Counterfactual Review-based Recommendation

Advisor : Jia-Ling, Koh
Speaker : Hsiao-Ting Huang
Source : CIKM'21
Date : 2022/08/16

Outline

- Introduction
- Method
- Experiment
- Conclusion

Recommendation

- **Matrix Facorization**

	i_1	i_2	i_3	i_4	i_5
u_1	1.4	X	1.1	0.7	X
u_2	X	0.3	X	0.7	0.5
u_3	0.4	0.3	X	X	0.3
u_4	1.4	X	1.2	X	0.8

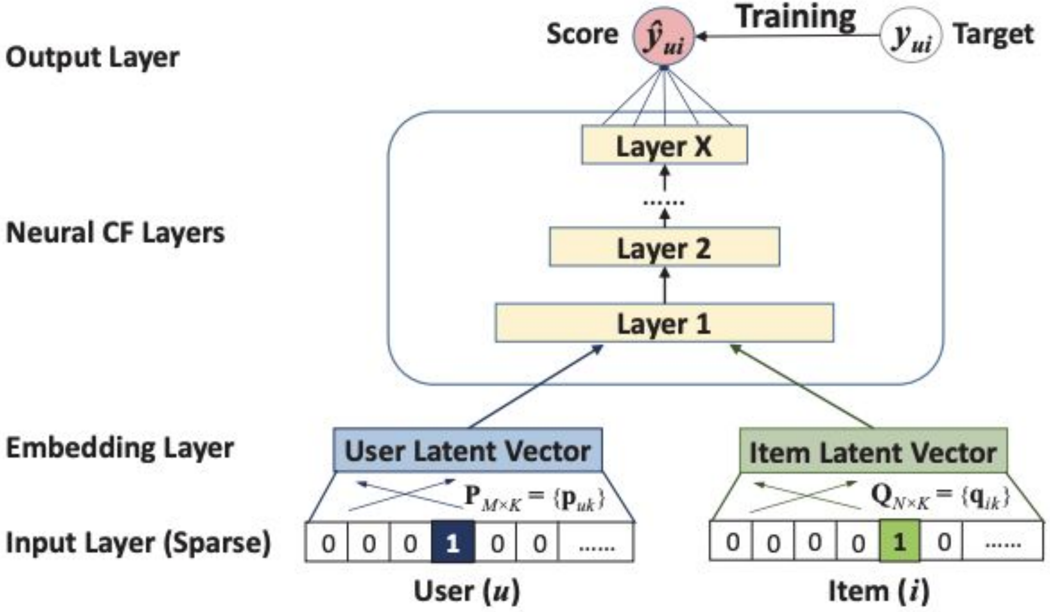
	f_1	f_2
u_1	0.8	0.6
u_2	0.9	0.1
u_3	0.1	0.3
u_4	0.9	0.5

X

	i_1	i_2	i_3	i_4	i_5
f_1	1.0	0.2	1.0	0.8	0.4
f_2	1.0	1.0	0.5	0.1	0.9

Recommendation

- NCF



Recommendation

- **Neural matrix factorization model**
    - Fusion of GMF and MLP

$$\hat{y}_{ui} = a_{out}(\mathbf{h}^T(\mathbf{p}_u \odot \mathbf{q}_i))$$

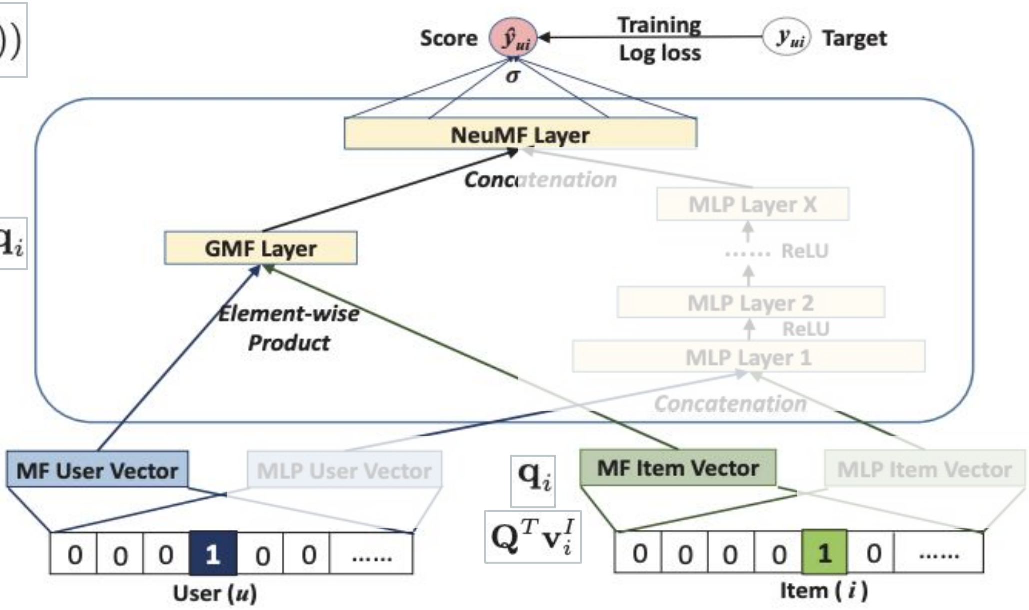
$$\phi_1(\mathbf{p}_u, \mathbf{q}_i) = \mathbf{p}_u \odot \mathbf{q}_i$$

$\mathbf{p}_u$

$\mathbf{P}^T\mathbf{v}_u^U$

$\mathbf{q}_i$

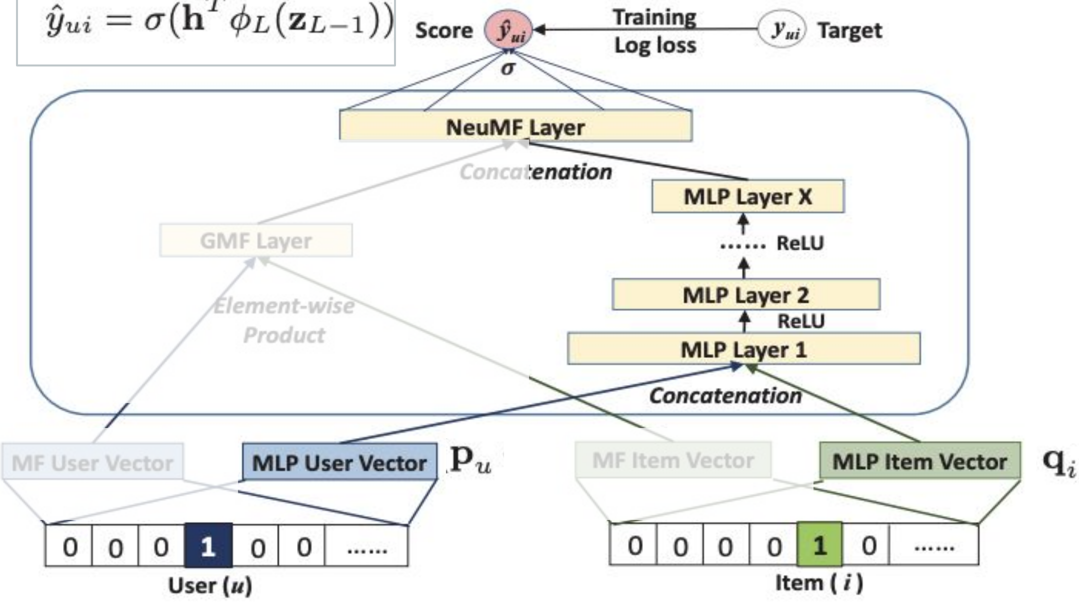
$\mathbf{Q}^T\mathbf{v}_i^I$



Recommendation

- **Neural matrix factorization model**
  - Fusion of GMF and MLP

$$\hat{y}_{ui} = \sigma(\mathbf{h}^T \phi_L(\mathbf{z}_{L-1}))$$



$$\phi_L(\mathbf{z}_{L-1}) = a_L(\mathbf{W}_L^T \mathbf{z}_{L-1} + \mathbf{b}_L)$$

......

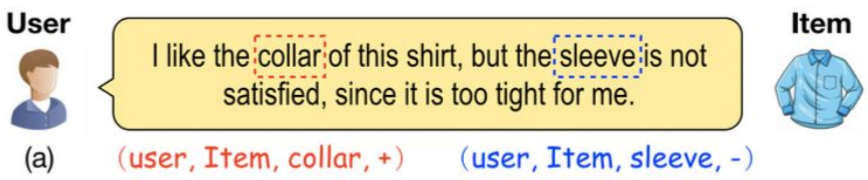
$$\phi_2(\mathbf{z}_1) = a_2(\mathbf{W}_2^T \mathbf{z}_1 + \mathbf{b}_2)$$

$$\mathbf{z}_1 = \phi_1(\mathbf{p}_u, \mathbf{q}_i) = \begin{bmatrix} \mathbf{p}_u \\ \mathbf{q}_i \end{bmatrix}$$

review-based Recommendation

- process the review information on the document level
  - All the review contents are squeezed into an embedding vector to improve the user or item representation

- feature-aware recommendation
  - utilize the review information by extracting user feature-level preferences



review

I like the collar of this shirt,
but the sleeve is not satisfied, since it is too tight for me.

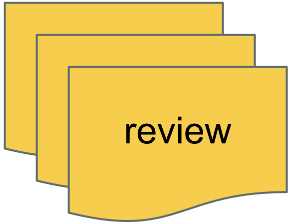
Sentires →

(u, i, feature, sentiment)

(userA, shirt, collar, positive)
(userA, shirt, sleeve, negative)

Input

- **user-feature attention matrix $\boldsymbol{A}$**

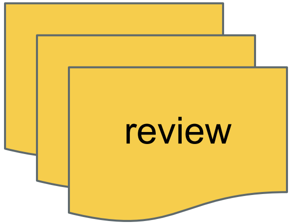


(userA, shirt, collar, +)
(userA, shirt, sleeve, -)
(userA, coat, collar, -)
(userA, sweater, collar, -)

$$\begin{array}{cc} & \begin{array}{ccc}\text{collar} & \text{sleeve} & \cdots\end{array} \\ \begin{array}{c}u_1\\u_2\\u_3\end{array} & \underbrace{\begin{bmatrix} \boxed{3/3 \quad 1/3 \quad \cdots} \\ \\ \\ \end{bmatrix}}_{F} \end{array} \quad \boldsymbol{A}_u \in \mathbb{R}^{1\times|\mathcal{F}|}$$

$$\boldsymbol{A} = [\boldsymbol{A}_u] \in \mathbb{R}^{|\mathcal{U}|\times|\mathcal{F}|}$$

- **item-feature quality matrix $\boldsymbol{B}$**



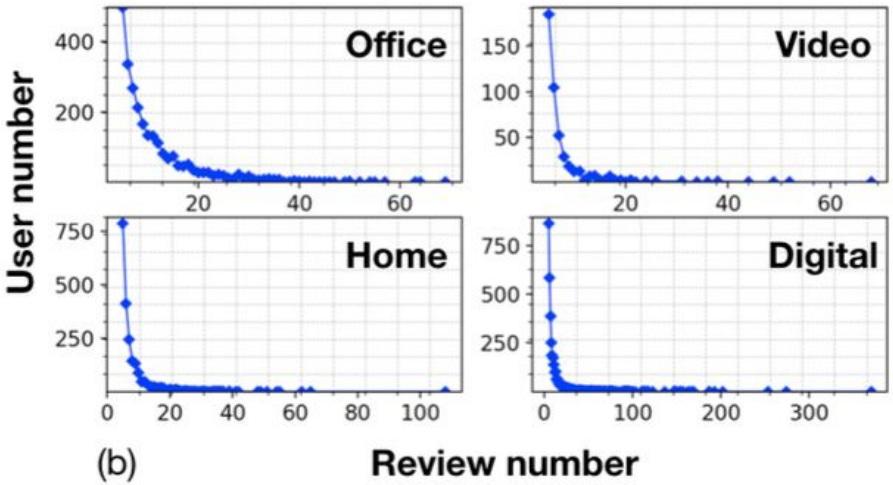
(userA, shirt, collar, +)
(userB, sweater, collar, +)
(userC, shirt, collar, +)
(userE, shirt, collar, +)
(userF, shirt, collar, -)
(userF, sweater,collar, +)

$$\begin{array}{ccc} & & \begin{array}{ccc}\text{collar} & \text{sleeve} & \cdots\end{array} \\ \begin{array}{c}\text{shirt}\\\text{sweater}\\ \\ \end{array} & \begin{array}{c}i_1\\i_2\\i_3\end{array} & \underbrace{\begin{bmatrix} \boxed{2/4 \qquad\qquad} \\ 2/2 \\ \\ \end{bmatrix}}_{F} \end{array} \quad \boldsymbol{B}_i \in \mathbb{R}^{1\times|\mathcal{F}|}$$

$$\boldsymbol{B} = [\boldsymbol{B}_i] \in \mathbb{R}^{|\mathcal{I}|\times|\mathcal{F}|}$$

review-based Recommendation

- **fundamental problem**
  - the review information can be not as ideal as expected.



(b)

Outline

- Introduction
- Method
- Experiment
- Conclusion

Algorithm 1: Learning algorithm of CF^2

```
Train the target model g with the original dataset O.
Initialize the counterfactual sample set O_c = ∅.
for i in [1, M] do
    W_t ∈ R^{d_t × d_{t-1}} with the probability of (1/n_u) / (Σ_{i=1}^{|U|} 1/n_i).
    d_T = 1 ... triplet (u, i, j) in O.
    ... formula (5) or (6) to get τ.
    Compute r*_ui and r*_uj based on τ.
    if r*_ui ≤ r*_uj then
        O_c ← O_c ∪ (u*, j, i)
    end
end
Train the target model g based on O ∪ O_c.
```

1. Implementation of g

$$r_{ui} = g(\mathbf{A}_u, \mathbf{B}_i)$$

$$r_{ui} = \mathbf{W}_T \sigma_T (\mathbf{W}_{T-1} \sigma_{T-1} (\ldots(\mathbf{W}_1 \sigma_1 (m(\mathbf{A}_u, \mathbf{B}_i)) + \boldsymbol{b}_1) + \ldots) + \boldsymbol{b}_{T-1}) + \boldsymbol{b}_T$$

- Element-wise product: $m(\mathbf{A}_u, \mathbf{B}_i) = \underbrace{\mathbf{W}_U^p}_{\mathbb{R}^{d_0 \times |\mathcal{F}|}\text{ embedding}} \underbrace{\mathbf{A}_u^T}_{\mathbb{R}^{|\mathcal{F}|}\text{ user feature}} \odot \mathbf{W}_I^p \mathbf{B}_i^T$

- Element-wise add: $m(\mathbf{A}_u, \mathbf{B}_i) = \mathbf{W}_U^a \mathbf{A}_u^T + \mathbf{W}_I^a \mathbf{B}_i^T$
- Hybrid method: $m(\mathbf{A}_u, \mathbf{B}_i) = [\mathbf{W}_U^{h1} \mathbf{A}_u^T \odot \mathbf{W}_I^{h1} \mathbf{B}_i^T, \mathbf{W}_U^{h2} \mathbf{A}_u^T + \mathbf{W}_I^{h2} \mathbf{B}_i^T]$
- Attention-based method: $m(\mathbf{A}_u, \mathbf{B}_i) = \underbrace{\mathbf{W}^{att}}_{\mathbb{R}^{d_0 \times |\mathcal{F}|}} [\boldsymbol{\alpha}_{ui} \odot (\mathbf{A}_u^T \odot \mathbf{B}_i^T)]$

$$\frac{\exp(w_1 A_{u,j} + w_2 B_{i,j})}{\sum_{k=1}^{|\mathcal{F}|} \exp(w_1 A_{u,k} + w_2 B_{i,k})}$$

Au= $\begin{pmatrix} 0 & 0.8 & 0.1 & 0.2 \end{pmatrix}$ w1=0.5

Bi= $\begin{pmatrix} 0 & 0.2 & 0.8 & 0.4 \end{pmatrix}$ w2=0.5

$\begin{pmatrix} 0 & 0.5 & 0.45 & 0.3 \end{pmatrix}$/1.25

1. Loss Function

$$L_{\text{BPR}} = -\sum_{(u,i,j)\in O} \log\left[\sigma(r_{ui} - r_{uj})\right] + \lambda||g||^2$$

Each element
(u,i,j) means user u prefers item i to item j

Algorithm 1: Learning algorithm of CF^2

Train the target model g with the original dataset O.
Initialize the counterfactual sample set $O_c = \emptyset$.
for i *in* $[1, M]$ **do**
  Sample a user u with the probability of $\frac{\frac{1}{n_u}}{\sum_{i=1}^{|\mathcal{U}|}\frac{1}{n_i}}$.
  Sample a triplet (u, i, j) in O.
  Optimize formula (5) or (6) to get τ.
  Compute r^*_{ui} and r^*_{uj} based on τ.
  if $r^*_{ui} \le r^*_{uj}$ **then**
    $O_c \leftarrow O_c \cup (u^*, j, i)$
  end
end
Train the target model g based on $O \cup O_c$.

2. Frequency-based sampling

Algorithm 1: Learning algorithm of CF^2

Train the target model g with the original dataset $\mathcal{O}$.

Initialize the counterfactual sample set $\mathcal{O}_c = \emptyset$.

for i *in* $[1, M]$ **do**

  Sample a user u with the probability of $\frac{\frac{1}{n_u}}{\sum_{i=1}^{|\mathcal{U}|} \frac{1}{n_i}}$.

  Sample a triplet (u, i, j) in $\mathcal{O}$.

  Optimize formula (5) or (6) to get τ.

  Compute r^*_{ui} and r^*_{uj} based on τ.

  if $r^*_{ui} \leq r^*_{uj}$ **then**

    $\mathcal{O}_c \leftarrow \mathcal{O}_c \cup (u^*, j, i)$

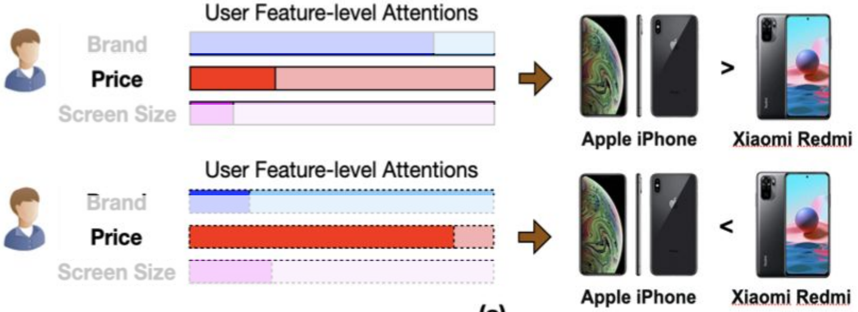
  end

end

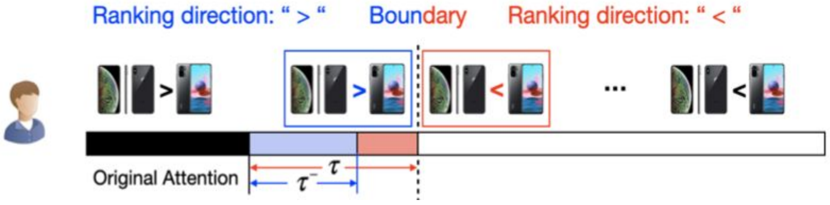
Train the target model g based on $\mathcal{O} \cup \mathcal{O}_c$.

- 建立空集合
- generate M counterfactual samples according to the user reviewing frequency
- fairly train different users

3. data generation



(a)



Algorithm 1: Learning algorithm of CF^2

Train the target model g with the original dataset $\mathcal{O}$.
Initialize the counterfactual sample set $\mathcal{O}_c = \emptyset$.
for i *in* $[1, M]$ **do**
    Sample a user u with the probability of $\frac{\frac{1}{n_u}}{\sum_{i=1}^{|\mathcal{U}|}\frac{1}{n_i}}$.
    Sample a triplet (u, i, j) in $\mathcal{O}$.
    Optimize formula (5) or (6) to get $\boldsymbol{\tau}$.
    Compute r^*_{ui} and r^*_{uj} based on $\boldsymbol{\tau}$.
    if $r^*_{ui} \leq r^*_{uj}$ **then**
        $\mathcal{O}_c \leftarrow \mathcal{O}_c \cup (u^*, j, i)$
    end
end
Train the target model g based on $\mathcal{O} \cup \mathcal{O}_c$.

- **objective:**

$$r^*_{ui} = g(\boldsymbol{A}_u + \boldsymbol{\tau}, \boldsymbol{B}_i) \qquad \boldsymbol{\tau} \in \mathbb{R}^{|\mathcal{F}|}$$

$$\min_{\boldsymbol{\tau}} ||\boldsymbol{\tau}||_2^2 + \alpha \log\left[\sigma(r^*_{ui} - r^*_{uj})\right]$$

$$\begin{cases} \text{Generate } (u^*, j, i). & \text{if } r^*_{ui} \leq r^*_{uj} \\ \text{No generation.} & \text{otherwise} \end{cases}$$

3. Constrained feature perturbation

Algorithm 1: Learning algorithm of CF^2

Train the target model g with the original dataset $\mathcal{O}$.
Initialize the counterfactual sample set $\mathcal{O}_c = \emptyset$.
for *i in [1, M]* **do**
  Sample a user u with the probability of $\frac{\frac{1}{n_u}}{\sum_{i=1}^{|\mathcal{U}|} \frac{1}{n_i}}$.
  Sample a triplet (u, i, j) in $\mathcal{O}$.
  Optimize formula (5) or (6) to get $\boldsymbol{\tau}$.
  Compute r^*_{ui} and r^*_{uj} based on $\boldsymbol{\tau}$.
  if $r^*_{ui} \leq r^*_{uj}$ **then**
    $\mathcal{O}_c \leftarrow \mathcal{O}_c \cup (u^*, j, i)$
  end
end
Train the target model g based on $\mathcal{O} \cup \mathcal{O}_c$.

- **hard**

$$\bar{r}_{ui} = g(A_u + \boldsymbol{k}^u \odot \boldsymbol{\tau}, B_i)$$

$$\min_{\boldsymbol{\tau}} ||\boldsymbol{k}^u \odot \boldsymbol{\tau}||_2^2 + \alpha \log\left[\sigma(\bar{r}_{ui} - \bar{r}_{uj})\right]$$

$\boldsymbol{k}^u \in \mathbb{R}^{|\mathcal{F}|}$ is a mask vector

- **soft**

$$\min_{\boldsymbol{\tau}} ||\boldsymbol{\tau}||_2^2 + ||\boldsymbol{\tau}||_1 + \alpha \log\left[\sigma(r^*_{ui} - r^*_{uj})\right]$$

3. Controlling the noisy information

Algorithm 1: Learning algorithm of CF^2

Train the target model g with the original dataset $\mathcal{O}$.
Initialize the counterfactual sample set $\mathcal{O}_c = \emptyset$.
for i *in* $[1, M]$ **do**
  Sample a user u with the probability of $\frac{\frac{1}{n_u}}{\sum_{i=1}^{|\mathcal{U}|}\frac{1}{n_i}}$.
  Sample a triplet (u, i, j) in $\mathcal{O}$.
  Optimize formula (5) or (6) to get τ.
  Compute r^*_{ui} and r^*_{uj} based on τ.
  if $r^*_{ui} \le r^*_{uj}$ **then**
    $\mathcal{O}_c \leftarrow \mathcal{O}_c \cup (u^*, j, i)$
  end
end
Train the target model g based on $\mathcal{O} \cup \mathcal{O}_c$.

- **original**

$$\begin{cases} \text{Generate } (u^*, j, i). & \text{if } r^*_{ui} \le r^*_{uj} \\ \text{No generation.} & \text{otherwise} \end{cases}$$

- **controlling the noisy information**

$$\begin{cases} \text{Generate } (u^*, j, i). & r^*_{ui} - r^*_{uj} \le \kappa \\ \text{No generation.} & r^*_{ui} - r^*_{uj} > \kappa \end{cases}$$

$$\kappa \in \mathbb{R}_-$$

the <u>number</u> and <u>reliability</u> of the generated samples

Outline

- Introduction
- Method
- Experiment
- Conclusion

Datasets

Table 1: Statistics of the datasets.

Dataset	#User	#Item	#Interaction	Density
Office Products	4905	2420	53258	0.45%
Digital Music	5541	3568	64706	0.33%
Tools & Home	16638	10217	134476	0.08%
Home & Kitchen	66519	28237	551682	0.03%
Yelp	4777	11774	187615	0.33%

Experiment-metric

- **NDCG**

Test set
(1, 2, 6) ➔ 1
(2, 4, 5) ➔ 4
(1, 3, 5) ➔ 2

Prediction
(1, 2, 6) ➔ 1, 2, 3, 4, 5, 6
(2, 4, 5) ➔ 2, 4, 1, 3, 6, 5
(1, 3, 5) ➔ 3, 4, 1, 5, 6, 2

- **NDCG @5 of (1, 2, 6) test pair**

$$= \left(\frac{2^1-1}{log(1+1)} + \frac{2^0-1}{log(2+1)} + \frac{2^0-1}{log(3+1)} + \frac{2^0-1}{log(4+1)} + \frac{2^0-1}{log(5+1)}\right) \quad \text{DCG}$$

$$/ \left(\frac{2^1-1}{log(1+1)} + \frac{2^0-1}{log(2+1)} + \frac{2^0-1}{log(3+1)} + \frac{2^0-1}{log(4+1)} + \frac{2^0-1}{log(5+1)}\right) \quad \text{IDCG}$$

- **NDCG @5 of (2, 4, 5) test pair**

$$= \left(\frac{2^0-1}{log(1+1)} + \frac{2^1-1}{log(2+1)} + \frac{2^0-1}{log(3+1)} + \frac{2^0-1}{log(4+1)} + \frac{2^0-1}{log(5+1)}\right) \quad \text{DCG}$$

$$/ \left(\frac{2^1-1}{log(1+1)} + \frac{2^0-1}{log(2+1)} + \frac{2^0-1}{log(3+1)} + \frac{2^0-1}{log(4+1)} + \frac{2^0-1}{log(5+1)}\right) \quad \text{IDCG}$$

Experiment-1

非review-based

review-based

feature-aware

Dataset	Office Products			Digital			Tools & Home			Home & Kitchen			Yelp		
Metric (@5)	F_1	NDCG	HR	F_1	NDCG	HR	F_1	NDCG	HR	F_1	NDCG	HR	F_1	NDCG	HR
BPR	0.088	0.110	0.420	0.086	0.147	0.429	0.050	0.071	0.263	0.081	0.122	0.409	0.190	0.290	0.755
NCF	0.102	0.127	0.464	0.081	0.115	0.352	0.058	0.080	0.303	0.082	0.128	0.429	0.178	0.236	0.783
MPCN	0.109	0.131	0.477	0.089	0.125	0.371	0.061	0.086	0.323	0.110	0.192	0.566	0.181	0.255	0.791
EFM	0.108	0.135	0.469	0.091	0.149	0.453	0.079	0.134	0.391	0.130	0.229	0.580	0.193	0.289	0.801
A2CF	0.113	0.171	0.550	0.092	0.155	0.461	0.080	0.138	0.413	0.133	0.238	0.590	0.197	0.292	0.805
$\text{CF}^2_{\text{base}}$-P	0.117	0.176	0.543	0.091	0.158	0.455	0.078	0.137	0.420	0.135	0.224	0.581	0.194	0.284	0.798
$\text{CF}^2_{\text{rand}}$-P	0.112	0.162	0.523	0.089	0.146	0.448	0.075	0.123	0.401	0.141	0.227	0.589	0.191	0.274	0.793
$\text{CF}^2_{\text{hard}}$-P	0.127	0.179	0.571	0.099	0.164	0.479	0.084	0.154	0.434	0.138	0.232	0.587	0.207	0.298	0.812
$\text{CF}^2_{\text{soft}}$-P	0.126	0.184	0.570	0.099	0.154	0.491	0.085	0.135	0.440	0.140	0.228	0.592	0.197	0.281	0.811
$\text{CF}^2_{\text{base}}$-A	0.114	0.165	0.534	0.088	0.135	0.445	0.080	0.125	0.430	0.140	0.233	0.584	0.204	0.284	0.810
$\text{CF}^2_{\text{rand}}$-A	0.110	0.160	0.523	0.090	0.143	0.447	0.080	0.135	0.421	0.139	0.232	0.581	0.204	0.283	0.803
$\text{CF}^2_{\text{hard}}$-A	0.120	0.170	0.547	0.100	0.166	0.482	0.084	0.138	0.431	0.144	0.238	**0.601**	0.207	0.286	0.812
$\text{CF}^2_{\text{soft}}$-A	0.123	0.174	0.564	0.094	0.141	0.458	0.083	0.138	0.428	0.143	0.234	0.592	0.208	0.288	0.812
$\text{CF}^2_{\text{base}}$-H	0.119	0.184	0.557	0.089	0.152	0.443	0.082	0.151	0.432	0.138	0.233	0.575	0.201	0.277	0.804
$\text{CF}^2_{\text{rand}}$-H	0.114	0.180	0.552	0.088	0.150	0.436	0.080	0.130	0.430	0.137	0.230	0.577	0.202	0.280	0.813
$\text{CF}^2_{\text{hard}}$-H	**0.127**	**0.193**	**0.575**	0.097	0.161	0.472	0.087	0.159	0.436	0.143	0.239	0.596	0.210	0.289	0.817
$\text{CF}^2_{\text{soft}}$-H	0.126	0.188	0.571	0.096	0.143	0.467	0.084	0.156	0.433	0.142	0.239	0.594	0.209	0.287	0.816
$\text{CF}^2_{\text{base}}$-AT	0.118	0.164	0.540	0.098	0.173	0.482	0.085	0.145	0.435	0.139	0.234	0.587	0.209	0.281	0.813
$\text{CF}^2_{\text{rand}}$-AT	0.113	0.165	0.530	0.100	0.176	0.493	0.087	0.147	0.444	0.138	0.226	0.586	0.205	0.284	0.809
$\text{CF}^2_{\text{hard}}$-AT	0.124	0.169	0.552	**0.106**	**0.183**	**0.504**	**0.093**	**0.158**	**0.474**	**0.148**	**0.246**	0.599	**0.216**	**0.301**	**0.827**
$\text{CF}^2_{\text{soft}}$-AT	0.121	0.181	0.557	0.104	0.176	0.486	0.089	0.154	0.454	0.143	0.241	0.592	0.213	0.291	0.819

Experiment-1

Dataset	Office Products			Digital			Tools & Home			Home & Kitchen			Yelp		
Metric (@5)	F_1	NDCG	HR	F_1	NDCG	HR	F_1	NDCG	HR	F_1	NDCG	HR	F_1	NDCG	HR
BPR	0.088	0.110	0.420	0.086	0.147	0.429	0.050	0.071	0.263	0.081	0.122	0.409	0.190	0.290	0.755
NCF	0.102	0.127	0.464	0.081	0.115	0.352	0.058	0.080	0.303	0.082	0.128	0.429	0.178	0.236	0.783
MPCN	0.109	0.131	0.477	0.089	0.125	0.371	0.061	0.086	0.323	0.110	0.192	0.566	0.181	0.255	0.791
EFM	0.108	0.135	0.469	0.091	0.149	0.453	0.079	0.134	0.391	0.130	0.229	0.580	0.193	0.289	0.801
A2CF	0.113	0.171	0.550	0.092	0.155	0.461	0.080	0.138	0.413	0.133	0.238	0.590	0.197	0.292	0.805
$\text{CF}^2_{\text{base}}$-P	0.117	0.176	0.543	0.091	0.158	0.455	0.078	0.137	0.420	0.135	0.224	0.581	0.194	0.284	0.798
$\text{CF}^2_{\text{rand}}$-P	0.112	0.162	0.523	0.089	0.146	0.448	0.075	0.123	0.401	0.141	0.227	0.589	0.191	0.274	0.793
$\text{CF}^2_{\text{hard}}$-P	0.127	0.179	0.571	0.099	0.164	0.479	0.084	0.154	0.434	0.138	0.232	0.587	0.207	0.298	0.812
$\text{CF}^2_{\text{soft}}$-P	0.126	0.184	0.570	0.099	0.154	0.491	0.085	0.135	0.440	0.140	0.228	0.592	0.197	0.281	0.811
$\text{CF}^2_{\text{base}}$-A	0.114	0.165	0.534	0.088	0.135	0.445	0.080	0.125	0.430	0.140	0.233	0.584	0.204	0.284	0.810
$\text{CF}^2_{\text{rand}}$-A	0.110	0.160	0.523	0.090	0.143	0.447	0.080	0.135	0.421	0.139	0.232	0.581	0.204	0.283	0.803
$\text{CF}^2_{\text{hard}}$-A	0.120	0.170	0.547	0.100	0.166	0.482	0.084	0.138	0.431	0.144	0.238	**0.601**	0.207	0.286	0.812
$\text{CF}^2_{\text{soft}}$-A	0.123	0.174	0.564	0.094	0.141	0.458	0.083	0.138	0.428	0.143	0.234	0.592	0.208	0.288	0.812
$\text{CF}^2_{\text{base}}$-H	0.119	0.184	0.557	0.089	0.152	0.443	0.082	0.151	0.432	0.138	0.233	0.575	0.201	0.277	0.804
$\text{CF}^2_{\text{rand}}$-H	0.114	0.180	0.552	0.088	0.150	0.436	0.080	0.130	0.430	0.137	0.230	0.577	0.202	0.280	0.813
$\text{CF}^2_{\text{hard}}$-H	**0.127**	**0.193**	**0.575**	0.097	0.161	0.472	0.087	0.159	0.436	0.143	0.239	0.596	0.210	0.289	0.817
$\text{CF}^2_{\text{soft}}$-H	0.126	0.188	0.571	0.096	0.143	0.467	0.084	0.156	0.433	0.142	0.239	0.594	0.209	0.287	0.816
$\text{CF}^2_{\text{base}}$-AT	0.118	0.164	0.540	0.098	0.173	0.482	0.085	0.145	0.435	0.139	0.234	0.587	0.209	0.281	0.813
$\text{CF}^2_{\text{rand}}$-AT	0.113	0.165	0.530	0.100	0.176	0.493	0.087	0.147	0.444	0.138	0.226	0.586	0.205	0.284	0.809
$\text{CF}^2_{\text{hard}}$-AT	0.124	0.169	0.552	**0.106**	**0.183**	**0.504**	**0.093**	**0.158**	**0.474**	**0.148**	**0.246**	0.599	**0.216**	**0.301**	**0.827**
$\text{CF}^2_{\text{soft}}$-AT	0.121	0.181	0.557	0.104	0.176	0.486	0.089	0.154	0.454	0.143	0.241	0.592	0.213	0.291	0.819

Experiment-1

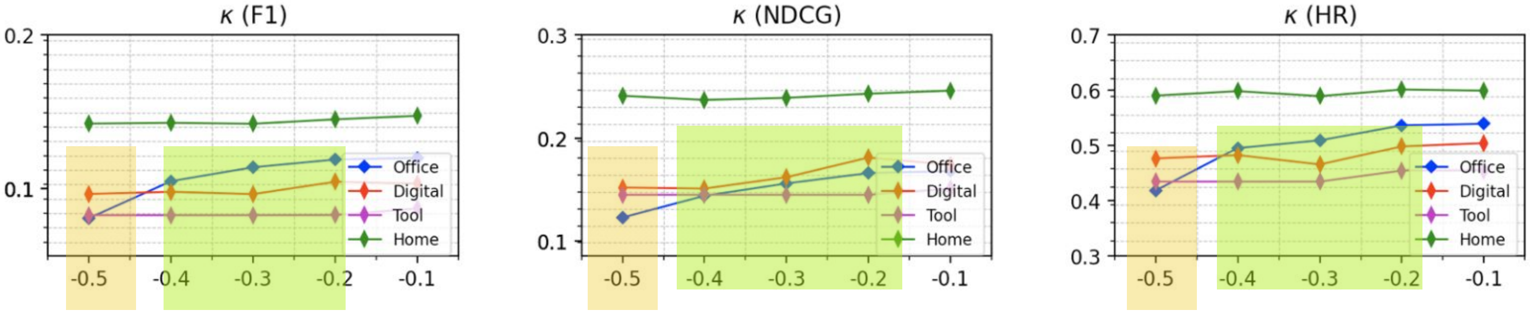
Dataset	Office Products			Digital			Tools & Home			Home & Kitchen			Yelp		
Metric (@5)	F_1	NDCG	HR	F_1	NDCG	HR	F_1	NDCG	HR	F_1	NDCG	HR	F_1	NDCG	HR
BPR	0.088	0.110	0.420	0.086	0.147	0.429	0.050	0.071	0.263	0.081	0.122	0.409	0.190	0.290	0.755
NCF	0.102	0.127	0.464	0.081	0.115	0.352	0.058	0.080	0.303	0.082	0.128	0.429	0.178	0.236	0.783
MPCN	0.109	0.131	0.477	0.089	0.125	0.371	0.061	0.086	0.323	0.110	0.192	0.566	0.181	0.255	0.791
EFM	0.108	0.135	0.469	0.091	0.149	0.453	0.079	0.134	0.391	0.130	0.229	0.580	0.193	0.289	0.801
A2CF	0.113	0.171	0.550	0.092	0.155	0.461	0.080	0.138	0.413	0.133	0.238	0.590	0.197	0.292	0.805
CF^2_{base}-P	0.117	0.176	0.543	0.091	0.158	0.455	0.078	0.137	0.420	0.135	0.224	0.581	0.194	0.284	0.798
CF^2_{rand}-P	0.112	0.162	0.523	0.089	0.146	0.448	0.075	0.123	0.401	0.141	0.227	0.589	0.191	0.274	0.793
CF^2_{hard}-P	0.127	0.179	0.571	0.099	0.164	0.479	0.084	0.154	0.434	0.138	0.232	0.587	0.207	0.298	0.812
CF^2_{soft}-P	0.126	0.184	0.570	0.099	0.154	0.491	0.085	0.135	0.440	0.140	0.228	0.592	0.197	0.281	0.811
CF^2_{base}-A	0.114	0.165	0.534	0.088	0.135	0.445	0.080	0.125	0.430	0.140	0.233	0.584	0.204	0.284	0.810
CF^2_{rand}-A	0.110	0.160	0.523	0.090	0.143	0.447	0.080	0.135	0.421	0.139	0.232	0.581	0.204	0.283	0.803
CF^2_{hard}-A	0.120	0.170	0.547	0.100	0.166	0.482	0.084	0.138	0.431	0.144	0.238	**0.601**	0.207	0.286	0.812
CF^2_{soft}-A	0.123	0.174	0.564	0.094	0.141	0.458	0.083	0.138	0.428	0.143	0.234	0.592	0.208	0.288	0.812
CF^2_{base}-H	0.119	0.184	0.557	0.089	0.152	0.443	0.082	0.151	0.432	0.138	0.233	0.575	0.201	0.277	0.804
CF^2_{rand}-H	0.114	0.180	0.552	0.088	0.150	0.436	0.080	0.130	0.430	0.137	0.230	0.577	0.202	0.280	0.813
CF^2_{hard}-H	**0.127**	**0.193**	**0.575**	0.097	0.161	0.472	0.087	0.159	0.436	0.143	0.239	0.596	0.210	0.289	0.817
CF^2_{soft}-H	0.126	0.188	0.571	0.096	0.143	0.467	0.084	0.156	0.433	0.142	0.239	0.594	0.209	0.287	0.816
CF^2_{base}-AT	0.118	0.164	0.540	0.098	0.173	0.482	0.085	0.145	0.435	0.139	0.234	0.587	0.209	0.281	0.813
CF^2_{rand}-AT	0.113	0.165	0.530	0.100	0.176	0.493	0.087	0.147	0.444	0.138	0.226	0.586	0.205	0.284	0.809
CF^2_{hard}-AT	0.124	0.169	0.552	**0.106**	**0.183**	**0.504**	**0.093**	**0.158**	**0.474**	**0.148**	**0.246**	0.599	**0.216**	**0.301**	**0.827**
CF^2_{soft}-AT	0.121	0.181	0.557	0.104	0.176	0.486	0.089	0.154	0.454	0.143	0.241	0.592	0.213	0.291	0.819

Experiment-ablation

- constrained feature perturbation
- frequency- based sampling
- noisy information control

Dataset	Office Products			Digital			Tools & Home			Home & Kitchen		
Metric (@5)	F_1	NDCG	HR	F_1	NDCG	HR	F_1	NDCG	HR	F_1	NDCG	HR
CF^2_{-cst}-H	0.124	0.180	0.562	0.093	0.140	0.461	0.082	0.147	0.431	0.142	0.238	0.593
$CF^2_{hard,-samp}$-H	0.122	0.172	0.566	0.095	0.149	0.468	0.083	0.138	0.438	0.141	0.236	0.587
$CF^2_{soft,-samp}$-H	0.121	0.169	0.555	0.094	0.144	0.463	0.078	0.124	0.428	0.131	0.205	0.576
$CF^2_{hard,-\kappa}$-H	0.112	0.168	0.535	0.087	0.142	0.425	0.074	0.117	0.399	0.137	0.229	0.577
$CF^2_{soft,-\kappa}$-H	0.117	0.176	0.562	0.087	0.142	0.425	0.081	0.148	0.424	0.127	0.215	0.542
CF^2-H	**0.127**	**0.193**	**0.575**	**0.097**	**0.161**	**0.472**	**0.087**	**0.159**	**0.436**	**0.143**	**0.239**	**0.596**
CF^2_{-cst}-AT	0.119	0.167	0.538	0.102	0.175	0.496	0.087	0.141	0.461	0.145	0.245	0.595
$CF^2_{hard,-samp}$-AT	0.117	0.166	0.537	0.101	0.174	0.492	0.089	0.153	0.447	0.143	0.239	0.596
$CF^2_{soft,-samp}$-AT	0.119	0.177	0.543	0.102	0.175	0.496	0.087	0.153	0.451	0.141	0.239	0.589
$CF^2_{hard,-\kappa}$-AT	0.078	0.109	0.399	0.097	0.170	0.479	0.067	0.112	0.345	0.135	0.227	0.583
$CF^2_{soft,-\kappa}$-AT	0.119	0.178	0.564	0.101	0.173	0.494	0.079	0.124	0.431	0.142	0.240	0.590
CF^2-AT	**0.124**	**0.181**	**0.557**	**0.106**	**0.183**	**0.504**	**0.093**	**0.158**	**0.474**	**0.148**	**0.246**	**0.599**

Experiment-hyper parameters-k



- **controlling the noisy information**

$$\begin{cases} \text{Generate } (u^*, j, i). & r^*_{ui} - r^*_{uj} \leq \kappa \\ \text{No generation.} & r^*_{ui} - r^*_{uj} > \kappa \end{cases}$$

$$\kappa \in \mathbb{R}_-$$

Conclusion

- enhance review-based recommendation based on the idea of data augmentation.

- design a learning-based method to discover boundary samples for better model optimization.