

Übung Künstliche Neuronale Netzwerke

Wintersemester 2011/2012

- ① Abgabe am 4. 11. 2011 bis Ende der Vorlesung handschriftlich auf Papier (1x pro Person)
② Abgabe bis 11. 11. 2011, 23:59 Uhr per E-Mail an markus.lessmann@ini.rub.de (1x pro Gruppe)

1. Übungsblatt vom 28. 10. 2011

1.1 Polynomregression

Eine Messreihe ist gegeben durch eine Folge von Eingabewerten x_1, \dots, x_P und Sollwerten t_1, \dots, t_P (Trainingsbeispiele). Diese soll durch Polynome vom Grad M angenähert werden.

$$y(\underline{w}, x) = \sum_{m=0}^M w_m x^m$$
$$E(\underline{w}) = \sum_{p=1}^P (y(\underline{w}, x_p) - t_p)^2$$

- a) Berechnen Sie die Ableitungen $\frac{\partial E}{\partial w_k}$ mit $0 \leq k \leq M$ für beliebige Modellkomplexität M . 3P ①

Um den minimalen Fehler zu finden, muss für alle k : $\frac{\partial E}{\partial w_k} = 0$ sein. Dies führt zu einem linearen Gleichungssystem der Form

$$\underline{\underline{A}} \underline{w} = \underline{b}$$

- b) Erstellen Sie das Gleichungssystem, indem Sie Berechnungsvorschriften für $A_{i,j}$ und b_j aufstellen. Welche Dimensionen haben $\underline{\underline{A}}$, \underline{b} und \underline{w} ? 4P ①
- c) Schreiben Sie eine Funktion, die aus einer gegebenen Trainingsmenge $\{(x_p, t_p) \mid p = 1, \dots, P\}$ für eine gegebene Modellkomplexität M die Matrix $\underline{\underline{A}}$ und den Vektor \underline{b} erstellt und das Gleichungssystem $\underline{\underline{A}} \underline{w} = \underline{b}$ löst. 4P ②

Wählen Sie als Trainingsbeispiele

$$x_p = \frac{(2p-1)\pi}{P} - \pi \quad t_p = \cos\left(\frac{x_p}{2}\right)$$

und als Testbeispiele

$$\hat{x}_p = \frac{2p\pi}{P} - \pi \quad \hat{t}_p = \cos\left(\frac{\hat{x}_p}{2}\right).$$

- d) Erzeugen Sie für $P = 11$ die Trainings- und Testbeispiele. Berechnen Sie den Restfehler auf der Trainingsmenge und den Fehler auf der Testmenge für die Modellkomplexitäten $M = 0, \dots, M_{max}$, indem Sie jeweils $\underline{\underline{A}}$, \underline{b} und \underline{w} erstellen. Was beobachten Sie? 5P ②
- e) Ersetzen Sie in der Erzeugung der Beispiele die cos-Funktion durch eine Gaußfunktion

$$g_{\mu, \sigma}(x) = \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right), \mu = 0, \sigma = 0.6$$

und tragen wieder die beiden Fehler gegen M auf. Wie hat sich das Verhalten verändert? 3P ②

bitte wenden

Tabelle 1:

x -Wert	\mathcal{C}_1	\mathcal{C}_2	\mathcal{C}_3					
$1.2 \pm .05$	1	6	1					
$1.3 \pm .05$	2	9	1					
$1.4 \pm .05$	3	10	2					
$1.5 \pm .05$	3	8	3					
$1.6 \pm .05$	2	11	5					
$1.7 \pm .05$	2	10	4					
$1.8 \pm .05$	1	7	8					
$1.9 \pm .05$	1	5	4					
$2.0 \pm .05$	0	4	5					
$2.1 \pm .05$	0	2	7					
$2.2 \pm .05$	0	0	10					
$2.3 \pm .05$	0	0	12					
$2.4 \pm .05$	0	0	16					
$2.5 \pm .05$	1	0	14					
$2.6 \pm .05$	2	0	12					
$2.7 \pm .05$	2	0	9					
$2.8 \pm .05$	3	0	5					
$2.9 \pm .05$	5	0	3					
$3.0 \pm .05$	7	0	1					
$3.1 \pm .05$	13	0	0					
$3.2 \pm .05$	17	0	0					
$3.3 \pm .05$	12	0	0					
$3.4 \pm .05$	7	0	0					

1.2 Bedingte Wahrscheinlichkeiten

Bei der Klassifikation von Mustern anhand eines eindimensionalen Merkmals x sind während der Trainingsphase die in Tabelle 1 aufgeführten Fälle aufgetreten (die leeren Zeilen und Spalten dienen zum Rechnen):

a) Berechnen Sie:

4P ①

- | | |
|--|---|
| (i) $P(\mathcal{C}_1), P(\mathcal{C}_2), P(\mathcal{C}_3)$ | (v) $P(1.45 > x \mid \mathcal{C}_2)$ |
| (ii) $P(1.65 \leq x < 2.65)$ | (vi) $P(\mathcal{C}_1 \mid 1.45 > x)$ |
| (iii) $P(1.65 \leq x < 2.65 \mid \mathcal{C}_3)$ | (vii) $P(\mathcal{C}_2 \mid 1.45 > x)$ |
| (iv) $P(3.05 \leq x \mid \mathcal{C}_1)$ | (viii) $P(\mathcal{C}_3 \mid 1.45 > x)$ |

b) Berechnen Sie für jedes Intervall die a posteriori-Wahrscheinlichkeiten $P(\mathcal{C}_i|x)$ zu den verschiedenen Klassen und zeichnen Sie diese in ein Diagramm. Welches sind die idealen Entscheidungsgrenzen?

4P ①