

Übungen zur Vorlesung  
**Wissensentdeckung in Datenbanken**  
Sommersemester 2009

Blatt 9

**Aufgabe 9.1 (3 Punkte)**

In der Vorlesung wurde zur Optimierung der SVM-Parameter das Verfahren SMO vorgestellt.

- (a) Warum optimiert man hierbei nicht alle Parameter gleichzeitig, und wie viele Parameter werden statt dessen gleichzeitig optimiert und warum?
- (b) Gehen Sie davon aus, dass  $C = 0, 1$  ist. Welchen Wert kann  $\alpha_2$  mindestens und welchen Wert maximal annehmen? Lösen Sie diese Aufgabe zeichnerisch und leiten Sie Ihre Lösung daraus ab. Gehen Sie einmal von  $y_1 = y_2$  und dann von  $y_1 \neq y_2$  aus.

**Aufgabe 9.2 (3 Punkte)**

Die SVM kann auf Beispielen trainiert werden, die in zwei Klassen fallen. Liegen mehrere Klassen vor, so müssen in der Regel mehrere SVMs trainiert werden.

- (a) Überlegen Sie sich mindestens eine (nicht unbedingt perfekte) Methode, wie sich über Training und Anwendung mehrerer SVMs Klassifikationsprobleme mit mehreren Klassen lösen lassen.
- (b) Nennen Sie Probleme/Nachteile Ihrer Methode.

**Aufgabe 9.3 (2 Punkte)**

Erläutern Sie, warum es bei der Optimierung des  $\gamma$ -Parameters einer SVM mit RBF-Kernel zu Overfitting kommen kann! Welche Eigenschaft der Lösung ist ein Hinweis darauf?

**Aufgabe 9.4 (2 Punkte)**

Die Regressions-SVM wurde für Zeitreihen vorgestellt. Bei Zeitreihen kann es zu Trends oder Zyklen kommen, die zu beachten sind. Geben Sie zusätzlich zu den in der Vorlesung und auf den Folien vorgestellten Zeitreihen drei Beispiele für Datenreihen, die Zyklen oder Trends enthalten.