Maschinelles Lernen II

2. Übung

Prof. Tobias Scheffer
Dr. Niels Landwehr
Christoph Sawade
Uwe Dick
Peter Haider

Sommer 2011

Ausgabe am: 05.05.11
Besprechung am: 12.05.11

Aufgabe 1 *Memory-basierte Empfehlung*

Gegeben sei folgende Matrix von Trainingsbeispielen:

$$B = \begin{array}{c} \\ \\ \\ \\ \end{array}\begin{array}{ccc|c} u_1 & u_2 & u_3 & \\ \end{array}$$

$$\begin{array}{cc} & \begin{array}{ccc} u_1 & u_2 & u_3 \end{array} \\ B = & \begin{bmatrix} 2 & 1 & 5 \\ 1 & 4 & 3 \\ 2 & ? & 3 \\ 5 & 4 & ? \end{bmatrix} \begin{array}{c} x_1 \\ x_2 \\ x_3 \\ x_4 \end{array} \end{array}$$

Berechnen Sie die Vorhersagen für die fehlenden Einträge mit der Formel aus der Vorlesung:

$$f(u, x) = \frac{\sum_{u' \neq u \wedge \exists i: u^i = u' \wedge x^i = x} w(u, u') y^i}{\sum_{u' \neq u \wedge \exists i: u^i = u' \wedge x^i = x} w(u, u')}$$

wobei sich das Gewicht für den Einfluss zwischen zwei Benutzern $w(u, u')$ aus der Stichprobenkovarianz $(\frac{1}{N-1} \sum_k (x_k - \bar{x})(y_k - \bar{y}))$ der von beiden Benutzern bewerteten Objekten ergibt.

Aufgabe 2 *Diskriminativ gelernte Gewichte*

Stellen Sie ein Optimierungsproblem basierend auf $f(u, x)$ aus Aufgabe 1 auf, das die symmetrischen Gewichte $w(u, u')$ so einstellt, dass der quadratische Verlust auf den Trainingsbeispielen minimiert wird.
Durch die Division durch die Summe der Gewichte gibt es keine eindeutige Lösung für das Optimierungsproblem, die Lösungen sind invariant gegenüber einer uniformen Skalierung

aller Gewichte. Erzwingen Sie eine eindeutige Lösung, indem Sie $w(u_1, u_2) = 1$ fixieren. Überlegen Sie sich, mit welchen Algorithmen man dieses Optimierungsproblem lösen könnte.
Berechnen Sie die optimalen Werte für $w(u_1, u_3)$ und $w(u_2, u_3)$.
Tipp: Optimierungsprobleme mit wenigen Variablen lassen sich mit www.wolframalpha.com lösen. Benennen Sie $w(u_1, u_3)$ und $w(u_2, u_3)$ in w und v um.
Beispiel: `minimize(5*w^2 + 2*v^2 - 2*w + w*v)`

Aufgabe 3 *Latente Features*

Zeigen Sie, dass beim Lernen latenter Features für kollaborative Empfehlung die separate Gewichtung der Regularisierungsterme für Benutzer- und Objektfeatures keinen Effekt hat: das Optimierungsproblem

$$\arg\min_{\psi,\phi} \sum_{i=1}^{n} L(y^i, \psi_{u^i}^T \phi_{x^i}) + \lambda \left[a \sum_j \|\psi_{u_j}\|^2 + \sum_l \frac{1}{a} \|\phi_{x_l}\|^2 \right]$$

führt unabhängig vom Wert der Konstanten a immer zu den gleichen Vorhersagen.
Hinweis: Zeigen Sie, dass $(\frac{1}{\sqrt{a}}\psi^*, \sqrt{a}\phi^*)$ eine Lösung des Optimierungsproblems ist, wenn (ψ^*, ϕ^*) eine Lösung für $a = 1$ ist.