# Accurate and Explainable Recommendation via Review Rationalization

Advisor : Jia-Ling, Koh

Speaker : Hsiao-Ting Huang

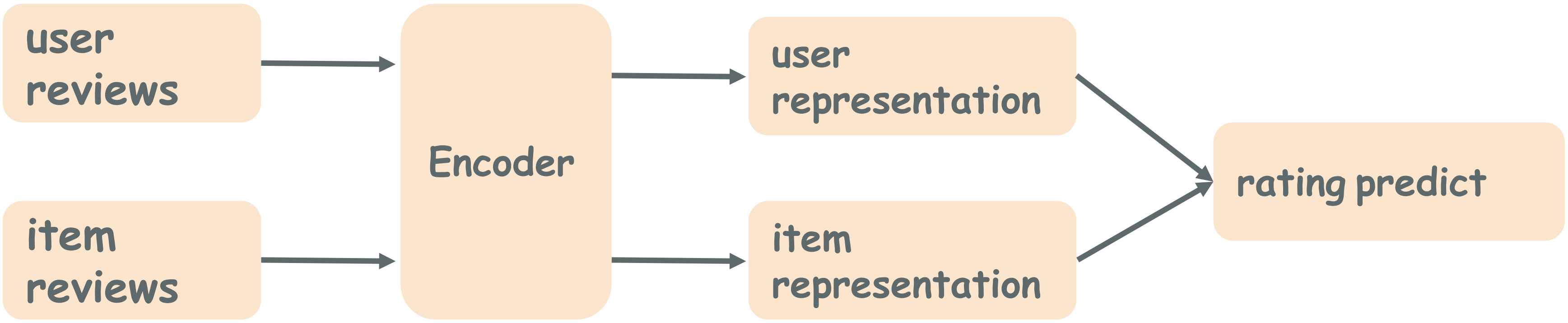
Source : WWW'22

Date : 2022/10/25

# Outline

- Introduction
- Motivation
- Method
- Experiment
- Conclusion

# Introduction ： Review-based recommendation



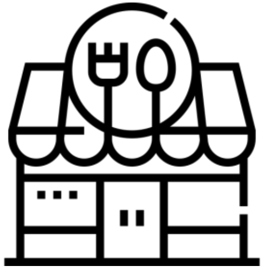
# Motivation

Noodles are truly great! The meat sauce and lamb are my favorites!

This is a perfect to-go place near the freeway for a work day including weekends! :-)

taste>convenience

rating predict : 4.3



# Input

- user u
- reviews of u : $W_{u,*|\sim tu,v} = \{W_{u,v1}, \ldots, W_{u,\sim v}\}$
- user u
- reviews of i : $W_{*,v\ |\sim tu,v} = \{W_{u1,v}, \ldots, W_{\sim u,v}\}$

# Output

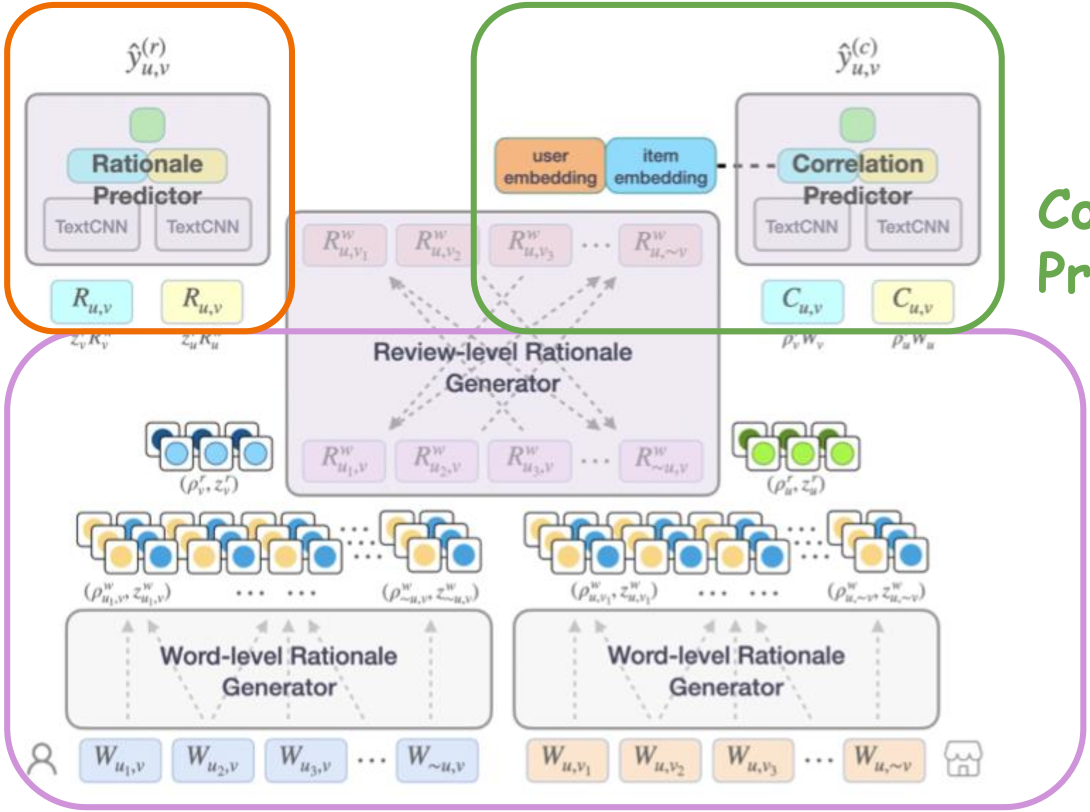
- ratings score

# Goal

- Extract the <u>rationales R</u> from reviews that is the <u>direct cause</u> of the rating Y .
- <u>Predict the rating</u> $\hat{y}_{u,v}$ that reflects how much u likes v.

# Method : RECOMMENDATION VIA REVIEW RATIONALIZATION(R3)

Rationale Predictor

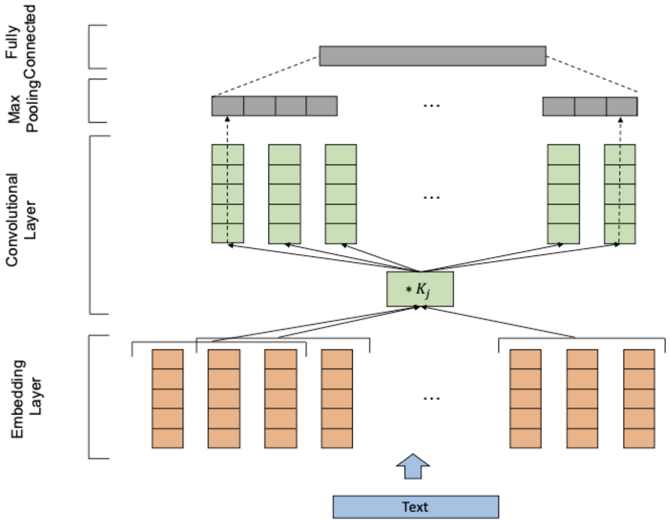



Correlation Predictor

Rationale Generator

# 1.1 Rationale Generator - Word level

- **Input** : each **word** of a **review**
- **output** : the **probabilities** of the words being selected as **rationales**
- **text processor**




$$\boxed{\rho} = \sigma(cW + b).$$

# 1.1 Rationale Generator - Word level

- To specifically choose rationales
- transform probabilities ρ to binary signals z
  - $\lfloor \cdot \rceil$ is the rounding function

$$\boxed{z} = \rho + f_d(\lfloor \rho \rceil - \rho).$$

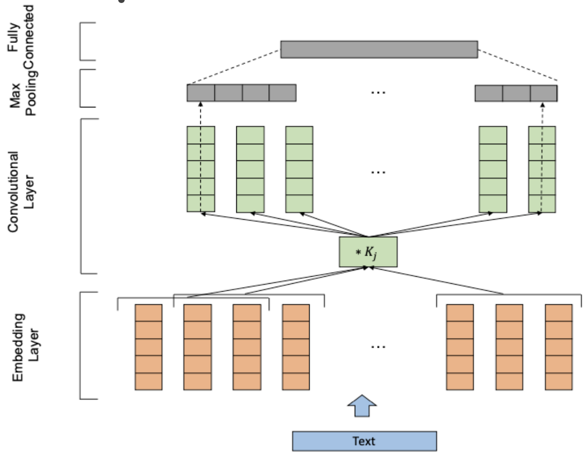
$$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

mask

# 1.2 Rationale Generator - Review level

- Input : word-level rationales from user historical reviews $R^{w}_{u,*}$

  word-level rationales from item historical reviews $R^{w}_{*,v}$
- output : user preference matrix / item preference matrix
- text processor

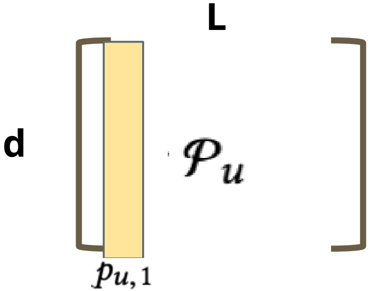



number of reviews from user u

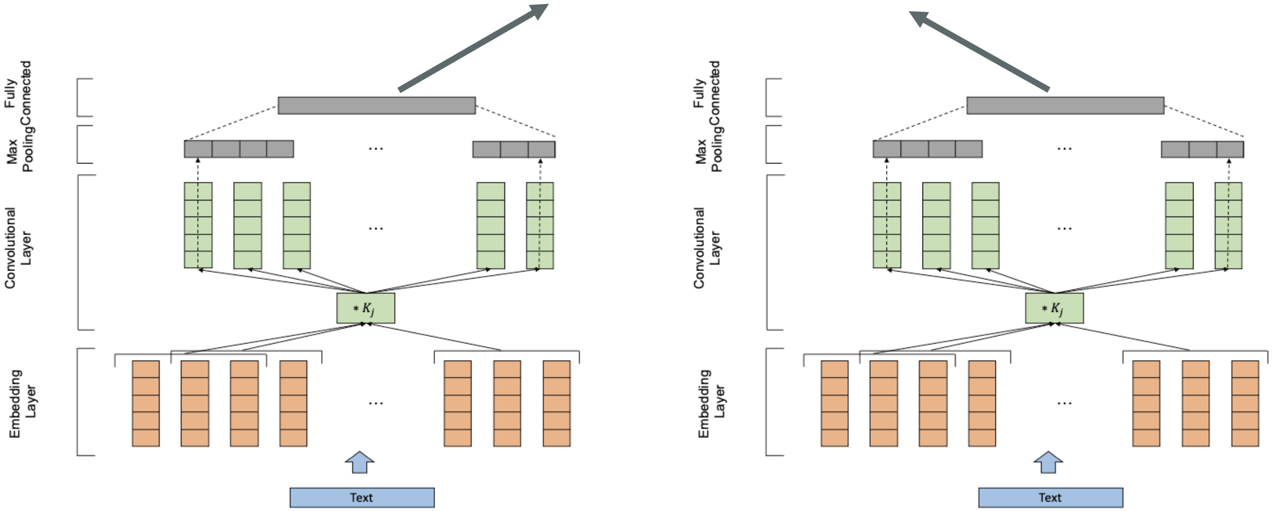
$$\begin{array}{cc} & L \\ d & \begin{bmatrix} \quad \mathcal{P}_u \quad \end{bmatrix} \end{array}$$

output size of the FC layer

# 1.2 Rationale Generator - Review level




FC layer : model user preference on items $p_{u,v}$




user u

item 1

# 1.2 Rationale Generator - Review level

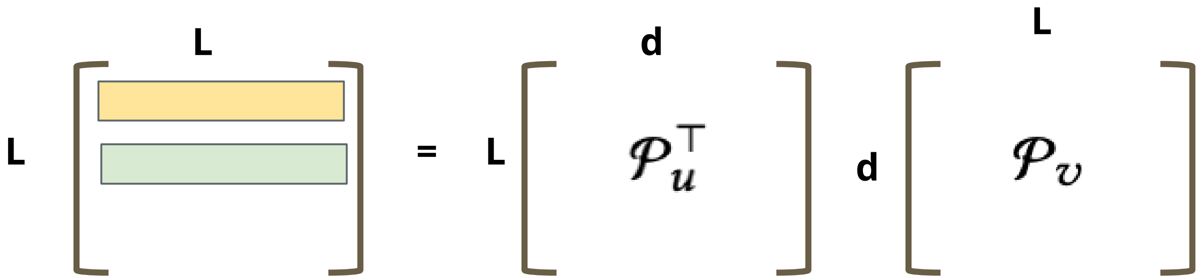
- Input : user preference matrix and item preference matrix
- output : affinity matrix
- to select rationales that match both user interests and item properties

$$s_{u,v} = \mathcal{P}_u^{\top} \mathcal{P}_v.$$

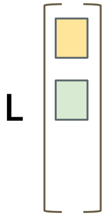
$$\begin{array}{cc} & \mathrm{L} \\ \mathrm{L} & \begin{bmatrix} \quad s_{u,v} \quad \end{bmatrix} \end{array} = \begin{array}{cc} & \mathrm{d} \\ \mathrm{L} & \begin{bmatrix} \quad \mathcal{P}_u^{\top} \quad \end{bmatrix} \end{array} \begin{array}{cc} & \mathrm{L} \\ \mathrm{d} & \begin{bmatrix} \quad \mathcal{P}_v \quad \end{bmatrix} \end{array}$$

# 1.2 Rationale Generator - Review level

- obtain the probabilities of potential rationales to become true rationales

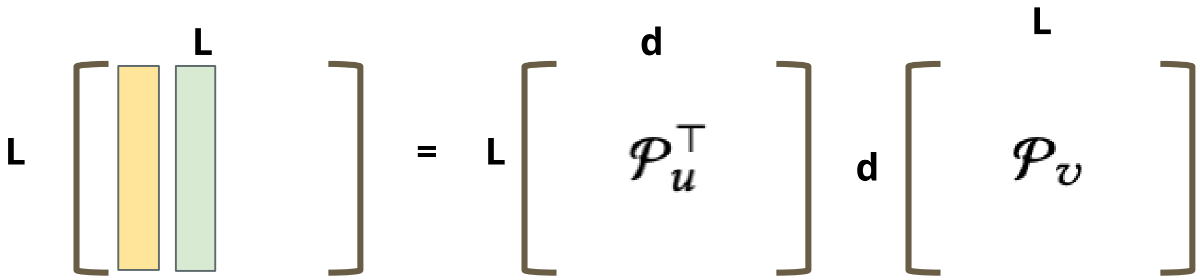



$$\rho_u^r = \sigma(\sum_{row} s),$$

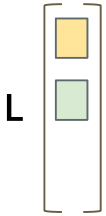


# 1.2 Rationale Generator - Review level

- obtain the probabilities of potential rationales to become true rationales




$$\rho_v^r = \sigma(\sum_{col} s)$$



# 2. Rationale Predictor

- predict user ratings on items <u>only by rationale features</u>
- user rationale features :

$$\gamma_u^{(r)} = {z_u^r}^\top \mathcal{P}_u$$

- item rationale features :

$$\gamma_v^{(r)} = {z_v^r}^\top \mathcal{P}_v$$

- rating prediction :

$$\hat{y}_{u,v}^{(r)} = h_r([\gamma_u^{(r)}, \gamma_v^{(r)}]) + b_u + b_v + \mu$$

# 3. Correlation Predictor

- predict user ratings by utilizing both rationale features and non-rationale features

$$\gamma_u^{(c)} = \rho_u^{r\top} \mathcal{P}_u, \quad \gamma_v^{(c)} = \rho_v^{r\top} \mathcal{P}_v$$

$$\hat{y}_{u,v}^{(c)} = h_c([\gamma_u^{(c)} + \gamma_u^{(e)}, \gamma_v^{(c)} + \gamma_v^{(e)}]) + b_u + b_v + \mu$$

user embedding and the item embedding trained by **matrix factorization**

# 4 Model Learning

- Loss:

$$\mathcal{L}_G = \mathcal{L}_R + \lambda \mathbf{ReLU}(\mathcal{L}_R - \mathcal{L}_C) + \alpha(\mathbb{E}[||R||_1] - \gamma).$$

- MSE:

$$\mathcal{L}_R = \sum_{u,v \in \mathcal{D}} \left(\hat{y}^{(r)}_{u,v} - y_{u,v}\right)^2,$$

$$\mathcal{L}_C = \sum_{u,v \in \mathcal{D}} \left(\hat{y}^{(c)}_{u,v} - y_{u,v}\right)^2.$$

# 4 Model Learning

- the rationale generator needs to construct a minimal feature set with maximum predictive ability

$$\mathcal{L}_G = \mathcal{L}_R + \lambda \mathrm{ReLU}(\mathcal{L}_R - \mathcal{L}_C) + \alpha(\mathbb{E}[||R||_1] - \gamma).$$

# 4 Model Learning

- ensures that the size of the selected rationales is small via a sparsity constraint:

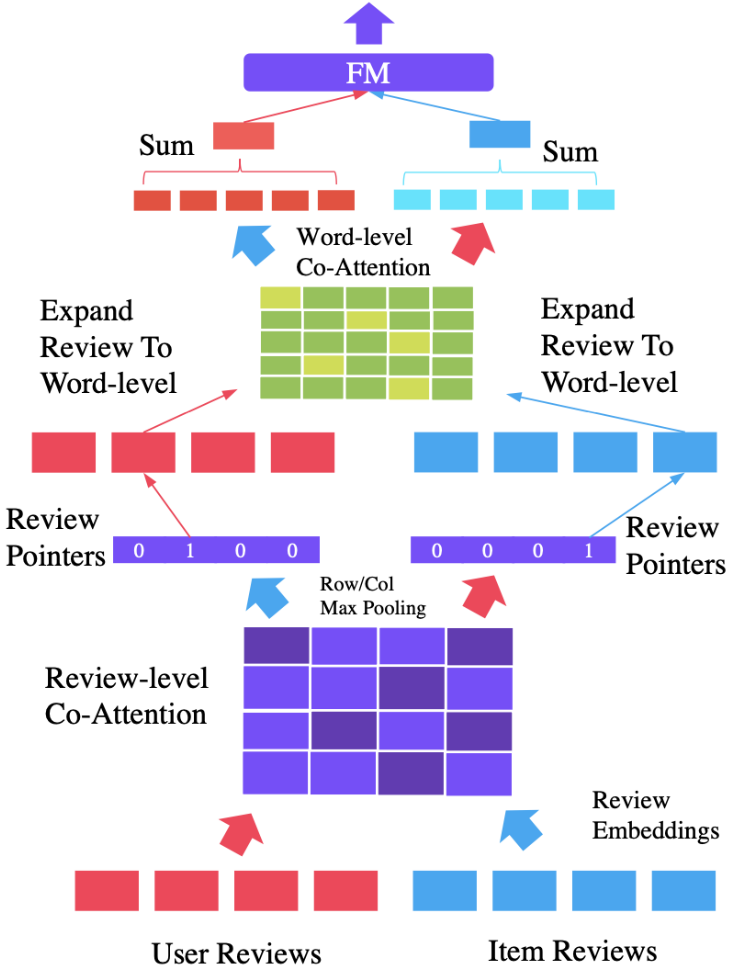
$$\mathcal{L}_G = \mathcal{L}_R + \lambda \mathbf{ReLU}(\mathcal{L}_R - \mathcal{L}_C) + \alpha(\mathbb{E}[||R||_1] - \gamma).$$

# Datasets

| Datasets | #Users | #Items | #Reviews |
|---|---|---|---|
| Home and Kitchen | 35,515 | 11,843 | 341,138 |
| Toys and Games | 19,385 | 11,912 | 167,328 |
| Health and Personal Care | 38,577 | 18,520 | 346,089 |
| Beauty | 22,348 | 12,095 | 198,378 |
| Yelp | 35,515 | 11,843 | 341,138 |

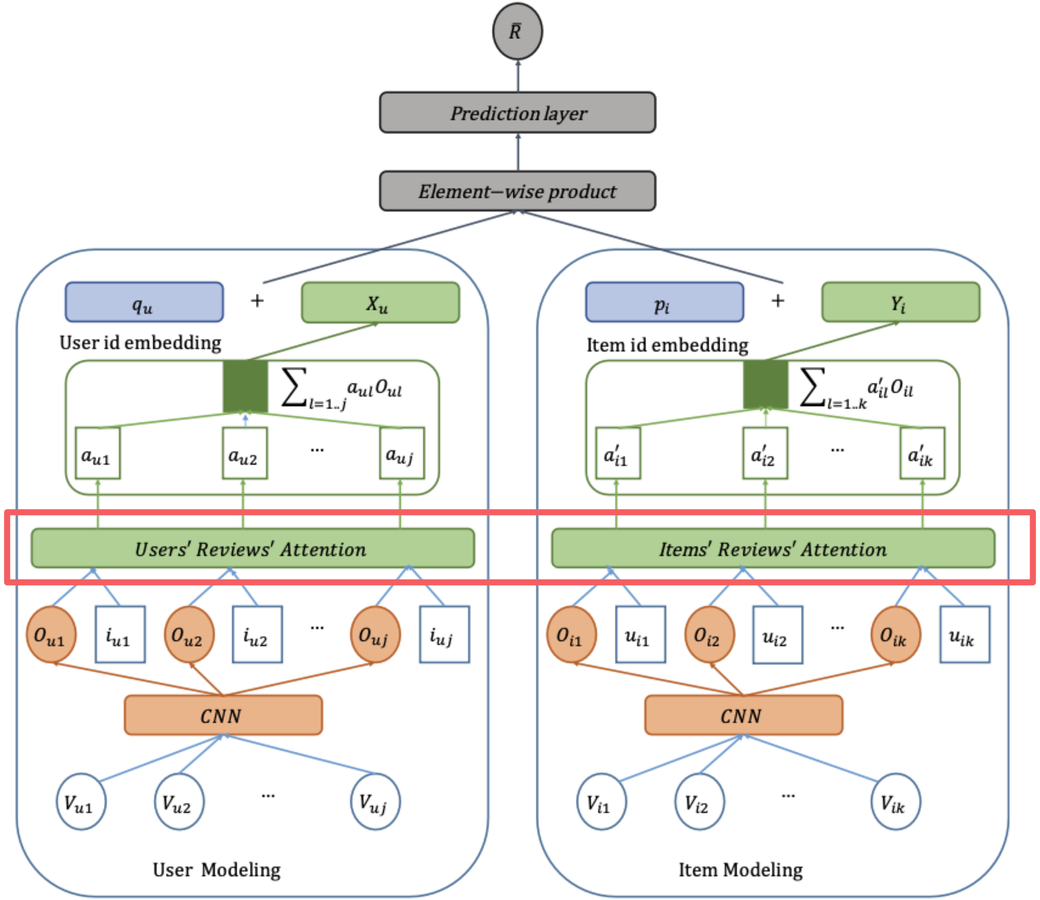
# Experiment-Baseline

- MPCN



# Experiment-Baseline

- NARRE



# Experiment-Baseline

| (a) Raw Dataset | MF | D-CoNN | D-CoNN++ | MPCN | NARRE | R3-R | R3-C |
|---|---|---|---|---|---|---|---|
| Home and Kitchen | 0.8882 | 0.8770 | 0.9895 | 0.8881 | 0.8815 | **0.8421**** | 0.8432 |
| Toys and Game | 0.6916 | 0.6799 | 0.6770 | 0.6925 | 0.6862 | 0.6760* | **0.6537** |
| Health and Personal Care | 0.9239 | 0.9188 | 1.0332 | 0.9531 | 0.9626 | 0.8510** | **0.8470** |
| Yelp | 1.0811 | 1.0796 | 1.0761 | 1.1097 | 1.0672 | **1.0603**** | 1.0638 |

# Experiment-the performance of R3 with data distribution shifts

| (a) Raw Dataset | MF | D-CoNN | D-CoNN++ | MPCN | NARRE | R3-R | R3-C |
|---|---|---|---|---|---|---|---|
| Home and Kitchen | 0.8882 | 0.8770 | 0.9895 | 0.8881 | 0.8815 | **0.8421**** | 0.8432 |
| Toys and Game | 0.6916 | 0.6799 | 0.6770 | 0.6925 | 0.6862 | 0.6760* | **0.6537** |
| Health and Personal Care | 0.9239 | 0.9188 | 1.0332 | 0.9531 | 0.9626 | 0.8510** | **0.8470** |
| Yelp | 1.0811 | 1.0796 | 1.0761 | 1.1097 | 1.0672 | **1.0603**** | 1.0638 |

| (b) Debiased Dataset | MF | D-CoNN | D-CoNN++ | MPCN | NARRE | R3-R | R3-C |
|---|---|---|---|---|---|---|---|
| Home and Kitchen | 3.9481 | 3.8327 | 4.4353 | 4.3538 | 3.8801 | **3.0957**** | 3.1048 |
| Toys and Game | 2.8575 | 2.8244 | 3.0766 | 2.7924 | 2.9847 | **2.0158**** | 2.1809 |
| Health and Personal Care | 3.7198 | 3.8870 | 4.5853 | 4.1551 | 4.1778 | **2.8640**** | 2.9443 |
| Yelp | 2.6131 | 2.8038 | 2.8559 | 3.1663 | 2.8603 | 2.3466** | **2.3432** |

test-u

# Experiment

$$\pi_{u,v} = \text{interest} + \text{bias}_{\text{user}} + \text{bias}_{\text{item}} + \text{bias}_{\text{global}}.$$
$$\pi = \text{interest}.$$
$$\pi_u = \text{interest} + \text{bias}_{\text{user}}.$$
$$\pi_v = \text{interest} + \text{bias}_{\text{item}}.$$

$$\text{PCC} = \frac{\sum_{u,v}(\hat{y}_{u,v} - \bar{\hat{y}}_{u,v})(y_{u,v} - \bar{y}_{u,v})}{\sqrt{\sum_{u,v}(\hat{y}_{u,v} - \bar{\hat{y}}_{u,v})^2}\sqrt{\sum_{u,v}(y_{u,v} - \bar{y}_{u,v})^2}},$$

| Model | MSE | PCC | | | |
|---|---|---|---|---|---|
| | | $\pi_{u,v}$ | $\pi$ | $\pi_u$ | $\pi_v$ |
| benchmark test set | | | | | |
| MF | 1.0811 | 0.3731 | 0.0470 | 0.2017 | 0.3325 |
| NARRE | 1.0672 | 0.3671 | 0.2810 | 0.2307 | 0.3490 |
| R3-R | **1.0603**** | 0.3696 | 0.1799 | 0.1965 | 0.3290 |
| R3-C | 1.0638 | 0.3677 | 0.1638 | 0.2027 | 0.3280 |

# Experiment

| Model | MSE | PCC | | | |
|---|---|---|---|---|---|
| | | $\pi_{u,v}$ | $\pi$ | $\pi_u$ | $\pi_v$ |
| benchmark test set | | | | | |
| MF | 1.0811 | 0.3731 | 0.0470 | 0.2017 | 0.3325 |
| NARRE | 1.0672 | 0.3671 | 0.2810 | 0.2307 | 0.3490 |
| R3-R | **1.0603**** | 0.3696 | 0.1799 | 0.1965 | 0.3290 |
| R3-C | 1.0638 | 0.3677 | 0.1638 | 0.2027 | 0.3280 |
| debiased test set | | | | | |
| MF | 2.6131 | 0.3257 | -0.0004 | -0.0004 | 0.3257 |
| NARRE | 2.8603 | 0.3237 | 0.2603 | 0.2603 | 0.3237 |
| R3-R | 2.3466** | 0.3290 | 0.3128 | 0.3128 | 0.3287 |
| R3-C | **2.3432** | 0.3323 | 0.2209 | 0.2209 | 0.3275 |

# Experiment-Explanation via Rationales

**Table 4: Rationales (words in boldface) from user and item reviews extracted by R3 for predicting the targeted review and rating. The targeted review is: food has improved recently but service is slow and always pesky flies - gross. its hit or miss.**

| user historical reviews | item historical reviews |
|---|---|
| - been here a few times - close location and "burger deals". **service** is **horrible** and **food** is so **slow**, could be better. cute burger ideas, but not worth full price - more **bar** like then food place.<br><br>- water as beverage - **horrid service** - to get water then food, oh boy then waiting for bill. **food** was **good** , but **service blew**! | have been there multiple times as 1/2 off coupons - **bad food** and **service** 3 out of 4 visits - 1/2 off coupons still available - no thank u. even when empty **service sucked**, **burgers overcooked**, side **dish errors** on plate, bad bad, was great place - bar seems **busy** for happy hour maybe, no more attempts by me. |

# Conclusion

- extracts rationales from user and item reviews via a rationale generator to alleviate the effects of spurious correlations in recommendation.

- R3 can achieve accurate recommendation and provide causal-aware explanation based on the rationales