

Non-Compensatory Psychological Models for Recommender Systems

Chen Lin, Xiaolin Shen, Si Chen¹
Muhua Zhu, Yanghua Xiao²

¹Xiamen University

²Alibaba Group

18-Dec-2018



Outline

① Compensatory Models

Background

Decision Process

② Non-Compensatory Models

Motivation

Rating Models

Ranking Models

③ Experimental Analysis

Rating Prediction

Ranking for Explicit Feedbacks

Ranking for Implicit Feedbacks

Research Question

- Q1: How do we explain existing recommendation models from a psychological perspective?
- Q2: How can we develop explainable and accurate recommendation models that operate differently from existing models?

Compensatory Models

Rating Models

- Goal: reconstruct ratings

$$\min_{\Theta=\{U,V\}} \mathcal{L}(\Theta) = \sum_{u,q} (\mathbf{X}_{u,q} - \hat{\mathbf{X}}_{u,q})^2 + \lambda_U \|\mathbf{U}\|_2^2 + \lambda_V \|\mathbf{V}\|_2^2 \quad (1)$$

- MF:

$$\hat{\mathbf{X}}_{u,q} = \sum_{k=1}^K \mathbf{q}_k \mathbf{u}_k \quad (2)$$

- AMF

$$\hat{\mathbf{X}}_{u,q} = \sum_{k=1}^K \mathbf{q}_k \left(\sum_{\mathbf{p} \in R(u)} \mathbf{p}_k / \sqrt{|R(u)|} \right), \quad (3)$$

- LLORMA

$$\hat{\mathbf{X}}_{u,q} = \sum_{t=1}^S \sum_{k=1}^K \mathbf{u}_{t,k} \frac{K((\mathbf{u}_t, \mathbf{i}_t), (\mathbf{u}, \mathbf{q}))}{\sum_{s=1}^S K((\mathbf{u}_s, \mathbf{i}_s), (\mathbf{u}, \mathbf{q}))} \mathbf{q}_{t,k}$$

Ranking Models

- Goal: reconstruct pairwise rankings $p \succ_u q$

$$\min_{\Theta=\{U,V\}} \mathcal{L}(\Theta) = -\sum_u \sum_{p,q} I(p \succ_u q) \log Pr(p \succ_u q) \quad (5)$$

$$+ \lambda_U \| \mathbf{U} \|_2^2 + \lambda_V \| \mathbf{V} \|_2^2,$$

- Thurstone Model

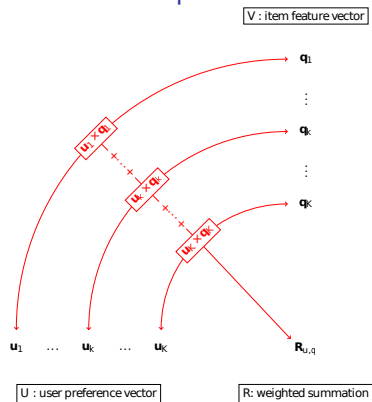
$$Pr(p \succ_u q) = \frac{1}{1 + \exp[-(\hat{\mathbf{X}}_{u,p} - \hat{\mathbf{X}}_{u,q})]} \quad (6)$$

- Bradley-Terry Model

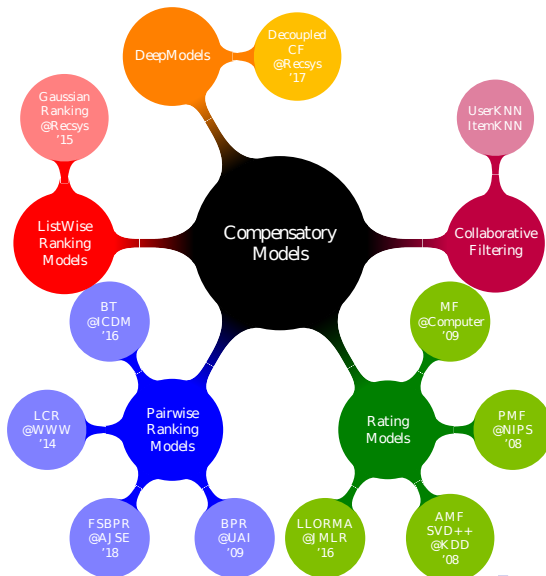
$$Pr(p \succ_u q) = \frac{\hat{\mathbf{X}}_{u,p}}{\hat{\mathbf{X}}_{u,p} + \hat{\mathbf{X}}_{u,q}},$$

Compensatory Decision Rule

The decision rule is compensatory because a good performance on one aspect compensates for bad performances on other aspects.



Summary: Related Work



Non-Compensatory Models

Non-Compensatory Decision Rules

- Lexicographic Rules: order aspects in terms of importance, evaluate items sequentially based on performances
- Conjunctive Rules: items must satisfy a cut-off point on each aspect
- Disjunctive Rules: eliminate items that do not satisfy requirement on an important aspect

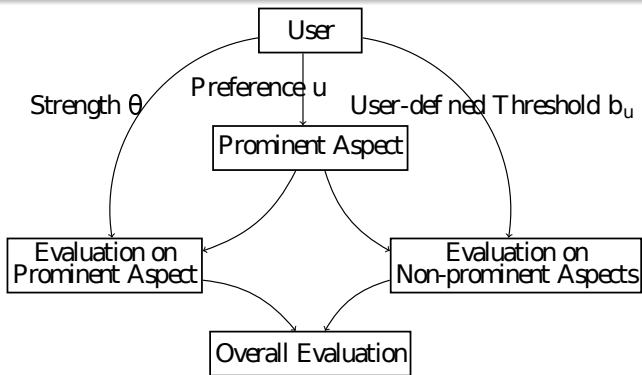
Table: A cellphone example to illustrate non-compensatory rules.

Item	Prominent aspect	Non-prominent aspects	
	Battery life	Price	Memory
iPhone SE	13 hours	700\$	64GB
Galaxy S8	9 hours	500\$	128GB
Honor 10	24 hours	589\$	128GB

Conceptual Model

Challenge

Turn symbolic rules to a data-driven learning framework



Non-Compensatory Rating Predictions

Combination of Lexicographic Rules and Conjunctive Rules

$$\hat{\mathbf{X}}_{u,q} = \sum_{k=1}^K \frac{\exp \mathbf{u}_k}{\sum_{k'} \exp \mathbf{u}_{k'}} [\mathbf{q}_k \exp \theta + \sum_{k' \neq k} (\mathbf{q}_{k'} - \mathbf{b}_{u,k'})]$$

Non-Compensatory Rating Predictions

Combination of Lexicographic Rules and Conjunctive Rules

$$\hat{\mathbf{X}}_{u,q} = \sum_{k=1}^K \frac{\exp \mathbf{u}_k}{\sum_{k'} \exp \mathbf{u}_{k'}} [\mathbf{q}_k \exp \theta + \sum_{k' \neq k} (\mathbf{q}_{k'} - \mathbf{b}_{u,k'})]$$

Conjunctive rule,
cut-off: $b_{u,k'}$

Lexicographic rule

Choice of prominent aspect

Realization to Rating Models

- Goal: reconstruct ratings

$$\min_{\Theta=\{U,V,\theta,b\}} \mathcal{L}(\Theta) = \sum_{u,q} (\mathbf{X}_{u,q} - \hat{\mathbf{X}}_{u,q})^2 + \lambda_U \|\mathbf{U}\|_2^2 + \lambda_V \|\mathbf{V}\|_2^2 \quad (8)$$

- MF-N:

$$\hat{\mathbf{X}}_{u,q} = \sum_{k=1}^K \frac{\exp \mathbf{u}_k}{\sum_{k'} \exp \mathbf{u}_{k'}} [\mathbf{q}_k \exp \theta + \sum_{k' \neq k} (\mathbf{q}_{k'} - \mathbf{b}_{u,k'})] \quad (9)$$

- AMF-N

$$\hat{\mathbf{X}}_{u,q} = \sum_{k=1}^K \frac{\exp(\sum_{\mathbf{p} \in R(u)} \mathbf{p}_k)}{\sum_{k'} \exp(\sum_{\mathbf{p} \in R(u)} \mathbf{p}_{k'})} [\mathbf{q}_k \exp \theta + \sum_{k' \neq k} (\mathbf{q}_{k'} - \mathbf{b}_{u,k'})]. \quad (10)$$

- LLORMA-N

$$\hat{\mathbf{X}}_{u,q} = \sum_{t=1}^S \sum_k \frac{\exp \mathbf{u}_k}{\sum_{k'} \exp \mathbf{u}_{k'}} \frac{K((\mathbf{u}_t, \mathbf{i}_t), (\mathbf{u}, \mathbf{q}))}{\sum_{s=1}^S K((\mathbf{u}_s, \mathbf{i}_s), (\mathbf{u}, \mathbf{q}))} [\mathbf{q}_{t,k} \exp \theta + \sum_{k' \neq k} (\mathbf{q}_{t,k'} - \mathbf{b}_{u,k'})]$$

Non-Compensatory Thurstone Model

$$Pr(p \succ_u q) = \frac{1}{1 + \exp[-(\hat{\mathbf{X}}_{u,p} - \hat{\mathbf{X}}_{u,q})]} \quad (12)$$

Inference

- Goal: reconstruct rankings

$$\begin{aligned} \min_{\Theta=\{U,V\}} \mathcal{L}(\Theta) = & -\sum_u \sum_{p,q} I(p \succ_u q) \log Pr(p \succ_u q) \\ & + \lambda_U \|\mathbf{U}\|_2^2 + \lambda_V \|\mathbf{V}\|_2^2, \end{aligned} \quad (13)$$

- Stochastic Gradient Descent

- $\frac{\partial \mathcal{L}}{\partial \Theta} = \frac{\partial \mathcal{L}'}{\partial \Theta} + \lambda_U \mathbf{U} + \lambda_V \mathbf{V}$
- $\frac{\partial \mathcal{L}'}{\partial \Theta} = \sum_u \sum_{p \succ_u q} \frac{\partial \mathcal{L}'}{\partial \Delta \hat{\mathbf{X}}_{u,p,q}} \frac{\partial \Delta \hat{\mathbf{X}}_{u,p,q}}{\partial \Theta}$, where $\Delta \hat{\mathbf{X}}_{u,p,q} = \hat{\mathbf{X}}_{u,p} - \hat{\mathbf{X}}_{u,q}$

Non-Compensatory Bradley Terry Model

Combination of Lexicographic Rules and Conjunctive Rules

$$Pr(p \succ_u q) = \sum_{k=1}^K u_k \left[\frac{\mathbf{p}_k}{\mathbf{p}_k + \theta \mathbf{q}_k} \prod_{k' \neq k} \frac{\theta \mathbf{p}_{k'}}{\mathbf{q}_{k'} + \theta \mathbf{p}_{k'}} \right]$$

Non-Compensatory Bradley Terry Model

Combination of Lexicographic Rules and Conjunctive Rules

$$Pr(p >_u q) = \sum_{k=1}^K \mathbf{u}_k \left[\frac{\mathbf{p}_k}{\mathbf{p}_k + \theta \mathbf{q}_k} \prod_{k' \neq k} \frac{\theta \mathbf{p}_{k'}}{\mathbf{q}_{k'} + \theta \mathbf{p}_{k'}} \right]$$

Choice of prominent aspect

Conjunctive rule
cut-off: $q_{k'}/\theta$

Lexicographic rule, $\mathbf{p}_k > \theta \mathbf{q}_k, \theta > 1$

Inference

- Stochastic EM
- Sample k

$$k \sim u_k^t \frac{\mathbf{p}_k^t}{\mathbf{p}_k^t + \theta^t \mathbf{q}_k^t} \prod_{k' \neq k} \left[\frac{\theta^t \mathbf{p}_{k'}^t}{\mathbf{q}_{k'}^t + \theta^t \mathbf{p}_{k'}^t} \right]. \quad (14)$$

- MM-bound

$$\mathcal{L} \geq \sum_u \sum_{d \in D(u)} \left\{ \log u_k + \left[1 - \frac{w_k + \theta l_k}{w_k^t + \theta^t l_k^t} + \log \frac{w_k}{w_k^t + \theta^t l_k^t} \right] \right. \\ \left. \sum_{k' \neq k} \left[1 - \frac{l_{k'} + \theta w_{k'}}{l_{k'}^t + \theta^t w_{k'}^t} + \log \frac{\theta w_{k'}}{l_{k'}^t + \theta^t w_{k'}^t} \right] \right\} \quad (15)$$

Experimental Analysis

Data Sets

Table: Statistics of Datasets with ratings

Dataset	#users	#items	#ratings	#pairs
Movielens	942	1,650	80,000	1,072,237
Filmtrust	1,235	2,062	35,497	623,516
CiaoDVD	2,665	14,280	72,665	2,478,836

Results

Method	AUC	Imp.	NDCG	Imp.	MRR	Imp.
MovieLens		(%)		(%)		(%)
MF	0.6729		0.6925		0.8300	
MF-N	0.7108	5.62	0.7166	3.48	0.8633	4.01
AMF	0.6901		0.7107		0.8747	
AMF-N	0.7027	1.83	0.7138	0.44	0.8790	0.49
LLORMA	0.7265		0.8734		0.7015	
LLORMA-N	0.7299	0.47	0.8999	3.03	0.7187	2.45
Filmtrust		(%)		(%)		(%)
MF	0.6507		0.5229		0.7011	
MF-N	0.6710	3.12	0.5241	0.23	0.7071	0.86
AMF	0.5971		0.5137		0.7411	
AMF-N	0.6133	2.71	0.5253	2.25	0.7619	2.80
LLORMA	0.6240		0.8596		0.7857	
LLORMA-N	0.6345	1.68	0.8684	1.02	0.8068	2.69
CiaoDVD		(%)		(%)		(%)
MF	0.7431		0.7949		0.8910	
MF-N	0.7903	6.34	0.8127	2.25	0.9154	2.74
AMF	0.6489		0.6612		0.8741	
AMF-N	0.6993	7.77	0.6878	4.02	0.8967	2.58
LLORMA	0.6752		0.7827		0.8267	
LLORMA-N	0.6845	1.38	0.7984	2.00	0.8384	1.42

Results

Method	AUC	Imp.	NDCG	Imp.	MRR	Imp.
Movielens		(%)		(%)		(%)
BT	0.6453		0.5329		0.8227	
BT-N	0.8511	31.89	0.5795	8.74	0.9256	12.51
BPR	0.7976		0.5674		0.8988	
BPR-N	0.8361	4.82	0.5761	1.53	0.9180	2.14
FSBPR	0.5048		0.5011		0.7524	
FSBPR-N	0.8272	63.86	0.5740	14.56	0.9136	21.42
LCR	0.7191		0.8555		0.9461	
LCR-N	0.7360	2.35	0.8605	0.58	0.9515	0.57
Filmtrust		(%)		(%)		(%)
BT	0.5405		0.5092		0.7702	
BT-N	0.6969	28.94	0.5446	6.95	0.8485	10.15
BPR	0.6412		0.5319		0.8206	
BPR-N	0.6729	4.94	0.5391	1.35	0.8364	1.93
FSBPR	0.4857		0.4968		0.7428	
FSBPR-N	0.6717	38.29	0.5388	8.47	0.8358	12.52
LCR	0.5977		0.9034		0.7511	
LCR-N	0.6144	2.79	0.9063	0.32	0.7635	1.65
CiaoDVD		(%)		(%)		(%)
BT	0.6063		0.5240		0.8031	
BT-N	0.9334	53.95	0.5981	14.1	0.9666	20.36
BPR	0.6344		0.5304		0.8172	
BPR-N	0.8987	41.66	0.5902	11.28	0.9493	16.17
FSBPR	0.7537		0.5574		0.8769	
FSBPR-N	0.8992	19.30	0.5903	5.91	0.9496	8.30
LCR	0.6260		0.9408		0.7889	
LCR-N	0.6349	1.42	0.9451	0.46	0.7988	1.25

Data Sets

Table: Statistics of Datasets with graded implicit feedback

Dataset	#users	#items	#pairs	#sessions
Tmall-single	33,815	176,231	5,682,833	364,844
Tmall-hybrid	62,101	198,344	6,072,061	475,503
Yoochoose	341,396	30,852	3,044,572	341,396

Results

Method	AUC	Imp.(%)	NDCG	Imp.(%)	MRR	Imp.(%)	MAP	Imp.(%)	Prec	Imp.(%)
Tmall-single										
BT	0.5304		0.2804		0.4870		0.4327		0.2778	
BT-N	0.5400	1.82	0.2840	1.28	0.4948	1.61	0.4386	1.34	0.2801	0.84
BPR	0.5181		0.2794		0.4854		0.4297		0.2767	
BPR-N	0.5349	3.24	0.2848	1.92	0.4960	2.18	0.4401	2.41	0.2806	1.41
FSBPR	0.5265		0.2824		0.4913		0.4350		0.2794	
FSBPR-N	0.5389	2.35	0.2863	1.39	0.4988	1.53	0.4432	1.90	0.2818	0.87
LCR	0.5200		0.8190		0.4277		0.3568		0.2534	
LCR-N	0.5290	1.73	0.8213	0.28	0.4360	1.94	0.3648	2.24	0.2586	2.05
Tmall-hybrid										
BT	0.5867		0.3015		0.5373		0.4929		0.2904	
BT-N	0.6568	11.94	0.3279	8.75	0.5990	11.48	0.5527	12.13	0.3036	4.53
BPR	0.6183		0.3183		0.5792		0.5318		0.2973	
BPR-N	0.6460	4.48	0.3276	2.92	0.5990	3.41	0.5524	3.87	0.3030	1.94
FSBPR	0.6334		0.3246		0.5916		0.5442		0.3026	
FSBPR-N	0.6544	3.31	0.3309	1.94	0.6062	2.48	0.5603	2.95	0.3047	0.69
LCR	0.5398		0.6644		0.4519		0.3745		0.2597	
LCR-N	0.5649	4.65	0.6790	2.20	0.4809	6.42	0.3988	6.49	0.2720	4.74
Yoochoose										
BT	0.6027		0.4734		0.7151		0.6361		0.4560	
BT-N	0.7000	16.15	0.5160	8.99	0.7869	10.04	0.7084	11.37	0.4785	4.92
BPR	0.6700		0.5065		0.7713		0.6895		0.4737	
BPR-N	0.6920	3.28	0.5131	1.31	0.7812	1.29	0.7027	1.91	0.4771	0.74
FSBPR	0.3272		0.3658		0.5062		0.4599		0.4006	
FSBPR-N	0.6198	89.45	0.4822	31.83	0.7169	41.62	0.6448	40.22	0.4650	16.08
LCR	0.5842		0.9725		0.8009		0.7934		0.7677	2.66
LCR-N	0.6315	8.10	0.9754	0.30	0.8231	2.77	0.8161	2.86	0.7881	2.66

Parameters

Table: AUC improvements with \mathbf{b} activated, scale of parameters $\mathbf{b}_{u,k}$ and θ .

Dataset	Imp.(%)	$\sigma(\mathbf{b}_u)$	θ
Movielens	5.37	0.0095 ± 0.0024	0.608 ± 0.105
FilmTrust	2.21	0.0095 ± 0.0023	0.667 ± 0.016
CiaoDVD	28.97	0.0093 ± 0.0022	0.773 ± 0.051

Summary

- Q1: How do we explain existing recommendation models from a psychological perspective?
- A1: A large body of existing recommendation models are based on compensatory decision rules.
- Q2: How can we develop explainable and accurate recommendation models that operate differently from existing models?
- A2: Applying non-Compensatory rules universally improve recommendation performances.

Paper and Codes

- Paper: Chen Lin et.al. Non-Compensatory Psychological Models for Recommender Systems. To appear in AAAI 2019
- Codes are available at <https://github.com/XMUDM/Non-Compensatory>