

Sprachtechnologie

4. Übung

Prof. Tobias Scheffer
Uwe Dick

Sommer 2016

Ausgabe am: 11.05.16
Besprechung am: 23.05.16

Aufgabe 1

Trigramm-Tagger

Bei der Wortarterkennung mit HMMs stellt jeder Zustand ein Pos-Tag (z.B. Nomen, Verben) dar. Um größere Kontextabhängigkeiten zu erkennen, müssen die Zustände mehrere Pos-Tags kodieren. Wie groß ist der Zustandsraum für einen Trigramm-Tagger bei m möglichen Pos-Tags?

Aufgabe 2

Viterbi-Algorithmus

Angenommen Sie hätten die folgenden Wahrscheinlichkeiten aus einem Korpus geschätzt:

$t \setminus t+1$	JJ	VBG	VBP	PRP	VBD	π	schlau	werden	wir
JJ	0.3	0.0	0.1	0.3	0.3	0.3	0.2	0.0	0.0
VBG	0.2	0.2	0.2	0.2	0.2	0.2	0.0	0.3	0.0
VBP	0.2	0.3	0.1	0.3	0.1	0.1	0.0	0.5	0.0
PRB	0.0	0.4	0.3	0.1	0.2	0.3	0.0	0.0	0.2
VBD	0.2	0.1	0.2	0.2	0.3	0.1	0.4	0.0	0.0

Zeichnen Sie das entstandene HMM und bestimmen Sie mit Hilfe des Viterbi-Algorithmus die wahrscheinlichste Zustandsfolge zu der Beobachtung „*Wir werden schlau*“.

Aufgabe 3

Lineare Modelle - Perzeptron

Ein Knoten, der die Vorzeichenfunktion $\sigma(h) = \text{sgn}(h)$, für $h = \theta^T \mathbf{x}$ als Aktivierungsfunktion verwendet, entspricht einem Perzeptron (Folie 73 Basic Models) und kann mit dem Algorithmus 1 trainiert werden. Visualisieren Sie das Neuronale Netz, welches einem Perzeptron entspricht (Was sind die Eingabe- und Ausgabeknoten?).

Algorithm 1 Perzeptron

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

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

Betrachten Sie die Trainingsbeispiele der folgenden Tabelle: 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

- Visualisieren Sie die Trainingsdaten in einem 2-dimensionalen Koordinatensystem und überlegen Sie sich, wie eine optimale Trennebene aussehen sollte.
- Simulieren Sie das Training eines Perzeptrons von Hand. Laufen Sie in jeder Iteration des Algorithmus von links nach rechts über die Tabelle. Klassifizieren Sie Testbeispiel $[3, 5; 3, 5]^T$ unter Verwendung der Klassifikationsfunktion $sgn(h)$. 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.
- Nehmen Sie an, Sie hätten ein zusätzliches positives Trainingsbeispiel $[5, 0; 1, 0]^T$. 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.