Learning both Vertical and Horizontal Dimensions of Feature Hierarchy for Effective Recommendation

Zhu Sun, **Jie Yang**, Jie Zhang, Alessandro Bozzon

AAAI 2017

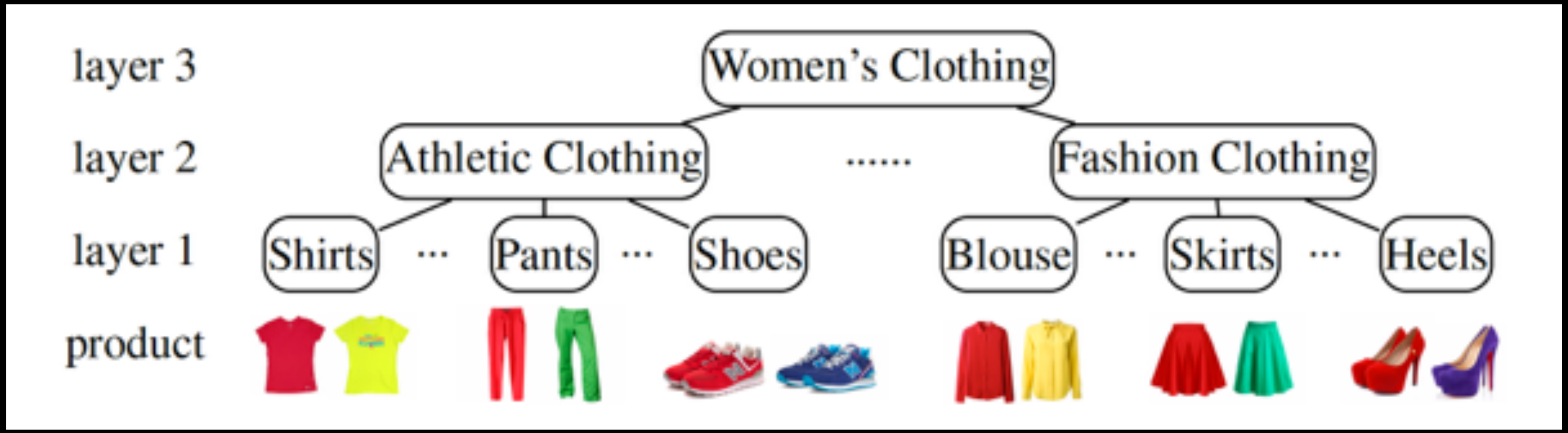
NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

TU Delft

Main contributions

- Model two types of semantically rich feature relationships in horizontal dimension, i.e., ***alternative*** and ***complementary***.
- Propose a unified framework (HieVH) that ***seamlessly*** fuses ***vertical*** and ***horizontal*** dimensions of feature hierarchy for effective recommendation.

Intuition



- A user prefers athletic style: tend to buy more Athletic clothing, e.g., athletic shoes and pants to match each other, instead of Fashion clothing, e.g., heels or skirts.

- Relationships of sibling/cousin features
  - Athletic and Fashion: alternative
  - Shoes and pants: complementary
  - Shoes and heels: alternative

Definition

- Item Relationships

Item i and j are alternative iff $P(e_i|e_j)<P(e_i)$ and $P(e_j|e_i)<P(e_i)$.

Item i and j are complementary iff $P(e_i|e_j)>P(e_i)$ and $P(e_j|e_i)>P(e_i)$.

Measure feature relationships

- Item Co-occurrence

$$\mathrm{IC}(i, j) = \frac{P(e_i \cap e_j)}{P(e_i) \times P(e_j)}$$

  [Theorem]
  i and j are alternative iff *IC(i, j)<1*
  i and j are complementary iff *IC(i, j)>1*

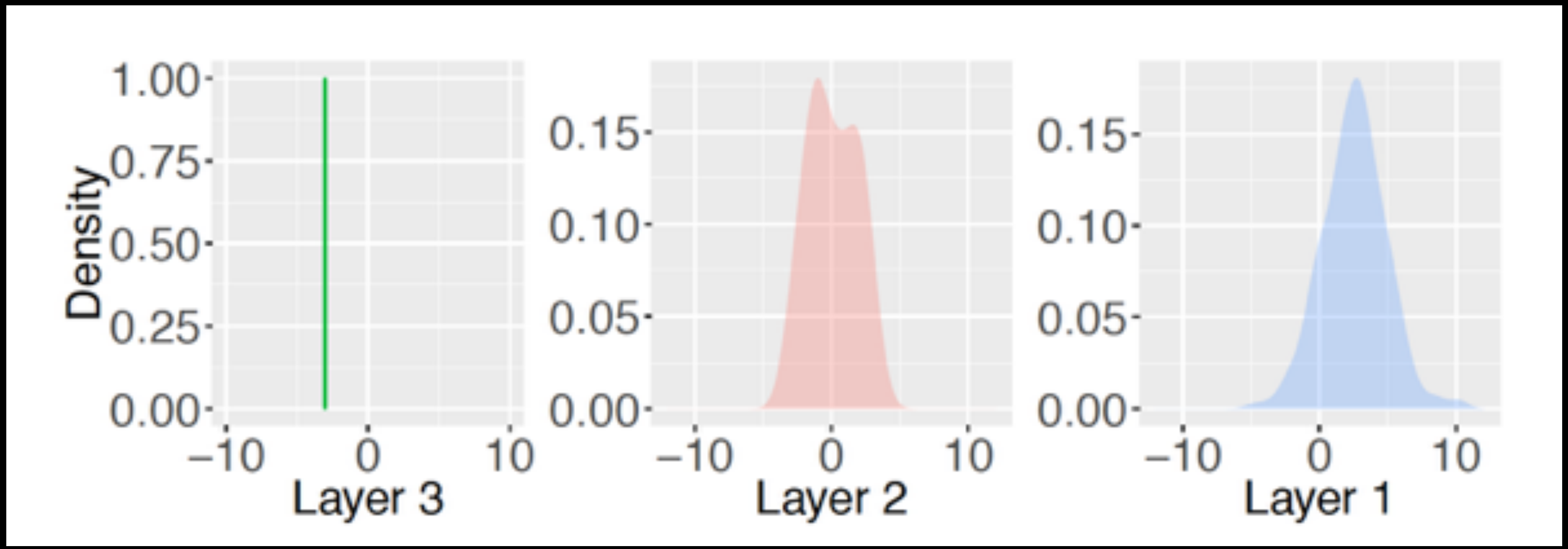
- Feature relationships

$$\mathrm{FR}(f, g) = \frac{1}{|f| \times |g|} \sum_{i \in f} \sum_{j \in g} \mathrm{IC}(i, j)$$

  FR(f, g)<1: alternative
  FR(f, g)>1: complementary

Empirical study

Distribution of feature relationships (*log(FR)*) in 3 layers (Amazon Clothing dataset).



Features have alternative relationships at Layer 3 (men vs. women); Feature relationships are evenly distributed at Layer 2; most features are complementary at Layer 1.

HieVH Framework

- The basic recommendation model

$$\mathcal{J} = \overbrace{\sum_{o_{ui}\neq 0} C\big(r_{ui}, <\theta_u, \overline{\theta}_i>\big)}^{\text{cost function}} + \overbrace{\alpha \sum_{f,g\in\mathcal{F}} \Psi(\theta_f, \theta_g) + \Omega(\Theta)}^{\text{regularizers}}$$

HieVH Framework

- Modeling Vertical Dimension: the latent factor of item is adapted by adding to it the latent factors of its affiliated features in the hierarchy

$$\overline{\theta}_i = \Phi(\theta_i, \theta_f, \vartheta_f) = \theta_i + \left[\vartheta_{f^1}, \vartheta_{f^2}, \cdots, \vartheta_{f^L}\right] \begin{bmatrix} -\theta_{f^1}- \\ -\theta_{f^2}- \\ \cdots \\ -\theta_{f^L}- \end{bmatrix}_{L\times d}$$

HieVH Framework

- Modeling Horizontal Dimension: if two features are alternative, the distance of their latent factors should be large; if complementary, the distance should be small

$$\Psi(\theta_f, \theta_g) = \sum_{l=1}^{L} \sum_{f,g \in \mathcal{F}^l, f<g} \sigma_{fg} \|\theta_f - \theta_g\|_F^2$$

Where $\sigma_{fg} = \log(FR(f,g))$

Validation

- Datasets: Amazon Web store (Clothing, Electronics, CDs and Home)

- Comparison Methods: MF[1], CMF[2], FM[3], TaxMF[4], Sherlock[5], ReMF[6]

  - [1] Salakhutdinov, R., and Hanjalic, A. 2007. Probabilistic matrix factorization.In NIPS.
  - [2] Singh, A.P., and Gordon, G. J. 2008. Relational learning via collective matric factorization. In KDD.
  - [3] Rendle, S. 2010. Factorization machine. In ICDM.
  - [4] Koenigstein, N.; Dror, G.; and Koren, Y. 2011. Yahoo! Music recommendations: modeling music ratings with temporal dynamics and item taxonomy. In RecSys.
  - [5] He, R.; Lin, C.; Wang, J.; and McAuley, J. 2016. Sherlock: sparse hierarchical embeddings for visually-aware one-class collaborative filtering. In IJCAI.
  - [6] Yang, J.; Sun Z.; Bozzon, A.; and Zhang, J.2016. Learning hierarchical feature influence for recommendation by recursive regularization. In RecSys.

Validation

Performance of comparative methods measured by AUC.

Datasets	Cases	MF	CMF	FM	TaxMF	Sherlock	ReMF	HieV	HieVC	HieVH	Improve
Clothing.	All Users	0.5455	0.5646	0.6826	0.6509	0.6747	0.7015*	0.7160	0.7291	**0.7375**	5.13%
	Cold Start	0.5426	0.5667	0.6629	0.6493	0.6702	0.7032*	0.7124	0.7284	**0.7352**	4.55%
Electronic.	All Users	0.5555	0.5762	0.6839	0.6569	0.6915	0.7337*	0.7512	0.7672	**0.7748**	5.60%
	Cold Start	0.5526	0.5735	0.6831	0.6475	0.6982	0.7305*	0.7474	0.7658	**0.7741**	5.97%
C.& V.	All Users	0.5478	0.5622	0.6356	0.6905	0.7082	0.7249*	0.7328	0.7516	**0.7600**	4.84%
	Cold Start	0.5433	0.5609	0.6231	0.6881	0.7076	0.7243*	0.7315	0.7514	**0.7588**	4.76%
H. & K.	All Users	0.5420	0.5545	0.6938	0.6469	0.6938	0.7279*	0.7456	0.7574	**0.7667**	5.33%
	Cold Start	0.5395	0.5562	0.6915	0.6511	0.6973	0.7275*	0.7412	0.7554	**0.7650**	5.15%

The incorporation of horizontal dimension enhances recommendation performance.

Validation

Percentage of complementarity *Cp* and alternativity *Ap* between rated and recommended items

Method	Layer 1		Layer 2	
	Cp	Ap	Cp	Ap
ReMF	87.62%	12.38%	75.23%	24.76%
ReV	88.89%	11.11%	76.19%	23.89%
ReVC	92.62%	7.38%	84.62%	15.38%
ReVH	95.65%	4.35%	89.35%	10.65%

By incorporating the two types of feature relationships, HieVH recommends to users more items complementing rated to items.

Thanks!

https://yangjiera.github.io/citrec2017/

CitRec2017

- Recommender Systems run ***by*** citizens, and ***for*** citizens.
- Benefit the society ***as a whole***: many challenges
  - Spatio-temporal context
  - Cross-domain
  - Conflicting interests
  - …