A Collective Variational Autoencoder for Top-N Recommendation with Side Information

Yifan Chen [1] Maarten de Rijke [2]

October 6, 2018

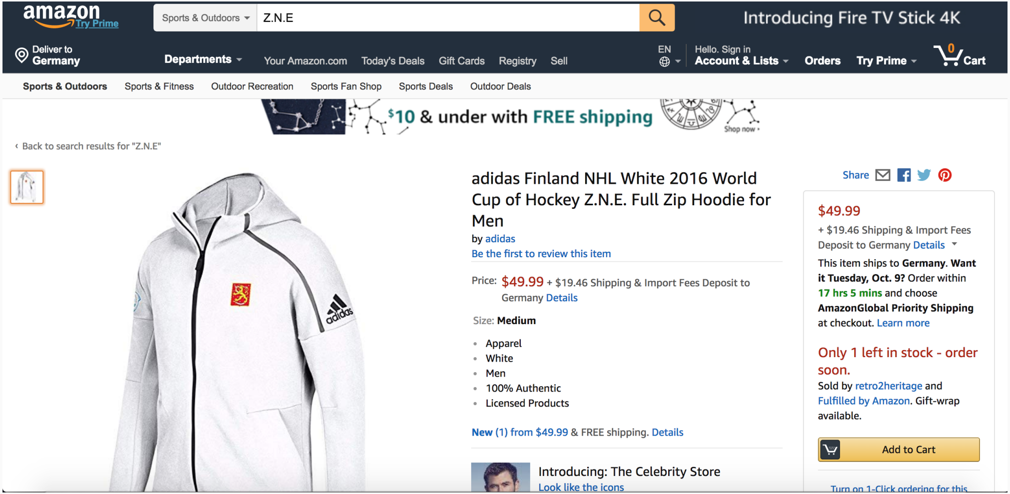
[1]University of Amsterdam
y.chen4@uva.nl

[2]University of Amsterdam
derijke@uva.nl

Roadmap

Introduction

Recommendation with side information



Recommendation with side information

Side information

- information associated with users or **items**
  - item-side information are more often utilized

Recommendation with side information

- increasingly **availability** of side information
- provide additional **information**
- overcome user rating **sparsity**

Related work

Recommendation with side information

Existing methods

- linear methods
- non-linear methods
    - deep autoencoder: using deep neural network to **extract item representation** from side information

Utilizing deep Autoencoders for recommendation

Denoise Autoencoder

- Collaborative Deep Learning (CDL) Wang et al. [2015]
- marginalized Denoising Autoencoder (mDA) Li et al. [2015]

Variational Autoencoder

- collaborative filtering Variational Autoencoder (cfVAE) Li and She [2017]
  - show state-of-the-art performance

Utilizing deep Autoencoders for recommendation

Suffer from high-dimensionality

- determine the input scale of network
- dominate the overall size of the model

Solution

- **collective Variational Autoencoder (cVAE)**
    - overcome the impact from **high-dimensionality**
    - take the advantage of deep learning

Method

Preliminaries

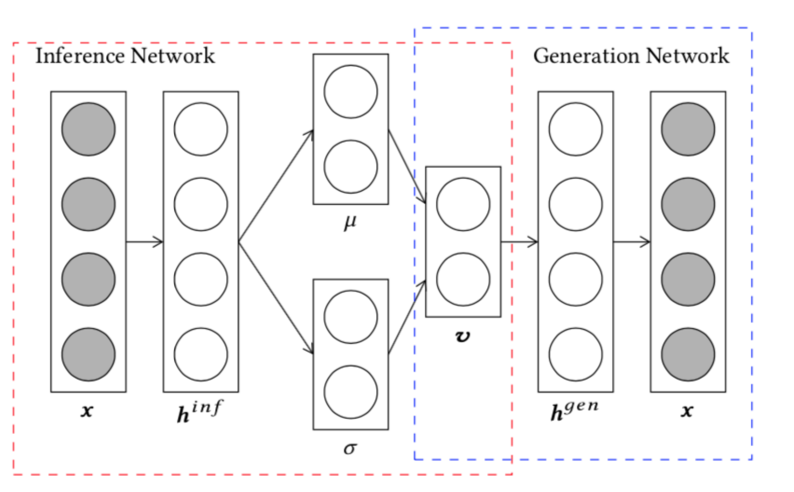
Notation

- suppose we have m users, n items and d dimensions for side information
- user rating: $Y \in \mathbb{R}^{m \times n}$
- item feature: $X \in \mathbb{R}^{d \times n}$

Assumption

1. do not distinguish item feature with side information
2. assume item feature is a vector with **numerical values**
3. side information is high-dimensional: $d \geq n$
4. assume user rating is **binarized** (typical assumption for implicit feedback)

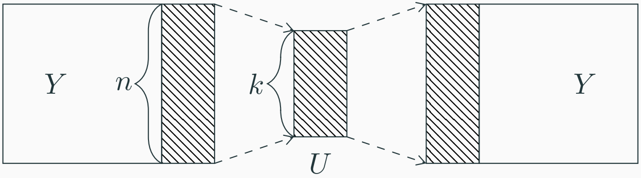
Variational Autoencoder (VAE) Kingma and Welling [2013]



Sparse LInear Method (SLIM) Ning and Karypis [2011]

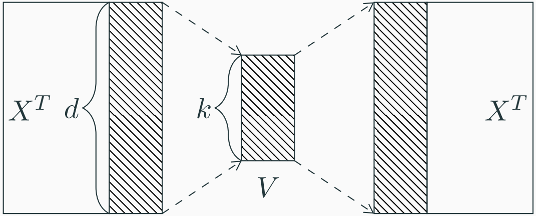
$$\begin{aligned} \min_{S} \quad & \|Y - YS\|_F^2 + \frac{\beta}{2}\|S\|_F^2 + \lambda\|S\|_1 \\ \text{s.t.} \quad & S \geq 0, \mathrm{diag}(S) = 0 \end{aligned}$$

- reproduce the rating matrix: $Y \sim YS$, which works similarly as User-side Autoencoder (UAE)

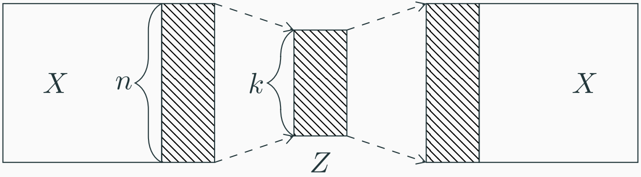


Motivation

Item-side Autoencoder (IAE)



Feature-side Autoencoder (FAE)



Motivation

collective Sparse LInear Method (cSLIM) Ning and Karypis [2012]

$$\min_{S} \quad \|Y - YS\|_F^2 + \alpha\|X - XS\|_F^2 + \frac{\beta}{2}\|S\|_F^2 + \lambda\|S\|_1$$
$$\text{s.t.} \quad S \geq 0, \operatorname{diag} S = 0$$

- $\|Y - YS\|_F^2$: works similarly as IAE
- $\|X - XS\|_F^2$: works similarly as FAE
- collective learning: both X and Y are recovered by learning S

We are inspired to propose **collective Variational Autoencoder (cVAE)**

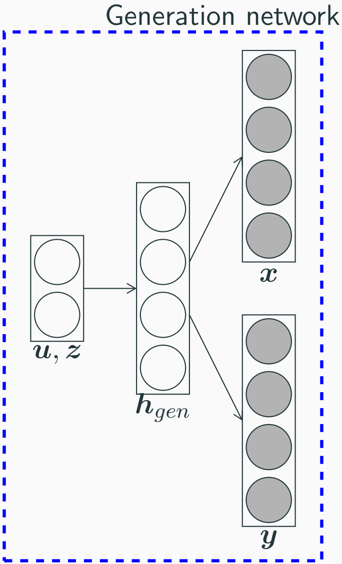
Generation network

for each user $j = 1, \ldots, m$:

1. draw $\boldsymbol{u}_j \sim \mathcal{N}(0, I)$;
2. draw $\boldsymbol{y}_j \sim Bernoulli\,(\varsigma(f_\theta(\boldsymbol{u}_j)))$

for each dimension of side information $j = 1, \ldots, d$:

1. draw $\boldsymbol{z}_j \sim \mathcal{N}(0, I)$;
2. draw $\boldsymbol{x}_j \sim \mathcal{N}(f_\theta(\boldsymbol{z}_u), I)$.



Inference network

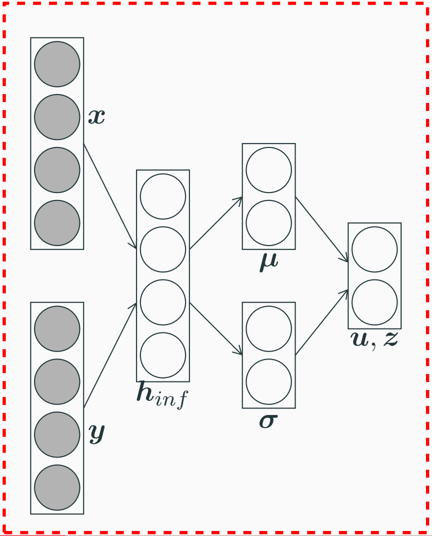
for each user $j = 1, \ldots, m$

1. $\boldsymbol{\mu}_j = \mu(f_\phi(\boldsymbol{y}_j))$
2. $\boldsymbol{\sigma}_j = \sigma(f_\phi(\boldsymbol{y}_j))$

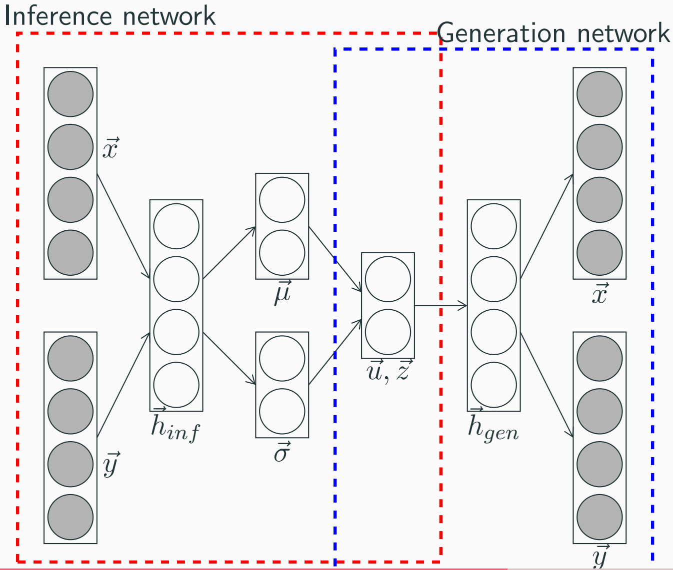
for each dimension of side information $j = 1, \ldots, d$

1. $\boldsymbol{\mu}_{m+j} = \mu(f_\phi(\boldsymbol{x}_j))$
2. $\boldsymbol{\sigma}_{m+j} = \sigma(f_\phi(\boldsymbol{x}_j))$

Inference network



collective Variational Autoencoder (cVAE)



collective Variational Autoencoder (cVAE)

Summarization

- **collective** learning: only one VAE
- **heterogeneous** input: utilize both user rating and side information
- overcome **high-dimensionality** of side information

Training: pre-train is important for DNN

- pre-train the network with item feature
- fine-tune the network with user rating

Experiment

Datasets

Table 1: Statistics of the datasets used.

Dataset	#User	#Item	#Rating	**#Dimension**	#Feature
Games	5,195	7,163	96,316	20,609	5,151,174
Sports	5,653	11,944	86,149	31,282	3,631,243

Compared method

Linear method

- cSLIM (Ning and Karypis [2012]): collective Sparse Linear Method

Non-linear methods

- cfVAE (Li and She [2017]): collaborative Variational Autoencoder, UAE+IAE;
- rVAE (Liang et al. [2018]): Variational Autoencoder using ratings only, UAE;
- fVAE (Our): Variational Autoencoder using side information only, FAE;
- cVAE (Our): collective Variational Autoencoder, UAE+FAE.

Experimental results

Table 2: Results on Games dataset

Method	Rec@5	Rec@10	Rec@15	Rec@20	MAP@5	MAP@10	MAP@15	MAP@20
cSLIM	0.0761	0.1162	0.1474	0.1734	0.0590	0.0337	0.0240	0.0188
cfVAE	0.0685	0.1065	0.1359	0.1608	0.0519	0.0298	0.0212	0.0165
rVAE	0.0137	0.0206	0.0270	0.0375	0.0106	0.0060	0.0043	0.0034
fVAE	0.0495	0.0796	0.1072	0.1276	0.0390	0.0230	0.0167	0.0131
cVAE	**0.0858***	**0.1376****	**0.1731****	**0.2081****	**0.0668***	**0.0394****	**0.0279****	**0.0218****

Conclusion

What have we done?

1. we propose a collective Variational Autoencoder (cVAE) to utilize high-dimensional side information to address rating sparsity for top-N recommendation
2. cVAE is the combination of a UAE and a FAE
3. cVAE can be regarded as the non-linear generalization of cSLIM

What should we do next?

1. utilize side information associated with user;
2. relax required assumption that side information is in accordance with user ratings for measuring item similarities
   - currently, cVAE performs poorly without this assumption
3. cVAE actually has two VAEs but they share the network parameters
   - $p_\theta(\boldsymbol{x} \mid \boldsymbol{z}), \quad p_\theta(\boldsymbol{y} \mid \boldsymbol{u})$
   - can we directly assume $p_\theta(\boldsymbol{y} \mid \boldsymbol{u}, \boldsymbol{x})$?

References i

D. P. Kingma and M. Welling. Auto-encoding variational bayes. *CoRR*, abs/1312.6114, 2013.

S. Li, J. Kawale, and Y. Fu. Deep collaborative filtering via marginalized denoising auto-encoder. In *CIKM*, pages 811–820. ACM, 2015.

X. Li and J. She. Collaborative variational autoencoder for recommender systems. In *SIGKDD*, pages 305–314. ACM, 2017.

D. Liang, R. G. Krishnan, M. D. Hoffman, and T. Jebara. Variational autoencoders for collaborative filtering. In *WWW*, pages 689–698. ACM, 2018.

X. Ning and G. Karypis. SLIM: sparse linear methods for top-n recommender systems. In *ICDM*, pages 497–506. IEEE, 2011.

X. Ning and G. Karypis. Sparse linear methods with side information for top-n recommendations. In *RecSys*, pages 155–162. ACM, 2012.

H. Wang, N. Wang, and D. Yeung. Collaborative deep learning for recommender systems. In *SIGKDD*, pages 1235–1244. ACM, 2015.

Acknowledgments

Source code. Source code to reproduce the experiments in this paper is available at `https://github.com/shikamaruChen/cVAE`.

Acknowledgements. This research was partially supported by NSFC Grant No. 61872446.