle an. Welche Funktion ermöglicht eine lineare Trennung der Daten?

x_1	x_2	y
-1.5	-2.0	+1
-1.0	0.0	+1
-0.5	1.0	+1
0.0	2.0	+1
0.5	1.0	+1
1.0	2.0	+1
1.5	3.75	+1
-1.0	-2.0	-1
-0.5	-1.0	-1
0.0	-3.0	-1
0.5	-0.5	-1
1.0	-2.0	-1
1.5	1.5	-1

(a) $\Phi_1(x_1, x_2) = (x_1^2, x_2)$

(b) $\Phi_2(x_1, x_2) = (x_1^3 - 2x_1, x_2)$

(c) $\Phi_3(x_1, x_2) = (x_1^3, x_2)$

Hinweis: Am besten erstellen Sie eine graphische Darstellung der Punkte. Diese ist bloß als Hilfe für Sie gedacht und braucht nicht mit abgegeben zu werden.

Aufgabe 8.2 (3 Punkte)

In dieser Aufgabe sollen Sie für $\vec{x} = (x_1, \ldots, x_p)$ zwei Funktionen $f_1(\vec{x})$ und $f_2(\vec{x})$ unter Nebenbedingungen $h_i(\vec{x})$ optimieren. Stellen Sie dazu, wie auf den Folien, die entsprechende Lagrange-Funktion $L(\vec{x}, \vec{\mu})$ auf. Die Lösung ergibt sich aus dem Gleichungssystem, das durch Nullsetzen der partiellen Ableitungen von L nach $\vec{x}$ und $\vec{\mu}$ entsteht.

(a) Maximiere $f_1(\vec{x}) = 1 - x_1^2 - x_2^2$ unter der Nebenbedingung $h(\vec{x}) = x_1 + x_2 - 1 = 0$.

(b) Maximiere $f_2(\vec{x}) = -x_1 - x_2 - x_3$ unter den Nebenbedingungen $h_1(\vec{x}) = x_1^2 + x_2 - 3 = 0$ und $h_2(\vec{x}) = x_1 + 3x_2 + 2x_3 - 7 = 0$.

Hinweis: Beachten Sie bitte, dass für die bei Gleichheitsbedingungen $h_i(\vec{x}) = 0$ eingeführten Multiplikatoren μ_i nur $\mu_i \neq 0$ (*nicht* $\mu_i \geq 0$!) gelten muss.

Aufgabe 8.3 (3 Punkte)
Sie haben die strukturelle Risikominimerung kennengelernt, die dazu dient, die Ausdrucksstärke von Modellklassen zu messen. Sie basiert auf der VC-Dimension.

(a) Das empirische Risiko zweier jeweils bester Modelle aus unterschiedlichen Modellklassen sei identisch. Die erste Modellklasse besitze endliche, die zweite Modellklasse unendliche VC-Dimension. Welches Modell empfiehlt sich hinsichtlich der strukturellen Risikominimierung und warum?

(b) Sei $\mathcal{APR}_2$ die Klasse der achsenparallelen Rechtecke in der Ebene (also im $\mathbb{R}^2$). Beweisen Sie (z. B. graphisch), dass diese Hypothesenklasse eine VC-Dimension von 4 besitzt. Dabei klassifiziert jedes Modell genau die Punkte innerhalb des Rechtecks als positiv.

(c) Die Klasse $\mathcal{K}_2$ der Kreisscheiben in der Ebene mit Parametern z (Mittelpunkt) und r (Radius) ist wie folgt definiert:

$$\mathcal{K}_2 := \{K(z,r) \mid K(z,r) = \{x \mid d(x,z) \leq r\}, z \in \mathbb{R}^2, r \in \mathbb{R}, r \geq 0\}\,.$$

Jedes Modell klassifiziert genau die Punkte innerhalb des Kreises als positiv. Beweisen Sie (z. B. graphisch), dass diese Hypothesenklasse eine VC-Dimension von 3 besitzt.