DCDIR: A Deep Cross-Domain Recommendation System for Cold Start Users in Insurance Domain

Ye Bi Liqiang Song Mengqiu Yao
Zhenyu Wu Jianming Wang Jing Xiao
Ping An Technology Shenzhen Co., Ltd

Speaker: **Ye Bi**

About the Speaker

- Senior Algorithm Engineer in Ping An Technology Shenzhen Co., Ltd
- **Major Research Direction:** Deep Learning, Recommendation System, Knowledge Graph

Introduction

PingAn Jinguanjia is one of the most popular comprehensive applications in China. In addition to traditional e-commerce products, it also provides financial products.

(a) Home Page

(b) Nonfinanacial Domain

(c) Insurance Domain

1. Introduction
   - Challenges
   - Contributions
   - Framework
   - Target Domain
   - Source Domain
   - Dataset
   - Performance

Challenges

Recommending insurance products online is challenging

- Insurance policies are so complex that ordinary users are relatively lack of knowledge to understand them.
- Insurance products are typically bought to be used for a long time period (e.g. one year for car insurance), so there exists data sparsity and cold start problem.

1. Introduction
   - Challenges
   - Contributions
   - Framework
   - Target Domain
   - Source Domain
   - Dataset
   - Performance

Contributions

Our contributions in this paper are as follows:

- To the best of our knowledge, this is the first work to utilize cross-domain mechanism to give personalized recommendations for cold start users in insurance domain.
- For the complexity of insurance products, we design a meta-path based method to learn more effective latent user and item features, revealing reasons behind recommendations.
- We conduct experiments on our company's scenarios, the results prove the efficacy of DCDIR over several baselines.

Framework

To provide recommendations to cold start users, we propose DCDIR. As shown in Figure 1, DCDIR contains three main parts: learning user latent features in two domain, mapping of user latent features.

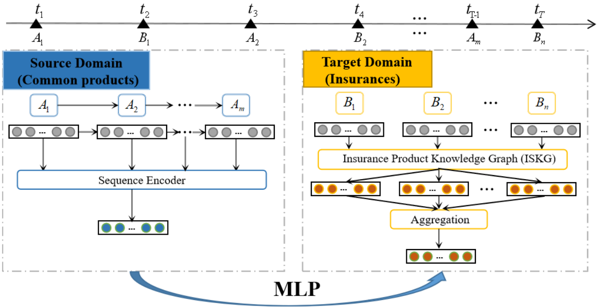
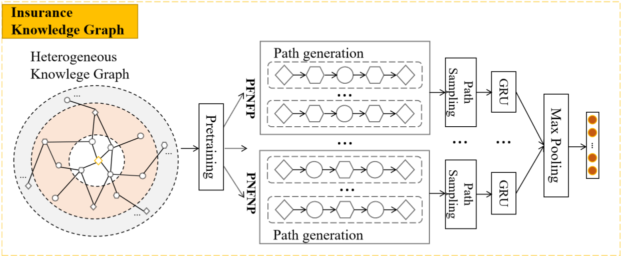


Figure 1: The Framework of DCDIR

Target Domain

- Pretrain KG by TransD
- Generate meta-paths connecting interacted items and target item
- Select high-quality meta-paths
- Model each meta-path by GRU



Target Domain

We properly design two meta-paths based on our scenario, where we fix entity type and path length.

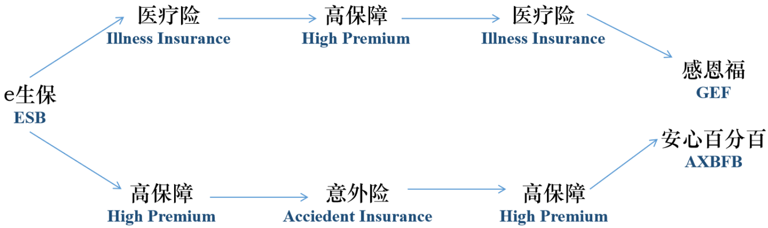


Figure 2: Meta-Path Example

Target Domain

We use *top-K* sampling module to select K useful paths. For a given path $p_{e_1,e_L} = [e_1, e_2, \ldots, e_L]$, we define a score function:

$$s_{e_1,e_L} = \mathrm{softmax}(\frac{P}{|\mathcal{NI}_u^t|}) + \frac{\boldsymbol{e}_L^T}{\|\boldsymbol{e}_L\|} \sum_{i=1}^{L-1} \frac{\boldsymbol{e}_i}{\|\boldsymbol{e}_i\|},$$

where P is e_1's position in $\mathcal{NI}_u^t$. The first part is to measure interaction time, the second part is to measure the similarity between the path and the target item.

Source Domain

Suppose there are n words in i's content c_i. We utilize word2vec to obtain word vectors, which are represented as $\{\boldsymbol{w}_k^i\}_{k=1}^n$. Then we get the final item embedding by:

$$\boldsymbol{i} = \text{max-pooling}(\text{concat}(\{\boldsymbol{w}_k^i\}_{k=1}^n)).$$

To model user latent feature $\boldsymbol{u}^s$, we employ GRU over $\mathcal{NI}_u^s$, and let $\boldsymbol{u}^s = \boldsymbol{h}_n^{\text{GRU}}(\mathcal{NI}_u^s)$

Dataset

We build and release a sub-dataset (named JGJISNF) from a comprehensive e-commerce dataset that contains about 20 million users pursue logs from June 1st 2018 to May 31th 2019.

Table 1: Statistics of the JGJISNF dataset.

IS-domain (Target domain)		NF-domain (Source domain)	
#Items	42	#Items	3,836
#Interactions	300,000	#Interactions	600,000
#KG relations	7		
#KG enitities	77		
#KG triples	282		
#Overlapped-users		21,016	
#Training-sequences		12,437	
#Test-sequences		4,218	
#Validation-sequences		4,298	

Performance

Table 2: Performance comparison in Recall@3 and NDCG.

η	10%		20%	
Method	NDCG	Recall@3	NDCG	Recall@3
BPR	0.27011	0.06418	0.27105	0.06518
GRU4REC	0.23923	0.02143	0.25964	0.07768
EMCDR-BPR	0.27343	0.07291	0.27342	0.07291
EMCDR-GRU	0.26775	0.11794	0.26801	0.11794
DCDIR-V1	0.34781(-4.66%)	0.17321(-6.28%)	0.35196	0.18016
DCDIR-V2	0.36278(-13.05%)	0.19159(-27.60%)	0.37021	0.19388
DCDIR	0.39394(-3.95%)	0.25185(-5.31%)	0.39741	0.25227
DCDIR vs. best	8.59%	26.23%	7.35%	24.96%

Performance

Table 3: Performance comparison in Recall@3 and NDCG (cont.).

η	50%		100%	
Method	NDCG	Recall@3	NDCG	Recall@3
BPR	0.27133	0.06451	0.27325	0.07124
GRU4REC	0.30725	0.09611	0.30623	0.08602
EMCDR-BPR	0.27342	0.07325	0.27347	0.07325
EMCDR-GRU	0.29056	0.11996	0.31288	0.12298
DCDIR-V1	0.35653	0.18078	0.36481	0.18481
DCDIR-V2	0.40273	0.24504	0.40925	0.26461
DCDIR	0.40773	0.26268	0.41016	0.26597
DCDIR vs. best	1.24%	7.20%	0.22%	0.51%

Performance

Table 4: Ablation study of path number and path strategy.

ISKG module		Metrics	
parameter	value	NDCG	Recall@3
path_num	10	0.36611	0.18207
	20	0.39394	0.25185
	30	0.38435	0.18541
path_strategy	‘topk’	0.39394	0.25185
	‘random’	0.34624	0.16065

Q&A

If you have any questions, please email to magicyebi@163.com

Acknowledgments

Thank you for your attention!