Method

Contrastive Collaborative Filtering for Cold-Start Item Recommendation

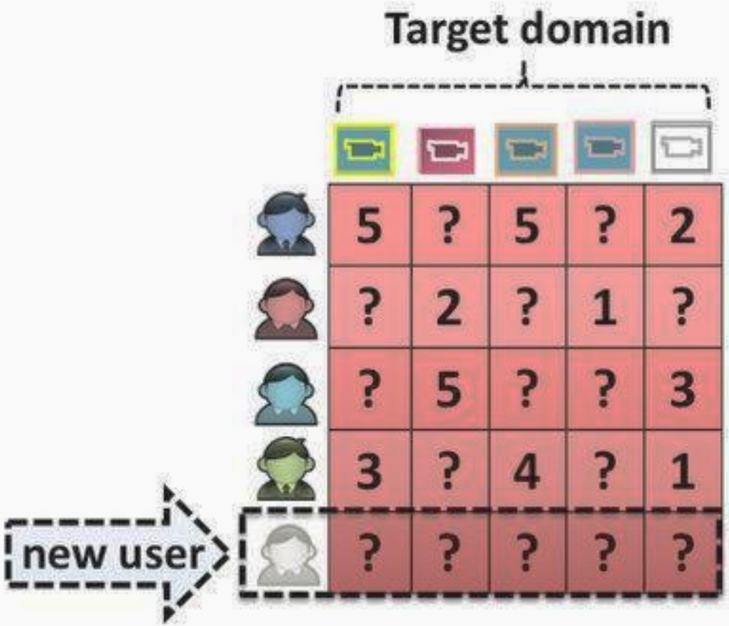
Task

Advisor : Jia-Ling, Koh
Speaker : Siao You, Chen
Source : WWW 2023
Date : 2024/02/20

Outline

- Introduction
- Method
- Experiment
- Conclusion

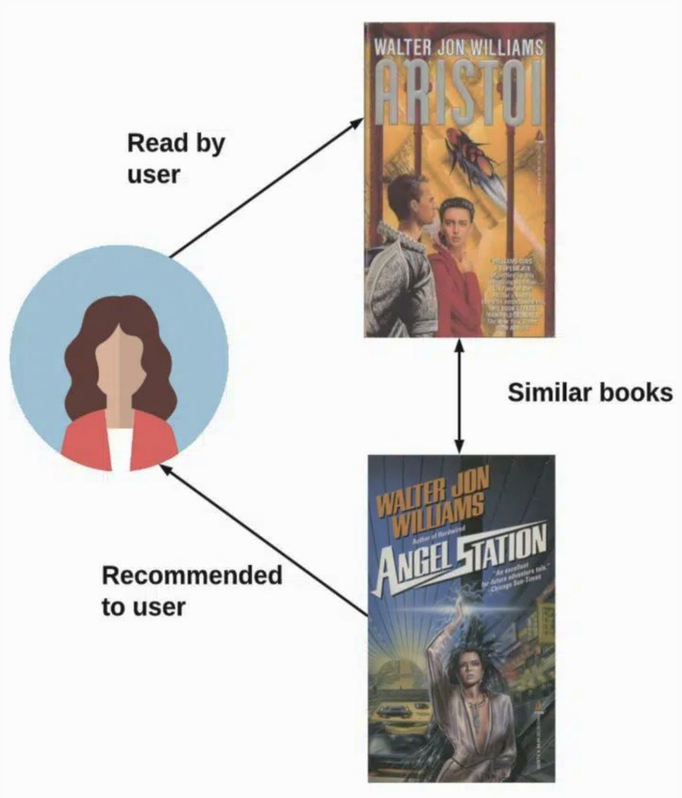
Cold-Start



Target domain
5 ? 5 ? 2
? 2 ? 1 ?
? 5 ? ? 3
3 ? 4 ? 1
? ? ? ? ?
new user

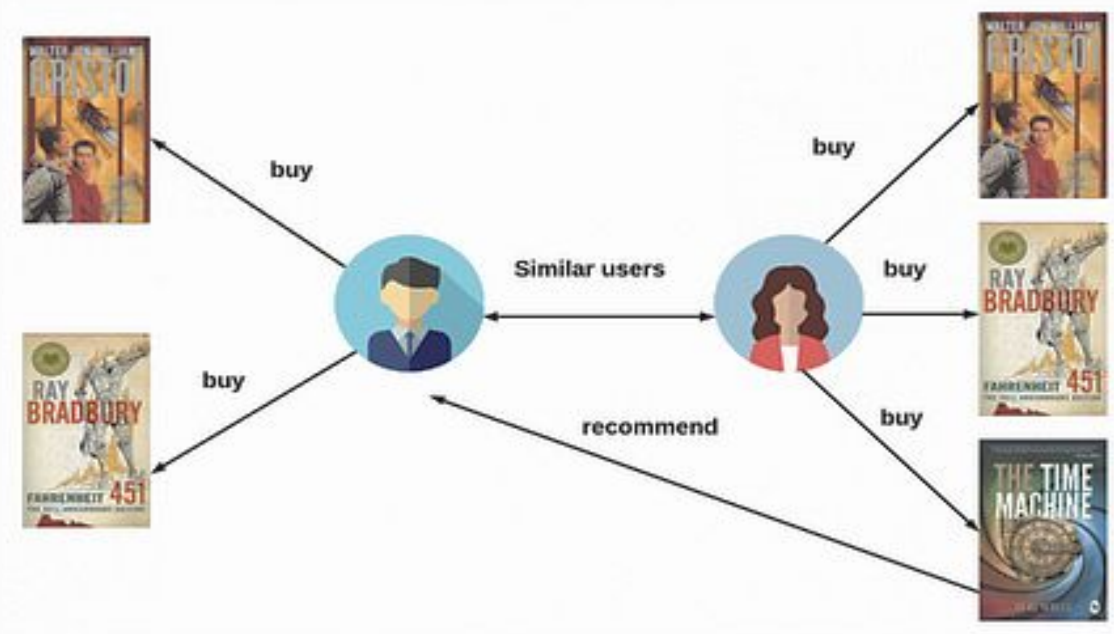
Target domain
new Item!
5 ? 5 ? ?
? 2 ? 1 ?
? 5 ? ? ?
3 ? 4 ? ?
? 2 ? 4 ?

Content-based Filtering



Based on what we like,
the algorithm will simply pick items with similar content to recommend us.

User-based Collaborative Filtering



Finding users who are most similar to the target user based on their historical interactions with items

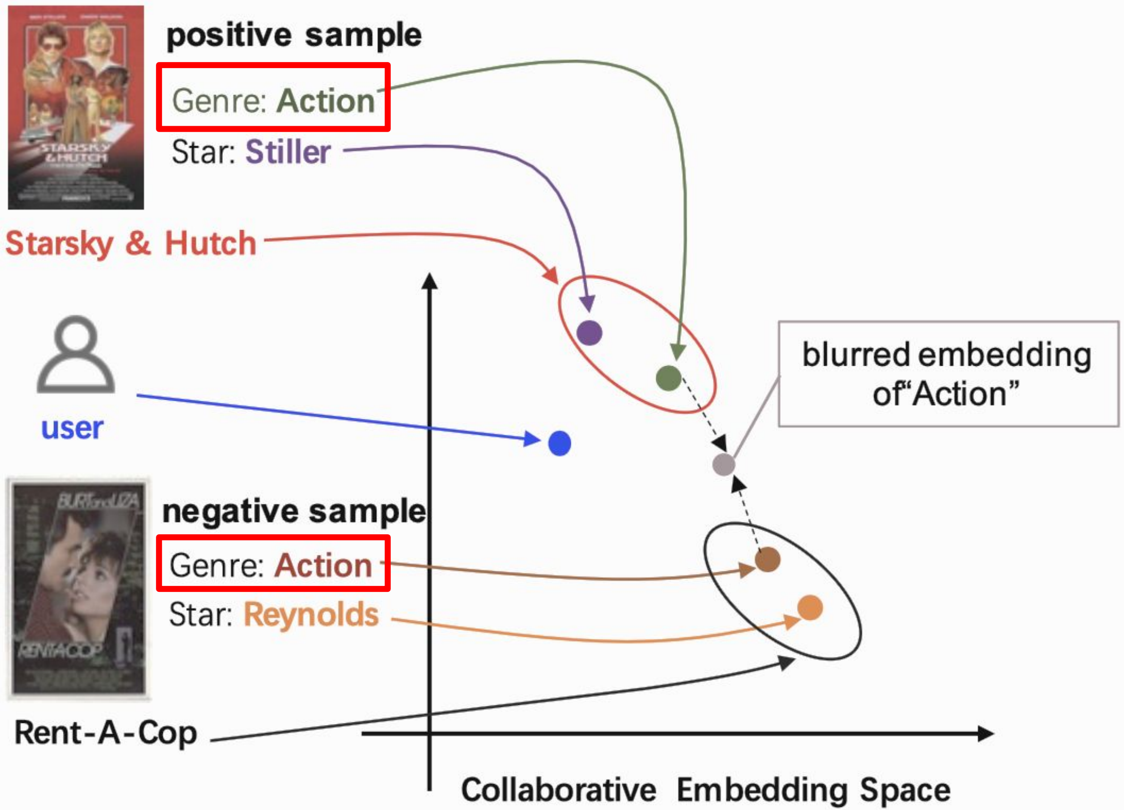
Item-based Collaborative Filtering

Identify items that are similar to the ones the target user has already interacted with.

Problem

Movie +

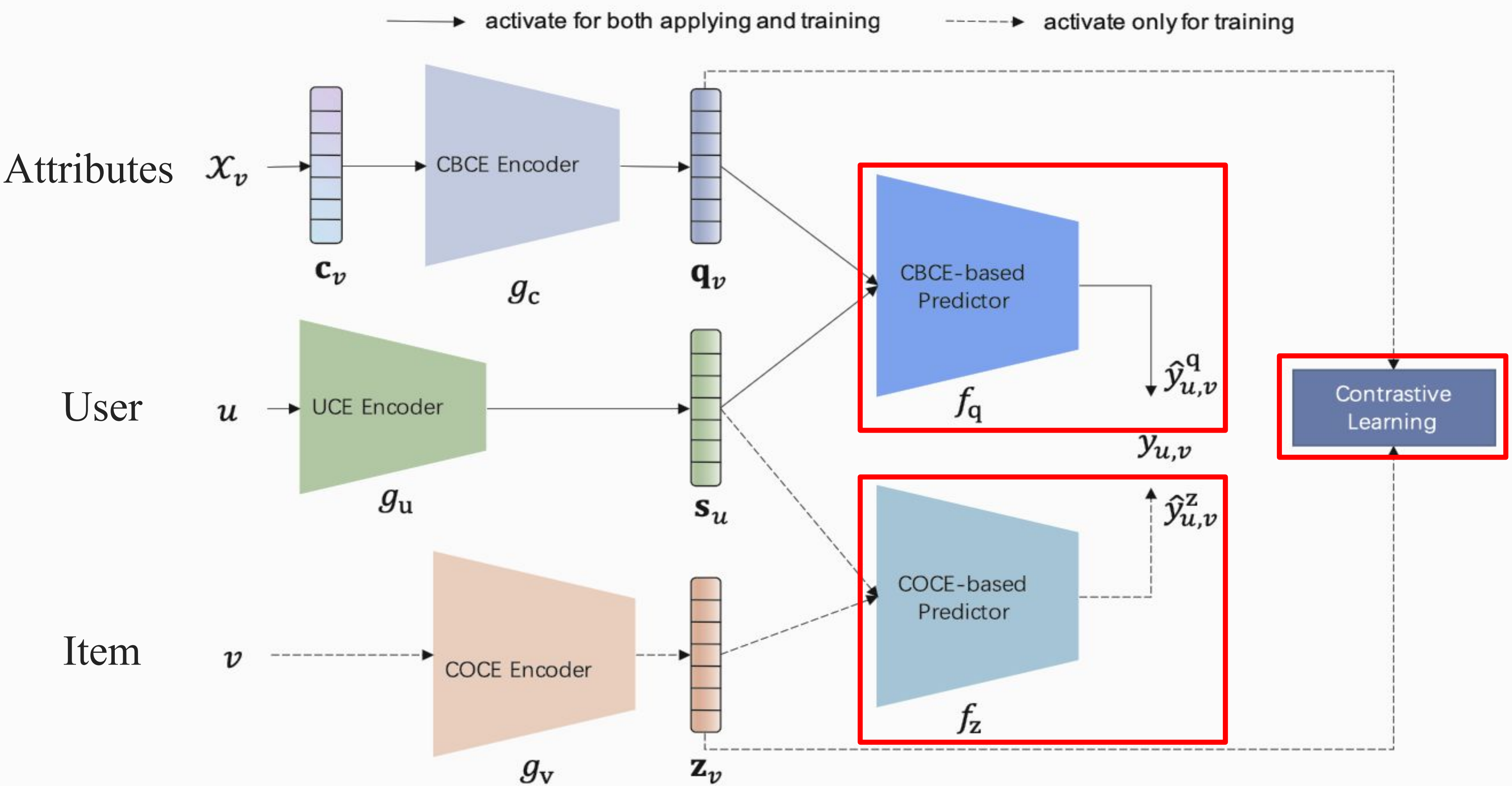
Movie –



Outline

- Introduction
- Method
- Experiment
- Conclusion

Model Structure



Notations

$\mathcal{U}$ and $\mathcal{V}$ be a set of users and a set of items

$\mathcal{O} = \{o_{u,v}\}$ is the interaction between a user $u \in \mathcal{U}$ and an item $v \in \mathcal{V}$

m attributes $\mathcal{X}_v = \{x_1^{(v)}, \cdots, x_m^{(v)}\}$

a training sample $(u, v, \mathcal{X}_v, y_{u,v})$

Model Structure

positive sample

Genre: **Action**

Star: David Soul,
Paul Michael Glaser

Starsky & Hutch

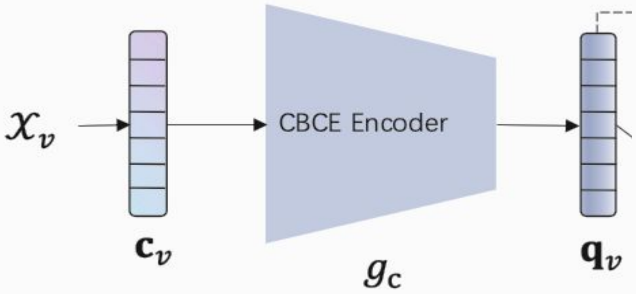
Action	Romantic	Animated
1	0	0

$x_1^{(v)}$

David	Reynolds	Paul
1	0	1

$x_2^{(v)}$

x_v



Model Structure

Action	Romantic	Animated
1	0	0

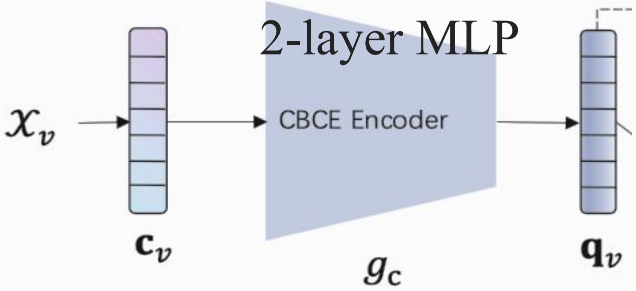
$x_1^{(v)}$

David	Reynolds	Paul
1	0	1

$x_2^{(v)}$

$$\mathbf{x}_i^{(v)} = \mathbf{W}_i x_i^{(v)}$$

$\mathbf{c}_v$ a concatenation of $\{\mathbf{x}_i^{(v)}\}$



Notations

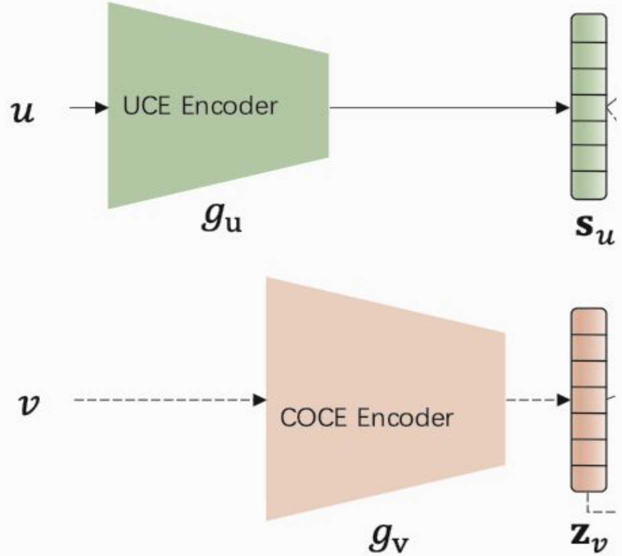
	Item 1	Item 2	Item 3	Item 4	…	
User 1	1	0	0	1	0	1
User 2	0	0	1	0	1	0
User 3	0	1	1	0	1	0
User 4	1	0	0	1	0	1

Model Structure

u

1	0	0	1	0	1

$$\mathbf{s}_u = g_{\mathrm{u}}(\mathbf{u}) = \mathbf{W}_u \mathbf{u}$$

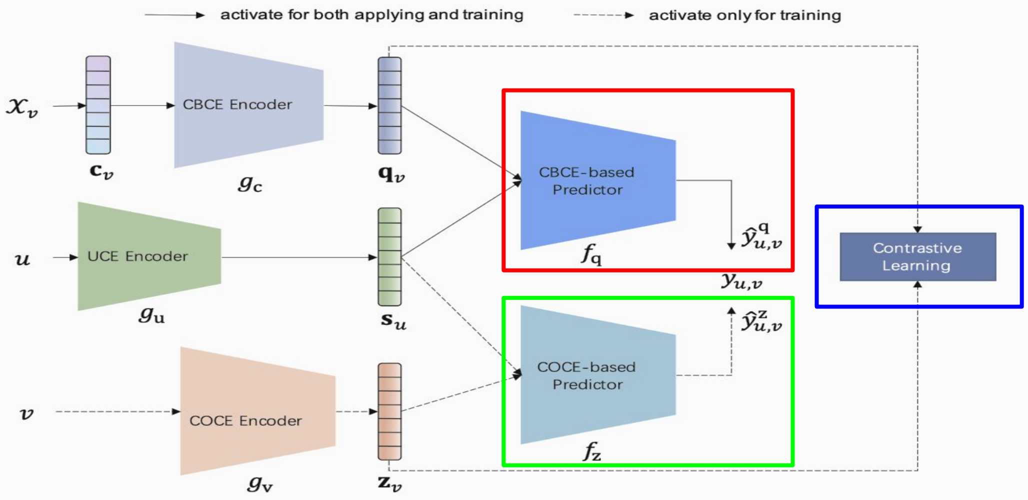


v

1	0	0	1

$$\mathbf{z}_v = g_{\mathrm{v}}(\mathbf{v}) = \mathbf{W}_v \mathbf{v}$$

Overall Loss



$$\mathcal{L} = \boxed{\mathcal{L}_q} + \boxed{\mathcal{L}_z} + \boxed{\lambda \mathcal{L}_c} + ||\Theta||$$

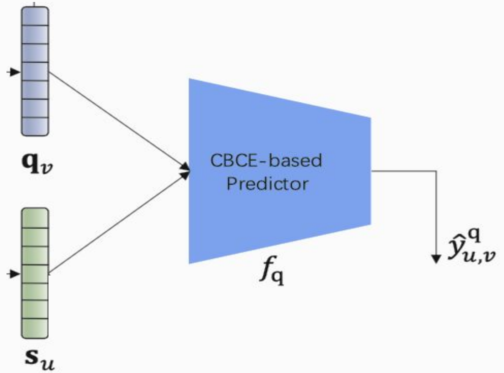
Prediction Loss

Contrastive Loss

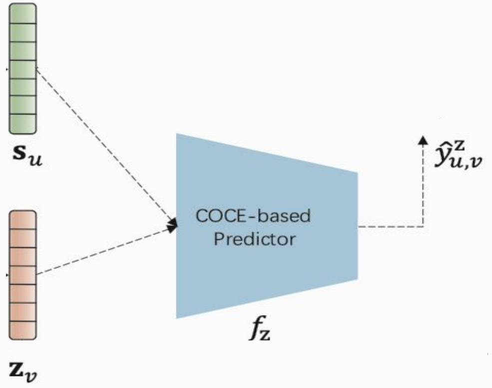
Learnable Parameters

Prediction Loss - Bayesian Personalized Ranking

$$\mathcal{L}_{\mathrm{q}} = - \sum_{(v,u^+,u^-)\in\mathcal{D}} \ln \sigma(\hat{y}^{\mathrm{q}}_{u^+,v} - \hat{y}^{\mathrm{q}}_{u^-,v})$$



$$\mathcal{L}_{\mathrm{z}} = - \sum_{(v,u^+,u^-)\in\mathcal{D}} \ln \sigma(\hat{y}^{\mathrm{z}}_{u^+,v} - \hat{y}^{\mathrm{z}}_{u^-,v})$$



σ is the sigmoid function

Contrastive Loss - InfoNCE

Attributes

Item
Interactions

$$\mathcal{L}_{\mathrm{c}} = -\mathbb{E}_{v \in \mathcal{D}, v^+ \in \mathcal{N}_v^+} \left[\ln \frac{\exp\left(\frac{\langle \mathbf{q}_v, \mathbf{z}_{v^+} \rangle}{\tau}\right)}{\exp\left(\frac{\langle \mathbf{q}_v, \mathbf{z}_{v^+} \rangle}{\tau}\right) + \sum_{v^- \in \mathcal{N}_v^-} \exp\left(\frac{\langle \mathbf{q}_v, \mathbf{z}_{v^-} \rangle}{\tau}\right)} \right]$$

$$\mathcal{N}_v^+ = \{v^+ : \mathcal{U}_v \cap \mathcal{U}_{v^+} \neq \emptyset\}$$

Outline

- Introduction
- Method
- Experiment
- Conclusion

Datasets

Dataset	#Interactions	#Users	#Items	Sparsity
ML-20M	19,904,260	138,493	24,003	0.598%
Amazon-VG	475,952	52,965	35,322	0.025%

- MovieLens-20M (ML-20M)
- Amazon Video Games (Amazon-VG)

Evaluation Metric

- HR@k

$$\mathrm{HR}@k = \frac{1}{|\mathcal{D}_t|} \sum_{v \in \mathcal{D}_t} \boxed{\frac{\sum_{u \in \mathcal{U}_v} \mathbb{I}(rank(u, l_v) \leq k)}{k}}$$

True	Pred	
A	A	1
B	C	1
C	D	1
D	F	0
E	G	0

= 3/5 = 0.6

Evaluation Metric

- NDCG@k $\quad \mathrm{NDCG}@k = \frac{1}{|\mathcal{D}_t|} \sum_{v \in \mathcal{D}_t} \frac{1}{|\mathcal{U}_v|} \boxed{\sum_{u \in \mathcal{U}_v} \frac{\mathbb{I}(rank(u, l_v) \le k)}{\log(1 + rank(u))}}$

True	Pred	
A	A	$\frac{1}{\log 1 + 1} \approx 3.3219$
B	C	$\frac{1}{\log 1 + 3} \approx 1.6609$
C	D	$\frac{1}{\log 1 + 4} \approx 1.4306$
D	F	0
E	G	0

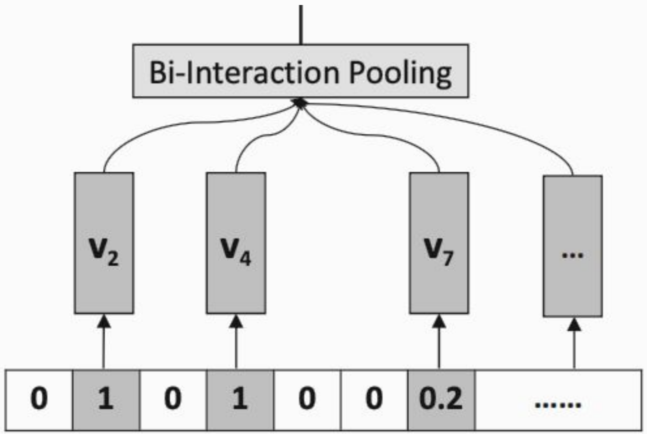
3.3219+1.6609+1.4306
= 6.4134

Baselines - NFM

$$\hat{y}(x) := w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{i=1}^{n} \sum_{j=i+1}^{n} \hat{w}_{i,j} x_i x_j$$

$$\hat{y}_{NFM}(\mathbf{x}) = w_0 + \sum_{i=1}^{n} w_i x_i + f(\mathbf{x})$$

Baselines - NFM

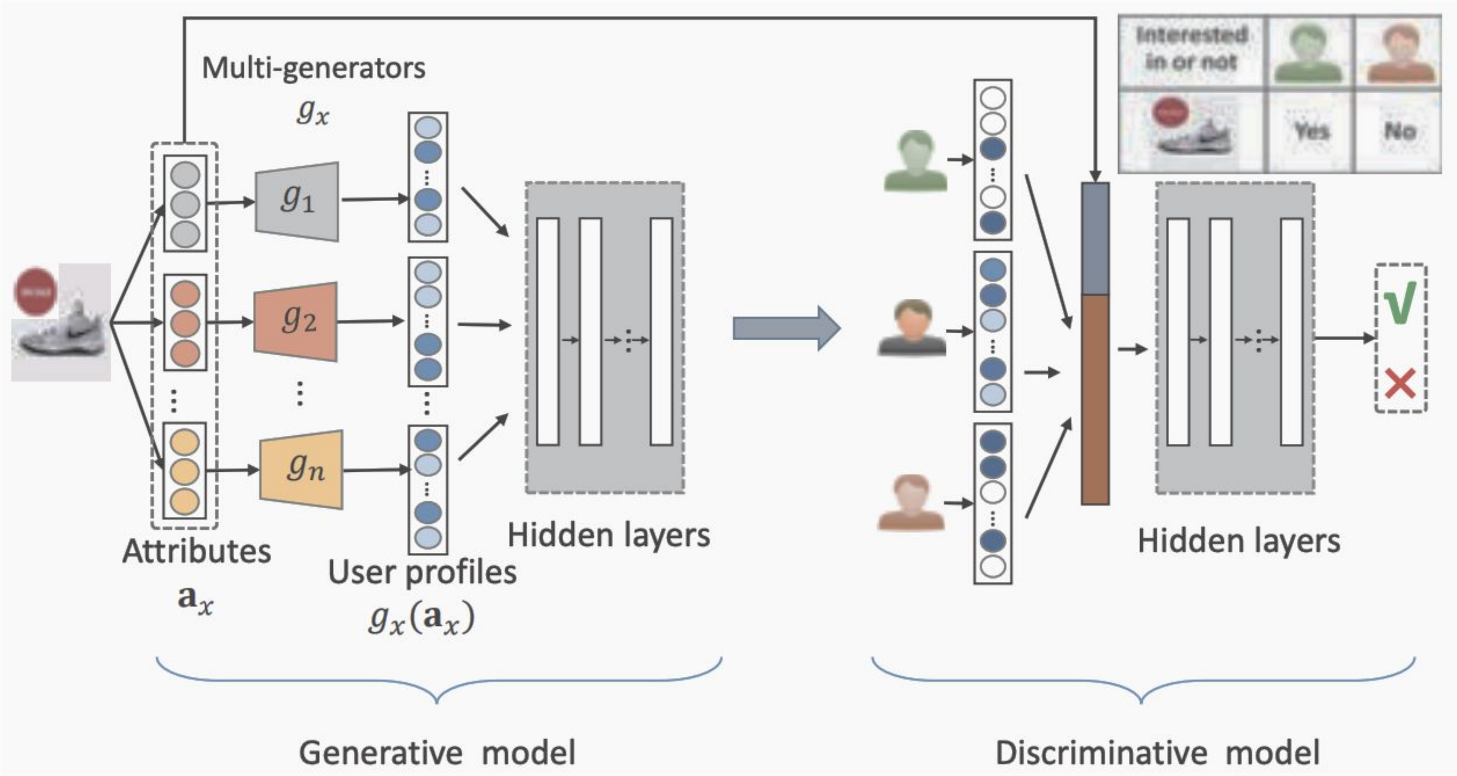


B-Interaction Layer $$f_{BI}(\mathcal{V}_x) = \sum_{i=1}^{n} \sum_{j=i+1}^{n} x_i \mathbf{v}_i \odot x_j \mathbf{v}_j$$

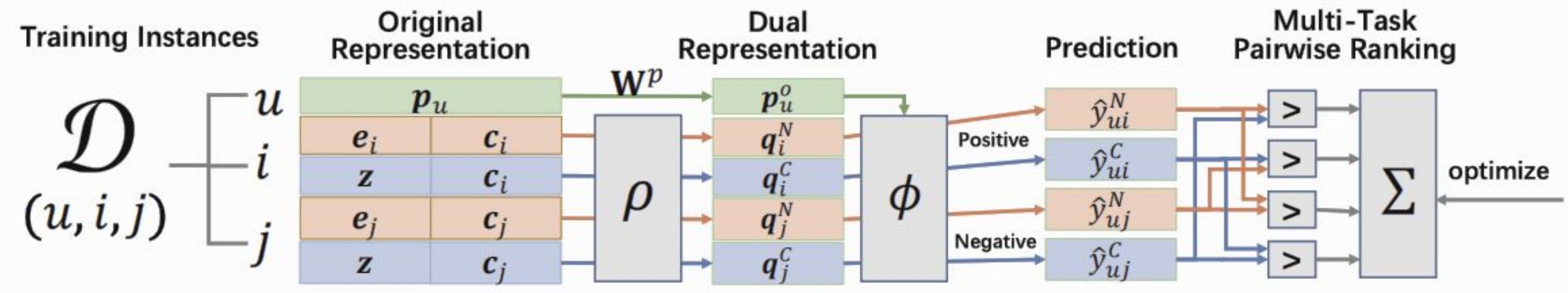
Embedding Layer

Input Feature Vector (sparse)

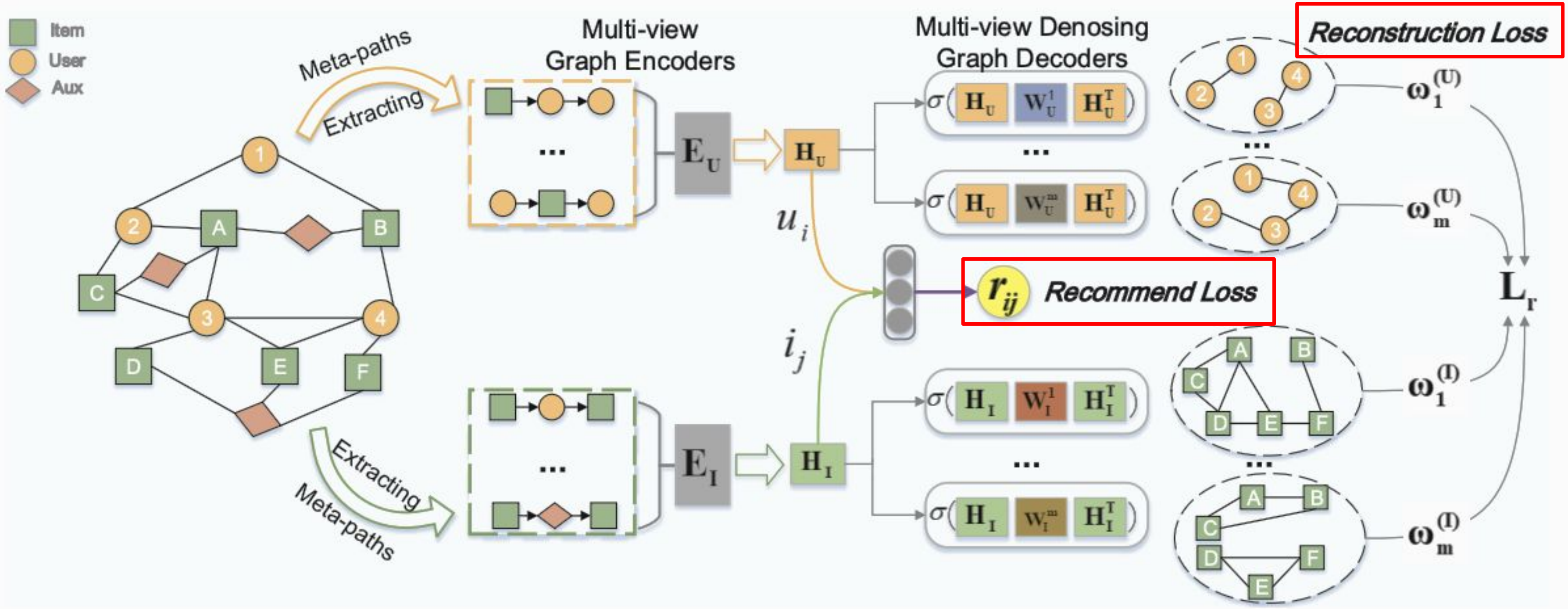
Baselines - LARA



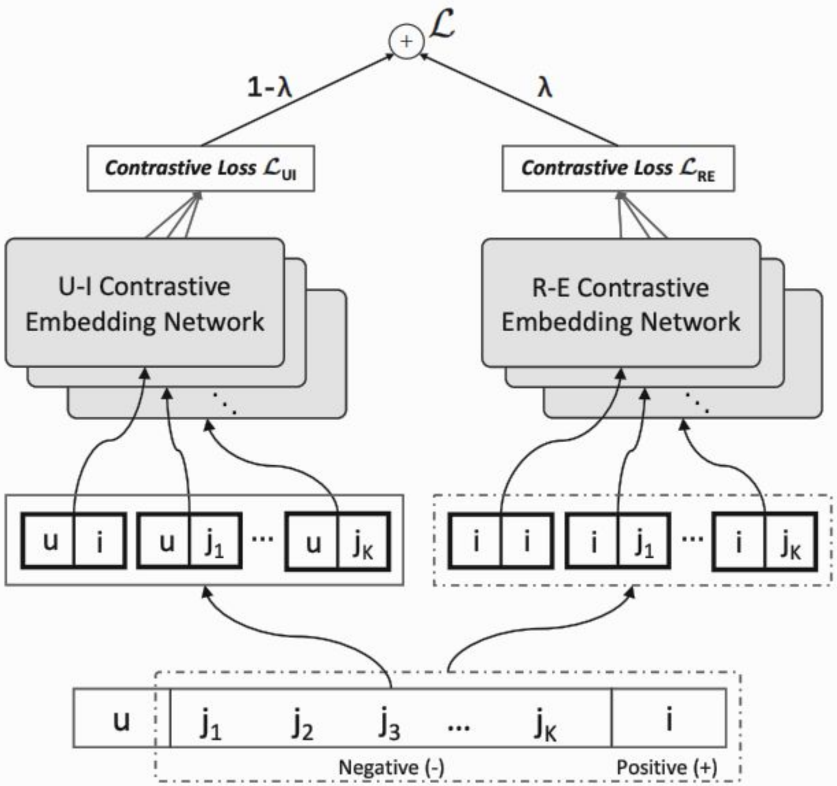
Baselines - MTPR



Baselines - MvDGAE



Baselines - CLCRec



Performances Comparison

NFM: neural FM

LARA: GAN based

MTPR:

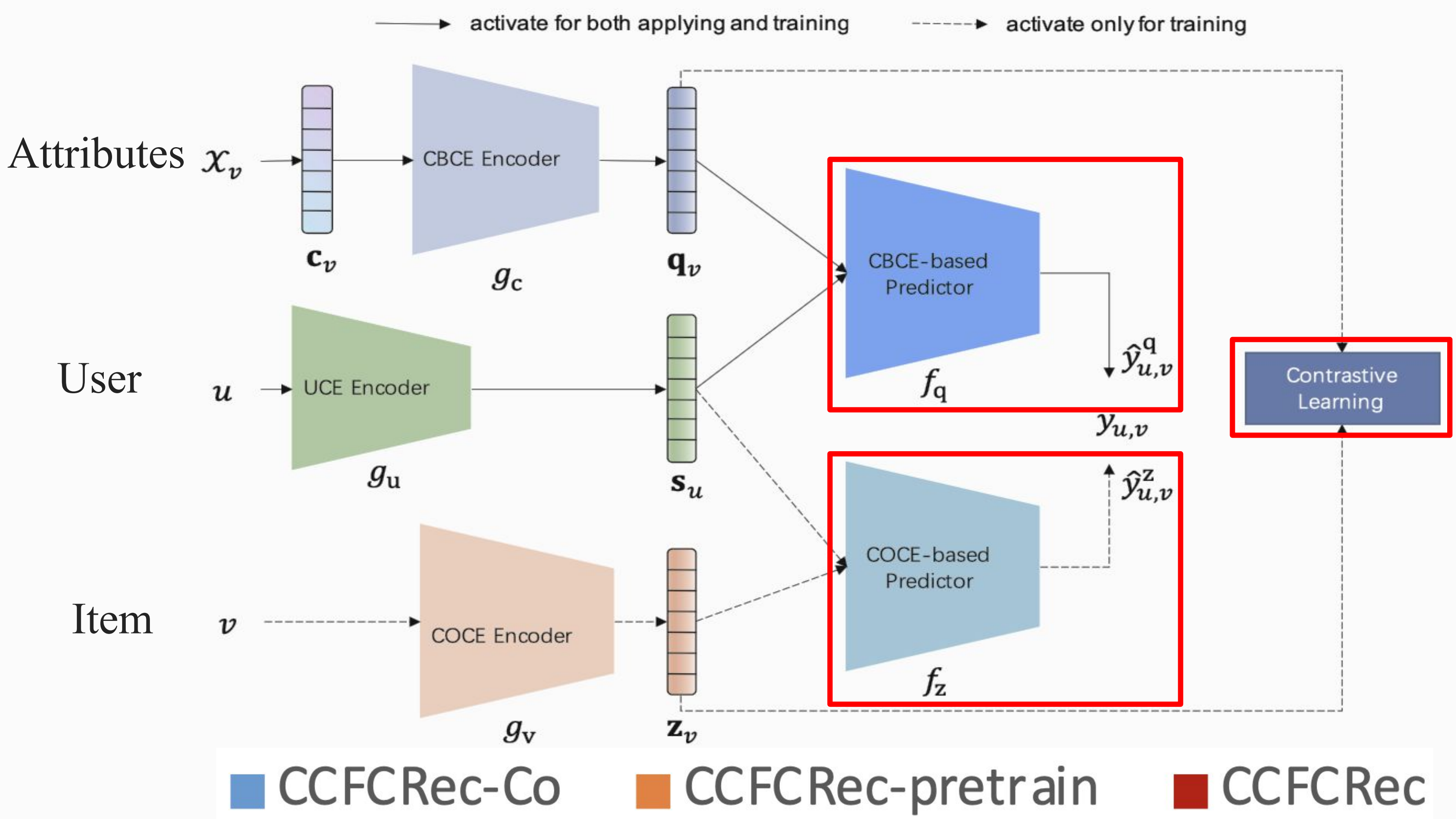
p_u	
e_i	c_i
z	c_i
e_j	c_j
z	c_j

MvDGAE: graph based

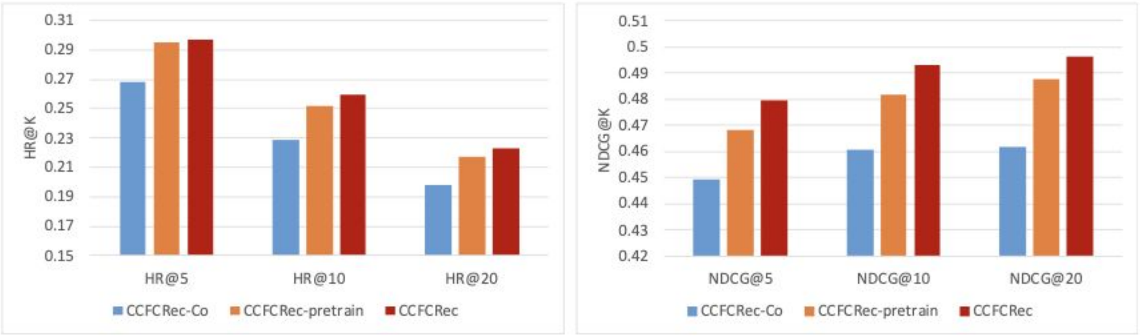
CLCRec: contrastive

Baseline	HR@5	HR@10	HR@20	NDCG@5	NDCG@10	NDCG@20
NFM	0.2119	0.1822	0.1531	0.3683	0.3842	0.3921
LARA	0.2425	0.2165	0.1829	0.4595	0.4541	0.4580
MTPR	0.2701	0.2393	0.2064	0.4504	0.4588	0.4721
MvDGAE	0.2789	0.2453	0.2128	0.4586	0.4660	0.4720
CLCRec	0.2677	0.2371	0.1971	0.4695	0.4820	0.4873
CCFCRec	**0.2969***	**0.2592***	**0.2230***	**0.4798***	**0.4933***	**0.4962***
Improv.	6.06%	5.36%	4.57%	2.15%	2.29%	1.79%
NFM	0.0162	0.0116	0.0089	0.0462	0.0532	0.0612
LARA	0.0140	0.0074	0.0044	0.0370	0.0381	0.0400
MTPR	0.0161	0.0112	0.0083	0.0457	0.0518	0.0587
MvDGAE	0.0161	0.0126	0.0089	0.0456	0.0539	0.0604
CLCRec	0.0229	0.0189	0.0150	0.0646	0.0769	0.0879
CCFCRec	**0.0326***	**0.0260***	**0.0201***	**0.0916***	**0.1074***	**0.1202***
Improv.	29.75%	27.30%	25.37%	29.48%	28.40%	26.87%

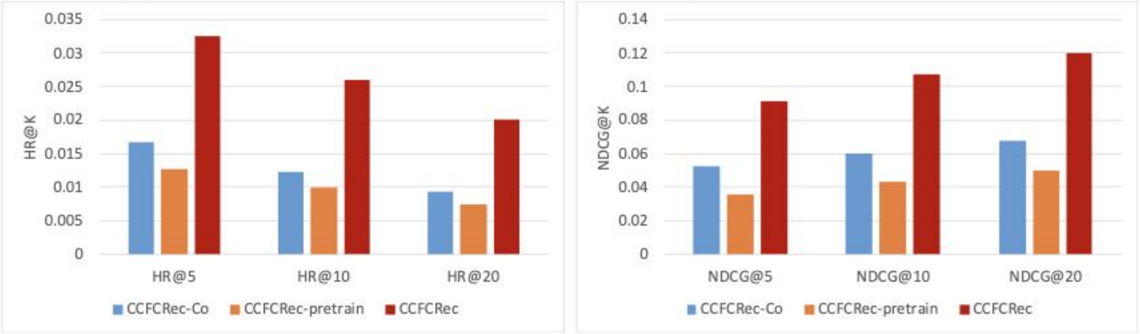
Ablation Study - Model Structure



Ablation Study



(a) HR on ML-20M (b) NDCG on ML-20M

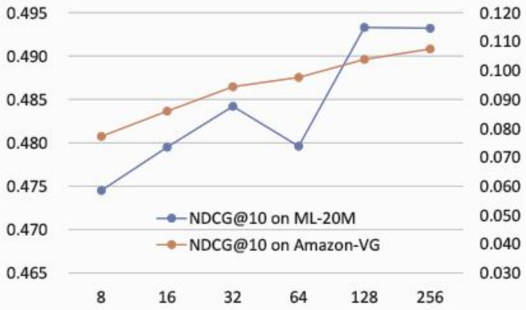


(c) HR on Amazon-VG (d) NDCG on Amazon-VG

CCFCRec-Co CCFCRec-pretrain CCFCRec

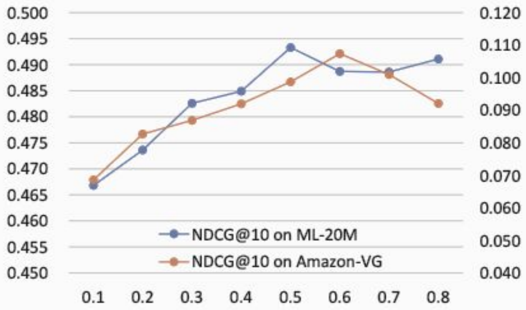
Tuning of Hyper-parameter

embedding dimensionality

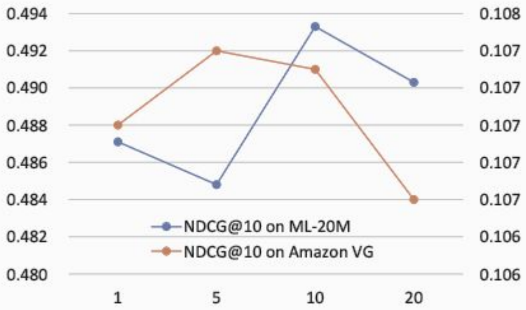


(a) d

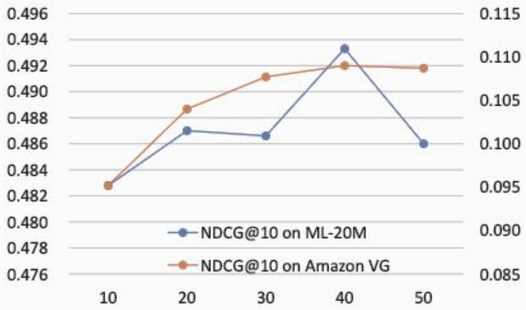
balance factor of the contrastive loss



(b) λ

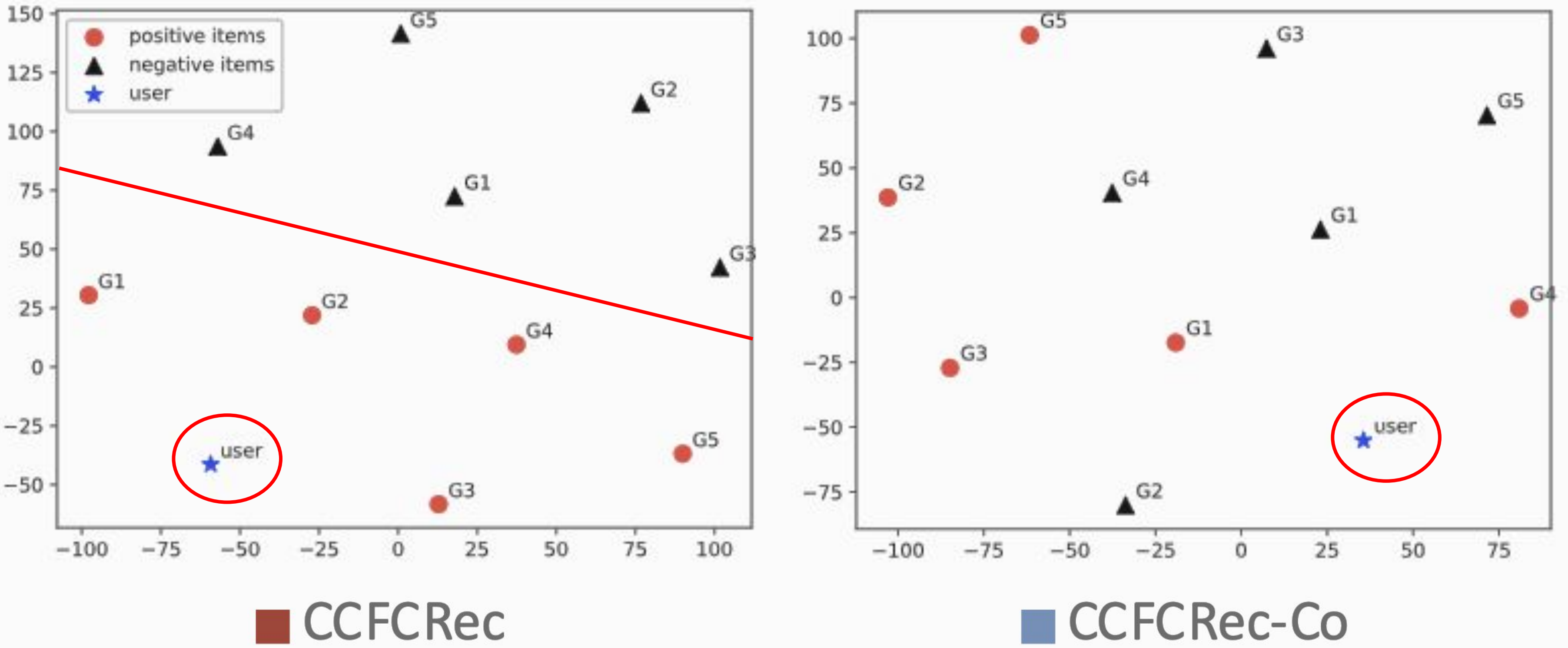


(c) The number of positive samples



(d) The number of negative samples

Case Study


positive items
negative items
user
G1
G2
G3
G4
G5
150
125
100
75
50
25
0
−25
−50
−100
−75
−50
−25
0
25
50
75
100
CCFCRec
CCFCRec-Co

Case Study

User 10430	d^+		d^-		$d^- - d^+$	
	w-Co	w/o-Co	w-Co	w/o-Co	w-Co	w/o-Co
G1	81.53	66.43	137.41	82.28	55.88	15.85
G2	70.97	167.23	204.98	73.66	134.01	-93.57
G3	74.05	123.58	181.43	153.77	107.38	30.19
G4	109.24	68.18	134.97	120.28	25.73	52.10
G5	149.3	184.17	192.45	130.46	43.15	-53.71

The distances from item embedding to the user embedding

w-Co CCFCRec

w/o-Co CCFCRec-Co

Outline

- Introduction
- Method
- Experiment
- Conclusion

Conclusion

Devise a contrastive collaborative filtering (CF) framework

- Content CF module
- Co-occurrence CF module