

Maschinelles Lernen

8. Übung

Prof. Tobias Scheffer
Dr. Niels Landwehr
Christoph Sawade
Jules Rasetaharison

WS10/11

Ausgabe am: 03.01.11
Besprechung am: 10.01.11

Aufgabe 1 (1/4 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

Wir betrachten den linearen Klassifikator $\text{sign}(f(\mathbf{x}))$ mit der Trennebene $f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + w_0$.

- Wie groß ist der euklidische Abstand eines Punktes \mathbf{z} von der Trennebene $f(\mathbf{x})$?
- Angenommen die Trennebene verläuft durch den Koordinatenursprung, welchen Einfluss hat dann die Länge des Gewichtsvektors \mathbf{w} auf das vorhergesagte Klassenlabel?
- Angenommen wir haben n Beispiele \mathbf{x}_i mit dazugehörigen Labels $y_i \in \{-1, +1\}$ welche linear separierbar sind, d.h. es gibt (mindestens) eine Trennebene f mit $\text{sign}(f(\mathbf{x}_i)) = y_i$ für $i = 1, \dots, n$. Als geometrischen *margin* bezeichnet man den Abstand zwischen einer solchen Trennebene f und dem Datenpunkt welcher am nächsten an der Trennebene liegt. Wie lässt sich dieser Abstand für eine gegebene Trennebene berechnen?
- Stellen sie die in der Tabelle gegebenen Datenpunkte in einem Diagramm dar und zeichnen sie eine beliebige Trennebene ein. Markieren sie in dem Diagramm den geometrischen margin und zeichnen sie \mathbf{w} und w_0 mit $\|\mathbf{w}\| = 1$ ein.

Aufgabe 2 (1/4 Punkt):

Wir betrachten erneut die Trainingsbeispiele aus Aufgabe 1.

- Simulieren sie das Training eines Rocchio-Klassifikators von Hand.
- Berechnen sie den Gewichtsvektor der Fisher-Diskriminate unter Verwendung der Formeln auf Seite 11 der Folien zu den Linearen Klassifikatoren. Bestimmen sie dazu zunächst die 2×2 -Matrix \mathbf{S}_W und deren Inverse. Tipp:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

- Klassifizieren sie das Testbeispiel $(3, 5; 3, 5)$ unter Verwendung der Klassifikationsfunktion $\text{sign}(f(\mathbf{x}))$ und den Gewichtsvektoren aus Aufgabe 2.a und 2.b. (Wir definieren $\text{sign}(a) = +1$ für $a \geq 0$ und $\text{sign}(a) = -1$ für $a < 0$.)
- Zusatzaufgabe (+1 Punkt): Unter welchen Voraussetzungen liefern Rocchio und die Fisher-Diskriminate gleiche Gewichtsvektoren (bzw. wann haben beide Gewichtsvektoren die gleiche Richtung)?

Aufgabe 3 (1/4 Punkt):

Wir betrachten erneut die Trainingsbeispiele aus Aufgabe 1.

Betrachten Sie den K-nearest neighbour (KNN) Ansatz. Welche Klassen werden vorhergesagt, wenn sie das Testbeispiel $(3, 5; 3, 5)$ für $K = 1, 2, \dots, 6$ klassifizieren? Benutzen sie dazu die euklidische Distanz $d(x,y)$:

$$d(x, y) = \|x - y\|_2 = \sqrt{\sum_{i=0}^n (x_i - y_i)^2}.$$

Aufgabe 4 (1/4 Punkt):

Betrachten wir die Trainingsbeispiele aus Aufgabe 1: Der Perzeptron-Algorithmus (ohne expliziten Bias b) ist wie folgt gegeben:

Algorithm 1 Perzeptron

Require: n Beispiele (\mathbf{x}_i, y_i)

```

1:  $\mathbf{w} := \mathbf{0}$ 
2: repeat
3:    $updates := 0$ 
4:   for  $i = 1$  to  $n$  do
5:     if  $y_i \mathbf{w}^\top \mathbf{x}_i \leq 0$  then
6:        $\mathbf{w} := \mathbf{w} + y_i \mathbf{x}_i$ 
7:        $updates := updates + 1$ 
8:     end if
9:   end for
10: until  $updates = 0$ 

```

- (a) Zeigen sie, dass ein Update $\mathbf{w}' := \mathbf{w} + y_i \mathbf{x}_i$ tatsächlich den Zielfunktionswert $\mathbf{w}'^\top \mathbf{x}_i$ in die gewünschte Richtung verändert, d.h. $y_i \mathbf{w}'^\top \mathbf{x}_i > y_i \mathbf{w}^\top \mathbf{x}_i$.
- (b) Simulieren sie das Training eines Perzeptrons $f(x) = \mathbf{w}^\top \mathbf{x}$ von Hand. Laufen sie in jeder Iteration des Algorithmus von links nach rechts über die Tabelle. Klassifizieren sie Testbeispiel $[3, 5; 3, 5]^\top$ unter Verwendung der Klassifikationsfunktion $sign(f(x))$. Tipp: Um Zeit zu sparen, brauchen sie den Zielfunktionswert 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.
- (c) Nehmen sie an, sie hätten ein zusätzliches positives Trainingsbeispiel $[5, 0; 1, 0]^\top$. 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.