New Insights in Preference Elicitation for Recommender Systems

Jill-Jênn Vie

iSWAG Symposium 2016

June 10, 2016

Death Note

Dog Days

Princesse Mononoké

The Place Promised in Our Early Days

- 2k users
- 14k works (anime & manga)
- 281k ratings (like / dislike / neutral / willsee / wontsee)
- October 2015: Student Demo Cup winner, Microsoft prize
- February 2016: Japanese Culture Embassy Prize, Paris

Collaborative Filtering

Problem

- Users $u = 1, \ldots, n$ and items $i = 1, \ldots, m$
- Every user u rates some items $\mathcal{R}_u$
  (r_{ui}: rating of user u on item i)
- How to guess unknown ratings?

k-nearest neighbor

- Similarity score between users:

$$score(u, v) = \frac{\mathcal{R}_u \cdot \mathcal{R}_v}{||\mathcal{R}_u|| \cdot ||\mathcal{R}_v||}.$$

- Let's find user's k nearest neighbors
- And recommend what they liked that he didn't rate

Objects: n vectors over m dimensions

Preference Elicitation

Problem

What questions to ask adaptively to a new user?

4 decks

- Popularity
- Controversy (Reddit)

$$controversy(L, D) = (L + D)^{min(L/D, D/L)}$$

- Most liked
- Precious pearls: few rates but almost no dislike

Problem: most people can't rate the controversial items

Matrix Completion

Let us assume that M has low rank r :

$$M = \begin{pmatrix} \mathcal{R}_1 \\ \mathcal{R}_2 \\ \vdots \\ \mathcal{R}_n \end{pmatrix} = \boxed{} = \boxed{C}\ \boxed{\qquad P \qquad}$$

Every line $\mathcal{R}_u$ is a linear combination of lines of P.

$$M : n \times m \qquad C : n \times r \qquad P : r \times m.$$
$$\mathcal{R}_1 = c_{11}P_1 + c_{12}P_2 + \ldots + c_{1r}P_r \qquad C_1 = (c_{11}, c_{12}, \ldots, c_{1r})$$

Example

If P	P_1 : adventure	P_2 : romance	P_3 : plot twist
And C_u	0,2	$-0,5$	0,6

it means :

u likes a bit adventure, dislikes romance,
really likes plot twists.

First Eigenvector's Top 30

Nausicaä of the Valley of the Wind
Princesse Mononoké
Le Château dans le ciel
Le Voyage de Chihiro
Toki wo Kakeru Shoujo
Tengen Toppa Gurren Lagann
Baccano!
Cowboy Bebop
Les Enfants Loups : Ame & Yuki
Mahou Shoujo Madoka Magica
Suzumiya Haruhi no Yuuutsu
Porco Rosso
Summer Wars
Neon Genesis Evangelion
Mon voisin Totoro

Ghost in the Shell
Kiki la petite sorcière
Suzumiya Haruhi no Shoushitsu
Le Château ambulant
Paprika
The Garden of Words
Barakamon
Steins;Gate
5 centimètres par seconde
Grave of the Fireflies
The Tale of The Princess Kaguya
Akira
Mushishi
Bakemonogatari
Durarara!!

First Eigenvector's Bottom 30

Zero no Tsukaima
To LOVE-Ru
Soul Eater
D.Gray-man
Another
Bleach
Rosario to Vampire Capu2
Vampire Knight
High School DxD
Naruto
Black Butler
Dragon Ball GT
Guilty Crown
Akame ga Kill!
Naruto the Movie 2: Legend of
the Stone of Gelel

Mirai Nikki
Tokyo Ghoul
Rosario to Vampire
L'Attaque des Titans
IS: Infinite Stratos
Fairy Tail
Sword Art Online II
Ao no Exorcist
One Piece
Highschool of the Dead
Sword Art Online
Bleach
Naruto
Fairy Tail
Naruto: Shippuuden

Second Eigenvector's Top 30

L'Attaque des Titans
Fullmetal Alchemist: Brotherhood
Death Note
Fullmetal Alchemist
Sword Art Online
Le Voyage de Chihiro
Princesse Mononoké
Ao no Exorcist
No Game No Life
Tokyo Ghoul
Mon voisin Totoro
FullMetal Alchemist
Psycho-Pass
Attaque Des Titans (l')
Code Geass: Hangyaku no Lelouch

Naruto
Fate/Zero
Les Enfants Loups : Ame & Yuki
Hunter x Hunter
Fullmetal Alchemist: Brotherhood OVA Collection
Mirai Nikki
Death note
Steins;Gate
Soul Eater
One Piece
Le Château ambulant
Le Château dans le ciel
Bleach
Durarara!!
Tokyo ghoul

Second Eigenvector's Bottom 30

Infinite Stratos 2
IS: Infinite Stratos
Ikkitousen: Dragon Destiny
The Severing Crime Edge
IS: Infinite Stratos Encore - Koi ni Kogareru ...
A Bridge to the Starry Skies
Sailor Moon R
Ikki Tousen
Vividred Operation
School Days: Magical Heart
Kokoro-chan
Papa to Kiss in the Dark
D.C. Da Capo
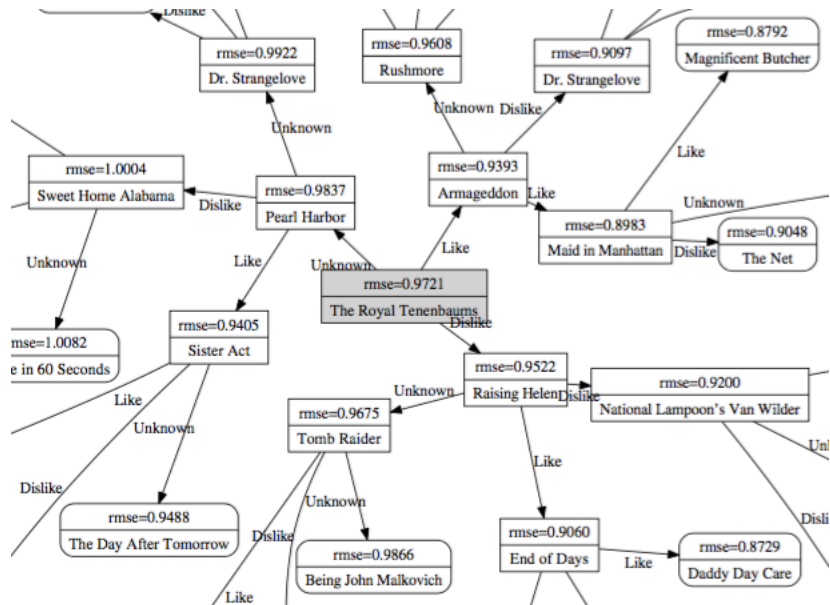Rail Wars!
Strawberry Panic
Freezing

To LOVE-Ru
School Days
Tokyo Mew Mew
Haruka Nogizaka's Secret
R-15
Wizard Barristers
Choujigen Game Neptune: The Animation
Yu-Gi-Oh! GX
Dragon Ball GT
Captain Earth
Astarotte's Toy
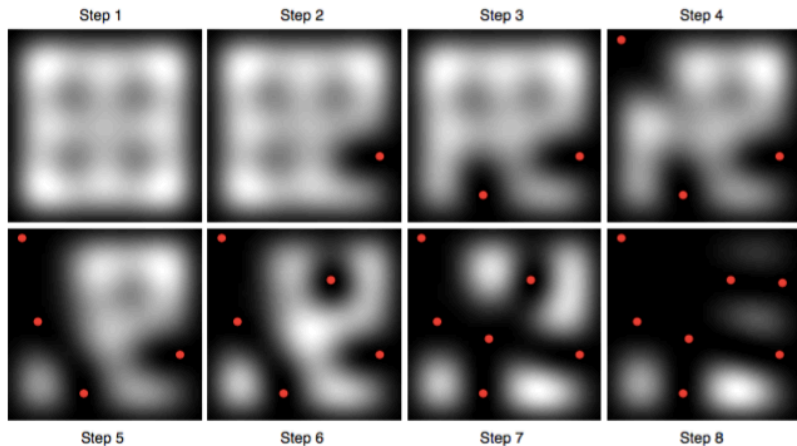Sakura Trick
Girls Bravo: First Season
Kiss x Sis
Dog Days
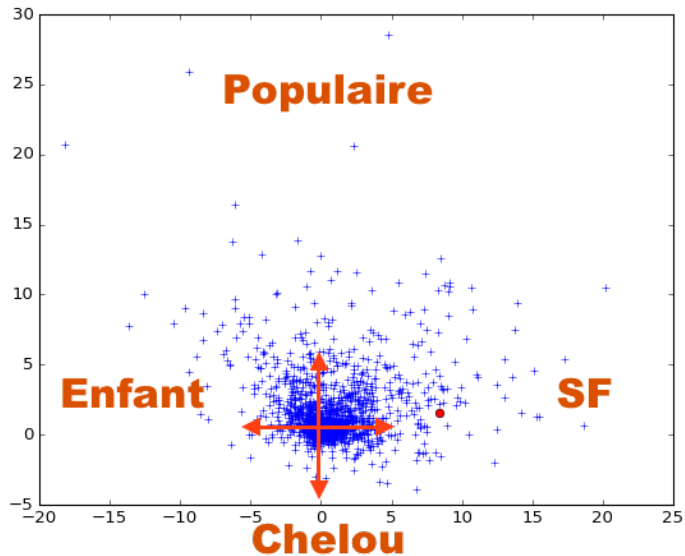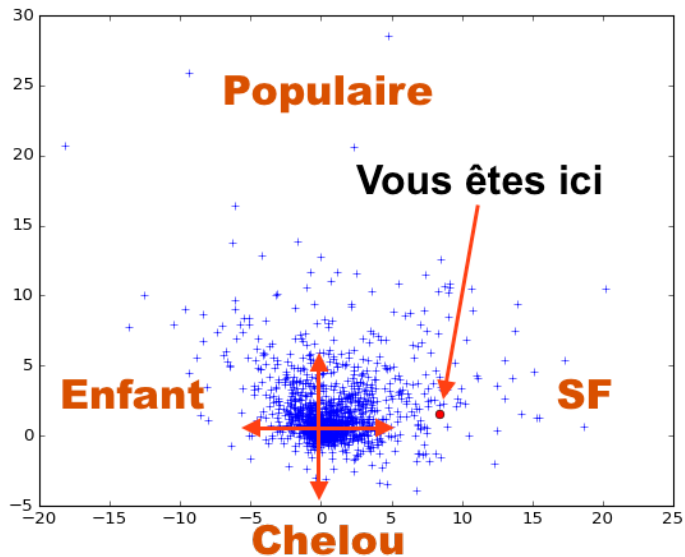
Yahoo's Bootstrapping Decision Trees

Determinantal Point Processes

We want to sample over *n* items
$K : n \times n$ similarity matrix over items
(semidefinite positive matrix)
P is a determinantal point process if *Y* is drawn with property:

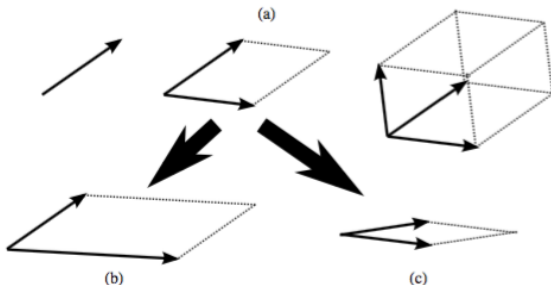$$\forall A \subset Y, \quad P(A \subseteq Y) \propto det(K_A)$$

Example

$$K = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 5 & 6 & 7 \\ 3 & 6 & 8 & 9 \\ 4 & 7 & 9 & 1 \end{pmatrix}$$

$A = \{1, 2, 4\}$ will be included
with probability prop. to

$$K_A = det \begin{pmatrix} 1 & 2 & 4 \\ 2 & 5 & 7 \\ 4 & 7 & 1 \end{pmatrix}$$

- The determinant is the volume of the vectors
- Non-correlated (diverse) vectors will increase the volume
- Algorithm samples with complexity $O(nk^3)$
  where k is the number of points that will be sampled.

MANGAKI

LA RECOMMANDATION FRANCOPHONE D'ANIME ET MANGA

- research.mangaki.fr

- jj@mangaki.fr

Fork us
on GitHub!