

Maschinelles Lernen

Übungsblatt 9

Dr. Steffen Rendle, Christoph Freudenthaler
Wirtschaftsinformatik und Maschinelles Lernen (ISMLL)
Universität Hildesheim

19. Januar 2010
Abgabe bis 28. Januar 2010

Support Vektor Maschinen I - Lineare Separierung

Nr.	Zielklasse	X_1	X_2
D1	+	-0.8	0.1
D2	+	1.2	-0.1
D3	+	0.3	-0.7
D4	+	0	1
D5	-	1.5	2
D6	-	-1.3	1.7
D7	-	-1.2	-1.9
D8	-	1.5	-1.5

- a) Handelt es sich bei den vorliegenden Daten um ein linear separables Problem?
- b) Stellen Sie die vorliegenden Daten in einem Streudiagramm dar und berechnen Sie für diese eine Kernel-SVM ohne Slackvariablen, dafür aber mit einem polynomiellen Kernel der Ordnung 4. Zeichnen Sie die Trenngerade in das Streudiagramm ein und berechnen Sie den Margin. Verwenden Sie einen geeigneten Startwert für ihre Modellparameter.

Support Vektor Maschinen II - Kernels

- a) Betrachten Sie die in der Vorlesung eingeführten Kernel-Funktionen des linearen, des polynomiellen Kernels der Ordnung 2 und einer Kernelfunktion mit radialer Basisfunktion (=RBF-Kernel, Gauß-Kernel). Erzeugen Sie in der 2-dimensionalen euklidischen Ebene Konturdiagramme für den linearen und den RBF-Kernel. Verwenden Sie als Referenzpunkt den Ursprung (1, 1). Was bedeutet es hinsichtlich der Modellkomplexität, wenn man den jeweiligen Kernel verwendet?
- b) Berechnen Sie die Darstellung des ursprünglichen Prediktorvektors $\mathbf{x} \in \mathbb{R}^p$ mit p erklärenden Variablen in den jeweiligen höherdimensionalen Kernelräumen, die durch die drei in a) angegebenen Kernel-funktionen induziert werden.

Hinweis: Jede Kernelfunktion kann als euklidisches Skalarprodukt zweier (höher-)dimensionaler Vektoren $\Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \in \mathbb{R}^q$ mit $q \geq p$ dargestellt werden, wobei $\Phi(\mathbf{x}_i)$ durch die Wahl der Kernelfunktion festgelegt wird. Somit gilt: $\langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle = K(\mathbf{x}_i, \mathbf{x}_j)$. Benutzen Sie zur Bestimmung der RBF-Kernel induzierten Darstellung $\Phi(\mathbf{x}_i)$ die Taylor-Polynom-Representation des Kernels.

- c) Welche Bedeutung hat der Hyperparameter γ bei Support-Vektor-Maschinen mit Slackvariablen? Was ist ein Hyperparameter? Gibt es noch andere Hyperparameter bei Support-Vektor-Maschinen?