# Supplementary information

**Figure S1. ROC curves of AMPred-MFG under 10-fold CV on the benchmark dataset.**



**Table S1. The hyper-parameters for AMPred-MFG.**

|  |  |  |
| --- | --- | --- |
| **Parameter** | **Description** | **Range** |
| batch\_size | Input batch size | {64, 128} |
| init\_lr | Initial learning rate | {0.00001} |
| weight\_decay | Weight decay for Adam | {1e-5} |
| epochs | Total number of epochs | {200} |
| heads | Multi-head num | {4} |
| feat\_dim | Output feature dimension | {128} |
| drop\_ratio | Dropout ratio | {0.2} |
| folds | K folds validation | {10,50} |

**Table S2. Comparison of AMPred-MFG with other methods on two external validation sets.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Methods | AUC | ACC | SPE | SEN | NPV | PPV | MCC |
| Test set Ⅰ | AMPred-MFG | **0.873** | 0.793 | **0.826** | 0.79 | 0.754 | **0.827** | **0.584** |
|  | HimGNN | 0.827 | **0.831** | 0.819 | **0.818** | **0.80** | 0.813 | 0.574 |
|  | DMFGAM | 0.829 | 0.754 | 0.718 | 0.785 | 0.746 | 0.76 | 0.505 |
|  | MolFPG | 0.772 | 0.691 | 0.807 | 0.604 | 0.605 | 0.806 | 0.412 |
|  | GAT | 0.826 | 0.76 | 0.676 | 0.834 | 0.782 | 0.745 | 0.519 |
|  | RF | 0.771 | 0.775 | 0.708 | 0.834 | 0.79 | 0.764 | 0.549 |
| Test set Ⅱ | AMPred-MFG | **0.902** | **0.819** | 0.837 | 0.805 | **0.774** | 0.861 | **0.639** |
|  | HimGNN | 0.81 | 0.765 | **0.844** | 0.814 | 0.71 | **0.875** | 0.511 |
|  | DMFGAM | 0.857 | 0.805 | 0.721 | **0.849** | 0.712 | 0.854 | 0.568 |
|  | MolFPG | 0.818 | 0.747 | 0.76 | 0.738 | 0.685 | 0.804 | 0.519 |
|  | GAT | 0.845 | 0.801 | 0.72 | 0.83 | 0.713 | 0.843 | 0.549 |
|  | RF | 0.743 | 0.751 | 0.721 | 0.766 | 0.605 | 0.836 | 0.469 |

The results for AMPred-MFG on the external test sets are shown in Table S2. For Test Set I, AMPred-MFG outperforms the second-best model by 4.4% in AUC, 0.7% in SPE, 1.4% in PPV, and 1% in MCC., but slightly lower than HimGNN in ACC, SEN, and NPV. AMPred-MFG excels in certain key metrics but falls slightly behind HimGNN in other aspects, likely due to distinct design choices and performance strengths of the two models in different areas. As for Test Set II, AMPred-MFG outperforms the second-best model by 4.5% in AUC, 1.4% in ACC, 6.1% in NPV, and 7.1% in MCC, but it is slightly lower than second-best model in SPE, SEN, and PPV. AMPred-MFG shows suboptimal performance in some metrics during validation, which may result from the different feature extraction methods used by comparison models. Some models may overly rely on specific features for good performance in certain metrics, whereas AMPred-MFG considers a holistic set of features, resulting in lower performance in specific metrics compared to the comparison models.

**Table S3. Comparison analysis of AMPred-MFG and its ablation experiments on three types of molecular fingerprints.**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Methods | | AUC | | ACC | | SPE | | SEN | | NPV | | PPV | | MCC | |
| AMPred-FMG | **0.912** | | **0.835** | | **0.847** | | 0.849 | | 0.811 | | **0.854** | | **0.665** | |
| MACCS | | 0.903 | | 0.817 | | 0.841 | | 0.78 | | 0.745 | | 0.849 | | 0.646 | |
| ECFP | | 0.901 | | 0.821 | | 0.736 | | 0.88 | | 0.84 | | 0.803 | | 0.63 | |
| RDKit | | 0.907 | | 0.826 | | 0.821 | | 0.836 | | 0.81 | | 0.85 | | 0.65 | |
| ECFP+MACCS | | 0.902 | | 0.773 | | 0.574 | | **0.93** | | **0.868** | | 0.733 | | 0.55 | |
| RDKit+MACCS | | 0.91 | | 0.823 | | 0.813 | | 0.832 | | 0.793 | | 0.848 | | 0.644 | |
| RDKit+ECFP | | 0.908 | | 0.832 | | 0.841 | | 0.824 | | 0.792 | | 0.847 | | 0.663 | |

We first validate the impact of three types of fingerprints on model performance, including three scenarios: one using a single molecular fingerprint, another using a combination of two molecular fingerprints, and lastly, AMPred-MFG containing three fingerprints. As shown in Table S3, when using MACCS, ECFP, or RDKit as input for the molecular fingerprint module, the model's AUCs are 0.903, 0.901, and 0.907, respectively. The combination of RDKit and MACCS achieves AUC of 0.91, which is 4% higher than using the single RDKit fingerprint. The combination of RDKit and ECFP reaches AUC of 0.908, which is 1% higher than using the single RDKit fingerprint. However, the combination of ECFP and MACCS performs lower in AUC than only MACCS fingerprint used, possibly due to redundancy in certain features between the two fingerprints, leading to decreased performance. Overall, the integration of the three fingerprint features results in more stable predictive performance, achieving the best AUC, ACC, SPE, PPV, and MCC. However, AMPred-MFG’s SEN and NPV are lower than those of ECFP and MACCS combination. This may be because these two fingerprints complement each other in certain features. Although they perform well on these two metrics, their performance on other metrics is comparatively poor, likely due to lower attention to other features. Therefore, it can be concluded that the integration of these three fingerprint features exhibits better overall performance and more stable predictive capabilities.

**Table S4.** Comparison analysis between AMPred-MFG and its ablation experiments on benchmark dataset.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Methods | AUC | ACC | SPE | SEN | NPV | PPV | MCC |
| AMPred-MFG | **0.912** | **0.835** | **0.847** | **0.849** | **0.811** | **0.854** | **0.665** |
| AMPred-FP | 0.824 | 0.801 | 0.748 | 0.788 | 0.773 | 0.814 | 0.612 |
| AMPred-MOTIF | 0.831 | 0.777 | 0.824 | 0.727 | 0.696 | 0.849 | 0.565 |
| AMPred-GAT | 0.817 | 0.795 | 0.77 | 0.816 | 0.782 | 0.806 | 0.587 |
| AMPred-FGAT | 0.857 | 0.817 | 0.78 | 0.847 | 0.805 | 0.827 | 0.63 |
| AMPred-FMOTIF | 0.886 | 0.803 | 0.826 | 0.791 | 0.779 | 0.85 | 0.617 |
| AMPred-MGAT | 0.873 | 0.83 | 0.791 | 0.841 | 0.809 | 0.838 | 0.656 |

As shown in Table S4, the AUC, ACC, SPE, SEN, NPV, PPV, and MCC for AMPred-FGAT are 0.857, 0.817, 0.78, 0.847, 0.805, 0.827, and 0.63, respectively. Compared to AMPred-FP, these metrics show improvements of 3.3%, 1.6%, 3.2%, 5.9%, 3.2%, 1.3%, and 1.8%, respectively. Compared to AMPred-GAT, the improvements are 4%, 2.2%, 1%, 3.1%, 2.3%, 2.1%, and 4.3%. This indicates that integrating molecular graph features and molecular fingerprint features yields better predictive performance than using a single feature alone. For AMPred-FMOTIF, the AUC, ACC, SPE, SEN, NPV, PPV, and MCC are 0.886, 0.803, 0.826, 0.785, 0.743, 0.85, and 0.607, representing improvements of 6.2%, 0.2%, 0.78%, 0.3%, 0.6%, 3.6%, and 0.5% compared to AMPred-FP. For AMPred-MGAT, the corresponding metrics are 0.873, 0.83, 0.791, 0.841, 0.809, 0.838, and 0.656, which show improvements of 5.6%, 3.5%, 2.1%, 2.5%, 2.7%, 3.2%, and 6.9% compared to AMPred-GAT. This demonstrates that motif-based graph features can compensate for the limitations of molecular representation, enhancing model predictive performance whether integrated with molecular graph features or molecular fingerprints.

# List of abbreviations

AUC: the area under ROC curve

ACC: accuracy

FPR: false positive rate

GAT: graph attention network

MCC: Matthew’s correlation coefficient

MLP: multilayer perceptron

NPV: negative predictive value

RF: random forest

SVM: support vector machine

SEN: sensitivity

TPR: true positive rate

PPV: positive predictive value

XGBoost: extreme gradient boosting

10-fold CV: 10-fold cross-validation