Top-N Recommendation Algorithms: A Quest for the State-of-the-Art

ANONYMOUS AUTHOR(S)

SUPPLEMENTAL MATERIAL

1 INTER-METRIC CORRELATIONS - BEYOND ACCURACY METRICS

In this section we report the correlations between each pair of beyond-accuracy metrics. For each dataset, the tables indicate the Pearson product-moment correlation coefficient, unveiling strong direct and inverse correlations. Please note we do not report same analysis for accuracy metrics here, since the topic of correlation among those metrics has been extensively studied in prior literature. Please refer to Valcarce et al. [2], and Anelli et al. [1] for further details.

Table 1. Detailed Metric Correlations. The tables show how much each beyond-accuracy metric (computed on recommendation lists of ten items for each user) correlates with each other. Specifically, the table shows the Pearson product-moment correlation coefficient for each dataset.

Movielens	EFD	Gini	IC	PopREO	PopRSP	ACLT	APLT	ARP
EPC	1.00	-0.73	-0.36	0.75	0.79	-0.79	-0.79	0.20
EFD		-0.74	-0.37	0.77	0.81	-0.80	-0.80	0.23
Gini			0.86	-0.99	-0.99	0.99	0.99	-0.81
IC				-0.86	-0.79	0.80	0.80	-0.95
PopREO					0.99	-0.99	-0.99	0.78
PopRSP						-1.00	-1.00	0.74
ACLT							1.00	-0.74
APLT								-0.74
Amazon	EFD	Gini	IC	PopREO	PopRSP	ACLT	APLT	ARP
EPC	1.00	-0.10	0.45	-0.26	0.09	-0.04	-0.04	-0.46
EFD		-0.05	0.49	-0.32	0.03	0.02	0.02	-0.50
Gini			0.78	-0.85	-0.96	0.88	0.88	-0.69
IC				-0.93	-0.74	0.70	0.70	-0.88
PopREO					0.87	-0.87	-0.87	0.84
PopRSP						-0.97	-0.97	0.60
ACLT							1.00	-0.58
APLT								-0.58
Epinions	EFD	Gini	IC	PopREO	PopRSP	ACLT	APLT	ARP
EPC	0.83	-0.05	-0.11	-0.94	-0.84	0.91	0.91	-0.81
EFD		-0.54	-0.59	-0.62	-1.00	0.97	0.97	-0.67
Gini			1.00	-0.25	0.55	-0.43	-0.43	-0.18
IC				-0.19	0.60	-0.49	-0.49	-0.12
PopREO					0.63	-0.74	-0.74	0.81
PopRSP						-0.99	-0.99	0.64
ACLT							1.00	-0.69
APLT								-0.69

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Manuscript submitted to ACM

^{© 2022} Association for Computing Machinery.

2 F1 SCORES - ADDITIONAL NUMERICAL EXAMPLES

The F1 score represents the harmonic mean of Precision and Recall. In the recommendation domain, when evaluating lists of k items (top-k evaluation), it is usually defined as follows:

$$F1 Score = \frac{1}{|U|} \sum_{u \in U} 2 * \frac{P_u @k * R_u @k}{P_u @k + R_u @k}$$

$$\tag{1}$$

where U is the set of the users in the population, and where Pu@k and Ru@k are the Precision and Recall values for a user u's top-k recommendations, respectively. In an alternative formulation, the F1 Score could be computed *after* obtaining the average Precision and Recall values across all users:

$$P@k = \frac{1}{|U|} \sum_{u \in U} P_u@k \tag{2}$$

$$R@k = \frac{1}{|U|} \sum_{u \in U} R_u@k \tag{3}$$

$$F1 \, Score = 2 * \frac{P@k * R@k}{P@k + R@k} \tag{4}$$

These alternative formulations may lead to different results, as we highlight in the following examples. Let us consider a population of five users for whom we have computed the Precision and Recall values for a recommendation system A (see Table 2a).

Table 2. Accuracy results for the toy recommendation systems. P@k, R@k, and F@k stands for individual Precision, Recall, and F1 Score with a list of k recommendations, respectively. *Average* reports the overall Precision and Recall values. Per-user F1 and average-based F1 indicates the F1 scores computed using Equation 1 and Equation 4, respectively.

Population	$P_u@k$	$R_u@k$	$F_u@k$
$user_0$	0.2	0.3	0.240
$user_1$	0.5	0.6	0.545
$user_2$	0.3	0.4	0.343
$user_3$	0.6	0.3	0.400
$user_4$	0.2	0.3	0.240
	P@k	R@k	F@k
Average	0.36	0.38	
Per-user F1			0.354
Average-based F1			0.370

⁽a) Toy recommendation system A.

_			
Population	$P_u@k$	$R_u@k$	$F_u@k$
$user_0$	0.2	0.4	0.267
$user_1$	0.5	0.2	0.286
$user_2$	0.4	0.4	0.400
user ₃	0.2	0.6	0.300
user ₄	0.5	0.4	0.444
	P@k	R@k	F@k
Average	0.36	0.40	
Per-user F1			0.339
Average-based F1			0.379

(b) Toy recommendation system B.

It is worth noticing that the F1 formulation from Equation 1, denoted as *Per-User F1*, returns an F1 score that is lower than the overall averaged values of Precision and Recall. This can happen due to the product of individual Precision and Recall values. If one of the two is small, it affects the result and impacts the F1 score. Conversely, this behavior is not likely to occur when the F1 is computed on already averaged Precision and Recall values (Average-based F1).

Furthermore, suppose that we evaluate the performance of two recommender systems, A and B (Table 2b). The two systems lead to the same average Precision value, and B leads to a higher Recall value than A. It may now be surprising to see that A has a higher *per-user* F1 score than B. As a consequence of the previously discussed phenomenon, it is indeed possible. That is, although the Precision value of system B is equal to system A, some individual Precision values

lead to poor individual F1 results that affect the overall value of the metric. Some examples of such cases can be found in the accuracy results of the paper.

3 HYPERPARAMETERS RANGE

Table 3. Generated by Spread-LaTeX

Algorithm	Hyperparameter	Range	Type	Distribution
UserKNN,	topK	5 - 1000	Integer	uniform
ItemKNN	similarity	cosine, jaccard, dice, pearson, euclidean	Categorical	
	topK	5 - 1000	Integer	uniform
$RP^3\beta$	alpha	0 - 2	Real	uniform
KF p	beta	0 - 2	Real	uniform
	normalization	True, False	Categorical	
	topK	5 -1000	Integer	uniform
SLIM	l1 ratio	0.00001 - 1	Real	log-uniform
	alpha	0.01 - 1	Real	uniform
$EASE^R$	l2 norm	1 - 10000000	Real	log-uniform
	num factors	8, 16, 32, 64, 128, 256	Integer	
	epochs	30 - 100	Integer	uniform
MF2020	learning rate	0.00001 - 1	Real	log-uniform
	reg	0.00001 - 0.1	Real	log-uniform
	negative sample	4,6,8	Integer	
	num factors	1 - 200	Integer	uniform
	scaling	linear, log	Categorical	
iALS	alpha	0.001 - 50	Real	uniform
	epsilon	0.001 - 10	Real	uniform
	reg	0.001 - 0.01	Real	uniform
	num factors	8, 16, 32, 64, 128, 256	Integer	
	learning rate	0.00001 - 1	Real	log-uniform
BPRMF	batch size	128, 256, 512	Integer	
BPRMF	reg user	0.00001 - 0.1	Real	log-uniform
	reg positive item	0.00001 - 0.1	Real	log-uniform
	reg negative item	0.00001 - 0.1	Real	log-uniform
	num factors	8, 16, 32, 64, 128, 256	Integer	
	epochs	30 - 100	Integer	uniform
NeuMF	learning rate	0.00001 - 1	Real	log-uniform
	batch size	128, 256, 512	Integer	
	negative sample	4,6,8	Integer	
MultiVAE	epochs	100 - 300	Integer	uniform
	learning rate	0.00001 - 1	Real	log-uniform
	batch_size	64, 128, 256	Integer	ū
	intermediate dim	400 - 800	Integer	uniform
	latent dim	100-400	Integer	uniform
	reg	0.00001 - 1	Real	log-uniform

Table 4. Generated by Spread-LaTeX

UserKNN topK similarity 117 correlation cosine cosine cosine ItemNN topK similarity 95 798 137 topK similarity cosine cosine cosine topK alpha 1.4350197 0.4973207 0.8719344 beta 0.3265517 0.2836938 0.2483698 normalization true false true topK 518 663 663 SLIM 11 ratio 0.0000420 0.0000108 0.0000108 alpha 0.2978543 0.0486771 0.0486771 EASE ^K 12 norm 238.5621338 238.5621338 238.5621338 num factors 128 64 16 16 epochs 72 92 97 97 MF2020 learning rate 0.1295965 0.1295965 0.0154435 reg 0.0087583 0.0125009 0.0223642 negative sample 4 8 4 num factors 51 200 178 <th>Algorithm</th> <th>Hyperparameter</th> <th>Movilens</th> <th>Amazon</th> <th>Epinions</th>	Algorithm	Hyperparameter	Movilens	Amazon	Epinions	
ItemNN topK 55 798 137	II INDI	topK	117	226	139	
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	UserKNN	similarity	correlation cosine		cosine	
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	It NINI	topK	95	798	137	
RP³β alpha beta beta 1.4350197 0.4973207 0.8719344 beta 0.3265517 0.2836938 0.2483698 0.2483698 normalization topK 518 663 663 663 663 SLIM I1 ratio 0.0000420 0.0000108 alpha 0.2978543 0.0486771 0.0486771 EASER I2 norm 238.5621338 238.5621338 238.5621338 238.5621338 num factors epochs 72 92 97 97 MF2020 learning rate 0.1295965 0.1295965 0.0154435 reg 0.0087583 0.0125009 0.0223642 negative sample 4 8 4 8 4 4 num factors epochs 27 70 145 scaling log log log alpha 6.3818930 9.1219718 2.8537184 epsilon 5.6496278 0.4921936 2.3098481 reg 0.0494734 0.4921936 0.0411491 num factors 256 64 256 epochs 73 86 63 learning rate 0.0378936 0.1265624 0.1004075 learning rate 0.0378936 0.1265624 0.1004075 learning rate 0.00378936 0.1265624 0.1004075 learning rate 0.005551 0.0052985 0.0034511 reg positive item reg negative item 0.0012779 0.0009577 0.0328127 num factors 16 128 32 epochs 93 100 39 learning rate 0.000366 0.0001365 0.0000465 learning rate 0.000366 0.0001365 0.0000465 learning rate 0.0000366 0.0001365 0.0000465 learning rate 0.0000366 0.0001365 0.0000465 learning rate 0.0000366 0.0001365 0.000000366 learning rate 0.0001545 0.0000723 0.0001003 learning rate batch size 128 128 128 learning rate batch size 128 intermediate dim latent dim 175 279 175 MultiVAE batch size 128 intermediate dim latent dim 175 279 175	ItemININ		cosine	cosine	cosine	
Deta		topK	158	803	144	
beta 0.3265517 0.2836938 0.2483698 normalization true false true topK 518 663 663 663 663 663 Alpha 0.2978543 0.0486771 0.0486771 0.0486771 0.0486771 0.0486771 0.0486771 0.0486771 0.0486771 0.0486771 EASE ^R 12 norm 238.5621338 238.5621	$pp3\rho$	alpha	1.4350197	0.4973207	0.8719344	
SLIM	RP ³ p	beta	0.3265517	0.2836938	0.2483698	
SLIM		normalization	true	false	true	
Alpha 0.2978543 0.0486771 0.0486771		topK	518	663	663	
EASE ^R 12 norm 238.5621338 238.5621338 238.5621338 num factors epochs 72 92 97 MF2020 learning rate reg 0.1295965 0.1295965 0.0154435 reg negative sample regative sample regochs 4 8 4 num factors repochs 51 200 178 regochs epochs regochs 27 70 145 regochs scaling alpha region 6.3818930 9.1219718 2.8537184 regochs epsilon reg 0.0494734 0.4921936 2.3098481 reg reg 0.0494734 0.4921936 0.398481 reg num factors reg 256 64 256 regochs epochs reg user 73 86 63 reguer learning rate reg positive item reg negative item 0.00378936 0.1265624 0.1004075 BPRMF batch size reg positive item reg negative item 0.0012779 0.0005673 0.0034511 reg negative item reg negative item 0.0012779 0.00058673 0.0034511 reg negative item reg negative item reg negative item regochs regochs 93 100 39 NeuMF learning rate pochs rego	SLIM	l1 ratio	0.0000420	0.0000108	0.0000108	
NF2020 NF20200 NF20200 NF20200 NF20200 NF202000 NF2020642 NF20200 NF202000 NF2020642 NF202000 NF2020642 NF20200 NF202000 NF2020642 NF202000 NF2020000 NF2020000000000000000000000000000000000		alpha	0.2978543	0.0486771	0.0486771	
Part	$EASE^R$	l2 norm	238.5621338	238.5621338	238.5621338	
MF2020 learning rate reg 0.1295965 0.1295965 0.0154435 negative sample 4 8 4 num factors 51 200 178 epochs 27 70 145 scaling log log log alpha 6.3818930 9.1219718 2.8537184 epsilon 5.6496278 0.4921936 2.3098481 reg 0.0494734 0.4921936 0.0411491 num factors 256 64 256 epochs 73 86 63 learning rate 0.0378936 0.1265624 0.1004075 BPRMF batch size 256 256 256 reg user 0.0157839 0.0058673 0.0002613 reg positive item 0.0005651 0.0052985 0.0034511 reg negative item 0.0012779 0.0009577 0.0328127 num factors 16 128 32 epochs 93 100 39		num factors	128	64	16	
reg negative sample 4		epochs	72	92	97	
negative sample	MF2020	learning rate	0.1295965	0.1295965	0.0154435	
iALS num factors epochs 51 200 178 epochs iALS scaling alpha log log log log log alpha log log log log log alpha epsilon 2.8537184 epsilon 5.6496278 0.4921936 2.3098481 reg 0.04921936 0.0411491 num factors reg 0.0494734 0.4921936 0.0411491 0.0411491 0.0494734 0.4921936 0.0411491 0.0411491 num factors epochs 73 86 63 learning rate 0.0378936 0.1265624 0.1004075 0.1004075 BPRMF batch size batch size batch size reg positive item reg positive item reg negative item 0.0012779 0.0052985 0.0034511 reg negative item 0.0012779 0.0009577 0.0328127 0.002613 0.000465 0.000465 0.0000465 0.0000465 0.0000465 0.0000465 0.0000465 0.0000465 0.0000465 0.0000465 0.00000465 0.00000000000000000000000000000000000		reg	0.0087583	0.0125009	0.0223642	
iALS epochs 27 70		negative sample	4	8	4	
iALS scaling log log log alpha 6.3818930 9.1219718 2.8537184 epsilon 5.6496278 0.4921936 2.3098481 reg 0.0494734 0.4921936 0.0411491 num factors 256 64 256 epochs 73 86 63 learning rate 0.0378936 0.1265624 0.1004075 learning rate 256 256 256 256 256 reg user 0.0157839 0.0058673 0.0002613 reg positive item 0.0005651 0.0052985 0.0034511 reg negative item 0.0012779 0.0009577 0.0328127 num factors 16 128 32 epochs 93 100 39 NeuMF learning rate 0.0000366 0.0001365 0.0000465 batch size 256 64 256 negative sample 6 6 6 8 epochs 100 205 200 learning rate 0.0001545 0.0000723 0.0001003 MultiVAE batch_size 128 128 128 128 intermediate dim 175 279 175		num factors	51	200	178	
ALS alpha epsilon 5.6496278 0.4921936 2.3098481 reg 0.0494734 0.4921936 0.0411491 num factors 256 epochs 173 86 63 learning rate 0.0378936 0.1265624 0.1004075 BPRMF batch size 256 256 256 256 reg user 0.0157839 0.0058673 0.0002613 reg positive item 0.0012779 0.0009577 0.0328127 num factors 16 128 296 epochs 93 100 39 NeuMF learning rate 0.000366 0.0001365 0.0000465 batch size 256 64 256 64 256 64 256 66 8 epochs 93 100 39 NeuMF learning rate 0.0000366 0.0001365 0.0000465 batch size 256 64 256 negative sample 6 6 6 8 epochs 100 205 200 learning rate 0.0001545 0.0000723 0.0001003 MultiVAE batch_size 128 128 128 128 intermediate dim 175 279 175		epochs	27	70	145	
alpha epsilon 5.6496278 0.4921936 2.3098481 reg 0.0494734 0.4921936 0.0411491 num factors 256 64 256 epochs 73 86 63 learning rate 0.0378936 0.1265624 0.1004075 batch size 256 256 256 256 reg user 0.0157839 0.0058673 0.0002613 reg positive item 0.0005651 0.0052985 0.0034511 reg negative item 0.0012779 0.0009577 0.0328127 num factors 93 100 39 NeuMF learning rate 0.000366 0.0001365 0.0000465 batch size 256 6 8 8 epochs 100 205 200 NeuMF learning rate 0.0001545 0.0000723 0.0001003 batch_size 128 128 128 128 intermediate dim 674 721 674 latent dim 175 279 175	iaic	scaling	log	log	log	
reg 0.0494734 0.4921936 0.0411491 num factors 256 64 256 epochs 73 86 63 learning rate 0.0378936 0.1265624 0.1004075 BPRMF batch size 256 256 256 reg user 0.0157839 0.0058673 0.0002613 reg positive item 0.0005651 0.0052985 0.0034511 reg negative item 0.0012779 0.0009577 0.0328127 num factors 16 128 32 epochs 93 100 39 NeuMF learning rate 0.0000366 0.0001365 0.0000465 batch size 256 64 256 negative sample 6 6 8 epochs 100 205 200 learning rate 0.0001545 0.0000723 0.0001003 batch_size 128 128 128 intermediate dim 674 721 674	IALS	alpha	6.3818930	9.1219718	2.8537184	
num factors 256 64 256 epochs 73 86 63 learning rate 0.0378936 0.1265624 0.1004075 BPRMF batch size 256 256 256 reg user 0.0157839 0.0058673 0.0002613 reg positive item 0.0005651 0.0052985 0.0034511 reg negative item 0.0012779 0.0009577 0.0328127 num factors 16 128 32 epochs 93 100 39 NeuMF learning rate 0.0000366 0.0001365 0.0000465 batch size 256 64 256 negative sample 6 6 8 epochs 100 205 200 learning rate 0.0001545 0.0000723 0.0001003 batch_size 128 128 128 intermediate dim 674 721 674 latent dim 175 279 175		epsilon	5.6496278	0.4921936	2.3098481	
BPRMF epochs learning rate 73 86 63 BPRMF batch size 256 256 256 reg user 0.0157839 0.0058673 0.0002613 reg positive item 0.0005651 0.0052985 0.0034511 reg negative item 0.0012779 0.0009577 0.0328127 num factors 16 128 32 epochs 93 100 39 NeuMF learning rate 0.0000366 0.0001365 0.0000465 batch size 256 64 256 negative sample 6 6 8 epochs 100 205 200 learning rate 0.0001545 0.0000723 0.0001003 batch_size 128 128 128 intermediate dim 674 721 674 latent dim 175 279 175		reg	0.0494734	0.4921936	0.0411491	
BPRMF batch size 256 256 256 256 reg user 0.0157839 0.0058673 0.0002613 reg positive item 0.0005651 0.0052985 0.0034511 reg negative item 0.0012779 0.0009577 0.0328127 num factors 16 128 32 epochs 93 100 39 NeuMF learning rate 0.0000366 0.0001365 0.0000465 batch size 256 64 256 negative sample 6 6 8 epochs 100 205 200 learning rate 0.0001545 0.0000723 0.0001003 batch_size 128 128 128 intermediate dim 674 721 674 latent dim 175 279 175		num factors	256	64	256	
BPRMF batch size 256 256 256 reg user 0.0157839 0.0058673 0.0002613 reg positive item 0.0005651 0.0052985 0.0034511 reg negative item 0.0012779 0.0009577 0.0328127 num factors 16 128 32 epochs 93 100 39 NeuMF learning rate 0.0000366 0.0001365 0.0000465 batch size 256 64 256 negative sample 6 6 8 epochs 100 205 200 learning rate 0.0001545 0.0000723 0.0001003 batch_size 128 128 128 intermediate dim 674 721 674 latent dim 175 279 175		epochs	73	86	63	
reg user 0.0157839 0.0058673 0.0002613 reg positive item 0.0005651 0.0052985 0.0034511 reg negative item 0.0012779 0.0009577 0.0328127 num factors 16 128 32 epochs 93 100 39 NeuMF learning rate 0.0000366 0.0001365 0.0000465 batch size 256 64 256 negative sample 6 6 8 epochs 100 205 200 learning rate 0.0001545 0.0000723 0.0001003 batch_size 128 128 128 intermediate dim 674 721 674 latent dim 175 279 175		learning rate	0.0378936	0.1265624	0.1004075	
reg positive item reg negative item 0.0005651 0.0052985 0.0034511 num factors 16 128 32 epochs 93 100 39 NeuMF learning rate 0.0000366 0.0001365 0.0000465 batch size 256 64 256 negative sample 6 6 8 epochs 100 205 200 learning rate 0.0001545 0.0000723 0.0001003 batch_size 128 128 128 intermediate dim 674 721 674 latent dim 175 279 175	BPRMF	batch size	256	256	256	
reg negative item 0.0012779 0.0009577 0.0328127 num factors 16 128 32 epochs 93 100 39 NeuMF learning rate 0.0000366 0.0001365 0.0000465 batch size 256 64 256 negative sample 6 6 8 epochs 100 205 200 learning rate 0.0001545 0.0000723 0.0001003 MultiVAE batch_size 128 128 128 intermediate dim 674 721 674 latent dim 175 279 175			0.0157839	0.0058673	0.0002613	
NeuMF 16 128 32 epochs 93 100 39 NeuMF learning rate 0.0000366 0.0001365 0.0000465 batch size 256 64 256 negative sample 6 6 8 epochs 100 205 200 learning rate 0.0001545 0.0000723 0.0001003 batch_size 128 128 128 intermediate dim 674 721 674 latent dim 175 279 175		reg positive item	0.0005651	0.0052985	0.0034511	
NeuMF epochs learning rate 93 100 39 NeuMF learning rate 0.0000366 0.0001365 0.0000465 batch size 256 64 256 negative sample 6 6 8 epochs 100 205 200 learning rate 0.0001545 0.0000723 0.0001003 batch_size 128 128 128 intermediate dim 674 721 674 latent dim 175 279 175			0.0012779	0.0009577	0.0328127	
NeuMF learning rate batch size regative sample 0.0000366 0.0001365 0.0000465 batch size regative sample 6 6 8 epochs learning rate 0.0001545 0.0000723 0.0001003 batch_size intermediate dim latent dim 674 721 674 latent dim 175 279 175	NeuMF		16	128	32	
batch size 256 64 256 negative sample 6 6 8 epochs 100 205 200 learning rate 0.0001545 0.0000723 0.0001003 batch_size 128 128 128 intermediate dim 674 721 674 latent dim 175 279 175						
negative sample 6 6 8 epochs 100 205 200 learning rate 0.0001545 0.0000723 0.0001003 batch_size 128 128 128 intermediate dim 674 721 674 latent dim 175 279 175						
MultiVAE epochs learning rate batch_size intermediate dim latent dim 100 100 205 10000723 200 00001003 00001003 MultiVAE batch_size 128 128 128 128 128 128 674 674 121 674				~ -		
MultiVAE learning rate 0.0001545 0.0000723 0.0001003 batch_size 128 128 128 intermediate dim 175 279 175						
MultiVAE batch_size intermediate dim latent dim 128 128 128 128 721 674 129 175 279 175	MultiVAE					
MultiVAE intermediate dim 674 721 674 1279 175 1						
intermediate dim 674 721 674 latent dim 175 279 175		_				
reg 0.0000105 0.1153400 0.0020018		latent dim		=		
		reg	0.0000105	0.1153400	0.0020018	

REFERENCES

^[1] Vito Walter Anelli, Tommaso Di Noia, Eugenio Di Sciascio, Claudio Pomo, and Azzurra Ragone. 2019. On the discriminative power of hyper-parameters in cross-validation and how to choose them. In *Proceedings of the 13th ACM Conference on Recommender Systems, RecSys 2019, Copenhagen, Denmark, September 16-20, 2019,* Toine Bogers, Alan Said, Peter Brusilovsky, and Domonkos Tikk (Eds.). ACM, 447–451.

[2] Daniel Valcarce, Alejandro Bellogín, Javier Parapar, and Pablo Castells. 2018. On the robustness and discriminative power of information retrieval metrics for top-N recommendation. In *Proceedings of the 12th ACM Conference on Recommender Systems, RecSys 2018.* 260–268.