Decision Support System for Chronic Diseases Based on Drug-Drug Interactions

Tian Bian1, Y uli Jiang1, Jia Li2*, Tingyang Xu3*, Y u Rong3,
Yi Su4, Timothy Kwok1, Helen Meng1, Hong Cheng1
1 The Chinese University of Hong Kong,
2 Hong Kong University of Science and Technology (Guangzhou),
3 Tencent AI Lab, 4 Hunan Normal University
{tbian,hmmeng,hcheng}@se.cuhk.edu.hk, yljiang@cse.cuhk.edu.hk, jialee@ust.hk,
tingyangxu@tencent.com, yu.rong@hotmail.com, alddle@hunnu.edu.cn, tkwok@cuhk.edu.hk

ICDE2023

Code: https://github.com/TianBian95/DSSDDI.

2023. 4. 22 • ChongQing

Reported by Yang Peng

1.Introduction

2.Method

3.Experiments

Introduction

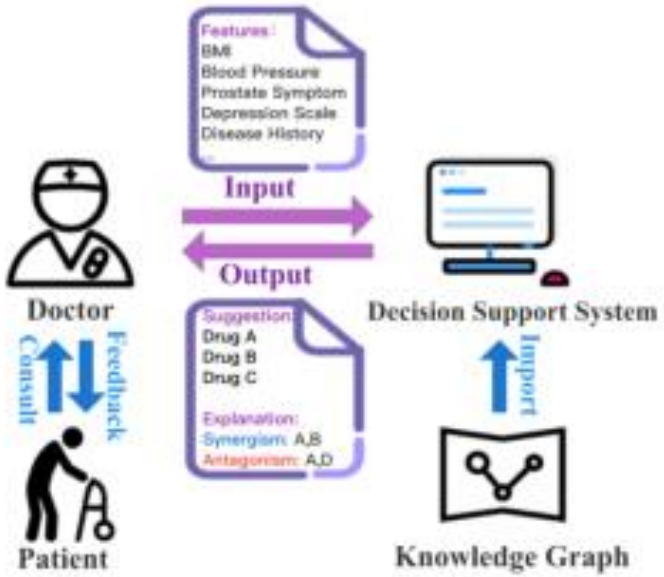


Fig. 1. Our proposed decision support system uses external knowledge of DDI. Given patient features as input, a doctor can obtain the medication suggestions from the system, as well as the corresponding DDI explanations.

1. leveraging DDI to avoid severe ADEs in medication suggestions.

2. employing the causal model to learn the potential causal relationships between DDI and medication suggestions to improve the accuracy.

Method

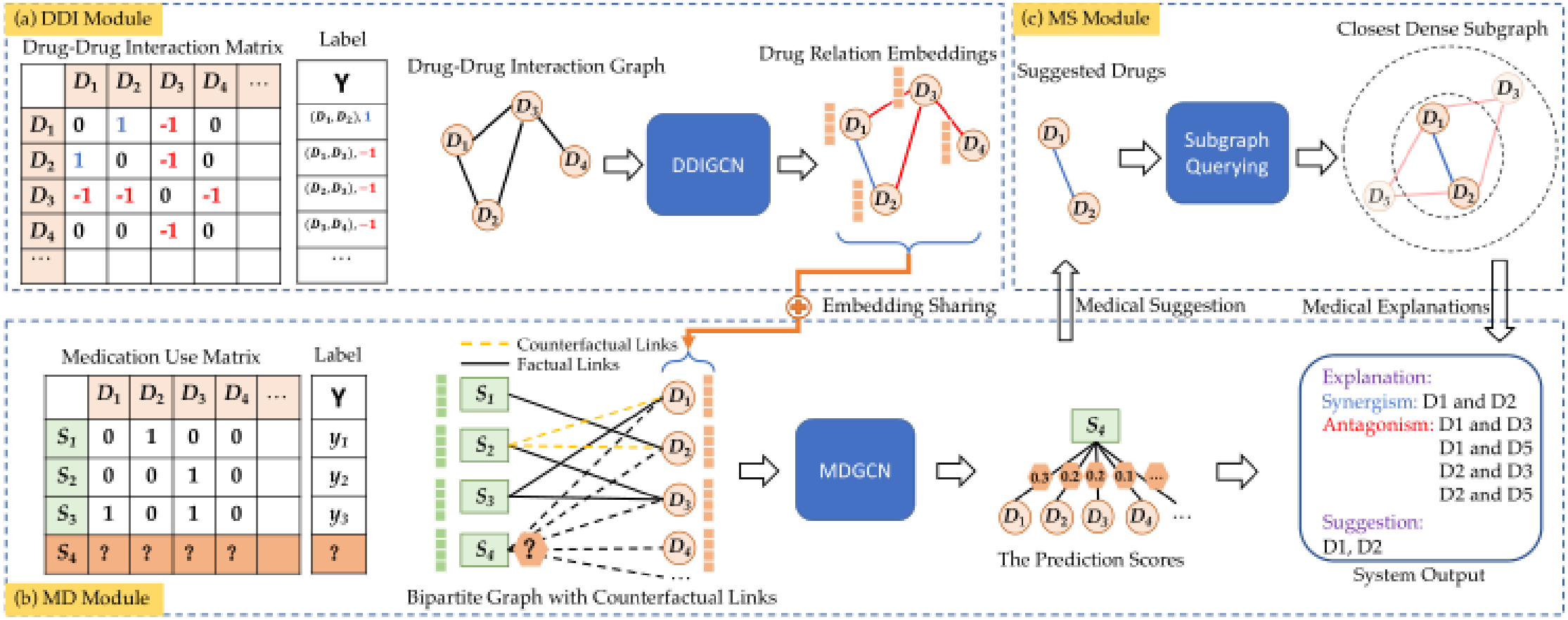
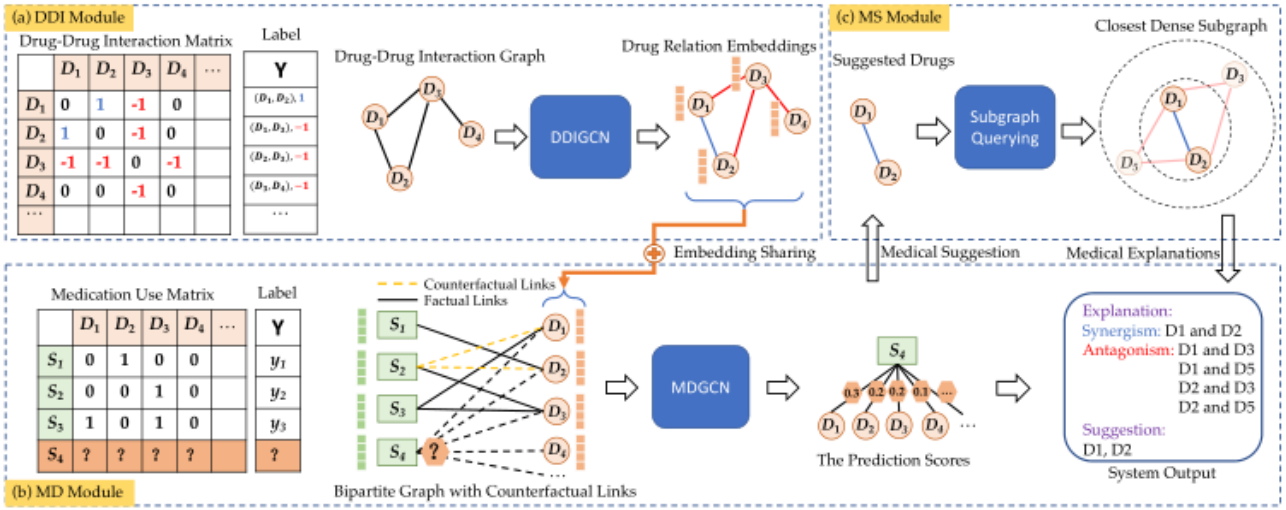
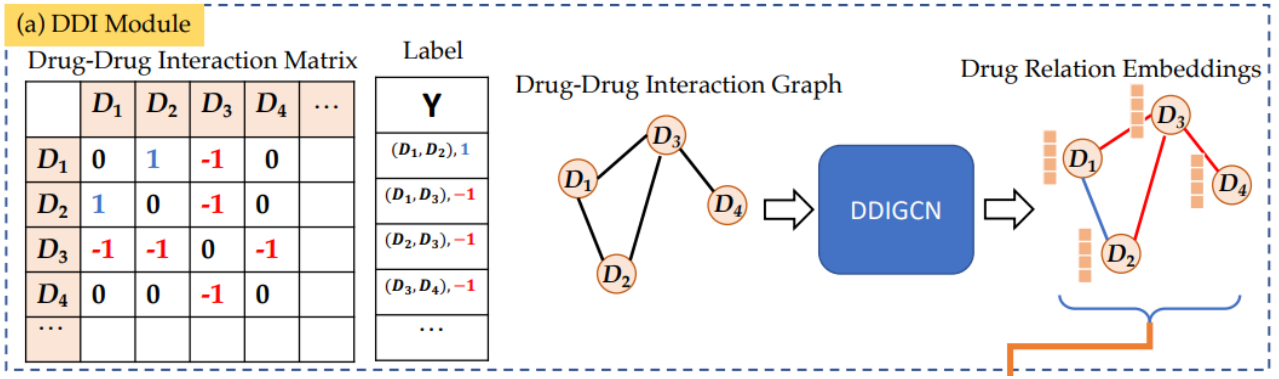


Fig. 4. The proposed decision support system (DSSDDI) consists of three modules: DDI module, MD module and MS module. (a) In the DDI module, we construct the drug-drug interaction matrix where the entries of the matrix represent the synergistic effects (blue) or antagonistic effects (red) among drugs. Through a proposed DDIGCN, we obtain drug relation representations shared with the MD module. (b) In the MD module, we first construct a bipartite graph from the medication use (black lines) and counterfactual links (yellow dashed lines). Then, we obtain the suggested drugs through the proposed MDGCN. (c) In the MS module, the explanation of the suggested drugs is extracted through a subgraph querying algorithm.

Method





Drug-Drug Interaction Module

DDI graph $G = (V, E)$

defined as:

$$\mathbf{z}_v^{(t)} = f_{\mathbf{\Theta_1}}^{(t)}\left((1+\epsilon^{(t)})\cdot \mathbf{z}_v^{(t-1)} + \frac{\sum_{u\in\mathcal{N}_v}\mathbf{z}_u^{(t-1)}}{|\mathcal{N}_v|}\right), \quad (1)$$

$$B_v(t) = \{D_u | e_{uv} = 1\} \quad U_v(t) = \{D_u | e_{uv} = -1\}$$

$$\mathbf{h}_v^{B(t)} = \sigma(\mathbf{W}^{B(t)}[\sum_{e_{iv}=1}\frac{\mathbf{h}_i^{B(t-1)}}{|B(t)|}, \sum_{e_{jv}=-1}\frac{\mathbf{h}_j^{U(t-1)}}{|U(t)|}, \mathbf{h}_v^{B(t-1)}]) \quad (2)$$

$$\mathbf{h}_v^{U(t)} = \sigma(\mathbf{W}^{U(t)}[\sum_{e_{iv}=1}\frac{\mathbf{h}_i^{U(t-1)}}{|U(t)|}, \sum_{e_{jv}=-1}\frac{\mathbf{h}_j^{B(t-1)}}{|B(t)|}, \mathbf{h}_v^{U(t-1)}]) \quad (3)$$

Finally, $\mathbf{z}_v^{(t)}$ is obtained by concatenating $\mathbf{h}_v^{B(t)}$ and $\mathbf{h}_v^{U(t)}$:

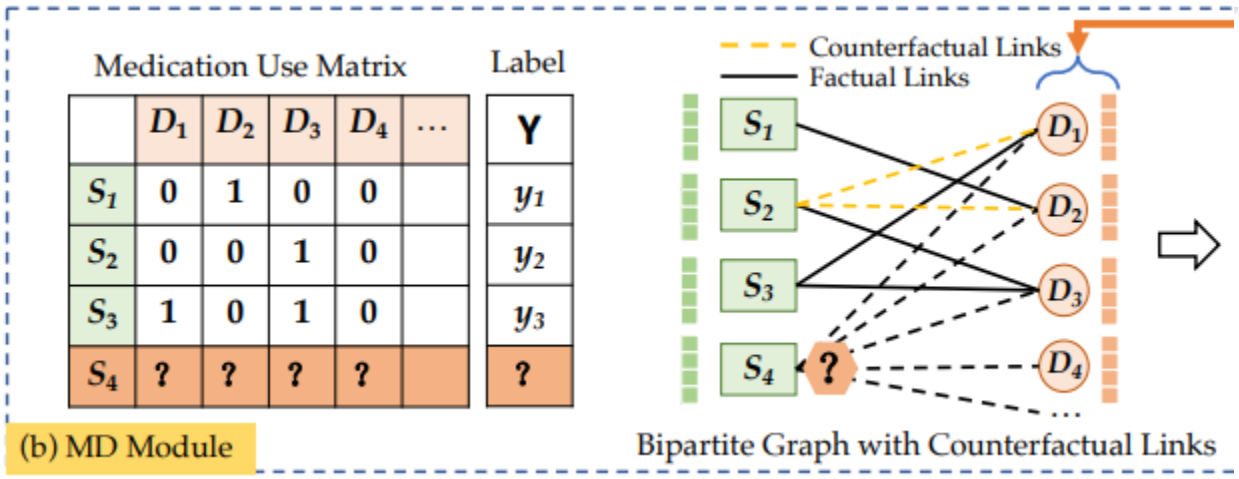
$$\mathbf{z}_v^{(t)} = [\mathbf{h}_v^{B(t)}, \mathbf{h}_v^{U(t)}], \quad (4)$$

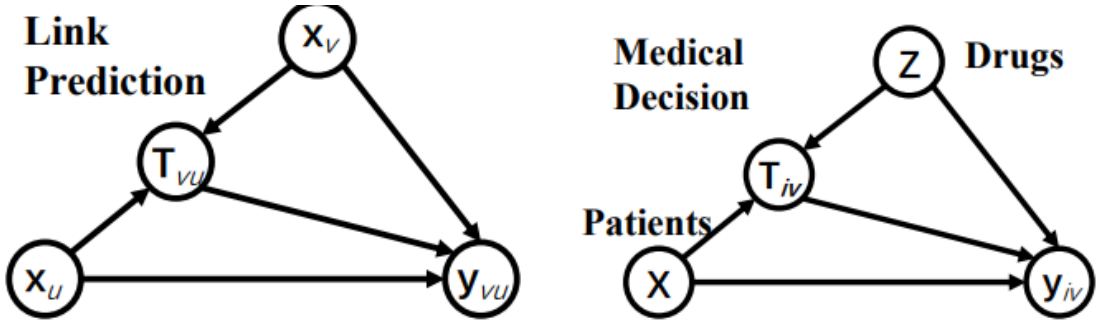
$$\hat{e}_{vu} = \mathbf{z}_v^{(t)\top}\mathbf{z}_u^{(t)}. \quad (5)$$

The MSE loss is defined as:

$$\mathcal{L}_M = \frac{1}{|E_{train}|}\sum_{v,u\in E_{train}}(\hat{e}_{vu} - e_{vu})^2, \quad (6)$$

Method





(a) Link prediction with causal model. The treatment $\mathbf{T}_{vu}$ is only related to the representations of node v and node u.

(b) Medical decision with causal model. The treatment $\mathbf{T}_{iv}$ is related to the representations of all the patients $\mathbf{X}$ and drugs $\mathbf{Z}$.

Fig. 5. Causal model improves patient and drug representation learning.

Medical Decision Module

$$\forall(S_i, D_v) \in S \times V$$

counterfactual link (S_j, D_u) is defined as

$$(S_j, D_u) = \underset{S_j \in S, D_u \in V}{\arg\min} \{\mathrm{dis}(\mathbf{x}_\mathrm{i}, \mathbf{x}_\mathrm{j}) + \mathrm{dis}(\mathbf{z}_\mathrm{v}, \mathbf{z}_\mathrm{u})| \tag{7}$$

$$\mathbf{T}_{ju} = 1 - \mathbf{T}_{iv}, \mathrm{dis}(\mathbf{x}_\mathrm{i}, \mathbf{x}_\mathrm{j}) < \gamma_\mathrm{p}, \mathrm{dis}(\mathbf{z}_\mathrm{v}, \mathbf{z}_\mathrm{u}) < \gamma_\mathrm{d}\},$$

the counterfactual treatment matrix $\mathbf{T}^{CF}$ and counterfactual adjacency matrix $\mathbf{Y}^{CF}$ as

$$\mathbf{T}_{iv}^{CF}, y_{iv}^{CF} = \begin{cases} 1 - \mathbf{T}_{iv}, y_{ju} & \text{, if } \exists(\mathrm{S_j}, \mathrm{D_u}) \in \mathrm{S} \times \mathrm{V} \\ & \text{satisfies Eq. (7);} \\ \mathbf{T}_{iv}, y_{iv} & \text{, otherwise.} \end{cases} \tag{8}$$

Method

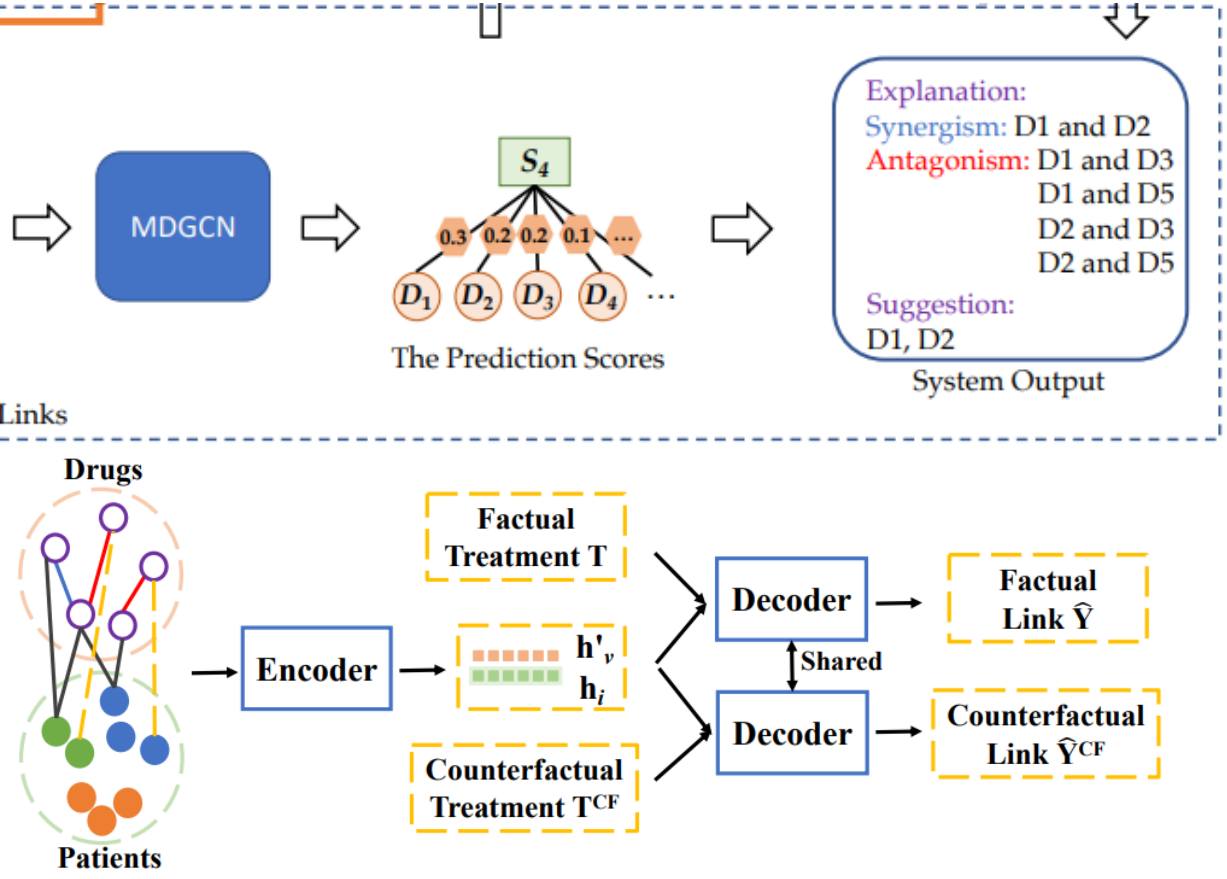


Fig. 6. Given the corresponding factual and counterfactual treatments, MDGCN is trained to predict factual and counterfactual links. The different colored patients indicate that they belong to different clusters. The red (blue) lines between drugs indicate antagonistic (synergistic) effects. The black (yellow) lines between patients and drugs indicate factual (counterfactual) links.

Medical Decision Module

$$\mathbf{h}_i = \sigma(\mathbf{W}_1\mathbf{x}_i + \mathbf{b}_1), i = 1, 2, \cdots, m, \tag{9}$$

$$\mathbf{h}_v = \sigma(\mathbf{W}_2\mathbf{z}_v + \mathbf{b}_2), v = 1, 2, \cdots, |V|, \tag{10}$$

The graph convolutional operation is defined as:

$$\mathbf{h}_i^{(t)} = \sum_{v \in \mathcal{N}_i} \frac{1}{\sqrt{|\mathcal{N}_i|}\sqrt{|\mathcal{N}_v|}} \mathbf{h}_v^{(t-1)}, \tag{11}$$

$$\mathbf{h}_v^{(t)} = \sum_{i \in \mathcal{N}_v} \frac{1}{\sqrt{|\mathcal{N}_v|}\sqrt{|\mathcal{N}_i|}} \mathbf{h}_i^{(t-1)}, \tag{12}$$

$$\mathbf{h}'_v = \sum_{t=0}^{T'} \beta_t \mathbf{h}_v^{(t)}, \tag{13}$$

$$\hat{y}_{iv} = f_{\mathbf{\Theta_2}}^{(t)}([\mathbf{h}_i^\top \odot \mathbf{h}'_v, \mathbf{T}_{iv}]), \tag{14}$$

$$\hat{y}_{iv}^{CF} = f_{\mathbf{\Theta_2}}^{(t)}([\mathbf{h}_i^\top \odot \mathbf{h}'_v, \mathbf{T}_{iv}^{CF}]), \tag{15}$$

Method

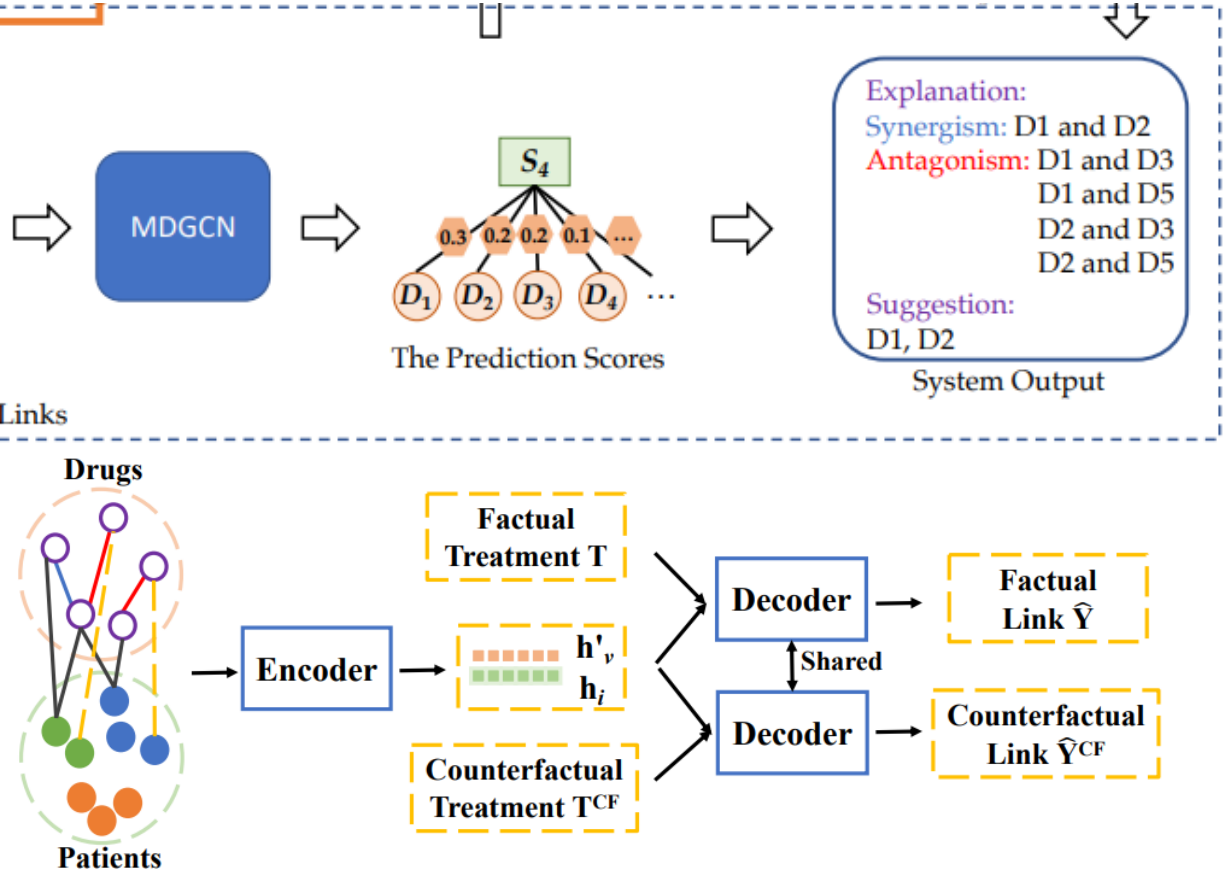


Fig. 6. Given the corresponding factual and counterfactual treatments, MDGCN is trained to predict factual and counterfactual links. The different colored patients indicate that they belong to different clusters. The red (blue) lines between drugs indicate antagonistic (synergistic) effects. The black (yellow) lines between patients and drugs indicate factual (counterfactual) links.

Medical Decision Module

$$\mathcal{L}_C = -\sum_{i}^{m}\sum_{v\in\mathcal{N}_i}[y_{iv}\log(\hat{y}_{iv}) + (1-y_{iv})\log(1-\hat{y}_{iv})], \quad (16)$$

$$\mathcal{L}_{CF} = -\sum_{i}^{m}\sum_{v\in\mathcal{N}_i}[y_{iv}^{CF}\log(\hat{y}_{iv}^{CF}) + (1-y_{iv}^{CF})\log(1-\hat{y}_{iv}^{CF})]. \quad (17)$$

The overall training loss of our model is:

$$\mathcal{L} = \mathcal{L}_C + \delta\mathcal{L}_{CF}, \quad (18)$$

Experiments

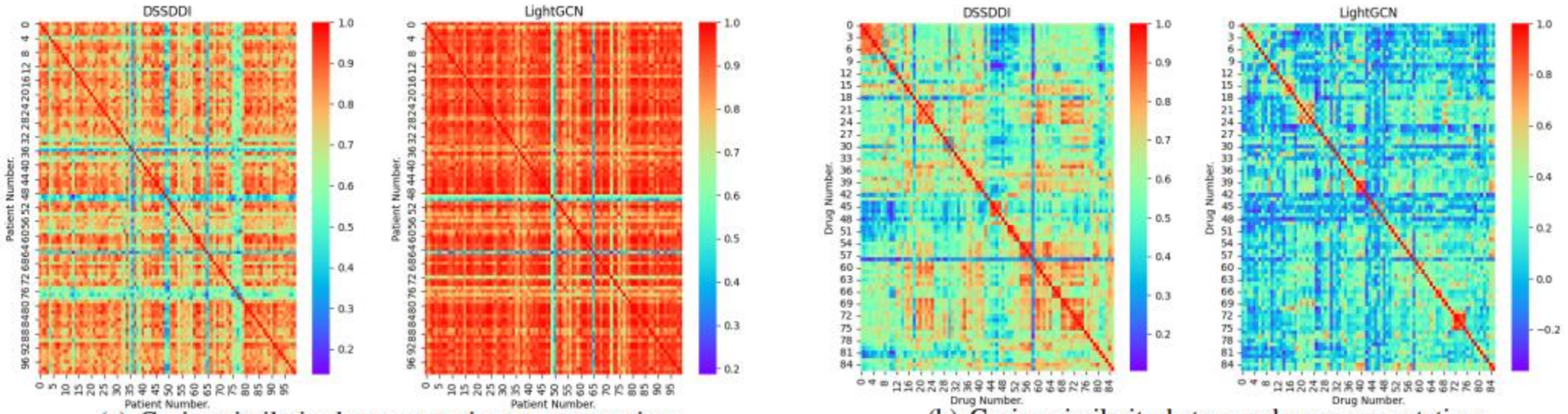
TABLE I

MEDICATION SUGGESTION PERFORMANCE COMPARISON BETWEEN THE PROPOSED METHOD AND BASELINE METHODS ON CHRONIC DATA SET (THE BEST RESULTS ARE IN BOLD AND THE SECOND RESULTS ARE UNDERLINED).

Method	Precision@6	Recall@6	NDCG@6	Precision@5	Recall@5	NDCG@5	Precision@4	Recall@4	NDCG@4
UserSim	0.0982	0.2227	0.1432	0.0971	0.2209	0.1426	0.0977	0.2181	0.1418
ECC	0.0214	0.0537	0.0328	0.0252	0.0519	0.0321	0.0060	0.0108	0.0127
SVM	0.0670	0.2166	0.2062	0.0635	0.1681	0.1847	0.0787	0.1653	0.1838
GCMC	0.1362	0.5310	0.3652	0.1447	0.4541	0.3181	0.1638	0.4057	0.3146
LightGCN	0.2073	0.7348	0.6012	0.2358	0.7245	0.5681	0.2581	0.6509	0.5436
SafeDrug	0.0863	0.3098	0.2267	0.1000	0.2952	0.2227	0.1250	0.2952	0.2233
Bipar-GCN	0.1741	0.6267	0.4817	0.1952	0.5911	0.4667	0.2172	0.5363	0.4418
CauseRec	0.1707	0.1025	0.5117	0.1124	0.4492	0.3030	**0.3186**	0.2468	0.1799
DSSDDI(SiGAT)	0.2214	0.8215	0.6482	0.2514	0.7834	0.6323	0.2876	0.7266	0.6076
DSSDDI(SNEA)	0.2192	0.7854	0.5949	0.2447	0.7364	0.5744	0.2740	0.6684	0.5442
DSSDDI(GIN)	0.2272	0.8407	0.6836	0.2534	0.8104	**0.6873**	0.2900	0.7704	0.6575
DSSDDI(SGCN)	**0.2348**	**0.8521**	**0.6850**	**0.2670**	**0.8153**	0.6717	0.3077	**0.7746**	**0.6680**

Method	Precision@3	Recall@3	NDCG@3	Precision@2	Recall@2	NDCG@2	Precision@1	Recall@1	NDCG@1
UserSim	0.0970	0.1970	0.1324	0.1370	0.1348	0.1033	0.0889	0.0088	0.0108
ECC	0.0072	0.0098	0.0123	0.0108	0.0098	0.0135	0.0192	0.0094	0.0204
SVM	0.1050	0.1649	0.1863	0.1575	0.1639	0.1970	0.3029	0.1552	0.2536
GCMC	0.1791	0.3437	0.2898	0.1971	0.2392	0.2406	0.1815	0.1428	0.2344
LightGCN	0.2925	0.5854	0.5436	0.3347	0.4021	0.4187	0.4231	0.2605	0.3786
SafeDrug	0.1206	0.2182	0.1872	0.1280	0.1536	0.1609	0.1406	0.0902	0.1406
Bipar-GCN	0.2484	0.4671	0.4118	0.2861	0.3672	0.3734	0.3377	0.2197	0.3377
CauseRec	0.2122	0.1595	0.1064	0.1665	0.1250	0.2494	0.1484	0.1484	0.1484
DSSDDI(SiGAT)	0.3361	0.6519	0.5745	0.3912	0.5261	0.5214	0.4531	0.3206	0.4531
DSSDDI(SNEA)	0.3133	0.5795	0.5059	0.3365	0.4242	0.4375	0.4038	0.2667	0.4038
DSSDDI(GIN)	0.3554	0.6918	0.6256	0.4261	**0.5926**	0.5842	0.4916	**0.3989**	**0.5565**
DSSDDI(SGCN)	**0.3670**	**0.7027**	**0.6378**	**0.4297**	0.5903	**0.5933**	**0.5300**	0.3743	0.5300

Experiments



(a) Cosine similarity between patient representations.

(b) Cosine similarity between drug representations.

Fig. 7. Comparison of cosine similarity between patient (drug) representations obtained from DSSDDI and LightGCN. Closer to red means that the representations are more similar and closer to blue means that they are less similar. (a) We sample 100 patients in the test set of patients. (b) 86 drugs are included in the comparison of their similarity to each other.

Experiments

TABLE II

ABLATION STUDIES WITH DIFFERENT DRUG EMBEDDINGS ON CHRONIC DATA SET (THE BEST RESULTS ARE IN BOLD). HERE, WE USE SGCN, THE BEST PERFORMING BACKBONE MODEL IN TABLE I, AS THE BACKBONE IN DDIGCN.

Method	Precision@6	Recall@6	NDCG@6	Precision@5	Recall@5	NDCG@5	Precision@4	Recall@4	NDCG@4
w/o DDI	0.2185	0.7974	0.6427	0.2490	0.7694	0.6301	0.2891	0.7256	0.6089
One-hot	0.2095	0.7952	0.6063	0.2365	0.7537	0.5891	0.2638	0.6830	0.5574
KG	0.2135	0.8170	0.6489	0.2411	0.7761	0.6319	0.2758	0.7187	0.6067
DDIGCN	**0.2348**	**0.8521**	**0.6850**	**0.2670**	**0.8153**	**0.6717**	**0.3077**	**0.7746**	**0.6680**

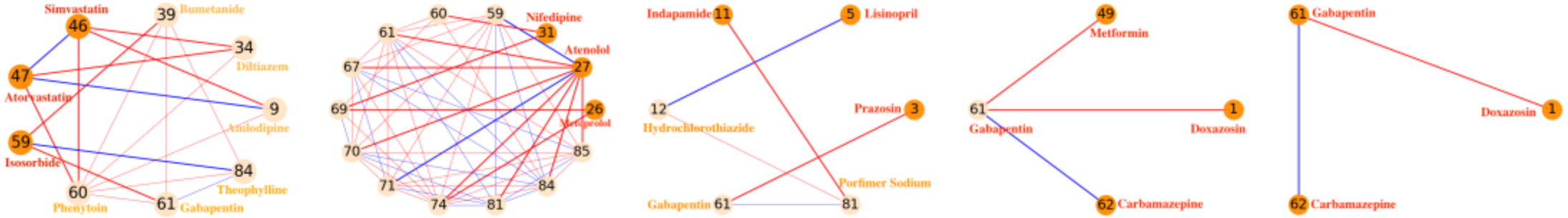
Method	Precision@3	Recall@3	NDCG@3	Precision@2	Recall@2	NDCG@2	Precision@1	Recall@1	NDCG@1
w/o DDI	0.3413	0.6499	0.5788	0.4032	0.5277	0.5292	0.4796	0.3204	0.4796
One-hot	0.2984	0.5904	0.5150	0.3462	0.4715	0.4635	0.4147	0.2855	0.4147
KG	0.3297	0.6609	0.5810	0.3918	0.5425	0.5304	0.4591	0.3338	0.4591
DDIGCN	**0.3670**	**0.7027**	**0.6378**	**0.4297**	**0.5903**	**0.5933**	**0.5300**	**0.3743**	**0.5300**

Experiments

TABLE III
SUGGESTION SATISFACTION COMPARISON BETWEEN THE PROPOSED METHOD AND BASELINE METHODS ON k MEDICATION SUGGESTIONS.

k	2	3	4	5	6
UserSim	0.4987	0.2506	0.0743	0.0470	0.0220
ECC	0.5000	0.2500	0.0952	0.0455	0.0339
SVM	0.5044	0.2695	0.1050	0.0469	0.0278
GCMC	0.4979	0.2533	0.1443	0.0491	0.0634
LightGCN	0.5046	0.2544	0.1631	0.0882	0.0575
SafeDrug	0.4412	0.1812	0.0741	0.0477	0.0329
Bipar-GCN	0.5139	0.2788	0.1735	0.1183	0.0866
CauseRec	0.4957	0.2462	0.0996	0.0482	0.0299
DSSDDI(SiGAT)	**0.5683**	0.3237	0.2043	0.1400	0.1042
DSSDDI(SNEA)	0.5522	0.2916	0.1811	0.1253	0.0899
DSSDDI(GIN)	0.5392	0.2767	0.1743	0.1227	0.0997
DSSDDI(SGCN)	0.5427	**0.3267**	**0.2158**	**0.1478**	**0.1083**

Experiments



(a) DSSDDI (b) LightGCN (c) GCMC (d) SVM (e) ECC

Fig. 8. A case of medication suggestion for a patient with cardiovascular disease. We use dark orange nodes to denote suggested drugs and light orange nodes to denote non-suggested nodes in the subgraph output by the MS module, red lines to indicate the antagonism between drugs, and blue lines to indicate the synergy between drugs. We make the edges between the non-suggested drugs transparent to highlight the interactions associated with the suggested drugs.

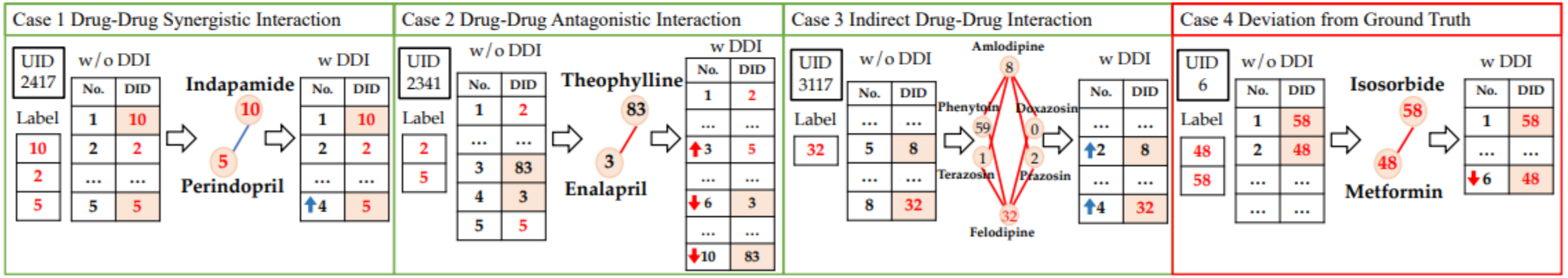


Fig. 9. Four case studies that demonstrate the superiority of DDI. The red numbers in the Label column indicate the drugs the patient took, the blue (red) line indicates the drug-drug synergistic (antagonistic) effect, the blue (red) arrow indicates an upward or downward movement brought by the synergistic (antagonistic) effect.

Experiments

TABLE IV
MEDICATION SUGGESTION PERFORMANCE COMPARISON BETWEEN THE PROPOSED METHOD AND BASELINE METHODS ON MIMIC-III DATA SET (THE BEST RESULTS ARE IN BOLD).

Method	Precision@8	Recall@8	NDCG@8	Precision@6	Recall@6	NDCG@6	Precision@4	Recall@4	NDCG@4
UserSim	0.5396	0.2349	0.6203	0.5699	0.1869	0.6534	0.7006	0.1551	0.7557
ECC	0.6957	0.3014	0.7360	0.7786	0.2562	0.7904	0.8116	0.1795	0.8111
SVM	0.7645	0.3343	0.7829	0.8234	0.2703	0.8182	0.8313	0.1833	0.8210
GCMC	0.7924	0.3283	0.8156	0.8186	0.2656	0.8208	0.8360	0.1764	0.8392
LightGCN	0.8099	0.3548	0.8252	0.8310	0.2738	0.8378	0.8449	0.1872	0.8495
SafeDrug	0.8038	0.3549	0.8215	0.8172	0.2701	0.8308	0.8434	0.1869	0.8494
Bipar-GCN	0.7939	0.3510	0.8135	0.8172	0.2718	0.8281	0.8327	0.1856	0.8390
CauseRec	0.1218	0.1428	0.1346	0.1196	0.1130	0.1238	0.1157	0.0818	0.1162
DSSDDI(GIN)	**0.8134**	**0.3611**	**0.8266**	**0.8352**	**0.2808**	**0.8408**	**0.8530**	**0.1932**	**0.8553**

Thank you!