**LncRNA-MiRNA Interaction Prediction Based on Multi-source Heterogeneous Graph Neural Network and Multi-level Attention Mechanism**

Ziyu Li 1, Kaibo Li1, Xuequan Lian1 and Jianwei Li 1,2\*

1 School of Artificial Intelligence, Hebei University of Technology, Tianjin 300401, China

2 Hebei Provincial Key Laboratory of Big Data Computing, Tianjin 300401, China

\* Correspondence: [lijianwei@hebut.edu.cn](mailto:lijianwei@hebut.edu.cn)

# **Baselines**

**RNAI-FRID** [1]**:** This is a feature representation method for RNA–RNA interactions that enhances information and reduces dimensionality by constructing arithmetic-based complex features from diverse base features, combined with feature ranking for optimal selection.

**MGCAT**[2]: This model leverages the graph attention network (GAT) model, integrating a multilayered attention mechanism to dynamically capture the topological features of the lncRNA-miRNA heterogeneous network

**CSGLMD**[3]: This model processes the intra-layer similarity matrix of lncRNA-miRNA-disease heterogeneous graph using a label instantiation mechanism, incorporates the contrastive self-supervised learning task, and jointly optimizes supervised and self-supervised objectives

**GCLMTP**[4]: A multi-task prediction method based on self-supervised learning, designed to simultaneously extract node embeddings from lncRNA-miRNA-disease heterogeneous networks using graph contrastive learning. It predicts LDAs, MDAs and LMIs using multiple classifiers.

**SSCLMD**[5]: A multi-task prediction model that introduces disease-specific attribute and topology graphs, utilizing specialized and general encoders, self-supervised contrastive learning, and view-level attention mechanisms for multi-task prediction

# **Analysis of hyper-parameters**

**Analysis of *k*-nearest neighbor graph**

The parameter of the KNN algorithm was selected from the set [6,8,10,12,15,18,20,25,30]. Through experimental verification, it is found that when exceeds 10, the model prediction effect basically does not change, sometimes there will be a small decrease, but the correlation edge in the data increases, and the model training time increases. So, we end up choosing =10.

**Table 1.** Analysis of *k*-nearest neighbor graph

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | AUC | AUPR | ACC | PRE | REC | F1 |
| 6 | 0.9224 | 0.9316 | 0.8575 | 0.8592 | **0.8551** | 0.8571 |
| 8 | 0.9252 | 0.9330 | 0.8598 | 0.8636 | 0.8551 | 0.8592 |
| **10** | **0.9283** | **0.9338** | **0.8738** | **0.9040** | 0.8364 | **0.8689** |
| 12 | 0.9280 | 0.9318 | 0.8618 | 0.8790 | 0.8314 | 0.8537 |
| 15 | 0.9280 | 0.9318 | 0.8618 | 0.8790 | 0.8314 | 0.8537 |
| 18 | 0.9252 | 0.9330 | 0.8598 | 0.8632 | 0.8551 | 0.8592 |
| 20 | 0.9252 | 0.9330 | 0.8598 | 0.8636 | 0.8551 | 0.8592 |
| 25 | 0.9252 | 0.9330 | 0.8598 | 0.8636 | 0.8551 | 0.8592 |
| 30 | 0.9252 | 0.9329 | 0.8621 | 0.8673 | 0.8551 | 0.8612 |

**Analysis of** **learning rate**

The learning rate was selected via grid search from the sets [10−5,5×10−5,10−4,5×10−4,10−3,5×10−3]. After evaluating model performance under each setting, 10−4 achieved the most stable and superior results, and was subsequently adopted for the final training.

**Table 2.** Analysis of learning rate

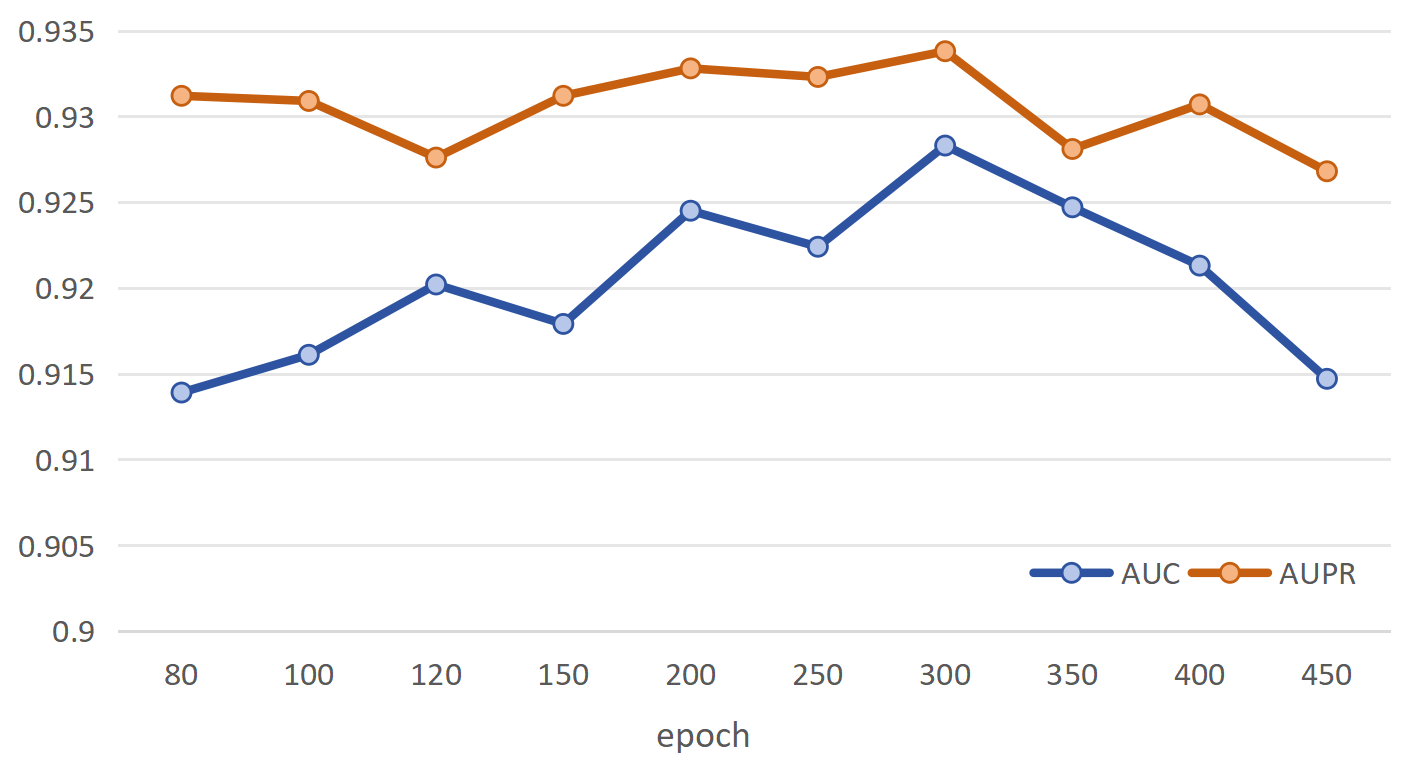
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | AUC | AUPR | ACC | PRE | REC | F1 |
| 10−5 | 0.8520 | 0.8870 | 0.7874 | 0.9489 | 0.6075 | 0.7407 |
| 5×10−5 | 0.9102 | 0.9264 | 0.8271 | 0.8333 | 0.8178 | 0.8255 |
| 10−4 | **0.9283** | **0.9338** | **0.8738** | **0.9040** | 0.8364 | **0.8689** |
| 5×10−4 | 0.9252 | 0.9307 | 0.8551 | 0.8486 | 0.8645 | 0.8565 |
| 10−3 | 0.9228 | 0.9293 | 0.8411 | 0.8230 | **0.8692** | 0.8455 |
| 5×10−3 | 0.9255 | 0.9321 | 0.8551 | 0.8519 | 0.8598 | 0.8558 |

**Analysis of the number of training epoch**

The number of training epoch was selected from the set [80,100,120,150,200,250,300,350,400,450]. The epoch number controls the training iterations, and selecting a suitable value helps balance sufficient learning and overfitting, ensuring good generalization. After experimental validation, 200 epochs were chosen for the final model.

**Table 3.** Analysis of the number of training epoch

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| epoch | AUC | AUPR | ACC | PRE | REC | F1 |
| 80 | 0.9139 | 0.9312 | 0.8458 | 0.8585 | 0.8458 | 0.8458 |
| 100 | 0.9161 | 0.9309 | 0.8435 | 0.8551 | 0.8224 | 0.8409 |
| 120 | 0.9202 | 0.9346 | 0.8411 | 0.8411 | 0.8271 | 0.8411 |
| 150 | 0.9179 | 0.9312 | 0.8341 | 0.8389 | 0.8271 | 0.8329 |
| 200 | 0.9245 | 0.9328 | 0.8575 | 0.8592 | 0.8551 | 0.8571 |
| 250 | 0.9224 | 0.9323 | 0.8435 | 0.8451 | 0.8381 | 0.8431 |
| 300 | **0.9283** | **0.9338** | **0.8738** | **0.9040** | 0.8364 | **0.8689** |
| 350 | 0.9247 | 0.9281 | 0.8679 | 0.8679 | **0.8679** | 0.8679 |
| 400 | 0.9213 | 0.9307 | 0.8435 | 0.8483 | 0.8364 | 0.8424 |
| 450 | 0.9147 | 0.9268 | 0.8341 | 0.8326 | 0.8364 | 0.8345 |



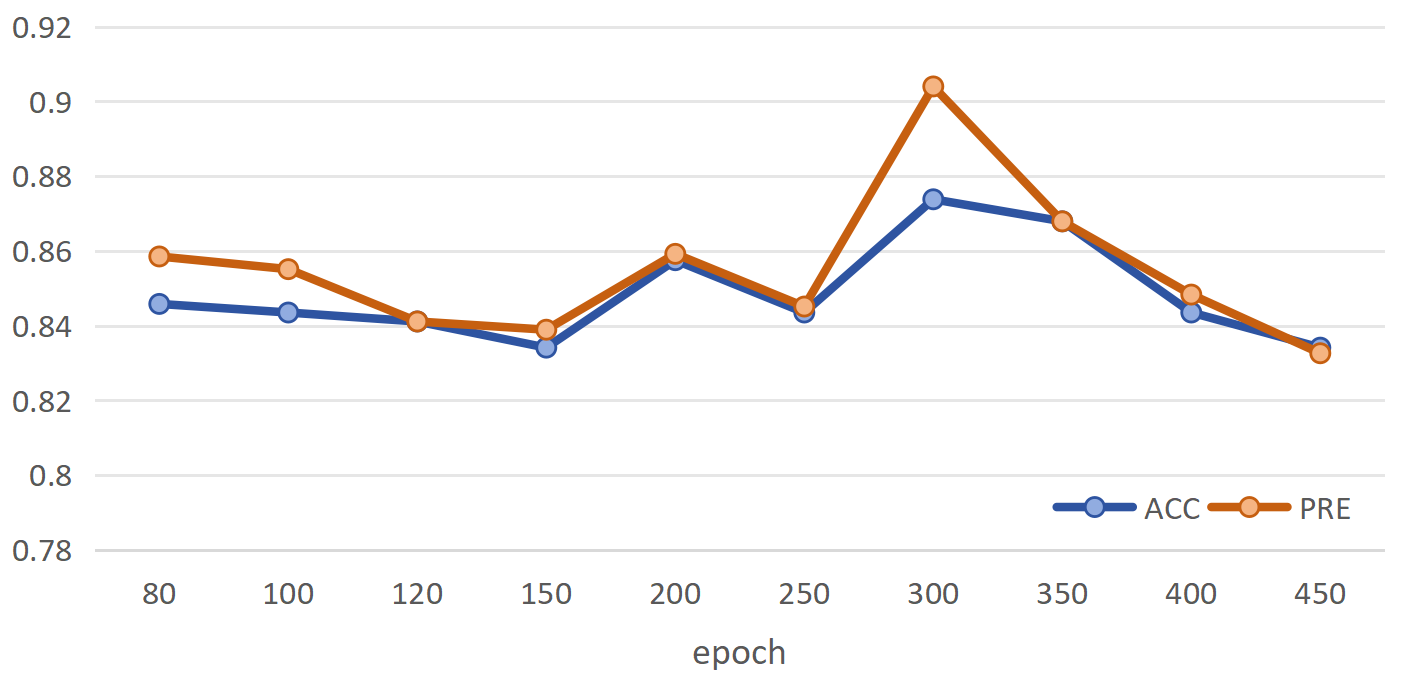
****

Figure 1. LMI-MM model evaluation indexes under different epoch parameters

# **Reference**

[1] Q. Kang, J. Meng and Y. Luan, "RNAI-FRID: novel feature representation method with information enhancement and dimension reduction for RNA-RNA interaction," *Brief Bioinform*, vol. 23, no. 3, pp. 1–10, May. 2022, doi: 10.1093/bib/bbac107.

[2] H. Li, et al., "Multi-view graph neural network with cascaded attention for lncRNA-miRNA interaction prediction," *Knowl. Based Syst.*, vol. 268, no. pp. 110492, May. 2023, doi: 10.1016/j.knosys.2023.110492.

[3] N. Sheng, et al., "Contrastive self-supervised graph convolutional network for detecting the relationship among lncRNAs, miRNAs, and diseases," in *Proc IEEE Int. Conf. Bioinform. Biomed*, Istanbul, Turkiye, 2023, pp. 622-629.

[4] W. Wang, L. Zhang, J. Sun, Q. Zhao and J. Shuai, "Predicting the potential human lncRNA-miRNA interactions based on graph convolution network with conditional random field," *Brief Bioinform*, vol. 23, no. 6, pp. bbac463, Nov. 2022, doi: 10.1093/bib/bbac463.

[5] L. Huang, et al., "Self-Supervised Contrastive Learning on Attribute and Topology Graphs for Predicting Relationships Among lncRNAs, miRNAs and Diseases," *IEEE J. Biomed. Health Inform.*, vol. 29, no.1, pp. 1 - 12, Sep. 2024, doi: 10.1109/JBHI.2024.3467101