On the discriminative power of Hyper-parameters in Cross-Validation and how to choose them

Vito Walter Anelli[1], Tommaso Di Noia[1], Eugenio Di Sciascio[1],
Claudio Pomo[1], Azzurra Ragone[2]

[1]Politecnico di Bari
Bari, Italy
firstname.lastname@poliba.it

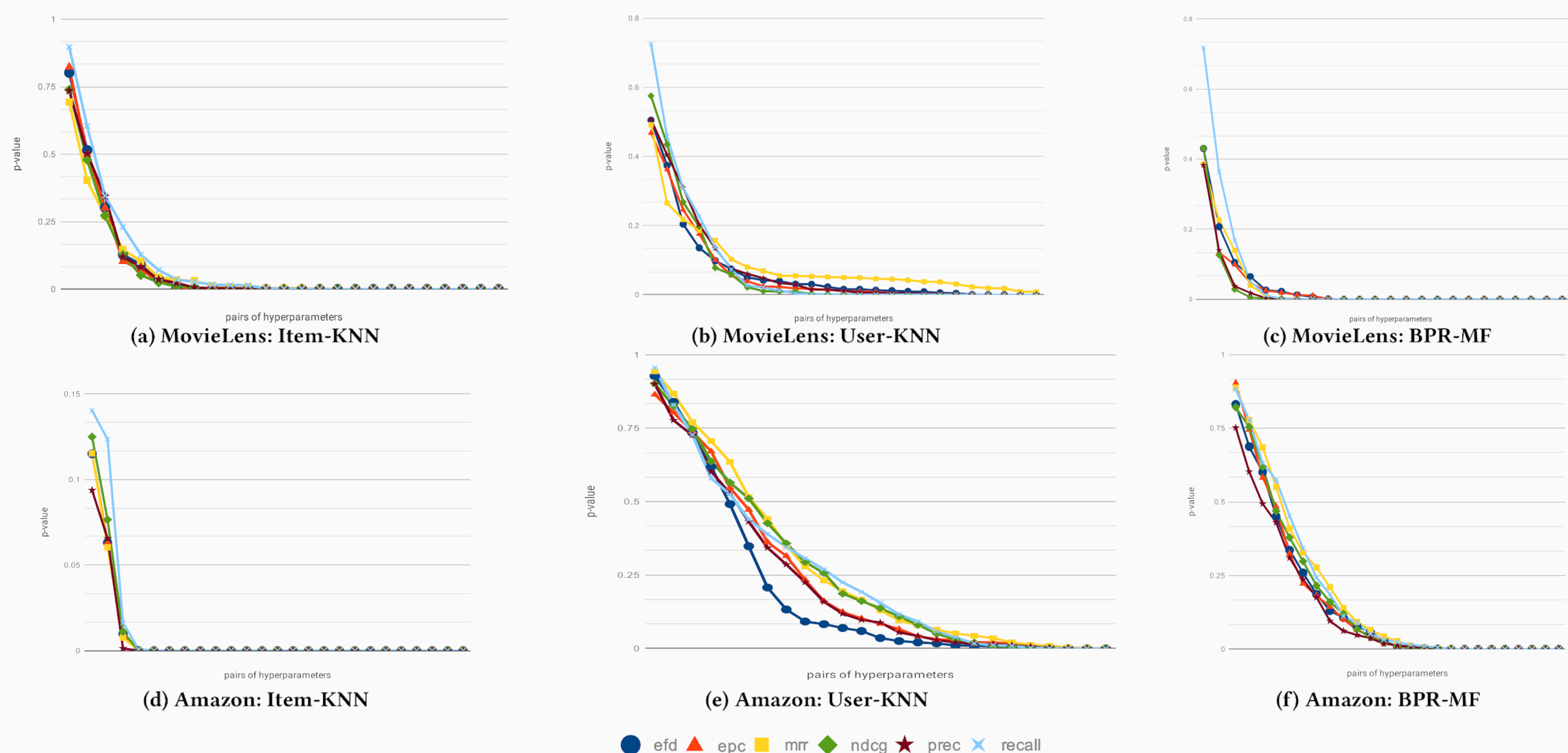[2]Independent Researcher
Milan, Italy
azzurra.ragone@gmail.com

Introduction

Hyper-parameters tuning is a crucial task to make a model perform at its best. We executed a extensive 10-Folds Cross-Validation on MovieLens and Amazon Movies for three well-known baselines: User-kNN, Item-kNN, BPR-MF. By means of a grid search strategy, we investigated the discriminative power of nDCG, Precision, Recall, MRR, EFD, EPC, and, finally, we analyzed the role of parameters on models' evaluation for Cross-Validation.

Contributions

• a study on the discriminative power of accuracy and novelty metrics for the k-Folds Cross-Validation hyperpsarameters tuning for three well-known collaborative models;

• a procedure to check if variations in only one of the parameters are particularly relevant for accuracy of recommendation;

• a study on the impact of "Number of latent factors/iterations", and "learning rate" on accuracy of recommendation for BPR-MF.

Discriminative Power

(a) MovieLens: Item-KNN

(b) MovieLens: User-KNN

(c) MovieLens: BPR-MF

(d) Amazon: Item-KNN

(e) Amazon: User-KNN

(f) Amazon: BPR-MF

● efd ▲ epc ■ mrr ◆ ndcg ★ prec ✕ recall

Parameters Impact

	MovieLens															Amazon				
Latent factors	10	14	20	28	40	56	80	113	160	226	320	452	640	905	1809	113	160	226	320	452
DP	0.703	0.664	0.602	0.512	0.592	0.707	0.575	0.535	0.635	0.639	0.572	0.499	0.303	0.565	**0.290**	2.775	2.681	2.348	**2.127**	2.542
Iterations	1	2	3	4	6	8	11	16	23	32	45	64	91	128		64	92	128		
DP	1.633	0.796	**0.707**	0.939	0.978	0.781	1.037	1.338	1.045	1.210	1.222	1.196	1.008	0.939		4.106	3.299	**2.357**		
Learning rate	0.20000	0.10000	0.05000	0.02500	0.01250	0.00625	0.00312	0.00156	0.00078	0.00039	0.00019	0.00009	0.00004	0.00002	0.00001	0.20000	0.10000	0.05000		
DP	0.765	**0.702**	1.918	2.667	2.131	2.429	2.811	1.552	1.044	0.803	1.083	1.225	0.950	2.335	3.548	3.941	**3.423**	4.803		

Conslusions

• nDCG@N and Precision@N represent a good choice for hyper-parameters tuning for Neighborhood-based models and BPR-MF;
• novelty metrics show also good *Discriminative Power* values suggesting that these metrics are very sensitive to changes in hyper-parameters values;
• in BPR-MF the number of latent factors is dominant with respect to the learning rate and the number of iterations.

Future Work

• explore other datasets;

• investigate different alghorithms in collaborative filtering scenarios;

• add side information to the items;

• exploit a different search algorithm and compare results in terms of performace and final reccomandation quality.