

Wstęp do rekomendacji

Na przykładzie serwisu streamingowego

Karol Chęciński

Plan prezentacji

1. Dlaczego potrzebujemy rekomendacji?
2. Czego oczekujemy od rekomendacji?
3. Trochę o danych
4. Zgrubna klasyfikacja systemów rekomendacyjnych
5. Przykładowe modele i algorytmy rekomendacyjne

Dlaczego potrzebujemy rekomendacji?



- Ogrom danych, możliwości, zróżnicowanych produktów
- Użytkownicy o zróżnicowanych gustach i potrzebach
- Potrzeba dostosowania przekazu

Przykłady



allegro

last.fm



Tripadvisor®

Dobre rekomendacje, czyli jakie?



Czego oczekujemy od rekomendacji?

1

Relevantność (ang. *relevance*)

- Rekomendacje mają odpowiadać preferencjom użytkownika
- Absolutna podstawa...
- ...Ale to nie wystarczy

Czego oczekujemy od rekomendacji?

2

Nowość
(ang. *novelty*)

- Rekomendacje mają dostarczać nowych produktów
- Nie możemy rekomendować ciągle tych samych produktów, nawet jeśli są idealne dla użytkownika
- Warto zadbać, by rekomendacje zawierały produkty, których użytkownik nie widział

Czego oczekujemy od rekomendacji?

3

**Zdolność do
przypadkowych
odkryć**
(ang. *serendipity*)

- Nie możemy szufladkować użytkowników
- Nie możemy zanudzać użytkowników
- Należy eksplorować i rozwijać upodobania użytkownika

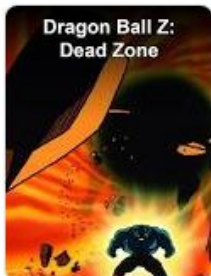
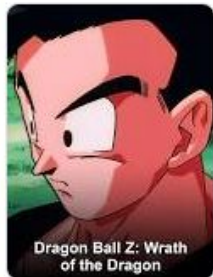
Czego oczekujemy od rekomendacji?

4

Różnorodność (ang. *diversity*)

- Różnorodność rekomendacji zwiększa prawdopodobieństwo, że użytkownik znajdzie coś dla siebie

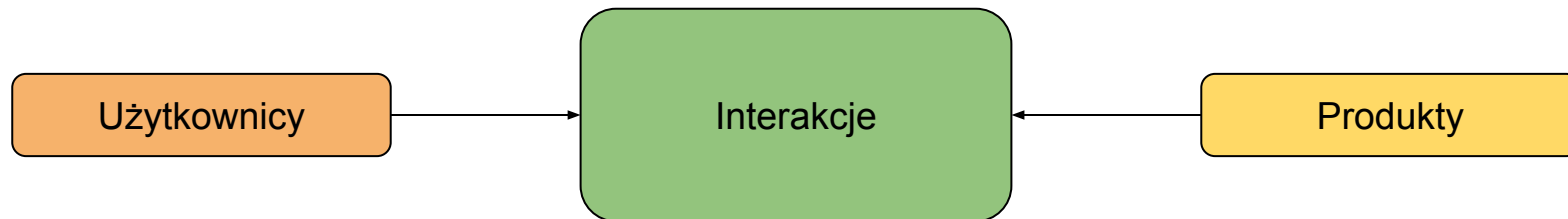
Super relewantny rekomender! (ale bez zadbania o różnorodność)



Dane



Dane



Interakcje

ID użytkownika	ID produktu	wartość
1	A	5
1	B	3
2	B	2
3	C	5
4	D	1
1	D	3

Interakcje - macierz rzadka

ID użytkownika	ID produktu	wartość
1	A	5
1	B	3
2	B	2
3	C	5
4	D	1
1	D	3

	1	2	3	4
A	5			
B	3	2		
C			5	
D	3			1

Oceny jawne (ang. explicit ratings)

- Powiązane ze świadomym wyrażeniem opinii przez użytkownika
- Pozwalają wyrażać zarówno aprobatę jak i dezaprobatę

Oceny jawne (ang. explicit ratings)



2M



7.8M



“interwałowa” (ang. interval-based)

- n poziomów
- różna semantyka

binarna (pozytywna / negatywna)

unarna (jedynie pozytywna, brak możliwości negatywnej oceny)

Oceny niejawne (ang. implicit rating)



More Details



Oceny niejawne - charakterystyka

Łatwiej dostępne

Brak negatywnej preferencji

Raczej prawdopodobieństwo
niż "siła" preferencji

Zaszumienie

Specyficzna ewaluacja

Algorytmy rekomendacyjne



Systemy rekomendacji - klasyfikacja

Collaborative filtering

- Bazuje jedynie na macierzy ocen

Content-based

- Bazuje na ocenach danego użytkownika oraz informacjach o produkcie

Knowledge-based

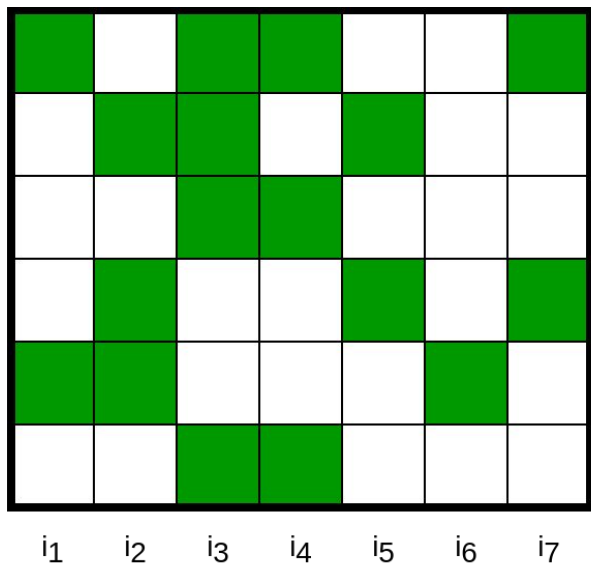
- Bazuje na jawnych preferencjach użytkownika oraz wiedzy dziedzinowej

Collaborative filtering

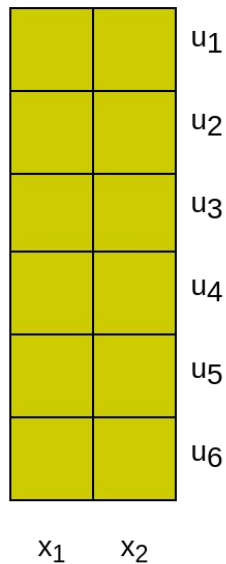
							u_1
							u_2
							u_3
?		?	?		?		u_4
							u_5
							u_6
i_1	i_2	i_3	i_4	i_5	i_6	i_7	

- Nie musimy wiedzieć nic o użytkownikach ani produktach
- Musimy wiedzieć coś o dotychczasowej aktywności użytkownika
- Musimy mieć dużą liczbę ocen innych użytkowników (problem “zimnego startu”)
- Promowanie popularnych produktów

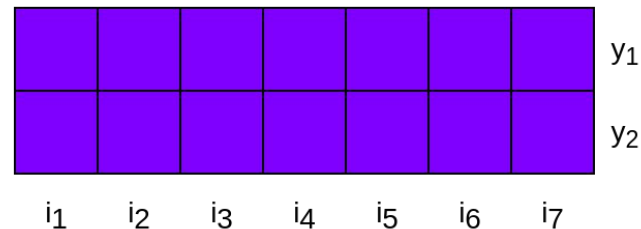
Model z faktoryzacją macierzy



≈



•



$$\hat{r}_{ui} = x_u^T y_i$$

Model z faktoryzacją macierzy (dla ocen jawnych)

$$\min_{x_*, y_*} \sum_{u, i} (r_{ui} - x_u^T y_i)^2 + \lambda (||x_u||^2 + ||y_i||^2)$$

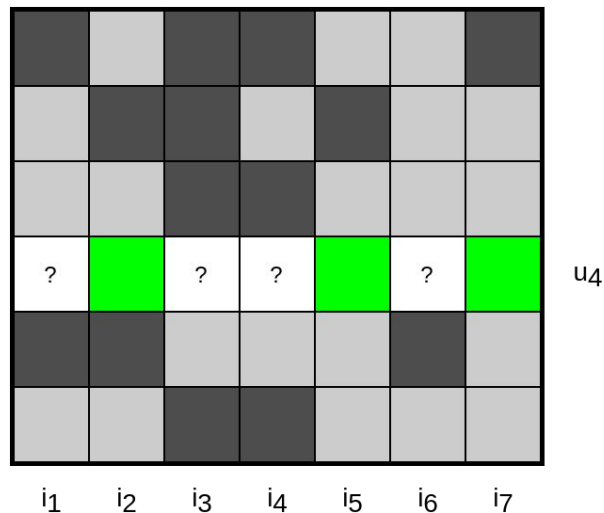
Optymalizacja metodą ALS (Alternating-Least-Squares)

Model z faktoryzacją macierzy (dla ocen niejawnych)

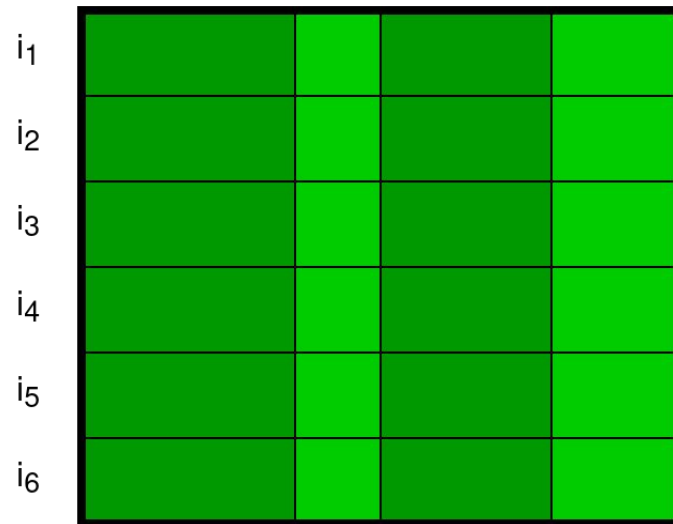
$$c_{ui} = r_{ui} \qquad p_{ui} = \begin{cases} 0 & \text{if } r_{ui} = 0 \\ 1 & \text{if } r_{ui} > 0 \end{cases}$$

$$\min_{x_*, y_*} \sum_{u, i} c_{ui} (p_{ui} - x_u^T y_i)^2 + \lambda (||x_u||^2 + ||y_i||^2)$$

Content-based filtering - wykorzystywane dane



oceny użytkownika



informacje o produktach

Content-based filtering

1. Budowanie cech produktów
2. Budowanie profilu użytkownika
3. Generowanie rekomendacji na podstawie profilu i cech produktów

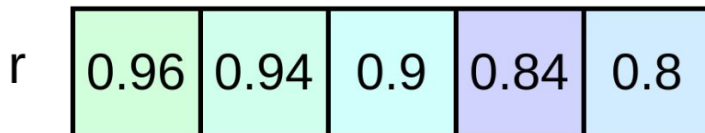
Content-based filtering - Budowanie cech produktów

- Cel - utworzenie reprezentacji produktu w przestrzeni cech
- Przykłady:
 - Budowa cech w oparciu o tagi
 - Wydobycie słów kluczowych z tekstu (Zabawa z NLP)
- Selekcja i ważenie atrybutów - nie każdy atrybut niesie równie wartościową informację

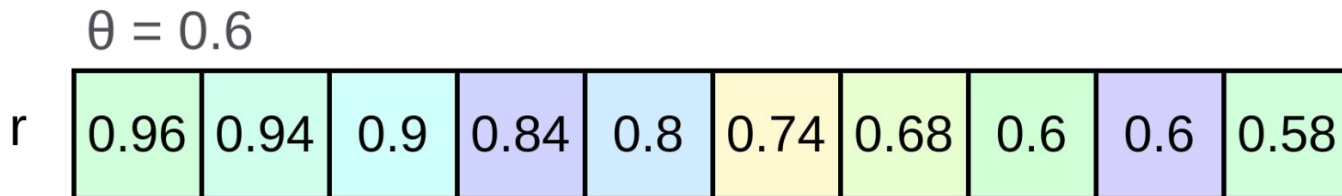
Content-based filtering - przykładowe podejście

- Sprowadzenie rekomendacji do problemu klasyfikacji / regresji
- Cechy ocenionych produktów i ich oceny - zbiór treningowy - wejście i pożądane wyjście
- Np. oceny binarne (łapka w górę / łapka w dół) => problem klasyfikacji binarnej
- Decyzja o rekomendacji produktu => predykcja klasy

A co z różnorodnością?

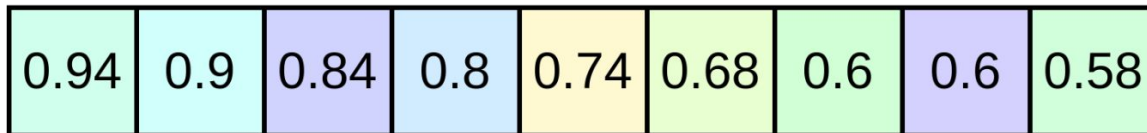


A co z różnorodnością?



A co z różnorodnością?

$\theta = 0.6$



A co z różnorodnością?

rank(r)	9	8	7	6	5	4	3	2	1
	0.94	0.9	0.84	0.8	0.74	0.68	0.6	0.6	0.58
d	0.06	0.25	1.25	0.69	0.86	0.28	0.01	1.34	0.0
rank(d)	5	3	8	6	7	4	2	9	1
$\theta \text{rank}(d) + (1 - \theta) \text{rank}(r)$	6.6	5	7.6	6	6.2	4	2.4	6.2	1
	0.96								

A co z różnorodnością?

rank(r)	9	8	7	6	5	4	3	2	1
	0.94	0.9		0.8	0.74	0.68	0.6	0.6	0.58
d	0.06	0.25	1.25	0.69	0.86	0.28	0.01	1.34	0.0
rank(d)	5	3	8	6	7	4	2	9	1
$\theta \text{rank(d)} + (1 - \theta) \text{rank(r)}$	6.6	5	7.6	6	6.2	4	2.4	6.2	1
	0.96	0.84							

A co z różnorodnością?

rank(r)	8	7	6	5	4	3	2	1
	0.94	0.9	0.8	0.74	0.68	0.6	0.6	0.58
d	0.47	0.39	0.42	1.43	1.06	0.67	0.67	0.61
rank(d)	3	1	2	8	7	5	6	4
$\theta \text{rank(d)} + (1 - \theta) \text{rank(r)}$	5	3.4	3.6	6.8	5.8	4.2	4.4	2.8
	0.96	0.84						

A co z różnorodnością?

rank(r)	7	6	5	4	3	2	1
	0.94	0.9	0.8	0.68	0.6	0.6	0.58
d	0.72	0.78	0.91	0.78	0.70	1.11	0.70
rank(d)	3	4	6	5	1	7	2
$\theta \text{rank}(d) + (1 - \theta) \text{rank}(r)$	4.6	4.8	5.6	4.6	1.8	5	1.6
	0.96	0.84	0.74				

A co z różnorodnością?

rank(r) 6 5

0.94	0.9
------	-----

d 0.63 0.62

rank(d) 2 1

$0.5 \text{rank}(d) + (1 - 0.5) \text{rank}(r)$ 3.6 2.6

rank(r) 4 3 2 1

0.68	0.6	0.6	0.58
------	-----	-----	------

0.94 0.72 0.88 0.69

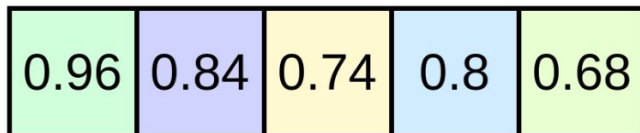
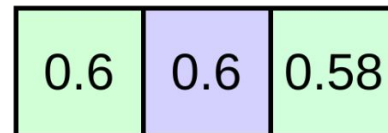
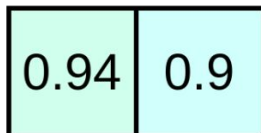
6 4 5 3

5.2 3.6 3.8 2.2

0.96	0.84	0.74	0.8
------	------	------	-----

Ziegler, Cai-Nicolas & Cai-Nicolas, & McNee, Sean & M, Sean & Konstan, & A, Joseph & Lausen, & Georg.. (2005). Improving recommendation lists through topic diversification.

A co z różnorodnością?



Ziegler, Cai-Nicolas & Cai-Nicolas, & McNee, Sean & M, Sean & Konstan, & A, Joseph & Lausen, & Georg,. (2005). Improving recommendation lists through topic diversification.

A co z różnorodnością?

przed	0.96	0.94	0.9	0.84	0.8
po	0.96	0.84	0.74	0.8	0.68

Ewaluacja modeli



Problemy z ewaluacją

- Wiele aspektów, które wpływają na jakość rekomendacji
- Ostateczny rezultat zależy od interakcji z żywym użytkownikiem - trudno to zasymulować

Podstawowy podział metod ewaluacji

Ewaluacja offline

- Podejście “klasyczne”
- Oparte na danych historycznych
- Istnieją zbiory benchmarkowe
- Jedynie pośrednie przełożenie na ostateczny wynik

Ewaluacja online

- Eksperymenty na żywym organizmie
- Możemy wyznaczać to co faktycznie ma znaczenie
- Stosunkowo powolne (gromadzenie danych)
- Wymaga odpowiedniej infrastruktury

Co możemy mierzyć?

- Dokładność - na ile relewantne są nasze rekomendacje?
2 podstawowe podejścia (w zależności od tego co zwraca nasz model)

Predykowane oceny

r_{ui}	\hat{r}_{ui}	$e_{ui} = \hat{r}_{ui} - r_{ui} $
5	3.71	1.29
4	4.11	0.11
3	2.01	2.99

Ranking produktów

1		1	
2		2	
3		3	
4		4	
5		5	

Uwaga na dane!

Dostępne dane wpływają na sens użycia poszczególnych metryk

implicit ratings => brak negatywnej preferencji => raczej czułość niż precyzja

Semantyka ocen, a rozróżnienie między pozytywną i negatywną preferencją

Metryki rankingowe

$$\overline{rank} = \frac{\sum_{u,i} rel_{ui}^t rank_{ui}}{\sum_{u,i} rel_{ui}^t}$$

average-rank

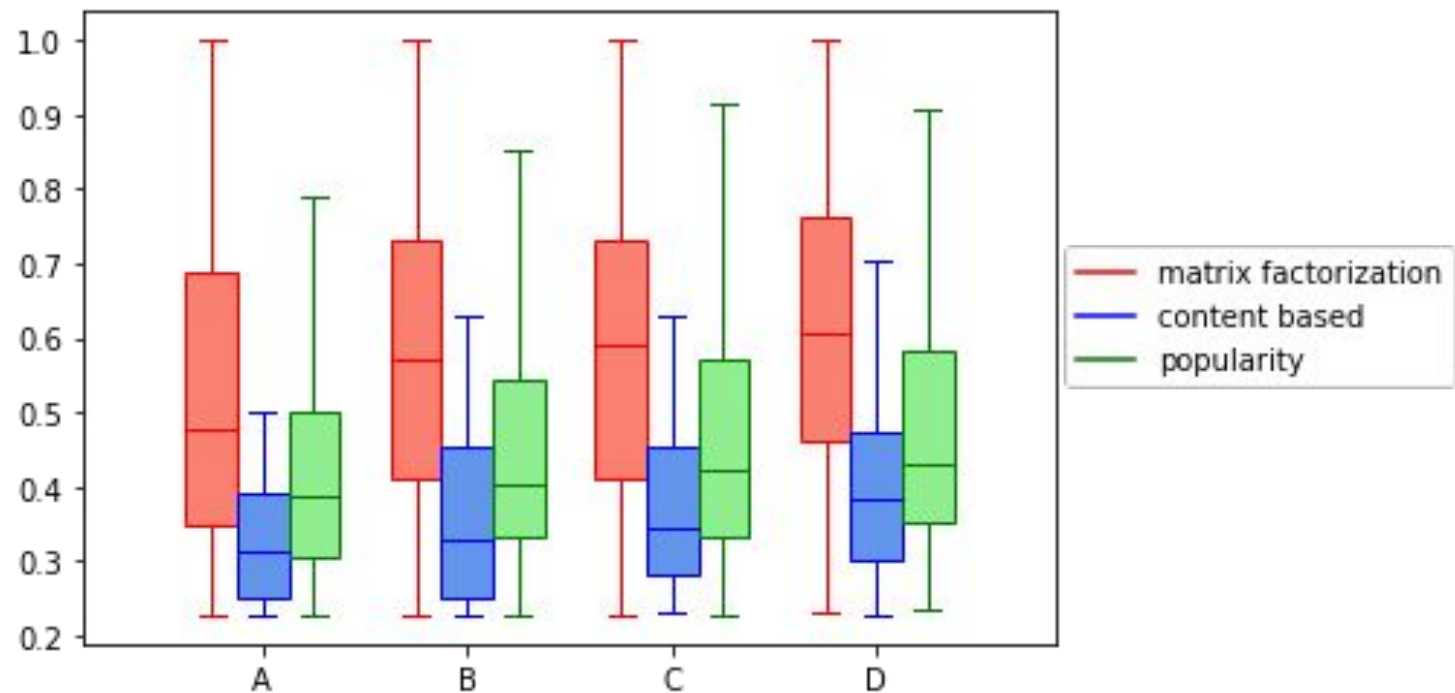
$$DCG_{@K}(Q_u) = \sum_{k=1}^{|Q_u|} \frac{2^{rel_{uk}} - 1}{\log_2(k + 1)}$$

discounted cumulative gain (DCG)

$$nDCG_{@K}(Q) = \frac{DCG_{@K}(Q)}{DCG_{@K}(Q_{ideal})}$$

normalized DCG (nDCG)

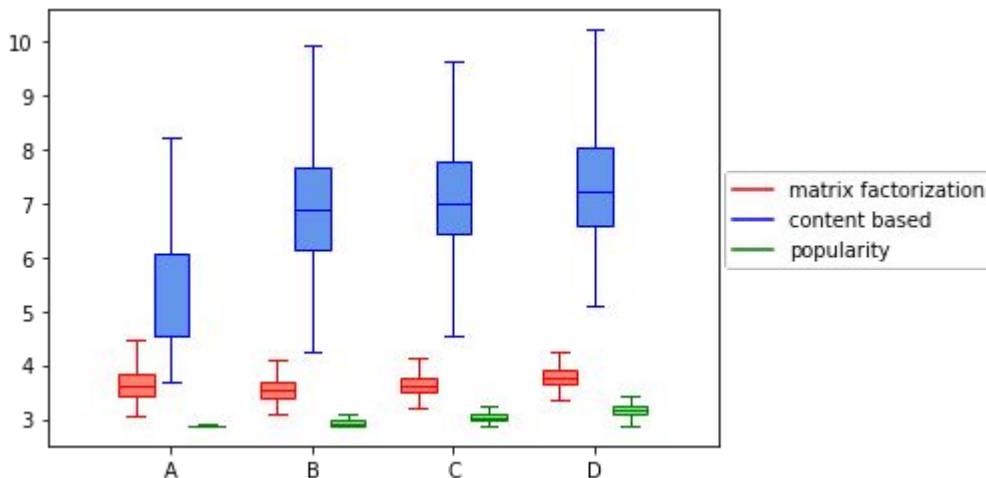
nDCG



Novelty

$$nov(R_u) = \sum_{k=1}^{|R_u|} disc(k) \frac{\log_2 pop(i_k)}{|R_u|}$$

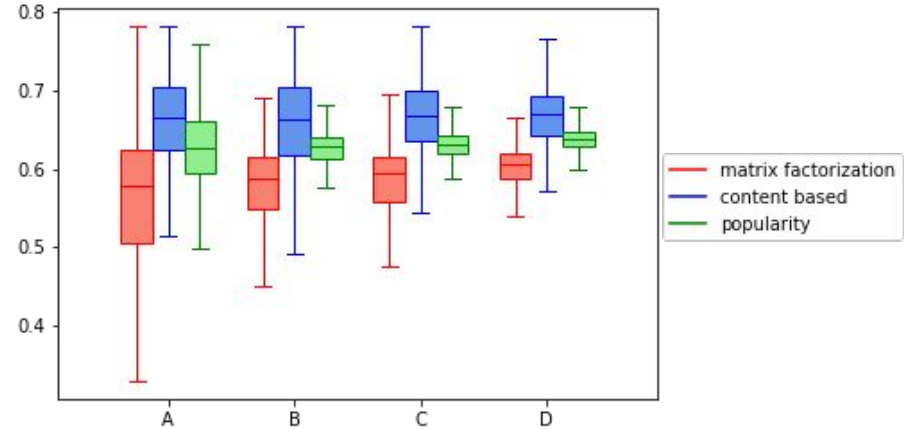
$$disc(k) = \frac{1}{\log_b(b+k)}$$



Serendipity

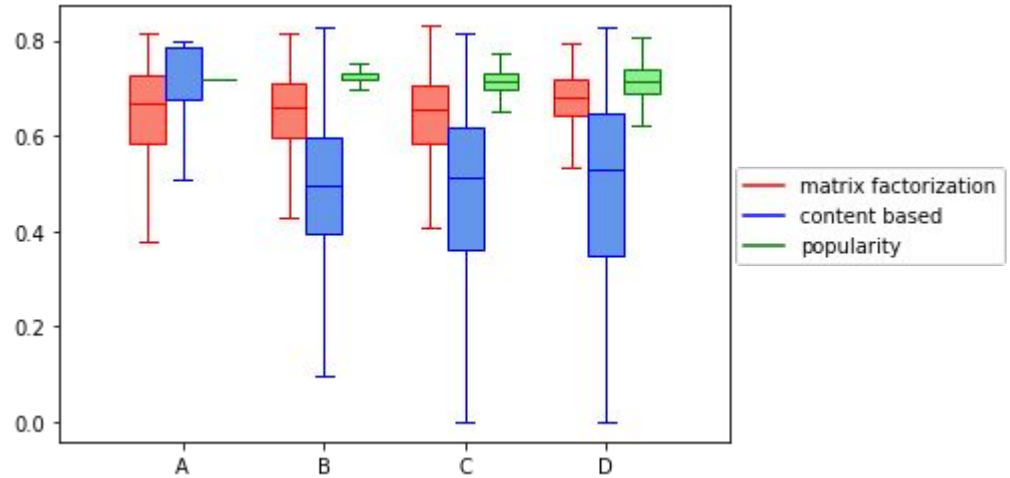
$$ser(R_u) = \frac{1}{|R_u|} \sum_{k=1}^{|R_u|} \frac{1}{|H_u|} \sum_{l \in H_u} disc(k) dist(i_k, i_l)$$

$$dist(i_k, i_l) = 1 - cossim(tfidf(i_k), tfidf(i_l))$$



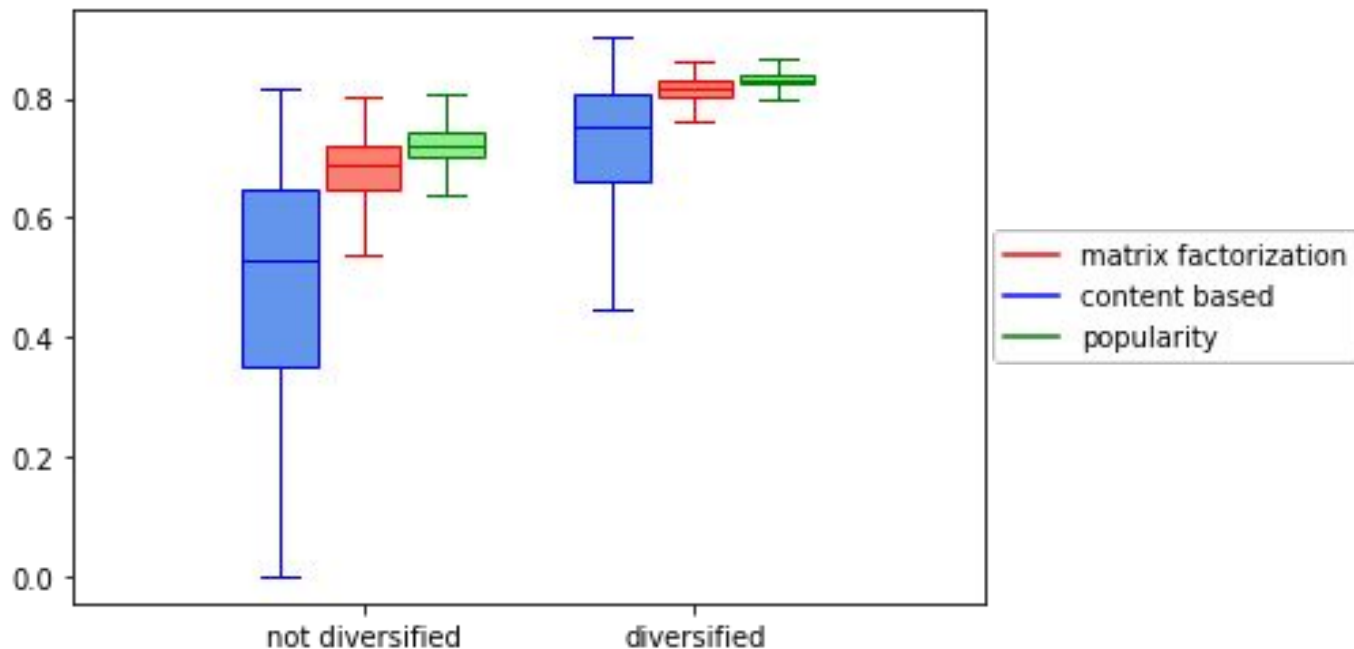
Diversity

$$div(R_u) = \frac{1}{R_u^2} \sum_{k=1}^{|R_u|} \sum_{l=k}^{|R_u|} dist(i_k, i_l)$$

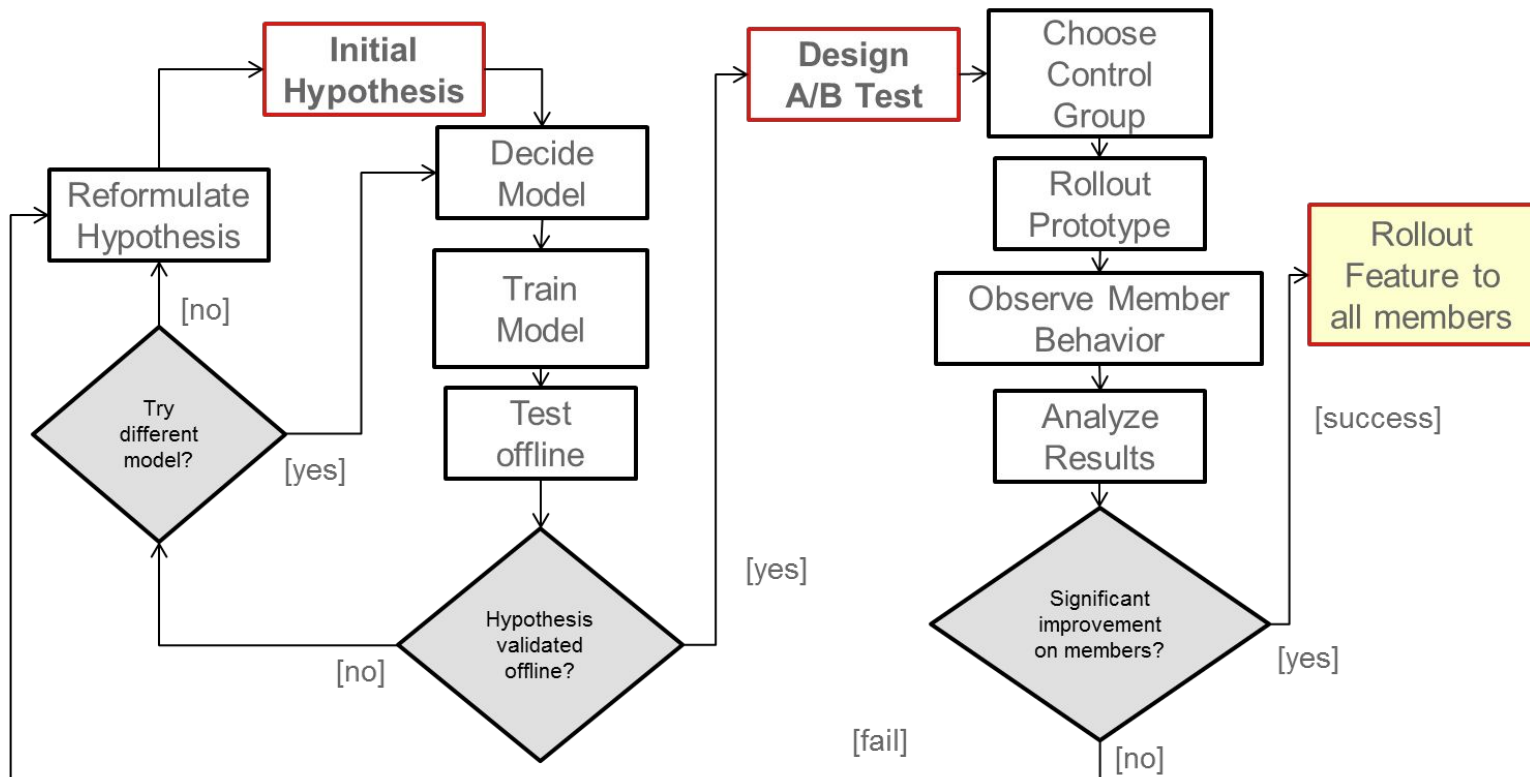


Silveira, T., Zhang, M., Lin, X., Liu, Y. and Ma, S., 2017. How good your recommender system is? A survey on evaluations in recommendation. *International Journal of Machine Learning and Cybernetics*, 10(5), pp.813-831.

Diversity - z modelem zwiększającym różnorodność



Ewaluacja online - Jak to robi Netflix?



Dziękuję za uwagę

Źródła

- Charu C. Aggarwal. 2016. *Recommender Systems: The Textbook* (1st. ed.). Springer Publishing Company.
- Y. Hu, Y. Koren and C. Volinsky. 2008. Collaborative Filtering for Implicit Feedback Datasets. Eighth IEEE International Conference on Data Mining, 2008, pp. 263-272, doi: 10.1109/ICDM.2008.22.
- Yuan Cao Zhang, Diarmuid Ó Séaghdha, Daniele Quercia, and Tamas Jambor. 2012. Auralist: introducing serendipity into music recommendation. In Proceedings of the fifth ACM international conference on Web search and data mining (WSDM '12). Association for Computing Machinery, New York, NY, USA, 13–22. doi: 10.1145/2124295.2124300
- Cai-Nicolas Ziegler, Sean M. McNee, Joseph A. Konstan, and Georg Lausen. 2005. Improving recommendation lists through topic diversification. In Proceedings of the 14th international conference on World Wide Web (WWW '05). Association for Computing Machinery, New York, NY, USA, 22–32. doi: 10.1145/1060745.1060754
- <https://netflixtechblog.com/netflix-recommendations-beyond-the-5-stars-part-2-d9b96aa399f5>