Ludwig-Maximilians-Universität München
Institut für Informatik
PD Dr. Peer Kröger, Dr. Arthur Zimek

Knowledge Discovery in Databases
WS 2010/11

Übungsblatt 7: Kernels und Support-Vektor-Machines

Aufgabe 7-1 *Support Vector Machines*

Angenommen, eine Support Vector Machine minimiert beim Lernen der Entscheidungsfunktion lediglich die Zahl der falsch klassifizierten Trainingsobjekte. Welches Problem kann sich daraus potentiell ergeben? Wie lässt sich dieses Problem beheben?

Aufgabe 7-2 *Support Vector Machines*

Betrachten wir für Support Vector Machines zur Klassifikation den minimalen Fall, in dem für jede Klasse nur ein Vektor gegeben ist. Dies impliziert zwangsläufig, dass die Klassen aufgrund der Trainingsmenge linear separierbar sind, d.h. wir können mit einem Hard Margin im Eingaberaum arbeiten. Als Trainingsmenge sollen uns die beiden Vektoren $(1, 1)$ für Klasse A ($y = -1$) und $(2, 2)$ für Klasse B ($y = 1$) dienen. Das verwendete Skalarprodukt sei das kanonische Skalarprodukt (vgl. Bsp. in der Vorlesung).

(a) Formulieren Sie das Problem zunächst als duales Optimierungsproblem mit Lagrange Multiplikatoren. Bestimmen Sie durch analytische Lösung die Werte der Lagrange Multiplikatoren.

(b) Zur Berechnung der Maximum Margin Hyperplane bestimmen wir den Normalenvektor $\vec{w}$. Das Inverse des Betrags dieses Vektors $\vec{w}$ stellt dabei die Breite des Randes (Margin) dar.
Benutzen Sie zur Bestimmung folgende Formel:

$$\vec{w} = \sum_{i=1}^{n} \alpha_i \cdot y_i \cdot \vec{x}_i$$

(c) Nachdem Sie den Normalenvektor $\vec{w}$ bestimmt haben, können wir daraus noch den Skalar b berechnen, um die Lage der Hyperebene $H(\vec{w}, b)$ endgültig festzulegen. Die Formel zur Berechnung ist:

$$b = -\frac{\max_{i,y_i=-1} \langle \vec{w}, \vec{x}_i \rangle + \min_{i,y_i=1} \langle \vec{w}, \vec{x}_i \rangle}{2}$$

(d) Nachdem Sie damit die trennende Hyperebene festgelegt haben, bestimmen Sie jetzt die Klasse der beiden Vektoren: $(3, 5)$ und $(0, 1)$. Benutzen Sie dazu entweder die Entscheidungsregel des primären OPs oder die des dualen OPs.

Aufgabe 7-3 *Kernel-Funktionen*

Eine Kernel-Funktion ("Kernel") zeichnet sich durch positive (Semi-)Definitheit aus. Eine Matrix A ist positiv definit, falls ihre Eigenwerte nichtnegativ sind, oder alternativ formuliert, falls für alle $x \in \mathbb{R}^d$ gilt: $x^\top \cdot A \cdot x \geq 0$

Zeigen Sie, dass folgende Funktionen Kernels sind, falls x und $\hat{x}$ Vektoren im $\mathbb{R}^d$ sind:

(a) $k_1(x, \hat{x}) = 1$

(b) $k_2(x, \hat{x}) = 3 \cdot x^\top \cdot \hat{x}$

(c) $k_3(x, \hat{x}) = 3 \cdot x^\top \cdot \hat{x} + 5$