Übungen Musterklassifikation
Sommersemester 2007

Abgabe am 28.6.07

Prof. Dr. Stefan Posch

Dr. Birgit Möller

(birgit.moeller@informatik.uni-halle.de)

Institut für Informatik
Universität Halle

Blatt 12

Aufgabe 12.1 (10 Punkte)

Gegeben sei ein 2-Klassen-Problem. Alle Mustervektoren einer Klasse wurden durch eine 2D-Normalverteilung erzeugt. Der klassifizierte Gesamtdatensatz ist auf der Website zur Übung zu finden (`2_class_mixture.dat`).

(a) Zerlegen Sie den Datensatz in einen *train*- und in einen *test*-Datensatz. Überlegen Sie sich, welche verhältnismässige Aufteilung günstig ist und begründen Sie dies.

(b) Schätzen Sie aus dem *train*-Datensatz die Parameter der 2D-Normalverteilungen für jede Klasse. Geben Sie die geschätzten Parameter an.

(c) Bauen Sie einen Likelihood-Klassifikator, mit dem Sie die Mustervektoren des *test*-Datensatzes klassifizieren können. Der Likelihood-Klassifikator entscheidet auf Klasse 1, wenn $\frac{P(\vec{c}\,|\,w_1)}{P(\vec{c}\,|\,w_2)} > \theta$ mit $\theta = 1$ gilt. Bestimmen Sie aus den Klassifikationsergebnissen die erzielte Sensitivität für Klasse 1 (Sn_1) und 2 (Sn_2)!

(d) Für welchen Schwellwert θ erhalten Sie einen Bayes-Klassifiaktor?

(e) Nun soll in "kleinen Schritten" vom Likelihood- zum Bayes-Klassifikator übergegangen werden, indem θ beginnend bei 1 entsprechend Ihren Überlegungen aus der vorangegangenen Teilaufgabe so verändert wird, dass am Schluß bei entsprechendem θ der Klassifikator, der nach $\frac{P(\vec{c}\,|\,w_1)}{P(\vec{c}\,|\,w_2)} > \theta$ auf Klasse 1 entscheidet, sich wie ein Bayes-Klassifikator verhält (entscheidet). Bestimmen Sie für jeden Schritt wiederum die beiden klassenspezifischen Sensitivitäten!

(f) Plotten Sie die erhaltenen Sensitivitäten Sn_1 über $1 - Sn_2$ und interpretieren Sie.

Aufgabe 12.2 (4 Punkte)

Beweisen Sie für $R[f] = E\{(y - f(\vec{x}))^2\}$:

$$\begin{aligned} R[f] &= E\{(f_0(\vec{x}) - f(\vec{x}))^2\} + E\{(y - f_o(\vec{x}))^2\} \\ \text{mit}\quad f_o(\vec{x}) &:= E\{y|\vec{x}\} = E_{P(y|\vec{x})}\{y\} = \int_y yP(y|\vec{x})dy \end{aligned}$$

Hinweis: $\int (a-b)^2 = \int (a - c + c - b)^2$