

# Maschinelles Lernen und Data Mining

## 8. Übung

Prof. Tobias Scheffer  
Steffen Bickel

WS06/07

Ausgabe am: 22.12.06  
Besprechung am: 12.01.07

### Aufgabe 1 (4/8 Punkt):

Sie haben zwei verschiedene Klassifikatoren trainiert und möchten deren Qualität mit Hilfe eines Testdatensatzes bewerten. Die gelernten Entscheidungsfunktionen sind  $f_1(x)$  und  $f_2(x)$ . Der Schwellwert ist  $\theta = 0$ , dies bedeutet, wenn  $f(x) \geq 0$  dann entscheidet sich der Klassifikator für die Klasse +, sonst für -. Die Testmenge besteht aus 10 Beispielen. Die resultierenden Werte der Entscheidungsfunktionen  $f_1(x_i)$  und  $f_2(x_i)$ ,  $i = 1, \dots, 10$ , können Sie zusammen mit der wahren Klasse  $y_i$  der Tabelle entnehmen.

| i          | 1   | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    | 10   |
|------------|-----|------|------|------|------|------|------|------|------|------|
| $y_i$      | +1  | -1   | -1   | +1   | +1   | -1   | -1   | -1   | +1   | +1   |
| $f_1(x_i)$ | 0.4 | -1.3 | -2.0 | -1.1 | 0.9  | -0.3 | -1.9 | 3.1  | -2.1 | -1.4 |
| $f_2(x_i)$ | 1.7 | 2.1  | -0.1 | 0.8  | -2.1 | 2.2  | -0.6 | -0.5 | 1.5  | 3.1  |

- Bestimmen Sie Precision, Recall und F-Measure für beide Klassifikatoren.
- Zeichnen sie die Precision-Recall-Kurven für  $f_1$  und  $f_2$  und bestimmen Sie die Precision-Recall-Breakeven-Points.
- Zeichnen Sie die ROC-Kurven für  $f_1$  und  $f_2$  und bestimmen Sie die jeweiligen AUC-Werte.
- Für welchen Klassifikator würden Sie sich entscheiden und warum?

### Aufgabe 2 (3/8 Punkt):

In der folgenden Tabelle sind sechs Trainingsbeispiele für einen Klassifikator abgebildet.  $y_i$  ist das Klassenlabel,  $x_{1i}$  und  $x_{2i}$  sind die Attribute.

| i        | 1   | 2   | 3   | 4   | 5   | 6   |
|----------|-----|-----|-----|-----|-----|-----|
| $y_i$    | +1  | +1  | +1  | -1  | -1  | -1  |
| $x_{1i}$ | 1.0 | 1.3 | 2.5 | 4.2 | 5.1 | 3.7 |
| $x_{2i}$ | 3.4 | 4.2 | 3.7 | 1.7 | 2.5 | 3.1 |

- Simulieren sie das Training eines Rocchio-Klassifikator von Hand und klassifizieren sie das Testbeispiel (3.5; 3.5). Verwenden sie  $\text{sign}(f(x))$  als Klassifikationsfunktion.
- Simulieren sie das Training eines Perzeptrons  $f(x) = \mathbf{w}\mathbf{x} + b$  von Hand. Setzen sie  $b = 0$  und ignorieren sie die Update-Schritte für  $b$ , damit sie schneller zu einem Ergebnis kommen. Laufen sie in jeder Iteration des Algorithmus von links nach rechts über die Tabelle. Klassifizieren sie wieder das Testbeispiel von Aufgabe (a). Tipp: Um Zeit zu sparen, brauchen sie den Entscheidungsfunktionswert nicht für jedes Beispiel explizit ausrechnen, oft sieht man schon an den Vorzeichen und Größenordnungen der Gewichtsvektoreinträge, dass ein Beispiel richtig klassifiziert wird und kein Update nötig ist.
- Nehmen sie an, sie hätten ein zusätzliches positives Trainingsbeispiel (5.0; 1.0). Sie möchten nun wieder ein Perzeptron trainieren. Stellen sie die Trainingsdaten grafisch in einem zweidimensionalen Diagramm dar und überlegen sie sich ohne ausprobieren, welcher Klasse das obige Testbeispiel nun zugeordnet werden würde.

### Aufgabe 3 (1/8 Punkt):

In der Vorlesung haben wir gelernt, dass das Minimum des primalen SVM-Optimierungskriteriums auch als MAP-Hypothese interpretiert werden kann. Überlegen sie sich, welcher Gewichtsvektor  $w$  der ML-Hypothese entspricht, wenn die Trainingsdaten linear separierbar sind. Kennen sie einen Algorithmus der die ML-Hypothese liefert für den linear separierbaren Fall? Was ist die ML-Hypothese, wenn die Trainingsdaten nicht-linear separierbar sind?