Predicting the Performance of Recommender Systems: An Information Theoretic Approach

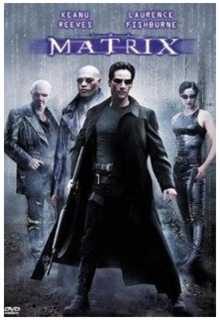
Alejandro Bellogín, Pablo Castells, Iván Cantador
Escuela Politécnica Superior
Universidad Autónoma de Madrid

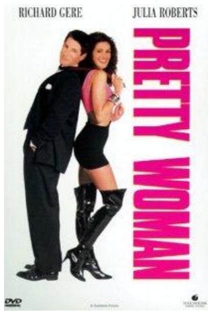
@abellogin
alejandro.bellogin@uam.es

Recommender Systems

- RS suggests “interesting” items to users
  - Most common: explicit ratings
  - Goal: predict rating r_{jk}

items





users

	i_1	…	i_k	…	i_m
u_1	★★★★★		★★★★★		★☆☆☆☆
⋮					
u_j	★★★☆☆		?		★★☆☆☆
⋮					
u_n	★★☆☆☆		★★★★☆		★★☆☆☆

Performance Prediction in IR

- Estimation of the system's performance in response to a specific query
- Predictors: query scope, query clarity, query drift, …
- We focus on **query clarity**:

$$\text{clarity}(q) = \sum_{w \in V} p(w \mid q) \log\left(\frac{p(w \mid q)}{p_c(w)}\right)$$

$$p(d \mid q) = p(q \mid d)\, p(d);\; p(q \mid d) = \prod_{w_q \in q} p(w_q \mid d)$$

$$p(w \mid q) = \sum_{d \in R} p(w \mid d)\, p(d \mid q)$$

$$p(w \mid d) = \lambda p_{ml}(w \mid d) + (1 - \lambda)\, p_c(w)$$

- Cronen-Townsend, S., Zhou, Y., Croft, W.B.: Predicting query performance. SIGIR 2002.
- He, B., Ounis, I.: Inferring query performance using pre-retrieval predictors. SPIRE 2004.
- Mitra, M.,Singhal, A., Buckley, C.: Improving automatic query expansion. SIGIR 1998.

Predictive Models of Recommendation Performance

- User clarity:
  - Distance between the user's and the system's probability model

$$\text{clarity}(u) = \sum_{x \in X} p(x|u) \log\left(\frac{p(x|u)}{p_c(u)}\right)$$

user model

background model

  - *X* may be: users, items, ratings, or a combination (*vocabulary* space)

Predictive Models of Recommendation Performance

- Three user clarity formulations:

Name	Vocabulary	User model	Background model
Rating-based	Ratings	$p(r \mid u)$	$p_c(r)$
Item-based	Items	$p(i \mid u)$	$p_c(i)$
Item-and-rating-based	Items rated by the user	$p(r \mid i,u)$	$p_{ml}(r \mid i)$

$$\text{clarity}(u) = \sum_{x \in X} p(x \mid u) \log\left(\frac{p(x \mid u)}{p_c(u)}\right)$$

user model

background model

Predictive Models of Recommendation Performance

- Seven user clarity models implemented:

Name	Formulation	User model	Background model
RatUser	Rating-based	$p_U(r \mid i,u); p_{UR}(i \mid u)$	$p_c(r)$
RatItem	Rating-based	$p_I(r \mid i,u); p_{UR}(i \mid u)$	$p_c(r)$
ItemSimple	Item-based	$p_R(i \mid u)$	$p_c(i)$
ItemUser	Item-based	$p_{UR}(i \mid u)$	$p_c(i)$
IRUser	Item-and-rating-based	$p_U(r \mid i,u)$	$p_{ml}(r \mid i)$
IRItem	Item-and-rating-based	$p_I(r \mid i,u)$	$p_{ml}(r \mid i)$
IRUserItem	Item-and-rating-based	$p_{UI}(r \mid i,u)$	$p_{ml}(r \mid i)$

Predictive Models of Recommendation Performance

- Seven user clarity models implemented:

Name	Formulation	User model	Background model
RatUser	Rating-based	$p_U(r\mid i,u); p_{UR}(i\mid u)$	$p_c(r)$
RatItem	Rating-based	$p_I(r\mid i,u); p_{UR}(i\mid u)$	$p_c(r)$
ItemSimple	Item-based	$p_R(i\mid u)$	$p_c(i)$
ItemUser	Item-based	$p_{UR}(i\mid u)$	$p_c(i)$
IRUser	Item-and-rating-based	$p_U(r\mid i,u)$	$p_{ml}(r\mid i)$
IRItem	Item-and-rating-based	$p_I(r\mid i,u)$	$p_{ml}(r\mid i)$
IRUserItem	Item-and-rating-based	$p_{UI}(r\mid i,u)$	$p_{ml}(r\mid i)$

$$p_U(r\mid i,u)\propto p(u\mid r,i) \qquad p_I(r\mid i,u)\propto p(i\mid r,u)$$

$$p_{UI}(r\mid i,u)\propto p(u,i\mid r)$$

- Wang, J., de Vries, A.P., Reinders, M.J.T.: Unified relevance models for rating prediction in collabotative filtering. ACM TOIS 2008.

Predictive Models of Recommendation Performance

- Seven user clarity models implemented:

Name	Formulation	User model	Background model
RatUser	Rating-based	$p_U(r\,|\,i,u); p_{UR}(i\,|\,u)$	$p_c(r)$
RatItem	Rating-based	$p_I(r\,|\,i,u); p_{UR}(i\,|\,u)$	$p_c(r)$
ItemSimple	Item-based	$p_R(i\,|\,u)$	$p_c(i)$
ItemUser	Item-based	$p_{UR}(i\,|\,u)$	$p_c(i)$
IRUser	Item-and-rating-based	$p_U(r\,|\,i,u)$	$p_{ml}(r\,|\,i)$
IRItem	Item-and-rating-based	$p_I(r\,|\,i,u)$	$p_{ml}(r\,|\,i)$
IRUserItem	Item-and-rating-based	$p_{UI}(r\,|\,i,u)$	$p_{ml}(r\,|\,i)$

$$p_R(i\,|\,u) = \sum_{r \in R} p_{ml}(i\,|\,r)\, p_{ml}(r\,|\,u)$$

$$p_{UR}(i\,|\,u) = \sum_{r \in R} p(i\,|\,u,r)\, p_{ml}(r\,|\,u)$$

Predictive Models of Recommendation Performance

- Seven user clarity models implemented:

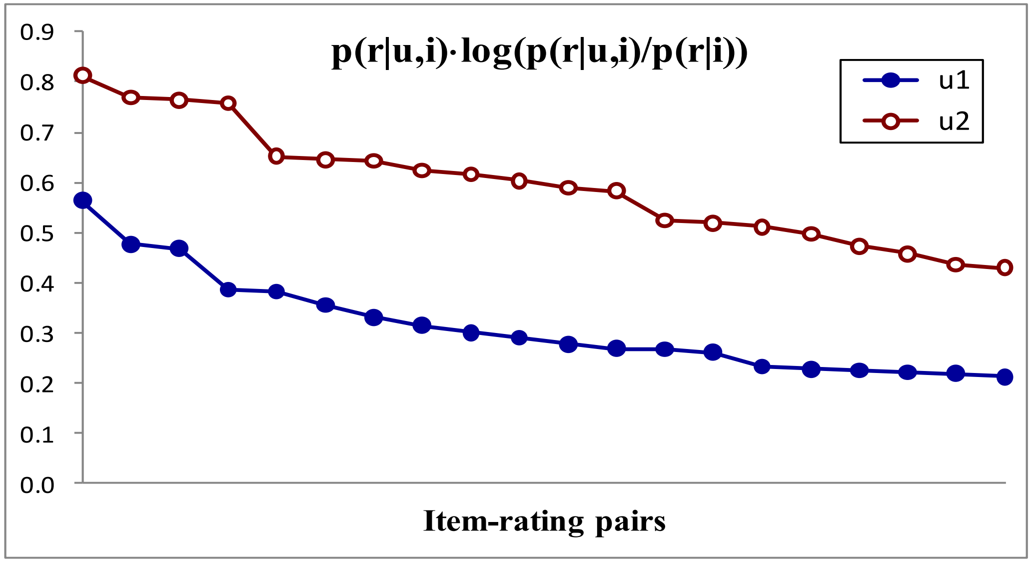
Name	Formulation	User model	Background model
RatUser	Rating-based	$p_U(r\|i,u); p_{UR}(i\|u)$	$p_c(r)$
RatItem	Rating-based	$p_I(r\|i,u); p_{UR}(i\|u)$	$p_c(r)$
ItemSimple	Item-based	$p_R(i\|u)$	$p_c(i)$
ItemUser	Item-based	$p_{UR}(i\|u)$	$p_c(i)$
IRUser	Item-and-rating-based	$p_U(r\|i,u)$	$p_{ml}(r\|i)$
IRItem	Item-and-rating-based	$p_I(r\|i,u)$	$p_{ml}(r\|i)$
IRUserItem	Item-and-rating-based	$p_{UI}(r\|i,u)$	$p_{ml}(r\|i)$

$$p(r|u) = \sum_{r(u,i)=r} p(r|i,u)\,p(i|u)$$

Examples

- Comparison of models for two users:

User	Number of ratings	ItemUser clarity	RatItem clarity	IRUserItem clarity
u1	51	216.01	28.60	6.85
u2	52	243.32	43.63	13.56



Term contributions for each user, ordered by their corresponding contribution to the user language model. **IRUserItem** clarity model.

Examples

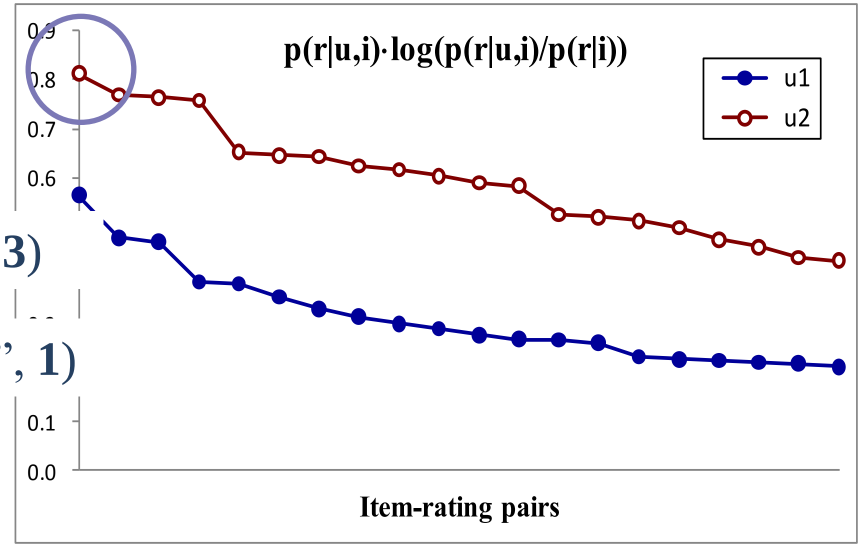
- Comparison of models for two users:

User	Number of ratings	ItemUser clarity	RatItem clarity	IRUserItem clarity
u1	51	216.01	28.60	6.85
u2	52	243.32	43.63	13.56

(4, “Mc Hale’s Navy”)

(u2, “Mc Hale’s Navy”, **3**)

(com, “Mc Hale’s Navy”, **1**)



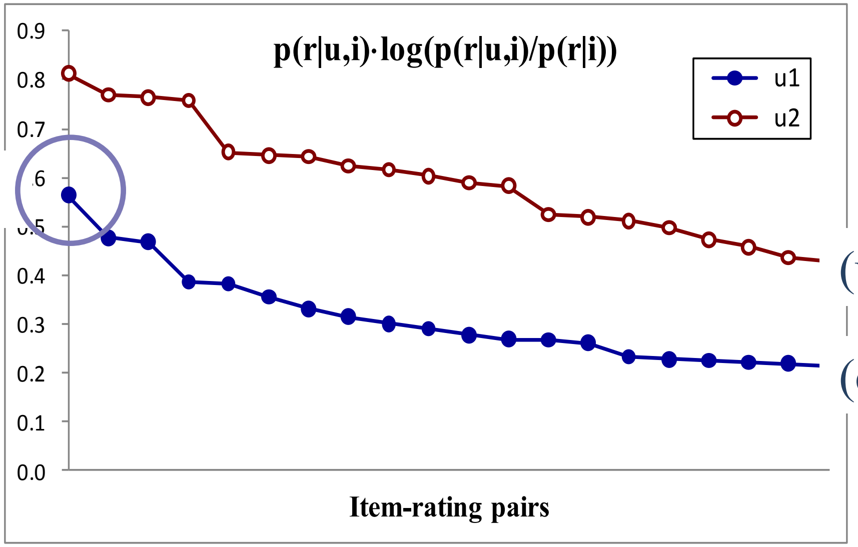
Term contributions for each user, ordered by their corresponding contribution to the user language model. **IRUserItem** clarity model.

Examples

- Comparison of models for two users:

User	Number of ratings	ItemUser clarity	RatItem clarity	IRUserItem clarity
u1	51	216.01	28.60	6.85
u2	52	243.32	43.63	13.56

(2, “Donnie Brasco”)

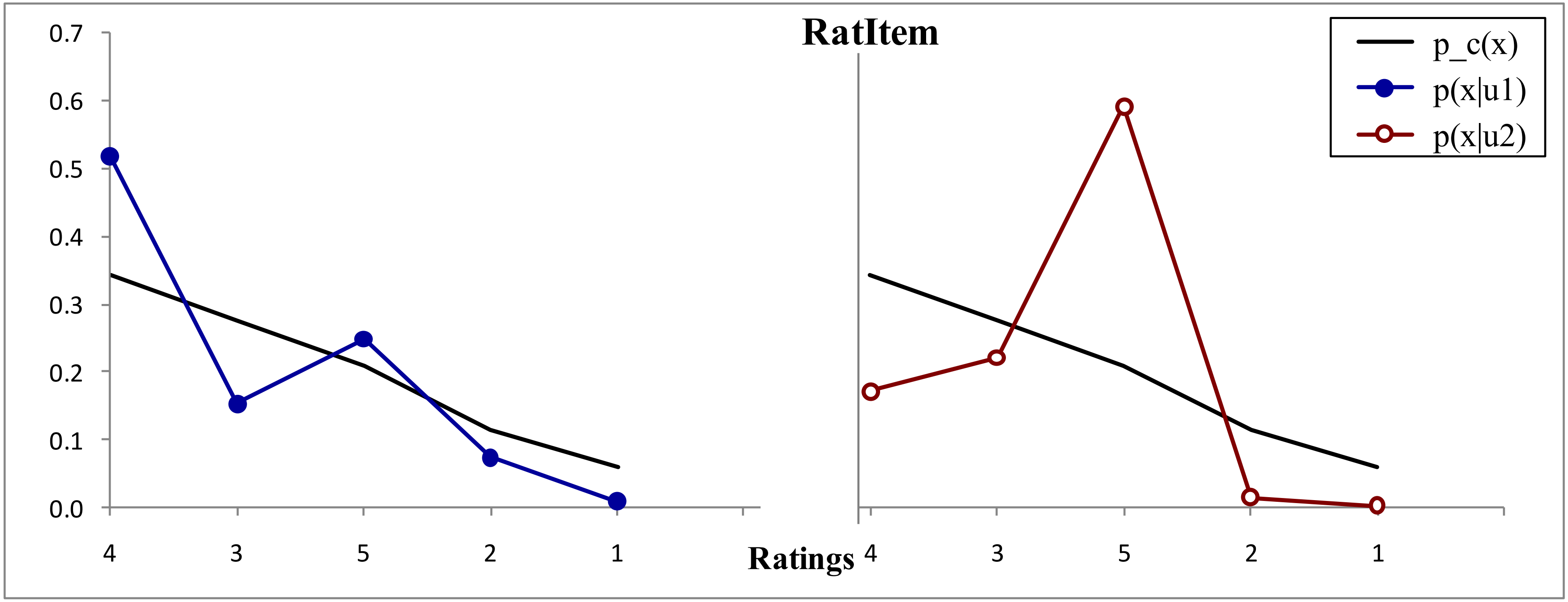


(u1, “Donnie Brasco”, **5**)

(com, “Donnie Brasco”, **4**)

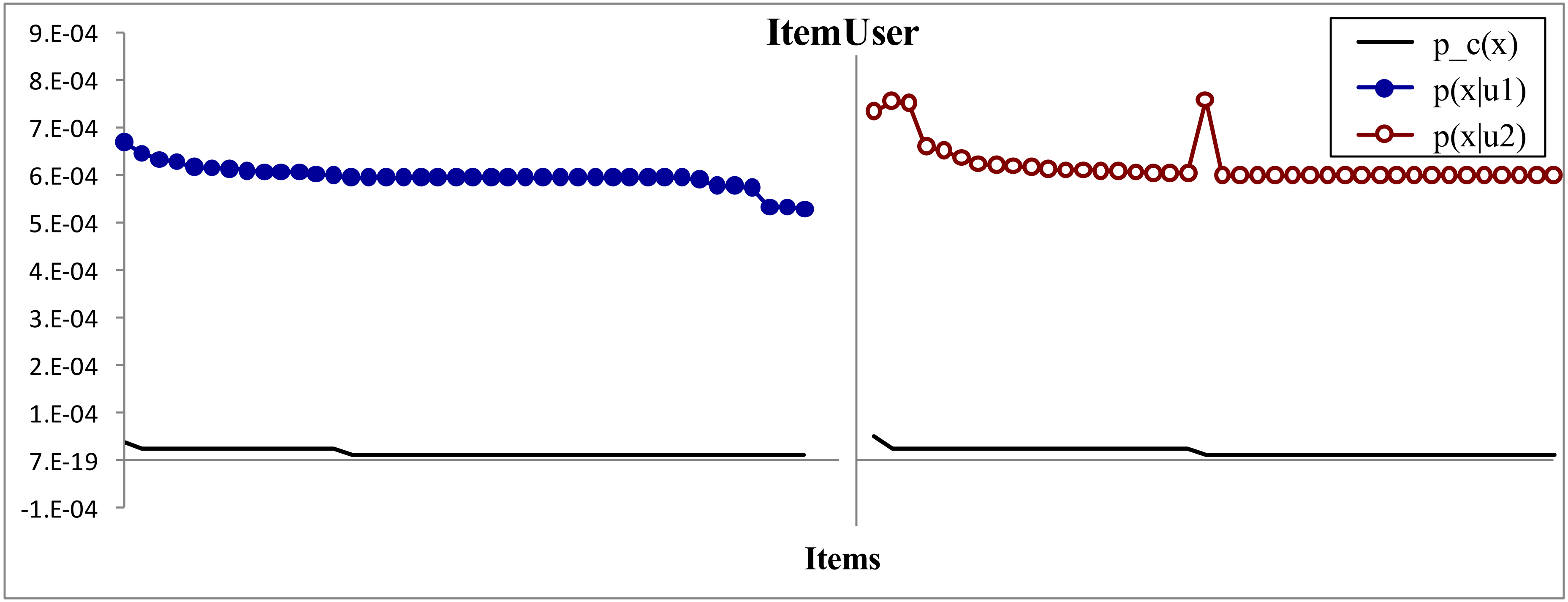
Term contributions for each user, ordered by their corresponding contribution to the user language model. **IRUserItem** clarity model.

Examples



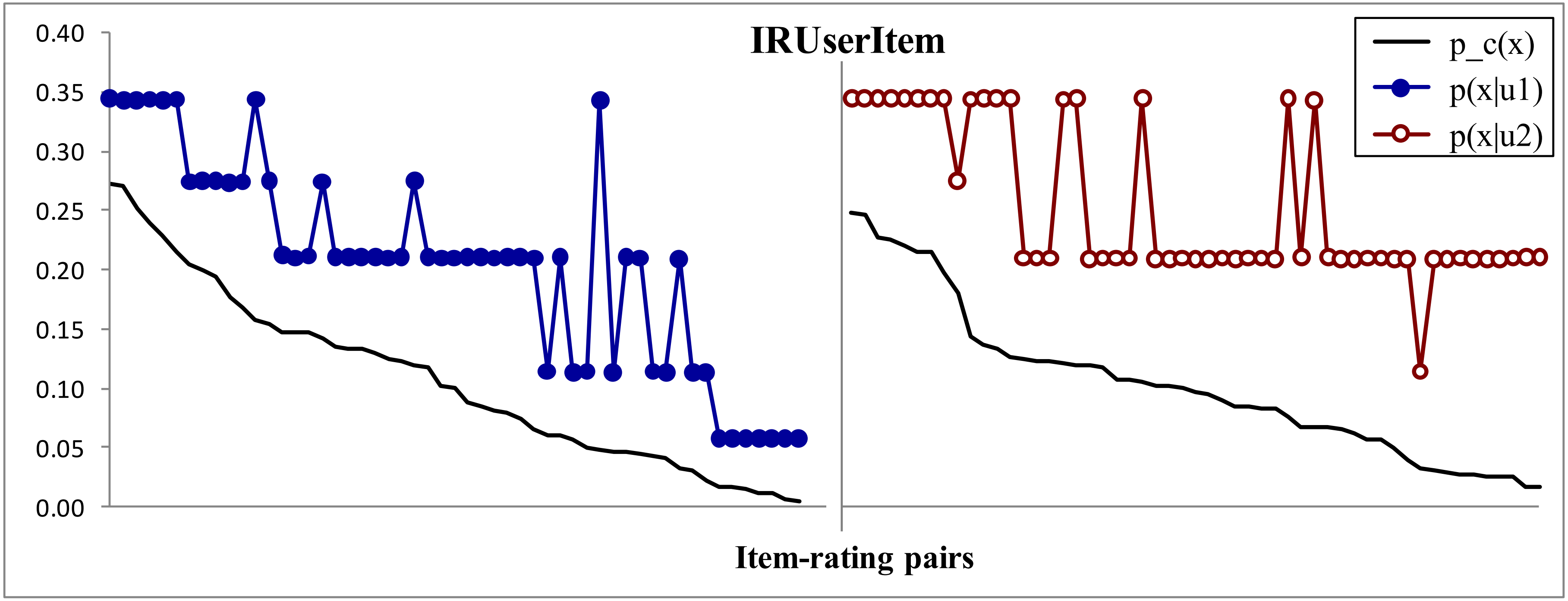
User language model sorted by collection probability

Examples



User language model sorted by collection probability

Examples



User language model sorted by collection probability

Experiments

- Dataset: Movielens 100K, 5-fold
- Recommenders:
  - four collaborative filtering (CF)
  - one content-based (CBF)
- Predictors: seven user clarity variations

- Analyse correlation between predictors and recommender performance
  - Against more than one recommender
  - Pearson / Spearman
  - nDCG / MAP (@N)

Results

- Pearson's correlation wrt. nDCG@50 for different recommenders

Predictor	CBF	IB	TF-L1	TF-L2	UB
ItemSimple	0.257	0.146	0.521	0.564	0.491
ItemUser	0.252	0.188	0.534	0.531	0.483
RatUser	0.234	0.182	0.507	0.516	0.469
RatItem	0.191	0.184	0.442	0.426	0.395
IRUser	0.171	-0.092	0.253	0.399	0.257
IRItem	0.218	0.152	0.453	0.416	0.372
IRUserItem	0.265	0.105	0.523	0.545	0.444

Results

- Pearson's correlation wrt nDCG@50 for different recommenders

Predictor	CBF	IB	TF-L1	TF-L2	UB
ItemSimple	0.257	0.146	0.521	0.564	0.491
ItemUser	0.252	0.188	0.534	0.531	0.483
RatUser	0.234	0.182	0.507	0.516	0.469
RatItem	0.191	0.184	0.442	0.426	0.395
IRUser	0.171	-0.092	0.253	0.399	0.257
IRItem	0.218	0.152	0.453	0.416	0.372
IRUserItem	0.265	0.105	0.523	0.545	0.444

Results

- Pearson's correlation wrt nDCG@50 for different recommenders

Predictor	CBF	IB	TF-L1	TF-L2	UB
ItemSimple	0.257	0.146	0.521	0.564	0.491
ItemUser	0.252	0.188	0.534	0.531	0.483
RatUser	0.234	0.182	0.507	0.516	0.469
RatItem	0.191	0.184	0.442	0.426	0.395
IRUser	0.171	-0.092	0.253	0.399	0.257
IRItem	0.218	0.152	0.453	0.416	0.372
IRUserItem	0.265	0.105	0.523	0.545	0.444
Performance	0.061	0.004	0.093	0.239	0.044

Results

- Performance prediction depends on
  - Actual recommender performance
  - Input sources used by the recommender

- In Information Retrieval, typically:
  - Only one system (or the mean/median of several) is reported
  - Language modelling retrieval systems are used

Conclusions and Future Work

- Adaptations of query clarity predictors in Recommender Systems
- Strong correlation values

- Revision of the grey sheep concept?
- Applications
  - Dynamic neighbour weighting
  - Dynamic adjustment of recommender ensembles
- Additional performance predictors
  - With explicit recommender dependence

Thank you!

Predicting the Performance of Recommender Systems: An Information Theoretic Approach

Alejandro Bellogín, Pablo Castells, Iván Cantador
Escuela Politécnica Superior
Universidad Autónoma de Madrid

@abellogin
alejandro.bellogin@uam.es