Considerations on Recommendation Independence for a Find-Good-Items Task

Toshihiro Kamishima[*] and Shotaro Akaho[*]

http://www.kamishima.net/

National Institute of Advanced Industrial Science and Technology (AIST), Japan

Workshop on Responsible Recommendation (FATREC)

in conjunction with RecSys2017 @ Como, Italy, Aug. 31, 2017

START

This Talk

Our previous work

- We advocated a concept of **recommendation independence**
- We developed two types of approaches to enhance recommendation independence for a **predicting-ratings task**
  - a **regularization approach** using a constraint term
  - a **model-based approach** using a special graphical model

This workshop paper

- A preliminary experiment of applying a regularization approach to a recommender for a **find-good-items task**
  - Independence of a preference score was enhanced, but that of relevance was not
  - Errors in AUC was not worsen by enhancing independence

Sensitive Feature

As in a case of standard recommendation, we use random variables
X: a user, Y: an item, and R: a recommendation outcome

We adopt a variable required for recommendation independence

S : sensitive feature

- This represents information to be ignored in a recommendation process
- Its values are determined depending on a user and/or an item

Ex. Sensitive feature = movie's popularity / user's gender

A sensitive feature is restricted to a binary type

Recommendation Independence

[Kamishima 12, Kamishima 13]

Recommendation Independence

the statistical independence

between a recommendation outcome, R, and a sensitive feature, S

$$\Pr[R \mid S] = \Pr[R]$$

$$\equiv$$

$$R \perp\!\!\!\perp S$$

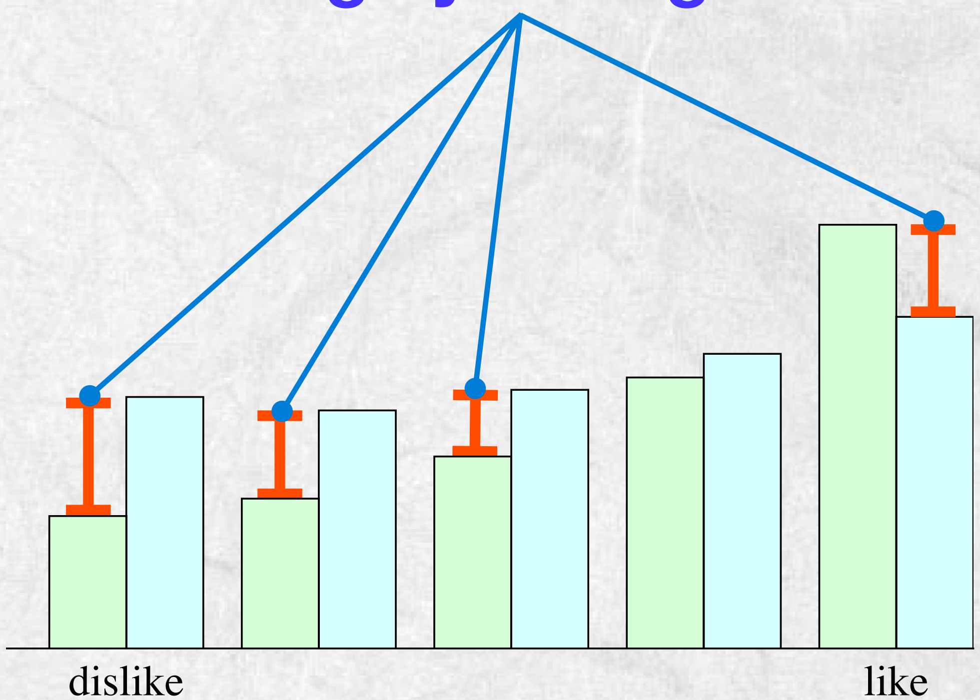
- No information about a sensitive feature influences the outcome
- The status of the sensitive feature is explicitly excluded from the inference of the recommendation outcome

Recommendation outcomes are predicted under the constraint of recommendation independence

Effect of Independence Enhancement

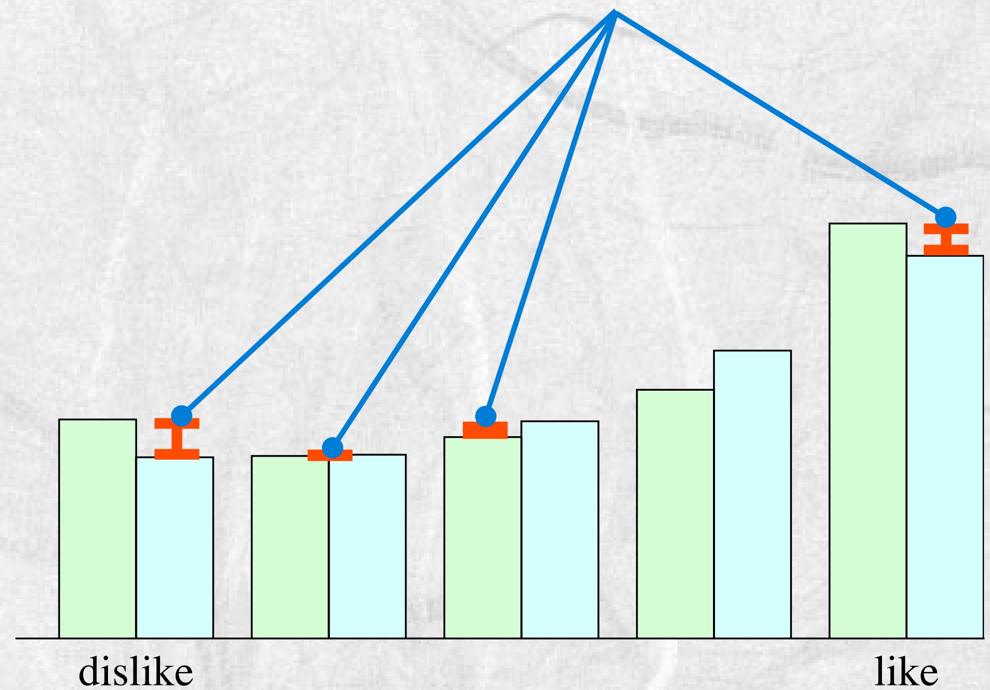
Standard

two distributions are largely diverged



Independence-enhanced

two distributions become closer



a sensitive feature = whether a movie is newer or older

✻ each bin of histograms of predicted scores for older and newer movies

The bias that older movies were rated higher could be successfully canceled by enhancing independence

Applications

Application
Adherence to Laws and Regulations

[Sweeney 13]

A recommendation service must be managed while adhering to laws and regulations

suspicious placement in keyword-matching advertisements

Advertisements indicating arrest records were more frequently displayed for names that are more popular among individuals of African descent than those of European descent

⬇

Socially discriminative treatments must be avoided

sensitive feature = users' demographic information

⬇

Legally or socially sensitive information can be excluded from the inference process of recommendation

Application
Fair Treatment of Content Providers

System managers should fairly treat their content providers

Fair treatment in search engines

[Bloomberg]

The US FTC has investigated Google to determine whether the search engine ranks its own services higher than those of competitors

Fair treatment in recommendation

A hotel booking site should not abuse their position to recommend hotels of its group company

sensitive feature = a content provider of a candidate item

⬇

Information about who provides a candidate item can be ignored, and providers are treated fairly

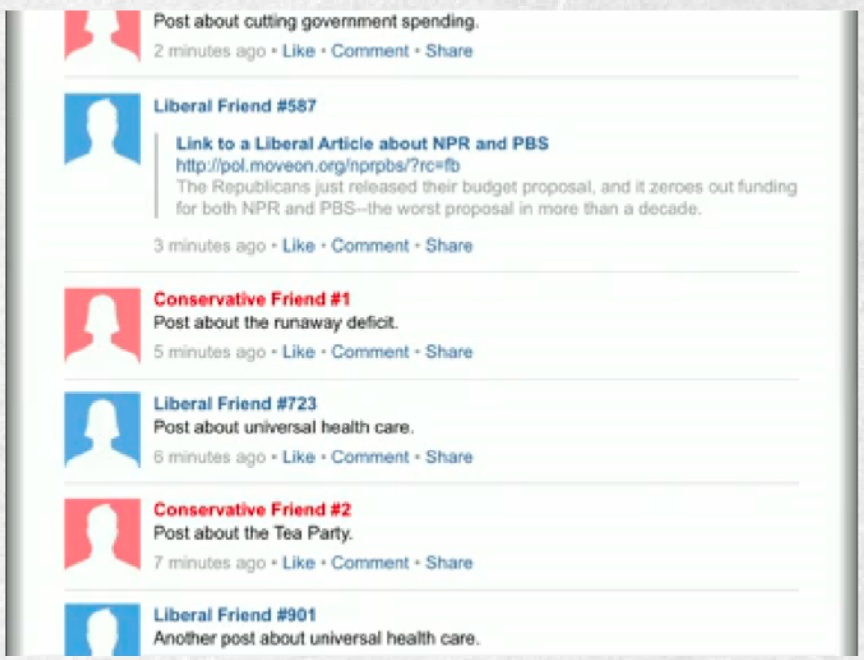
Application

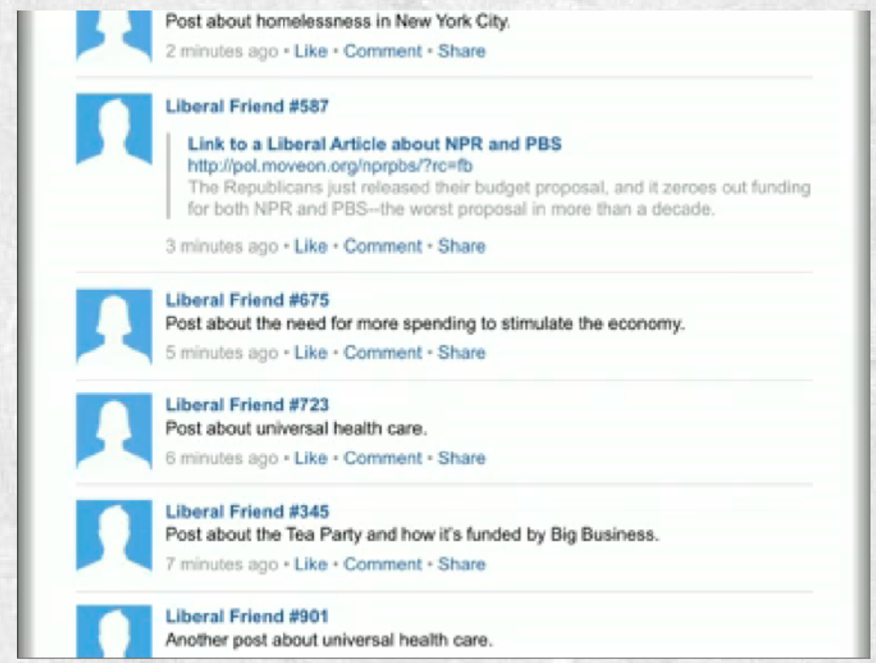
Exclusion of Unwanted Information

[TED Talk by Eli Pariser, http://www.filterbubble.com/]

Information unwanted by a user is excluded from recommendation

Filter Bubble: To fit for Pariser's preference, conservative people are eliminated from his friend recommendation list in FaceBook





sensitive feature = a political conviction of a friend candidate

⬇

Information about whether a candidate is conservative or progressive can be ignored in a recommendation process

Independence-Enhanced Recommendation Algorithms

Independence-Enhanced Recommendation Task

[Kamishima+ 12, Kamishima+ 13, Kamishima+ 14, Kamishima+16]

Predicting Ratings: a task to predict a rating value that a user would provide to an item

Random variables: user X, item Y, rating R, sensitive feature S

Standard Recommendation

Dataset

$$\mathcal{D} = \{(x_i, y_i, r_i)\}$$

Prediction Function

$$\hat{r}(x, y)$$

Independence-Enhanced Rec.

Dataset

$$\mathcal{D} = \{(x_i, y_i, r_i, s_i)\}$$

Prediction Function

$$\hat{r}(x, y, s)$$

Probabilistic Matrix Factorization

[Salakhutdinov 08, Koren 08]

Probabilistic Matrix Factorization Model

predict a preference rating of an item y rated by a user x

well-performed and widely used

Prediction Function

$$\hat{r}(x,y) = \underbrace{\mu}_{\text{global bias}} + \underbrace{b_x}_{\text{user-dependent bias}} + \underbrace{c_y}_{\text{item-dependent bias}} + \underbrace{\mathbf{p}_x \mathbf{q}_y^\top}_{\text{cross effect of users and items}}$$

Objective Function

$$\underbrace{\textstyle\sum_D (r_i - \hat{r}(x_i, y_i))^2}_{\text{squared loss function}} + \underbrace{\lambda}_{\text{regularization parameter}} \underbrace{\|\boldsymbol{\Theta}\|^2}_{L_2 \text{ regularizer}}$$

For a given training dataset, model parameters are learned by minimizing the squared loss function with an L_2 regularizer.

Independence-Enhaned PMF

[Kamishima+ 13]

Prediction Function

a prediction function is selected according to a sensitive value

$$\hat{r}(x, y, s) = \mu^{(s)} + b_x^{(s)} + c_y^{(s)} + \mathbf{p}_x^{(s)} \mathbf{q}_y^{(s)\top}$$

sensitive feature

Objective Function

independence parameter: control the balance between the independence and accuracy

$$\sum_D \left(r_i - \hat{r}(x_i, y_i, s_i)\right)^2 - \eta \, \boxed{\mathrm{indep}(R, S)} + \lambda \|\Theta\|^2$$

independence term: a regularizer to constrain independence

- The larger value indicates that ratings and sensitive values are more independent
- Matching means of predicted ratings for two sensitive values

For a Find-Good-Items Task

[Kamishima+ 17]

Find Good Items: a task to find some items preferred by a user

⬇

making a preference score independent, instead of a predicted rating

Preference Score: How strongly a user prefers an item

sigmoid function

$$\hat{r}(x, y, s) = \operatorname{sig}\left(\mu^{(s)} + b_x^{(s)} + c_y^{(s)} + \mathbf{p}_x^{(s)} {\mathbf{q}_y^{(s)}}^\top\right)$$

$$\textstyle\sum_D \mathrm{CE}\left(r_i - \hat{r}(x_i, y_i, s_i)\right) - \eta \ \mathrm{indep}(R, S) + \lambda \|\Theta\|^2$$

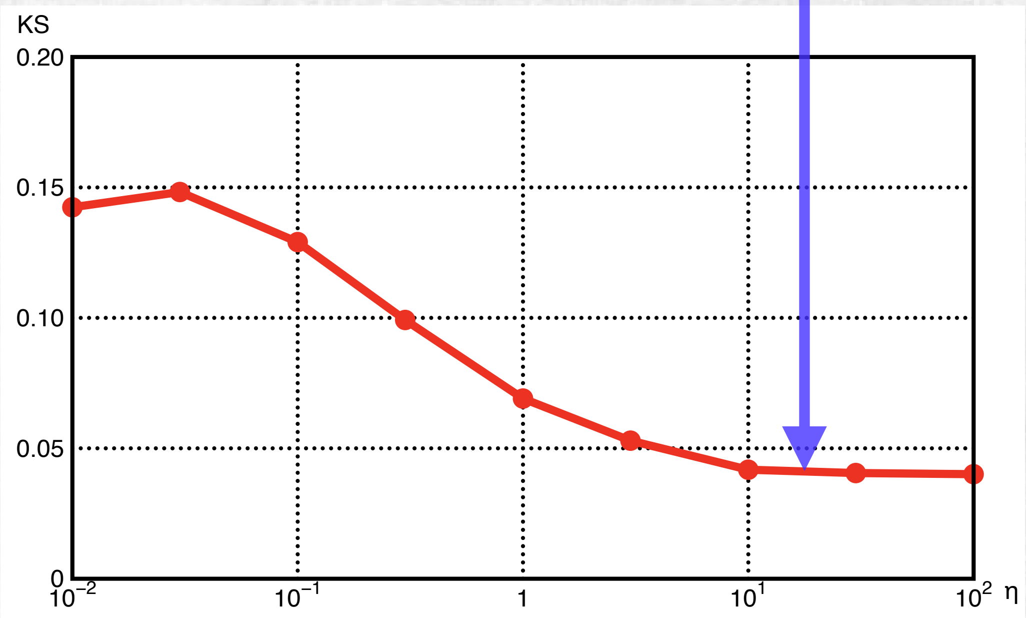
cross-entropy loss

enhancing the independence between a preference score and a sensitive feature

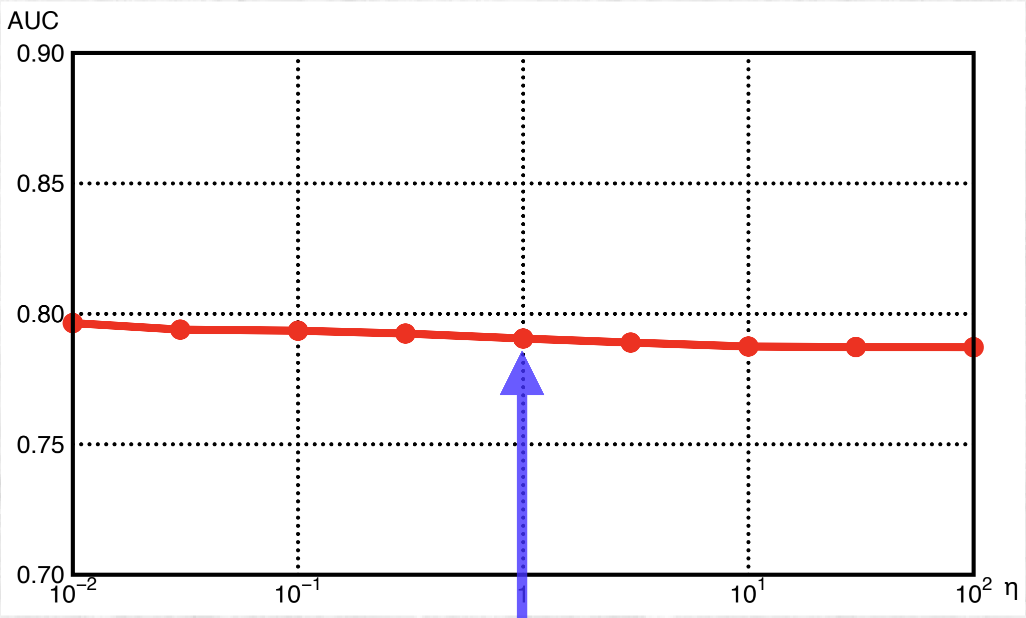
Preference Score vs Sensitive Feature

[Kamishima+ 17]

Observation 1: **A preference score could be successfully made independent from a sensitive feature**



Kolmogolov-Smirnov statistic between $\Pr[R \mid S=0]$ and $\Pr[R \mid S=1]$



ranking accuracy (AUC)

Observation 2: **A ranking accuracy (AUC) did not worsen so much** by enhancement of the recommendation independence
This is was contrasted with the increase of a prediction error (MAE) in a predicting ratings task

Relevance and Sensitive Feature

[Kamishima+ 17]

Recommending top-k items whose preference scores are the largest

sort according to their preference scores

select top-k items

1.0 1.0 0.9 | 0.7 0.5 0.3 0.1

relevant items | irrelevant items

check the independence from a relevance, not from a preference score

Observation 3: **The relevance of items was not independent from a sensitive feature** for some values of k, in particular, small k

A need for a new method that fits for a ranked item list

Conclusions

Contributions

- We advocated a notion of recommendation independence and developed methods to enhance it
- We tested a preliminary approach to enhance independence, but it was not effective for a ranked item list

Future work

- Developing an independence-enhancement method being fit for a ranked item list

Acknowledgment

- We would like to thank for providing datasets for the Grouplens research lab
- This work is supported by MEXT/JSPS KAKENHI Grant Number JP24500194, JP15K00327, and JP16H02864

Survey Slide of Fairness-Aware Data Mining

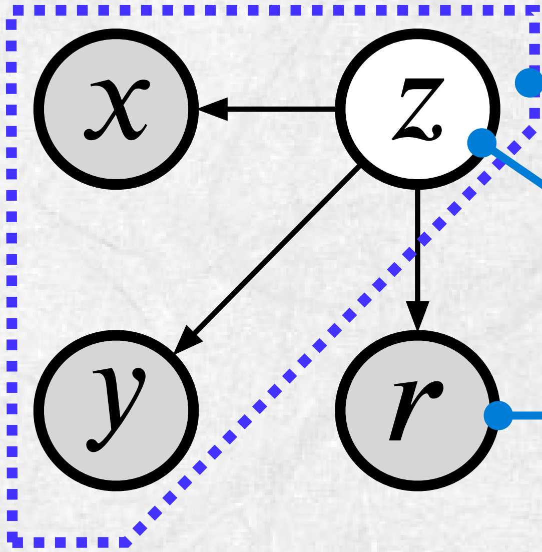
http://www.kamishima.net/archive/fadm.pdf

Extra Slides

Latent Class Model

[Hofmann 99]

Latent Class Model: A probabilistic model for collaborative filtering



A basic topic model, pLSA

latent topic variable

extended so as to be able to deal with ratings r given by users x to items y

Model parameters can be learned by an EM algorithm

Prediction:

$$
\begin{aligned}
\hat{r}(x, y) &= \mathrm{E}_{\Pr[r|x,y]}[\mathrm{level}(r)] \\
&= \textstyle\sum_r \Pr[r|x, y]\, \mathrm{level}(r)
\end{aligned}
$$

the r-th rating value

A rating value can be predicted by the expectation of ratings

Independence-Enhanced LCM

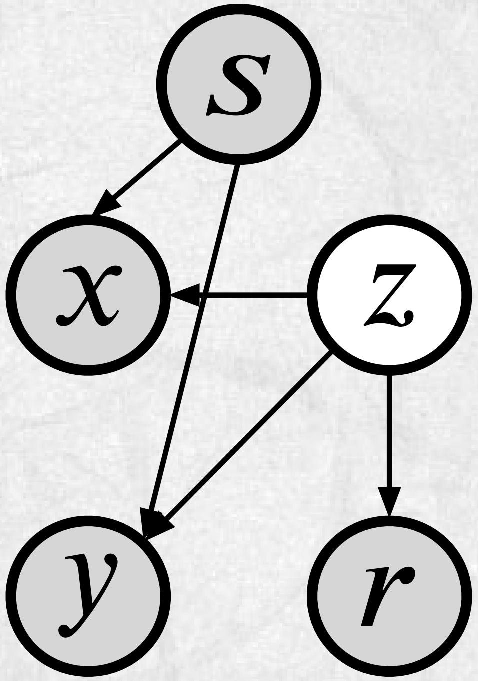
[Kamishima+ 16]

Independence-Enhancement by a Model-based Approach

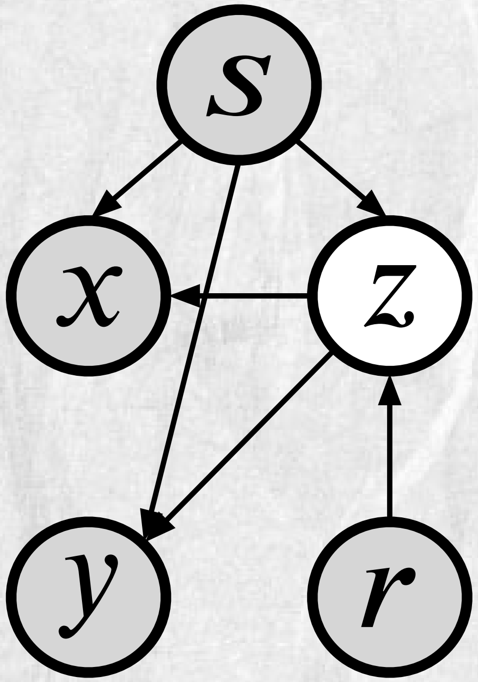
A sensitive variable is embedded into the original LCM

- A rating and a sensitive variable are mutually independent
- A user, an item, and a rating are conditionally independent given Z

Type 1 model



Type 2 model



A type 2 model can more strictly enhance recommendation independence, because in addition to X and Y, Z depends on a sensitive variable

Popularity Bias

[Celma 08]

Popularity Bias
the tendency for popular items to be recommended more frequently

[Jamali+ 10]

Flixster data
The degree popularity of an item is measured
by the number of users who rated the item

short-head (top 1%)
share in ratings: 47.2%
mean rating: **3.71**

↔

long-tail (bottom 99%)
share in ratings: 52.8%
mean rating: **3.53**

Short-head items are frequently and highly rated

sensitive feature = popularity of items

⬇

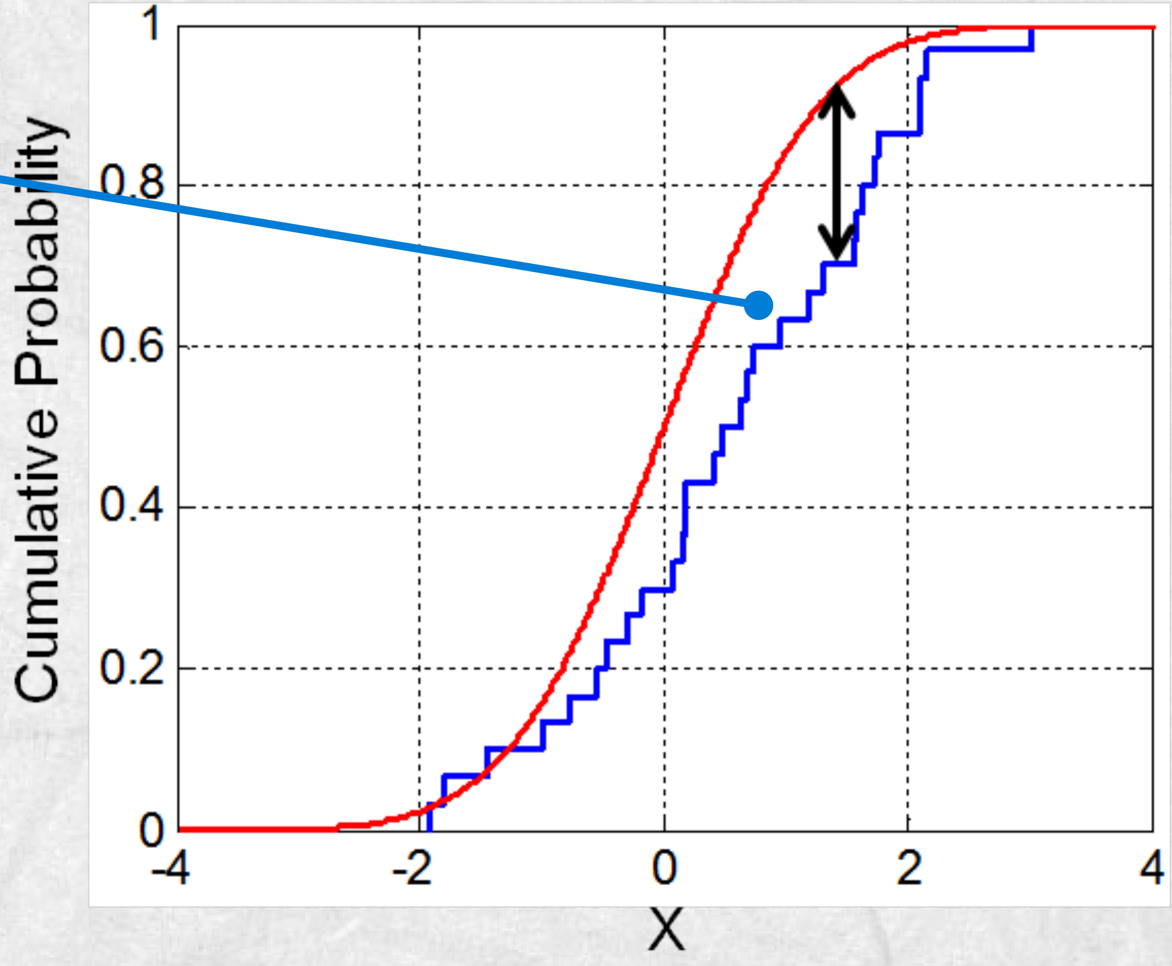
Popularity bias can be corrected

Kolmogorov-Smirnov Statistic

The statistic of the two-sample Kolmogorov-Smirnov test
a nonparametric test for the equality of two distribution

Evaluating the degree of independence
by measuring the equality between $\Pr[R \mid S{=}0]$ and $\Pr[R \mid S{=}1]$

Kolmogorov-Smirnov statistic
the area between two empirical cumulative distributions



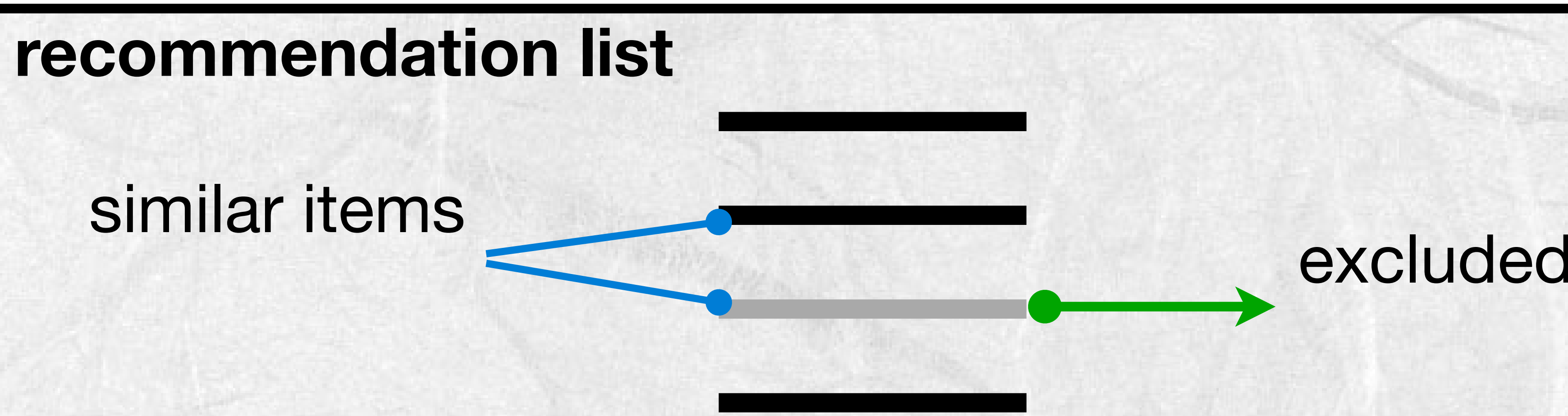
Wikipedia

Recommendation Diversity

[Ziegler+ 05, Zhang+ 08, Latha+ 09, Adomavicius+ 12]

Recommendation Diversity

Similar items are not recommended in a single list, to a single user, to all users, or in a temporally successive lists



Diversity

Items that are similar in a specified metric are excluded from recommendation results

The mutual relations among results

Independence

Information about a sensitive feature is excluded from recommendation results

The relations between results and sensitive values

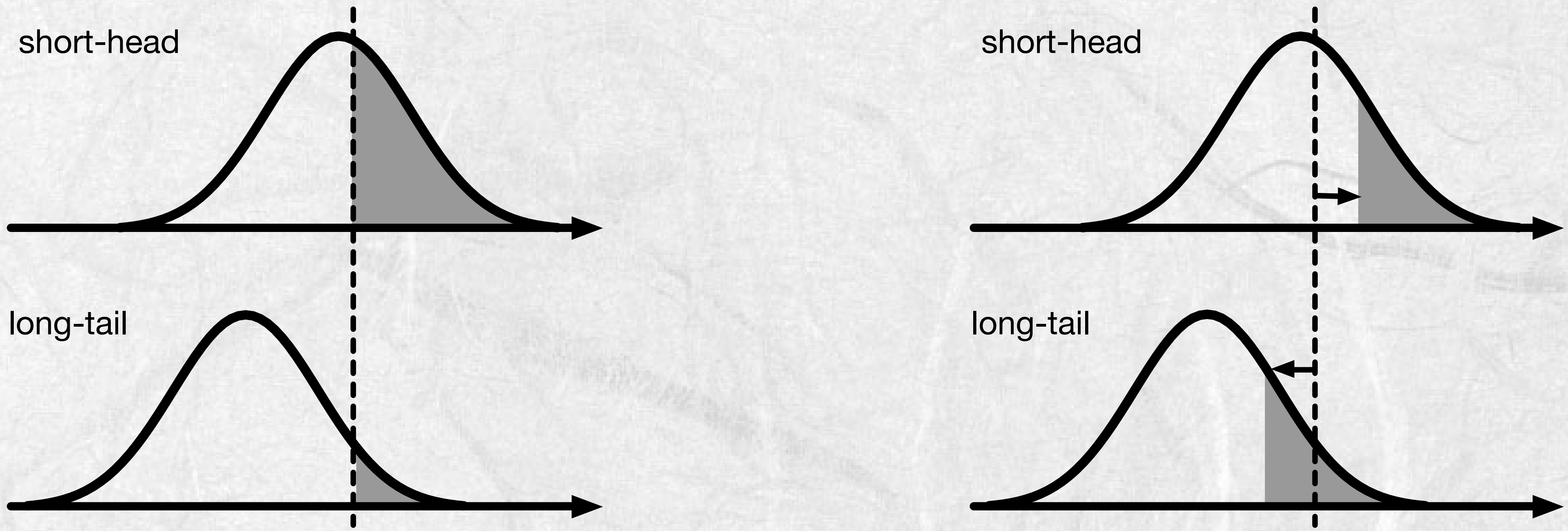
Diversity vs Independence

Diversity	**Independence**
Depending on the definition of similarity measures	Depending on the specification of sensitive feature
⬇	⬇
Similarity	**Sensitive Feature**
A function of **two items**	A function of **a user-item pair**

Because a sensitive feature depends on a user, neutrality can be applicable for coping with users' factor, such as, users' gender or age, which cannot be straightforwardly dealt by using diversity

Diversity vs Independence

standard recommendation | diversified recommendation



Because a set of recommendations are diversified by abandoning short-head items, **predicted ratings are still biased**

Prediction ratings themselves are unbiased by enhancing recommendation independence

Privacy-Preserving Data Mining

recommendation results, R, and sensitive features, S,
are statistically independent

⬇

mutual information between a recommendation result, R,
and a sensitive feature, S, is zero

$$\mathrm{I}(R; S) = 0$$

⬇

In a context of privacy-preservation
Even if the information about R is disclosed,
the information about S will not exposed

In particular, a notion of the t-closeness has strong connection