

Kollaboratives Filtern

Beim **kollaborativen Filtern** (*collaborative filtering*) werden Verhaltensmuster von Benutzergruppen ausgewertet, um auf die Interessen Einzelner zu schließen. Dabei handelt es sich um eine Form des **Data-Mining**, die eine explizite Nutzereingabe überflüssig macht.

1 Ziel

Die Anwendung von kollaborativem Filtern erfolgt meistens für sehr große Datenmengen. Kollaboratives Filtern wird für die verschiedensten Bereiche angewandt wie z.B. im Finanzdienstleistungssektor zur Integration finanzieller Quellen oder in Anwendungen im **eCommerce** und **Web 2.0**. Dieser Artikel beschäftigt sich mit dem kollaborativen Filtern für Benutzerdaten, auch wenn manche Methoden und Ansätze auf andere Bereiche übertragen werden können.

Das Ziel der Methode ist eine automatische Vorhersage (Filtern) von Benutzerinteressen. Zu diesem Zweck werden Informationen über das Verhalten und Vorlieben von möglichst vielen Nutzern gesammelt (**Kollaboration**). Die zugrunde liegende Annahme von kollaborativem Filtern ist, dass jemand der sich für etwas (z.B. eine Fernsehsendung) in der Vergangenheit interessiert hat, sich auch in der Zukunft dafür interessieren wird. Durch kollaboratives Filtern kann für das Fernsehprogramm eine Vorhersage gemacht werden, welche Fernsehsendung einem Zuschauer gefallen könnte. Der Output wäre eine Liste mit möglichen favorisierten Fernsehsendungen. Es gilt zu beachten, dass diese Vorhersage für jeden einzelnen Zuschauer individuell gemacht wird. Die Datengrundlage zur Vorhersage wird von der Gesamtheit der Nutzer gesammelt. Hier unterscheidet sich das kollaborative Filtern von einfacheren Methoden, bei denen ein unspezifischer Mittelwert errechnet wird.

Ein spezifisches Problem kollaborativer Filter besteht in ihrer Latenzzeit: Ein neuer Nutzer tritt mit einem leeren Benutzerprofil in das System ein. Da seine Interessen noch nicht bekannt sind, kann er zu Beginn keine sinnvollen Empfehlungen erhalten. Gleiches gilt für neu in das System eintretende Elemente (z. B. Produkte in einem Online-Shop). Sie weisen keine quantifizierbare Ähnlichkeit mit anderen Elementen auf und können damit nicht sinnvoll empfohlen werden. Es handelt sich bei kollaborativen Filtern also um lernende Systeme und damit um eine Form der künstlichen Intelligenz.

2 Methodik

Kollaboratives Filtern läuft meistens in zwei Schritten ab.

1. Suche nach Nutzern, die das gleiche Verhaltensmuster wie der aktive Nutzer haben. (= der Nutzer für den die Vorhersage getroffen wird)
2. Verwendung der Verhaltensmuster um eine Vorhersage für den aktiven Nutzer zu treffen.

Alternativ dazu gibt es das artikelbasierte kollaborative Filtern, das durch Amazon.com bekannt wurde („Das könnte Sie auch interessieren.“) und erstmals von Vucetic und Obradovic im Jahre 2000 vorgestellt wurde.

1. Erstellen einer Ähnlichkeitsmatrix zur Bestimmung von Beziehungen zwischen Artikeln.
2. Aus der Matrix werden die Vorlieben des aktiven Nutzers abgeleitet.

Weitere Formen des Kollaborativen Filterns können auf impliziter Beobachtung der Nutzerverhalten beruhen. Bei diesen Formen des Filterns wird das Verhalten des einzelnen Benutzers mit dem Verhalten aller anderen Benutzer verglichen (Welche Musik haben sie gehört? Welche Produkte haben sie gekauft?). Diese Daten werden genutzt um das zukünftige Verhalten des Nutzers vorhersagen zu können. Dabei ist es nicht sinnvoll, einem Nutzer ein bestimmtes Musikstück anzubieten, wenn er durch sein Verhalten deutlich gemacht hat, dass er es bereits besitzt. Ebenso ist es nicht sinnvoll, einem Nutzer weitere Paris-Reiseführer anzubieten, wenn er bereits einen Reiseführer für diese Stadt besitzt.

Im heutigen Informationszeitalter stellen sich diese und ähnliche Technologien als äußerst hilfreich für die Produktauswahl heraus, gerade wenn bestimmte Produktgruppen (z.B. Musik, Filme, Bücher, Nachrichten, Internetseiten) so groß geworden sind, dass einzelne Personen das gesamte Angebot nicht überblicken können.

3 Anwendung

3.1 In kommerziellen Systemen

Kommerzielle Webseiten, die kollaboratives Filtern nutzen:

- Amazon
- Amie Street
- Barilliance
- Barnes and Noble
- Baynote
- ChoiceStream
- Collarity
- Digg.com
- eBay
- Google News ^[1]
- Gravity R&D
- half.ebay.com
- Heeii
- Hollywood Video
- Hulu
- iLike - Musik
- Internet Movie Database - Filme
- iTunes - Musik
- Last.fm - Musik
- LibraryThing - Bücher
- Loomia – Software as a Service Provider von Empfehlungs Technologien
- Musicmatch
- MyStrands - Entwickler von social recommendation Technologien
- Netflix
- Simania - Buchempfehlungsseite
- Strands – Strands nutzt und wirbt für seine eigene Empfehlungs Engine für Soziale Netzwerke und eCommerce
- StumbleUpon - Internetseiten
- Threadless - T-Shirt
- TiVo
- Yelp
- Ramkol - Sophisticated recommendation for local search in Israel

3.2 In nicht kommerziellen Systemen

- AmphetaRate – RSS Artikel
- Everyone's a Critic - Filme
- GiveALink.org - Internetseiten
- Gnomoradio - Musik
- MovieLens - Filme
- Rate Your Music Musik

4 Literatur

- Andreas Meier (Autor), Henrik Stormer (Autor): *eBusiness & eCommerce: Management der digitalen Wertschöpfungskette*. Springer, Berlin 2009, 3540850163.
- Robert Buchberger(Autor): *Contentmanager.de* http://www.contentmanager.de/magazin/artikel_47-print_wenn_es_persoendlich_wird_- .html, abgerufen am 14. April 2010
- Goldberg, David; David Nichols, Brain M. Oki, Douglas Terry: *Using collaborative filtering to weave an information tapestry* 1992, Communications of the ACM 35 (12): 61–70.
- Direkt Marketing. „Re-Targeting soll Interessenten zurückholen.“ November 2007: 29.
- Torben Brodt: *Collaborative Filtering: für automatische Empfehlungen*. VDM Verlag, Saarbrücken 2010. ISBN 3-639-25509-7.

5 Quellen

- [1] Google news personalization: scalable online collaborative filtering, Abhinandan S. Das, Mayur Datar, Ashutosh Garg, Shyam Rajaram, Proceeding WWW '07 Proceedings of the 16th international conference on World Wide Web ACM New York, NY, USA ©2007, ISBN 978-1-59593-654-7 doi>10.1145/1242572.1242610

Normdaten (Sachbegriff): GND: 7512839-1

6 Text- und Bildquellen, Autoren und Lizenzen

6.1 Text

- **Kollaboratives Filtern** *Quelle:* https://de.wikipedia.org/wiki/Kollaboratives_Filtern?oldid=141759332 *Autoren:* Aka, Andim, Nepomucki, Cepheiden, Kam Solusar, Kaffeefan, Rax, Flominator, Hydro, Zahnradzacken, Zuphilip, Gary Dee, Wikidienst, Knase, Krd, Xqbot, Rr2000, Renepick, Iltu, ZéroBot, Fmuessig, Hänz ab, Toru10, KLBot2, Boshomi, Dexbot und Anonyme: 5

6.2 Bilder

6.3 Inhaltslizenz

- Creative Commons Attribution-Share Alike 3.0