Disentangling User Interest and Conformity for Recommendation with Causal Embedding

Yu Zheng¹, Chen Gao¹, Xiang Li², Xiangnan He³, Depeng Jin¹, Yong Li¹

Tsinghua University¹
University of Hong Kong²
University of Science and Technology of China³

liyong07@tsinghua.edu.cn
Background

• What are the **causes** behind each user-item interaction?

There are two main causes:

• **Interest**
• **Conformity**

  • How users tend to follow other people

**Goal:** Learn disentangled representations for interest and conformity
Motivation

• Why learning disentangled representations?
  • Causal recommendation under **non-IID situations**!
  • IID: independent and identically distributed

• Robustness
  • Recommenders are trained and updated in real-time
  • Training data and test data are not IID

• Interpretability
  • Improve user-friendliness
  • Facilitates algorithm developing
Causal Recommendation

- Inverse Propensity Scoring (IPS)\[^{[1]}\]

\[
\hat{R}_{\text{IPS}}(\hat{Z}|P) = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \frac{1}{|\mathcal{S}_u|} \sum_{i \in \mathcal{S}_u^*} \frac{c(\hat{Z}_{u,i})}{P_{u,i}}
\]

\[= \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \frac{1}{|\mathcal{S}_u|} \sum_{i \in \mathcal{S}_u} \frac{c(\hat{Z}_{u,i})}{P_{u,i}} \cdot O_{u,i}
\]

- Propensity score is estimated from item popularity
- Intuition: impose lower weights on popular items, and boost unpopular items
- Interest and popularity are **bundled** as one unified representation

Two factors are **entangled**! 😞

Causal Recommendation

• Causal Embeddings (CausE)[1]

\[ L^\text{prod}_{\text{CausE}} = L(U\Theta_t, Y_t) + \Omega(\Theta_t) + L(U\Theta_c, Y_c) + \Omega(\Theta_c) + \Omega(\Theta_t - \Theta_c) \]

- **treatment task loss**
- **control task loss**
- regularizer between tasks

• Require a large fraction of **biased** data and a small fraction of **unbiased** data
• Perform two MF on biased and unbiased data, respectively
• Impose L1/L2 regularization on two MF

Still **entangled** representations! 😞

Disentangling interest and conformity

- Variety of conformity
- Conformity depends on **both users and items**
- One user’s conformity varies on **different items**, and conformity towards one item varies for **different users**

- Learning disentangled representations is intrinsically hard
- Only observational data is accessible.
- **No ground-truth** for user interest.

- An interaction can come from one or both factor
- Careful designs are needed for **combining the two factors** to make recommendations.
Methodology: Our DICE Model

• **Disentangling** **Interest and Conformity with Causal Embedding (DICE)**

• **Challenge 1:** Variety of conformity

• **Our proposal:** Adopt separate embeddings of interest and conformity for users and items

• **Benefit 1:** Embedding proximity in high dimensional space can express the variety of conformity (**challenge 1 addressed**)

• **Benefit 2:** Independent modeling of interest and conformity
Methodology: Our DICE Model

• **Disentangling Interest and Conformity with Causal Embedding (DICE)**

• **Challenge 2:** Learning disentangled representations is intrinsically hard

• **Our proposal:** Utilize the colliding effect from causal inference to obtain cause-specific data.

**Intuition:**
Train interest/conformity embeddings with interactions that are caused by interest/conformity
Methodology: Our DICE Model

• **Disentangling Interest and Conformity with Causal Embedding (DICE)**

• **Challenge 3**: Aggregation of the two factors is complicated

• **Our proposal**: Leverage multi-task curriculum learning to combine the two causes.

![Diagram of the DICE Model](image-url)
Methodology: Our DICE Model

- **Disentangling Interest and Conformity with Causal Embedding (DICE)**

- Causal Embedding
- Disentangled Representation Learning
- Multi-task Curriculum Learning

(a) Causal Graph

(b) Causal Embedding
Methodology: Our DICE Model

- Causal graph and Structural Causal Model (SCM)

\[
X_{ui}^{int} := f_1(u, i, N^{int}), \\
X_{ui}^{con} := f_2(u, i, N^{con}), \\
Y_{ui}^{click} := f_3(X_{ui}^{int}, X_{ui}^{con}, N^{click}),
\]

Causal graph

SCM
Methodology: Our DICE Model

- Causal embedding
- **Separate embeddings** for interest and conformity
  - User: $u^{(int)}, u^{(con)}$
  - Item: $i^{(int)}, i^{(con)}$
- Use **inner product** to compute matching score
- Predict click by combining two causes

\[
s^{int}_{ui} = \langle u^{(int)}, i^{(int)} \rangle, \quad s^{con}_{ui} = \langle u^{(con)}, i^{(con)} \rangle,
\]
\[
s^{click}_{ui} = s^{int}_{ui} + s^{con}_{ui},
\]
Methodology: Our DICE Model

- Mining *cause-specific* data with *causal inference*
- Immorality and collider

- **Colliding effect**
  - A and B are independent
  - A and B are **NOT** independent when *conditioned on C*
Methodology: Our DICE Model

- Mining **cause-specific** data with **causal inference**
- e.g.
  - A: whether a student is talented
  - B: whether a student is hard-working
  - C: whether a student passes an exam

- **Bob passes the exam, and Bob is not talented**
  - He is hard-working with high probability

- **Alice doesn’t pass the exam, and Alice is talented**
  - She is most likely not hard-working
Methodology: Our DICE Model

- Mining **cause-specific** data with **causal inference**
- The **colliding effect** can come to help!
- Click is the **collider** of interest and conformity!

![Diagram showing causal relationships between interest, conformity, and click]

- Use popularity as a **proxy** for conformity
  - A clicked item with low popularity
  - high interest
  - An unclicked item with high popularity
  - low interest
Methodology: Our DICE Model

- Notation
- \( M^I \): interest matching probability matrix
- \( M^C \): conformity matching probability matrix

Case 1: \( u \) clicks a popular item \( a \), doesn’t click an unpopular item \( b \)
\[
\begin{align*}
M^C_{ua} &> M^C_{ub}, \\
M^I_{ua} + M^C_{ua} &> M^I_{ub} + M^C_{ub}.
\end{align*}
\]

Case 2: \( u \) clicks an unpopular item \( c \), doesn’t click a popular item \( d \)
\[
\begin{align*}
M^I_{uc} &> M^I_{ud}, M^C_{uc} < M^C_{ud}, \\
M^I_{uc} + M^C_{uc} &> M^I_{ud} + M^C_{ud}.
\end{align*}
\]
Methodology: Our DICE Model

- $\mathcal{O}$: whole training set $(u, i, j)$: user, pos item, neg item
- $\mathcal{O}_1$: negative samples more popular than positive samples
- $\mathcal{O}_2$: negative samples less popular than positive samples

$$\mathcal{O} = \mathcal{O}_1 + \mathcal{O}_2$$

$\mathcal{O}_1$

$$M_{ua}^C > M_{ub}^C,$$

$$M_i^I + M_u^C > M_{ub}^I + M_{ub}^C.$$  

$\mathcal{O}_2$

$$M_{uc}^I > M_{ud}^I, M_{uc}^C < M_{ud}^C,$$

$$M_{uc}^I + M_{uc}^C > M_{ud}^I + M_{ud}^C.$$  

Solution: train different embeddings with different cause-specific data
Methodology: Our DICE Model

• Main task: estimating clicks

\[ O_1 \]
\[ M_{ua}^C > M_{ub}^C, \]
\[ M_{ua}^I + M_{ua}^C > M_{ub}^I + M_{ub}^C. \]

\[ O_2 \]
\[ M_{uc}^I > M_{ud}^I, M_{uc}^C < M_{ud}^C, \]
\[ M_{uc}^I + M_{uc}^C > M_{ud}^I + M_{ud}^C. \]

\[ u^t = u^{(int)} \| u^{(con)}, i^t = i^{(int)} \| i^{(con)}, j^t = j^{(int)} \| j^{(con)}, \]

\[ L_{\text{click}}^{O_1+O_2} = \sum_{(u,i,j) \in O} \text{BPR}(\langle u^t, i^t \rangle, \langle u^t, j^t \rangle). \]
Methodology: Our DICE Model

- Interest modeling
  - Only use interest embedding

\[ \mathcal{O}_2 \]

\[ M_{uc}^I > M_{ud}^I, M_{uc}^C < M_{ud}^C, \]
\[ M_{uc}^I + M_{uc}^C > M_{ud}^I + M_{ud}^C. \]

\[ L_{interest}^{O_2} = \sum_{(u, i, j) \in \mathcal{O}_2} \text{BPR}(\langle u^{(int)}, i^{(int)} \rangle, \langle u^{(int)}, j^{(int)} \rangle). \]
Methodology: Our DICE Model

• Conformity modeling
• Only use conformity embedding

\[ O_1 \]

\[ M_{ua}^C > M_{ub}^C, \]

\[ M_{ua}^I + M_{ua}^C > M_{ub}^I + M_{ub}^C. \]

\[ L_{\text{conformity}}^{O_1} = \sum_{(u, i, j) \in O_1} \text{BPR}(\langle u^{(\text{con})}, i^{(\text{con})}, \langle u^{(\text{con})}, j^{(\text{con})} \rangle), \]

\[ O_2 \]

\[ M_{uc}^I > M_{ud}^I, M_{uc}^C < M_{ud}^C, \]

\[ M_{uc}^I + M_{uc}^C > M_{ud}^I + M_{ud}^C. \]

\[ L_{\text{conformity}}^{O_2} = \sum_{(u, i, j) \in O_2} -\text{BPR}(\langle u^{(\text{con})}, i^{(\text{con})}, \langle u^{(\text{con})}, j^{(\text{con})} \rangle), \]

\[ L_{\text{conformity}}^{O_1} + L_{\text{conformity}}^{O_2} = L_{\text{conformity}}^{O_1} + L_{\text{conformity}}^{O_2}. \]
Methodology: Our DICE Model

- Discrepancy task
  - direct supervision on disentanglement
- L1-inv: \(-L1(E^{(int)}, E^{(con)})\)
- L2-inv: \(-L2(E^{(int)}, E^{(con)})\)
- distance correlation:

\[
dCor(E^{(int)}, E^{(con)}) = \frac{dCov(E^{(int)}, E^{(con)})}{\sqrt{dVar(E^{(int)}) \cdot dVar(E^{(con)})}}
\]
Methodology: Our DICE Model

- Multi-task learning
  \[ L = L_{\text{click}}^{O_1+O_2} + \alpha(L_{\text{interest}}^{O_2} + L_{\text{conformity}}^{O_1+O_2}) + \beta L_{\text{discrepancy}}. \]

- Popularity based Negative Sampling with Margin (PNSM)
  - Popularity of the positive item: \( p \)
  - Sample negative items with popularity:
    - Larger than \( p + m \)
    - Lower than \( p - m \)
  - Large \( m \): high confidence on inequalities, easy
  - Small \( m \): low confidence on inequalities, hard

- Curriculum learning: an easy-to-hard strategy
  - \textbf{decay} \( m, \alpha \) and \( \beta \) by a factor of 0.9 after each epoch
Experiments

- Datasets:
  - Movielens-10M
  - Netflix

- Evaluation: non-IID protocol (same with CausE[1]):
  - Train: 60% normal+ 10% intervened
  - Validation: 10% intervened
  - Test: 20% intervened

- Metrics:
  - Recall, Hit Ratio, NDCG

- Recommendation models
  - MF[2]
  - LightGCN[3]

---

<table>
<thead>
<tr>
<th>Dataset</th>
<th>User</th>
<th>Item</th>
<th>Interaction</th>
<th>Ent. Train</th>
<th>Ent. Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Movielens-10M</td>
<td>37962</td>
<td>4819</td>
<td>1371473</td>
<td>6.22</td>
<td>7.97</td>
</tr>
<tr>
<td>Netflix</td>
<td>32450</td>
<td>8432</td>
<td>2212690</td>
<td>6.85</td>
<td>8.54</td>
</tr>
</tbody>
</table>

Experiments

- **RQ1**: How does our proposed DICE framework perform compared with state-of-the-art causal recommendation methods under **non-IID circumstances**?

- **RQ2**: Can the proposed DICE framework guarantee **interpretability**?

- **RQ3**: Can the proposed DICE framework guarantee **robustness**?
Experiments

- Overall Comparison

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Model</th>
<th>Method</th>
<th>Recall</th>
<th>HR</th>
<th>NDCG</th>
<th>Recall</th>
<th>HR</th>
<th>NDCG</th>
<th>Recall</th>
<th>HR</th>
<th>NDCG</th>
<th>Recall</th>
<th>HR</th>
<th>NDCG</th>
</tr>
</thead>
<tbody>
<tr>
<td>TopK = 20</td>
<td>Movielens-10M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>0.1286</td>
<td>0.4429</td>
<td>0.0846</td>
<td>0.2346</td>
<td>0.6295</td>
<td>0.1170</td>
<td>0.1122</td>
<td>0.5194</td>
<td>0.0943</td>
<td>0.1928</td>
<td>0.6749</td>
<td>0.1185</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPS</td>
<td>0.1335</td>
<td>0.4434</td>
<td>0.0852</td>
<td>0.2376</td>
<td>0.6288</td>
<td>0.1174</td>
<td>0.1058</td>
<td>0.4882</td>
<td>0.0864</td>
<td>0.1855</td>
<td>0.6562</td>
<td>0.1112</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPS-C</td>
<td>0.1367</td>
<td>0.4564</td>
<td>0.0875</td>
<td>0.2429</td>
<td>0.6383</td>
<td>0.1203</td>
<td>0.1119</td>
<td>0.5046</td>
<td>0.0919</td>
<td>0.1938</td>
<td>0.6700</td>
<td>0.1174</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPS-CN</td>
<td>0.1412</td>
<td>0.4700</td>
<td>0.0925</td>
<td>0.2509</td>
<td>0.6477</td>
<td>0.1264</td>
<td>0.1080</td>
<td>0.5042</td>
<td>0.0935</td>
<td>0.1912</td>
<td>0.6621</td>
<td>0.1185</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPS-CNSR</td>
<td>0.1365</td>
<td>0.4588</td>
<td>0.0895</td>
<td>0.2419</td>
<td>0.6366</td>
<td>0.1219</td>
<td>0.1110</td>
<td>0.5159</td>
<td>0.0948</td>
<td>0.1937</td>
<td>0.6713</td>
<td>0.1192</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CausE</td>
<td>0.1157</td>
<td>0.4066</td>
<td>0.0744</td>
<td>0.2121</td>
<td>0.5924</td>
<td>0.1037</td>
<td>0.0935</td>
<td>0.4641</td>
<td>0.0782</td>
<td>0.1651</td>
<td>0.6272</td>
<td>0.0994</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DICE</td>
<td><strong>0.1634</strong></td>
<td><strong>0.5197</strong></td>
<td><strong>0.1084</strong></td>
<td><strong>0.2872</strong></td>
<td><strong>0.6975</strong></td>
<td><strong>0.1468</strong></td>
<td><strong>0.1258</strong></td>
<td><strong>0.5545</strong></td>
<td><strong>0.1070</strong></td>
<td><strong>0.2164</strong></td>
<td><strong>0.7090</strong></td>
<td><strong>0.1345</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TopK = 50</td>
<td>Netflix</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>0.1378</td>
<td>0.4625</td>
<td>0.0898</td>
<td>0.2513</td>
<td>0.6505</td>
<td>0.1247</td>
<td>0.1026</td>
<td>0.4908</td>
<td>0.0870</td>
<td>0.1842</td>
<td>0.6609</td>
<td>0.1112</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPS</td>
<td>0.1394</td>
<td>0.4645</td>
<td>0.0919</td>
<td>0.2538</td>
<td>0.6473</td>
<td>0.1275</td>
<td>0.1101</td>
<td>0.5091</td>
<td>0.0950</td>
<td>0.1941</td>
<td>0.6657</td>
<td>0.1203</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPS-C</td>
<td>0.1478</td>
<td>0.4829</td>
<td>0.0971</td>
<td>0.2654</td>
<td>0.6632</td>
<td>0.1339</td>
<td>0.1157</td>
<td>0.5219</td>
<td>0.1004</td>
<td>0.2037</td>
<td>0.6816</td>
<td>0.1270</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPS-CN</td>
<td>0.1119</td>
<td>0.3997</td>
<td>0.0701</td>
<td>0.2281</td>
<td>0.6112</td>
<td>0.1057</td>
<td>0.0726</td>
<td>0.3991</td>
<td>0.0643</td>
<td>0.1472</td>
<td>0.5841</td>
<td>0.0866</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPS-CNSR</td>
<td>0.1300</td>
<td>0.4427</td>
<td>0.0852</td>
<td>0.2336</td>
<td>0.6282</td>
<td>0.1171</td>
<td>0.0826</td>
<td>0.4337</td>
<td>0.0715</td>
<td>0.1589</td>
<td>0.6124</td>
<td>0.0940</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CausE</td>
<td>0.1027</td>
<td>0.3729</td>
<td>0.0632</td>
<td>0.2044</td>
<td>0.5811</td>
<td>0.0941</td>
<td>0.0838</td>
<td>0.4289</td>
<td>0.0677</td>
<td>0.1569</td>
<td>0.6119</td>
<td>0.0902</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DICE</td>
<td><strong>0.1812</strong></td>
<td><strong>0.5563</strong></td>
<td><strong>0.1228</strong></td>
<td><strong>0.3100</strong></td>
<td><strong>0.7216</strong></td>
<td><strong>0.1629</strong></td>
<td><strong>0.1420</strong></td>
<td><strong>0.5910</strong></td>
<td><strong>0.1217</strong></td>
<td><strong>0.2367</strong></td>
<td><strong>0.7340</strong></td>
<td><strong>0.1499</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Experiments

- Observations

Table 2: Overall performance on Movielens-10M dataset and Netflix dataset.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Model</th>
<th>Method</th>
<th>TopK = 20</th>
<th>TopK = 50</th>
<th>Netflix</th>
<th>TopK = 20</th>
<th>TopK = 50</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Recall</td>
<td>HR</td>
<td>NDCG</td>
<td>Recall</td>
<td>HR</td>
<td>NDCG</td>
</tr>
<tr>
<td>Movielens-10M</td>
<td></td>
<td>Recall</td>
<td>HR</td>
<td>NDCG</td>
<td>Recall</td>
<td>HR</td>
<td>NDCG</td>
</tr>
<tr>
<td>MF</td>
<td>None</td>
<td>0.1286</td>
<td>0.4429</td>
<td>0.0846</td>
<td>0.2346</td>
<td>0.6295</td>
<td>0.1170</td>
</tr>
<tr>
<td></td>
<td>IPS</td>
<td>0.1335</td>
<td>0.4434</td>
<td>0.0852</td>
<td>0.2376</td>
<td>0.6288</td>
<td>0.1174</td>
</tr>
<tr>
<td></td>
<td>IPS-C</td>
<td>0.1367</td>
<td>0.4564</td>
<td>0.0875</td>
<td>0.2429</td>
<td>0.6383</td>
<td>0.1203</td>
</tr>
<tr>
<td></td>
<td>IPS-CN</td>
<td>0.1412</td>
<td>0.4700</td>
<td>0.0925</td>
<td>0.2509</td>
<td>0.6477</td>
<td>0.1264</td>
</tr>
<tr>
<td></td>
<td>IPS-CNSR</td>
<td>0.1365</td>
<td>0.4588</td>
<td>0.0895</td>
<td>0.2419</td>
<td>0.6366</td>
<td>0.1219</td>
</tr>
<tr>
<td></td>
<td>CausalE</td>
<td>0.1157</td>
<td>0.4066</td>
<td>0.0744</td>
<td>0.2121</td>
<td>0.5924</td>
<td>0.1037</td>
</tr>
<tr>
<td>DICE</td>
<td></td>
<td>0.1634</td>
<td>0.5197</td>
<td>0.1084</td>
<td>0.2872</td>
<td>0.6975</td>
<td>0.1468</td>
</tr>
<tr>
<td>GCN</td>
<td>None</td>
<td>0.1378</td>
<td>0.4629</td>
<td>0.0898</td>
<td>0.2513</td>
<td>0.6505</td>
<td>0.1247</td>
</tr>
<tr>
<td></td>
<td>IPS</td>
<td>0.1394</td>
<td>0.4645</td>
<td>0.0919</td>
<td>0.2538</td>
<td>0.6473</td>
<td>0.1275</td>
</tr>
<tr>
<td></td>
<td>IPS-C</td>
<td>0.1478</td>
<td>0.4829</td>
<td>0.0971</td>
<td>0.2654</td>
<td>0.6632</td>
<td>0.1339</td>
</tr>
<tr>
<td></td>
<td>IPS-CN</td>
<td>0.1119</td>
<td>0.3997</td>
<td>0.0701</td>
<td>0.2281</td>
<td>0.6112</td>
<td>0.1057</td>
</tr>
<tr>
<td></td>
<td>IPS-CNSR</td>
<td>0.1300</td>
<td>0.4427</td>
<td>0.0852</td>
<td>0.2336</td>
<td>0.6282</td>
<td>0.1171</td>
</tr>
<tr>
<td></td>
<td>CausalE</td>
<td>0.1027</td>
<td>0.3729</td>
<td>0.0632</td>
<td>0.2044</td>
<td>0.5811</td>
<td>0.0941</td>
</tr>
<tr>
<td>DICE</td>
<td></td>
<td>0.1812</td>
<td>0.5563</td>
<td>0.1228</td>
<td>0.3100</td>
<td>0.7216</td>
<td>0.1629</td>
</tr>
</tbody>
</table>

- Our proposed DICE framework outperforms baselines with significant improvements with respect to all metrics on both datasets.
• DICE is a highly general framework which can be combined with various recommendation models.

### Table 2: Overall performance on Movielens-10M dataset and Netflix dataset.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Model</th>
<th>Method</th>
<th>Recall</th>
<th>HR</th>
<th>NDCG</th>
<th>Recall</th>
<th>HR</th>
<th>NDCG</th>
<th>Recall</th>
<th>HR</th>
<th>NDCG</th>
<th>Recall</th>
<th>HR</th>
<th>NDCG</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MF</td>
<td>None</td>
<td>0.1286</td>
<td>0.4429</td>
<td>0.0846</td>
<td>0.2346</td>
<td>0.6295</td>
<td>0.1170</td>
<td>0.1122</td>
<td>0.5194</td>
<td>0.0943</td>
<td>0.1928</td>
<td>0.6749</td>
<td>0.1185</td>
</tr>
<tr>
<td></td>
<td>MF</td>
<td>IPS</td>
<td>0.1335</td>
<td>0.4434</td>
<td>0.0852</td>
<td>0.2376</td>
<td>0.6288</td>
<td>0.1174</td>
<td>0.1058</td>
<td>0.4882</td>
<td>0.0864</td>
<td>0.1855</td>
<td>0.6562</td>
<td>0.1112</td>
</tr>
<tr>
<td></td>
<td>MF</td>
<td>IPS-C</td>
<td>0.1367</td>
<td>0.4564</td>
<td>0.0875</td>
<td>0.2429</td>
<td>0.6383</td>
<td>0.1203</td>
<td>0.1119</td>
<td>0.5046</td>
<td>0.0919</td>
<td>0.1938</td>
<td>0.6700</td>
<td>0.1174</td>
</tr>
<tr>
<td></td>
<td>MF</td>
<td>IPS-CN</td>
<td>0.1412</td>
<td>0.4700</td>
<td>0.0925</td>
<td>0.2509</td>
<td>0.6477</td>
<td>0.1264</td>
<td>0.1080</td>
<td>0.5042</td>
<td>0.0935</td>
<td>0.1912</td>
<td>0.6621</td>
<td>0.1185</td>
</tr>
<tr>
<td></td>
<td>MF</td>
<td>IPS-CNSR</td>
<td>0.1365</td>
<td>0.4588</td>
<td>0.0895</td>
<td>0.2419</td>
<td>0.6366</td>
<td>0.1219</td>
<td>0.1110</td>
<td>0.5159</td>
<td>0.0948</td>
<td>0.1937</td>
<td>0.6713</td>
<td>0.1192</td>
</tr>
<tr>
<td></td>
<td>MF</td>
<td>CausE</td>
<td>0.1157</td>
<td>0.4066</td>
<td>0.0744</td>
<td>0.2121</td>
<td>0.5924</td>
<td>0.1037</td>
<td>0.0935</td>
<td>0.4641</td>
<td>0.0782</td>
<td>0.1651</td>
<td>0.6272</td>
<td>0.0994</td>
</tr>
<tr>
<td></td>
<td>MF</td>
<td>DICE</td>
<td>0.1634</td>
<td>0.5197</td>
<td>0.1084</td>
<td>0.2872</td>
<td>0.6975</td>
<td>0.1468</td>
<td>0.1258</td>
<td>0.5545</td>
<td>0.1070</td>
<td>0.2164</td>
<td>0.7090</td>
<td>0.1345</td>
</tr>
<tr>
<td></td>
<td>GCN</td>
<td>None</td>
<td>0.1378</td>
<td>0.4625</td>
<td>0.0898</td>
<td>0.2513</td>
<td>0.6505</td>
<td>0.1247</td>
<td>0.1026</td>
<td>0.4908</td>
<td>0.0870</td>
<td>0.1842</td>
<td>0.6609</td>
<td>0.1112</td>
</tr>
<tr>
<td></td>
<td>GCN</td>
<td>IPS</td>
<td>0.1394</td>
<td>0.4645</td>
<td>0.0919</td>
<td>0.2538</td>
<td>0.6473</td>
<td>0.1275</td>
<td>0.1101</td>
<td>0.5091</td>
<td>0.0950</td>
<td>0.1941</td>
<td>0.6657</td>
<td>0.1203</td>
</tr>
<tr>
<td></td>
<td>GCN</td>
<td>IPS-C</td>
<td>0.1478</td>
<td>0.4829</td>
<td>0.0971</td>
<td>0.2654</td>
<td>0.6632</td>
<td>0.1339</td>
<td>0.1157</td>
<td>0.5219</td>
<td>0.1004</td>
<td>0.2037</td>
<td>0.6816</td>
<td>0.1270</td>
</tr>
<tr>
<td></td>
<td>GCN</td>
<td>IPS-CN</td>
<td>0.1119</td>
<td>0.3997</td>
<td>0.0701</td>
<td>0.2281</td>
<td>0.6112</td>
<td>0.1057</td>
<td>0.0726</td>
<td>0.3991</td>
<td>0.0643</td>
<td>0.1472</td>
<td>0.5841</td>
<td>0.0866</td>
</tr>
<tr>
<td></td>
<td>GCN</td>
<td>IPS-CNSR</td>
<td>0.1300</td>
<td>0.4427</td>
<td>0.0852</td>
<td>0.2336</td>
<td>0.6282</td>
<td>0.1171</td>
<td>0.0826</td>
<td>0.4337</td>
<td>0.0715</td>
<td>0.1589</td>
<td>0.6124</td>
<td>0.0940</td>
</tr>
<tr>
<td></td>
<td>GCN</td>
<td>CausE</td>
<td>0.1027</td>
<td>0.3729</td>
<td>0.0632</td>
<td>0.2044</td>
<td>0.5811</td>
<td>0.0941</td>
<td>0.0838</td>
<td>0.4289</td>
<td>0.0677</td>
<td>0.1569</td>
<td>0.6119</td>
<td>0.0902</td>
</tr>
<tr>
<td></td>
<td>GCN</td>
<td>DICE</td>
<td>0.1812</td>
<td>0.5563</td>
<td>0.1228</td>
<td>0.3100</td>
<td>0.7216</td>
<td>0.1629</td>
<td>0.1420</td>
<td>0.5910</td>
<td>0.1217</td>
<td>0.2367</td>
<td>0.7340</td>
<td>0.1499</td>
</tr>
</tbody>
</table>
Experiments

• Interpretability

• Conformity embeddings of items with different popularity form layers.
Experiments

• Interpretability

• Interest embeddings of items with different popularity are uniformly distributed in the space.
Experiments

- Interpretability

- Conformity embeddings largely captures conformity, and interest embeddings squeeze out conformity
Experiments

• Robustness

• Test data with different strength of intervention
• DICE is more robust than IPS based method under different levels of intervention
Conclusion and Future Work

• We propose to learn disentangled representations of user interest and conformity for recommendation with tools of causal inference. A general framework DICE is developed which shows great robustness and interpretability under non-IID situations.

• Future work
  • Extend DICE to incorporate more features.
  • Learn disentangled representations for finer-grained user interest, e.g. price preference, brand preference...
  • Codes can be found at: https://github.com/tsinghua-fib-lab/DICE
Thanks for listening!

Contact: liyong07@tsinghua.edu.cn
Lab Info: http://fi.ee.tsinghua.edu.cn