

Multi-Domain Collaborative Filtering

Formel-Sezierung

Multi-Domain Collaborative Filtering

- Bewertungslücken kompensieren
- Algorithmus um viele Domänen abzudecken
- Übertragung von Bewertungskorrelationsproblemen von einer Domäne (z. B. Bücher) auf eine andere Domäne (z. B. Elektronik)
- Automatisch lernendes System
- Hoch komplex ...

Lernfunktion mit Linkfunktion

$$\begin{aligned}
 & J_1(\{\mathbf{U}^i\}, \{\mathbf{V}^i\}, \sigma, \lambda, \eta, \boldsymbol{\Omega}, \theta) \\
 &= \sum_{i=1}^K \left(\frac{1}{2\sigma_i^2} \right) \times \sum_{j=1}^m \left(\sum_{k=1}^{n_i} I_{jk}^i \times \left(g(\mathbf{X}_{jk}^i) - (\mathbf{U}_j^i)^T \mathbf{V}_k^i \right)^2 \right) \\
 &+ \sum_{i=1}^K \left(\frac{1}{2\lambda_i^2} \right) \times \sum_{j=1}^m \left((\mathbf{U}_j^i)^T \mathbf{U}_j^i \right) + \sum_{i=1}^K \left(\frac{1}{2\eta_i^2} \right) \times \sum_{j=1}^m \left((\mathbf{V}_j^i)^T \mathbf{V}_j^i \right) \\
 &+ \frac{1}{2} \times \sum_{i=1}^K \left(\ln \left(\sigma_i^2 \times \sum_{j=1}^m \left(\sum_{k=1}^{n_i} I_{jk}^i \right) \right) \right) + \frac{md}{2} \times \sum_{i=1}^K (\ln(\lambda_i^2)) \\
 &+ \sum_{i=1}^K \left(\frac{dn_i}{2} \times \ln(\eta_i^2) \right) + \frac{1}{2} \times \text{tr}(\mathbf{U}\boldsymbol{\Omega}^{-1}\mathbf{U}^T) + \frac{md}{2} \times \ln|\boldsymbol{\Omega}| \\
 &- \sum_{i=1}^K \left(\sum_{j=1}^m \left(\sum_{k=1}^{n_i} (I_{jk}^i \times \ln(g'(\mathbf{X}_{jk}^i)) + \text{Konstante}) \right) \right)
 \end{aligned}$$

Die Linkfunktion

- Ungenauigkeit entgegenwirken
- Umwandlung in Binärform in 1. Definition
- Verlust von Bewertungsgenauigkeit
 - z. B. Bewertungsskala von 1 bis 10
- Erneut hoch komplex ...

Bedingte Bewertungsverteilung in der i -ten Domäne

(1. Definition)

$$p(\mathbf{X}^i | \mathbf{U}^i, \mathbf{V}^i, \sigma_i) = \prod_{j=1}^m \prod_{k=1}^{n_i} \left[\mathcal{N} \left(X_{jk}^i | (\mathbf{U}_j^i)^T \mathbf{V}_k^i, \sigma_i^2 \right) \right]^{I_{jk}^i}$$

| | |
|--|--|
| \mathbf{X}^i | ... Bewertungsmatrix für die i -te Domäne |
| \mathbf{U}^i | ... Latente Benutzermatrix |
| \mathbf{V}^i | ... Latente Objekteigenschaftsmatrix |
| σ_i | ... Standardabweichung (<i>Sigma</i>) |
| j | ... Benutzerindex / -iterator |
| m | ... alle Benutzer, aller Domänen |
| k | ... Objektindex / -iterator |
| n_i | ... alle Objekte der i -ten Domäne |
| $[\cdot]^{I_{jk}^i}$ | ... Indikatorfunktion für Binärbewertung (1 falls Bewertung vorliegt, sonst 0) |
| $\mathcal{N}(\mathbf{m}, \mathbf{\Sigma})$ | ... Multivarianz der Normalverteilung, vom mittleren \mathbf{m} und der Kovarianzmatrix $\mathbf{\Sigma}$ (<i>Sigma</i>) |
| X_{jk}^i | ... Bewertung des j -ten Benutzers des k -ten Objektes in \mathbf{X}^i |
| $(\mathbf{U}_j^i)^T$ | ... Transponierter, latenter Benutzervektor |
| \mathbf{V}_k^i | ... Latenter Objekteigenschaftsvektor |
| σ_i^2 | ... Varianz (<i>Sigma</i>) |