Multi-Modal Recommender Systems:
Towards Addressing Sparsity, Comparability, and Explainability


Trung-Hoang Le
SMU Singapore Management University | School of Computing and Information Systems


Quoc-Tuan Truong
amazon


Aghiles Salah
Rakuten Institute of Technology


Hady W. Lauw
SMU Singapore Management University | School of Computing and Information Systems

PREFERRED.AI

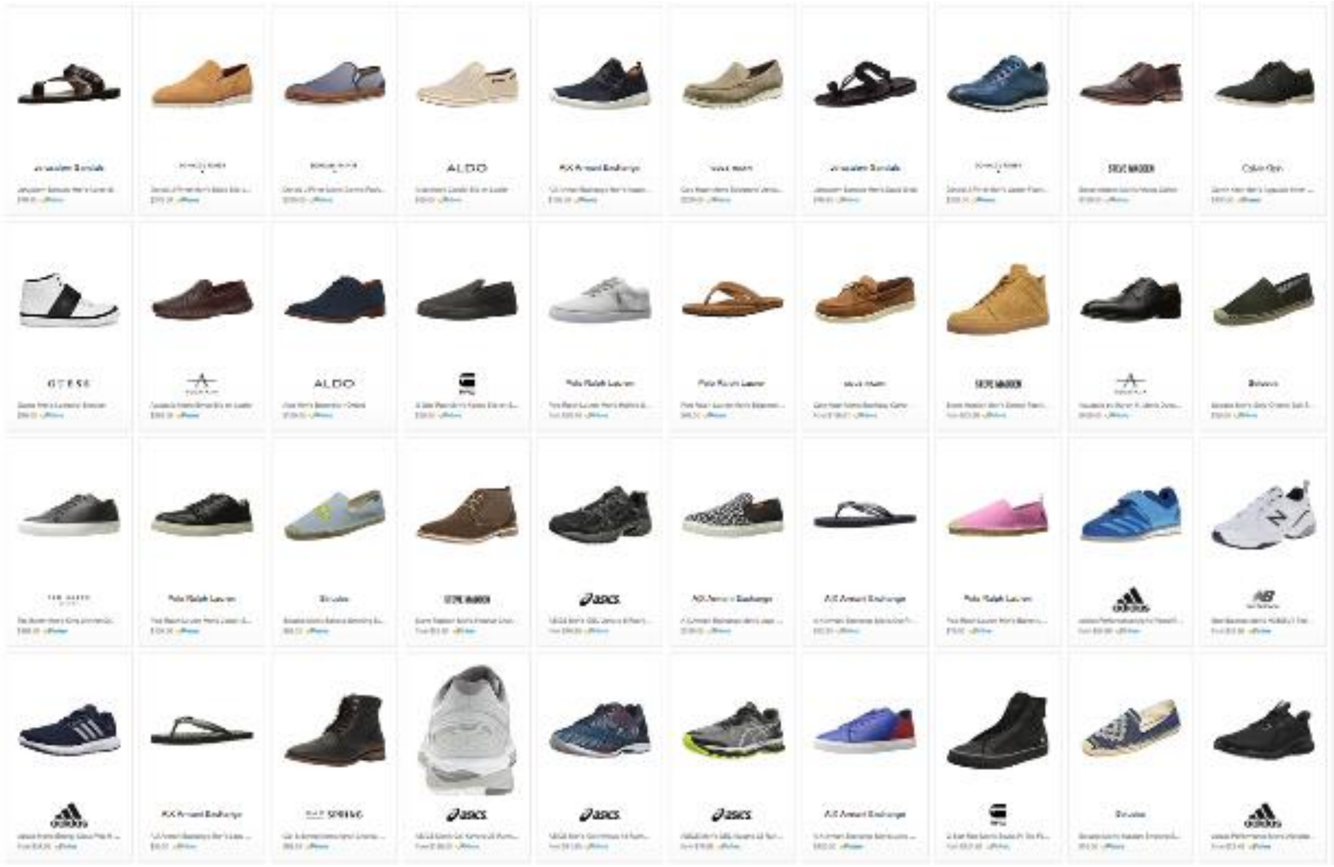
Outline (180 min)

- Multimodal Recommender Systems (15 mins)
- Hands-on 1: Cornac framework (15 mins)
- Text and Image Modalities (30 mins)
- Graph Modality (30 mins)
- Break
- Hands-on 2: Multi-modal & Cross-modal (30 mins)
- Explainability (30 mins)
- Hands-on 3: Explainability (20 mins)
- Q & A (10 mins)

PREFERRED.AI

Choice Explosion

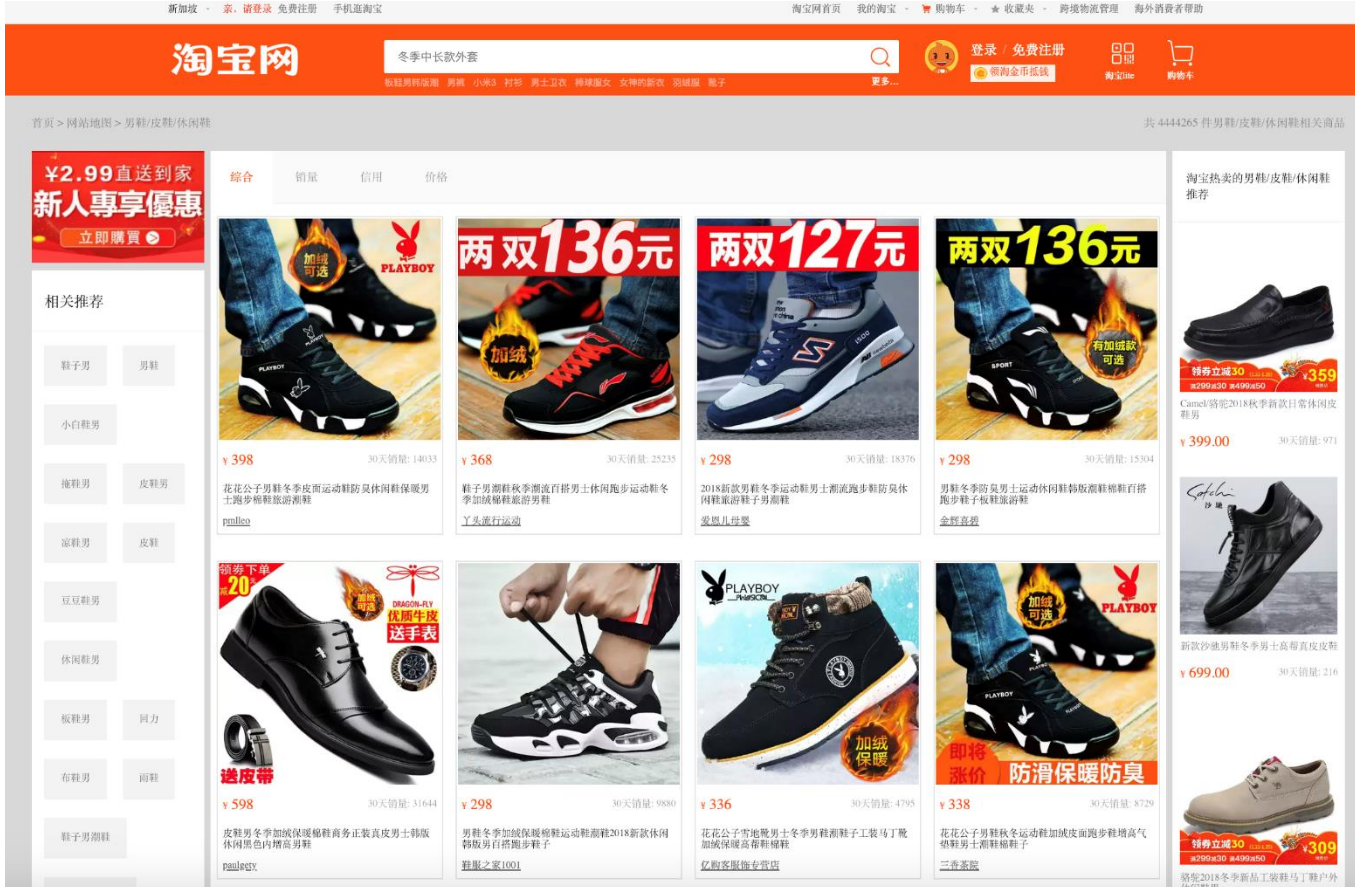
133,520 results for "Men’s Shoes"



amazon

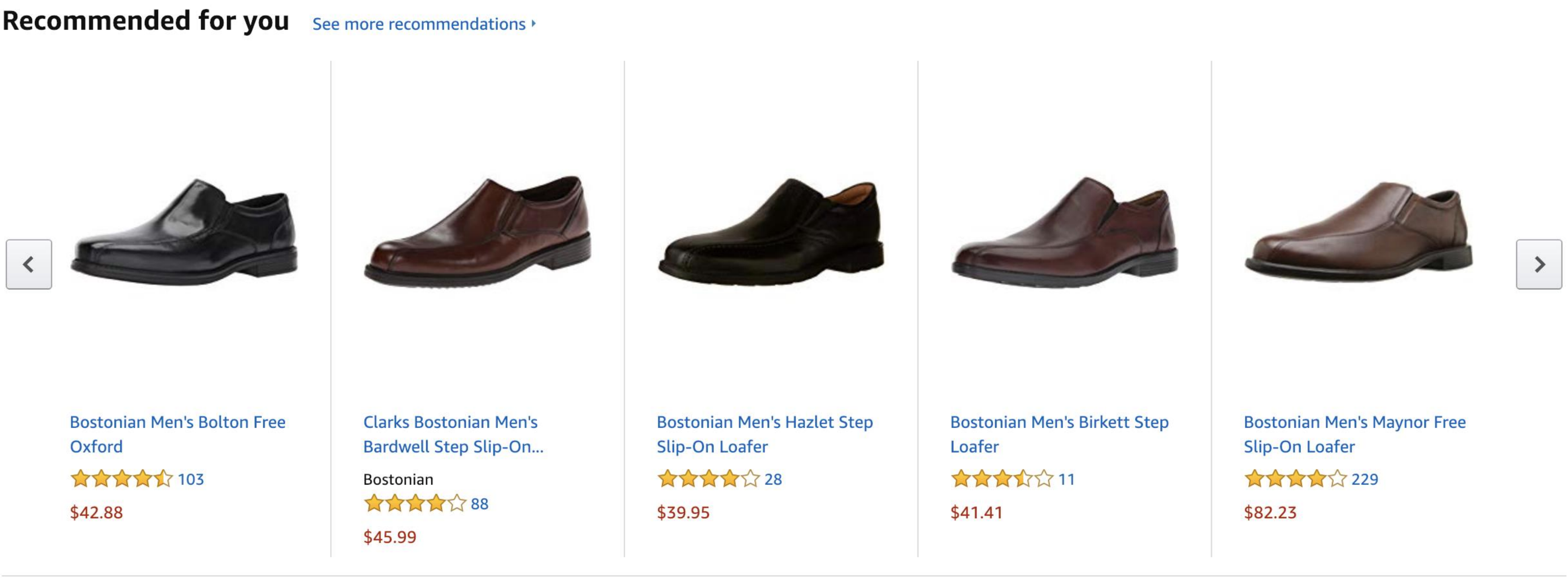
Choice Explosion

4 million results results for ”男鞋/皮鞋/休闲鞋"



淘宝网
Taobao.com

Solution: Recommender Systems



Where do we see a recommender system?

Formulation

Rating Prediction

- Given a rating dataset $\boldsymbol{R}$
  - each $r_{ui} \in \boldsymbol{R}$ indicates a rating, or the degree of preference user u has for item i
  - where $r_{ui} \notin \boldsymbol{R}$, predict rating $\hat{r}_{ui}$ as a function $f(u, i)$
- "Is the user going to like this item?"

Ranking

- Given a ranking dataset $\boldsymbol{T}$
  - each triple $t_{uij} \in \boldsymbol{T}$ indicates whether/how much u prefers i to j
  - where $t_{uij} \notin \boldsymbol{T}$, predict rating $\hat{t}_{uij}$ as a function $g(u, i, j)$
- "Which items are the user most likely to have an interest?"

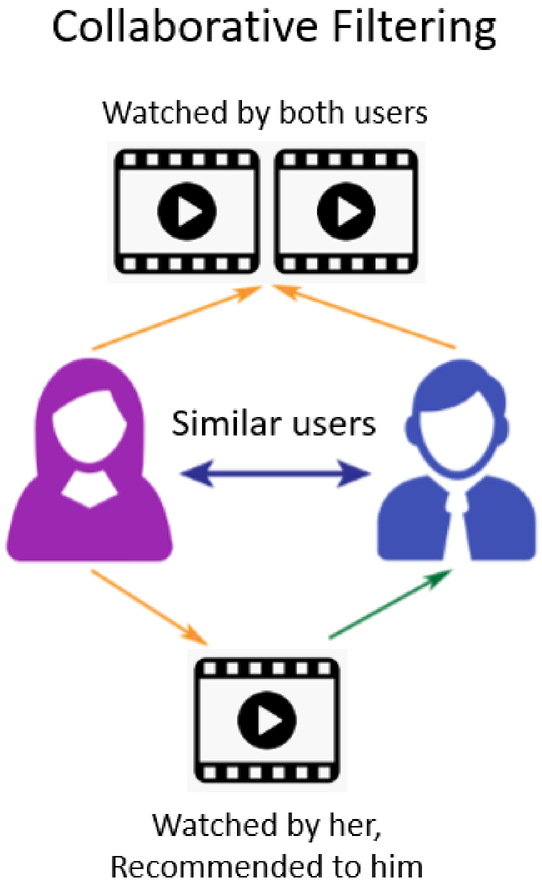
Basis for User-Item Relations

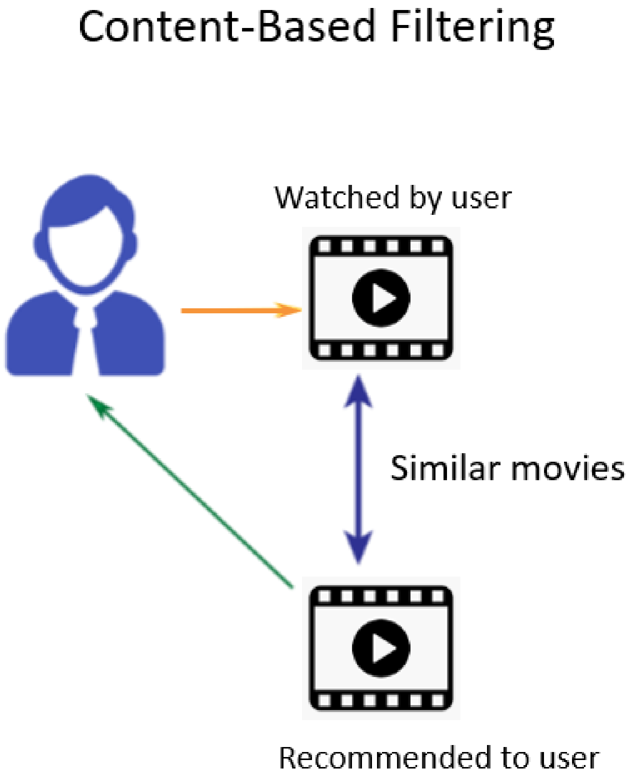
Collaborative Filtering

- A user tends to have similar consumption behavior to other 'like-minded' users

Content-Based Filtering

- A user tends to like items with similar contents to those previously consumed





https://www.kdnuggets.com/2019/11/content-based-recommender-using-natural-language-processing-nlp.html

Matrix Factorization

Koren, Bell, and Volinsky, "Matrix Factorization Techniques for Recommender Systems", IEEE Computer, 2009

Our winning entries consist of more than 100 different predictor sets, the majority of which are factorization models using some variants of the methods described here. Our discussions with other top teams and postings on the public contest forum indicate that these are the most popular and successful methods for predicting ratings.

~BellKor

https://datajobs.com/data-science-repo/Recommender-Systems-[Netflix].pdf

Example Rank-2 Matrix Factorization

R

		NERO	JULIUS CAESAR	CLEOPATRA	SLEEPLESS IN SEATTLE	PRETTY WOMAN	CASABLANCA
HISTORY	1	1	1	1	0	0	0
HISTORY	2	1	1	1	0	0	0
HISTORY	3	1	1	1	0	0	0
BOTH	4	1	1	1	1	1	1
ROMANCE	5	-1	-1	-1	1	1	1
ROMANCE	6	-1	-1	1	1	1	1
ROMANCE	7	-1	-1	-1	1	1	1

≈

U

	HISTORY	ROMANCE
1	1	0
2	1	0
3	1	0
4	1	1
5	-1	1
6	-1	1
7	-1	1

X

V^T

	NERO	JULIUS CAESAR	CLEOPATRA	SLEEPLESS IN SEATTLE	PRETTY WOMAN	CASABLANCA
HISTORY	1	1	1	0	0	0
ROMANCE	0	0	1	1	1	1

R

		NERO	JULIUS CAESAR	CLEOPATRA	SLEEPLESS IN SEATTLE	PRETTY WOMAN	CASABLANCA
HISTORY	1	0	0	0	0	0	0
HISTORY	2	0	0	0	0	0	0
HISTORY	3	0	0	0	0	0	0
BOTH	4	0	0	-1	0	0	0
ROMANCE	5	0	0	-1	0	0	0
ROMANCE	6	0	0	1	0	0	0
ROMANCE	7	0	0	-1	0	0	0

Residual matrix: $\boldsymbol{R} - \boldsymbol{U}\boldsymbol{V}^T$

Estimating Latent Factors

- Minimize loss function:

$$\mathcal{L}(\boldsymbol{U},\boldsymbol{V}) = \frac{1}{2}\|\boldsymbol{R} - \boldsymbol{U}\boldsymbol{V}^T\|^2 = \frac{1}{2}\sum_{r_{ij}\in\boldsymbol{R}}\left(r_{ij} - \sum_{k=1}^{K} u_{ik}\cdot v_{jk}\right)^2$$

  - defined only over observed ratings

Regularization

- Due to sparsity, some users or items may have very few ratings
- To prevent overfitting, we can introduce regularization:

$$\mathcal{L}(\boldsymbol{U},\boldsymbol{V}|\lambda) = \frac{1}{2}\sum_{r_{ij}\in\boldsymbol{R}}\left(r_{ij} - \sum_{k=1}^{K} u_{ik}\cdot v_{jk}\right)^2 + \frac{\lambda}{2}\sum_{i=1}^{N}\|\boldsymbol{u}_i\|^2 + \frac{\lambda}{2}\sum_{j=1}^{M}\|\boldsymbol{v}_j\|^2$$

$$= \frac{1}{2}\sum_{r_{ij}\in\boldsymbol{R}}\left(r_{ij} - \sum_{k=1}^{K} u_{ik}\cdot v_{jk}\right)^2 + \frac{\lambda}{2}\sum_{i=1}^{N}\sum_{k=1}^{K}(u_{ik})^2 + \frac{\lambda}{2}\sum_{j=1}^{M}\sum_{k=1}^{K}(v_{jk})^2$$

User and Item Biases

- Users have different ranges of ratings
  - some are generous with their ratings mainly in the upper range, others are strict
- Items differ in popularity or likeability
  - some have mainly high ratings, others mainly low
- Introduce bias parameters:

$$\hat{r}_{ij} = \mu + b_{u_i} + b_{v_j} + \boldsymbol{u}_i^T \boldsymbol{v}_j$$

  - μ is the global average in the training data, to be computed
  - b_{u_i} is a bias term for a specific user, to be learnt
  - b_{v_j} is a bias term for a specific item, to be learn
- Loss function:

$$\mathcal{L}(\boldsymbol{U}, \boldsymbol{V}|\lambda) = \frac{1}{2} \sum_{r_{ij} \in \boldsymbol{R}} \left(r_{ij} - (\mu + b_{u_i} + b_{v_j} + \boldsymbol{u}_i^T \boldsymbol{v}_j)\right)^2 + \frac{\lambda}{2} \sum_{i=1}^{N} \left(\|\boldsymbol{u}_i\|^2 + \left(b_{u_i}\right)^2\right) + \frac{\lambda}{2} \sum_{j=1}^{M} \left(\|\boldsymbol{v}_j\|^2 + \left(b_{v_j}\right)^2\right)$$

Explicit vs. Implicit Feedback

Explicit

- Stated clearly and readily observable
- Ratings, thumbs up/down
- Some notion of positive/negative

Implicit

- Preferences are unclear
  - suggested, not directly expressed
- Noisy
  - clicks, viewing time
  - observed values may have very small or wide ranges
- Absence of negative signals
  - only positive examples observed

Weighted Matrix Factorization (WMF)

Hu, Koren, and Volinsky, " Collaborative Filtering for Implicit Feedback Datasets", ICDM 2009

- Adoption $\hat{p}_{ij} = \boldsymbol{u}_i^T \boldsymbol{v}_j$ by user i on item j
  - $p_{\text{ij}} = \begin{cases} 1, \text{if } r_{ij} \geq \text{threshold} \\ 0, \ \text{otherwise} \end{cases}$
- Confidence
  - $c_{\text{ij}} = \begin{cases} a, \text{if } r_{ij} \geq \text{threshold} \\ b, \ \text{otherwise} \end{cases}$

  Customarily $b \ll a$ as we have less confidence on non-observations
- Loss function

$$\mathcal{L}(\boldsymbol{U},\boldsymbol{V}|\lambda) = \frac{1}{2}\sum_{i,j} c_{ij}\left(p_{ij} - \boldsymbol{u}_i^T\boldsymbol{v}_j\right)^2 + \frac{\lambda}{2}\sum_{i=1}^{N}\|\boldsymbol{u}_i\|^2 + \frac{\lambda}{2}\sum_{j=1}^{M}\left\|\boldsymbol{v}_j\right\|^2$$

PREFERRED.AI

Bayesian Personalized Ranking (BPR)

Rendle, Freudenthaler, Gantner and Schmidt-Thieme, "BPR: Bayesian Personalized Ranking from Implicit Feedback", UAI 2009

- Ordinal triple

$$j >_i l \begin{cases} 1, \text{if } r_{ij} \in \boldsymbol{R}^+ \wedge r_{il} \in \boldsymbol{R}^- \\ 0, \text{if } r_{ij} \in \boldsymbol{R}^- \wedge r_{il} \in \boldsymbol{R}^+ \\ \text{otherwise unspecified} \end{cases}$$

- Triple probability as sigmoid function

$$P(j >_i l) = \text{sigmoid}(\boldsymbol{u}_i^T \boldsymbol{v}_j - \boldsymbol{u}_i^T \boldsymbol{v}_l)$$

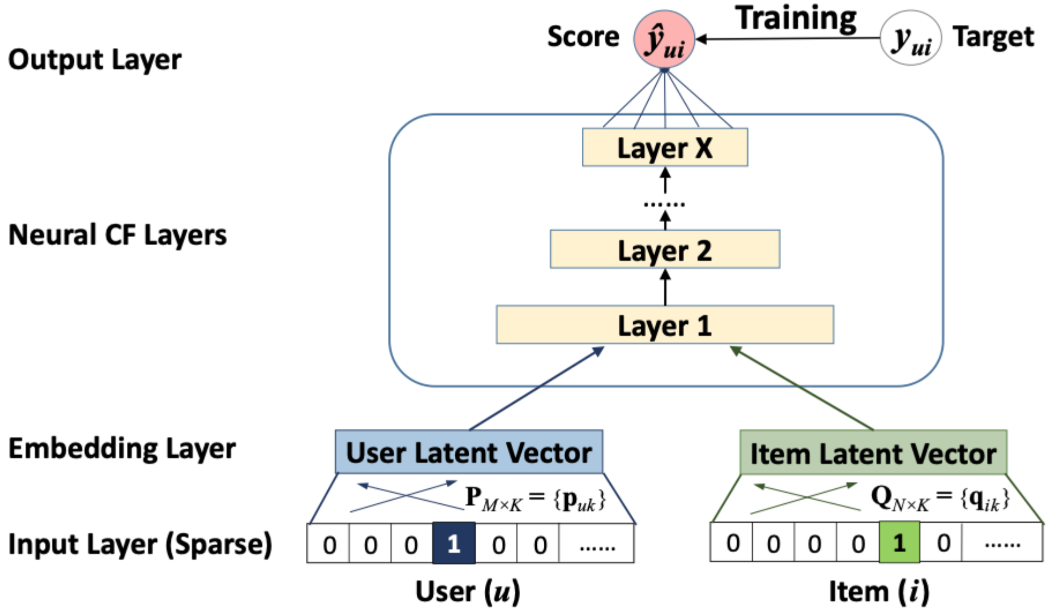
- Maximize the joint probability across all triples $\prod_{(j >_i l) \in \boldsymbol{S}} P(j >_i l)$

- With regularization, minimize regularized negative log-likelihood function:

$$\mathcal{L}(\boldsymbol{U}, \boldsymbol{V} | \lambda) = \sum_{(j >_i l) \in \boldsymbol{S}} \ln(1 + \exp\{-(\boldsymbol{u}_i^T \boldsymbol{v}_j - \boldsymbol{u}_i^T \boldsymbol{v}_l)\}) + \frac{\lambda}{2} \sum_{i=1}^{N} \|\boldsymbol{u}_i\|^2 + \frac{\lambda}{2} \sum_{j=1}^{M} \|\boldsymbol{v}_j\|^2$$

Neural Collaborative Filtering

He et al., "Neural collaborative filtering", WWW 2017

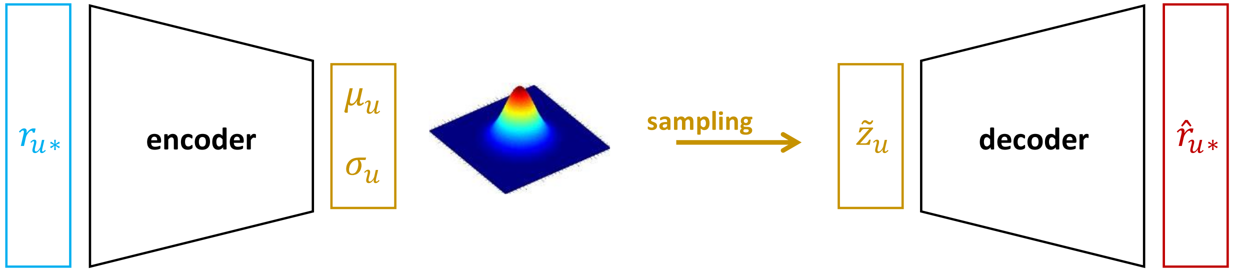


- NeuMF unifies the strengths of MF and MLP in modeling user preference
  - MF captures interaction via inner product (simple yet effective)
  - MLP is more capable of capturing complex user intention

(Bilateral) Variational Auto-Encoder

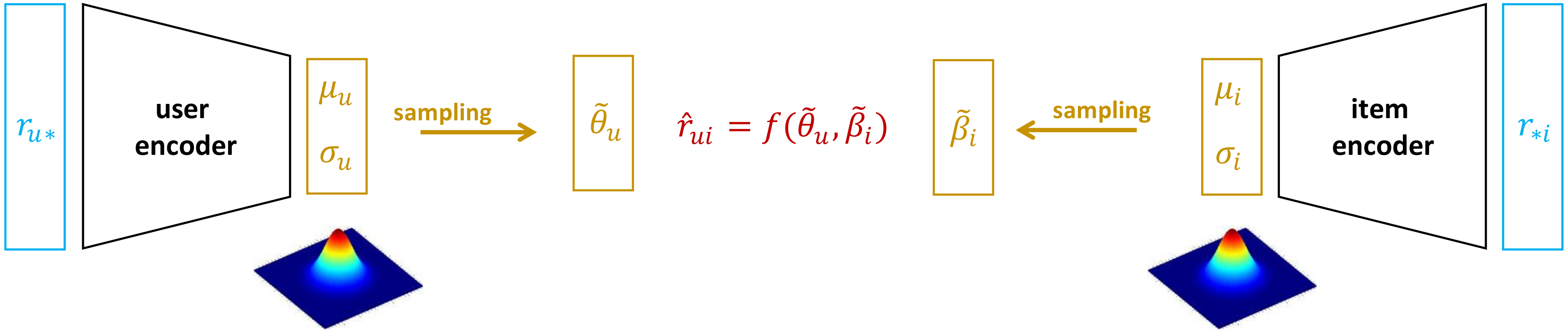
Liang, Krishnan, Hoffman, and Jebara, "Variational autoencoders for collaborative filtering", WWW 2018.

VAE



Truong, Salah, and Lauw, "Bilateral Variational Autoencoder for Collaborative Filtering", WSDM, 2021.

BiVAE

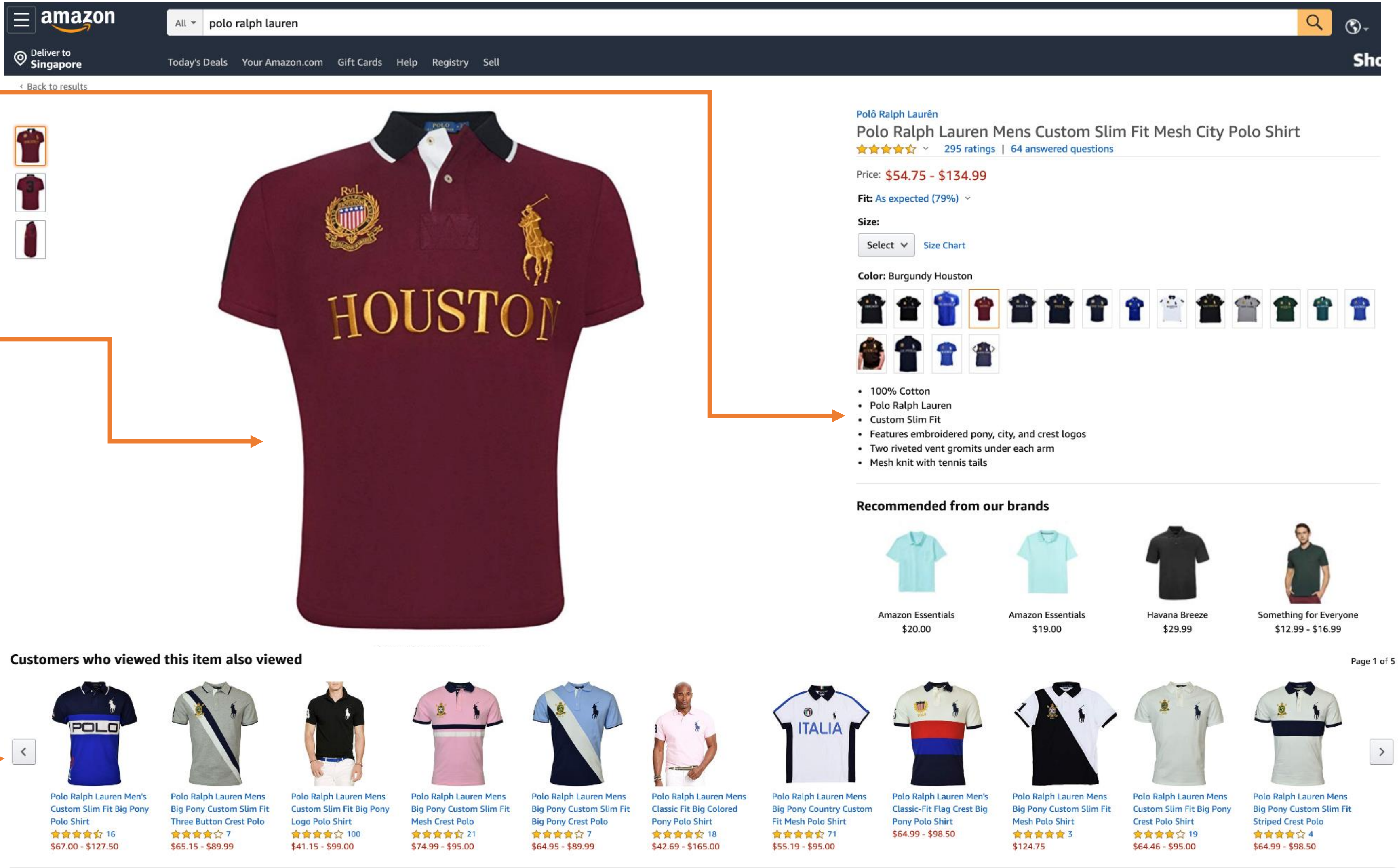


- Respect the two-way nature of dyadic data
- Can accommodate auxiliary data from both sides

Product description

Product image

Related products



Collaborative vs. Content-Based Filtering

Collaborative Filtering

- Presumes that behavior drives consumption
- Greater capacity for personalization, especially for matrix factorization with user latent vectors

Content-Based Filtering

- Presumes that content drives consumption
- Caters to cold-start items

Can we do both?

Yes, since matrix factorization or deep learning algorithms are highly extensible

Multimodality

- Key idea: learn from multiple modalities at the same time

- First modality: preference feedback
  - Explicit feedback
  - Implicit feedback

- Second (and third, etc.) modality
  - Item content in the form of text, graph, or image
  - Alternatively, user content

Hands-on #1: Cornac

https://github.com/PreferredAI/tutorials/blob/master/multimodal-recsys/01_getting_started.ipynb

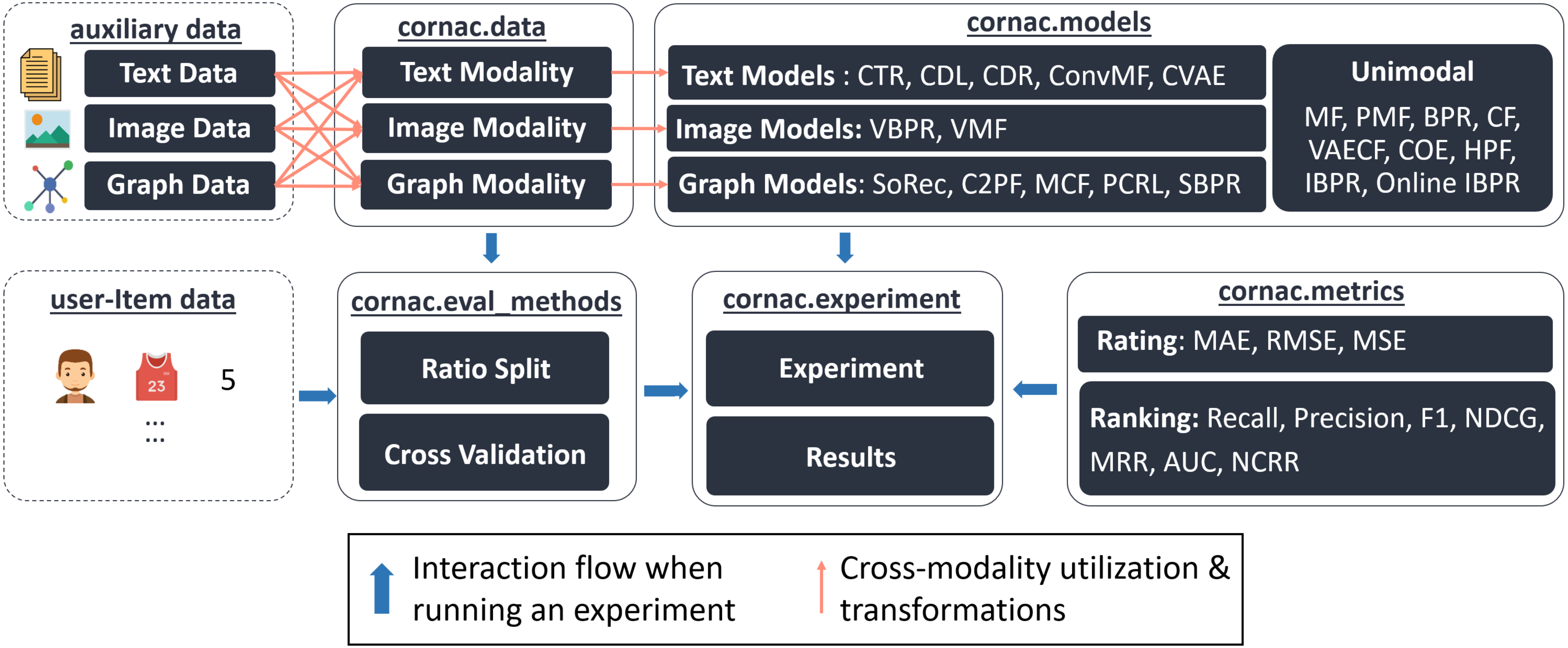
Cornac

- Python-based recommender framework

- An open-source project by Preferred.AI
  - https://cornac.preferred.ai
  - Go to our GitHub and star it

- Publication at Journal of Machine Learning Research:
  - https://jmlr.org/papers/v21/19-805.html

Overview



Fast **experimentation**, **exploration**, and **comparisons**

Cornac Key Features

- **Multimodality Support**
  - Reading, transforming, formatting and representing different types of data
  - Convenient development of new models
  - Broadening the use cases of existing models
- **Scalability**
  - A collection of Iterators for easy stochastic optimization
  - Harnessing the Python ecosystem, e.g., NumPy and Scipy, for efficient operations
  - Leverage Cython to achieve C/C++ performance
- **Accessibility & Reproducibility**
  - Open-access to a rich collection of **models** (**>40**)
  - Straightforward usage of real-world benchmark datasets
  - Full control over random number generators

Experiment oriented

Text Modality

Text-Modality

Text Modeling	MF or PMF or FM	WMF	BPR	NCF	Auto-Encoder
Term Vector	SVDFeature				
Matrix Factorization	CMF				
Topic Model	HFT	CTR	CTRank		
Auto-Encoder	AutoSVD	CDL, CVAE	CDR		AddVAE, MD-CVAE
CNN	ConvMF, DeepCoNN	DeepMusic			
RNN, LSTM				MRG, NRT	

References

- Wang, C., & Blei, D. M. (2011, August). Collaborative topic modeling for recommending scientific articles. In Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 448-456).
- Wang, H., Wang, N., & Yeung, D. Y. (2015, August). Collaborative deep learning for recommender systems. In Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining (pp. 1235-1244).
- Kim, D., Park, C., Oh, J., Lee, S., & Yu, H. (2016, September). Convolutional matrix factorization for document context-aware recommendation. In Proceedings of the 10th ACM conference on recommender systems (pp. 233-240).
- Ying, H., Chen, L., Xiong, Y., & Wu, J. (2016, April). Collaborative deep ranking: A hybrid pair-wise recommendation algorithm with implicit feedback. In Pacific-asia conference on knowledge discovery and data mining (pp. 555-567). Springer, Cham.
- Li, X., & She, J. (2017, August). Collaborative variational autoencoder for recommender systems. In Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 305-314).
- Yao, W., He, J., Wang, H., Zhang, Y., & Cao, J. (2015, February). Collaborative topic ranking: Leveraging item meta-data for sparsity reduction. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 29, No. 1).
- Van Den Oord, A., Dieleman, S., & Schrauwen, B. (2013). Deep content-based music recommendation. In Neural Information Processing Systems Conference (NIPS 2013) (Vol. 26). Neural Information Processing Systems Foundation (NIPS).
- McAuley, J., & Leskovec, J. (2013, October). Hidden factors and hidden topics: understanding rating dimensions with review text. In Proceedings of the 7th ACM conference on Recommender systems (pp. 165-172).
- Singh, A. P., & Gordon, G. J. (2008, August). Relational learning via collective matrix factorization. In Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 650-658).
- Chen, T., Zhang, W., Lu, Q., Chen, K., Zheng, Z., & Yu, Y. (2012). SVDFeature: a toolkit for feature-based collaborative filtering. The Journal of Machine Learning Research, 13(1), 3619-3622.
- Zhang, S., Yao, L., & Xu, X. (2017, August). Autosvd++ an efficient hybrid collaborative filtering model via contractive auto-encoders. In Proceedings of the 40th International ACM SIGIR conference on Research and Development in Information Retrieval (pp. 957-960).
- Zheng, L., Noroozi, V., & Yu, P. S. (2017, February). Joint deep modeling of users and items using reviews for recommendation. In Proceedings of the tenth ACM international conference on web search and data mining (pp. 425-434).
- Li, P., Wang, Z., Ren, Z., Bing, L., & Lam, W. (2017, August). Neural rating regression with abstractive tips generation for recommendation. In Proceedings of the 40th International ACM SIGIR conference on Research and Development in Information Retrieval (pp. 345-354).
- Zhu, Yaochen, and Zhenzhong Chen. "Mutually-regularized dual collaborative variational auto-encoder for recommendation systems.", WWW 2022.
- Tran, Nhu-Thuat, and Hady W. Lauw. "Aligning Dual Disentangled User Representations from Ratings and Textual Content.", KDD 2022.

SVDFeature

Chen, Zhang, Lu, Chen, Zheng, & Yu, "SVDFeature: a toolkit for feature-based collaborative filtering", JMLR, 2012.

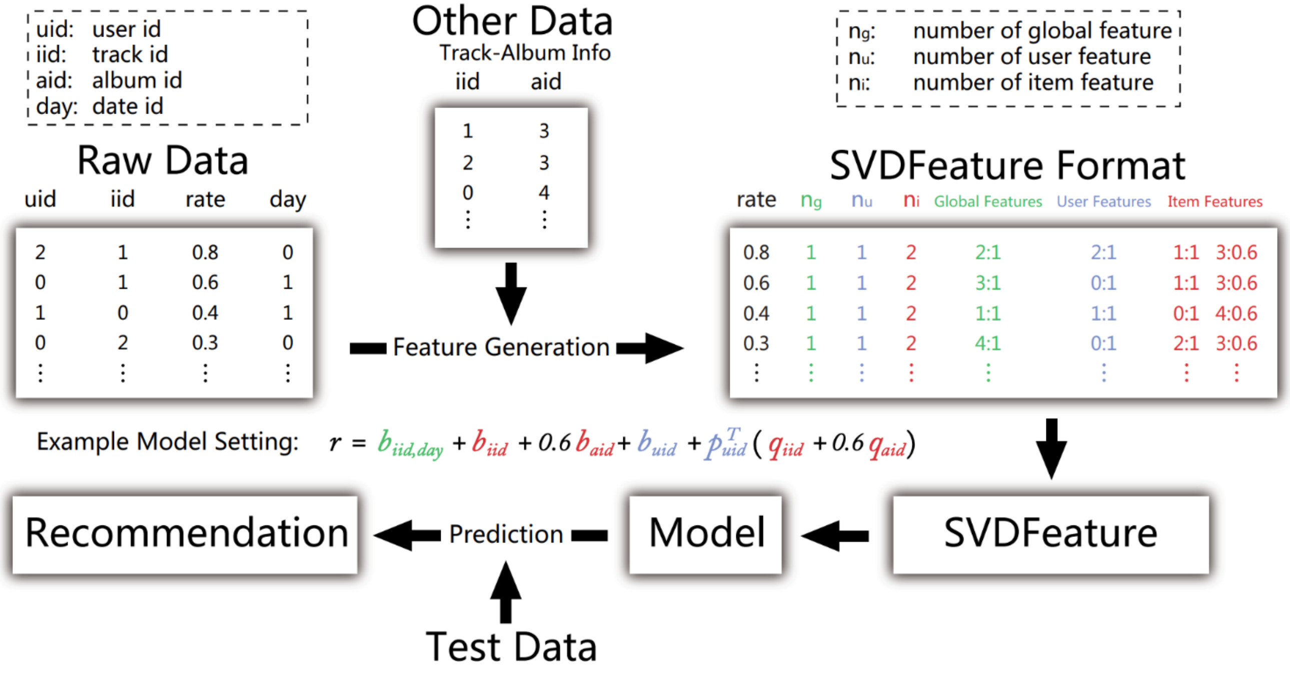


Figure 1: Usage flow example of SVDFeature

Collective Matrix Factorization

Singh and Gordon, "Relational learning via collective matrix factorization", KDD, 2008.

- Let $\boldsymbol{R} \in \mathbb{R}^{N \times M}$ be a sparse rating matrix for N users and M items
  - $r_{ij} \in \boldsymbol{R}$ is an observed rating by user i on item j
  - For each user i, a real-valued vector $\boldsymbol{u}_i \in \mathbb{R}^K$. For all users, collectively $\boldsymbol{U} \in \mathbb{R}^{N \times K}$.
  - For each item j, a real-valued vector $\boldsymbol{v}_j \in \mathbb{R}^K$. For all items, collectively $\boldsymbol{V} \in \mathbb{R}^{M \times K}$.
- Let $\boldsymbol{D} \in \mathbb{R}^{M \times L}$ be a sparse matrix for M items and L words
  - $d_{jl} \in \boldsymbol{D}$ is an observed occurrence/importance of word l in item j
  - For each word l, a real-valued vector $\boldsymbol{z}_l \in \mathbb{R}^K$. For all words, collectively $\boldsymbol{Z} \in \mathbb{R}^{L \times K}$.
- Overall loss function (with regularization)

$$\mathcal{L}(\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{Z} | \lambda) = \frac{1}{2} \sum_{r_{ij} \in \boldsymbol{R}} \left(r_{ij} - \sum_{k=1}^{K} u_{ik} \cdot v_{jk} \right)^2 + \frac{1}{2} \sum_{d_{jl} \in \boldsymbol{D}} \left(d_{jl} - \sum_{k=1}^{K} v_{jk} \cdot z_{lk} \right)^2 + \frac{\lambda}{2} \sum_{i=1}^{N} \|\boldsymbol{u}_i\|^2 + \frac{\lambda}{2} \sum_{j=1}^{M} \|\boldsymbol{v}_j\|^2 + \frac{\lambda}{2} \sum_{l=1}^{L} \|\boldsymbol{z}_l\|^2$$

  - and λ is the regularization weight

Topic Model

Blei, Ng, and Jordan, "Latent Dirichlet Allocation", JMLR 2003.

- Text Document

The William Randolph Hearst Foundation will give $1.25 million to Lincoln Center, Metropolitan Opera Co., New York Philharmonic and Juilliard School. "Our board felt that we had a real opportunity to make a mark on the future of the performing arts with these grants an act every bit as important as our traditional areas of support in health, medical research, education and the social services," Hearst Foundation President Randolph A. Hearst said Monday in announcing the grants. Lincoln Center's share will be $200,000 for its new building, which will house young artists and provide new public facilities. The Metropolitan Opera Co. and New York Philharmonic will receive $400,000 each. The Juilliard School, where music and the performing arts are taught, will get $250,000. The Hearst Foundation, a leading supporter of the Lincoln Center Consolidated Corporate Fund, will make its usual annual $100,000 donation, too.

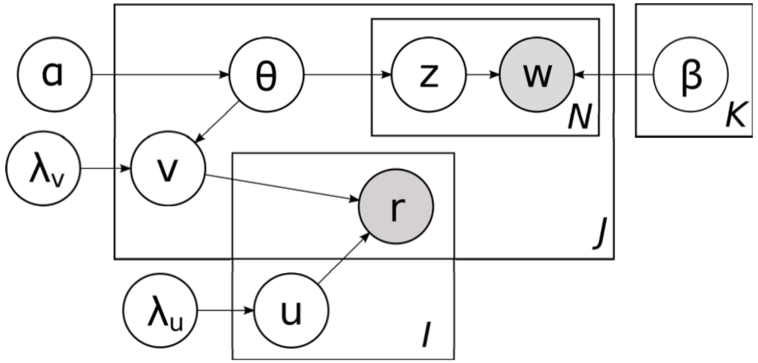
- Topics

"Arts"	"Budgets"	"Children"	"Education"
NEW	MILLION	CHILDREN	SCHOOL
FILM	TAX	WOMEN	STUDENTS
SHOW	PROGRAM	PEOPLE	SCHOOLS
MUSIC	BUDGET	CHILD	EDUCATION
MOVIE	BILLION	YEARS	TEACHERS
PLAY	FEDERAL	FAMILIES	HIGH
MUSICAL	YEAR	WORK	PUBLIC
BEST	SPENDING	PARENTS	TEACHER
ACTOR	NEW	SAYS	BENNETT
FIRST	STATE	FAMILY	MANIGAT
YORK	PLAN	WELFARE	NAMPHY
OPERA	MONEY	MEN	STATE
THEATER	PROGRAMS	PERCENT	PRESIDENT
ACTRESS	GOVERNMENT	CARE	ELEMENTARY
LOVE	CONGRESS	LIFE	HAITI

- A document of many words may discuss a relatively small number of "topics"
- A topic can be described by a series of words that frequently co-occur with one another

Collaborative Topic Regression (CTR)

Wang and Blei, "Collaborative Topic Modeling for Recommending Scientific Articles", KDD 2011.



- Combines weighted matrix factorization and topic model (LDA)
- Item j has a document with a distribution $\boldsymbol{\theta}_j$ over K topics
- Item j also has K-dimensional latent vector $\boldsymbol{v}_j$

$$\hat{p}_{ij} = \boldsymbol{u}_i^T \boldsymbol{v}_j$$

- Intuition: two items with similar topics would behave similarly
- Assume that $\boldsymbol{v}_j$ is drawn from a Normal distribution, with mean $\boldsymbol{\theta}_j$

$$\boldsymbol{v}_j \sim \mathcal{N}\left(\boldsymbol{\theta}_j, \lambda^{-1}\mathbf{I}\right)$$

- Equivalently:

$$\boldsymbol{v}_j = \boldsymbol{\theta}_j + \boldsymbol{\epsilon}_j$$
$$\boldsymbol{\epsilon}_j \sim \mathcal{N}(\mathbf{0}, \lambda^{-1}\mathbf{I})$$

PREFERRED.AI

CTR: Generative Process

- Generative process:
  - For each user i
    - Draw latent vector $\boldsymbol{u}_i \sim \mathcal{N}(\mathbf{0}, \lambda^{-1}\mathbf{I})$
  - For each item j
    - Draw topic proportions $\boldsymbol{\theta}_j \sim \mathrm{Dirichlet}(\alpha)$
    - Draw item offset $\boldsymbol{\epsilon}_j \sim \mathcal{N}(\mathbf{0}, \lambda^{-1}\mathbf{I})$
    - For the n^{th} word in the article:
      - Draw topic assignment $z_{jn} \sim \mathrm{Multinomial}(\boldsymbol{\theta}_j)$
      - Draw word $w_{jn} \sim \mathrm{Multinomial}(\beta_{jn})$
  - For each user-item pair (i, j)
    - Draw adoption $p_{ij} \sim \mathcal{N}(\boldsymbol{u}_i^T \boldsymbol{v}_j, c_{ij}^{-1})$

- Adoption:
  - Binary adoption or
  - Valued adoption

$$p_{ij} = \begin{cases} r_{ij}, \text{if } r_{ij} \geq \text{threshold} \\ 0, \text{otherwise} \end{cases}$$

- Confidence:

$$c_{ij} = \begin{cases} a, \text{if } r_{ij} \geq \text{threshold} \\ b, \text{otherwise} \end{cases}$$

  - More confidence on positive adoption, i.e., $a \gg b$

CTR Learning

- Minimizes the negative log-likelihood of ratings and item descriptions

$$\mathcal{L}(\boldsymbol{U},\boldsymbol{V},\boldsymbol{\theta},\boldsymbol{\beta}|\lambda) = \frac{1}{2}\sum_{i,j} c_{ij}\left(p_{ij} - \sum_{k=1}^{K} u_{ik}\cdot v_{jk}\right)^2 - \sum_{j}\sum_{n}\log\left(\sum_{k=1}^{K}\theta_{jk}\cdot\beta_{k,w_{jn}}\right) + \frac{\lambda}{2}\sum_{i=1}^{N}\|\boldsymbol{u}_i\|^2 + \frac{\lambda}{2}\sum_{j=1}^{M}(\boldsymbol{v}_j - \boldsymbol{\theta}_j)^T(\boldsymbol{v}_j - \boldsymbol{\theta}_j)$$

- Iteratively:
  - Optimize for user and item latent vectors $\boldsymbol{u}_i$ and $\boldsymbol{v}_j$ based on the current $\boldsymbol{\theta}_j$
  - Optimize for topic proportions $\boldsymbol{\theta}_j$ based on the current vectors $\boldsymbol{u}_i$ and $\boldsymbol{v}_j$ and topic words $\boldsymbol{\beta}_k$
  - Optimize for topic words $\boldsymbol{\beta}_k$ based on the current topic proportions $\boldsymbol{\theta}_j$
- Prediction
  - For existing items: $\hat{p}_{ij} = \boldsymbol{u}_i^T\boldsymbol{v}_j$
  - For new items: $\hat{p}_{ij} = \boldsymbol{u}_i^T\boldsymbol{\theta}_j$

Generalizing to Unseen Items

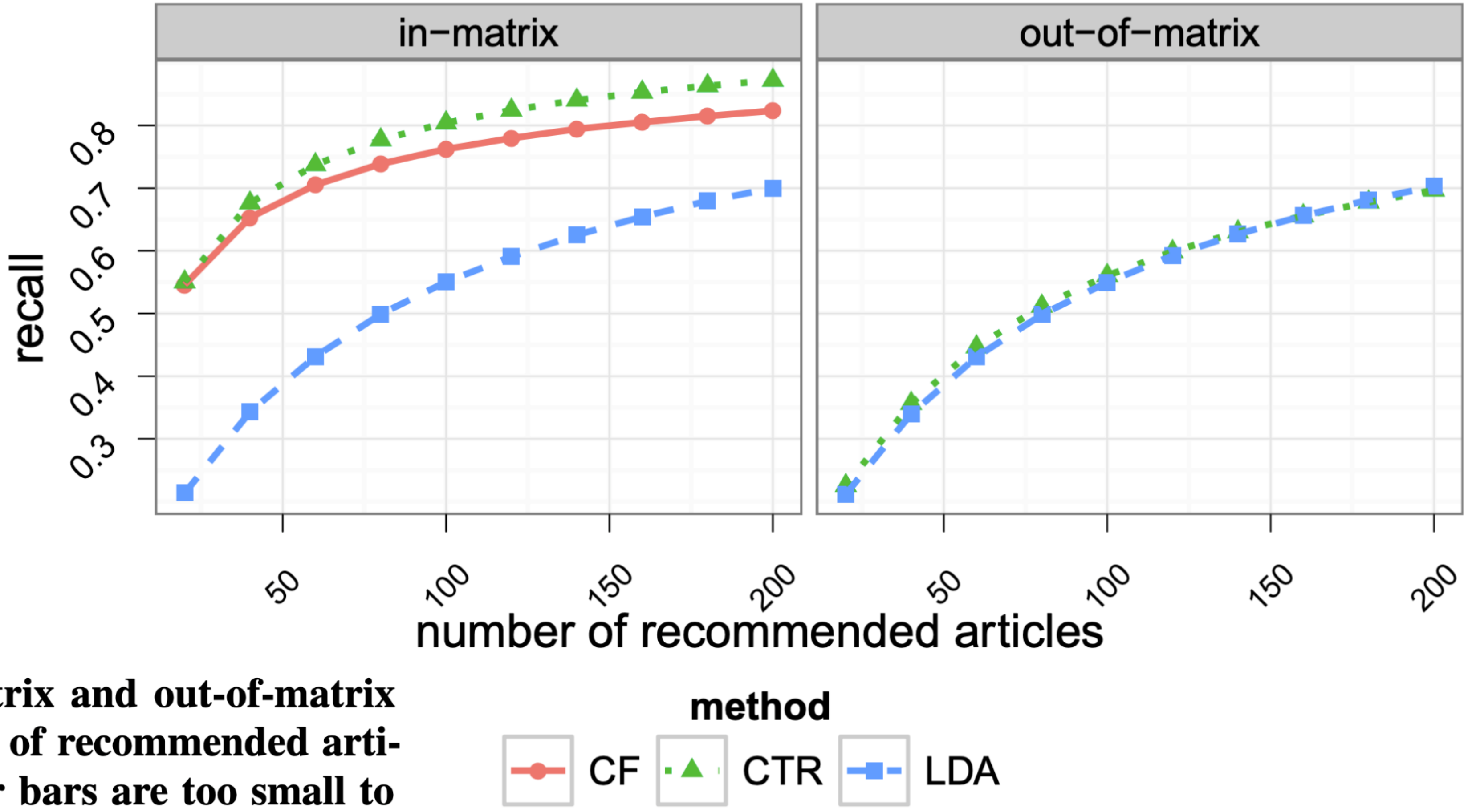


Figure 3: Recall comparison on in-matrix and out-of-matrix prediction tasks by varying the number of recommended articles. For CTR, we set $\lambda_v = 100$. Error bars are too small to show. The maximum expected recall for random recommendation is about 6%. CF can not do out-of-matrix prediction. CTR performs best.

Stacked Denoising Autoencoder (SDAE)

Wang, Wang, & Yeung et al., "Collaborative deep learning for recommender systems ", KDD 2015.

- Let the content for item j be a vector $\boldsymbol{x}_j \in \mathbb{R}^S$ where S is the vocabulary
  - each element of the vector indicates the 'importance' (e.g., tfidf) of word in document j
  - document can be plot of movie, abstract of paper, etc.
- Let the <u>corrupted</u> content be a vector $\boldsymbol{x}_j^0 \in \mathbb{R}^S$
  - with some probability an element of $\boldsymbol{x}_j$ is set to zero
- Encode the corrupted version and decode into the original (clean) version

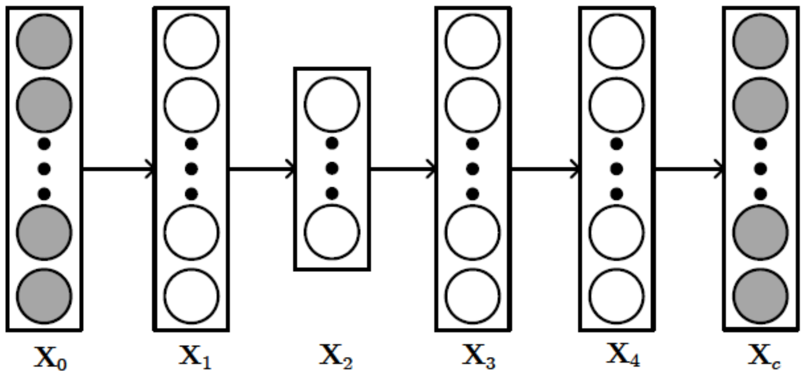
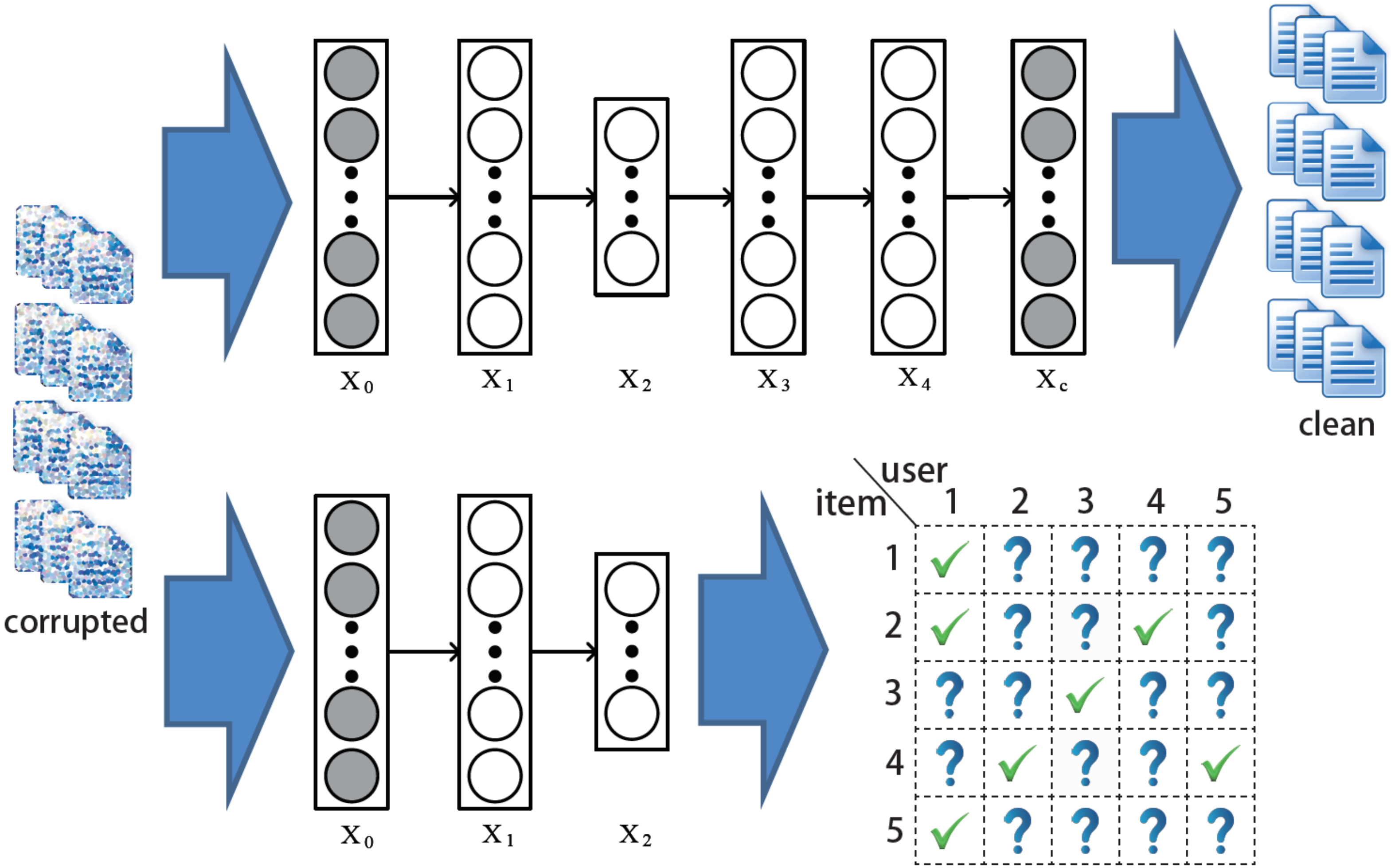


Figure 2: A 2-layer SDAE with $L = 4$.

Collaborative Deep Learning (CDL)

Wang, Wang, & Yeung et al., “Collaborative deep learning for recommender systems ", KDD 2015.



CDL Learning

$$\mathcal{L} =$$

$$\frac{\lambda_u}{2}\sum_i \|\boldsymbol{u}_i\|^2 +$$

- Regularize user vectors

$$\frac{\lambda_w}{2}\sum_l (\|\boldsymbol{W}_l\|^2 + \|\boldsymbol{b}_l\|^2) +$$

- Regularize autoencoder parameters

$$\frac{\lambda_v}{2}\sum_j \left\|\boldsymbol{v}_j - \boldsymbol{x}_j^{L/2}\right\|^2 +$$

- Minimize offset so item vector is close to its middle encoding

$$\frac{\lambda_n}{2}\sum_i \left\|\boldsymbol{x}_j^{\boldsymbol{L}} - \boldsymbol{x}_j^{c}\right\|^2 +$$

- Ensure the autoencoder reconstruction is similar to the clean output

$$\sum_{i,j} \frac{c_{ij}}{2}\left(p_{ij} - \boldsymbol{u}_i^T \boldsymbol{v}_j\right) +$$

- Minimize error in rating prediction

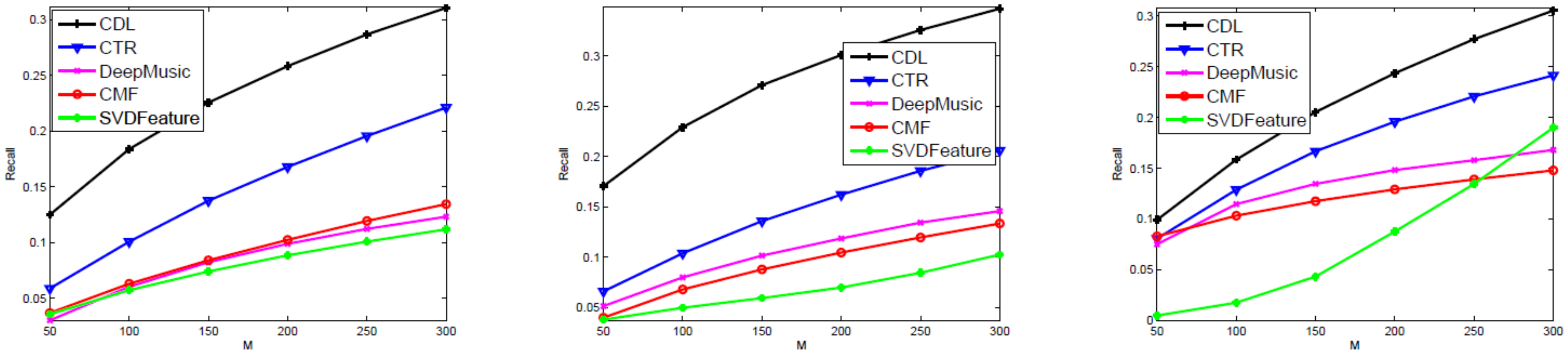


Figure 4: Performance comparison of CDL, CTR, DeepMusic, CMF, and SVDFeature based on recall@M for datasets *citeulike-a*, *citeulike-t*, and *Netflix* in the sparse setting. A 2-layer CDL is used.

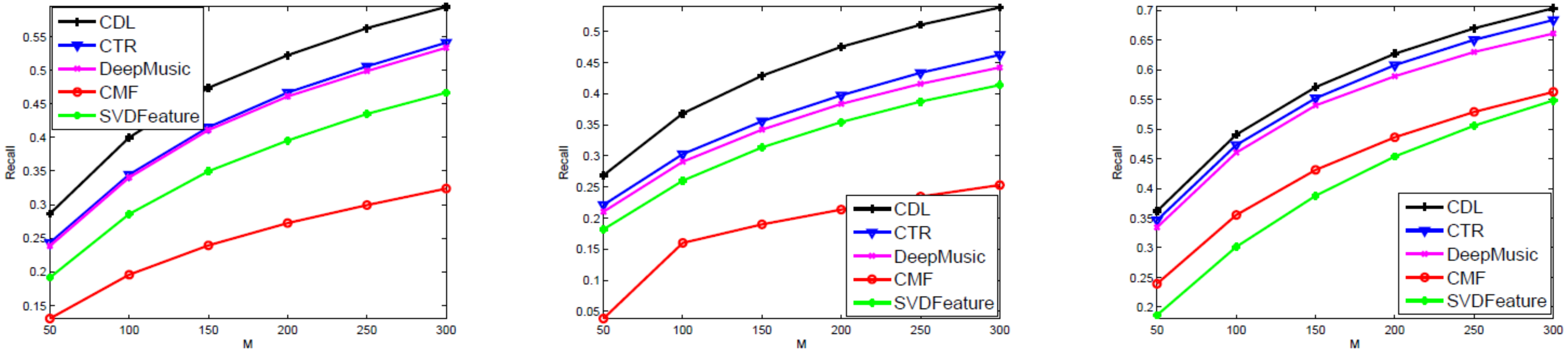
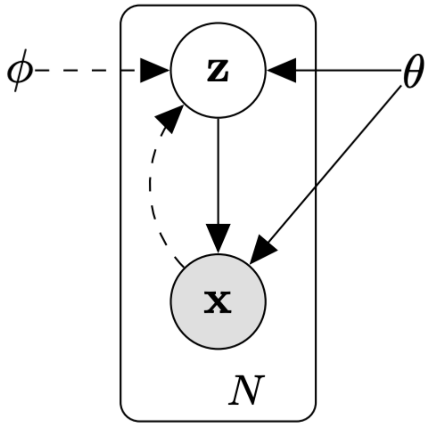
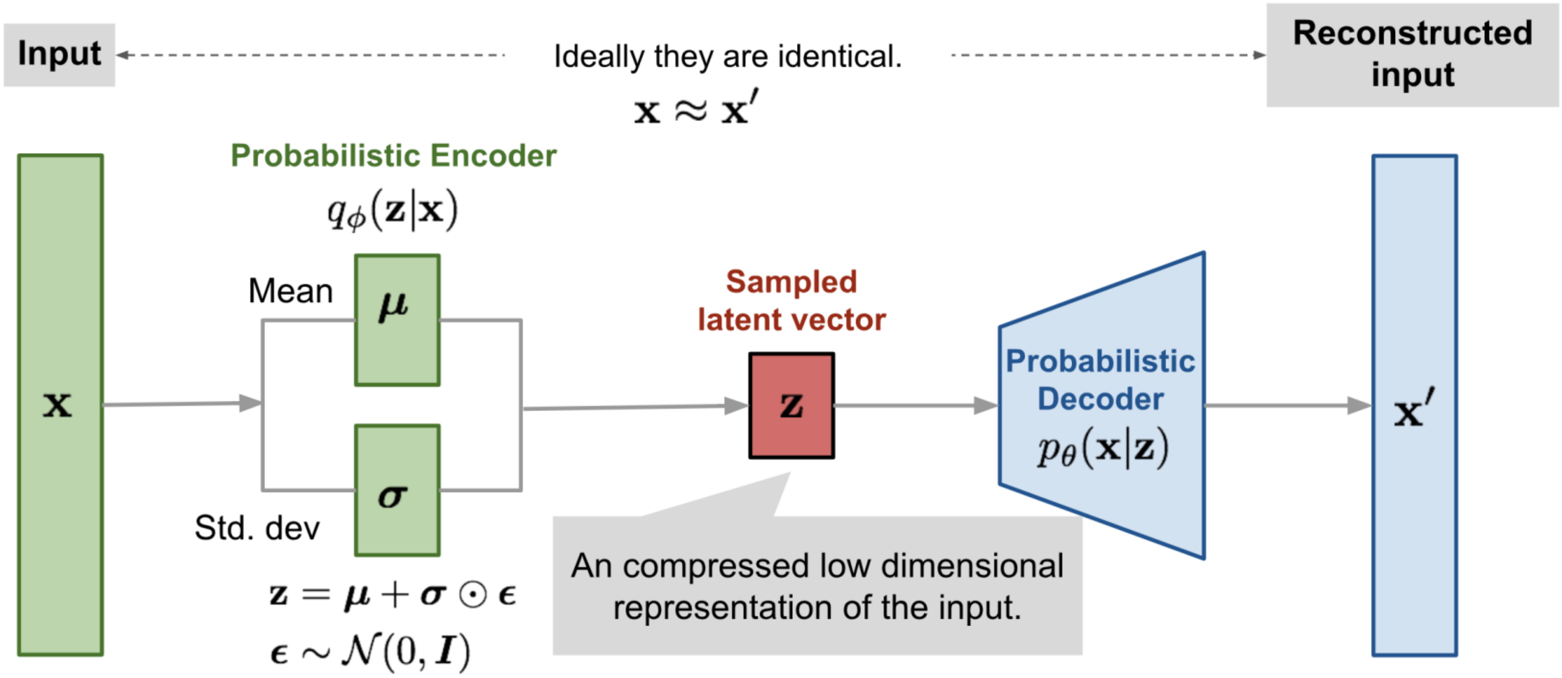


Figure 5: Performance comparison of CDL, CTR, DeepMusic, CMF, and SVDFeature based on recall@M for datasets *citeulike-a*, *citeulike-t*, and *Netflix* in the dense setting. A 2-layer CDL is used.

Variational Auto-Encoder (VAE)

Kingma, Diederik P., and Max Welling. "Auto-encoding variational bayes.", ICLR 2014.





https://lilianweng.github.io/posts/2018-08-12-vae/

VAE for Text Modeling

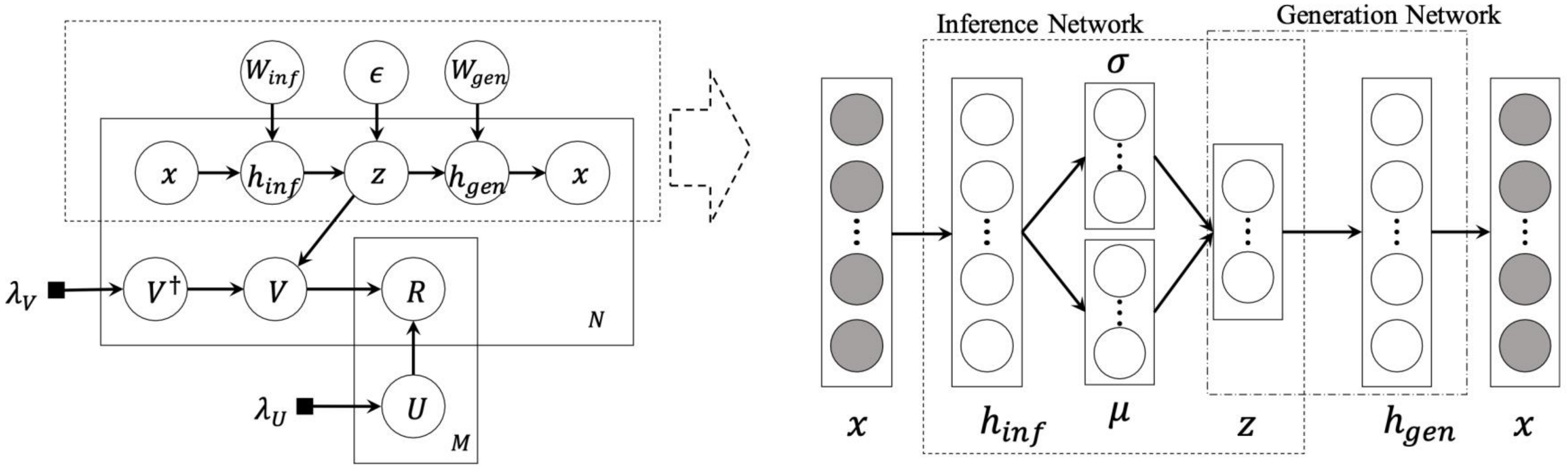


Figure 2: On the left is the proposed Collaborative Variational Autoencoder (CVAE); on the right is the zoom-in of the inference network and generation network in CVAE.

- Li, Xiaopeng, and James She. "Collaborative variational autoencoder for recommender systems.", KDD 2017
- Zhu, Yaochen, and Zhenzhong Chen. "Mutually-regularized dual collaborative variational auto-encoder for recommendation systems.", WWW 2022

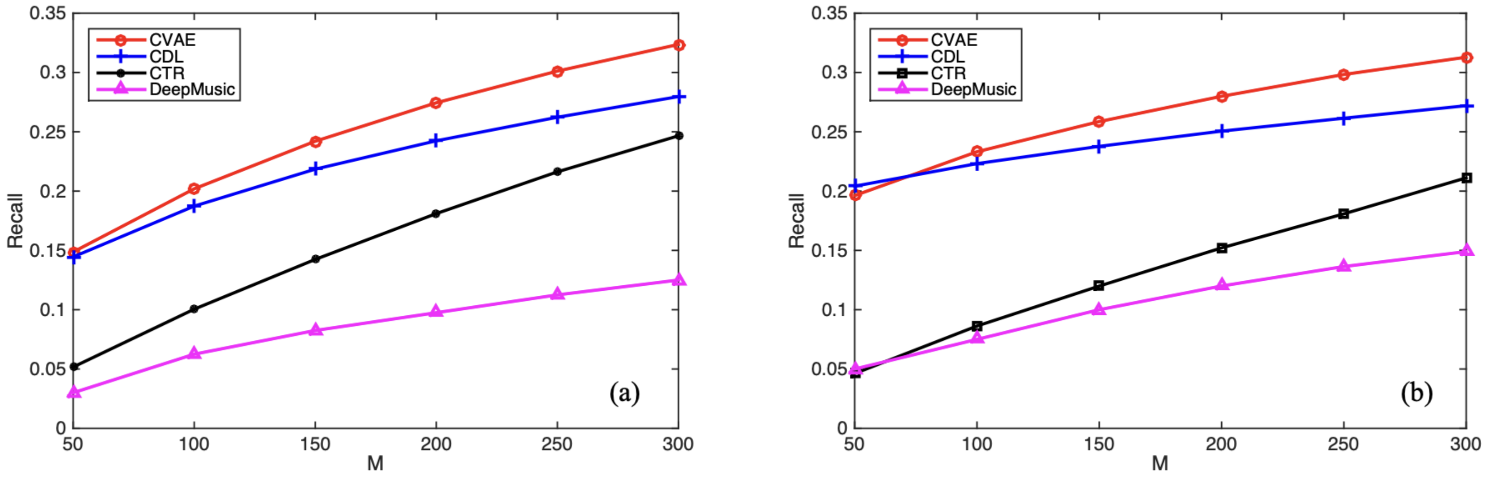


Figure 4: Performance comparison of CVAE, CDL, CTR and DeepMusic based on recall in the sparse setting for dataset (a) *citeulike-a* and (b) *citeulike-t*.

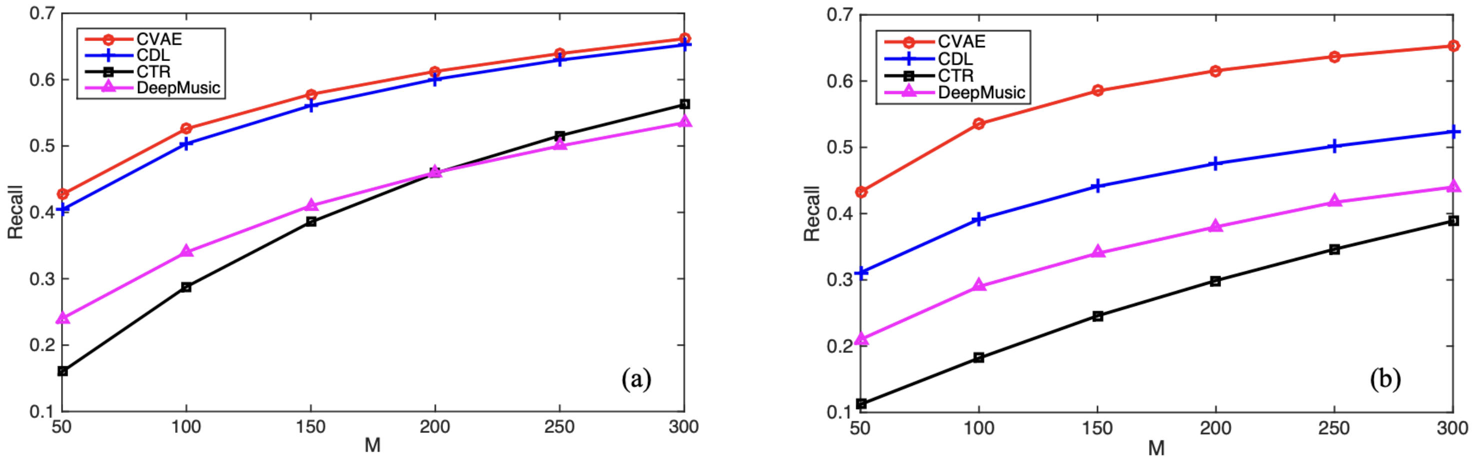
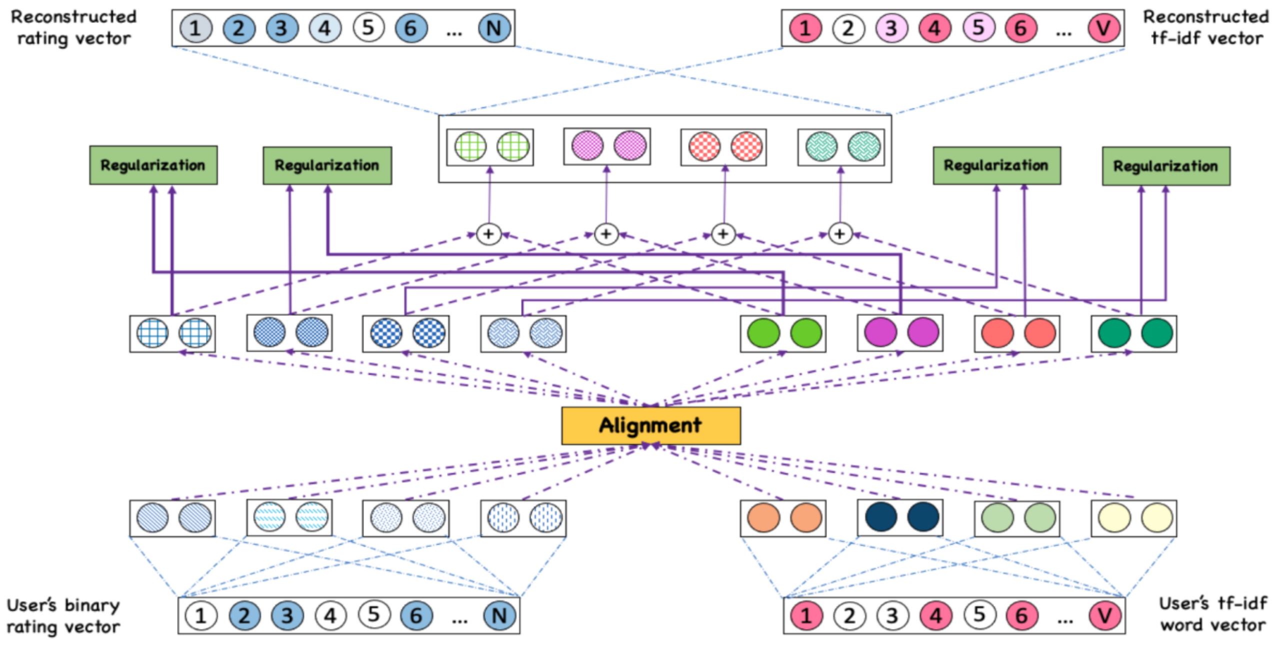


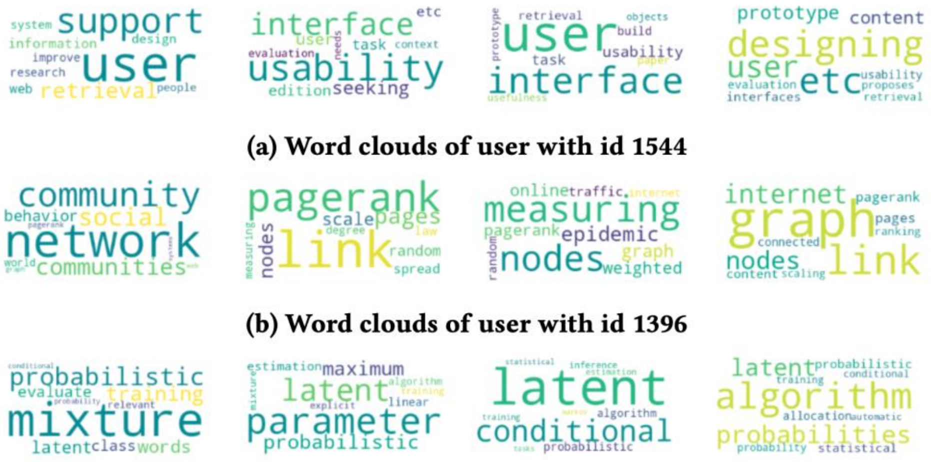
Figure 5: Performance comparison of CVAE, CDL, CTR and DeepMusic based on recall in the dense setting for dataset (a) *citeulike-a* and (b) *citeulike-t*.

Disentangled User Representations

Tran, Nhu-Thuat, and Hady W. Lauw. "Aligning Dual Disentangled User Representations from Ratings and Textual Content.", KDD 2022.



AddVAE Model



(a) Word clouds of user with id 1544

(b) Word clouds of user with id 1396

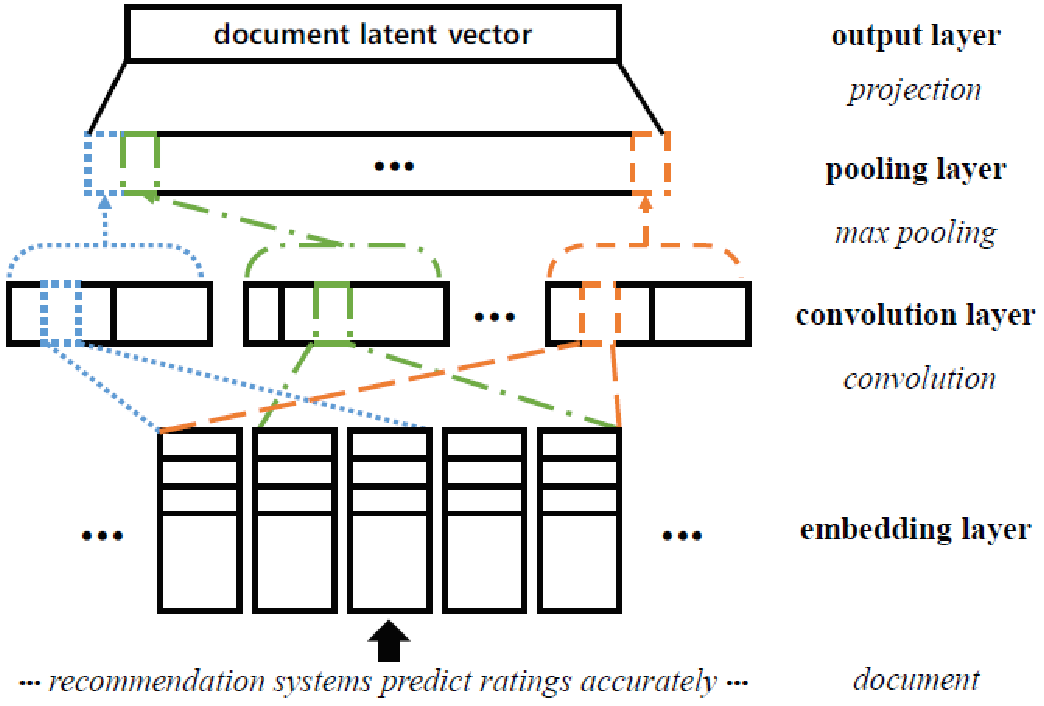
(c) Word clouds of user with id 5542

Figure 2: On Citeulike-a dataset with $K = 4$, items that user interacts with are scientific articles. Each user has four word clouds representing their topic of interest corresponding to four factors. Size of word encodes their predicted score given by AddVAE. Best viewed in color.

Convolutional Matrix Factorization (ConvMF)

Kim et al., “Convolutional matrix factorization for document context-aware recommendation ", RecSys 2016.

- Represent each product as a text document
  - e.g., product description, combine reviews into a single document
  - a term vector $\boldsymbol{x}_j \in \mathbb{R}^S$ where S is the vocabulary
- Embedding layer
  - each column initialized with p-dimensional word vector (e.g., Glove) of each word
- Derive the hidden representation using CNN



Model	Ratio of training set to the entire dataset (density)						
	20% (0.93%)	30% (1.39%)	40% (1.86%)	50% (2.32%)	60% (2.78%)	70% (3.25%)	80% (3.71%)
PMF	1.0168	0.9711	0.9497	0.9354	0.9197	0.9083	0.8971
CTR	1.0124	0.9685	0.9481	0.9337	0.9194	0.9089	0.8969
CDL	1.0044	0.9639	0.9377	0.9211	0.9068	0.8970	0.8879
ConvMF	**0.9745**	**0.9330**	**0.9063**	**0.8897**	**0.8726**	**0.8676**	**0.8531**
Improve	2.98%	3.20%	3.36%	3.41%	3.77%	3.27%	3.92%

Table 4: Test RMSE over various sparseness of training data on ML-1m dataset

Cornac-Supported Text-Based Models

- Collaborative Topic Modeling (CTR)
- Hidden Factors and Hidden Topics (HFT)
- Collaborative Deep Learning (CDL)
- Collaborative Deep Ranking (CDR)
- Collaborative Variational Autoencoder (CVAE)
- Convolutional Matrix Factorization (ConvMF)
- …

Image Modality

Image-Modality

Image Modeling	Explicit (MF or PMF)	Implicit (BPR or WMF)
Pre-trained Embedding	VMF, VPOI	VBPR, ACF, DMRL, NPR
Convolutional Neural Nets		DVBPR, CF2Label, CKE, CDL, JRL

References

- Chanyoung Park, Donghyun Kim, Jinoh Oh, and Hwanjo Yu. 2017. Do" Also-Viewed" Products Help User Rating Prediction?. In WWW. 1113–1122.
- Suhang Wang, Yilin Wang, Jiliang Tang, Kai Shu, Suhas Ranganath, and Huan Liu. 2017. What your images reveal: Exploiting visual contents for point-of-interest recommendation. In WWW. 391–400.
- Ruining He and Julian McAuley. 2016. VBPR: visual bayesian personalized ranking from implicit feedback. In AAAI.
- Jingyuan Chen, Hanwang Zhang, Xiangnan He, Liqiang Nie, Wei Liu, and Tat-Seng Chua. 2017. Attentive collaborative filtering: Multimedia recommendation with item-and component-level attention. In SIGIR. 335–344.
- Wei Niu, James Caverlee, and Haokai Lu. 2018. Neural personalized ranking for image recommendation. In WSDM. 423–431.
- Wang-Cheng Kang, Chen Fang, Zhaowen Wang, and Julian McAuley. 2017. Visually-aware fashion recommendation and design with generative image models. In ICDM. 207–216.
- Chenyi Lei, Dong Liu, Weiping Li, Zheng-Jun Zha, and Houqiang Li. 2016. Comparative deep learning of hybrid representations for image recommendations. In CVPR. 2545–2553.
- Fuzheng Zhang, Nicholas Jing Yuan, Defu Lian, Xing Xie, and Wei-Ying Ma. 2016. Collaborative knowledge base embedding for recommender systems. In SIGKDD. 353–362.
- Yongfeng Zhang, Qingyao Ai, Xu Chen, and W Bruce Croft. 2017. Joint representation learning for top-n recommendation with heterogeneous information sources. In CIKM. 1449–1458.
- Ying, R., He, R., Chen, K., Eksombatchai, P., Hamilton, W. L., & Leskovec, J. (2018, July). Graph convolutional neural networks for web-scale recommender systems. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 974-983.
- Liu, F., Chen, H., Cheng, Z., Liu, A., Nie, L., & Kankanhalli, M. (2022). Disentangled Multimodal Representation Learning for Recommendation. *IEEE Transactions on Multimedia*.
- Bibas, K., Sar Shalom, O., & Jannach, D. (2022, October). Collaborative Image Understanding. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management* (pp. 77-87).

Visual BPR

He and McAuley, "VBPR: visual bayesian personalized ranking from implicit feedback ", AAAI 2016.

- Item j has K-dimensional latent vector $\boldsymbol{v}_j$ (in matrix factorization sense)

$$\hat{r}_{ij} = \boldsymbol{u}_i^T \boldsymbol{v}_j$$

- Item j also has an image feature vector $\boldsymbol{f}_j \in \mathbb{R}^D$, which is of length $D \neq K$
  - D can be very high-dimensional
  - To reduce its dimensionality, introduce a projection matrix $\boldsymbol{E} \in \mathbb{R}^{Q \times D}$, where $Q \ll D$
  - The projection $\boldsymbol{E} \times \boldsymbol{f}_j \in \mathbb{R}^Q$ is now of length Q

- We have two representations of items
  - From collaborative filtering $\boldsymbol{v}_j$
  - From image features $\boldsymbol{E} \times \boldsymbol{f}_j$

- Intuition: user's preference is influenced by both representations

Prediction

- The preference by user i on item j is modeled as follows:

$$\hat{r}_{ij} = b_j + \boldsymbol{u}_i^T \boldsymbol{v}_j + \boldsymbol{\rho}_i^T \left(\boldsymbol{E} \times \boldsymbol{f}_j\right) + \boldsymbol{\Theta}^T \boldsymbol{f}_j$$

  - $\boldsymbol{f}_j \in \mathbb{R}^D$ is the image feature vector (input to the model)
  - $\boldsymbol{\Theta} \in \mathbb{R}^D$ is visual bias vector (learned)
  - $\boldsymbol{\Theta}^T \boldsymbol{f}_j$ indicates the general popularity of item with particular visual features

  - $\boldsymbol{u}_i \in \mathbb{R}^K$ is user latent vector and $\boldsymbol{v}_j \in \mathbb{R}^K$ is item latent vector (learnt)
  - $\boldsymbol{u}_i^T \boldsymbol{v}_j$ indicates preference by user i on item j based on collaborative filtering

  - $\boldsymbol{\rho}_i \in \mathbb{R}^Q$ is user visual preference and $\left(\boldsymbol{E} \times \boldsymbol{f}_j\right) \in \mathbb{R}^Q$ is item visual representation (learned)
  - $\boldsymbol{\rho}_i^T \left(\boldsymbol{E} \times \boldsymbol{f}_j\right)$ indicates preference by user i on item j based on visual signals

Learning

- Input:
  - Positive triples $\boldsymbol{S} = \{j >_i l | r_{ij} \in \boldsymbol{R}^+ \wedge r_{il} \in \boldsymbol{R}^-\}$
  - Image features $\{\boldsymbol{f}_j\}$

- Minimize regularized negative log-likelihood function:

$$\mathcal{L}(\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{b}, \boldsymbol{\Theta}, \mathbf{E}, \mathbf{P} | \lambda)$$
$$= \sum_{(j>_i l) \in \boldsymbol{S}} \ln(1 + \exp\{-(\hat{r}_{ij} - \hat{r}_{il})\}) + \frac{\lambda}{2} \sum_{i=1}^{N} (\|\boldsymbol{u}_i\|^2 + \|\boldsymbol{\rho}_i\|^2) + \frac{\lambda}{2} \sum_{j=1}^{M} \left(b_j^2 + \|\boldsymbol{v}_j\|^2\right) + \frac{\lambda}{2} \|\boldsymbol{\Theta}\|^2 + \frac{\lambda}{2} \|\boldsymbol{E}\|^2$$

Experiments on Fashion Datasets

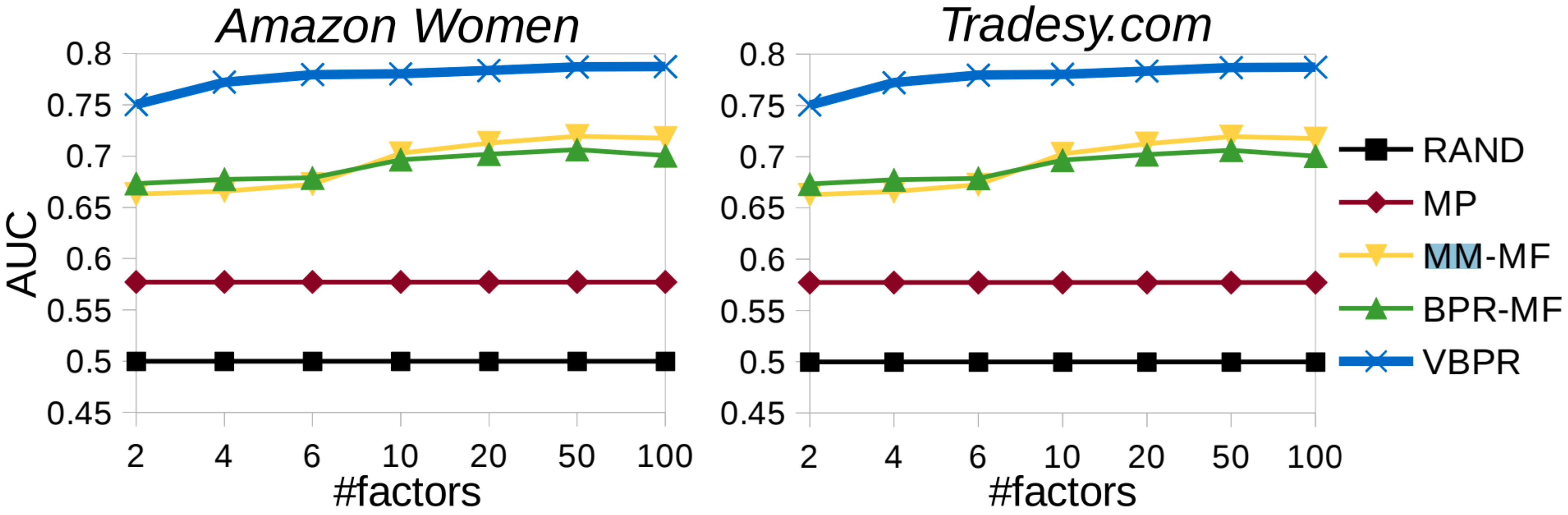


Figure 2: AUC with varying dimensions.

Figure 4: 2-D visualization (with t-SNE (?)) of the 10-D visual space learned from *Amazon Women*. All images are from the test set. For clarity, the space is discretized into a grid and for each grid cell one image is randomly selected among overlapping instances.

Deep Bayesian Personalized Ranking (DVBPR)

Visually-Aware Fashion Recommendation and Design with Generative Image Models (ICDM'17)

- VBPR preference estimate: $\hat{r}_{ij} = b_j + \boldsymbol{u}_i^T \boldsymbol{v}_j + \boldsymbol{\rho}_i^T (\boldsymbol{E} \times \boldsymbol{f}_j) + \boldsymbol{\Theta}^T \boldsymbol{f}_j$

- DVBPR preference estimate: $\hat{r}_{ij} = \boldsymbol{\rho}_i^T \Phi(\boldsymbol{X}_j)$

- Pre-trained visual features and embedding matrix are replaced with a CNN network $\Phi(.)$ to extract visual features directly from the images themselves

- Item bias terms b_j and non-visual latent factors $\boldsymbol{u}_i, \boldsymbol{v}_j$ are excluded. Empirically, doing so improves performance (the remaining terms are sufficient to capture these factors implicitly).

Visual Matrix Factorization (VMF)

Do “Also-Viewed” Products Help User Rating Prediction? (WWW’17)

- VMF preference estimate is almost the same as VBPR, except bias removed:

$$\hat{r}_{ij} = \boldsymbol{u}_i^T \boldsymbol{v}_j + \boldsymbol{\rho}_i^T \left(\boldsymbol{E} \times \boldsymbol{f}_j \right)$$

- Learning via minimizing MSE instead of BPR criteria

$$\mathcal{L}(\boldsymbol{U}, \boldsymbol{V} | \lambda) = \frac{1}{2} \sum_{i,j} \left(r_{ij} - \hat{r}_{ij} \right)^2 + \frac{\lambda}{2} \sum_{i=1}^{N} \|\boldsymbol{u}_i\|^2 + \frac{\lambda}{2} \sum_{j=1}^{M} \left\| \boldsymbol{v}_j \right\|^2$$

Attentive Collaborative Filtering (ACF)

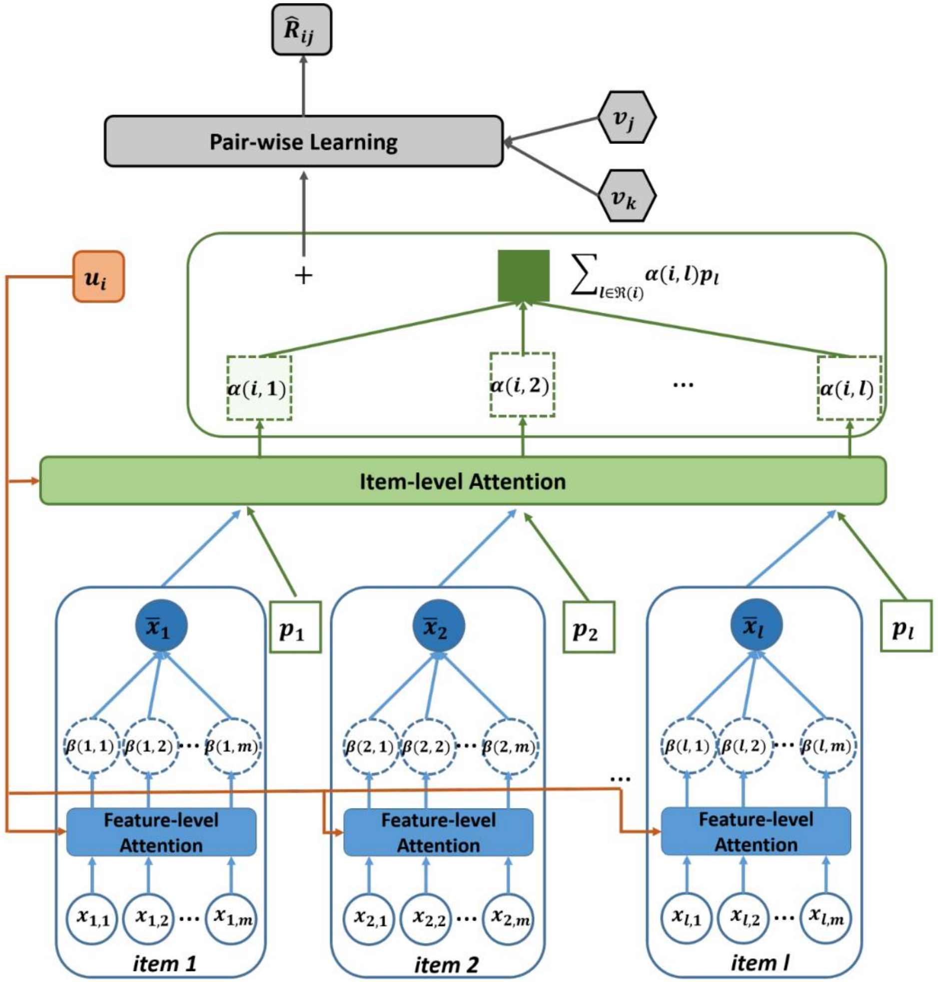
Attentive Collaborative Filtering: Multimedia Recommendation with Item- and Component-Level Attention (SIGIR'17)

- Item visual components:
  - Spatial regions of an image
  - Frames of a video

- Neighborhood-based model:

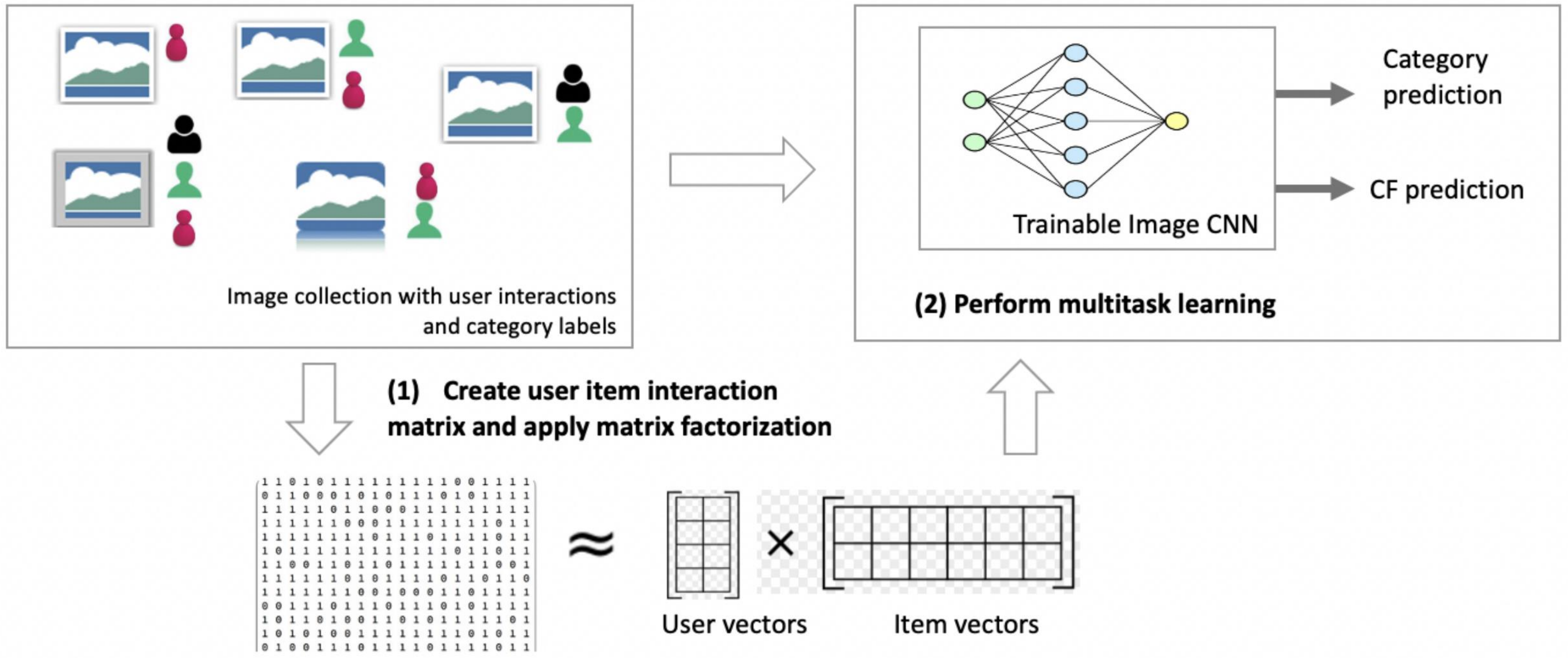
$$\hat{R}_{ij} = \overbrace{\mathbf{u}_i^T \mathbf{v}_j}^{latent\ factor\ model} + \underbrace{\sum_{l \in \mathcal{R}(i)} \alpha(i,l)\mathbf{p}_l^T \mathbf{v}_j}_{neighborhood\ model}$$

- ACF optimizes BPR objective



CF2Label (Multitask Learning)

Collaborative Image Understanding (CIKM'22)



Disentangled Multimodal Representation Learning (DMRL)

Disentangled Multimodal Representation Learning for Recommendation (IEEE Trans. on Multimedia 2022)

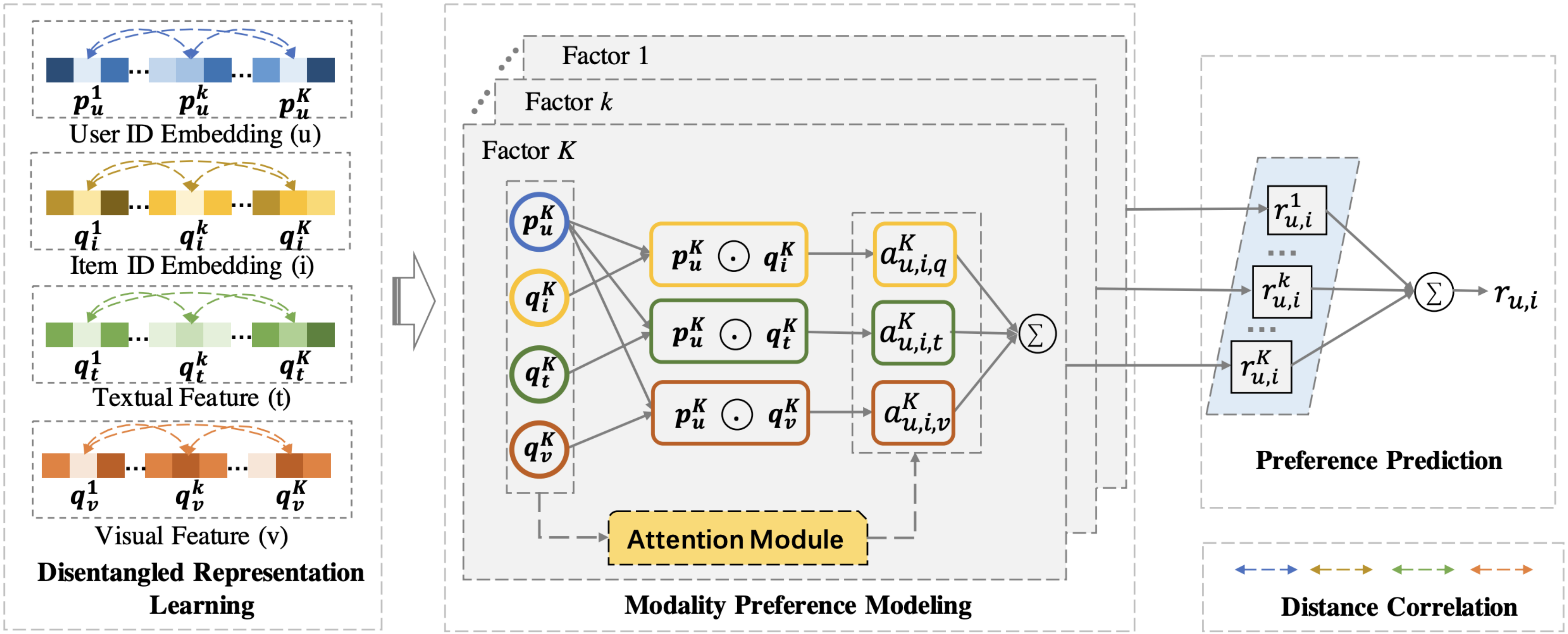
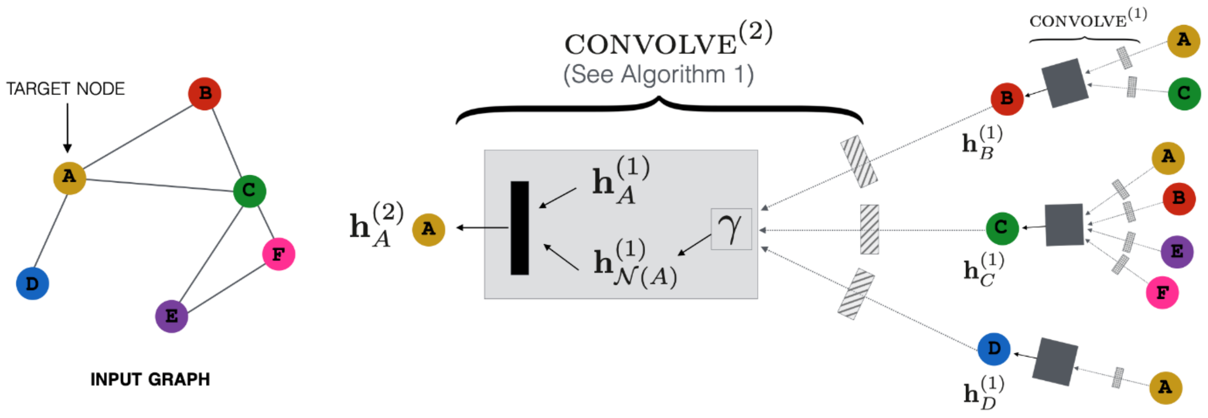
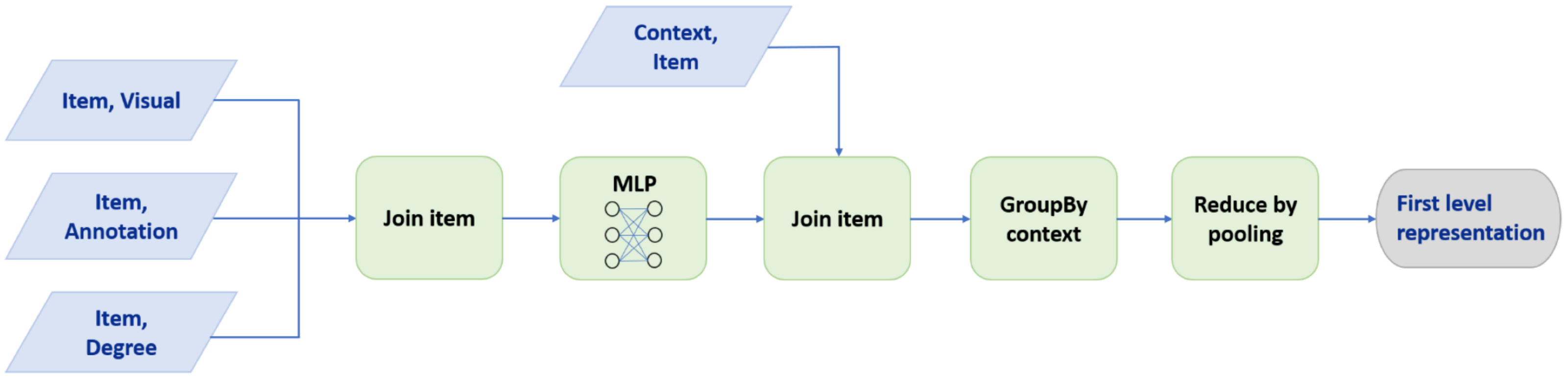


Fig. 1: Overview of our DMRL model. Best viewed in color.

PinSage



- Learning node embedding with GNN
- Node embedding includes visual features



- Recommendation by nearest-neighbor lookup in the learned embedding space

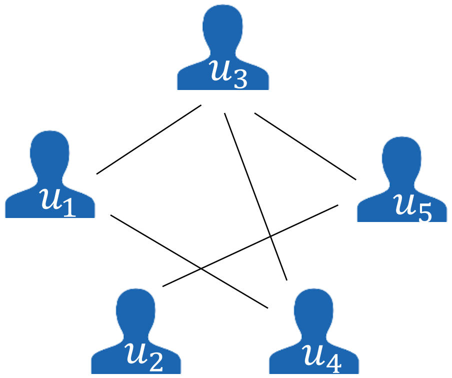
Graph Convolutional Neural Networks for Web-Scale Recommender Systems (KDD’18)

Cornac-Supported Visual-Based Models

- Visual Bayesian Personalized Ranking (VBPR)
- Visual Matrix Factorization (VMF)
- Adversarial Multimedia Recommendation (AMR)
- …

Graph Modality

User-Side: Social Relationships



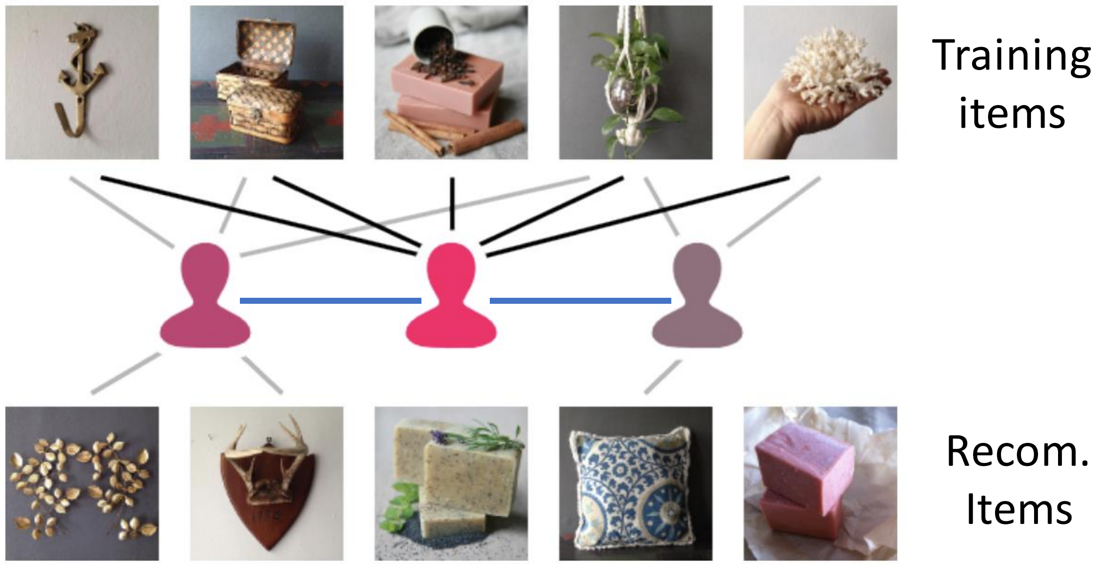
Undirected (e.g., friendship)

$$
\begin{array}{c|ccccc}
 & u_1 & u_2 & u_3 & u_4 & u_5 \\
\hline
u_1 & 0 & 0 & 1 & 1 & 0 \\
u_2 & 0 & 0 & 0 & 0 & 1 \\
u_3 & 1 & 0 & 0 & 1 & 0 \\
u_4 & 1 & 0 & 1 & 0 & 0 \\
u_5 & 0 & 1 & 1 & 0 & 0
\end{array}
$$

Adjacency Matrix, Symmetric

Social Collaborative Filtering: Intuition

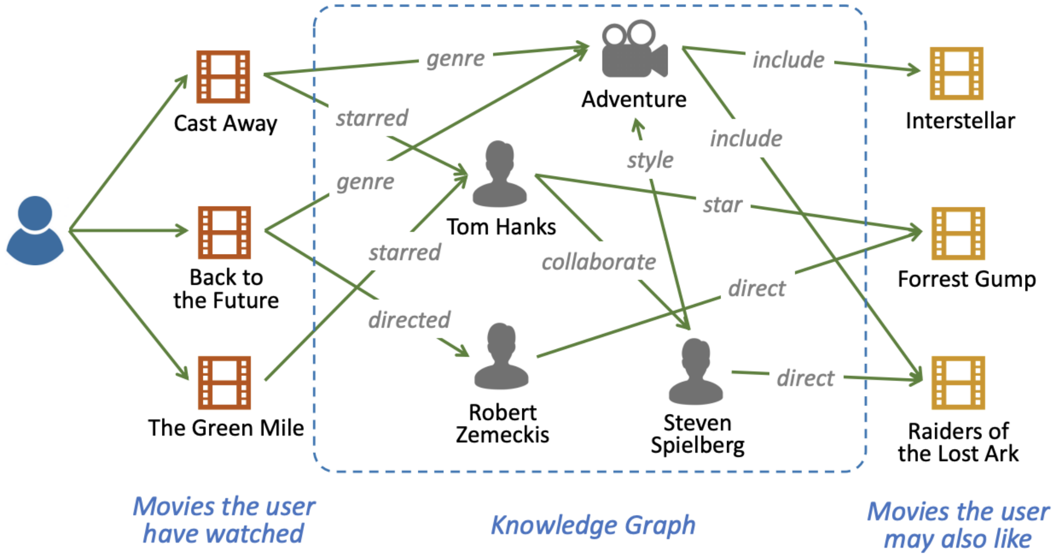
- Our consumption behavior is biased by our social connections
- Two signals driving user preferences
  - Collaborative, capturing general user preferences
  - Social, capturing influence from friends
- Social Collaborative Filtering aims to capture user preferences from both signals



Target user (center) along with her training items (top), recommendations (bottom), and her friends (left and right users). [Chaney, A. J. et al, RecSys'15]

Item-Side

- We are interested in ***item-relatedness***



Knowledge Graph

Browsed/Bought Together

Why is Item-Relatedness Important

- We tend to consume items which
  - can complement each other
  - are alternatives to each other (substitutes)
- Item-relatedness goes beyond feature similarities
  - Difficult to capture with the text or image modalities

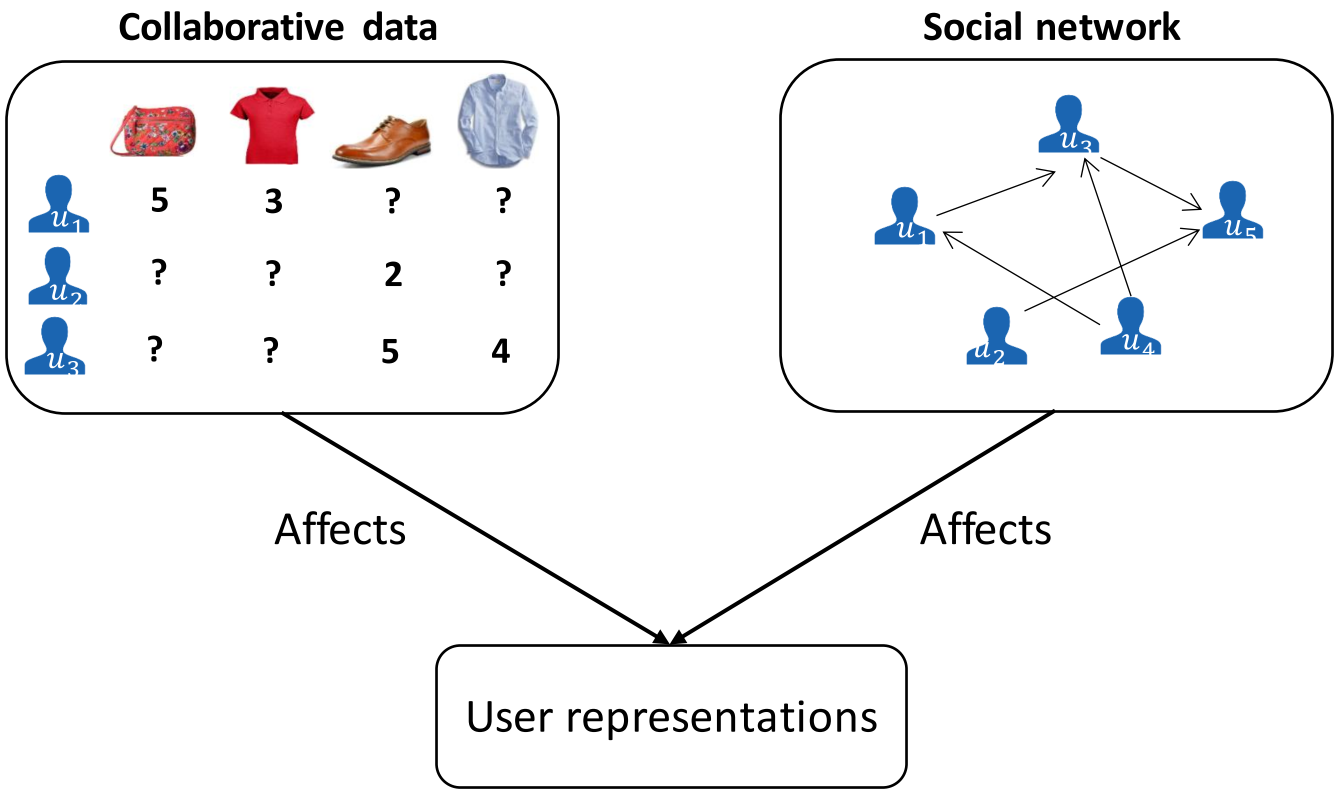
Substitutes & Complements

Main Families: Model-Centric

- **Feature-based**
  - Derive user/item representations from graph and preference data
  - Existing methods differ mainly in the models used for each modality
- **Regularization-based**
  - Regularize the latent space of the CF model using the graph modality
  - Users/items which are connected are encouraged to have similar latent representations
- **Graph-aware architectures**
  - Embed the graph structure into the CF model's architecture
  - Learning is driven by the collaborative signal only

Feature-Based Approaches

Any adequate models can be used to represent the collaborative and graph signals

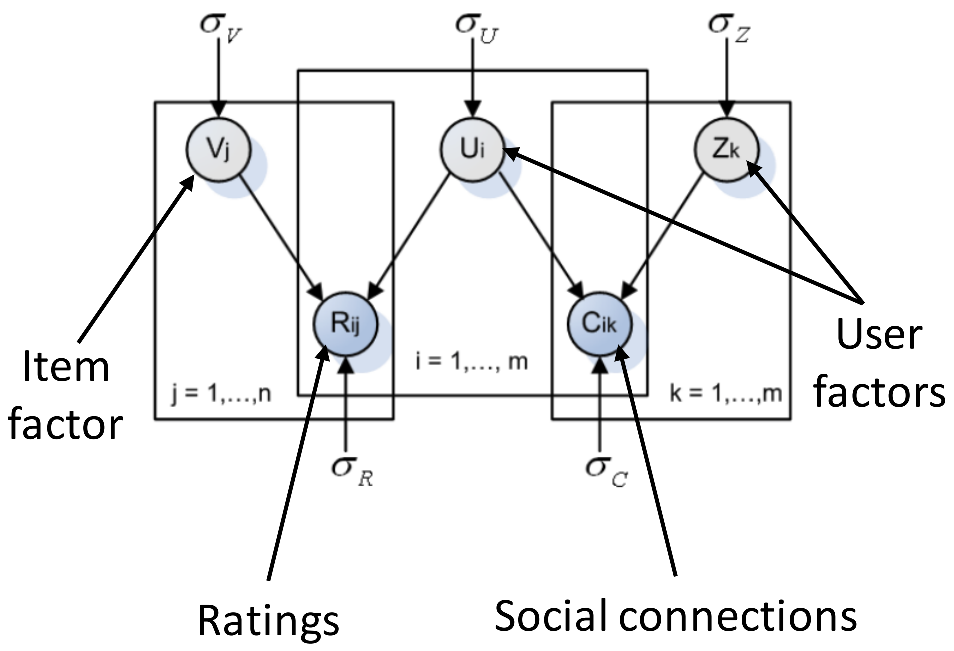


The big picture with User-User graph

SoRec Model

Ma, Yang, Lyu, and King, "SoRec: Social Recommendation Using Probabilistic Matrix Factorization", CIKM 2009.

SoRec graphical model



- Modeling assumptions

$$\boldsymbol{U}_i \sim N(\boldsymbol{0}, \sigma_u^2 \boldsymbol{I}) \qquad \boldsymbol{C}_{ik} | \boldsymbol{U}, \boldsymbol{Z} \sim N(g(\boldsymbol{U}_i^\top \boldsymbol{Z}_k), \sigma_C^2)$$

$$\boldsymbol{Z}_k \sim N(\boldsymbol{0}, \sigma_z^2 \boldsymbol{I}) \qquad \boldsymbol{R}_{ij} | \boldsymbol{U}, \boldsymbol{V} \sim N(g(\boldsymbol{U}_i^\top \boldsymbol{V}_j), \sigma_R^2)$$

$$V_j \sim N(\boldsymbol{0}, \sigma_v^2 \boldsymbol{I})$$

- Loss function

$$\sum_{i=1}^{m} \sum_{j=1}^{n} I_{ij}^{R} \left(r_{ij} - g\left(\boldsymbol{U}_i^\top \boldsymbol{V}_j \right) \right)^2 + \lambda_C \sum_{i=1}^{m} \sum_{k=1}^{m} I_{ik}^{C} \left(c_{ik} - g\left(\boldsymbol{U}_i^\top \boldsymbol{Z}_k \right) \right)^2 + \ Reg.$$

- Gradient with respect to $\boldsymbol{U}_i$ components

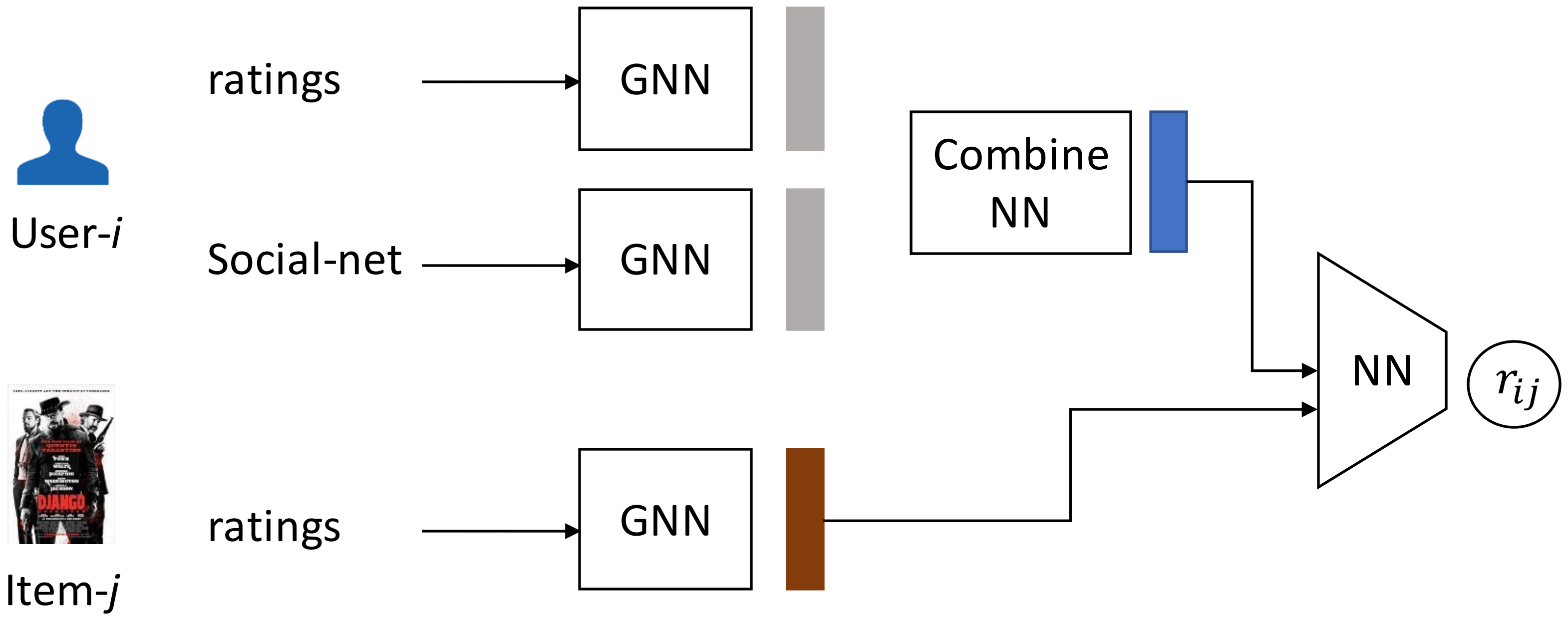
$$\frac{d\mathcal{L}}{du_{ik}} = \sum_{j:\, r_{ij} \in \boldsymbol{R}} -v_{jk} \cdot \left(r_{ij} - \boldsymbol{U}_i^\top \boldsymbol{V}_j \right) + \lambda_C \sum_{h:\, g_{ih} \in \boldsymbol{G}} -z_{hk} \cdot \left(c_{ik} - \boldsymbol{U}_i^\top \boldsymbol{Z}_k \right) + \lambda \cdot u_{ik}$$

PREFERRED.AI

GraphRec

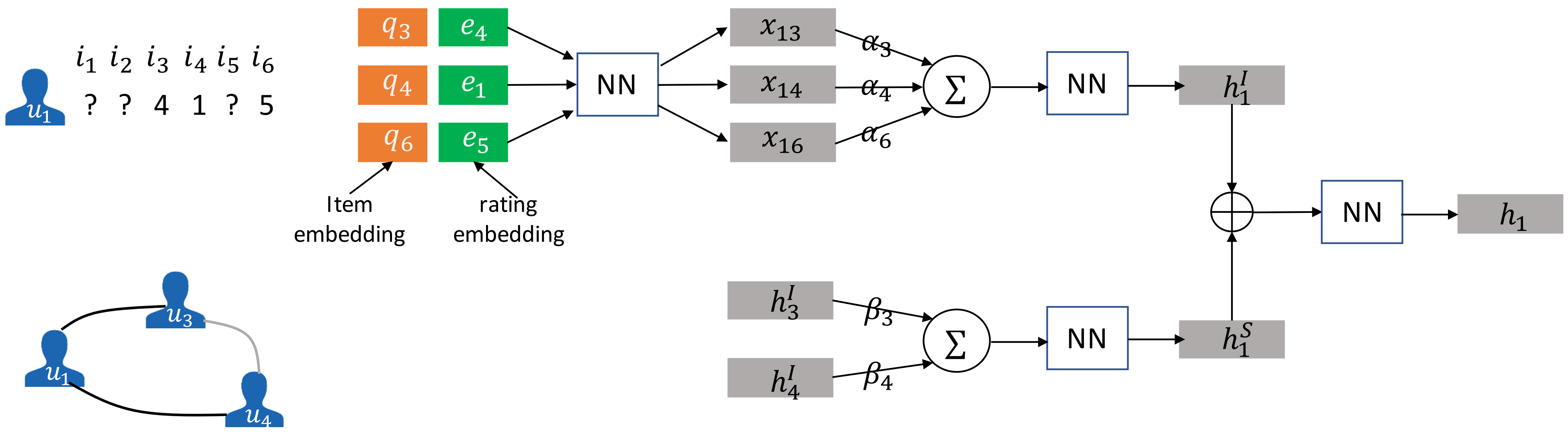
Fan, Wenqi, et al. "Graph neural networks for social recommendation." *WWW*. 2019.

- Uses GNNs to learn user/item representations from user-item preferences and social network data

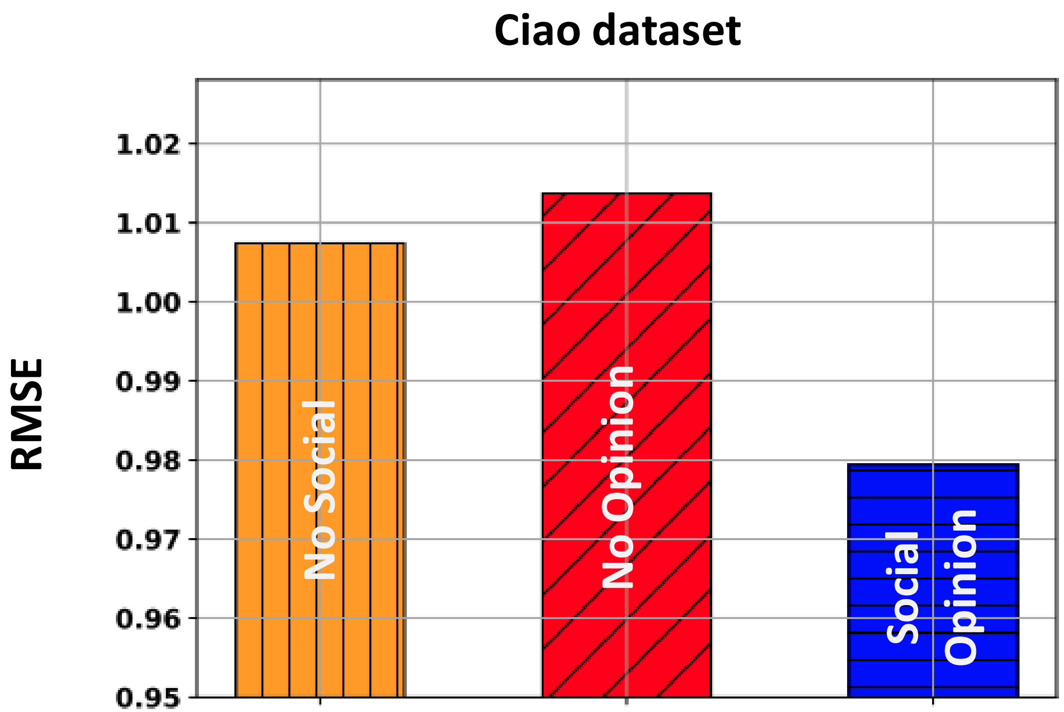


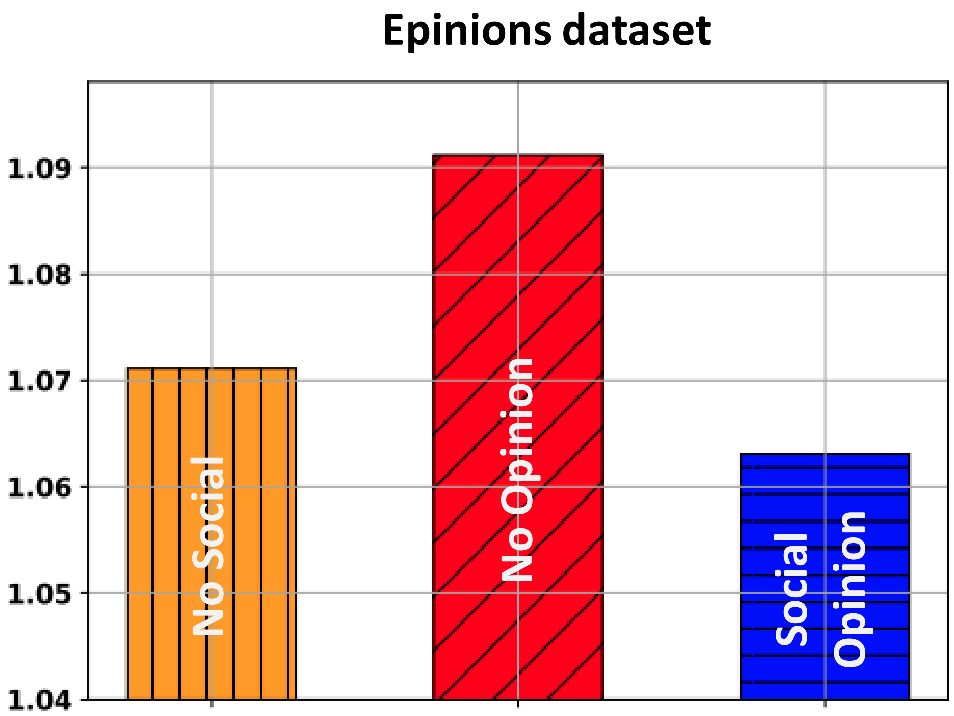
GNN-based user representation

Fan, Wenqi, et al. "Graph neural networks for social recommendation." *WWW*. 2019.



GraphRec: Results





RMSE: Root Mean Square Error

Effect of Social Network and User Opinions

Regularization-based methods

- **Core idea:** bring the latent representations of connected users/items closer to each other
- Act at the objective function level by introducing a suitable regularization term

Objective = Collaborative-Objective + $\lambda \times$ *Graph-Aware Regularization*

Regularization paramter

Pull closer the representations of connected users/items

SoReg: Social Regularized Matrix Factorization

Ma, Hao, et al. "Recommender systems with social regularization." *WSDM*. 2011.

- SoReg's objective function

$$\min_{U,V} \mathcal{L}_1(R,U,V) = \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n} I_{ij}(R_{ij} - U_i^T V_j)^2$$

- Collaborative MF term

$$+ \frac{\alpha}{2}\sum_{i=1}^{m} \| U_i - \frac{\sum_{f \in \mathcal{F}^+(i)} Sim(i,f) \times U_f}{\sum_{f \in \mathcal{F}^+(i)} Sim(i,f)} \|_F^2$$

- Average-based regularization

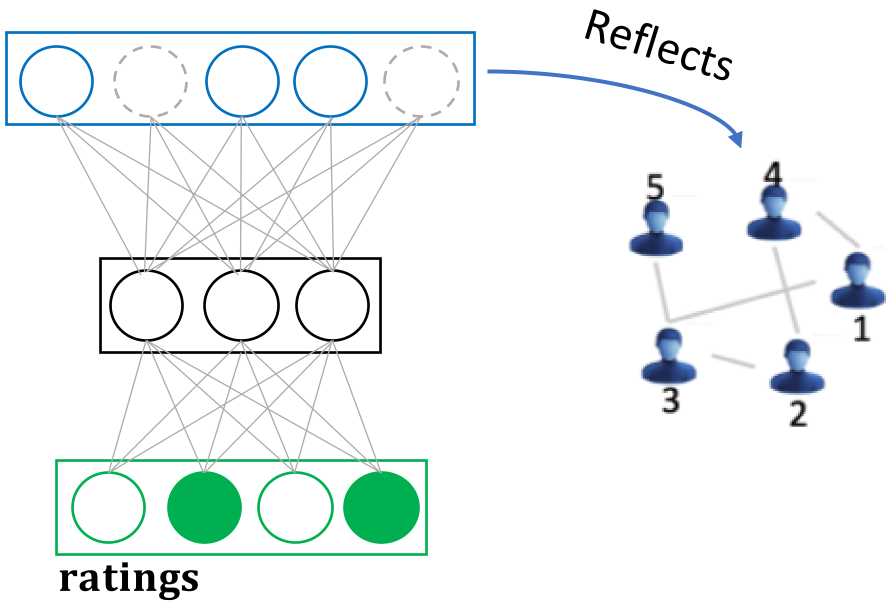
Set of friends of user i

$$+ \frac{\beta}{2}\sum_{i=1}^{m}\sum_{f \in \mathcal{F}^+(i)} Sim(i,f) \| U_i - U_f \|_F^2$$

- Individual-based regularization

Graph-Aware Architecture

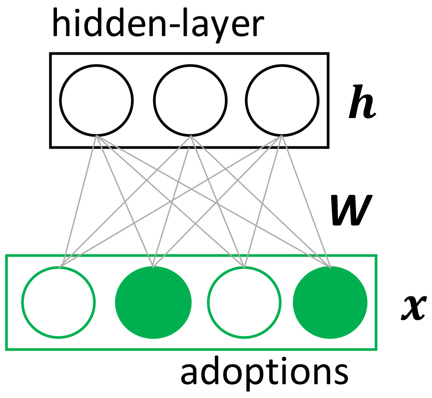
The graph structure is reflected in the collaborative model's architecture



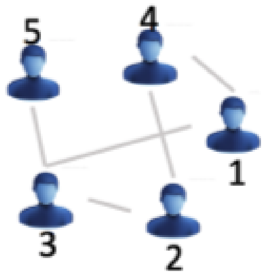
Social RBM-Deep

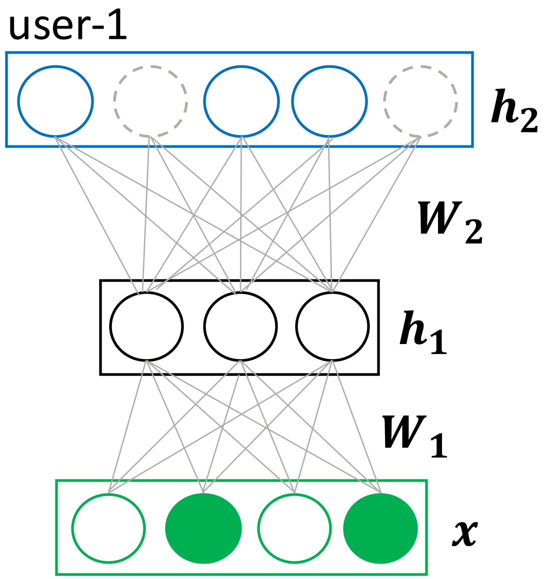
RBM: Restricted Boltzmann Machine

Nguyen, Trong T., and Hady W. Lauw. "Representation learning for homophilic preferences." *RecSys*. 2016.



Vanilla RBM



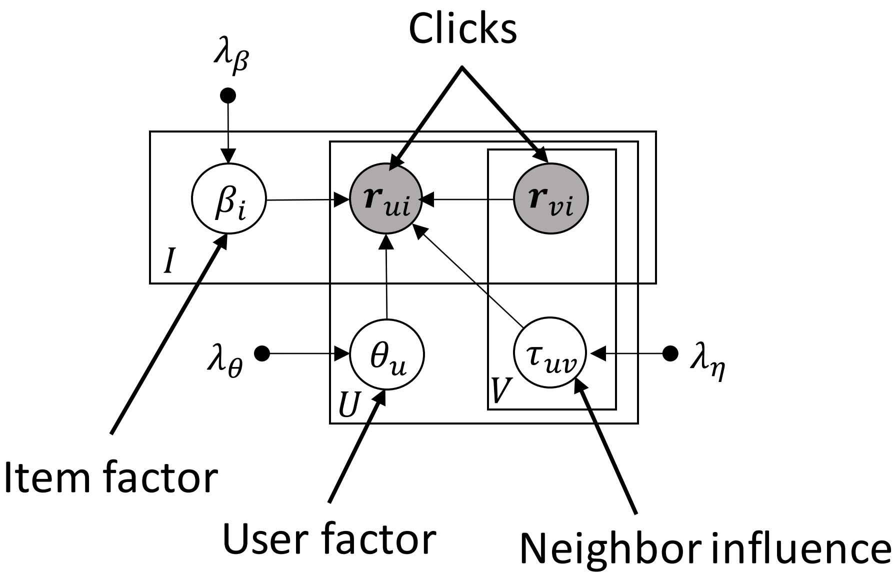


SocialRBM-Deep

- The top layer $\boldsymbol{h}_2$ has U hidden units, corresponding to U users
- User-1 has connections to users 3 and 4, thus the hidden units h_2^1, h_2^3 and h_2^4 are available for encoding user-1 adoptions

Social Poisson Factorization (SPF)

Chaney, A. J., Blei, D. M., & Eliassi-Rad, T. A probabilistic model for using social networks in personalized item recommendation. RecSys. 2015. (pp. 43-50).

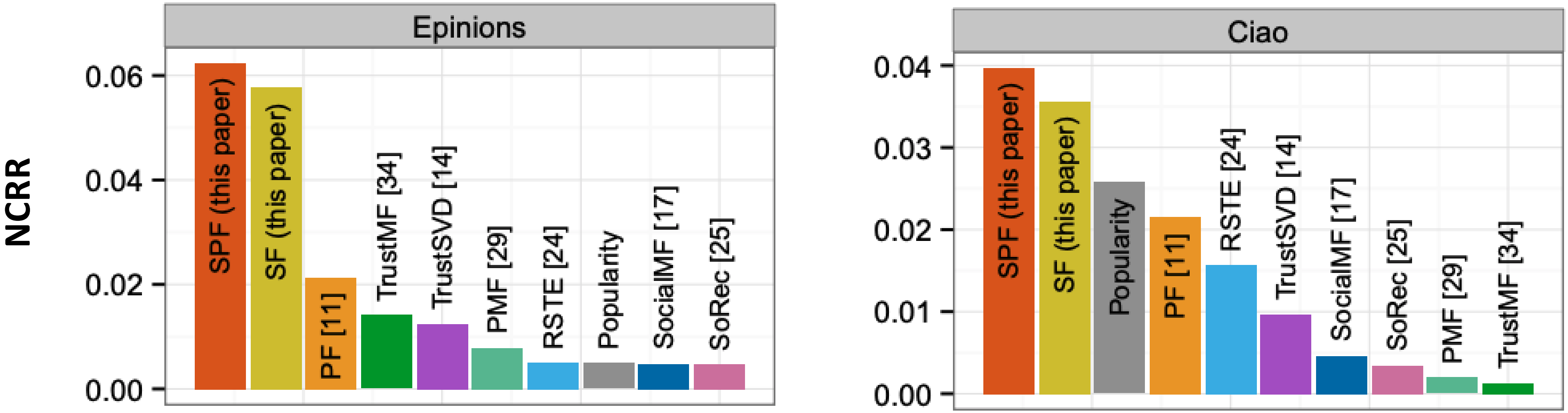


- Two signal driving each user's clicks
  - Latent preference for items captured by β_i and θ_i
  - The latent influence of her friends represented by τ_{uv} and r_{vi}
- The model is specified conditionally

$$r_{ui} \mid r_{-u,i} \sim \text{Poisson}\left(\theta_u^\top \beta_i + \sum_{v \in N(u)} \tau_{uv} r_{vi}\right)$$

- This is an improper model, but offers strong performances
- Scalable fitting to data with Variational Inference

SPF vs other Models on Epinions and Ciao

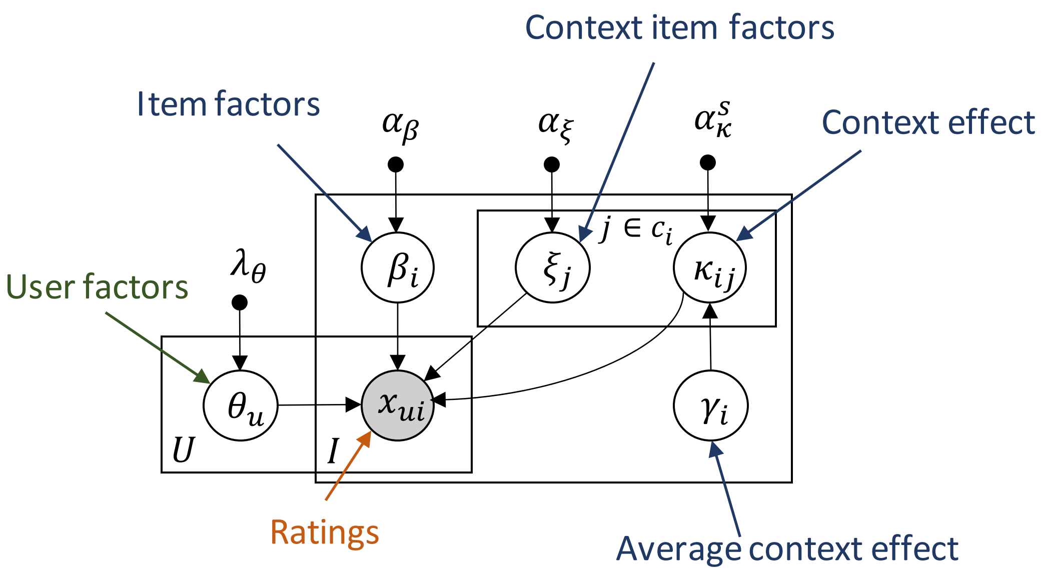


NCRR: normalized cumulative reciprocal rank

- The base model of the clicks is important
- Popularity-based recommendation is a strong baseline

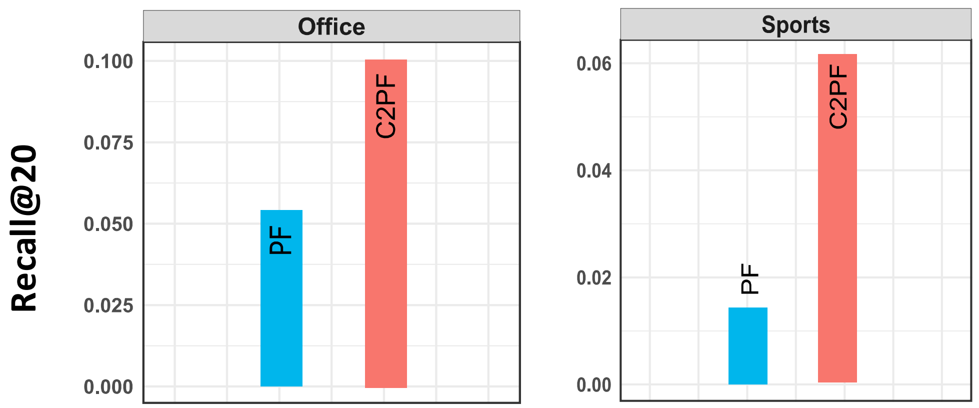
Collaborative Context PF (C2PF)

Salah, Aghiles, and Hady W. Lauw. "A bayesian latent variable model of user preferences with item context." IJCAI, 2018.



C2PF in plate notations

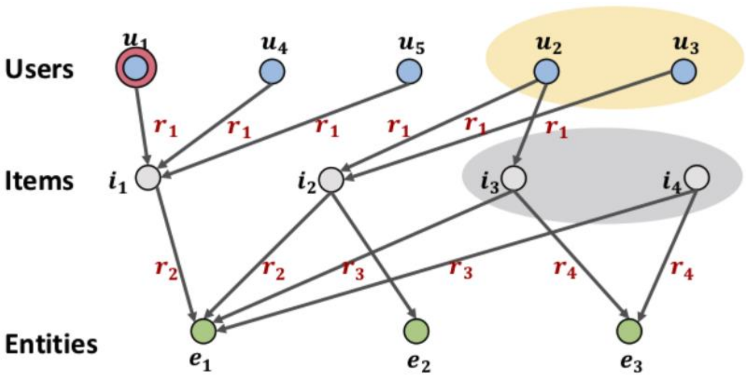
- Side-Information embedded into model's architecture
- Leaning driven by rating signals only
- Strong improvements over PF on Amazon datasets



C2PF vs PF on Amazon data

Knowledge Graph Attention Network (KGAT)

Wang, X., He, X., Cao, Y., Liu, M., & Chua, T. S. Kgat: Knowledge graph attention network for recommendation. *SIGKDD.* 2019 (pp. 950-958).



Collaborative Knowledge Graph (CKG)
User-Item graph ∪ Knowledge Graph

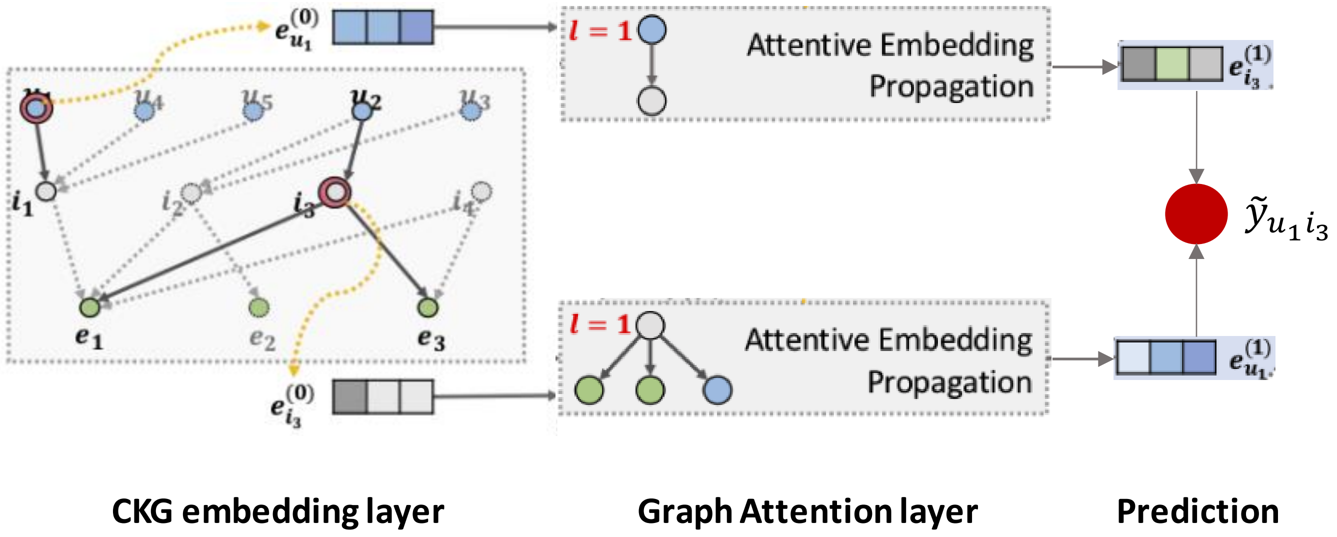


Illustration of KGAT model

- Embed entities and relations
- Graph Attention based representation of entities
- Predict user-item interactions

Which Family to Choose?

- Which family to choose may be **context-dependent**.

Use case 1. Assume that we have a good preference model that we seek to improve with some graph information, and that we cannot afford any extra computational overhead at inference time due to some real-time constraints. In this the case, a regularization-based approach maybe a good fit as it will not affect our original inference procedure.

Use case 2. We have access to a social network, and we are interested in making both item and friend recommendations. We can take the feature-based approach to model user-item and social data jointly.

Members of the graph-aware architecture family tend to offer better performance according to the literature

Cornac-Supported Graph-Based Models

- User network:
  - Social Recommendation using PMF (SoRec)
  - Social Bayesian Personalized Ranking (SBPR)
  - …
- Item network:
  - Collaborative Context Poisson Factorization (C2PF)
  - Probabilistic Collaborative Representation Learning (PCRL)
  - Matrix Co-Factorization (MCF)
  - …

More on Graph-related Recommendation

- Social Recommendation (Users)
  - Social regularized von Mises–Fisher mixture model for item recommendation. (*Data Min. Knowl. Discov, 2017)*
  - Deep Social Collaborative Filtering (RecSys'19)
  - Deep Adversarial Social Recommendation (IJCAI'19)
  - Graph Neural Network for Social Recommendation (WWW'19)
  - A Neural Influence Diffusion Model for Social Recommendation (SIGIR'19)
  - A Graph Neural Network Framework for Social Recommendations (TKDE'20)
  - Multi-channel Hypergraph Convolutional Network (WWW'*21)*
- Knowledge-graph-aware Recommendation (Items)
  - Leveraging Meta-path based Context for Top-N Recommendation with A Neural Co-Attention Model (KDD'18)
  - Knowledge Graph Convolutional Networks for Recommender Systems with Label Smoothness Regularization (KDD'19 and WWW'19)
  - RippleNet: Propagating User Preferences on the Knowledge Graph for Recommender Systems (CIKM'18)
  - Knowledge graph contrastive learning for recommendation (SIGIR'22)

Hands-on #2: Multi-Modal & Cross-Modal

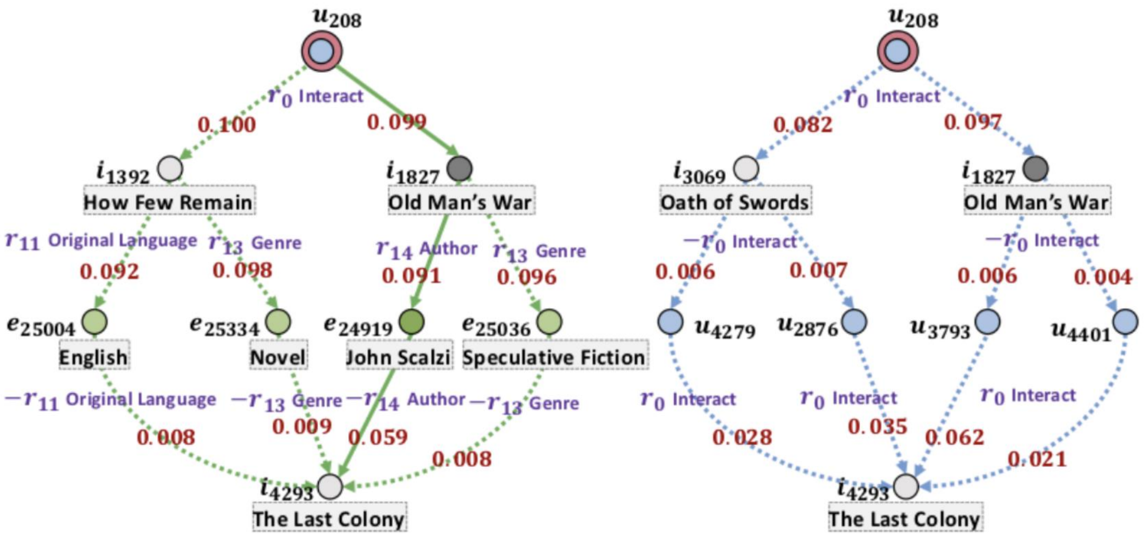
https://github.com/PreferredAI/tutorials/blob/master/multimodal-recsys/02_multimodality.ipynb

Explainability

Multi-Modal Explanations

- Explanation may come from different data sources, e.g., text, image, graph, etc. and may appear in many different forms.

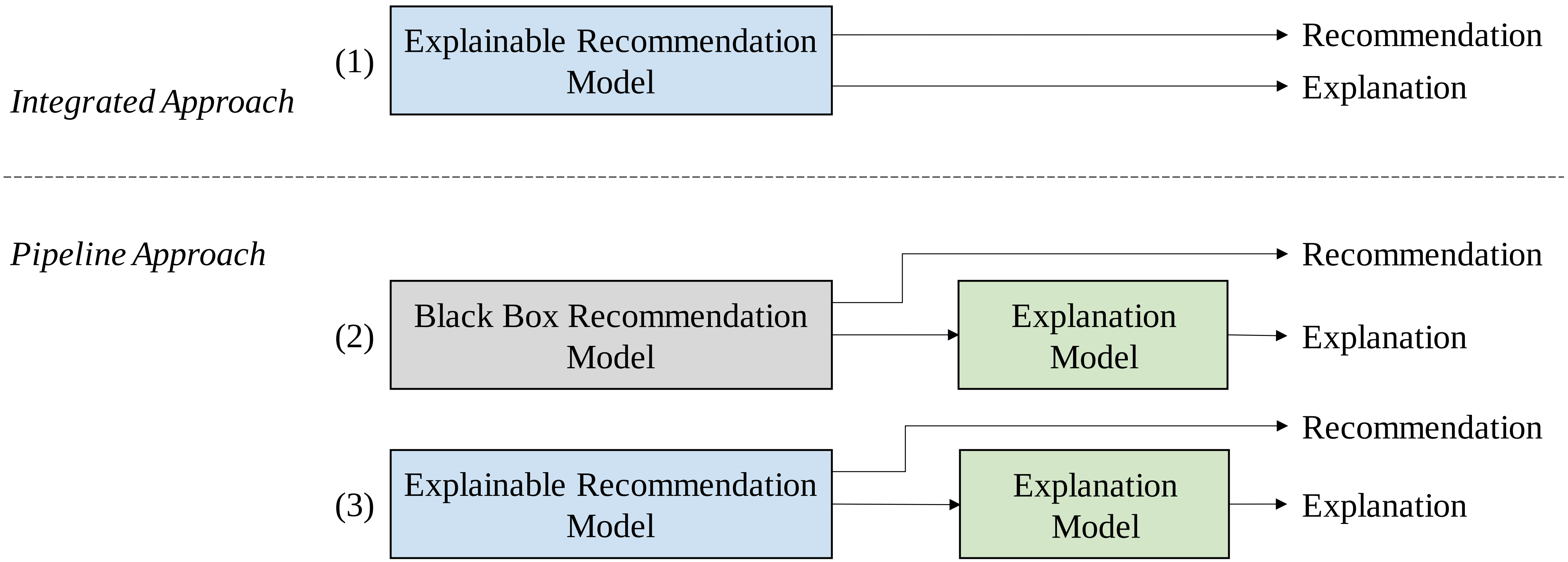
	Photo	Rating	Review
Philz Coffee		**4.6**	**i 've had a few times for the best breakfast sandwich .**
		4.0	the avocado toast was surprisingly good .
A16		**4.2**	**i was n't sure to try the pizza .**
		3.0	i think i might want to try their other pizzas which might be better tasting than their funghi .



Truong *et. al.*, 2019

Wang *et. al.*, 2019

Multi-Modal Explanations



Explanations rely on Additional Modalities

Text	**Graph**	**Image**
Review selection: NARRE, HRDR + Question selection: QuestER	KGAT	Visual Explanation: VECF
Template: EFM, MTER, ComparER	RippleNet	Visual Sentiment: VistaNet
Review synthesis: SEER	HAGERec	Image + generated review: MRG
Review generation: DAML		

References

- Zhang, Y., Lai, G., Zhang, M., Zhang, Y., Liu, Y., & Ma, S. (2014, July). Explicit factor models for explainable recommendation based on phrase-level sentiment analysis. In SIGIR'14 (pp. 83-92).
- Chong Chen, Min Zhang, Yiqun Liu, and Shaoping Ma. 2018. Neural Attentional Rating Regression with Review-level Explanations. In WWW'18
- Wang, N., Wang, H., Jia, Y., & Yin, Y. (2018, June). Explainable recommendation via multi-task learning in opinionated text data. In SIGIR'18.
- Donghua Liu, Jing Li, Bo Du, Jun Chang, and Rong Gao. 2019. DAML: Dual Attention Mutual Learning between Ratings and Reviews for Item Recommendation. In KDD '19
- Wang, X., He, X., Cao, Y., Liu, M., & Chua, T. S. (2019, July). Kgat: Knowledge graph attention network for recommendation. In SIGKDD'19.
- Wang, H., Zhang, F., Wang, J., Zhao, M., Li, W., Xie, X., & Guo, M. (2018, October). Ripplenet: Propagating user preferences on the knowledge graph for recommender systems. In CIKM (pp. 417-426).
- Chen, X., Chen, H., Xu, H., Zhang, Y., Cao, Y., Qin, Z., & Zha, H. (2019, July). Personalized fashion recommendation with visual explanations based on multimodal attention network: Towards visually explainable recommendation. In SIGIR'19 (pp. 765-774).
- Truong, Q. T., & Lauw, H. W. (2019, May). Multimodal review generation for recommender systems. In WWW'19 (pp. 1864-1874).
- Truong, Q. T., & Lauw, H. W. (2019, July). Vistanet: Visual aspect attention network for multimodal sentiment analysis. In AAAI'19.
- Yang, Z., & Dong, S. (2020). HAGERec: Hierarchical attention graph convolutional network incorporating knowledge graph for explainable recommendation. Knowledge-Based Systems, 204, 106194.
- Le, T. H., & Lauw, H. W. (2020). Synthesizing aspect-driven recommendation explanations from reviews. In IJCAI'20. **Distinguished Paper Award.**
- Liu, H., Wang, Y., Peng, Q., Wu, F., Gan, L., Pan, L., & Jiao, P. (2020). Hybrid neural recommendation with joint deep representation learning of ratings and reviews. Neurocomputing, 374, 77-85.
- Le, T. H., & Lauw, H. W. (2021). Explainable Recommendation with Comparative Constraints on Product Aspects. In WSDM'21.
- Le, T. H., & Lauw, H. W. (2022). Question-Attentive Review-Level Recommendation Explanation. In BigData'22.

Review-Level Explanation

Item: "Iron Man" on Amazon

Less useful

A

An Awesome Movie!
By Jokerz Wild on October 9, 2017
Format: Amazon Video | Verified Purchase

I love Iron Man!

rough preference of the consumer

B

Comic book characters... making millions of horrible movies these days.
By TylerVogt3329 on November 14, 2008
Format: DVD

You people these days consider this a good movie? Haha. Who in their right mind believes that a rich playboy can save the world from evil?
For good and REAL action check out WWE, ECW, or TNA.
For good classic Wrestling.. check out WCW and WWF.

talking about something else, not about the film

Useful

C

Good solid film
By M-M on July 30, 2013
Format: Amazon Video | Verified Purchase

It turned out to be entertaining and at the end I enjoyed the film. Good special effects, nice story line for "actions" and "comics". The protagonist (Tony Stark) looks natural: arrogant, brash, but at the same time clever, intelligent and ethic. The villain is a little bit overreacting, and annoying, as most of antagonists :) Overall that's a good movie.

8 people found this helpful. Was this review helpful to you? Yes No Report abuse

detailed information, more helpful for user's potential consumption

[Chen *et. al.*, 2018]

Review Selection: NARRE

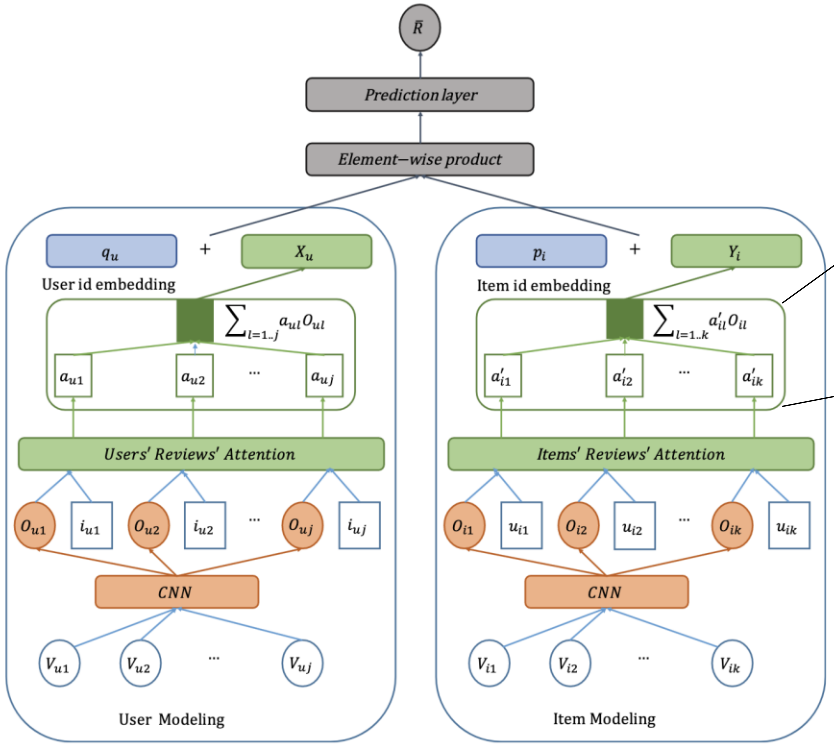
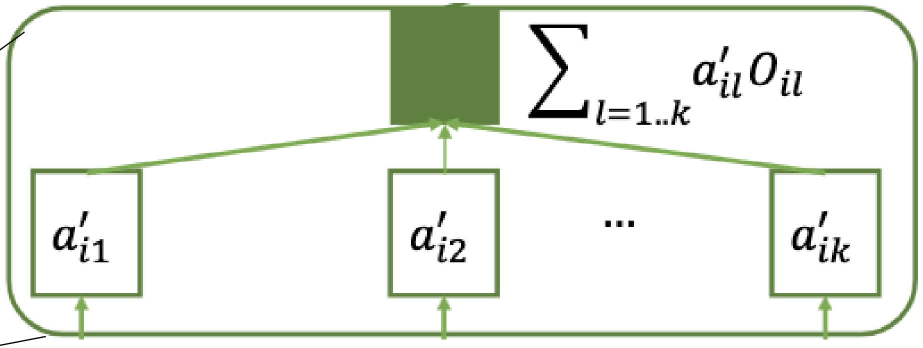
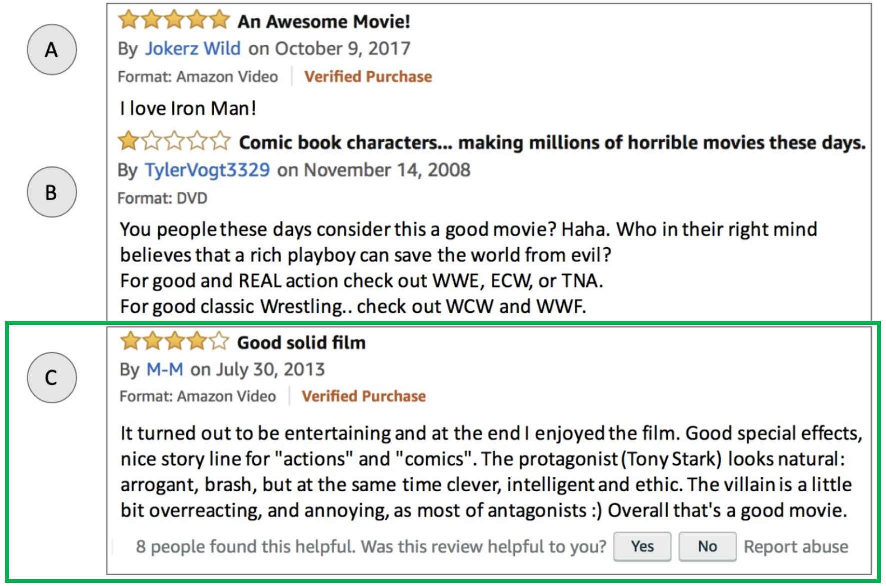


Figure 3: The neural network architecture of NARRE. Our attention model uses both IDs(i_{uj}, u_{jk}) and review content(O_{uj}, O_{jk}) to automatically assign weights to reviews.





Useful review

Chen *et. al.*, 2018. Neural Attentional Rating Regression with Review-level Explanations. In WWW'18

Review Selection: HRDR

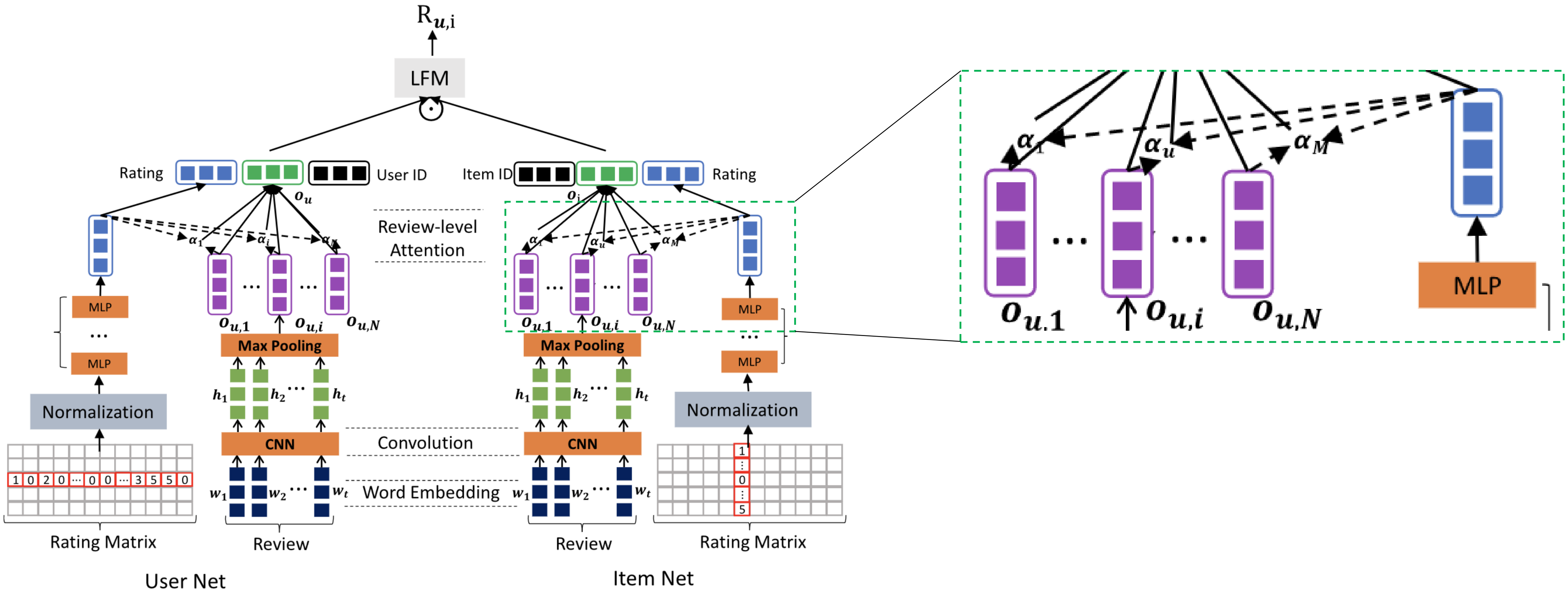


Fig. 2. Overview of HRDR: two parallel neural networks for users and items.

Liu *et. al.*, (2020). Hybrid neural recommendation with joint deep representation learning of ratings and reviews. In Neurocomputing

Question-Level Explanation

Le, T. H., & Lauw, H. W. (2022). Question-Attentive Review-Level Recommendation Explanation. In BigData'22.

Asin: B07P15K8Q7

Canon EOS Rebel T7 DSLR Camera Bundle with Canon EF-S 18-55mm f/3.5-5.6 is II Lens + 2pc SanDisk 32GB Memory Cards + Accessory Kit

Question voting

8 votes

Question: does this camera have wifi ablity>?

Answer: Ya

By Melaku D on May 14, 2019

5 votes

Question: Why do you sell a flash with this bundle that does not work with the camera?

Answer: It works fine, the flash is used on the bar mount that is also provided. The bar mounts to the tripod thread on the bottom of the Camera and then the Flash mounts to the bar. It is for a secondary flash, hence it senses your popup flash and goes off. Technically you can have the flash behind the subject that way the ba… see more

By Ultimaron on November 24, 2019

2 votes

Question: Can this camera send photos to smartphones using WiFi?

Answer: Yes. You have to set up the Wi-Fi functions, nickname your camera. I really recommend the Canon EOS Rebel T7/2000D book (for dummies) book as a companion translation to the manual.

By Wily Girl on June 26, 2019

2 votes

Question: does it have a port for an external microphone?

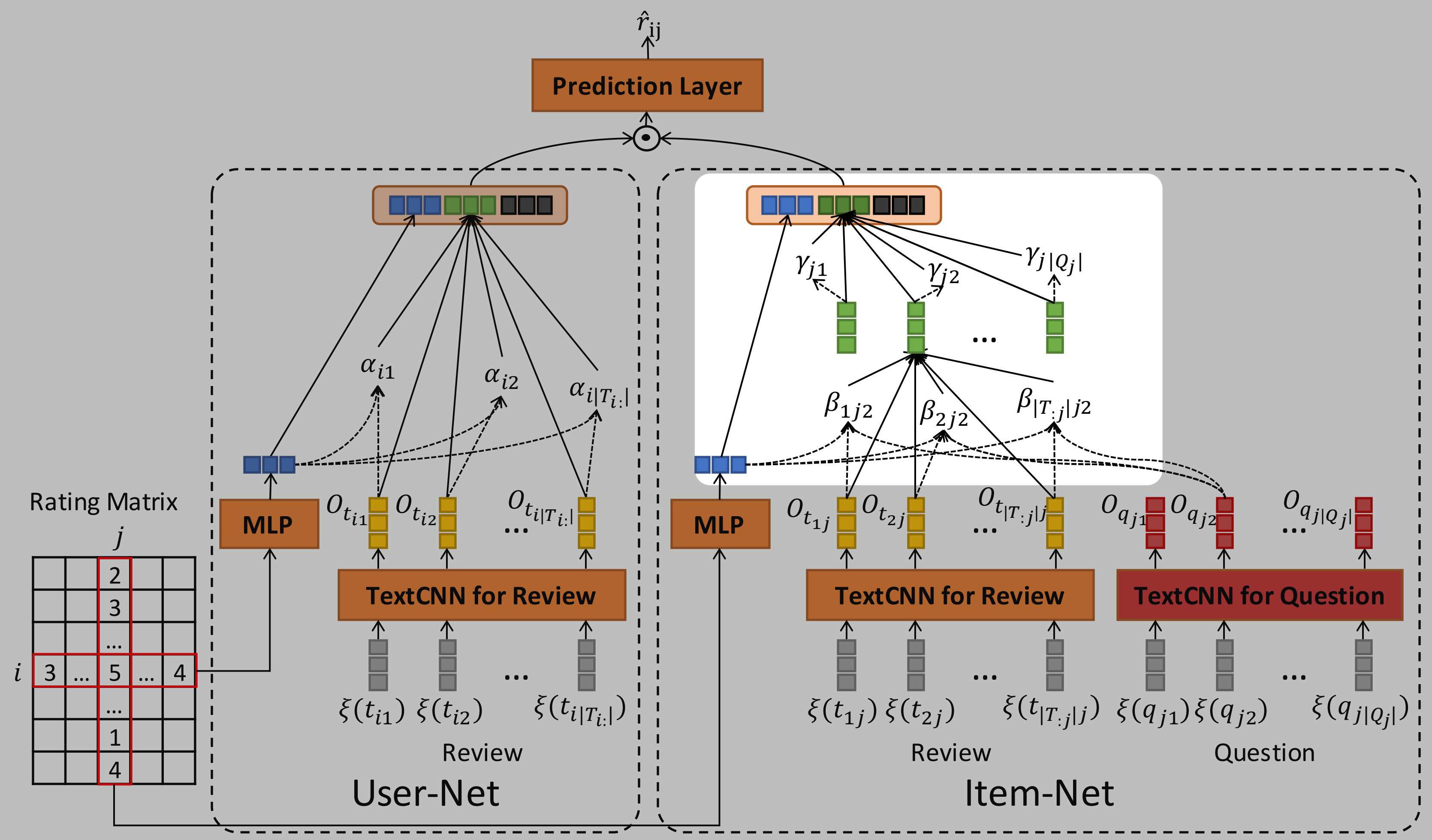
Answer: No, only the t7i

By Lane Wallen on January 20, 2021

See more answers (1)

See more answered questions (216)

Question-Attentive Review-Level Explanation for Neural Rating Regression



Le, T. H., & Lauw, H. W. (2022). Question-Attentive Review-Level Recommendation Explanation. In BigData'22.

Template Explanation: EFM

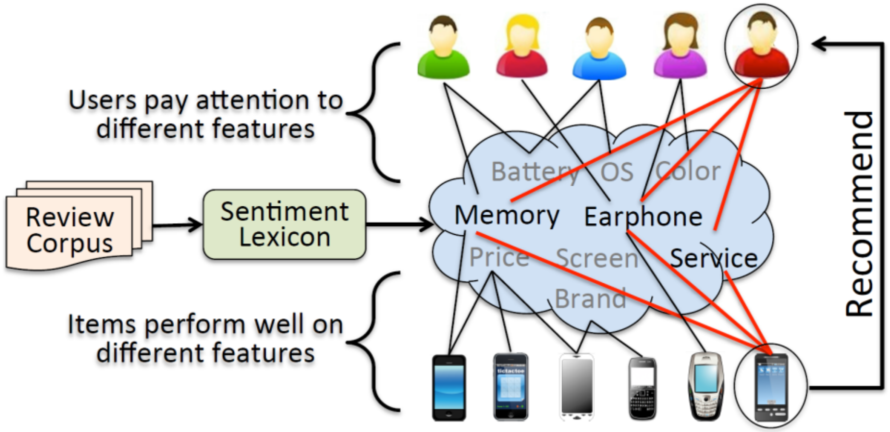


Figure 1: The product feature word and user opinion word pairs are extracted from user review corpus to construct the sentiment lexicon, and the feature word set further serves as the explicit feature space. An item would be recommended if it performs well on the features that a user cares.

`You might be interested in [feature], on which this product performs well.`
`You might be interested in [feature], on which this product performs poorly.`

Figure 2.7: Generating sentence explanations with template-based methods.

Zhang et al., "Explicit factor models for explainable recommendation based on phrase-level sentiment analysis", SIGIR 2014.

Extracting Aspects & Sentiments from Reviews

- Let $\mathcal{F}$ be the set of aspect words, e.g., screen, earphone
  - Let F be the number of aspect in this set, i.e., $\mathrm{F} = |\mathcal{F}|$

- Let O be the set of opinion words, e.g., perfect, good

- Let $S = \{-1, +1\}$ be a binary sentiment value

- A tuple (f, o, s) is extracted from a product review if the aspect $f \in \mathcal{F}$ appears with opinion phrase $\mathrm{o} \in O$ within the review with sentiment $s \in S$
  - an opinion phrase usually has an associated sentiment, e.g., *good* is +1
  - an opinion phrase can be reversed by negation, e.g., *not good* is -1

Extracting Aspects & Sentiments from Reviews

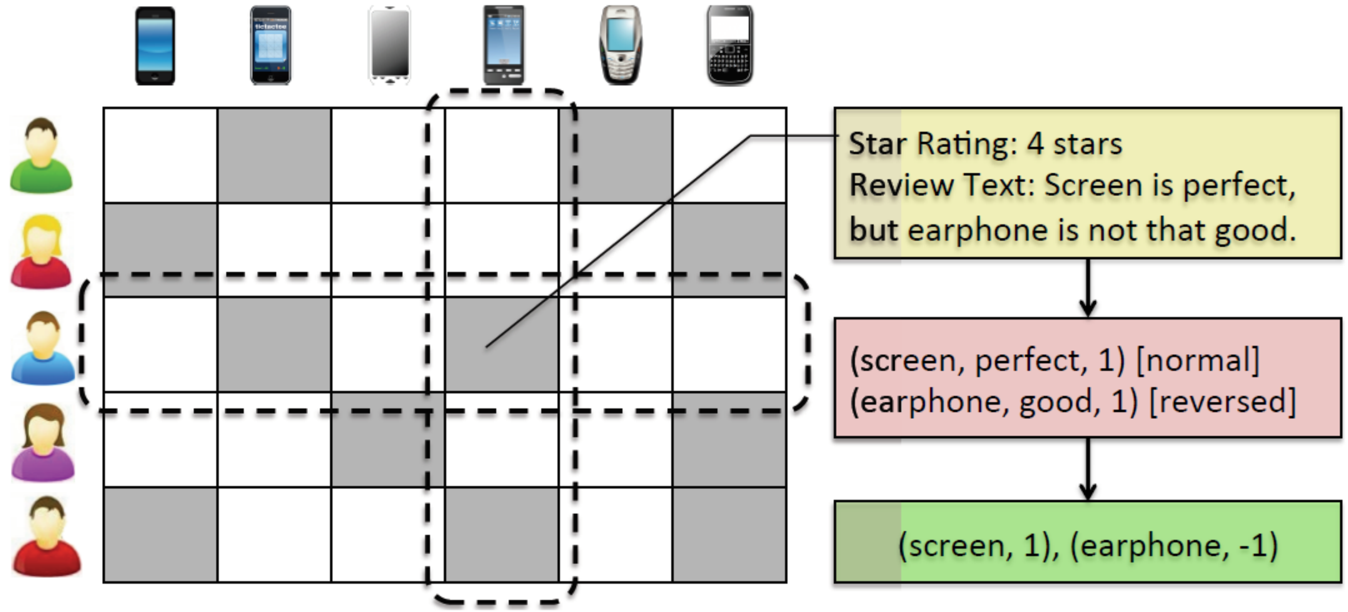


Figure 2: An example of user-item review matrix, where each shaded block is a review made by a user towards an item; the entries included in the review are extracted, and further transformed to feature scores while considering the negation words.

User-Aspect Attention Matrix

- Let $\boldsymbol{X} \in \mathbb{R}^{N \times F}$ be a sparse aspect matrix for N users and F aspects
- $x_{if} \in \boldsymbol{X}$ indicates the degree of attention by user i on aspect f

$$x_{if} = \begin{cases} 0, \text{if user } i \text{ never mentions aspect } f \\ 1 + (maxvalue - 1)\left(\dfrac{2}{1 + \exp\{-t_{if}\}} - 1\right), \text{otherwise} \end{cases}$$

  - the t_{if} is the frequency with which user i mentions aspect f across all her reviews
  - the above rescales t_{if} to the range of 1 to $maxvalue$ (highest rating value, e.g., 5)
  - If $t_{if} \to \infty$, user i mentions aspect f frequently, then $x_{if} \to maxvalue$
  - If $t_{if} \to 0$, user i mentions aspect f rarely, then $x_{if} \to 1$

User-Aspect Attention Matrix

- The observations x_{if} are non-negative
- Since $\boldsymbol{X}$ is sparse, predict the missing attention scores not in $\boldsymbol{X}$ through non-negative matrix factorization
  - the L latent factors are constrained to be non-negative
  - For each user i, a L-dimensional vector $\tilde{\boldsymbol{u}}_i \in \mathbb{R}_+^L$. For all users, collectively $\tilde{\boldsymbol{U}} \in \mathbb{R}_+^{N \times L}$.
  - For each aspect f, a L-dimensional vector $\boldsymbol{z}_f \in \mathbb{R}_+^L$. For all items, collectively $\boldsymbol{Z} \in \mathbb{R}_+^{F \times L}$.
- Loss function:

$$\mathcal{L}_{\boldsymbol{X}}(\tilde{\boldsymbol{U}}, \boldsymbol{Z}) = \frac{1}{2} \sum_{x_{if} \in \boldsymbol{X}} \left(x_{if} - \sum_{l=1}^{L} \tilde{u}_{il} \cdot z_{fl} \right)^2 + \frac{\lambda}{2} \sum_{i=1}^{N} \|\tilde{\boldsymbol{u}}_i\|^2 + \frac{\lambda}{2} \sum_{f=1}^{F} \|\boldsymbol{z}_f\|^2$$

such that $\forall_{i,l}\ \tilde{u}_{il} \geq 0, \forall_{f,l}\ z_{fl} \geq 0$

PREFERRED.AI

Item-Aspect Quality Matrix

- Let $\boldsymbol{Y} \in \mathbb{R}^{M \times F}$ be a sparse aspect matrix for M items and F aspects

- $y_{jf} \in \boldsymbol{Y}$ indicates the quality of item j on aspects f

$$y_{jf} = \begin{cases} 0, \text{if item } j \text{ was never reviewed on aspect } f \\ 1 + (maxvalue - 1)\left(\dfrac{1}{1 + \exp\{-s_{jf}\}}\right), \text{otherwise} \end{cases}$$

  - s_{jf} is the sum of sentiment values with which item j has been mentioned with regards to aspect f across all its reviews
  - the above rescales s_{jf} to the range of 1 to $maxvalue$ (highest rating value, e.g., 5)
  - If $s_{jf} \rightarrow \infty$, item j has been commented on aspect f positively, then $y_{jf} \rightarrow maxvalue$
  - If $s_{jf} \rightarrow -\infty$, item j has been commented on aspect f negatively, then $y_{jf} \rightarrow 1$

Item-Aspect Quality Matrix

- The observations y_{jf} are non-negative
- Since $\boldsymbol{Y}$ is sparse, predict the missing quality scores not in $\boldsymbol{Y}$ through non-negative matrix factorization
  - the L latent factors are constrained to be non-negative
  - For each item j, a L-dimensional vector $\tilde{\boldsymbol{v}}_j \in \mathbb{R}_+^L$. For all users, collectively $\tilde{\boldsymbol{V}} \in \mathbb{R}_+^{M \times L}$.
  - For each aspect f, a L-dimensional vector $\boldsymbol{z}_f \in \mathbb{R}_+^L$. For all aspects, collectively $\boldsymbol{Z} \in \mathbb{R}_+^{F \times L}$.
- Loss function:

$$\mathcal{L}_{\boldsymbol{Y}}(\tilde{\boldsymbol{V}}, \boldsymbol{Z}) = \frac{1}{2} \sum_{y_{jf} \in \boldsymbol{Y}} \left(y_{jf} - \sum_{l=1}^{L} \tilde{v}_{jl} \cdot z_{fl} \right)^2 + \frac{\lambda}{2} \sum_{j=1}^{M} \left\| \tilde{\boldsymbol{v}}_j \right\|^2 + \frac{\lambda}{2} \sum_{f=1}^{F} \left\| \boldsymbol{z}_f \right\|^2$$

$$\text{such that } \forall_{j,l}\ \tilde{v}_{jl} \geq 0, \forall_{f,l}\ z_{fl} \geq 0$$

Explicit Factor Model (EFM)

- Joining three non-negative matrix factorizations:
  - Due to rating matrix $\boldsymbol{R}$
  - Due to user-aspect attention matrix $\boldsymbol{X}$
  - Due to item-aspect quality matrix $\boldsymbol{Y}$

- Into one combined loss function

$$\mathcal{L}(\boldsymbol{U},\boldsymbol{V},\widetilde{\boldsymbol{U}},\widetilde{\boldsymbol{V}},\boldsymbol{Z})$$

$$= \sum_{r_{ij}\in\boldsymbol{R}}\left(r_{ij} - \left(\sum_{k=1}^{K} u_{ik}\cdot v_{jk} + \sum_{l=1}^{L}\tilde{u}_{il}\cdot\tilde{v}_{jl}\right)\right)^2 + \lambda_x\sum_{x_{if}\in\boldsymbol{X}}\left(x_{if} - \sum_{l=1}^{L}\tilde{u}_{il}\cdot z_{fl}\right)^2 + \lambda_y\sum_{y_{jf}\in\boldsymbol{Y}}\left(y_{jf} - \sum_{l=1}^{L}\tilde{v}_{jl}\cdot z_{fl}\right)^2$$

$$+ \lambda_h\sum_{i=1}^{N}\|\boldsymbol{u}_i\|^2 + \lambda_h\sum_{j=1}^{M}\|\boldsymbol{v}_j\|^2 + \lambda_u\sum_{i=1}^{N}\|\tilde{\boldsymbol{u}}_i\|^2 + \lambda_u\sum_{j=1}^{M}\|\tilde{\boldsymbol{v}}_j\|^2 + \lambda_v\sum_{f=1}^{F}\|\boldsymbol{z}_f\|^2$$

such that $\forall_{i,k}\, u_{ik} \geq 0, \forall_{j,k}\, v_{jk} \geq 0, \forall_{i,l}\, \tilde{u}_{il} \geq 0, \forall_{j,l}\, \tilde{v}_{jl} \geq 0, \forall_{f,l}\, z_{fl} \geq 0$

Prediction

- To identify aspects that a user i is most interested in, we take the k aspects C_i with the highest predicted $\hat{x}_{if} = \tilde{\boldsymbol{u}}_i^T \boldsymbol{z}_f$ attention scores among $\{\hat{x}_{if}\}_{f=1}^{F}$
- The predicted ranking score by a user i on item j is given by:

$$\alpha \cdot \frac{\sum_{f \in C_i} \hat{x}_{if} \cdot \hat{y}_{jf}}{k \cdot maxvalue} + (1 - \alpha) \cdot \hat{r}_{ij}$$

where:

- $\hat{x}_{if} = \tilde{\boldsymbol{u}}_i^T \boldsymbol{z}_f$ is the predicted user attention on aspect f
- $\hat{y}_{jf} = \tilde{\boldsymbol{v}}_j^T \boldsymbol{z}_f$ is the predicted item quality on aspect f
- $\hat{r}_{ij} = \boldsymbol{u}_i^T \boldsymbol{v}_j + \tilde{\boldsymbol{u}}_i^T \tilde{\boldsymbol{v}}_j$ is the predicted rating

Evaluative vs. Comparative Explanation

Time

Reference item: B00AI5V3CQ
The OontZ Angle Ultra Portable Wireless Bluetooth Speaker

Recommendation item: B00F6AVFK8
The Oontz XL - Cambridge SoundWorks Most Powerful Portable, Wireless, Bluetooth Speaker

`Product B00F6AVFK8 is better at` **`quality`** `than B00AI5V3CQ. But worse at` **`sound.`**

Le, T. H., & Lauw, H. W. (2021). Explainable Recommendation with Comparative Constraints on Product Aspects. In WSDM'21.

Comparative Explainable Recommendation

Reference item j	Recommendation j'
	

ComparER

ASPECT	q_{jk}		$q_{j'k}$		$\hat{q}_{jk}$		$\hat{q}_{j'k}$
quality	4	<	5		4.3	<	4.8
battery	3	<	4		3.5	<	4.2
wireless	3		3		3		3
sound	5	>	4		4.2	>	3.8
…	-		-		-		-

Product y is dominant by product x, if x is at least as good as y in all aspects and better in at least one aspect

Le, T. H., & Lauw, H. W. (2021). Explainable Recommendation with Comparative Constraints on Product Aspects. In WSDM'21.

Aspect-Level Quality

Subjective

design
color
...

Objective

battery life
storage
...

Objective Aspect-Level Quality

Y				
quality				
battery				
wireless				
sound				
…				

$$Y_{jk} = \begin{cases} 0, \text{if aspect } k \text{ is not discussed in the reviews of } j \\ 1 + \dfrac{N-1}{1+e^{-s'_{jk}}}, \text{otherwise} \end{cases}$$

$$L_{\mathrm{ComparER}_{obj}} = -\sum_{(j,j')\in\bigcup_{i\in U} S_i} \overbrace{\left(1+\ln\left(c_{jj'}\right)\right)}^{\text{scaling factor}} \sum_{\{k \mid y_{jk} < y_{j'k}\}} \ln\sigma\left(\hat{y}_{j'k} - \hat{y}_{jk}\right)$$

ComparER: Generating Explanation

[recommended item] is better at [an aspect] than

$$\hat{y}_{jk} < \hat{y}_{j'k}$$

[reference item], but worse at [another aspect]

$$\hat{y}_{jk'} > \hat{y}_{j'k'}$$

Synthesizing Explanation for Explainable Recommendation

Inputs

Review sentences

A great mouse

This is a good mouse

The size is good

Very good mouse

The mouse is very comfortable and nice looking

The side button is not handy

The scroll design is bad

…

Demanded aspects

2 x mouse (4.7)

1 x scroll (3.5)

1 x size (4.9)

Predicted aspect sentiment scores

Output

The mouse is very comfortable and nice looking

great

This is a ~~good~~ mouse

acceptable

The scroll design is ~~bad~~

perfect

The size is ~~good~~

Le, T. H., & Lauw, H. W. (2020). Synthesizing Aspect-Driven Recommendation Explanations from Reviews. In *IJCAI 2020.* **Distinguished Paper Award**

Selecting Representative Sentences

- Representativeness: a sentence s represents another sentence s' by a cost $\delta_{ss'}$

$$\mathrm{r\_cost}(\tau) = \sum_{s \in \tau} \sum_{s' \in S_j} \delta_{ss'} \cdot \Gamma_{ss'}$$

Solution set

s_1: A great mouse

s_2: This is a great mouse

s_3: Very good mouse

s_4: The mouse is very comfortable and nice looking

…

$\delta_{ss'}$	s_1	s_2	s_3	s_4	…	
s_1	-	0.20	0.40	0.44	…	=1.04
s_2	0.20	-	0.44	0.39	…	=1.03
s_3	0.40	0.44	-	0.37	…	=1.21
s_4	0.44	0.39	0.37	-	…	=1.20
…	…	…	…	…	…	

Selecting Coherence Sentences

- Intuitively, a document was written by fewer authors would be more coherent than by many
- Sentences from a review $\mathrm{t}_{\mathrm{i}'}$ will associated with a cost $\theta_{\mathrm{i}'}$

$$\mathrm{c\_cost}(\tau) = \sum_{\mathrm{t}_{\mathrm{i}'\mathrm{j}} \in \mathrm{T}_{\mathrm{j}}} \theta_{\mathrm{i}'} \cdot \zeta_{\mathrm{i}'}$$

SEER: Optimization Objective

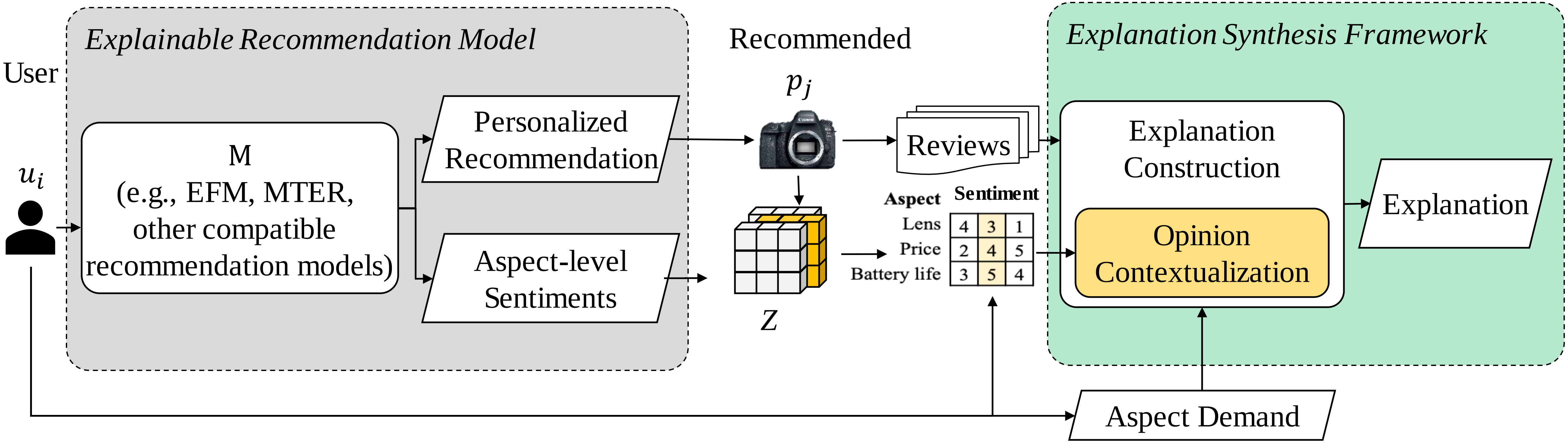
- Overall Cost:

$$\mathrm{cost}(\tau) = \mathrm{c\_cost}(\tau) + \mathrm{r\_cost}(\tau)$$

- Inherent trade off between c_cost and r_cost
- Optimize the overall cost is an NP-hard problem

- Enhancing template-based (e.g., EFM explanation) closer to natural language → SEER
  - E.g., The mouse is very comfortable and nice looking. This is a great mouse. The size is perfect.

Overall Architecture of SEER



Visually Explainable Recommendation

Figure 1: Example of users with different fine-grained visual preferences. User reviews may have partial correspondences to the fashion image. The pink italic and green bold fonts indicate the review information that can and cannot be aligned with some specific image regions.

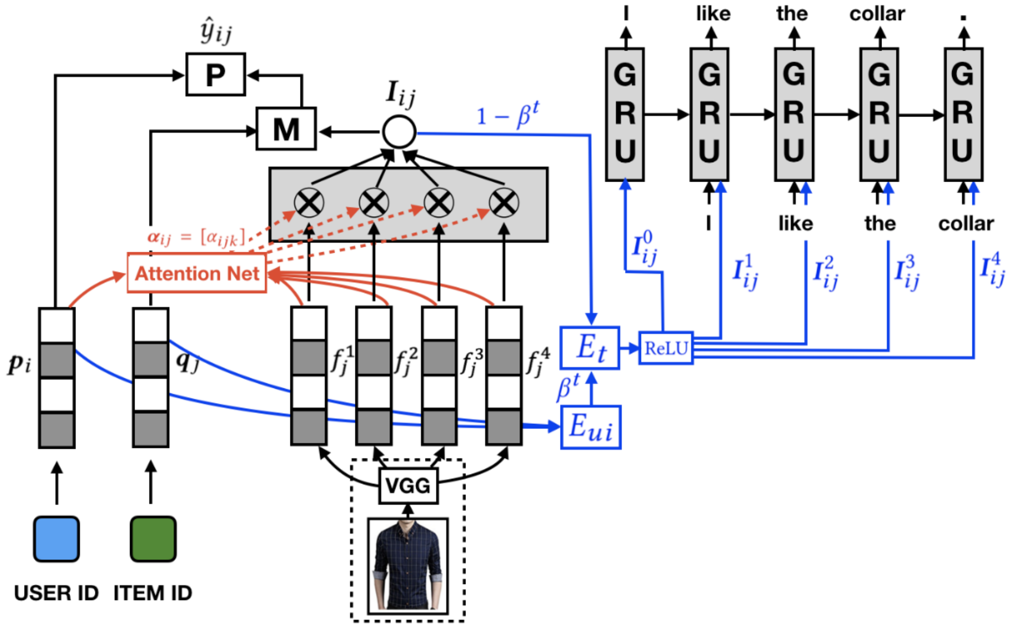
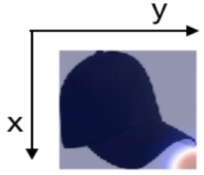
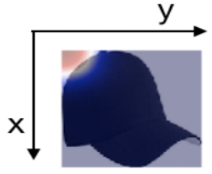
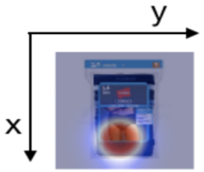
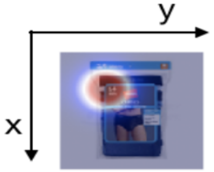
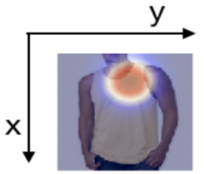
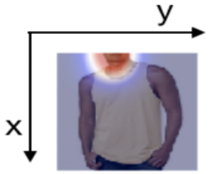


Figure 2: The proposed VECF model. The red lines indicate the attention mechanism designed for fashion image modeling. The blue lines highlight the modeling of user reviews.

Chen et. al., 2019 "Personalized Fashion Recommendation with Visual Explanations based on Multimodal Attention Network: Towards Visually Explainable Recommendation". SIGIR 2019

Visually Explainable Recommendation

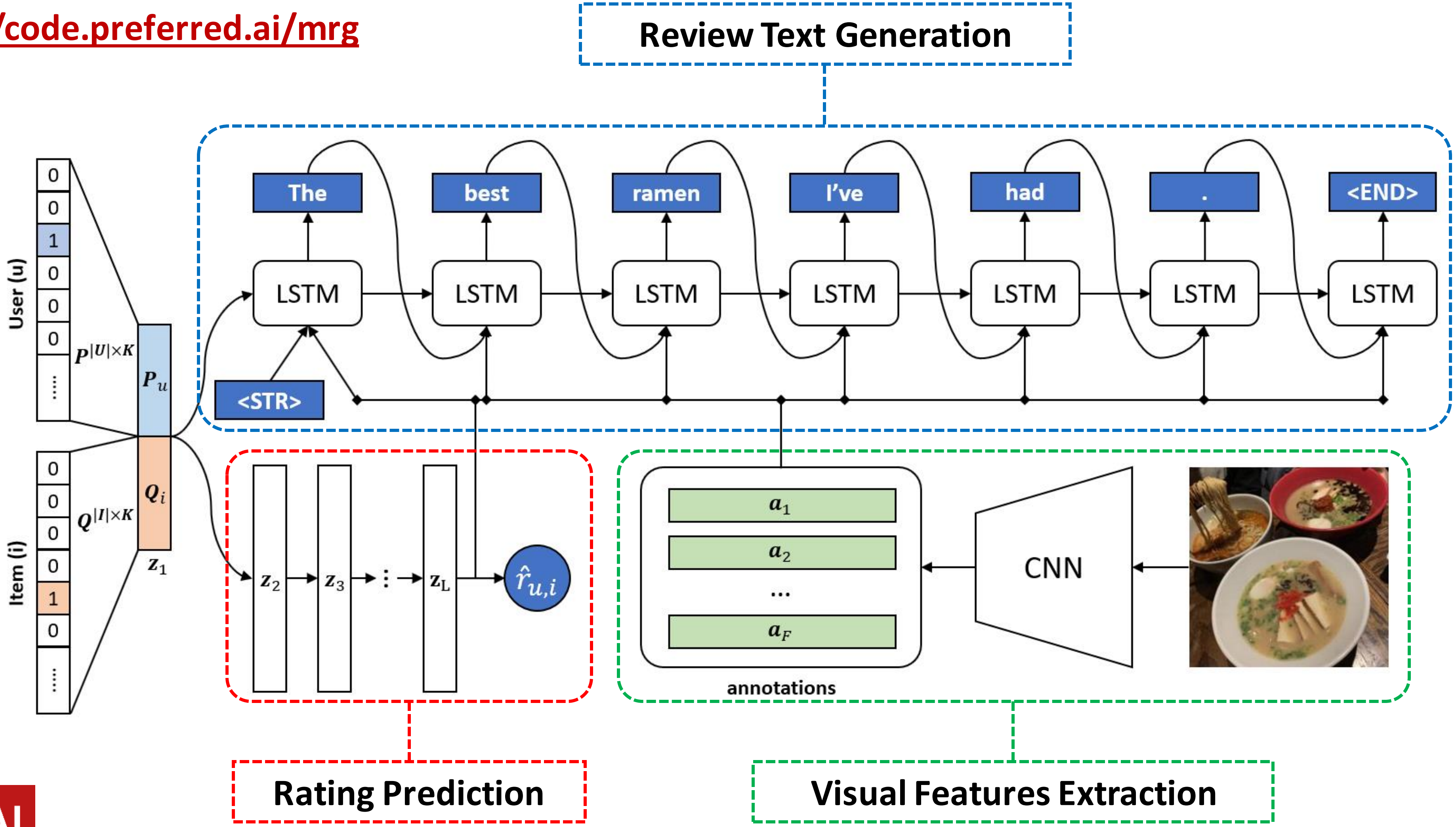
	Target Item	Textual Review	Visual Explanation	
			VECF(-rev)	VECF
4		The cap, which is made of a fairly heavy fabric, makes the head feel hot when worn for several hours in a warm gym or outside on a warm day. I, therefore, tend to wear it only when it is cold outside . -bi	 [x=13, y=13]	 [x=1, y=1]
5		These are comfortable and are a great value. I like the waist band and they are so so so (more words) comfortable....; -)-bi	 [x=10, y=6]	 [x=4, y=4]
6		The fabric is amazingly soft and the fit is perfect. I own several items from next level and will continue to add to my collection with different colors and styles. Amazing company, Amazing product.	 [x=3, y=7]	 [x=1, y=5]

Chen et. al., 2019 “Personalized Fashion Recommendation with Visual Explanations based on Multimodal Attention Network: Towards Visually Explainable Recommendation”. SIGIR 2019

Multimodal Review Generation

Truong and Lauw, “Multimodal Review Generation for Recommender Systems", WWW 2019.

https://code.preferred.ai/mrg



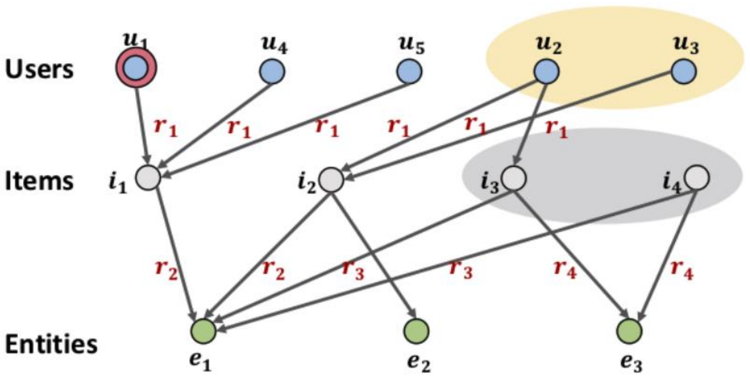
Multimodal Review Generation

Truong and Lauw, "Multimodal Review Generation for Recommender Systems", WWW 2019.

	Photo	Rating	Review
Ellen "FuZe" Z.		**4.5**	**the clam chowder was good .**
		5.0	best clam chowder i 've ever had .
Young Y.		**3.4**	**the clam chowder was a bit too salty .**
		3.0	the boston clam chowder was pretty salty and i 've had lots of clam chowder before .

Knowledge Graph Attention Network (KGAT)

Wang, X., He, X., Cao, Y., Liu, M., & Chua, T. S. Kgat: Knowledge graph attention network for recommendation. *SIGKDD.* 2019 (pp. 950-958).



Collaborative Knowledge Graph (CKG)
User-Item graph ∪ Knowledge Graph

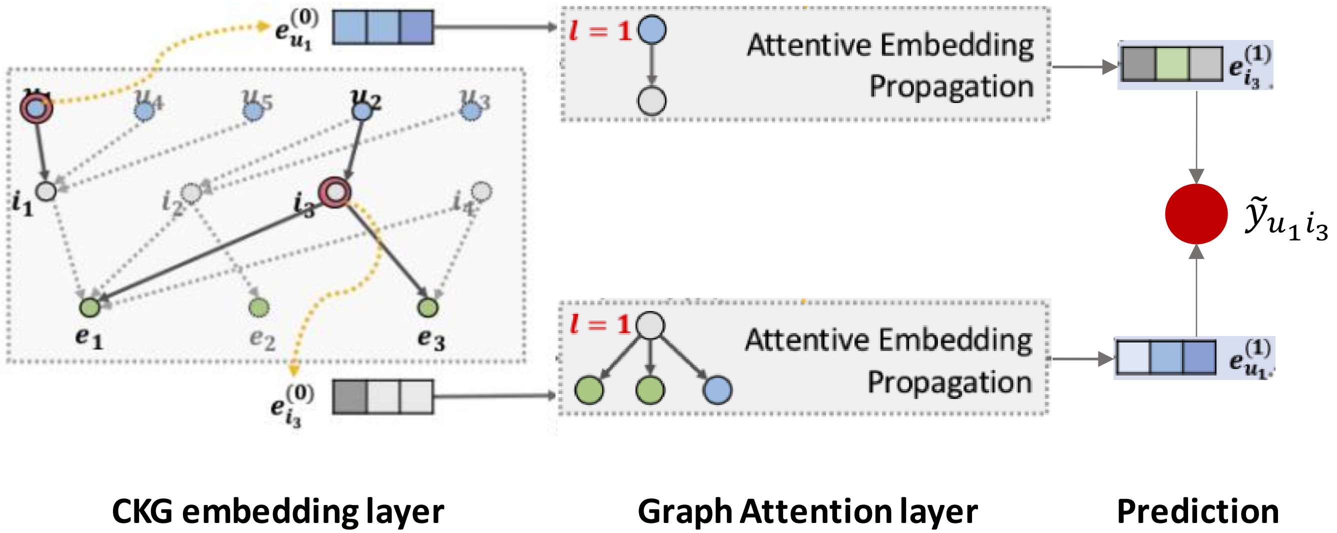
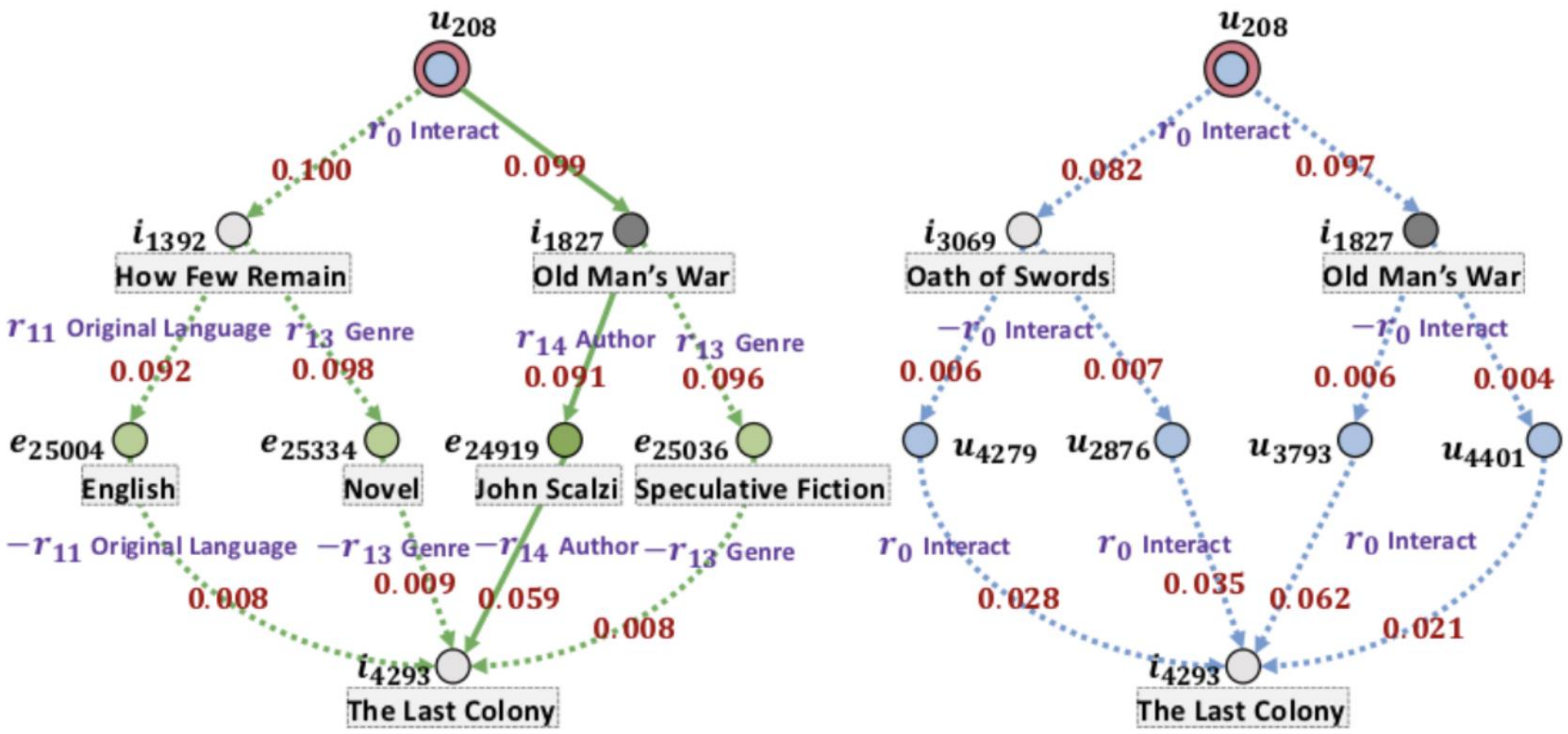


Illustration of KGAT model

- Embed entities and relations
- Graph Attention based representation of entities
- Predict user-item interactions

KGAT Explanations



Wang *et. al.*, 2019

Hands-on #3: Explainability

https://github.com/PreferredAI/tutorials/blob/master/multimodal-www23/03_explainability.ipynb

Future Directions

Leverage Generative Models

A who likes the Steve Jobs biography book receive the following recommendation list. Can you explain it to her in a very concise way and say if the recommendation list makes sense.
Recommendation list:

- Who is bill gates
- Inside apple
- Tim Cook: The Genius Who Took Apple to the Next Level
- Business-Do: The Way to Successful Leadership
- Datapoint: The Lost Story of the Texans Who Invented the Personal Computer Revolution

12:31

Rakuten Viber AI Chat

Powered by OpenAi

AI Chat & Create 🤖

The recommendation list consists of books that are related to technology, business, and leadership, which might interest someone who read the Steve Jobs biography. The books on this list are biographies of influential people and offer insights into the corporate culture and strategies of Apple. Additionally, there are books that provide practical advice on leadership and historical accounts of personal computing that could provide a broader understanding of the technology industry's history. Overall, the recommendation list seems relevant and sensible for someone looking to explore more books on the topic of technology, business, and leadership.

12:31

Future Directions

Leverage Generative Models

A who likes the Steve Jobs biography book receive the following recommendation list. Can you explain it to her in a very concise way and say if the recommendation list makes sense.
Recommendation list:
- Contemporary Color
- Unsettled Account: The Evolution of Banking in the Industrialized World Since
- Beyond Surface Appeal: Literalism, Sensibilities, and Constituencies in the Work of James Carpenter
- Designer Babies
- Performance Practice of the Instrumental-Vocal Works of Johann Sebastian Bach
- Social Economics: Market Behavior in a Social Environment

12:44

Rakuten Viber AI Chat
Powered by OpenAi

AI Chat & Create 🤖

The recommendation list includes books that are seemingly unrelated to the topic of technology, business, or leadership, making it not very relevant to someone who likes the Steve Jobs biography book. It may not make sense for someone looking for books that complement their interests in technology, innovation, and entrepreneurship.

12:45