

Übungen zur Vorlesung
Wissensentdeckung in Datenbanken
Sommersemester 2008

Blatt 9

Wiederholung Die *Support Vector Machine (SVM)* oder auch *Stützvektormethode* ist ein komplexes Lernverfahren, das nahe an der Linearen Regression liegt. Im Folgenden sollten Sie sich einige Gedanken zum Prinzip der Klassifikation mit dem SVM-Verfahren machen:

- (a) Geben Sie einen kurzen Überblick über das *Funktionenlernen*, den zugehörigen Hypothesenraum und die Entsprechung im SVM-Verfahren!
- (b) Wozu wird die SVM verwendet und was ist das Ergebnis des Lernschrittes?
- (c) Erklären Sie den Zusammenhang der separierenden Hyperebene mit dem *maximum margin problem*!
- (d) Wie wird das *maximum margin problem* gelöst? Was ist das *duale Problem*?
- (e) Was drückt die Bedingung

$$\sum_{i=1}^n \alpha_i y_i = 0$$

in Bezug auf die Berücksichtigung der Stützvektoren aus?

Hinweis: Von den in der Vorlesung vorgestellten Arbeiten finden Sie einige wieder direkt auf der Web-Seite mit den Übungsblättern, das Letztgenannte ist in der Uni-Bibliothek verfügbar. Falls Interesse an dem zweiten oder dritten Papier besteht, bitte bei mir melden.

- “*A Tutorial on Support Vector Machines for Pattern Recognition*”, Christopher Burges in: “Data Mining and Knowledge Discovery”, 1998, pp. 121-167
- “*Fast Training of Support Vector Machines using Sequential Minimal Optimization*”, John C. Platt in: “Advances in Kernel-Methods”, 1999, MIT Press, pp. 185-208
- “*Advances in Kernel-Methods*”, Schölkopf, Burges, Smola (eds), 1999, MIT Press
- “*Theorie der Zeichenerkennung*”, W.N. Vapnik, A. Tscherwonkis, 1979, Akademie Vg. [UB-Dortmund: Signatur M20947]

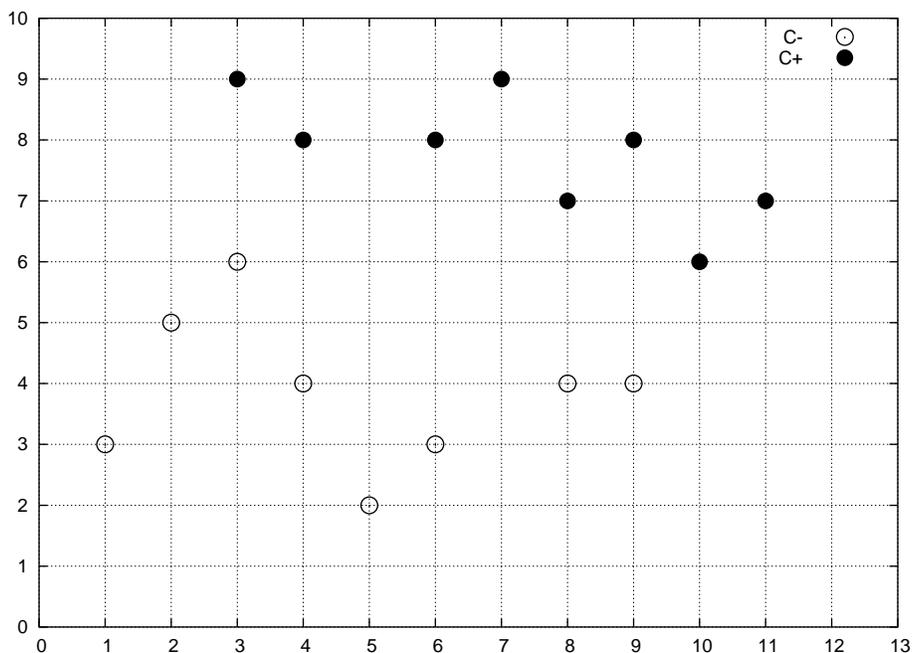


Abbildung 1: Datenpunkte im \mathbb{R}^2

Aufgabe 9.1

In dieser Aufgabe sollen Sie den naiven Vorläufer des SVM-Algorithmus “von Hand” ausprobieren. Gegeben ist die folgende Menge D von Datenpunkten (vgl. Abbildung 1):

$$\begin{aligned}
 D = \{ & (1, 3, -1), (2, 5, -1), (3, 6, -1), (4, 4, -1), \\
 & (5, 2, -1), (6, 3, -1), (8, 4, -1), (9, 4, -1), \\
 & (3, 9, +1), (4, 8, +1), (6, 8, +1), (7, 9, +1), \\
 & (8, 7, +1), (9, 8, +1), (10, 6, +1), (11, 7, +1) \}
 \end{aligned}$$

Dabei bezeichnen die Komponenten jedes Tripels (x_1, x_2, y) aus D die erste und zweite Koordinate des Punktes sowie die zugehörige Klasse $y \in \{-1, 1\}$.

- (a) Bestimmen Sie für $j \in \{-1, +1\}$ jeweils die Mittelpunkte \vec{c}_+ und \vec{c}_- der Menge

$$C_j = \{(x_1, x_2, y) \in D \mid y = j\}$$

- (b) Bestimmen Sie \vec{w} und den Mittelpunkt \vec{c} .
- (c) Zu welchen Klassen werden anhand dieses einfachen Verfahrens die folgenden Punkte zugeordnet?

$$(4, 6), (7, 6), (12, 4), (-1, 8), (-4, 11)$$

Aufgabe 9.2

Als nächstes sollen Sie sich etwas mehr mit der Stützvektormethode anfreunden. Betrachten Sie dazu erneut die Datenpunkte aus Aufgabe 1.

- (a) Wählen Sie sowohl aus C_{-1} als auch aus C_{+1} je zwei geeignete Stützvektoren und stellen Sie die Geradengleichungen durch diese Vektoren auf.
- (b) Ermitteln Sie die optimale separierende Hyperebene (Gerade) zwischen den gewählten Stützvektoren und geben Sie diese in Form einer normalen Geradengleichung an.

Aufgabe 9.3

In dieser Aufgabe geht es darum, den Effekt der Φ -Transformation der 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 transformierte Tabelle sowie eine graphische Darstellung der neuen Punkte an.

Welche Funktion ermöglicht eine lineare Trennung der Daten?

- (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: Die graphische Darstellung ist als Hilfe für Sie gedacht und braucht nicht mit abgegeben zu werden. Sie finden auf der letzten Seite ein Koordinatensystem, welches Sie für ihre Notizen benutzen können.

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

