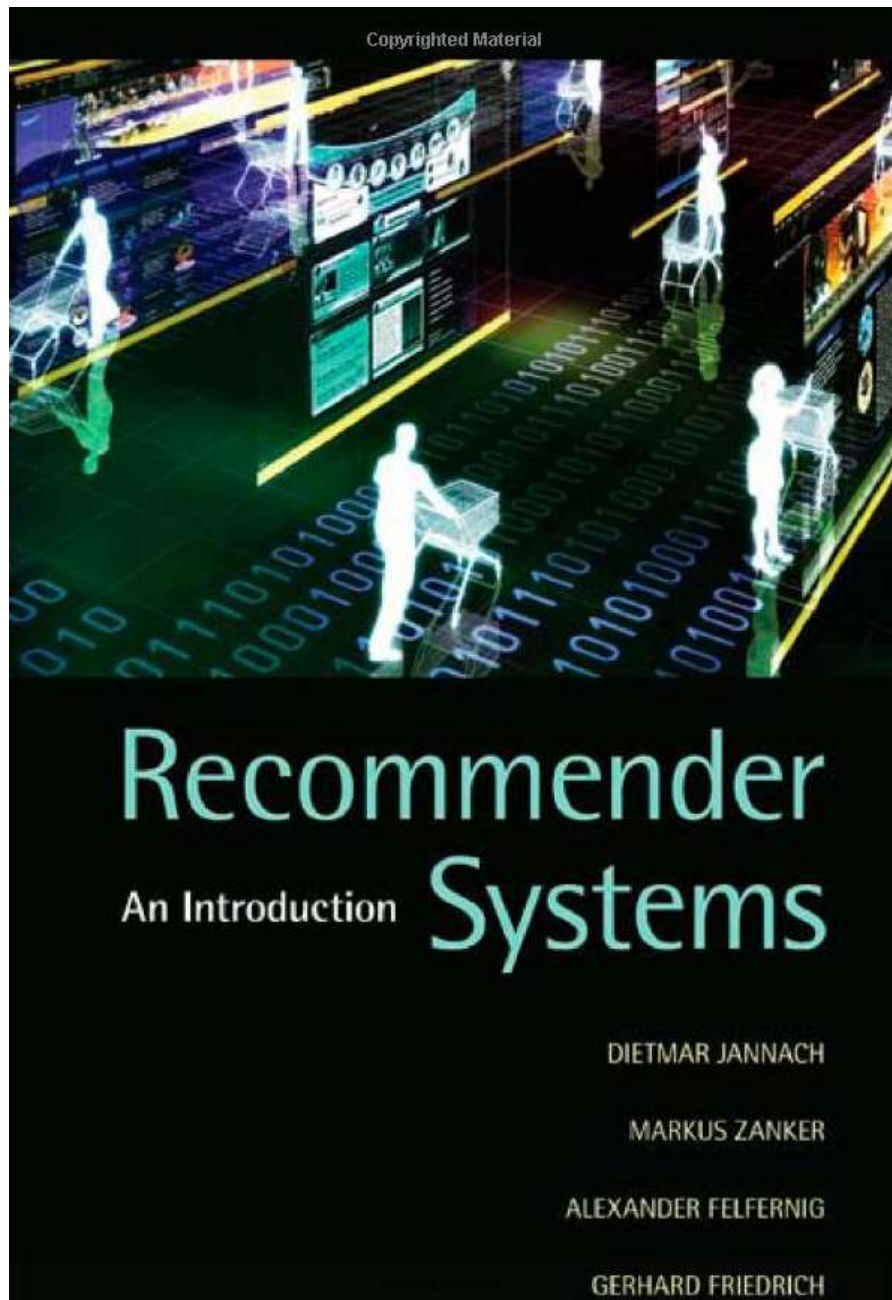


AJÁNLÓ RENDSZEREK

Werner Ágnes



- a szabványos, egyenvásárlóra tervezett tömegtermékek ideje lejárt
- többféle vásárló, többféle igényét kielégítő, heterogén (sokféle) termékek gyártásának irányába kell elmozdulni

Motiváció

A heterogenitás dilemmája:

- Megvan a sokféle termék, változatos kínálat 😊
- A vásárlónak rengeteg lehetőségből kell választani (információ túlterhelés) ☹️

A kínálatot szűrni/rangsorolni kell!

Ajánló rendszerek:

- A vásárlónak csak a releváns termékeket mutassuk 😊
- Egyénre/testreszabott vásárlás 😊

Webes vásárlásokat (vásárlói kosarak) könnyű nyomon követni 😊

Visszatérő vásárlók 😊

Ajánló rendszerek

← → ↻ 🏠 ⓘ localhost/recommendation/recommendations.php 🔍 ☆ ⬇️ 🌐 👤

Alkalmazások Daniel Bognar (@b... Darkfly Video (@da... GDPR Rico x Miss Mood -... Next Level ft. Quer... BLR x Rave & Crave... Dj Free & Purebeat... Kygo & Ellie Gouldi...

Moviee ©



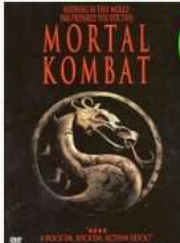
firstuser

- My profile
- Rating history
- Recommendations
- About
- Logout

Current location: **Veszprém, HU**

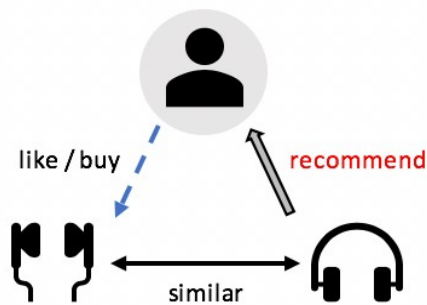
Do you want to change? [CHANGE LOCATION](#)

The recommended movies list in order of your predicted taste:

		
Hudson Hawk	Mortal Kombat	The Golden Compass
1991	1995	2007
★ 3.6/10	★ 4.4/10	★ 5.6/10
  	  	 

Definíció

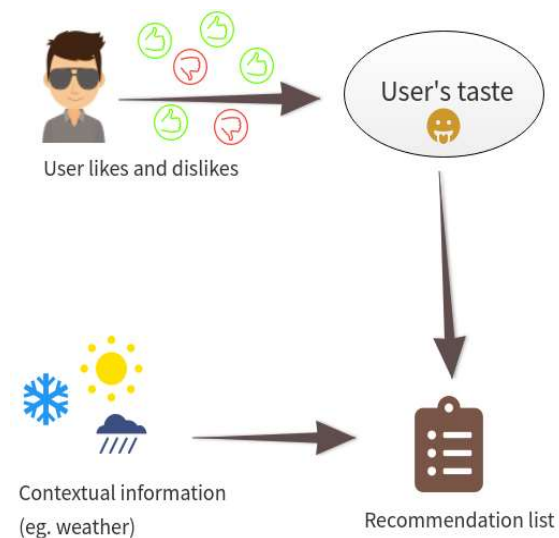
Az ajánlórendszerek speciális **információsűrő rendszerek**, amelyek olyan **tartalmat kínálnak a felhasználónak**, amely nagy valószínűséggel érdekes lesz a számára.



Tartalomalapú ajánlás

A		✓	✗	✓	✓
B			✓	✗	✗
C		✓	✓	✗	
D		✗		✓	
E		✓	✓	?	✗

Kollaboratív ajánlás



Kontextus függő ajánlás

Megközelítések

- **Kollaboratív ajánlás:** a predikció inputjához hasonló adatok is felhasználásra kerülnek a predikció során
 - user-based: a cél felhasználóhoz hasonló ízlésű felhasználók szavazatainak felhasználásával történik a predikció
 - item-based: a cél termékhez hasonló termékekre tett értékelések használatával készül a predikció
- **Tartalomalapú módszerek:** a célfelhasználó korábbi értékelései alapján történik a predikció
- **Kontextus függő:** figyelembe vesszünk valamilyen jellemzőt, amely csak egy adott kontextusban érdekes
- **Hibrid módszerek:** az előző módszerek valamilyen szintű ötvözése

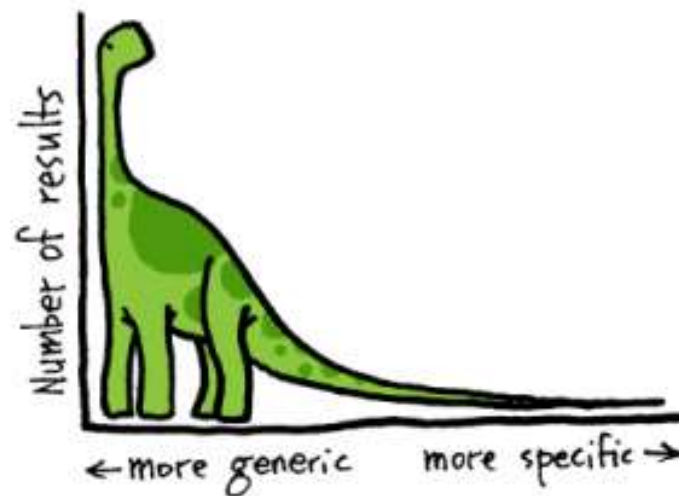
Ajánlórendszer megközelítések összehasonlítása

Ajánlórendszer típusa	Előny	Hátrány
Kollaboratív	Bőséges értékelésekkel és sok felhasználóval pontos ajánlás.	Hideg-indítás probléma az új felhasználók esetében (Hideg indítás problémáról akkor beszélünk, amikor egy új felhasználóról nincs informáciánk és nem tudja az algoritmus, hogy mit ajánljon neki).
Tartalomalapú	Nincs szükség a többi felhasználó értékeléseire az ajánlás elkészítéséhez.	Nincs váratlan termék az ajánlások között, túlságosan kiszámítható ajánlások.
Tudásalapú	Determinisztikus ajánlások, nincs hidegindítás.	A tudás direkt megszerzése túl kényelmetlen.
Hibrid	A vegyes alkalmazás miatt, elkerüli néhány hátrányát az egyes megközelítéseknek.	Komplex, bonyolult megvalósítási mód.
Demográfiai	Az ajánlások csak az adott régióra érvényesek, így elkerülve az érvénytelen ajánlásokat.	Kevés a kutatás még ebben a típusú megközelítésben.

Kiértékelés

Valós célfüggvények:

- Ügyfél elégedettsége!
 - Ajánlott elemek megvásárlása (likeolása)
 - Többet vásárolnak mintha nem lenne ajánlás?
-
- A Long Tail jelentkezésének klasszikus példája az Amazon.com sikere. A webáruház olyan **közösségi ajánlórendszer**t épített árukatalógusa és megrendelő oldalai köré, amely minden professzionális marketingnél hatékonyabban hívja fel a vásárlók figyelmét a számukra csábító termékekre.
 - A "hosszú farok" szerint **megfelelően hatékony, korszerű logisztikával a pár embert érdeklő termékeket is ugyanolyan hatásfokkal és haszonnal lehet értékesíteni** (tehát támogatni, forgalmazni, piacra dobni is), mint a hagyományosan sikerbiztos termékeket.



Számítanak-e a generációs különbségek?

- Általában véve: **NEM**
- Egyes területeken **IGEN**:
 - eszközhasználat
 - kommunikáció módja
 - megértés
 - ügyfélkezelés (reakciók, interakciók)
 - üzenet a felhasználónak (marketing)



Kívánalmak általában

A VEVŐ

- Olcsóbban (legolcsóbban)
- Gyorsan (gyorsabban)
- Könnyebben
- Megbízható eladótól
- Egyéb szempontok:
 - Fizetési mód
 - Ügyfélkezelés
 - Presztízs
-

A KERESKEDŐ

- Több vevő
- Gyorsabb kiszolgálás
- Nagyobb bevétel (-profit)
-

Ajánlási feladat

U - felhasználók (users),

I - termékek (items),

egy $f: U \times I \rightarrow \mathbf{R}$ leképezésből **néhány** példa
($\mathbf{R}$ halmaz, teljes rendezéssel).

- explicit értékelés: **milyen skálán?**
- implicit értékelés: **kattintások, oldallátogatások, eltöltött idő stb.**

Tanulási feladat:

határozzuk meg azt az

$f': U \times I \rightarrow \mathbf{R}$ leképezést, amely

- a lehető legjobban közelíti f -et és
- teljesen definiált a teljes $U \times I$ téren.

Ajánló rendszer:
$$i_u = \arg \max_{i \in I, f(i,u) \text{ definiálatlan}} f'(u, i)$$

1. Kollaboratív ajánlás

					
A		✓	✗	✓	✓
B			✓	✗	✗
C		✓	✓	✗	
D		✗		✓	
E		✓	✓	?	✗

Alapgondolat:

- Felhasználók (explicit vagy implicit) értékelik a tartalmakat
- Azok a felhasználók akik a múltban hasonló dolgokat kedveltek a jövőben is hasonlóan fognak viselkedni

Típusok:

Felhasználó-alapú legközelebbi szomszéd alapú ajánlás

Felhasználó-alapú legközelebbi szomszéd alapú ajánlás

	Item1	Item2	Item3	Item4	Item5
Alice	5	3	4	4	?
User1	3	1	2	3	3
User2	4	3	4	3	5
User3	3	3	1	5	4
User4	1	5	5	2	1

1. Válasszunk ki hasonló felhasználókat (peer), akik értékelték a kérdéses tartalmat
2. Becslésünk az értékelésre legyen a peerek értékelésének átlaga (aggregáció)

Felhasználó-alapú legközelebbi szomszéd alapú ajánlás

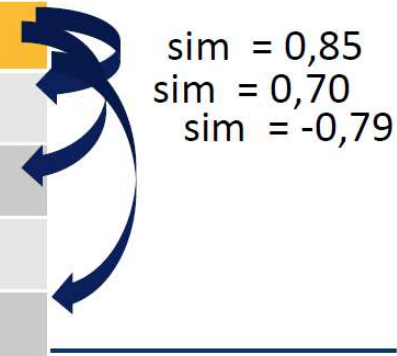
- Hogyan mérjük a hasonlóságot?
- Hány peert válasszunk?
- Átlagolás?

Hasonlósági mérték: korreláció

$$\frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

- két felhasználó hasonlósága
- csak azokon az elemeken ahol mindketten értékelték

	Item1	Item2	Item3	Item4	Item5
Alice	5	3	4	4	?
User1	3	1	2	3	3
User2	4	3	4	3	5
User3	3	3	1	5	4
User4	1	5	5	2	1



sim = 0,85
sim = 0,70
sim = -0,79

Felhasználó-alapú legközelebbi szomszéd alapú ajánlás

Becslés az értékelésre

$$pred(a, p) = \bar{r}_a + \frac{\sum_{b \in N} sim(a, b) * (r_{b,p} - \bar{r}_b)}{\sum_{b \in N} sim(a, b)}$$

- hasonlósággal súlyozunk
- a felhasználók saját átlagától való eltérésre normalizálunk (pesszimista vs. optimista felhasználók)

Felhasználó-alapú legközelebbi szomszéd alapú ajánlás

Problémák az értékeléssel

- Nem minden értékelés ugyanolyan súlyú
 - ahol mindenki egyetért az kevésbé számít, mint ahol vegyesek a vélemények
 - szórást vegyük figyelembe!
- Ha a közösen értékelt elemek száma alacsony (minden párra más az N) az nagyon torzíthat
 - súlyozzuk a hasonlósági metrikát a metszet elemszámával!
- Szomszédság meghatározása:
 - fix K , vagy küszöbérték

Elem-alapú legközelebbi szomszéd alapú ajánlás

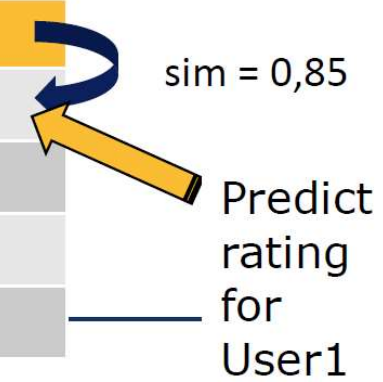
	Item1	Item2	Item3	Item4	Item5
Alice	5	3	4	4	?
User1	3	1	2	3	3
User2	4	3	4	3	5
User3	3	3	1	5	4
User4	1	5	5	2	1

Elem-alapú legközelebbi szomszéd alapú ajánlás

A „hidegindítás” probléma

- Ha érkezik egy új felhasználó vagy elem akkor ahhoz nincs értékelésünk. Hogyan találunk hasonló egyedeket?
 - „Kényszerítjük” az új felhasználót néhány ajánlásra vagy
 - Rekurzív kollaboratív ajánlás:

	Item1	Item2	Item3	Item4	Item5
Alice	5	3	4	4	?
User1	3	1	2	3	?
User2	4	3	4	3	5
User3	3	3	1	5	4
User4	1	5	5	2	1



The diagram illustrates the recommendation process for User1. A blue curved arrow points from the rating of Item5 for Alice (5) to the missing rating for User1 (indicated by a question mark). This arrow is labeled with $\text{sim} = 0,85$, representing the similarity between Alice and User1. A yellow arrow points from the rating of Item5 for User2 (5) to the same missing rating for User1. A blue line points from the text 'Predict rating for User1' to the missing rating cell.

Gráf-alapú ajánlás

- szomszédsági mátrix $\{a\}$
- felhasználó reprezentációs mátrix $\{cr\}$: „melyik user mennyire reprezentatív a másokra nézve”
- termék reprezentációs mátrix $\{pr\}$: „melyik termék mennyire reprezentatív egy user számára”

$$cr_{it} = \frac{\sum_{a_{tj}=1} pr_{ij}}{\sum_{a_{tj}=1} a_{tj}}$$

$$pr_{ik} = \frac{\sum_{a_{jk}=1} cr_{ij}}{\sum_{a_{jk}=1} a_{jk}}$$

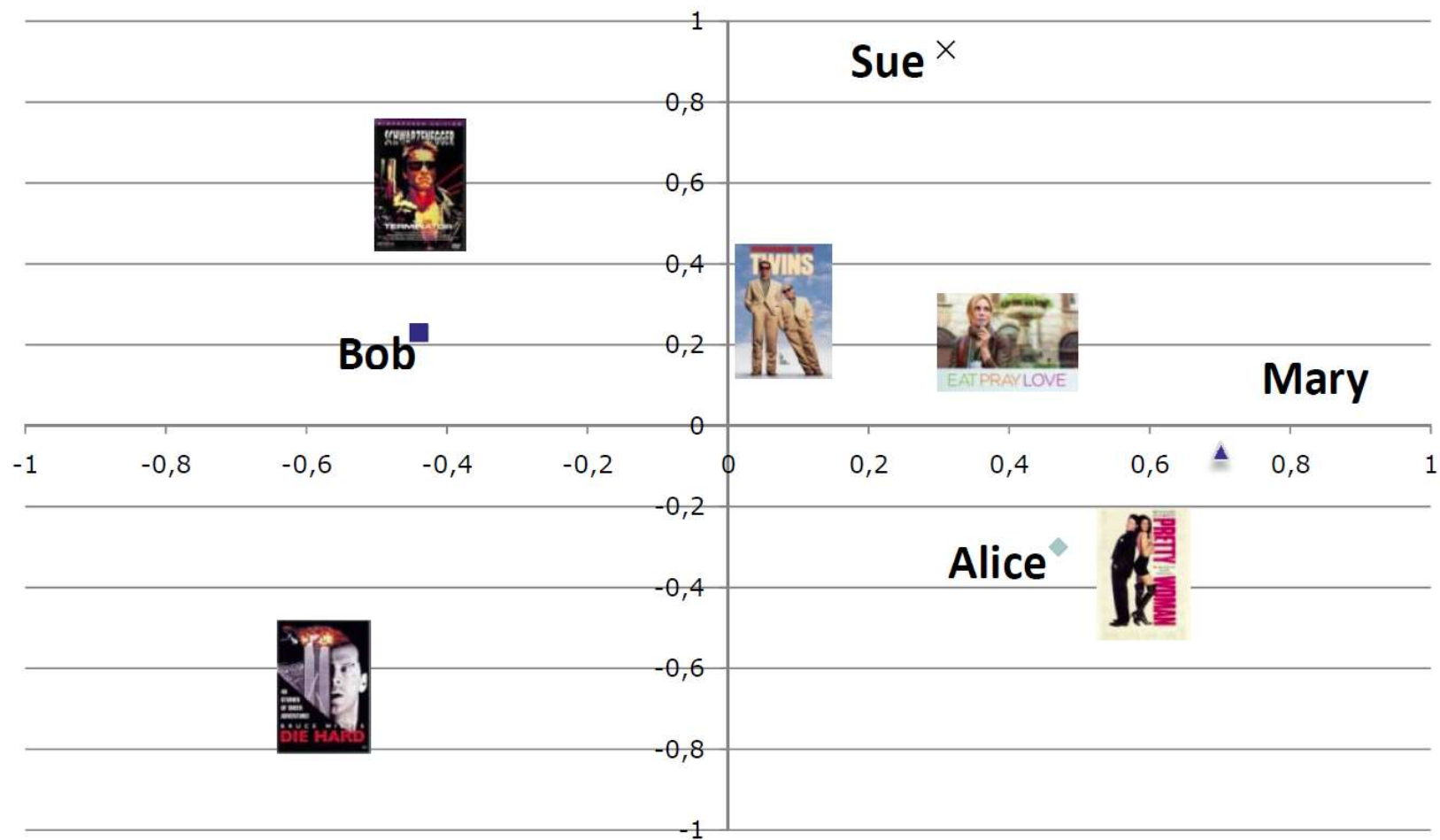
Regresszió alapú aggregáció

- felhasználók közti hasonlósági mérték (heurisztika) helyett
- minden egyes felhasználóhoz tanítsunk egy regressziós modellt, ami a többi felhasználó értékelései, mint jellemzők alapján jósol
- regresszor tanítása: olyan elemeken, ahol a célfelhasználónak ismertek az értékelései

	Item1	Item2	Item3	Item4	Item5
Alice	5	3	4	4	?
User1	3	1	2	3	3
User2	4	3	4	3	5
User3	3	3	1	5	4
User4	1	5	5	2	1

megjegyzés: nagyon kicsik az egyes tanító adatbázisok, ezért egyszerű modell (lineáris regresszió)

Dimenzió csökkentés alapú ajánlás



Dimenzió csökkentés alapú ajánlás

- SVD: $M_k = U_k \times \Sigma_k \times V_k^T$

U_k	Dim1	Dim2
Alice	0.47	-0.30
Bob	-0.44	0.23
Mary	0.70	-0.06
Sue	0.31	0.93

V_k^T					
Dim1	-0.44	-0.57	0.06	0.38	0.57
Dim2	0.58	-0.66	0.26	0.18	-0.36

Σ_k	Dim1	Dim2
Dim1	5.63	0
Dim2	0	3.23

- Prediction: $\hat{r}_{ui} = \bar{r}_u + U_k(Alice) \times \Sigma_k \times V_k^T(EPL)$
 $= 3 + 0.84 = 3.84$

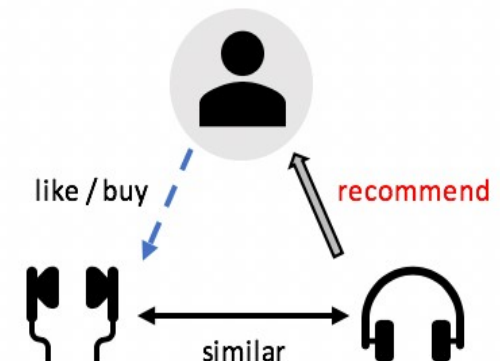
Kollaboratív ajánló rendszerek

- Előnyök:
 - Intuitív
 - Jól működik számos környezetben
 - Nem kell hozzá „jellemzőtér-tervezés”
- Hátrányok:
 - Sok felhasználó, sok értékelés kell
 - Ritka mátrix probléma...
 - Külső információt nehéz bevonni

Az egyes változatok teljesítményére különböző eredmények (nincs konszenzus)

2. Tartalom alapú ajánló rendszerek

- Kizárólag a célfelhasználó korábbi értékelései alapján ajánlunk
- Az elemeket jellemzőkkel írjuk le
pl: színészek, rendező, műfaj, főszereplő stb.
- A jellemzőkre egy osztályozót/regresszort építhetünk
- Tanítás a célfelhasználó korábbi értékelései alapján



Tartalom alapú ajánló rendszerek

Előnyök

- Független a többi felhasználótól (megjegyzés: jellemzőtérbe beépíthetjük az információt)
- Egyetlen felhasználó esetén is életképes a rendszer
- Hidegindítás problémája kisebb
- Ritkaság nem okoz gondot
- „Egyéniségek”-nél is működik
- Új vagy nem felkapott elemeket is ajánlhat (long tail)
- Értelmezhető modell (pl. legnagyobb súlyú tüzelő jellemzők)

Tartalom alapú ajánlás

Hátrányok

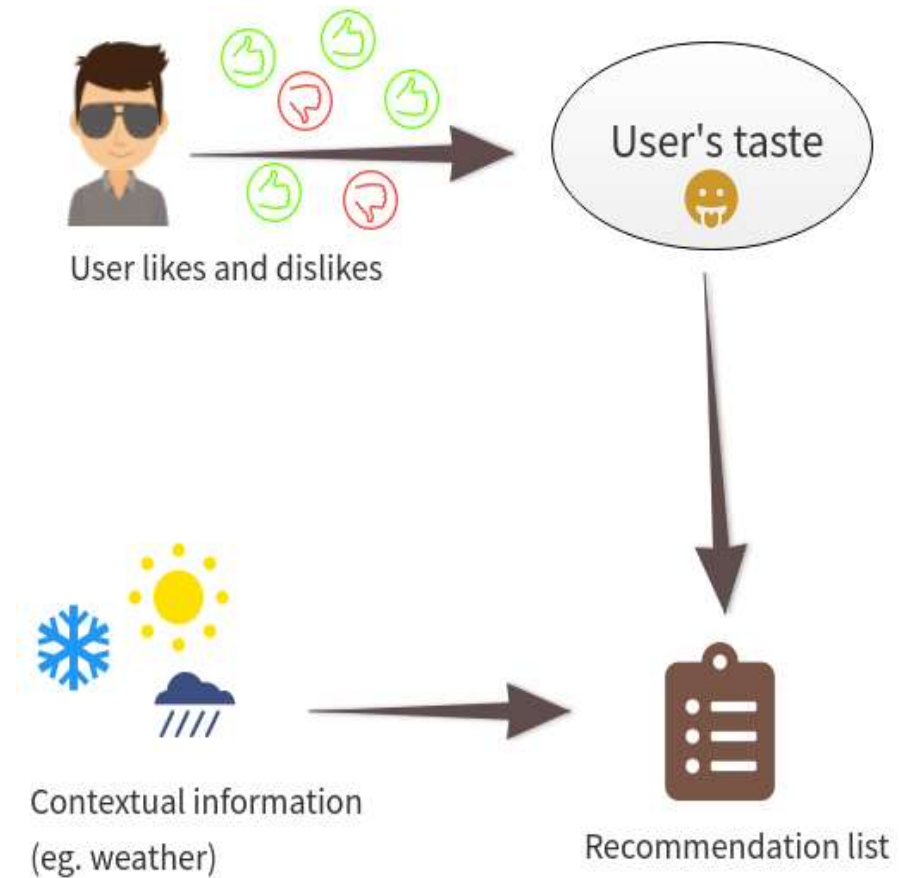
- Jellemzőtér kialakítása probléma specifikus és gyakran nehézségekbe ütközik
- Túl specializáció:
Elképzелhető, hogy nem mindig a már magasan értékeltekhez leginkább hasonló elem ajánlása az ideális
Pl. Étterem ajánló rendszer: amennyiben a felhasználó még csak kínai és magyar ételeket értékelt, soha nem fogunk neki görög éttermet ajánlani (még ha az a legjobb is a városban), hiszen túlságosan különböző az eddig (jónak) értékelt elemektől
- Egy új felhasználónak egy ideig nem tudunk értékelést adni (**megfelelő méretű tanító adatbázis kell**)

3. Hibrid rendszerek létrehozásának lehetőségei

- **tartalom alapú → kollaboratív**
 - „tanulható” felhasználóknál hiányzó értékelésekre becslés tartalom alapú módszerekkel
 - a becsült értéket (kis súllyal) felhasználjuk kollaboratív módszerekben
- **kollaboratív → tartalom alapú**
 - jellemzőket nyerünk ki a többi felhasználó és azok értékelései alapján
Pl: szakirodalom ajánlásnál hivatkozások alapján

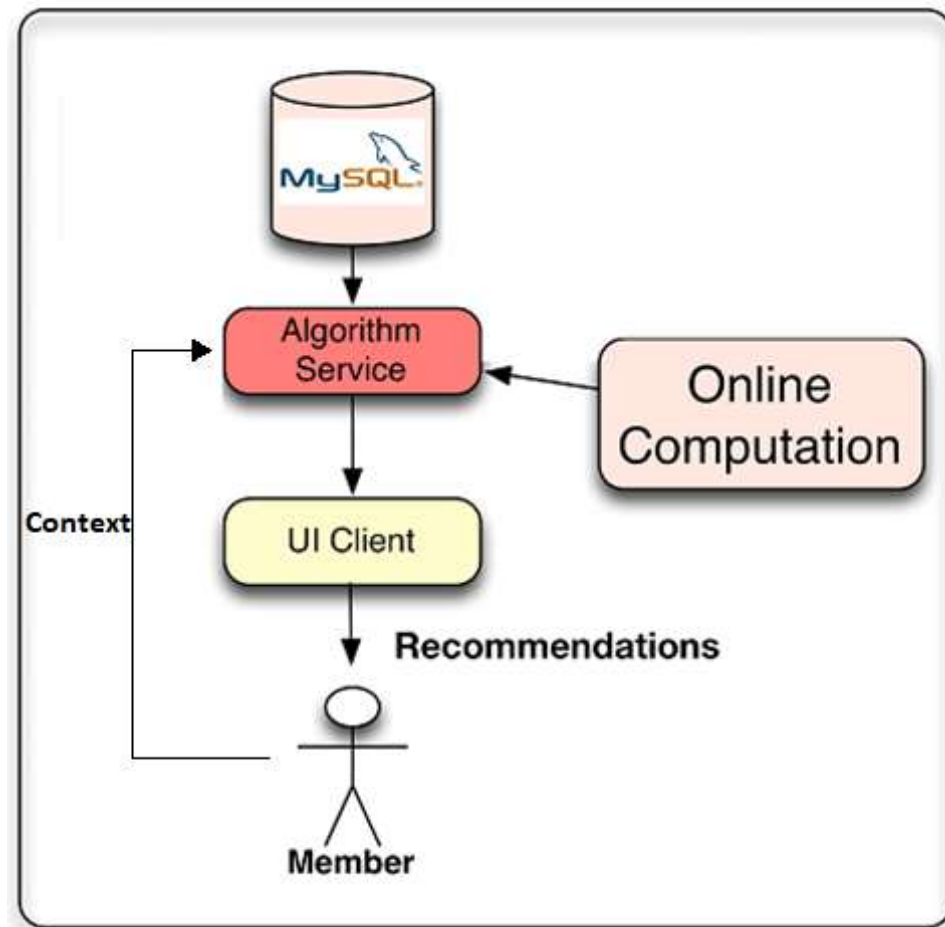
4. Kontextus függő ajánló rendszerek

- Utazási ajánlatok helyszínekre
- Télen és nyáron különböző
- Kontextus: évszak



Kontextus függő ajánló rendszerek

- Filmek ajánlása a felhasználó ízlése alapján, amely felhasználja a felhasználó lokációját
- A felhasználó lokációja körüli filmek ajánlási sorrendbe állítása a cél



Kontextus függő ajánló rendszerek

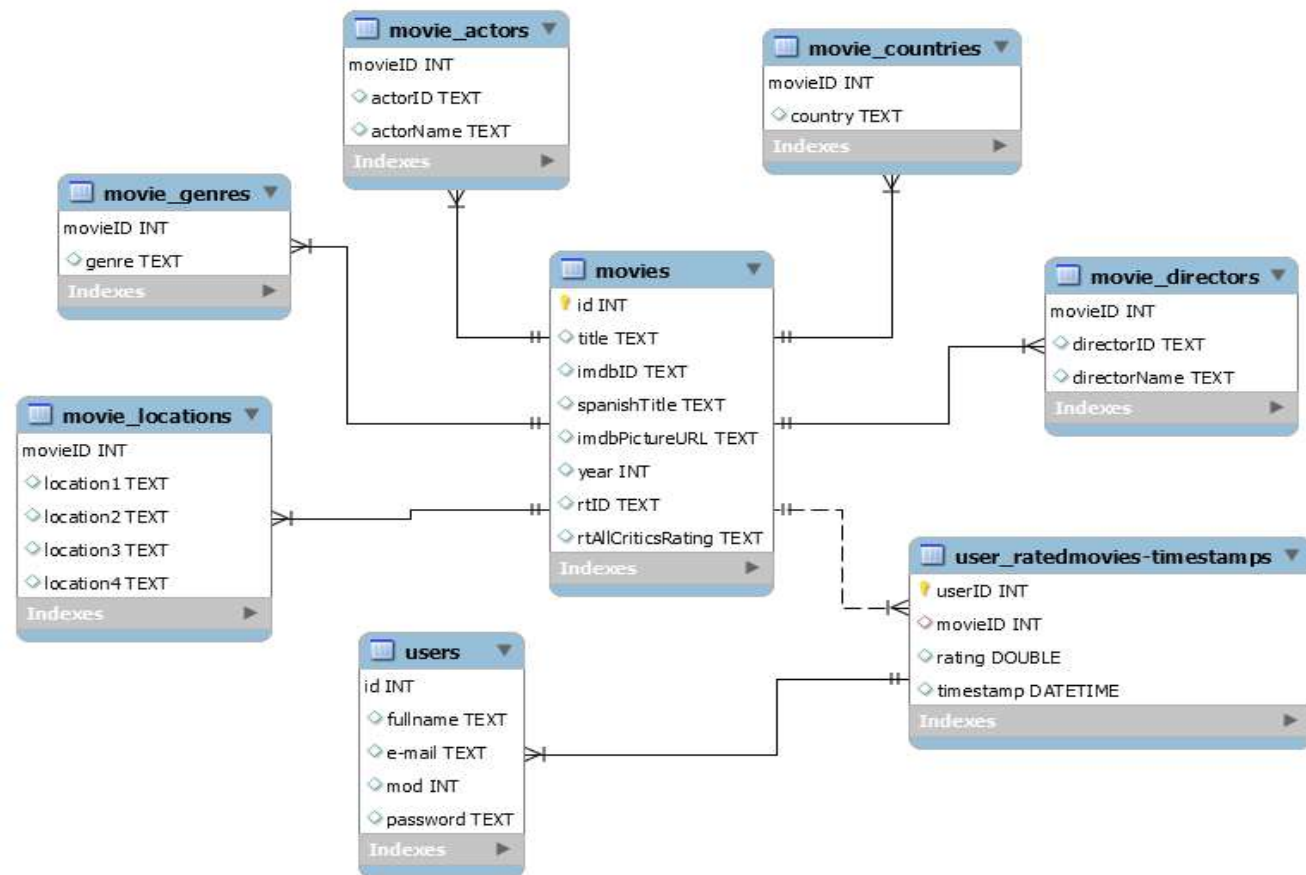
Adatbázis

MovieLens adatbázis:

- 2113 felhasználó
- 10197 film
- 20 műfaj
- 4060 rendező
- 95321 színész
- 72 ország
- 47899 helyszín
- **855598 értékelés**

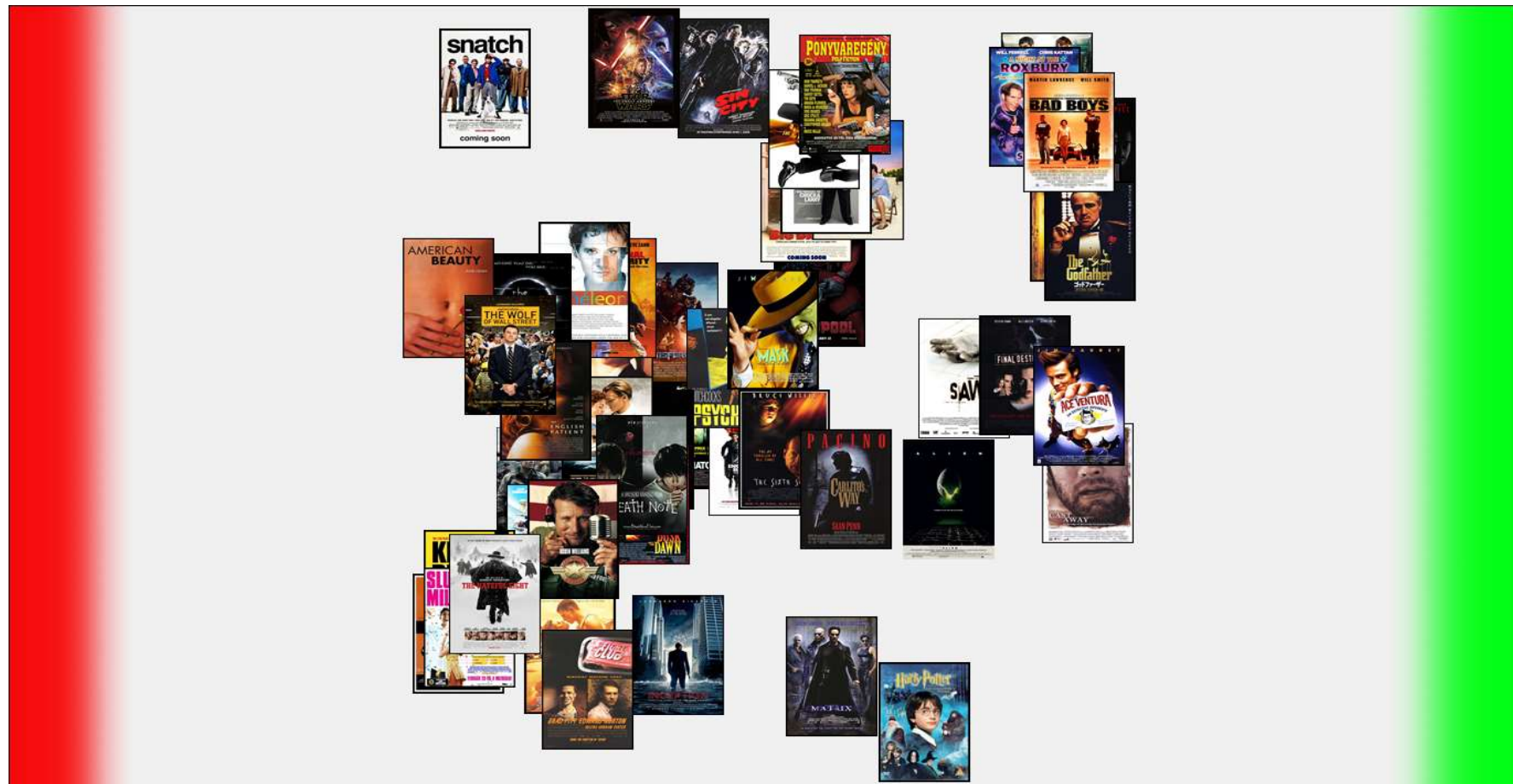
Filmek értékelése:

0.5 - 5



Kontextus függő ajánló rendszerek

Felhasználói interfész



Kontextus függő ajánló rendszerek

Felhasználói interfész

localhost/recommendation/recommendations.php

Alkalmazások Daniel Bogнар (@b... Darkfly Video (@da... GDPR Rico x Miss Mood -... Next Level ft. Querl... BLR x Rave & Crave... Dj Free & Purebeat... Kygo & Ellie Gouldi...

Moviee ©


firstuser

[My profile](#)

[Rating history](#)

Recommendations

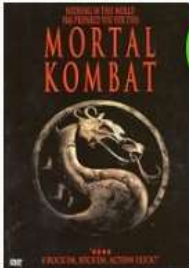
[About](#)

[Logout](#)

Current location: **Veszprém, HU**

Do you want to change? [CHANGE LOCATION](#)

The recommended movies list in order of your predicted taste:

		
Hudson Hawk	Mortal Kombat	The Golden Compass
1991	1995	2007
★ 3.6/10	★ 4.4/10	★ 5.6/10
 	 	 

My ratings

localhost/recommendation/myratings.php

AlkalmazásokDaniel Bogнар (@boDarkfly Video (@darAnswerMiner - PrediGDPRRico x Miss Mood - VNext Level ft. QueryBLR x Rave & CraveDj Free & Purebeat

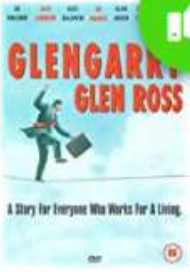
Moviee 🎬

My details

My rated movies

The recommendations

About



Glengarry Glen Ross

1992

★ 8.5/10



Strange Days

1995

★ 6/10



Stalker

1979

★ 8.2/10



Crouching Tiger, Hidden Dragon



One Million Years B.C.



Repo Man