

Modelowanie i analiza sieci złożonych

XIII. Systemy rekomendacyjne.

Grzegorz Siudem

Politechnika Warszawska



Politechnika
Warszawska

Unia Europejska
Europejski Fundusz Społeczny



Zadanie 10 pn.

„Przygotowanie i uruchomienie nowego kierunku studiów na studiach II stopnia
- Inżynieria i Analiza Danych (IAD)”

realizowane jest w ramach projektu
„NERW PW. Nauka – Edukacja – Rozwój – Współpraca”
współfinansowanego ze środków Unii Europejskiej
w ramach Europejskiego Funduszu Społecznego

Wykład

Czym są?

- to algorytmy i techniki próbujące przewidzieć czego chce użytkownik.
- poszukujemy najlepiej dopasowanych par użytkownik-rzecz.
- optymalna krawędź w grafie dwudzielnym?

Rozróżniamy podejścia:

- bazujące na historycznych wyborach wszystkich użytkowników (ang. *Collaborative Filtering*),
- bazujące na cechach użytkownika i produktu (ang. *Content-based recommending*),
- mieszane (ang. *Hybrid approaches*).

Polecam lekturę <http://www.prem-melville.com/publications/recommender-systems-empl2010.pdf>

		<i>Items</i>												
		<i>1</i>	<i>2</i>	<i>...</i>	<i>i</i>	<i>...</i>	<i>m</i>							
<i>Users</i>	<i>1</i>	5	3		1	2								
	<i>2</i>		2				4							
	:			5										
	<i>u</i>	3	4		2	1								
	:					4								
	<i>n</i>			3	2									
		<table border="1" style="border-collapse: collapse; margin: 0 auto;"> <tr> <td style="padding: 2px 10px;"><i>a</i></td> <td style="padding: 2px 10px; text-align: center;">3</td> <td style="padding: 2px 10px; text-align: center;">5</td> <td style="padding: 2px 10px;"></td> <td style="padding: 2px 10px; text-align: center;">?</td> <td style="padding: 2px 10px; text-align: center;">1</td> <td style="padding: 2px 10px;"></td> </tr> </table>						<i>a</i>	3	5		?	1	
<i>a</i>	3	5		?	1									

Recommender Systems. Figure 1. User ratings matrix, where each cell $r_{u,j}$ corresponds to the rating of user u for item i . The task is to predict the missing rating $r_{a,i}$ for the active user a

Metody typu *Collaborative Filtering* dodatkowo dzielimy na

- *neighborhood-based*, gdzie poszukujemy użytkowników podobnych do analizowanego, a następnie na podstawie ich wyborów opracowujemy predykcję według schematu:
 - Obliczamy wagi dla każdego użytkownika zgodnie z ich podobieństwem (mierzonym np. współczynnikiem Pearsona albo Spearmana albo Kendalla albo błędem średniokwadratowym, etc.) do wybranego użytkownika.
 - Wybieramy k najbliższych sąsiadów.
 - Predykcję otrzymujemy jako ważoną średnią odpowiedzi sąsiadów.
- *model-based*, w którym estymujemy parametry modelu charakteryzującego użytkowników, a rekomendację przeprowadzamy w oparciu o model.

Metody typu *Content-based Recommending* to problem klasyfikacji, a zatem możemy wykorzystać

- naiwne klasyfikatory Bayesowskie,
- algorytm k -najbliższych sąsiadów,
- drzewa decyzyjne,
- sieci neuronowe,
- etc.

Dziękuję za uwagę!



**Politechnika
Warszawska**

Unia Europejska
Europejski Fundusz Społeczny



Zadanie 10 pn.

„Przygotowanie i uruchomienie nowego kierunku studiów na studiach II stopnia
- Inżynieria i Analiza Danych (IAD)”

realizowane jest w ramach projektu
„NERW PW. Nauka – Edukacja – Rozwój – Współpraca”
współfinansowanego ze środków Unii Europejskiej
w ramach Europejskiego Funduszu Społecznego