

Maschinelles Lernen und Data Mining
Sommersemester 2013
Übungsblatt 6

Besprechung des Übungsblattes am 25.06.2013

Aufgabe 6-1 Generatives Modell

- a) Wenn $P(c = j)$ und $P(\mathbf{x}|c = j)$ bekannt sind, lässt sich der optimale Klassifikator nach der Bayes'schen Regel berechnen. Geben Sie ihn als Entscheidungsfunktion an. b.w.
- b) Nun sei $P(\mathbf{x}|c = j)$ für alle j mit identischer Kovarianz aber unterschiedlichen Zentren normalverteilt. Formulieren Sie das Problem aus.

Hinweis: Die mehrdimensionale Normalverteilung ist definiert als

$$\mathcal{N}(\mathbf{x}_i|c = j) = \frac{1}{(2\pi)^{\frac{p}{2}} |\Sigma|^{\frac{1}{2}}} e^{(-\frac{1}{2}(\mathbf{x}_i - \mu_j)^T \Sigma^{-1} (\mathbf{x}_i - \mu_j))},$$

wobei p die Dimensionalität der Daten, μ_j der Mittelwertsvektor zu Klasse j , Σ die Kovarianzmatrix über alle Klassen, und $|\Sigma|$ die Determinante von Σ ist.

- c) Dieses Problem lässt sich optimieren zu einem Schätzer auf den μ_j und Σ :

$$\hat{\mu}_j = \frac{1}{N_j} \sum_{i:y_i=j} \mathbf{x}_i$$
$$\hat{\Sigma} = \frac{1}{N - M} \sum_{j=1}^C \sum_{i:y_i=j} (\mathbf{x}_i - \hat{\mu}_j)(\mathbf{x}_i - \hat{\mu}_j)^T,$$

wobei N die Größe des Trainingsdatensatzes ist, M die Dimension der Daten und N_j die Anzahl der Trainingsdaten, die zu Klasse j gehören. Die Grund-Wahrscheinlichkeiten für die Klassen, $P(c = j)$, lassen sich anhand ihrer Häufigkeit im Datensatz abschätzen.

Trainieren Sie das in b) entwickelte Modell auf den Daten aus `bayesianData.txt` und bestimmen Sie anschließend den Trainingsfehler. (Die Datei besteht aus 2 Tab-separierten Variablenpalten mit vorangehender Klassenlabelspalte.)

Aufgabe 6-2 Bestimmung der optimal-trennenden Hyperebene

Bestimmen Sie die optimale, separierende Hyperebene des zwei-Klassen-Datensatzes (A, B) mit:

$$A = \left\{ p_1 = \begin{pmatrix} 2 \\ 4 \end{pmatrix}, p_2 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}, p_3 = \begin{pmatrix} 1 \\ 0.5 \end{pmatrix}, p_4 = \begin{pmatrix} 2.5 \\ 3 \end{pmatrix}, p_5 = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right\},$$
$$B = \left\{ p_6 = \begin{pmatrix} 0.5 \\ 1.5 \end{pmatrix}, p_7 = \begin{pmatrix} -1 \\ 3 \end{pmatrix}, p_8 = \begin{pmatrix} 0 \\ 0.5 \end{pmatrix} \right\}$$

Die Klassenlabels y der Klasse A sind deklariert mit 1, die von B mit -1 .

Visualisieren Sie das Ergebnis und geben Sie die Supportvektoren an. Wie breit ist die Margin?