Übung zur Vorlesung

Maschinelles Lernen 1

Wintersemester 2009/2010

Abteilung Maschinelles Lernen
Institut für Softwaretechnik und
theoretische Informatik
Fakultät IV, Technische Universität Berlin
Prof. Dr. Klaus-Robert Müller
Email: klaus-robert.mueller@tu-berlin.de

Blatt 12

Abgabe bis Abgabe bis **Mittwoch, den 20. Januar 2010, 14 Uhr** an Mikio Braun (FR6058, notfalls unter der Türe durchschieben).

Aufgaben

Auf diesem Aufgabenzettel wird eine Kernversion der linearen Diskriminanzanalyse (LDA) hergeleitet. Zur Erinnerung:

Wir haben eine Datensatz X_i, Y_i, $1 \leq i \leq n$, mit $Y_i = \pm 1$. Sei D_k die Menge aller Punkte X_i so dass $Y_i = k$, und $m_k = \sum_{x \in D_k} x / |D_k|$ der Mittelpunkt der Klasse k.

Die FDA-Kostenfunktion lautet

$$J(w) = \frac{w^\top S_B w}{w^\top S_W w},$$

wobei

$$S_W = S_1 + S_2 \qquad \text{mit} \qquad S_i = \sum_{x \in D_i} (x - m_i)(x - m_i)^\top, \qquad \text{(der “within scatter”)}$$
$$S_B = (m_1 - m_2)(m_1 - m_2)^\top \qquad \text{(der “between scatter”).}$$

Für die Kernversion wird nun eine Kernfunktion $k(x, y)$ eingeführt, die eine Merkmalstransformation $\phi : \mathbb{R}^d \to \mathcal{F}$ induziert, so dass $k(x, y) = \langle \phi(x), \phi(y) \rangle_{\mathcal{F}}$ gilt.

In den Aufgaben soll die Kostenfunktion für die Kernversion hergeleitet werden, d.h. eine Formel für $J(w)$, in der nur noch $k(x, y)$, aber keine ϕ mehr vorkommen, da diese nicht ausgerechnet werden können. Da w dann in $\mathcal{F}$ liegt, welcher potentiell unendliche Dimension hat, wird w zunächst durch die Datenpunkte dargestellt:

$$w = \sum_{i=1}^{n} \alpha_i \phi(X_i).$$

Grob gesagt, erhält man die Lösung für die Kernversion, indem man für w diese Formel einsetzt und jedes Vorkommen eines x mit $\phi(x)$ ersetzt, und dann solange umformt, bis nur noch Kernauswertungen $k(x, y)$ vorkommen. In dem Endresultat dürfen nur Objekte mit endlicher Dimensionalität vorkommen!

Gehe dabei wie folgt vor:

1. Zeige zunächst, dass

$$w^\top m_i = \alpha^\top q_i, \qquad \text{mit} \qquad q_i = \left[\frac{1}{|D_i|} \sum_{x \in D_i} k(x, x_j) \right]_{j=1}^{n}.$$

   (10 Punkte)

2. Verwende das Ergebnis aus Aufgabe 1 und leite eine Formulierung für $w^\top S_B w$ her, in der keine Auswertungen von ϕ mehr einzeln vorkommen. **(10 Punkte)**

3. Verwende das Ergebnis aus Aufgabe 2 und leite eine Formulierung für $w^\top S_W w$ her, in der keine Auswertungen von ϕ mehr einzeln vorkommen. **(10 Punkte)**

Für Fragen zum Übungsblatte bitte in der Google Group http://groups.google.com/group/mikiobraun-lehre registrieren und die Frage an die Mailingliste stellen.