

Multi-view Substructure Learning for Drug-Drug Interaction Prediction

Table S1: The statistics of the datasets.

Datasets	Drugs	DDIs	DDI types	Train DDIs	Valid DDIs	Test DDIs
DrugBank	1,706	191,808	86	115084	38362	38362
Twosides	645	4,576,287	963	2745773	915257	915257

Table S2: Grid-search Space for hyper-parameter settings.

Hyper-parameters	Search Space
The number of MSN encoder	3, 4, 5
Learning rate	0.01, 0.001, 0.005
Batch size	512, 1024, 2048, 4096
The number of multi-head	1, 2, 4

Table S3: Performance evaluation between MSN-DDI and its five variants on DrugBank dataset in transductive setting. The highest value in each column is shown in bold.

variant	ACC(%)	AUROC(%)	AP(%)	P(%)	R(%)
wo_inter	94.64	98.34	98.07	94.08	95.26
wo_intra	94.56	98.50	98.29	93.81	95.42
wo_update	95.18	98.63	98.50	95.04	95.33
wo_SAGPool	95.62	98.84	98.97	96.14	95.06
wo_co-attention	95.07	98.69	98.49	94.69	95.51
MSN-DDI	96.94	99.47	99.37	97.14	96.73

Table S4: Performance evaluation between MSN-DDI and its five variants on Twosides dataset in transductive setting. The highest value in each column is shown in bold.

variant	ACC(%)	AUROC(%)	AP(%)	P(%)	R(%)
wo_inter	80.69	88.04	84.96	77.35	86.81
wo_intra	97.89	99.77	99.75	97.51	98.29
wo_update	85.37	92.66	91.15	83.22	88.60
wo_SAGPool	98.66	99.89	99.87	98.52	98.81
wo_co-attention	97.92	99.71	99.67	97.79	98.05
MSN-DDI	98.83	99.90	99.89	98.61	99.05

Table S5: The statistics of the drugbank dataset for new drug application

Drug split	Drugs	DDIs
Old drugs before 2017	1681	188341
New drugs after 2017	25	3467
Total	1706	191808

Table S6: Performance evaluation between MSN-DDI,SSI-DDI,GMPNN-CS on Drugbank dataset for new drug application. The highest value in each column is shown in bold.

Method	ACC(%)	AUROC(%)	AP(%)	F1(%)
SSI-DDI	67.85	74.34	73.18	68.32
GMPNN-CS	68.10	77.62	73.64	64.25
MSN-DDI	74.07	81.95	78.95	77.22