Recommender Systems

from Preference Learning to Useful Recommendations

Francesco Ricci

Free University of Bozen-Bolzano

fricci@unibz.it

unibz

Fakultät für Informatik

Facoltà di Scienze e Tecnologie informatiche

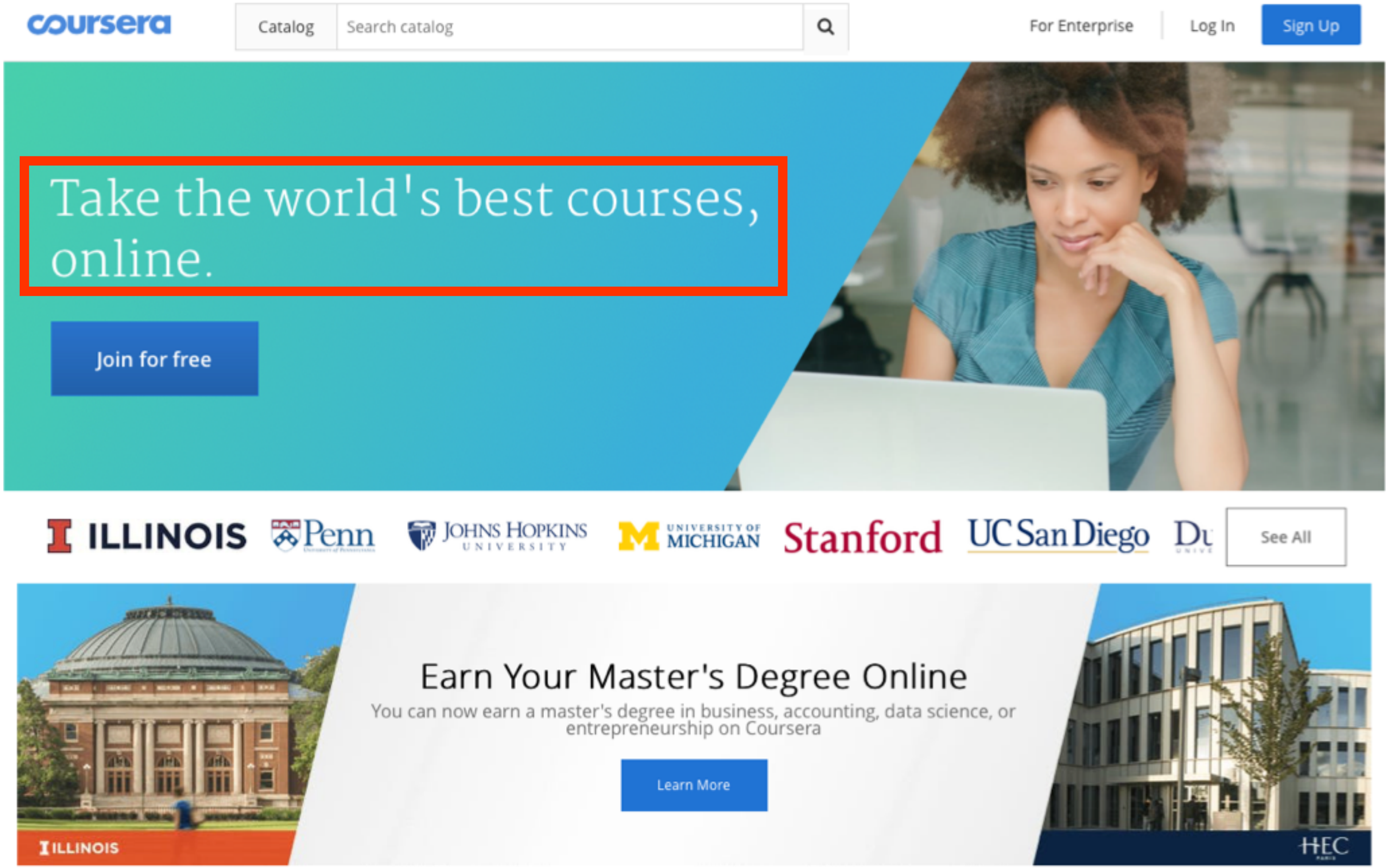
Faculty of Computer Science

IRCDL

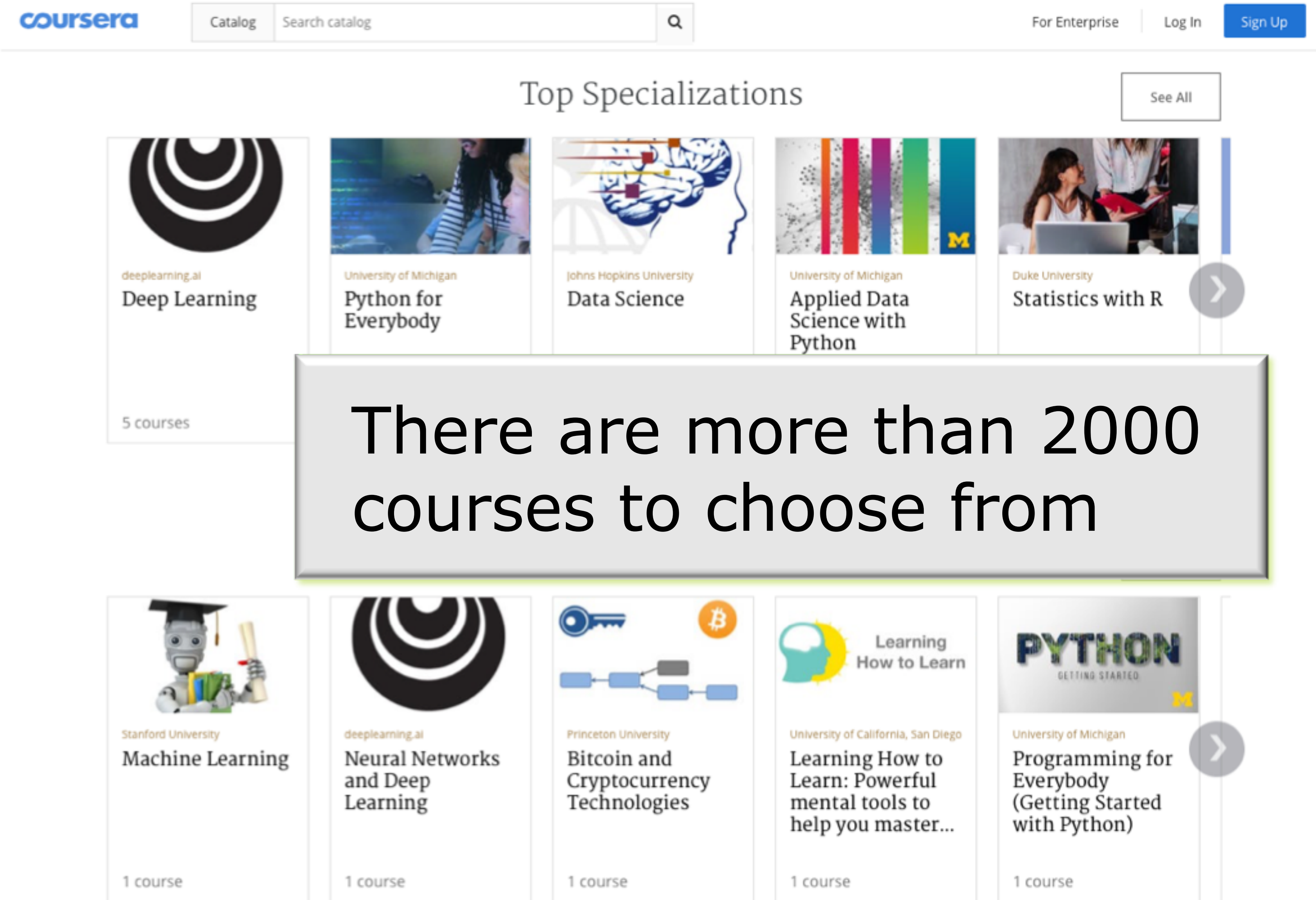
Content

- Motivations – information filtering
- Recommender system
- Critical Assumptions of RSs
- Preference modeling
- Choice modeling
- Data dynamics
- Revisiting Recommendations

Coursera


coursera
Catalog
Search catalog
For Enterprise
Log In
Sign Up
Take the world's best courses, online.
Join for free
ILLINOIS
Penn
JOHNS HOPKINS UNIVERSITY
UNIVERSITY OF MICHIGAN
Stanford
UC San Diego
See All
Earn Your Master's Degree Online
You can now earn a master's degree in business, accounting, data science, or entrepreneurship on Coursera
Learn More
ILLINOIS
HEC

What course?



There are more than 2000 courses to choose from

Choice and Well-Being

- We have **more choice**, more freedom, autonomy, and self determination
- Increased choice **should improve well-being:**
  - *added options can only make us better off: those who care will benefit, and those who do not care can always ignore the added options*
- Various assessment of well-being have shown that **increased affluence** have accompanied by **decreased well-being**.

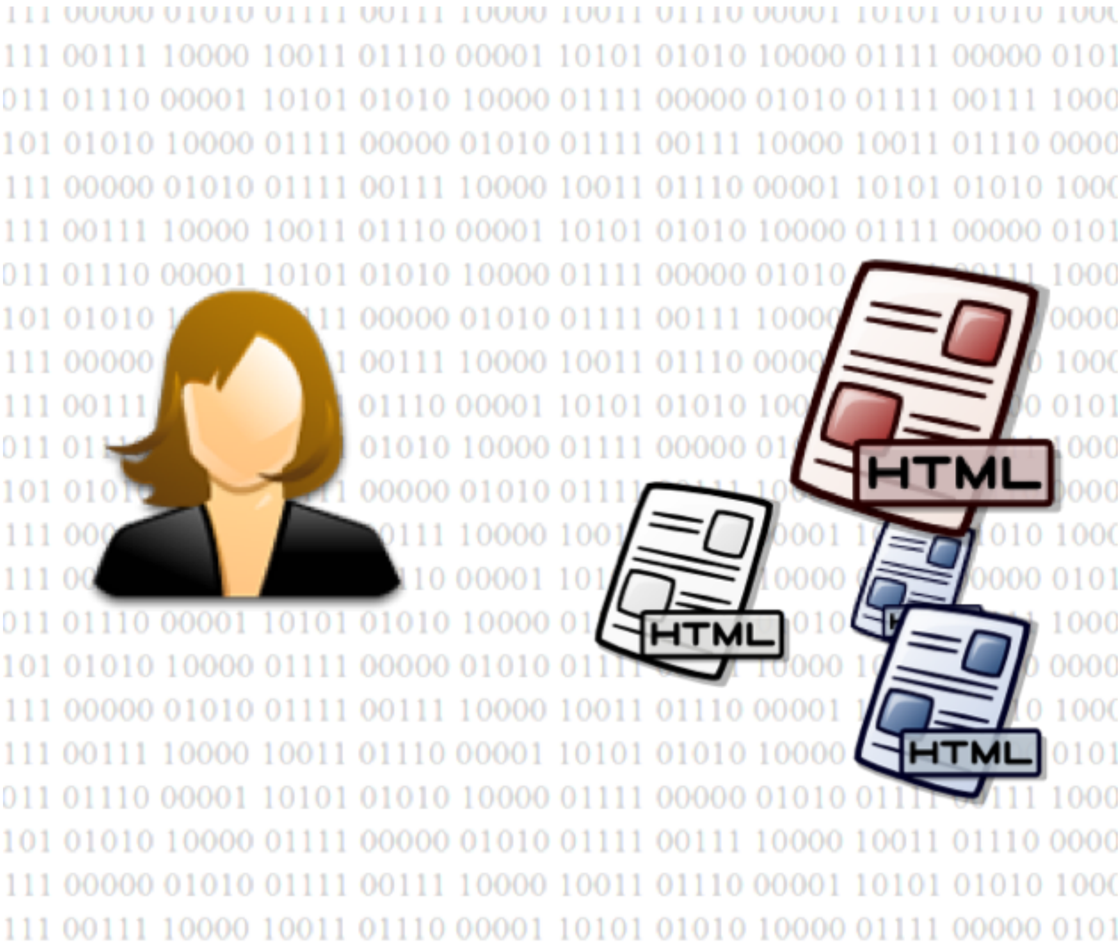
B. Schwartz, “The Paradox of Choice”, Harper Perennial, 2004.

Successful Queries are the Minority

Engines:	Total(%)	Organic(%)	PPC(%)
Google	19.41%	18.67%	0.74%
Yahoo	29.59%	29.59%	0.00%
Bing	15.06%	15.06%	0.00%
AOL	40.48%	40.48%	0.00%
ASK	50.98%	50.98%	0.00%
Other	50.66%	50.66%	0.00%

Source: http://www.keyworddiscovery.com/

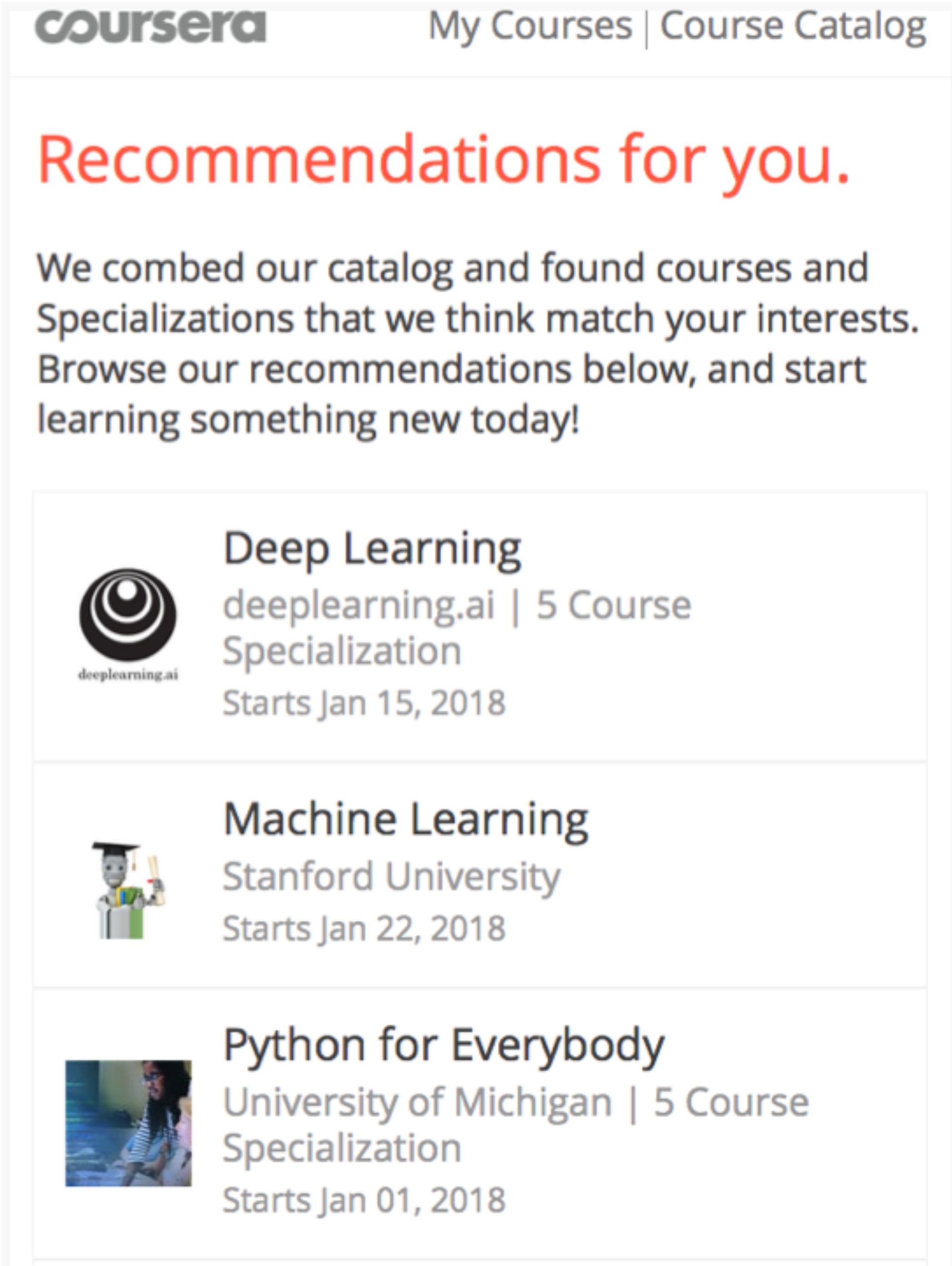
Queries will disappear



Leverage multiple signals to get rid of queries

Recommender Systems

Coursera Recommendations



Movie Recommendation – YouTube

Recommendations account for about 60% of all video clicks from the home page.

1. Preference Elicitation

2. Preference prediction

3. Selecting and presenting the recommendations

Classical Recommendation Model

Two types of entities: **Users and Items**

1. A **background knowledge**:
   - A set of **ratings – preferences** - is a map
     - *r: Users x Items → [0,1] U {?}*
   - A set of "features" of the Users and/or Items
2. A method for **predicting** the preference function *r* on (user, item) pairs where it is unknown

$$r^*(u, i) = \text{Average}_{u' \text{ is similar to } u} \{r(u', i)\}$$

3. A method for **selecting** the items to recommend (choice):
   - Recommend to u the item $i^* = \arg\max_{i \in Items} \{r^*(u,i)\}$

G. Adomavicius, A. Tuzhilin: Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. IEEE Trans. Knowl. Data Eng. 17(6): 734-749 (2005)

Movie rating data

Training data

user	movie	date	score
1	21	5/7/02	1
1	213	8/2/04	5
2	345	3/6/01	4
2	123	5/1/05	4
2	768	7/15/02	3
3	76	1/22/01	5
4	45	8/3/00	4
5	568	9/10/05	1
5	342	3/5/03	2
5	234	12/28/00	2
6	76	8/11/02	5
6	56	6/15/03	4

Test data

user	movie	date	score
1	62	1/6/05	?
1	96	9/13/04	?
2	7	8/18/05	?
2	3	11/22/05	?
3	47	6/13/02	?
3	15	8/12/01	?
4	41	9/1/00	?
4	28	8/27/05	?
5	93	4/4/05	?
5	74	7/16/03	?
6	69	2/14/04	?
6	83	10/3/03	?

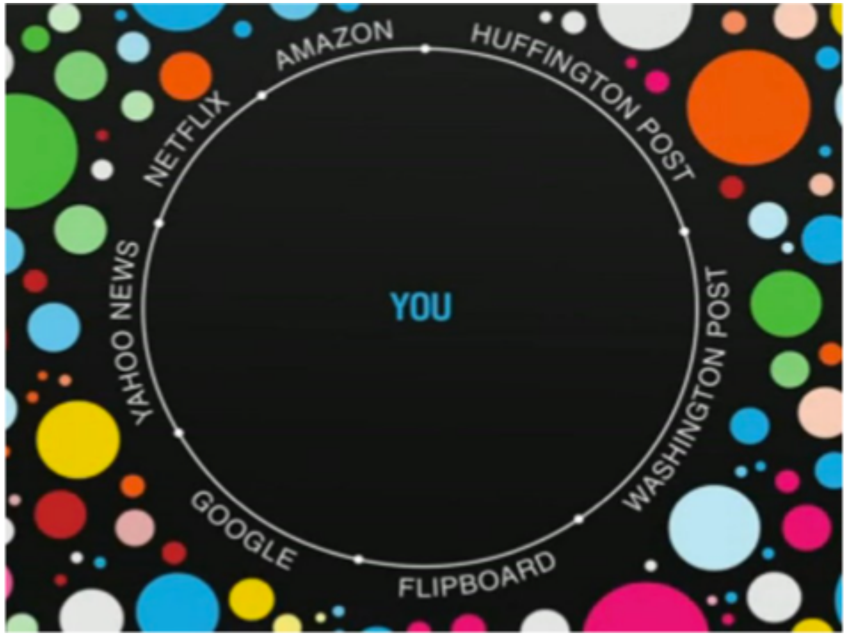
RecSys 2017 Facts and Numbers

- **627** Attendees
- **43** Countries
- **247** Submissions
- **23** Sessions
- **46** Scientific papers
- **12** Industry papers
- **14** Sponsors

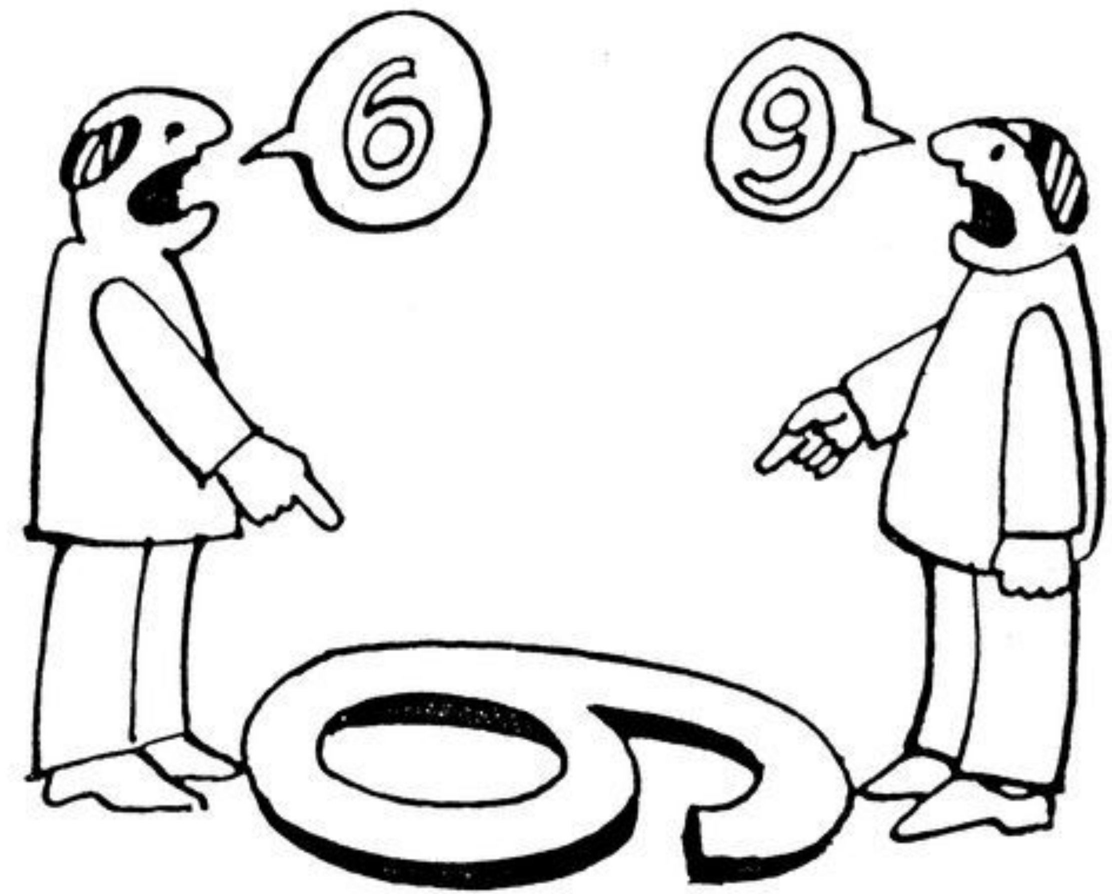
Eleventh ACM Conference on Recommender Systems, RecSys 2017, Como, Italy, August 27-31, 2017

Problems and Issues

- Cold Start (new user and new item)
- Filter Bubble
- How much to personalize
- How to contextualize
- Learning to interact and proactivity
- Recommendations for Groups
- Scalability and big data
- Privacy and security
- Diversity and serendipity
- Stream based recommendations

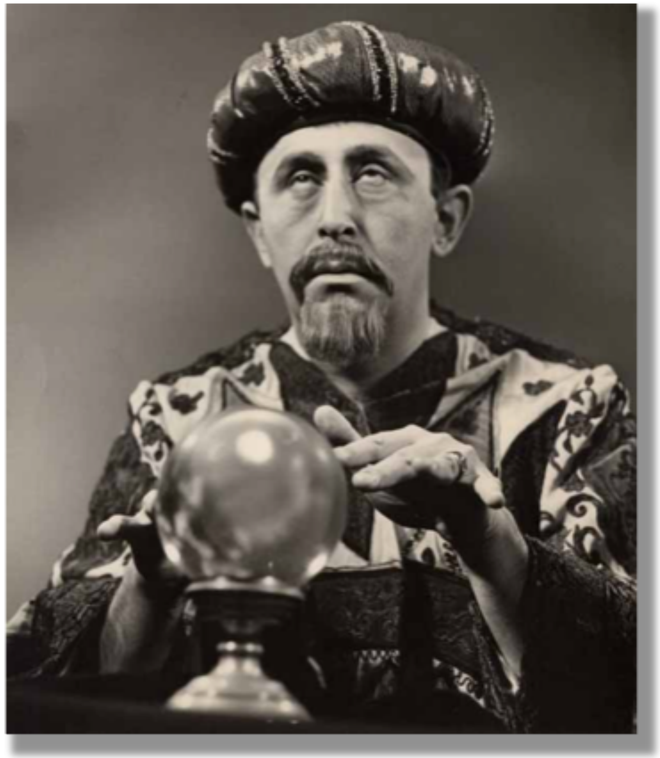


Critical Assumptions



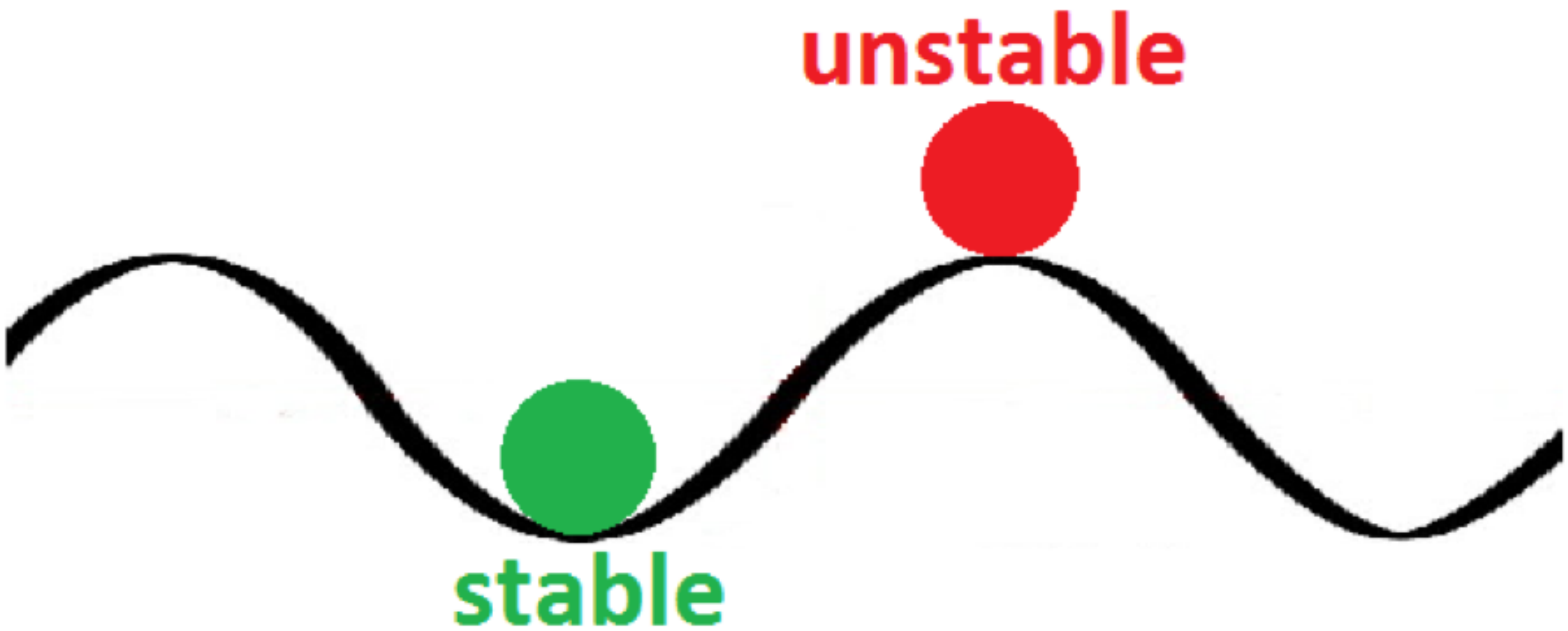
Predictability

- **Predictability:** by observing the user's expressed preferences or behavior (choices) the recommender can build a concise algorithmic model of what the user prefers or chooses



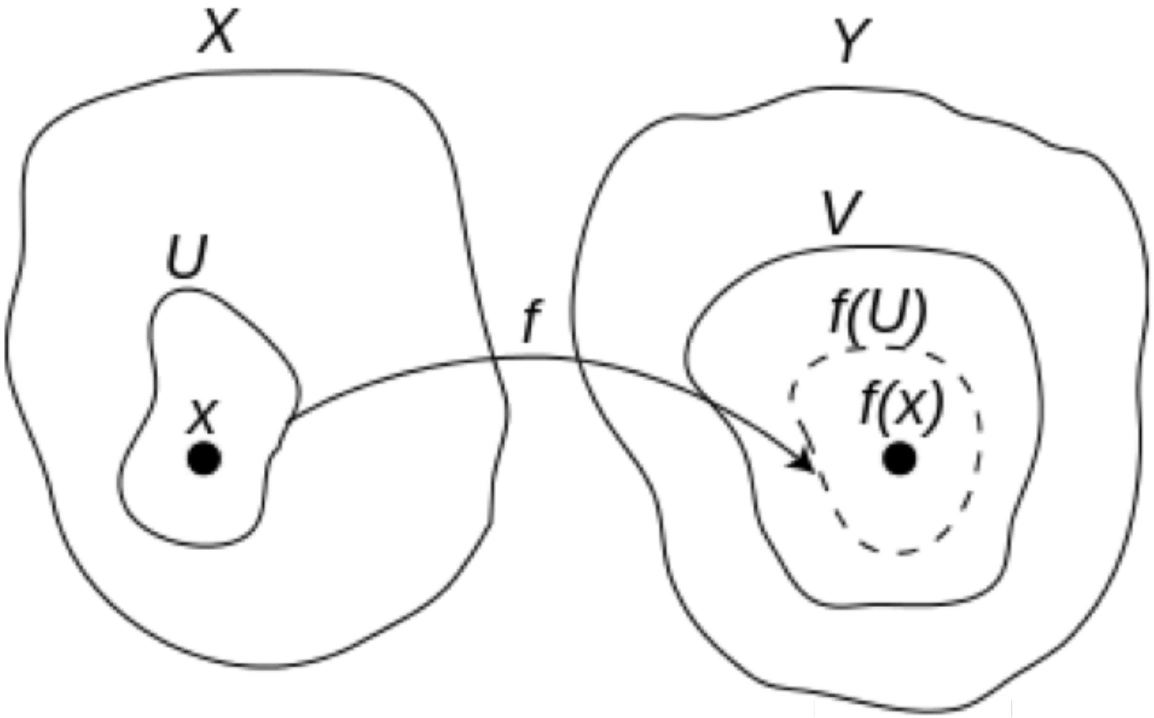
Stability of User Preferences

- User **preferences** are supposed to be rather **stable –** models are built by using historical data



Continuity

- User preference function is "**continuous**": there exist a notion of item-to-item similarity such that similar items generate similar reactions in a user




FALSE

Violation of predictability, stability and continuity assumptions

- *Today I read **drama** while last month I was preferably reading **adventure***
- *The user is rating high Fellini's movies but is often watching Star Wars*
- *I like Pustertal but I do not like Vinshgau*

Pustertal

Vinshgau

Issues

- We have excessively simplified the user preference and choice models
  - *We need more sophisticated models*
- We blindly rely on the observed data
  - *Some data should be ignored*
- Preferences and models are dynamic and also the recommender influences the observed data.

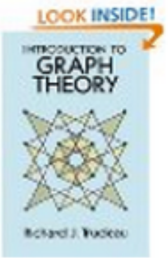
Preferences

Ratings (recommendations)

amazon.com

Help | Close window

Recommended for You



Introduction to Graph Theory (Dover Books on Mathematics)
by Richard J. Trudeau (February 9, 1994)
In Stock
List Price: ~~$14.95~~
Price: $3.99
59 used & new from $3.26

Add to Cart | Add to Wish List

Rate this item
I own it
Not interested

Because you purchased...

Patterns of Software: Tales from the Software Community (Hardcover)
by Richard P. Gabriel (Author)

This was a gift
Don't use for recommendations

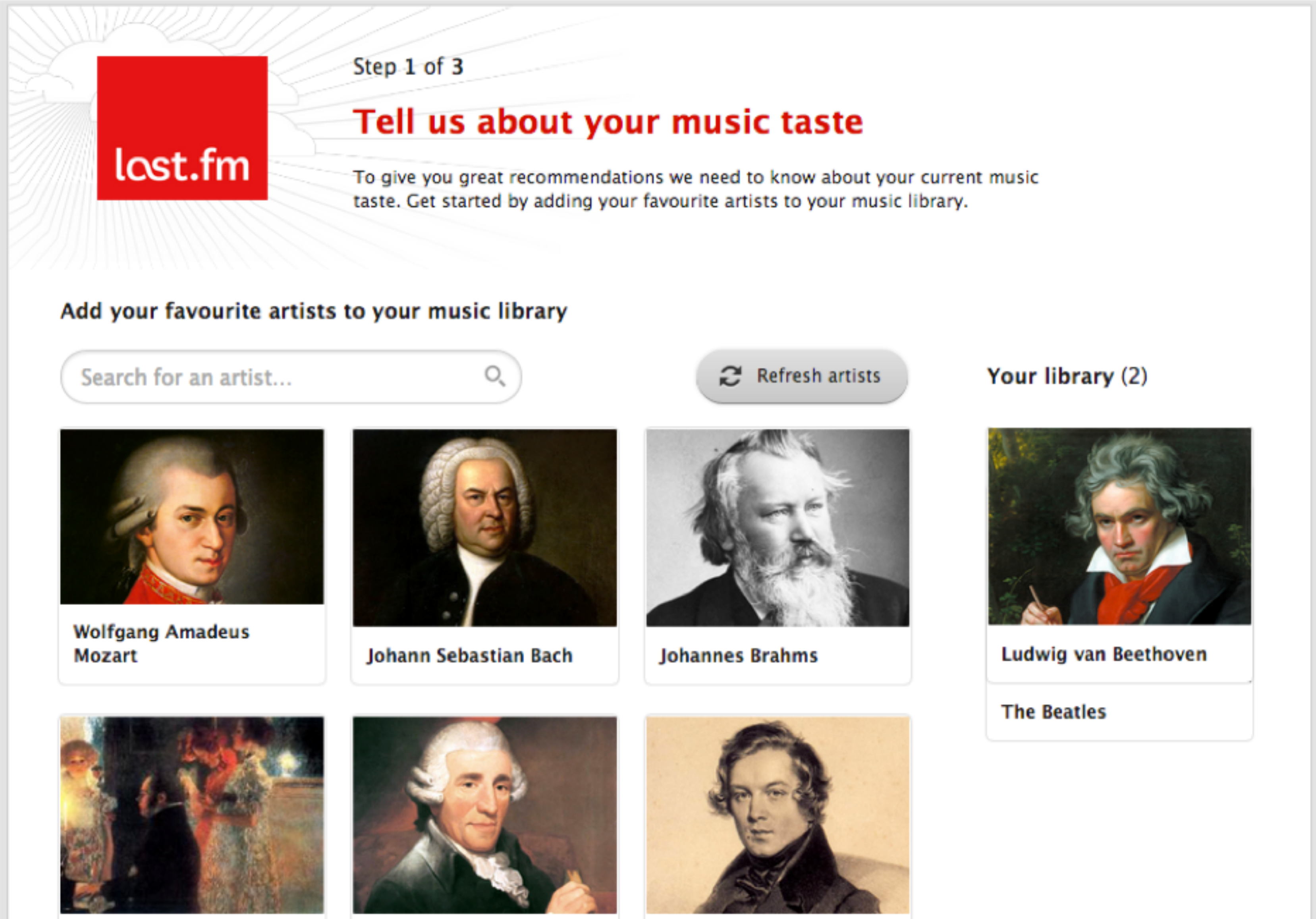
Machine Learning (Hardcover)
by Tom M. Mitchell (Author)

This was a gift
Don't use for recommendations

Reinforcement Learning: An Introduction (Adaptive Computation and Machine Learning) (Hardcover)
by Richard S. Sutton (Author), Andrew G. Barto (Author)

This was a gift
Don't use for recommendations

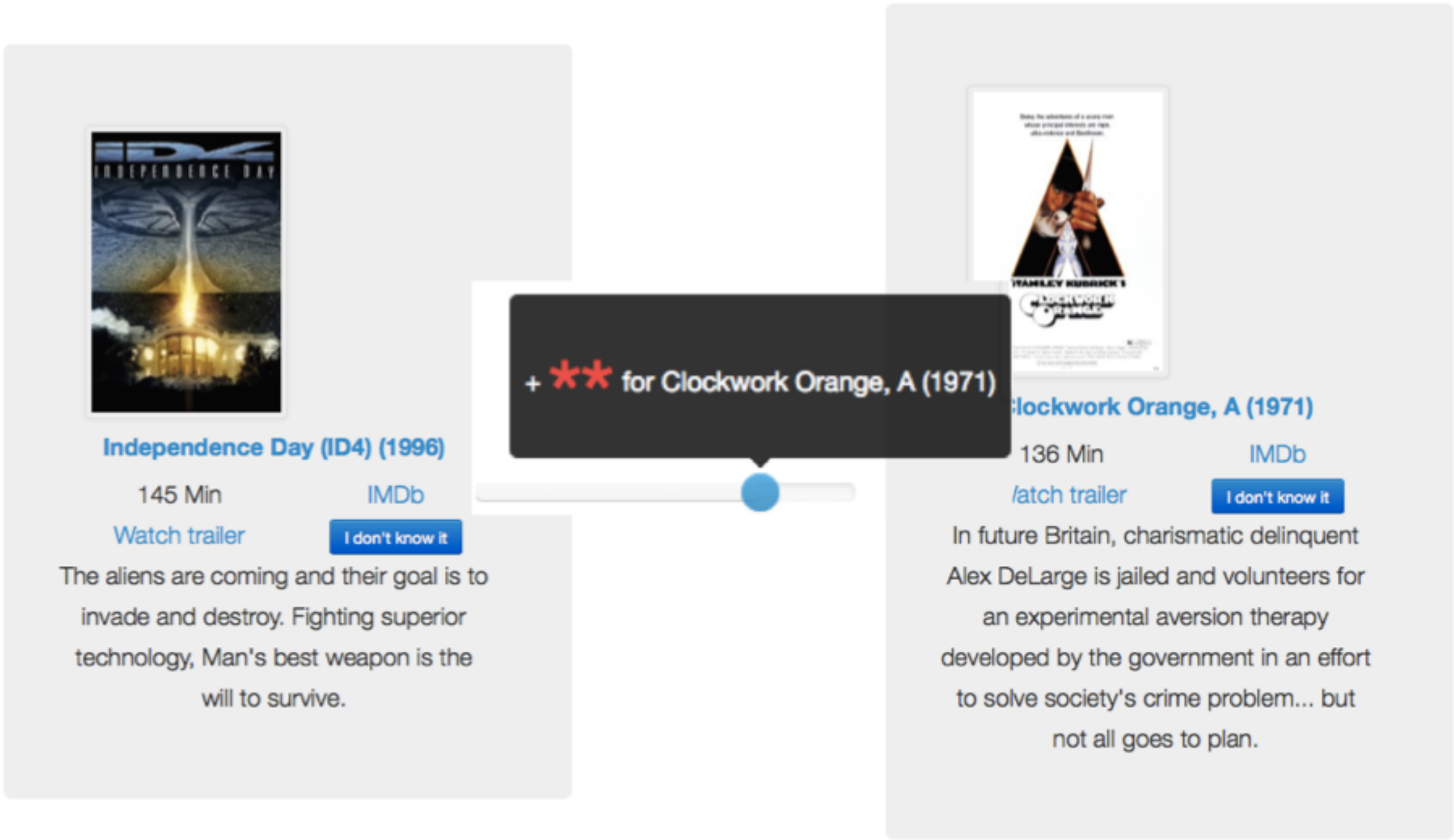
Likes



last.fm
Step 1 of 3
Tell us about your music taste
To give you great recommendations we need to know about your current music taste. Get started by adding your favourite artists to your music library.
Add your favourite artists to your music library
Search for an artist...
Refresh artists
Your library (2)
Wolfgang Amadeus Mozart
Johann Sebastian Bach
Johannes Brahms
Ludwig van Beethoven
The Beatles

Likes

Pairwise Preferences



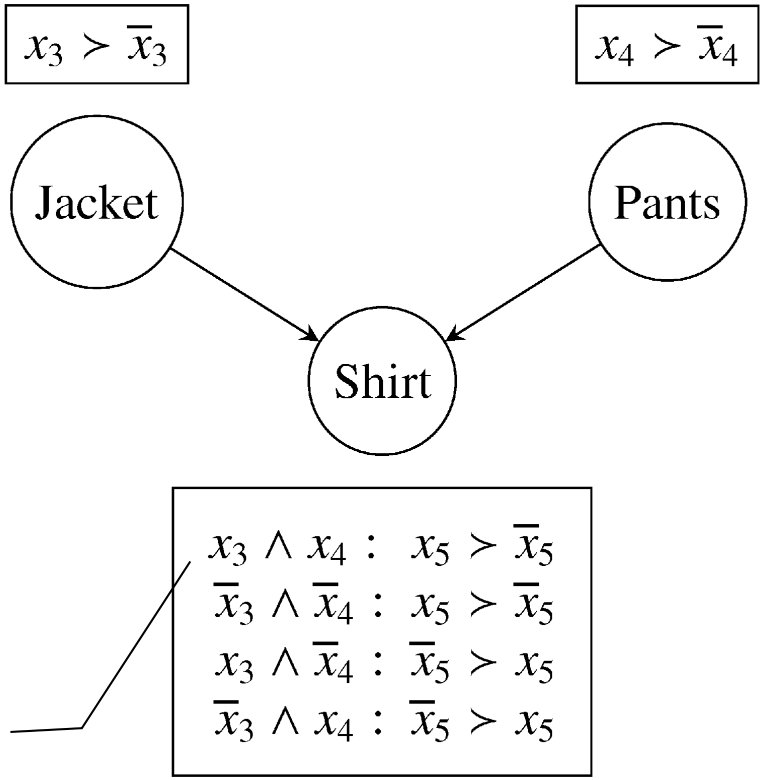
Pairwise-Based Recsys

- System that uses pairwise preferences for eliciting user preferences makes users **more aware of their choice options**
- A system variant based on pairwise preferences **outperformed a rating-based variant** in terms of recommendation accuracy measured by nDCG and precision
- **Nearest-neighbor** approaches are effective, but the user-to-user similarity must be computed with specific metrics (e.g. Goodman Kruskal gamma correlation)

- L. Blédaité, F. Ricci: Pairwise Preferences Elicitation and Exploitation for Conversational Collaborative Filtering. Hyper Text 2015: 231-236
- S. Kalloori, F. Ricci, M. Tkalcic: Pairwise Preferences Based Matrix Factorization and Nearest Neighbor Recommendation Techniques. RecSys 2016: 143-146

CP-Network

Variable	Literals
Coat	trench (x_1) and parka ($\overline{x}_1$)
Hat	large (x_2) and small ($\overline{x}_2$)
Jacket	black (x_3) and white ($\overline{x}_3$)
Pants	black (x_4) and white ($\overline{x}_4$)
Shirt	red (x_5) and white ($\overline{x}_5$)
Shoes	heels (x_6) and flats ($\overline{x}_6$)



Red shirt is preferred to white shirt if jacket and pants are black

Frédéric Koriche, Bruno Zanuttini: Learning conditional preference networks. Artif. Intell. 174(11): 685-703 (2010)

Choice Modeling

The recommender is an agent that can take decision on behalf of the user (for the user)

Decision Making

- A **decision maker** DM selects a single alternative (or action) $a \in A$
- An **outcome** (or consequence) $x \in X$ of the chosen action depends on the **state of the word** $s \in S$
- **Consequence function**:

$$c: A \times S \rightarrow X$$

- **User preferences** are expressed by a value or **utility function** – desirability of outcomes:

$$v: X \rightarrow \mathbb{R}$$

- **Goal:** *select the action $a \in A$ that leads to the best outcome*

D. Brazunas, Computational Approaches to Preference Elicitation, Tech Rep University of Toronto, 2006

Example – one user - certainty

- **Actions** = *{read, run}*
- **States** = *Contexts* = *{sun, rain}*
- **Outcomes** *X* = *Contexts x Items* = *{(read, sun), (read, rain), (run, sun), (run, rain)}*
- **Preferences in context**:
  - *v(read, sun) = 3, v(read, rain) = 4, v(run, sun) = 5, v(run, rain) = 1*
- *Context is know*
  - If it is *sun* then recommend: *run*
  - If it is *rain* then recommend: *read*

Recommender

- **If** the **context is know**
- **And** we **know** – or we can fully predict - the **preferences** of the user *u* over the space of outcomes *X* (items in context) - either as pairwise comparisons or as an ordinal function (rating):

$$r: U \times I \times C \rightarrow R$$

- *Then we can predict the user choice*

$$i^* = \arg\max_{i \in Items} \{r(u, i, c)\}$$

- **Unfeasible!**
  - Context space is huge
  - We do not fully know the relevant context
  - It is hard to accurately predict the preferences in all the possible user contexts.

G. Adomavicius, A. Tuzhilin: Context-Aware Recommender Systems. Recommender Systems Handbook 2015: 191-226

Context Aware RSs Algorithms

- Reduction-based Approach, 2005
- Exact and Generalized PreFiltering, 2009
- Item Splitting, 2009
- Tensor Factorization, 2010
- User Splitting, 2011
- Context-aware Matrix Factorization, 2011
- Factorization Machines, 2011
- Differential Context Relaxation, 2012
- Differential Context Weighting, 2013
- UI splitting, 2014
- Similarity-Based Context Modeling, 2015

Preference Knowledge

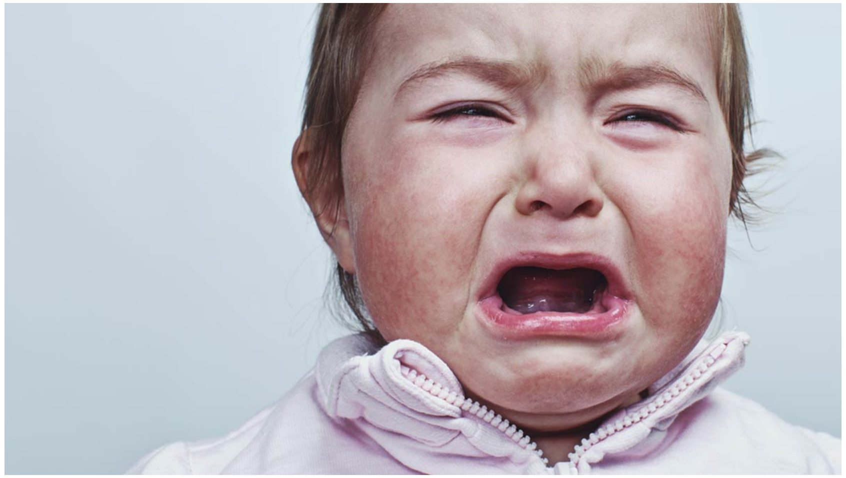
- The system knowledge of the user preferences is not only incomplete but it is also **largely inaccurate**

Remembering

- D. Kahneman (nobel prize): what we remember about an experience is determined by (**peak-end rule**)
  - *How the experience felt when it was at its peak (best or worst)*
  - *How it felt when it ended*
- *So how well do we rate or compare?*
  - It is doubtful that we prefer an experience to another very similar just because the first ended better.

Summing Up – so far

- Preferences are **context dependent**
- But it is practically **impossible to know/predict preferences** in all the potentially relevant contexts
- Preferences judgements acquired **after the experience** of the item may be **unreliable**



Irrelevant Context

- It is hard to say what is really irrelevant

Attraction Effect

- Alternative options:
  - You could get access to all our **web content** for $59,
  - A subscription to the **print edition** for $125,
  - Or a **combined print and web** subscription, also for $125.
- D. Ariely surveyed students about which option they preferred
  - Predictably, nobody chose print subscription alone;
  - 84% opted for the **combination deal**,
  - and 16% for the web subscription.

Ariely, Dan. Predictably Irrational: The Hidden Forces That Shape Our Decisions. New York: Harper Perennial, 2010.

Without Attraction

- Alternative options:
  - You could get access to all our **web content** for $59,
  - Or a **combined print and web** subscription, also for $125.

- D. Ariely surveyed again students about which option they preferred
  - 32% wanted the **print and web** subscription (vs 84% in the previous experiment)
  - while 68% preferred to go **web-only** (vs 16% in the previous experiment).

Irrelevant context does matter

- Modeling the alternative options as context

$$r: U \times I \times C \rightarrow R$$

- With the dominated option
  - r(u, web, (print, print+web)) = 4
  - **r(u, print+web, (web, print)) = 5**
  - r(u, print, (web, print+web)) = 0
- Without the dominated option
  - r(u, web, (print+web)) = 4
  - **r(u, print+web, (web)) = 3**

Context space explodes: we must consider even apparently irrelevant context wen estimating preferences.

Random Choice

- A model of choice gives the **probability** of **choosing** an item i from a set of choices X: $p(i|X)$
- If i is represented by a feature vector v_i the **multinomial logit model** (MLM) state that:

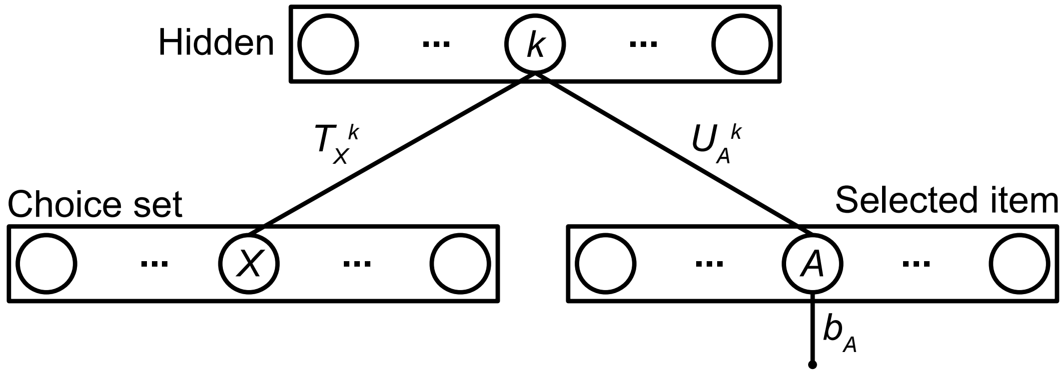
$$p(i\,|X) = \frac{\exp(w^T v_i)}{\sum_{j \in X} \exp(w^T v_j)}$$

- w is a vector of weights and $w^T v_i$ is the attractiveness of i (modelled by v_i)
- $w^T v_i = r(u,i)$ – assuming w is the vector modeling u
- This is a step ahead from the assumption that u will choose the item i that maximizes $r(u,i)$.

T. Osogami, Human choice and good choice, in The role and importance of mathematics in innovation, Springer, 2017.

Restricted Boltzmann Machine

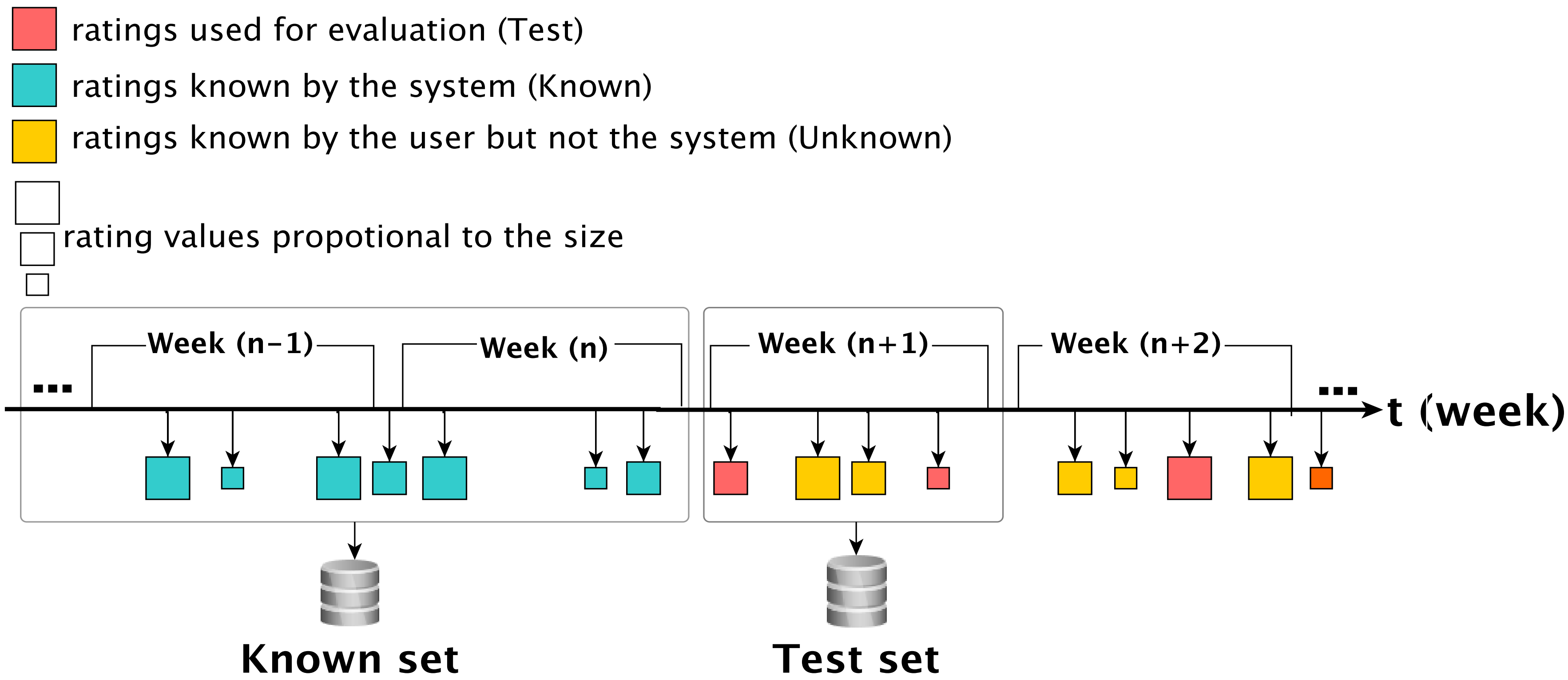
- MLM choice model cannot explain "attraction" since the ratio of $p(i|X)$ and $p(j|X)$ does not change if we remove an item k from the choice set X
- In a restricted Boltzmann machine the attractiveness of an item depends on the attractiveness of the other items



T. Osogami, M. Otsuka: Restricted Boltzmann machines modeling human choice. NIPS 2014: 73-81

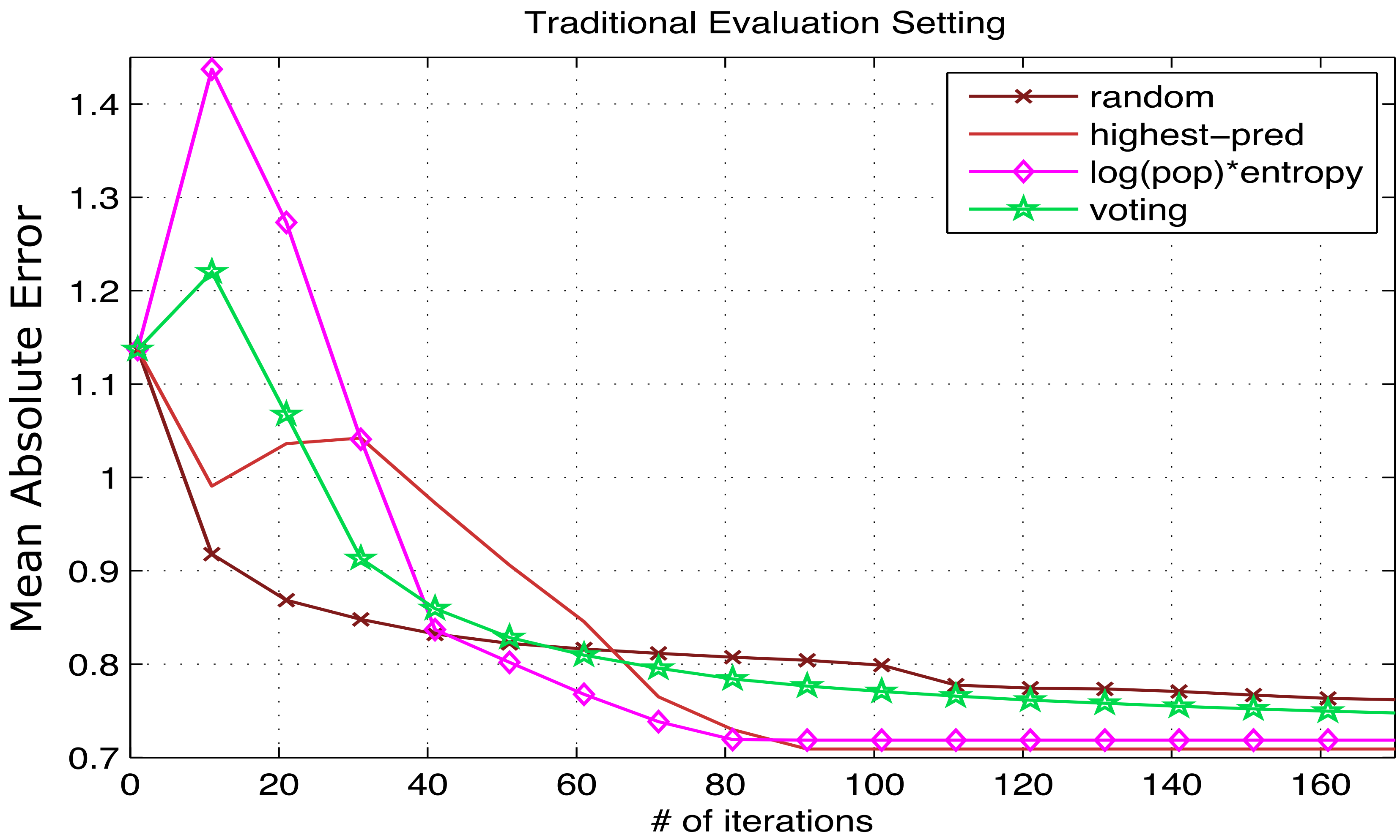
System Dynamics

Simulating Rating Acquisition



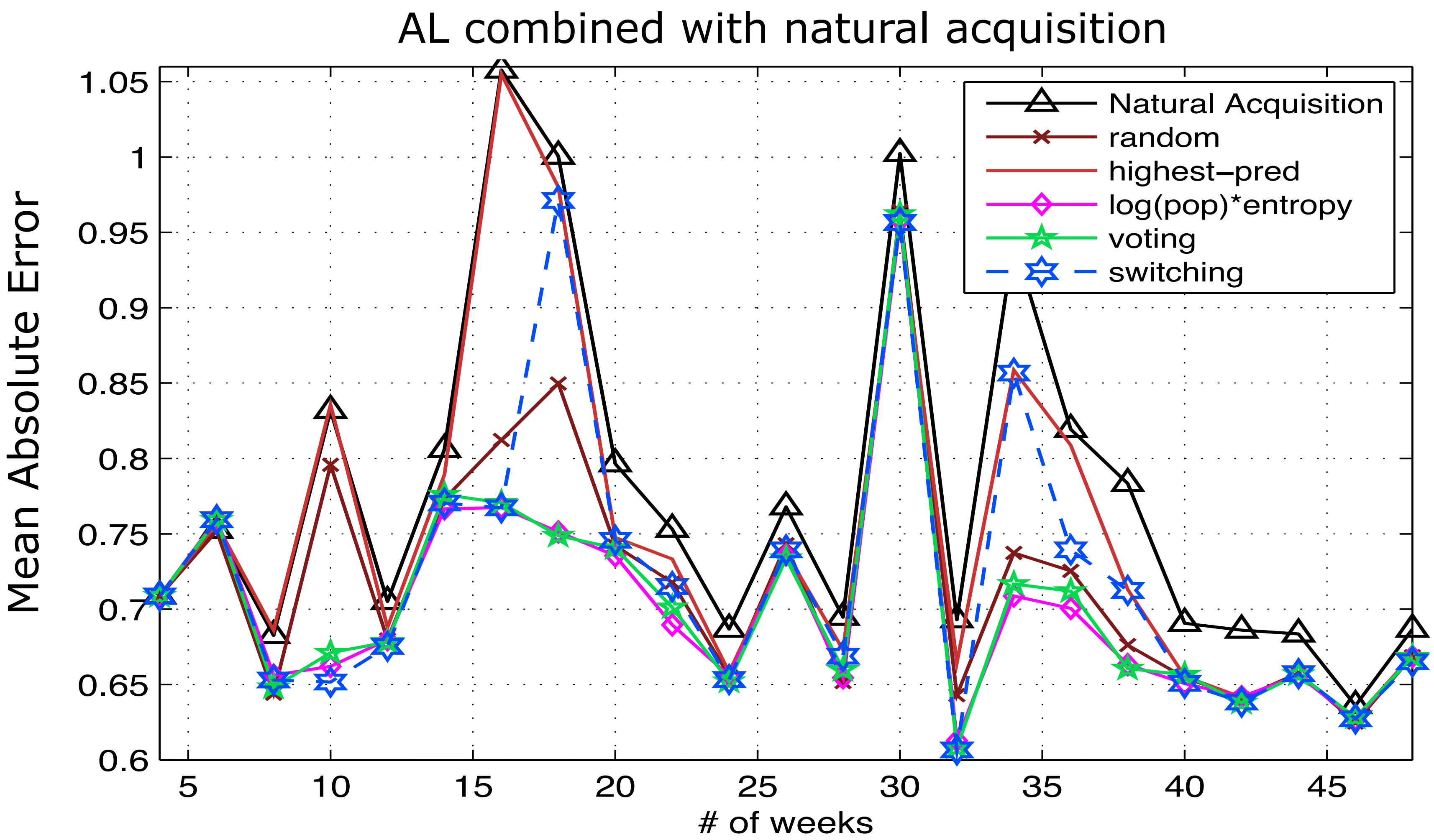
[M. Elahi, F. Ricci, N. Rubens: Active learning strategies for rating elicitation in collaborative filtering: A system-wide perspective. ACM TIST 5(1): 13:1-13:33 (2013)]

Active Learning Strategies



M. Elahi, F. Ricci, N. Rubens: Active learning strategies for rating elicitation in collaborative filtering: A system-wide perspective. ACM TIST 5(1): 13:1-13:33 (2013)

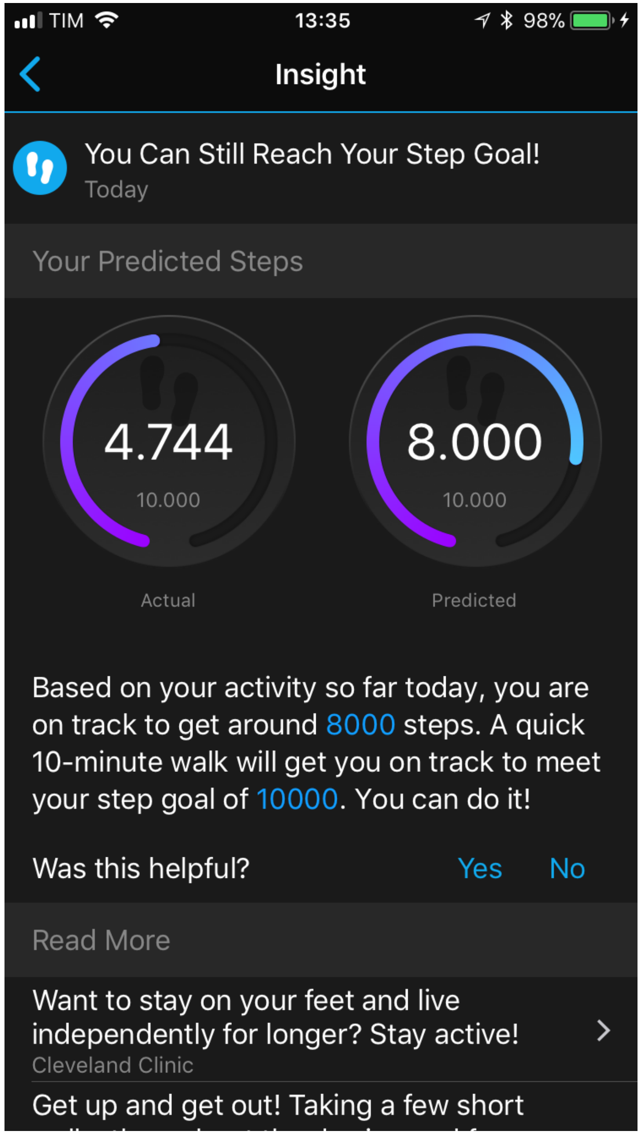
Active Learning and Natural Acquisition



Do we really need 100% precise preference and choice prediction models?

Choices and Recommendations

- Recommendations should not only tell the user what is the target (choice)
- The target may be given – e.g. by the user
- How to smartly achieve the target may be more useful.



Behavioural Model Learning

- Learning the choice model of the user
- Determining the rationale for the decisions
- Generating “non trivial” recommendations that intelligently deviates from the learned behavioral model
  - *The user is predicted to take a coffee at 8:00 at Walter Bar; let us suggest to get it at Rosy Bar – it is cheaper and better*
  - *The user is getting back home; let us suggest to visit a Photo Exhibition along the path – he likes photography and will still be able to get home on time.*
- Transparent behavioral model is learnt using Inverse Reinforcement Learning.

D. Massimo, M. Elahi, F. Ricci: Learning User Preferences by Observing User-Items Interactions in an IoT Augmented Space. UMAP 2017.

Lesson Learned

- Preferences are **contextual**, **dynamic** and **hard to predict**
- Predicting preferences does not suffice for supporting decision making with recommendations - **choice model**
- **Preference are dynamic** and it is important to control preference elicitation and understand the effect of the recommender on the elicited preferences
- **Recommendations** could be more “intelligently” generated on top of the learned preferences and choice models.

Thanks

- In particular to my students and collaborators who contributed to develop these ideas:
  - David Massimo
  - Linas Baltrunas
  - Laura Bledaite
  - Marius Kaminskas
  - Marko Gasparic
  - Marko Tkalcic
  - Matthias Braunhofer
  - Mehdi Elahi
  - Saikishore Kalloori
  - Tural Gurbanov
  - Thuy Ngoc Nguyen