TPR: Text-aware Preference Ranking for Recommender Systems

Yu-Neng Chuang

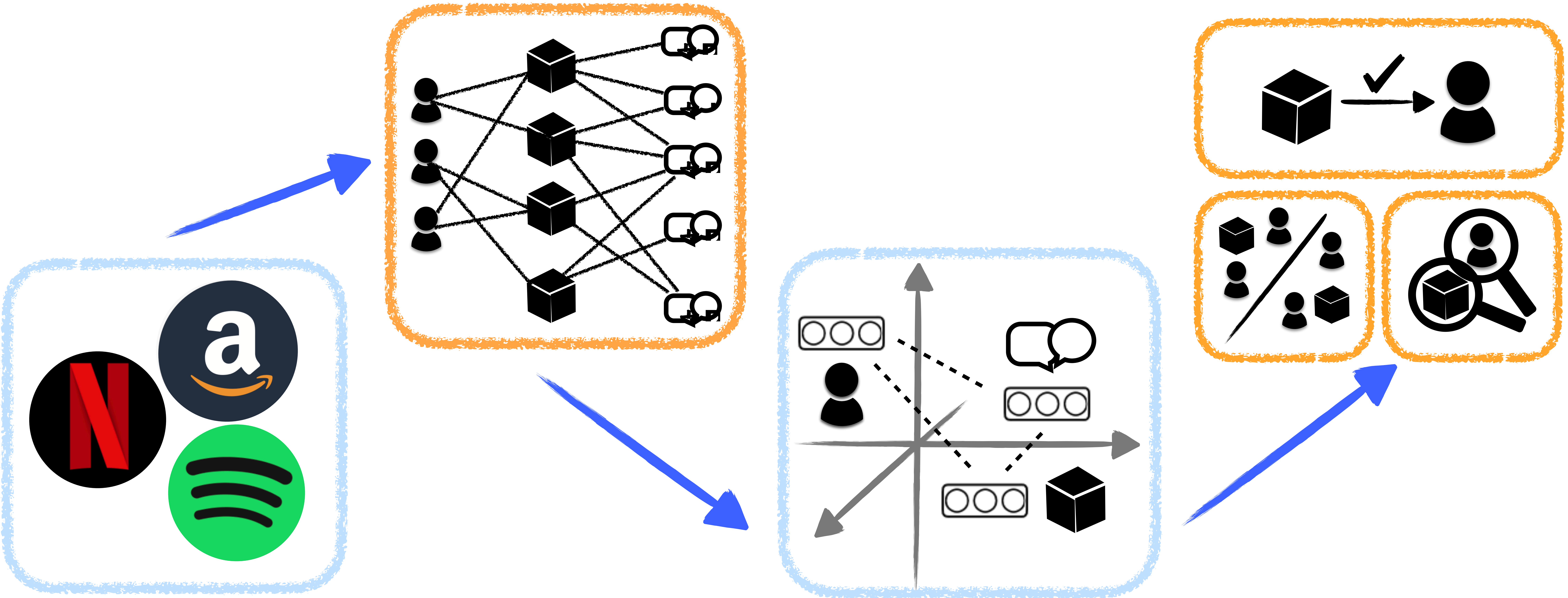
Dept. Computer Science, National Chengchi University

Joint work with Chih-Ming Chen

Prof. Chuan-Ju Wang, Prof. Ming-Feng Tsai,

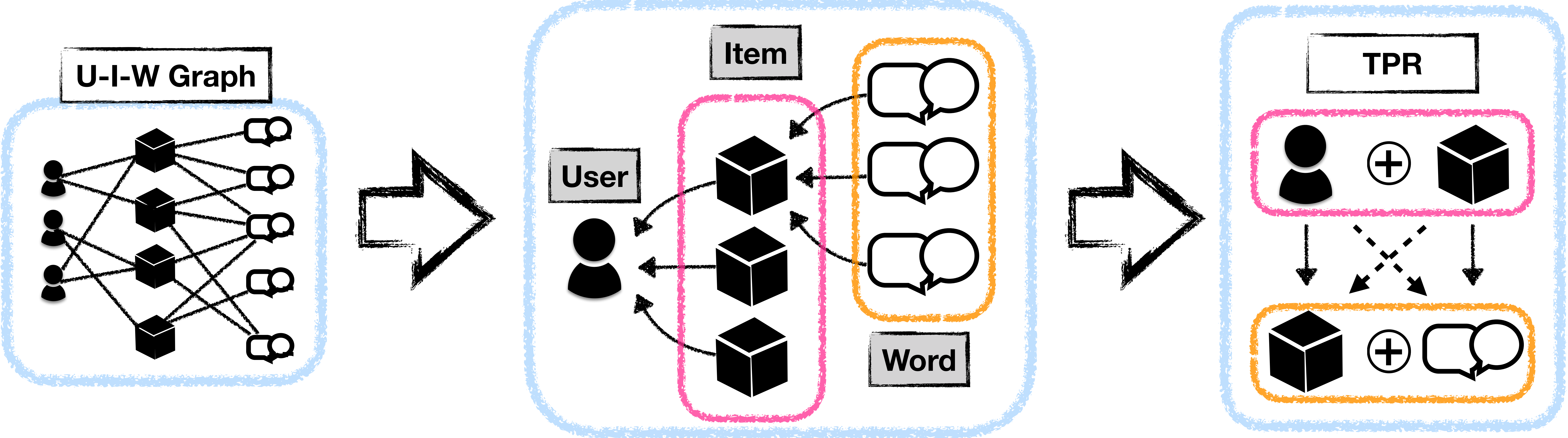
Prof. Yuan Fang, and Prof. Ee-Peng Lim

Introduction



Text-aware Preference Ranking

- Motivation: Seeking a method to jointly learn the feature from user log and text



Text-aware Preference Ranking

- Two ranking structures, IPR and WRR, model the relation of U-I and I-W pairs

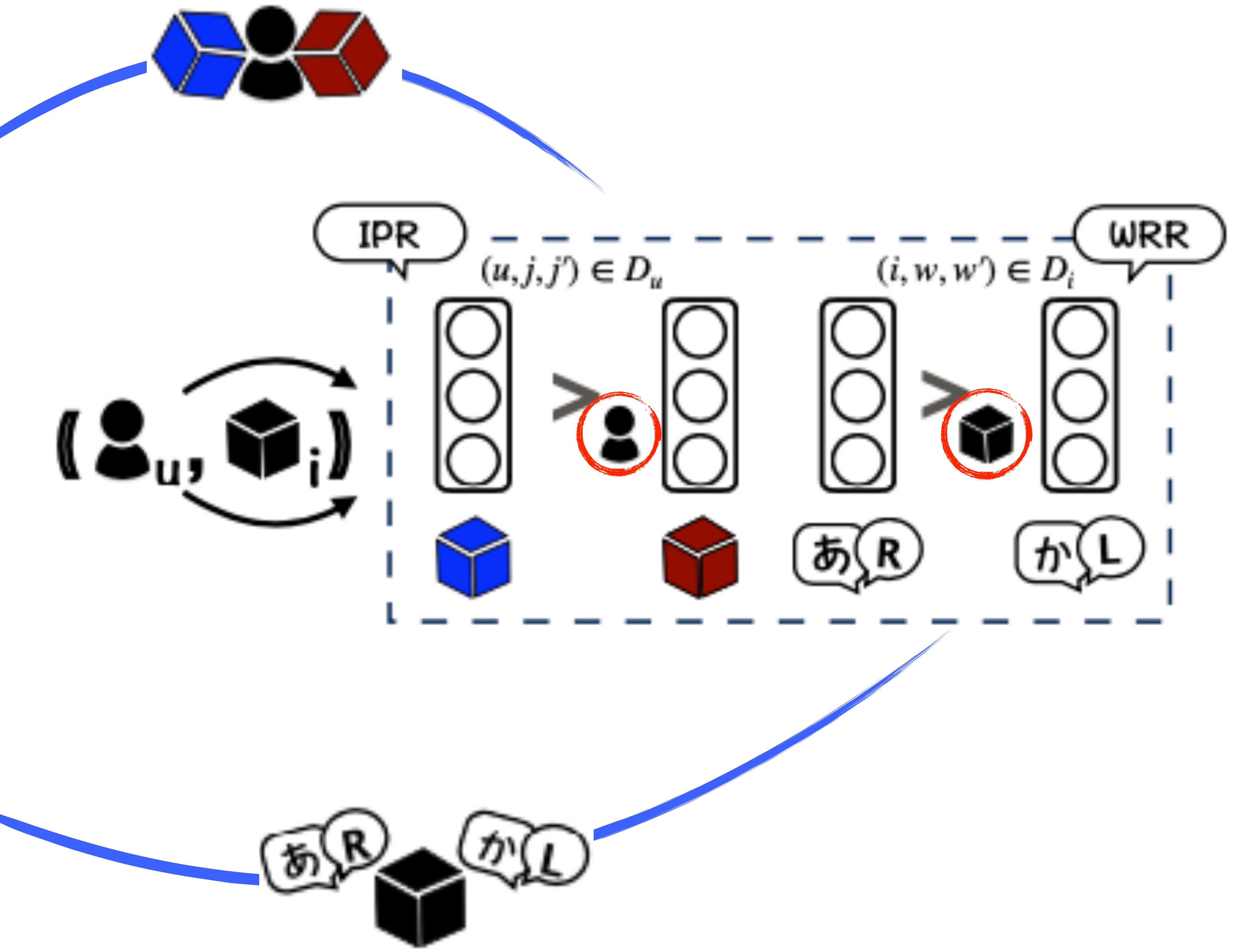
Item Preference Ranking (IPR)

$$p(j >_u j'|\Theta) = \sigma\left(\langle \Theta_u, \Theta_j - \Theta_{j'} \rangle\right)$$

$$O_{\mathrm{TPR}} \equiv \max \prod_{(u,i)\in E_{u,i}} p(\overbrace{>_u}^{\mathrm{IPR}}, \overbrace{>_i}^{\mathrm{WRR}} |\Theta)$$

$$p(w >_i w'|i) = \sigma\left(\langle \Theta_i, \Theta_w - \Theta_{w'} \rangle\right)$$

Word Relatedness Ranking (WRR)



$$\begin{aligned} &p(j >_u j', w >_i w'|\Theta) \\ &= \sigma\left(\langle \Theta_u + \Theta_i, (\Theta_j - \Theta_{j'}) + (\Theta_w - \Theta_{w'}) \rangle\right) \\ &= \sigma\left(\langle \Theta_u, (\Theta_j - \Theta_{j'}) \rangle + \langle \Theta_u, (\Theta_w - \Theta_{w'}) \rangle + \right. \\ &\qquad \left. \langle \Theta_i, (\Theta_j - \Theta_{j'}) \rangle + \langle \Theta_i, (\Theta_w - \Theta_{w'}) \rangle \right), \end{aligned}$$

Text-aware Preference Ranking

- Seeking a method to jointly learn two ranking structures : IPR and WRR

Item Preference Ranking (IPR)

$$p(j >_u j' | \Theta) = \sigma(\langle \Theta_u, \Theta_j - \Theta_{j'} \rangle)$$

$$O_{\mathrm{TPR}} \equiv \max \prod_{(u,i) \in E_{u,i}} p(\overbrace{>_u}^{\mathrm{IPR}}, \overbrace{>_i}^{\mathrm{WRR}} | \Theta)$$

$$p(w >_i w' | i) = \sigma(\langle \Theta_i, \Theta_w - \Theta_{w'} \rangle)$$

Word Relatedness Ranking (WRR)

TPR-OPT

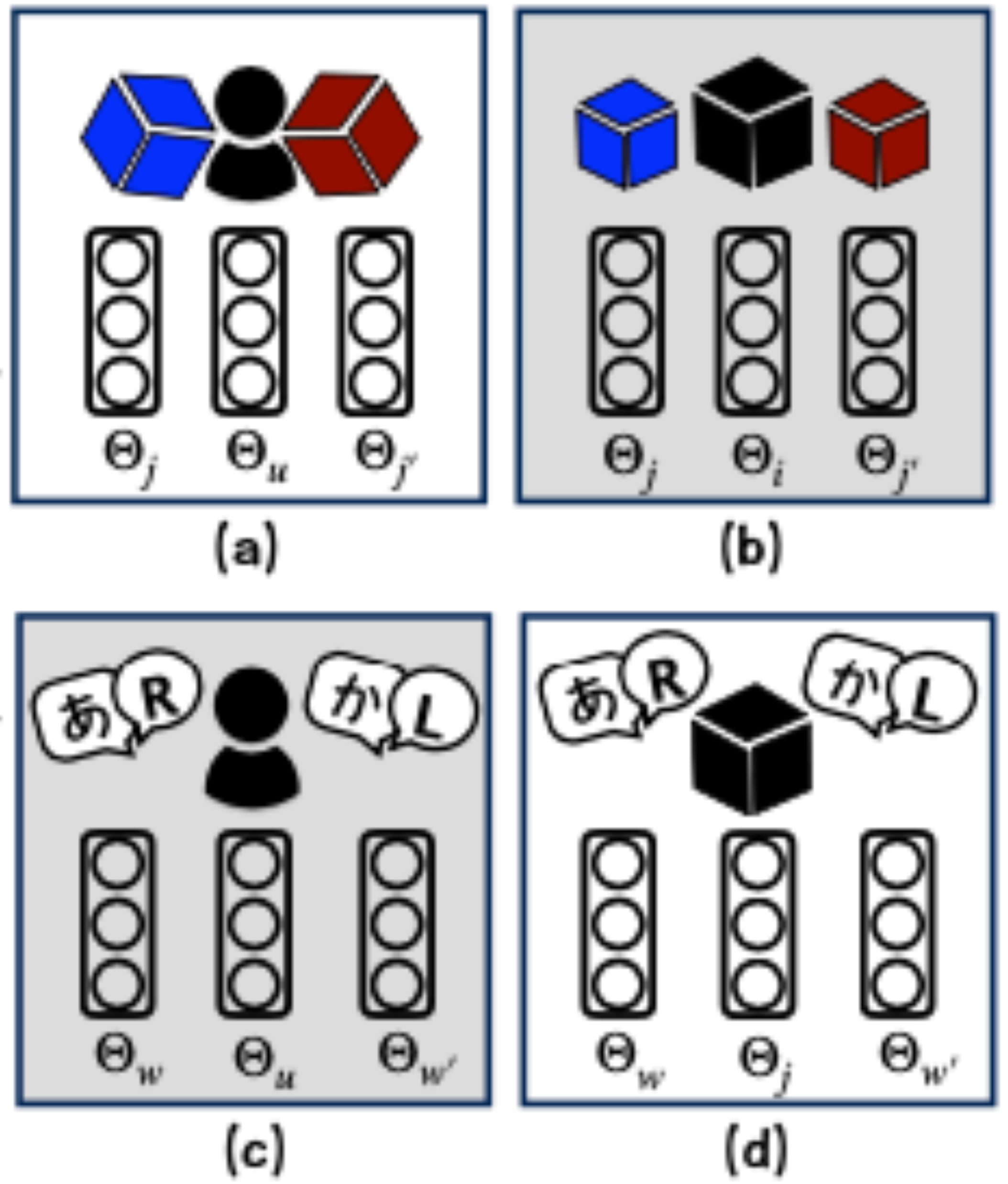
$$:= \ln p(\Theta | >_u, >_i) \propto \ln p(>_u, >_i | \Theta) p(\Theta)$$

$$= \ln \prod_{\substack{(u,j,j') \in D_u, \\ (i,w,w') \in D_i}} p(j >_u j', w >_i w' | \Theta) p(\Theta)$$

$$= \sum_{\substack{(u,j,j') \in D_u, \\ (i,w,w') \in D_i}} \ln \sigma(\langle \Theta_u + \Theta_i, (\Theta_j - \Theta_{j'}) + (\Theta_w - \Theta_{w'}) \rangle) - \lambda_\Theta \|\Theta\|^2$$

λ_{IPR}

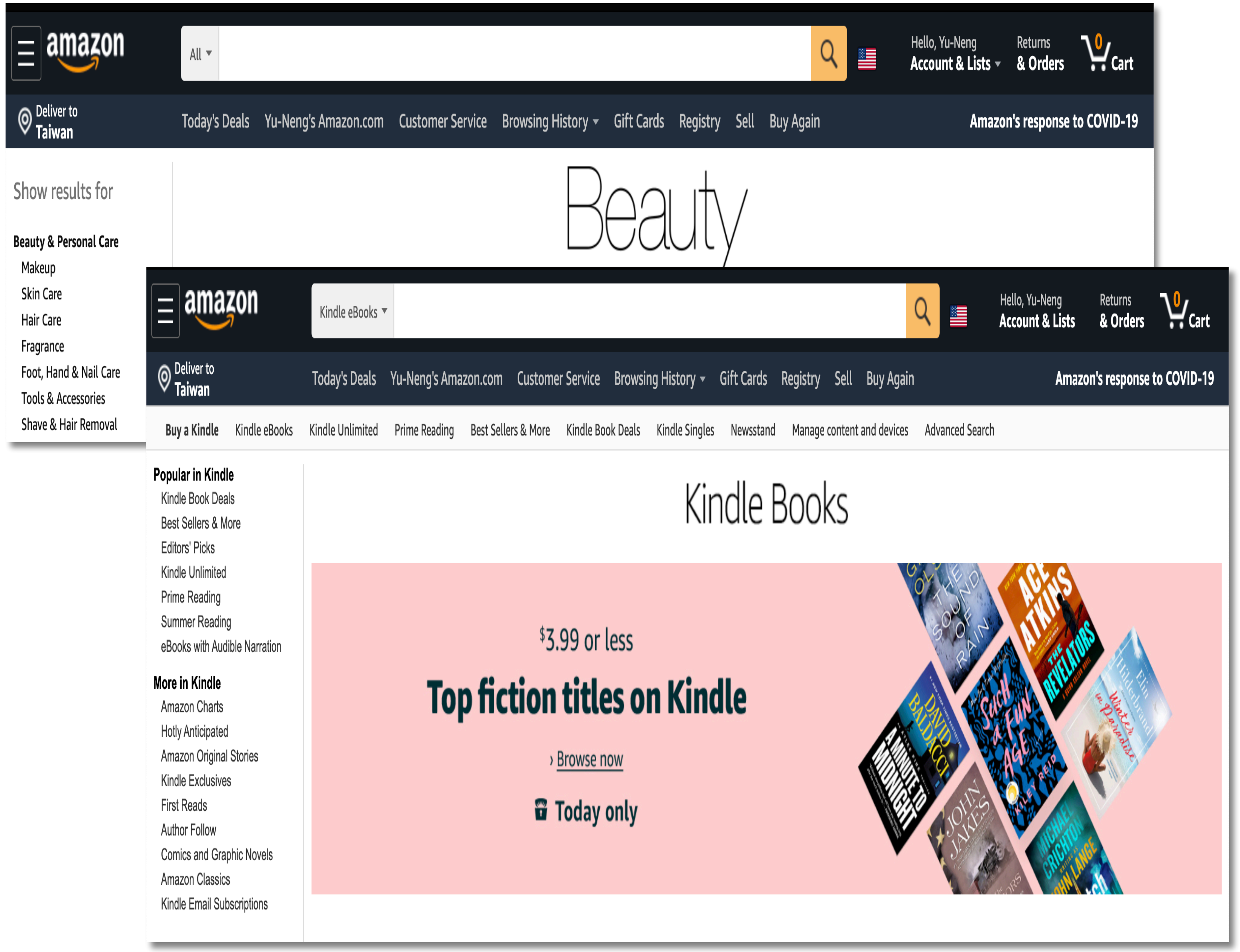
λ_{WRR}



Experiment Results

- **Datasets : Six different public real-world datasets.**

  - Each of the dataset contains
    - A. User-item interaction log (U-I)
    - B. Item with its textual description (I-W)



	Users	Items	Words	U-I edges	I-W edges
Amazon-Magazine	2,825	1,299	6,740	11,685	9,4381
Amazon-Beauty	4,801	4,865	4,115	11,685	159,475
Amazon-Application	11,823	5,554	9,712	42,675	410,079
Amazon-Software	13,634	9,325	11,111	57,793	766,112
Amazon-Fashion	19,875	36,080	5,076	75,596	442,136
Amazon-Kindle	363,303	356,634	36,445	3,334,521	6,794,209

Experiment Results

- Top-N recommendation performance

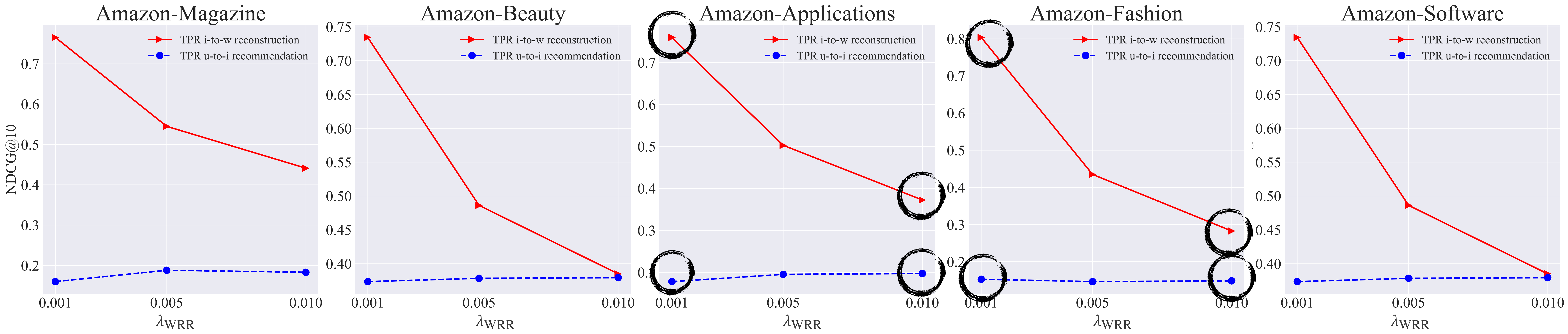
	Amazon-Magazine		Amazon-Beauty		Amazon-Applications		Amazon-Fashion	
	Recall@10	NDCG@10	Recall@10	NDCG@10	Recall@10	NDCG@10	Recall@10	NDCG@10
BPR [20]	0.3306	0.1734	0.4278	0.3468	0.3035	0.1590	0.1563	0.1223
WARP [23]	0.3435	0.1892	0.3468	0.3437	0.3016	0.1655	0.1815	0.1298
SINE [25]	0.0360	0.0083	0.0549	0.0157	0.1283	0.0280	0.0865	0.0181
HPE [3]	0.3419	0.1377	† 0.4773	† 0.3652	† 0.3552	0.1736	† 0.2126	† 0.1393
GATE [13]	0.2720	0.0489	0.3940	0.0812	0.1336	0.0225	0.0819	0.0186
CKE [26]	0.3838	0.2061	0.4208	0.3450	0.2933	0.1562	0.1581	0.1230
KGAT [22]	† **0.4156**	† **0.2156**	0.4321	0.3558	0.3213	† 0.1862	0.1862	0.1268
TPR ($\lambda_{\text{WRR}} = 0.001$)	0.3681	0.1599	***0.4950**	*0.3735	*0.3937	*0.1779	*0.2394	*0.1525
TPR ($\lambda_{\text{WRR}} = 0.005$)	0.4101	0.1880	*0.4925	*0.3783	***0.4097**	*0.1951	***0.2270**	*0.1462
TPR ($\lambda_{\text{WRR}} = 0.01$)	0.4182	0.1840	*0.4840	***0.3793**	*0.3997	***0.1971**	*0.2258	***0.1482**
Improv. (%)	−0.62%	−12.80%	+3.70%	+3.86%	+15.34%	+5.85%	+6.77%	+6.38%

	Amazon-Software		Amazon-Kindle		Course		SG-OPN	
	Recall@10	NDCG@10	Recall@10	NDCG@10	Recall@10	NDCG@10	Recall@10	NDCG@10
BPR [20]	† 0.3669	0.1779	0.4414	0.2097	0.5731	0.4129	0.1008	0.0339
WARP [23]	0.3423	0.1556	† 0.5461	† 0.3392	0.5340	0.3639	† 0.2623	† 0.1131
SINE [25]	0.0976	0.0257	0.2812	0.1394	0.0357	0.0168	0.0412	0.0150
HPE [3]	0.3658	0.1405	0.5228	0.2803	0.3391	0.2294	0.0047	0.0040
GATE [13]	0.1326	0.0202	-	-	0.4477	0.3170	0.0010	0.0035
CKE [26]	0.3448	0.1497	-	-	† 0.6094	† **0.4583**	0.1050	0.0837
KGAT [22]	0.3907	† 0.1847	-	-	0.5902	0.4294	0.1473	0.0512
TPR ($\lambda_{\text{WRR}} = 0.001$)	*0.3898	0.1615	*0.5682	*0.3448	0.5735	0.4177	***0.3110**	***0.1411**
TPR ($\lambda_{\text{WRR}} = 0.005$)	*0.4252	0.1844	*0.6065	*0.3722	0.6014	0.4422	*0.3094	*0.1392
TPR ($\lambda_{\text{WRR}} = 0.01$)	***0.4319**	***0.1956**	***0.6164**	***0.3804**	***0.6155**	0.4468	*0.3086	*0.1434
Improv. (%)	+17.71%	+5.90%	+12.87%	+12.14%	+1.00%	−2.50%	+18.56%	+24.75%

Experiment Results

Item-to-word reconstruction
User-to-item recommendation

- Sensitivity check on regularization term (λ_{WRR} & λ_{IPR})



Allen Chung, National Chengchi University

- A smaller λ_{WRR} can benefit modeling the relation between item and text
- A larger λ_{IPR} can prevent overfitting problems on modeling User-Item relation
- A trade-off parameter provides the flexibility on modeling different tasks.

Conclusion

- Design a framework on joint association of user-item interaction and relations between items and associated text
- TPR comprehensively modeling four types of ranking relations on the six different tasks to attest the effectiveness of the learned embeddings
- TPR achieves high modeling efficiency in terms of execution time and memory usage.

TPR Implementation

- TPR is now publicly available on GitHub:
  - Repo: https://github.com/cnclabs/codes.tpr.rec
- TPR is implemented on the framework of SMORe:
  - Repo: https://github.com/cnclabs/smore

Thanks For Your Listening

Any Question ?