Supplementary materials for Multi-view Contrastive Learning Hypergraph Neural Network for Drug-Microbe-Disease Association Prediction

1 The distribution of data node degrees

We conducted a statistical analysis of node degree about the data set, the abscissa represents the degree interval of the node, and the ordinate represents how many nodes are located in this degree interval. As can be seen from Figure 1, the degree of most nodes is relatively low, for example, 274 nodes have a degree between [1,10], 131 nodes have a degree between [11,20], and only about 22 nodes have a degree greater than 60, indicating that the correlation data of drug-microbe-disease are generally sparse.

2 Implementation of the baseline methods

All the baseline methods were implemented by their publicly available source codes. The best or default parameters of each method were used. For the parameters of CoSTCo, NeurTN, and HypergraphSynergy, we preserved the settings provided in original papers; for CP, Tucker, RF, MLP, and GCN, we carefully tuned them to achieve their optimal performance. Notably, for the method of GCN, we broke the positive DMD triple-wise associations into two pair-wise associations (i.e. drug-microbe and microbe-disease association), then we adopted negative sampling strategies for the two pair-wise sets in the same way as in the MCHNN, generating the same proportion of negative samples, and we modeled them using two GCNs separately. Finally, the two sets of pair-wise predicted values were multiplied as the final DMD triple-wise association prediction value.

3 Parameter settings for MCHNN

Our model is composed of GIN for the drug features, the FCN for the microbe and disease features, the HGNN, the contrastive learning, and the scoring function. The hyper-parameter α, that balances the contributions of the prediction task and the contrastive learning task, and the learning rate lr need to be tuned within a fixed range. Here, we choose α from $\{0.1, 0.2, ..., 0.9, 1.0\}$ and lr from $\{0.0001, 0.001, 0.005, 0.01, 0.1\}$ on the 5-CV to build MCHNN, respectively. According to the results (Figure 2), we can find when α was fixed to 0.8, and the lr was fixed to 0.005, the MCHNN can produce the best performance.

4 Performance of independent test set

The independent test results of MCHNN and baselines under the four scenarios are shown in Figure 3.

5 Details and results of t-SNE visualization analysis

We used t-sne visualization to verify the quality of embeddings learned by the models on the independent test set. Among all methods, only MCHNN, HypergraphSynergy, and NeurTN can be extracted the learned embeddings, so we only choose the embeddings learned by these three methods for visualization. In addition, the ratio of positive and negative samples in our test set is 1:29, in order to reflect the ability of the model to distinguish positive and negative samples more clearly, we use a random sample strategy to extract the embeddings of 1:1 positive and negative samples for visualization, and the results are shown in Figure 4.

6 Other results were placed in Supplementary Excel tables

Due to the space limitation of the IJCAI paper, we put the remaining relevant results in the supplementary Excel tables.

- "compare_methods" includes the results of the numerical values corresponding to the 5-fold mean values, 5-fold standard deviation, 5-fold P value, and independent test. It is worth noting that the values of AUC, AUPR, and ACC under scenario 4 (random negative sampling) are also included.
- "ablation_results" includes the results of the numerical values corresponding to the ablation study results.
- "computational_performance" includes the results of running time and memory usage based on MCHNN and other compared models.

We run MCHNN and other compared methods on our workstation with 64 Intel(R) Xeon(R) Gold 6146 3.20GHZ CPUs.

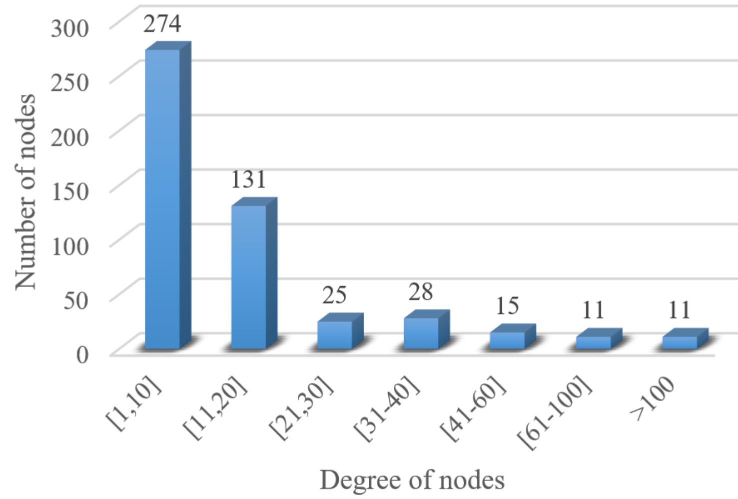

Figure 1: The distribution of data node degrees.

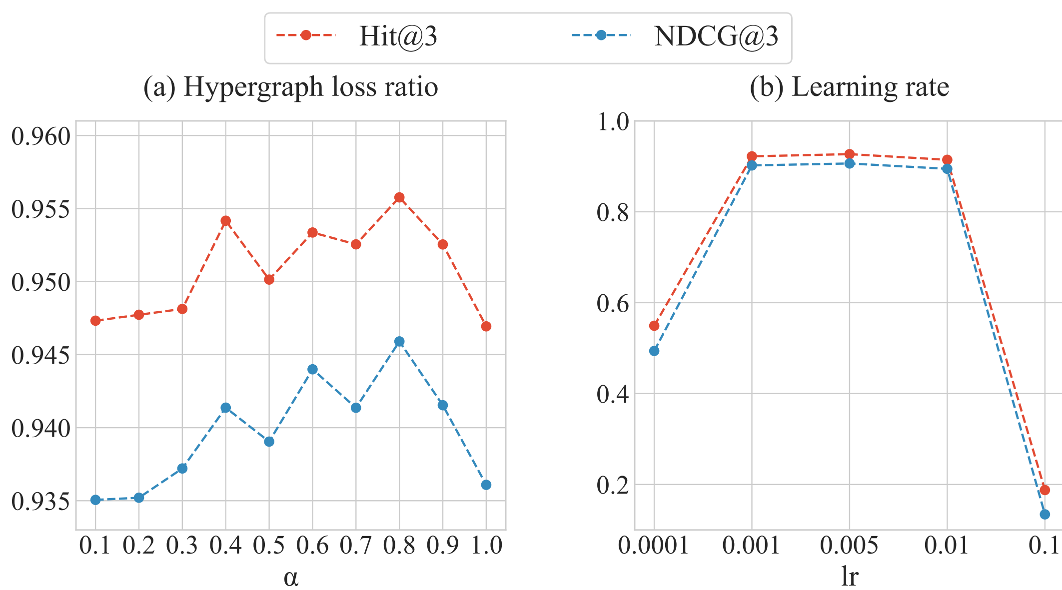

Figure 2: The performance of MCHNN on 5-CV under scenario 4 with different hyper-parameters. (a) The impact of the hypergraph loss ratio α. (b) The impact of the learning rate lr.

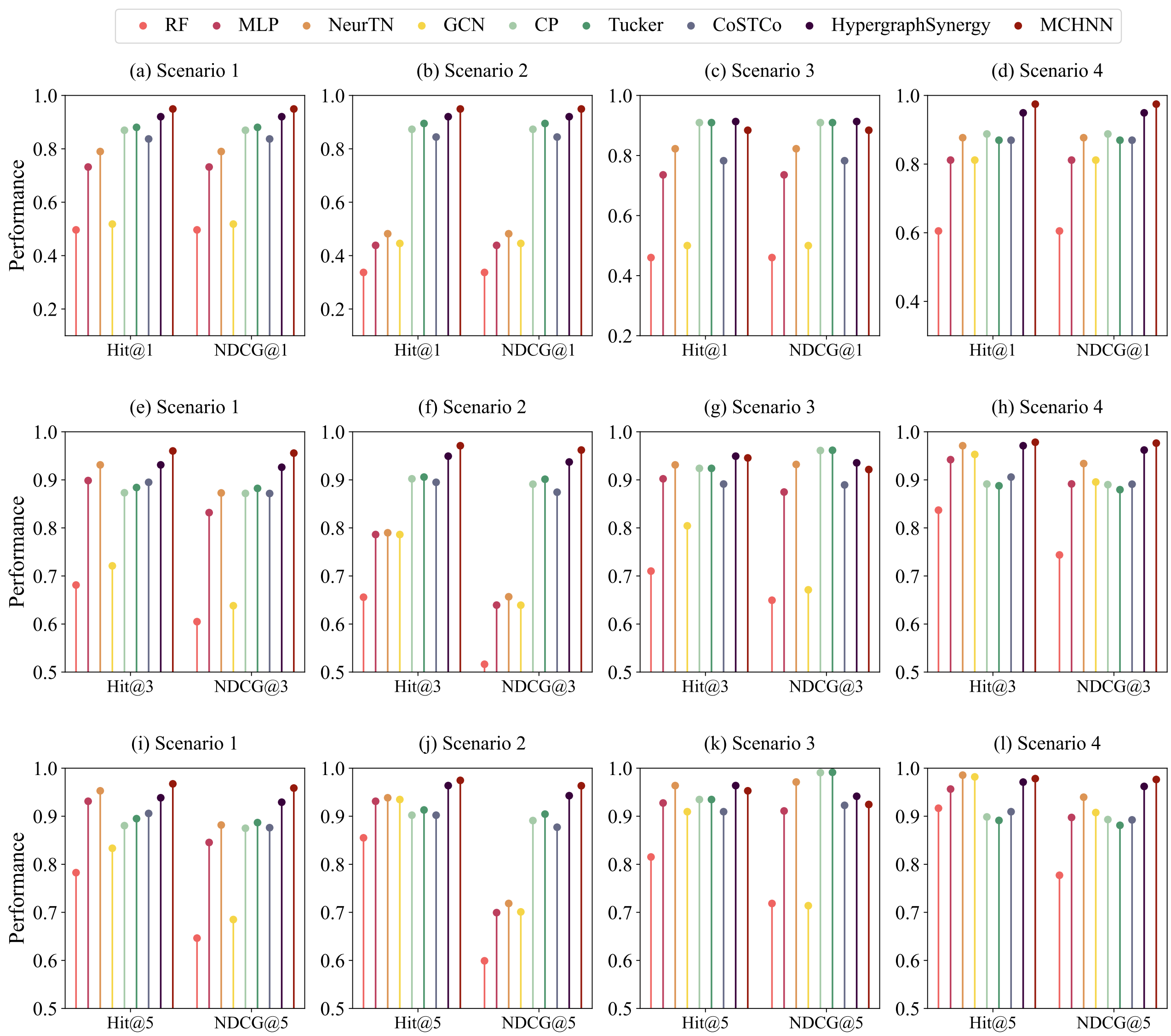

Figure 3: Independent test performance of MCHNN and baselines in four scenarios in terms of Hit@1 and NDCG@1(a-d), Hit@3 and NDCG@3(e-h), Hit@5 and NDCG@5(i-l).

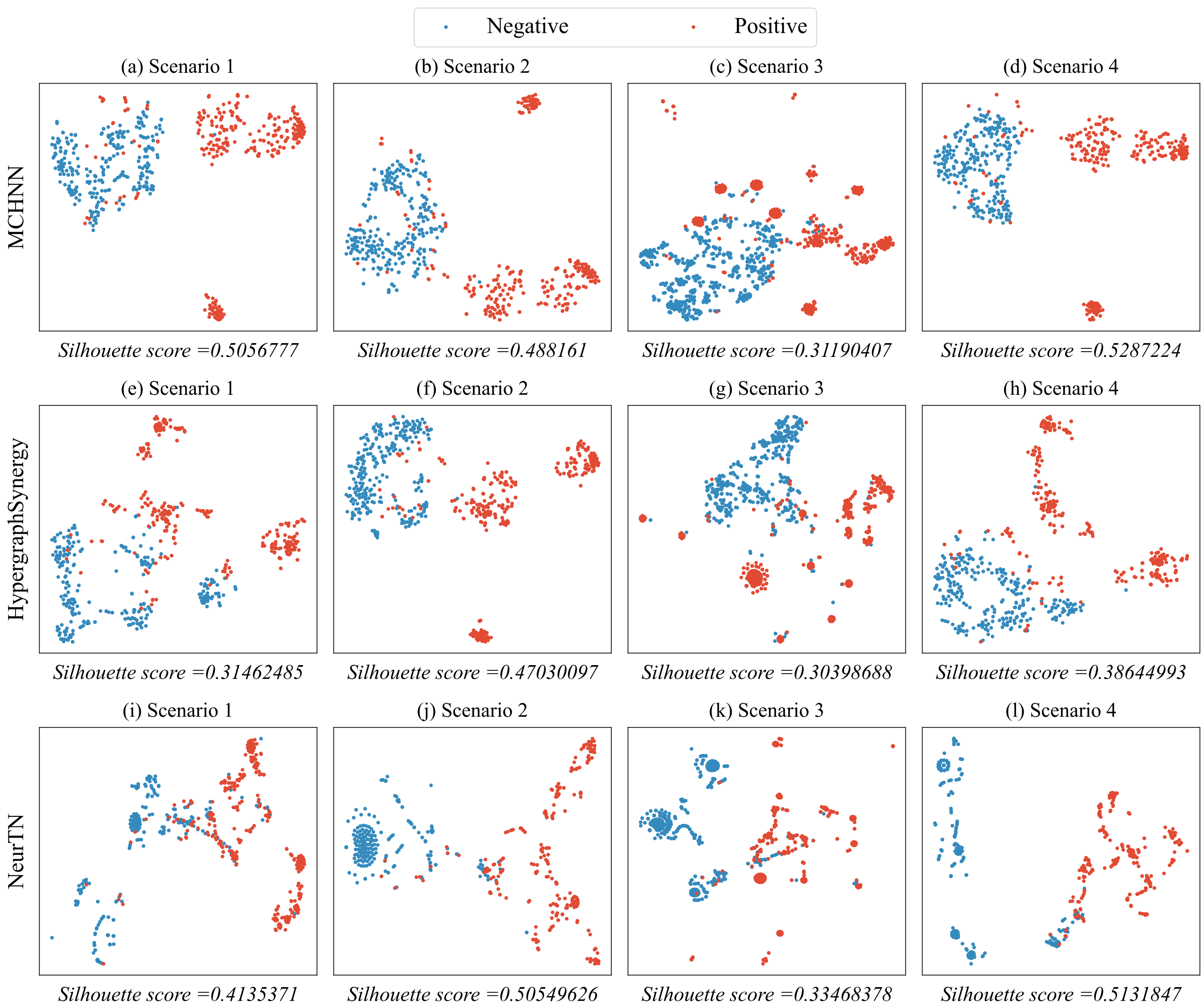

Figure 4: The t-SNE visualization of three models on the four scenarios, MCHNN(a-d), HypergraphSynergy(e-h), NeurTN(i-l).