Supplementary Material for "MVA-DDI: Interpretable attention network with multi-view learning for drug-drug interaction prediction"

1 Metrics

ACC represents the ratio of correctly predicted samples to the total number of samples. Specifically, ACC can be defined as follows:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

where TP (True Positive) and TN (True Negative) represent the numbers of correctly predicted positive and negative instances, respectively, while FP (False Positive) and FN (False Negative) represent the numbers of incorrectly predicted positive and negative instances. A higher accuracy indicates a more precise prediction by the classifier.

AUROC reflects the probability of the model correctly distinguishing between positive and negative samples. It is calculated as the area under the ROC curve, where the horizontal axis represents the False Positive Rate (FPR) and the vertical axis represents the True Positive Rate (TPR). A larger AUROC value indicates better performance of the model. FPR and TPR can be calculated as follows:

$$FPR = \frac{FP}{FP + TN} \tag{2}$$

$$TPR = \frac{TP}{TP + FN} \tag{3}$$

AUPR is commonly used to evaluate the performance of classifiers, especially in cases where the dataset is extremely imbalanced. It measures the area under the Precision-Recall curve as an evaluation metric to describe the performance of the classifier at different recall levels. Precision and recall can be calculated as follows:

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

The F1 score is an evaluation metric that combines the analysis of a model's precision and recall values. The core idea of F1 is to maximize both precision and recall while minimizing the difference between them.

$$F1 - score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{6}$$

2 Interpretable analysis

In the MVA-DDI experiments, we utilized attention-based fusion to combine two channels: sequence features and graph features. We encoded the drug feature representations from two different perspectives and then fused them using the attention mechanism to leverage their complementarity and interaction. This approach allows us to capture different aspects of drug information and improve the prediction of drug interactions.

We utilized violin plots to display the weights of sequence features and graph features for four selected drug classes in the MVA-DDI network, as shown in Figure 1. We randomly chose three specific drugs for each of the Antioxidants, Hormone Antagonists, Anilides, and Enzyme Inhibitors classes. The results indicate that drugs in the Antioxidants and Enzyme Inhibitors classes have a higher proportion of weights on sequence features compared to graph features, indicating that sequence features are more characteristic for these classes. On the other hand, Hormone Antagonists and Anilides have a proportion of weights on graph features that is mostly above 0.5, suggesting that graph features are more indicative for these classes.

The attention mechanism provides feature weights for drugs from different perspectives, allowing us to understand the basis of model predictions. It also helps identify important features by emphasizing weighted attention on different feature encodings. This can assist drug researchers in better understanding drug interactions and discovering new associations. Adopting attention-based fusion for dual-channel feature integration can enhance model performance and provide interpretability and transparency in DDI prediction. This approach can also be extended to feature fusion and prediction in other areas of bioinformatics and medical data.

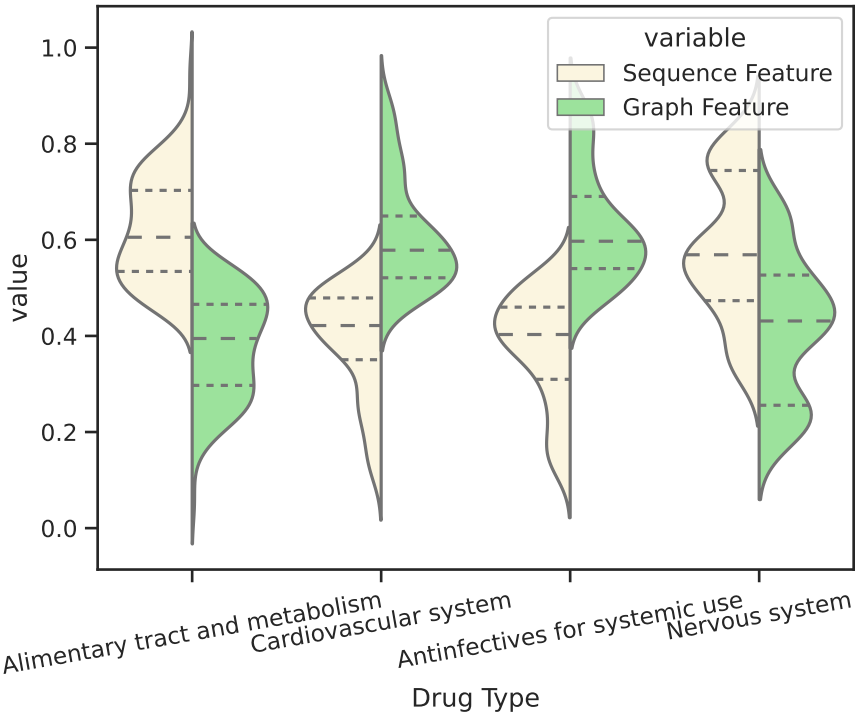

Figure 1: Violin weight analysis. We implemented violin plots using the SNS tool to visualize feature weights. These plots allow for a clear analysis of the performance of different drug classes in the dual-channel setting. In these plots, the cornsilk color represents the weight of sequence features, while green represents the weight of graph features.

3 Case study

To validate the effectiveness of the MVA-DDI method, we used two drugs for the treatment of cardiovascular and cerebrovascular diseases and COVID-19 (Arbutamine and Favipiravir). Table 1 and Table 2 respectively report the top 50 drugs predicted to have DDI with Arbutamine and Favipiravir.

Table 1: The top 50 predicted Arbutamine-associated drugs.

Rank	DrugBank ID	Drug Name	Verification Method
1	DB00598	Labetalol	DrugBank,PubChem
2	DB01037	Selegiline	DrugBank,PubChem
3	DB00612	Bisoprolol	DrugBank,PubChem
4	DB08807	Bopindolol	DrugBank,PubChem
5	DB09289	Tianeptine	DrugBank,PubChem
6	DB04861	Nebivolol	DrugBank,PubChem
7	DB00866	Alprenolol	DrugBank,PubChem
8	DB00805	Minaprine	DrugBank,PubChem
9	DB09245	Toloxatone	DrugBank,PubChem
10	DB00264	Metoprolol	DrugBank,PubChem
11	DB01580	Oxprenolol	DrugBank,PubChem
12	DB08808	Bupranolol	DrugBank,PubChem
13	DB01162	Terazosin	DrugBank,PubChem
14	DB00696	Ergotamine	DrugBank,PubChem
15	DB00590	Doxazosin	DrugBank,PubChem
16	DB01392	Yohimbine	DrugBank,PubChem
17	DB00734	Risperidone	DrugBank,PubChem
18	DB00818	Propofol	DrugBank,PubChem
19	DB00601	Linezolid	DrugBank,PubChem
20	DB00741	Hydrocortisone	DrugBank,PubChem
21	DB01338	Pipecuronium	DrugBank,PubChem
22	DB00543	Amoxapine	DrugBank,PubChem
23	DB01216	Finasteride	DrugBank,PubChem
24	DB00780	Phenelzine	DrugBank,PubChem
25	DB00788	Naproxen	DrugBank,PubChem
26	DB00420	Promazine	DrugBank,PubChem
27	DB00749	Etodolac	DrugBank,PubChem
28	DB00091	Cyclosporine	DrugBank,PubChem
29	DB01348	Spirapril	DrugBank,PubChem
30	DB06698	Betahistine	DrugBank,PubChem
31	DB01224	Quetiapine	DrugBank,PubChem
32	DB00159	Icosapent	DrugBank,PubChem
33	DB01009	Ketoprofen	DrugBank,PubChem
34	DB01244	Bepridil	DrugBank,PubChem
35	DB00810	Biperiden	DrugBank,PubChem
36	DB13781	Xamoterol	DrugBank,PubChem
37	DB00620	Triamcinolone	DrugBank,PubChem
38	DB06654	Safinamide	DrugBank,PubChem
39	DB00482	Celecoxib	DrugBank,PubChem
40	DB01021	Trichlormethiazide	DrugBank,PubChem
41	DB00226	Guanadrel	DrugBank,PubChem
42	DB00572	Atropine	DrugBank,PubChem
43	DB00354	Buclizine	DrugBank,PubChem
44	DB02925	Piretanide	DrugBank,PubChem
45	DB01108	Trilostane	DrugBank,PubChem
46	DB00939	Meclofenamic acid	DrugBank,PubChem
47	DB00821	Carprofen	DrugBank,PubChem
48	DB00376	Trihexyphenidyl	DrugBank,PubChem
49	DB00661	Verapamil	DrugBank,PubChem
50	DB01336	Metocurine	DrugBank,PubChem

Table 2: The top 50 predicted Favipiravir-associated drugs.

Rank	DrugBank ID	Drug Name	Verification Method
1	DB00227	Lovastatin	DrugBank,PubChem
2	DB11753	Rifamycin	DrugBank,PubChem
3	DB14055	(S)-Warfarin	DrugBank,PubChem
4	DB08496	(R)-warfarin	DrugBank,PubChem
5	DB11362	Selexipag	DrugBank,PubChem
6	DB00468	Quinine	DrugBank,PubChem
7	DB00900	Didanosine	DrugBank,PubChem
8	DB04951	Pirfenidone	DrugBank
9	DB00203	Sildenafil	DrugBank,PubChem
10	DB00316	Acetaminophen	DrugBank
11	DB08816	Ticagrelor	DrugBank,PubChem
12	DB09080	Olodaterol	DrugBank,PubChem
13	DB06772	Cabazitaxel	DrugBank,PubChem
14	DB00659	Acamprosate	DrugBank
15	DB00495	Zidovudine	DrugBank,PubChem
16	DB12001	Abemaciclib	DrugBank
17	DB06403	Ambrisentan	DrugBank,PubChem
18	DB05015	Belinostat	DrugBank,PubChem
19	DB11712	Tezacaftor	DrugBank,PubChem
20	DB08893	Mirabegron	DrugBank,PubChem
21	DB00741	Hydrocortisone	DrugBank,PubChem
22	DB01263	Posaconazole	DrugBank,PubChem
23	DB00818	Propofol	DrugBank,PubChem
24	DB00359	Sulfadiazine	DrugBank,PubChem
25	DB00759	Tetracycline	DrugBank,PubChem
26	DB01053	Benzylpenicillin	DrugBank,PubChem
27	DB01414	Cefacetrile	DrugBank,PubChem
28	DB01234	Dexamethasone	DrugBank,PubChem
29	DB00228	Enflurane	DrugBank
30	DB13956	Estradiol valerate	DrugBank,PubChem
31	DB01229	Paclitaxel	DrugBank,PubChem
32	DB09068	Vortioxetine	DrugBank,PubChem
33	DB11693	Voclosporin	DrugBank
34	DB08907	Canagliflozin	DrugBank,PubChem
35	DB01197	Captopril	DrugBank,PubChem
36	DB11682	Daprodustat	DrugBank
37	DB01415	Ceftibuten	DrugBank,PubChem
38	DB11652	Tucatinib	DrugBank,PubChem
39	DB00347	Trimethadione	DrugBank,PubChem
40	DB00857	Terbinafine	DrugBank,PubChem
41	DB09256	Tegafur	DrugBank,PubChem
42	DB00139	Succinic acid	DrugBank,PubChem
43	DB12026	Voxilaprevir	DrugBank,PubChem
44	DB00586	Diclofenac	DrugBank,PubChem
45	DB00783	Estradiol	DrugBank,PubChem
46	DB00250	Dapsone	DrugBank,PubChem
47	DB12836	Grapiprant	DrugBank,PubChem
48	DB00398	Sorafenib	DrugBank,PubChem
49	DB15822	Pralsetinib	DrugBank
50	DB06616	Bosutinib	DrugBank,PubChem