Supplementary Material for

MDGP-Forest: A Novel Deep Forest for Multi-class Imbalanced Learning Based on Multi-class Disassembly and Feature Construction by Genetic Programming

This supplementary material delves deeper into several aspects.

1. Experimental setting

1.1 Definition of evaluation metrics

In multi-class imbalanced learning, conventional metrics, such as accuracy, exhibit inherent bias due to majority class dominance, thereby necessitating unbiased evaluation metrics for comprehensive performance assessment. This section introduces the specific definitions of the main metrics used in the experiments. The primary evaluation metrics we employ include F1, G-means, and MCC, which are calculated using the macro-average. F1 metric is the harmonic mean of Precision and Recall, calculated using Formulas (1) to (3).

$$\text{Recall} = \text{avg}\left(\sum_{i=1}^{c} \frac{TP_i}{TP_i+FN_i}\right) \tag{1}$$

$$\text{Precision} = \text{avg}\left(\sum_{i=1}^{c} \frac{TP_i}{TP_i+FP_i}\right) \tag{2}$$

$$\text{F1} = \text{avg}\left(\sum_{i=1}^{c} \frac{2\times Precision_i \times Recall_i}{Precision_i + Recall_i}\right) \tag{3}$$

where TP_i, FN_i, FP_i, and TN_i represent true positives, true negatives, false positives, and false negatives of class i respectively. The F1 effectively quantifies a classifier's prediction capability for minority classes by harmonizing precision and recall metrics.

Geometric Mean (G-mean) metric considers the TP_i and TN_i of the algorithm, balancing the prediction capabilities for both positive and negative instances. It is calculated using Formulas (4) to (5). The G-mean ensuring balanced recognition performance across all classes and preventing neglect of severely underperforming classes.

$$\text{Specificity} = avg\left(\sum_{i=1}^{c} \frac{TN_i}{TN_i+FP_i}\right) \tag{4}$$

$$\text{G} - \text{mean} = avg\left(\sum_{i=1}^{c} \sqrt{Recall_i \times \text{Specificity}_i}\right) \tag{5}$$

Matthews correlation coefficient (MCC) evaluation metric takes into account the number of various instances simultaneously. It is calculated using Formula (6).

$$\text{MCC} = \frac{cp\times n - \sum_i^c p_i \times t_i}{\sqrt{(n^2 - \sum_i^c p_i{}^2)\times(n^2 - \sum_i^c t_i{}^2)}} \tag{6}$$

where t_i is the number of times class i truly occurred, p_i is the number of times class i was predicted, cp is the total number of instances correctly predicted, and n is the total number of instances. The MCC holistically incorporates all elements of the confusion matrix to compute the global correlation between predicted and true labels, thereby providing robust evaluation performance even under class imbalance.

The three employed metrics collectively address evaluation bias induced by data skewness from distinct perspectives: F1 focuses on minority class performance, G-mean emphasizes inter-class fairness, and MCC captures the holistic correlation between predictions and true labels. Their combined usage effectively compensates for the limitations inherent in any single metric.

1.2 Details about comparative algorithms

In the experiments, SelfPacedEnsemble, BalancedRandomForest, SMOTEBoost, RUSBoost, BalanceCascade, and EasyEnsemble methods are implemented using the imbalanced-ensemble[1] toolkit, RandomForest is implemented using the scikit-learn[2] toolkit, Deep Forest and imbalance-DF are implemented by reproducing the original papers, XGBoost[3], lightGBM[4], CatBoost[5], OpenFE[6] are implemented by calling the toolkits provided in the corresponding papers.

In terms of hyperparameter settings, the default values for "n_estimators" vary across different toolkits. For example, algorithms implemented using the imbalanced-ensemble toolkit have a default "n_estimators" setting of 50, whereas in scikit-learn, the default value for "n_estimators" is 100. To ensure the comparability of the experiments, we set the "n_estimators" parameter to 50 for all algorithms. Additionally, the experiments maintained identical parameter configurations for the same type of base classifiers across all algorithms. For example, we ensured that BalancedRandomForest and RandomForest used the same settings for their DecisionTree components. To prevent generation errors in SMOTE when processing minority classes with insufficient instances, we set "k_neighbors" as 3 for all algorithms incorporating SMOTE modules.

This study aims to systematically compare the performance differences of algorithms designed based on different principles in multi-class imbalanced scenarios. Considering the following factors: (1) the large scale of experiments (involving 10 repetitions of 5-fold cross-validation across 35 datasets and

[1] https://github.com/ZhiningLiu1998/imbalanced-ensemble
[2] https://scikit-learn.org/stable/
[3] https://github.com/dmlc/xgboost
[4] https://github.com/microsoft/LightGBM
[5] https://catboost.ai/
[6] https://github.com/IIIS-Li-Group/OpenFE

14 algorithms), conducting grid parameter search for all algorithms would incur prohibitively high computational costs; (2) maintaining parameter consistency helps avoid the impact of tuning bias on cross-algorithm comparisons; (3) default parameters are typically rigorously validated by toolkit developers. Thus, all remaining parameters were set to their recommended default settings.

The base classifiers used in SelfPacedEnsemble, BalanceCascade, BalancedRandomForest, and RandomForest are all decision trees. SelfPacedEnsemble is an ensemble learning framework for massive highly imbalanced classification. The hyperparameter settings of SelfPacedEnsemble in the experiments are shown in Table S1.

Table S1 Parameter settings for SelfPacedEnsemble

SelfPacedEnsemble Parameter	Settings
estimator	DecisionTree
n_estimators	50
k_bins	5
criterion	gini
max_depth	None
min_samples_split	2
min_samples_leaf	1
min_weight_fraction_leaf	0
max_features	sqrt
max_leaf_nodes	None
min_impurity_decrease	0
ccp_alpha	0

BalanceCascade iteratively drops majority class samples that were already well-classified by the current ensemble. After that, it performs random under-sampling on the remaining majority class samples and train a new base estimator. The hyperparameter settings of BalanceCascade in the experiments are shown in Table S2.

Table S2 Parameter settings for BalanceCascade

BalanceCascade Parameter	Settings
estimator	DecisionTree
n_estimators	50
criterion	gini
max_depth	None
min_samples_split	2
min_samples_leaf	1
min_weight_fraction_leaf	0
max_features	sqrt
max_leaf_nodes	None
min_impurity_decrease	0

ccp_alpha	0

BalancedRandomForest randomly under-samples each bootstrap sample to balance it. The hyperparameter settings of BalancedRandomForest in the experiments are shown in Table S3.

Table S3 Parameter settings for BalancedRandomForest

BalancedRandomForest Parameter	Settings
estimator	DecisionTree
n_estimators	50
criterion	gini
max_depth	None
min_samples_split	2
min_samples_leaf	1
min_weight_fraction_leaf	0
max_features	sqrt
max_leaf_nodes	None
min_impurity_decrease	0
bootstrap	TRUE
warm_start	FALSE
ccp_alpha	0
max_samples	None

RandomForest fits a number of decision tree classifiers on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting. The hyperparameter settings of BalancedRandomForest in the experiments are shown in Table S4.

Table S4 Parameter settings for RandomForest

RandomForest Parameter	Settings
estimator	DecisionTree
n_estimators	50
criterion	gini
max_depth	None
min_samples_split	2
min_samples_leaf	1
min_weight_fraction_leaf	0
max_features	sqrt
max_leaf_nodes	None
min_impurity_decrease	0
bootstrap	TRUE
warm_start	FALSE
ccp_alpha	0
max_samples	None

SMOTEBoost, RUSBoost, and EasyEnsemble are AdaBoost-based algorithms that all employ decision stumps as base learners in their default configurations. To ensure a more equitable comparison, our experiments substitute these with full-depth decision trees. SMOTEBoost alleviates the problem of class balancing by SMOTE over-sampling the sample at each iteration of the boosting algorithm. The hyperparameter settings of SMOTEBoost in the experiments are shown in Table S5.

Table S5 Parameter settings for SMOTEBoost

SMOTEBoost Parameter	Settings
estimator	DecisionTree(max_depth=None)
n_estimators	50
k_neighbors	5
learning_rate	1
algorithm	SAMME
early_termination	FALSE

RUSBoost alleviates the problem of class balancing is alleviated by random under-sampling the sample at each iteration of the boosting algorithm. The hyperparameter settings of RUSBoost in the experiments are shown in Table S6.

Table S6 Parameter settings for RUSBoost

RUSBoost Parameter	Settings
estimator	DecisionTree(max_depth=None)
n_estimators	50
learning_rate	1
algorithm	SAMME
early_termination	FALSE

EasyEnsemble is an ensemble of AdaBoost learners trained on different balanced boostrap samples. The balancing is achieved by random under-sampling. The hyperparameter settings of EasyEnsemble in the experiments are shown in Table S7.

Table S7 Parameter settings for EasyEnsemble

EasyEnsemble Parameter	Settings
estimator	AdaBoost
n_estimators	50
estimator of AdaBoost	DecisionTree(max_depth=None)
n_estimators of AdaBoost	10
max_samples	1
max_features	1
bootstrap	TRUE
bootstrap_features	FALSE
warm_start	FALSE
learning_rate	1
algorithm	SAMME
early_termination	FALSE

XGBoost, LightGBM, and CatBoost are all algorithmic frameworks based on gradient boosting implementations. XGBoost (eXtreme Gradient Boosting) is an optimized distributed gradient boosting framework designed for efficiency, flexibility, and portability. The hyperparameter settings of XGBoost in the experiments are shown in Table S8.

Table S8 Parameter settings for XGBoost

XGBoost Parameter	Settings
n_estimators	50
objective	multi:softprob
base_score	0.5
booster	gbtree
colsample_bylevel	1
colsample_bynode	1
colsample_bytree	1
gamma	0
gpu_id	-1
grow_policy	depthwise
learning_rate	0.3
max_bin	256
max_cat_to_onehot	4
max_delta_step	0
max_depth	6
max_leaves	0
min_child_weight	1
num_parallel_tree	1
random_state	0
reg_alpha	0
reg_lambda	1
subsample	1
tree_method	exact
validate_parameters	1

LightGBM (Light Gradient Boosting Machine) is a high-performance gradient boosting framework developed by Microsoft that uses histogram-based algorithms and leaf-wise tree growth for exceptional speed and memory efficiency. The hyperparameter settings of LightGBM in the experiments are shown in Table S9.

Table S9 Parameter settings for LightGBM

lightGBM Parameter	Settings
boosting_type	gbdt
num_leaves	31
max_depth	-1
learning_rate	0.1
n_estimators	50

subsample_for_bin	200000
objective	multiclass
min_split_gain	0
min_child_weight	1e-3
min_child_samples	20
subsample	1
subsample_freq	0
colsample_bytree	1
reg_alpha	0.
reg_lambda	0
importance_type	split

CatBoost is a gradient boosting library developed by Yandex that specializes in handling categorical features natively through ordered boosting and permutation-driven encoding, while employing oblivious decision trees for robust prediction. The hyperparameter settings of CatBoost in the experiments are shown in Table S10.

Table S10 Parameter settings for CatBoost

CatBoost Parameter	Settings
nan_mode	Min
eval_metric	MultiClass
n_estimators	50
sampling_frequency	PerTree
leaf_estimation_method	Newton
random_score_type	NormalWithModelSizeDecrease
grow_policy	SymmetricTree
penalties_coefficient	1
boosting_type	Plain
model_shrink_mode	Constant
feature_border_type	GreedyLogSum
bayesian_matrix_reg	0.1
eval_fraction	0
force_unit_auto_pair_weights	FALSE
l2_leaf_reg	3
random_strength	1
rsm	1
boost_from_average	FALSE
model_size_reg	0.5
use_best_model	FALSE
random_seed	0
depth	6
posterior_sampling	FALSE
border_count	254
bagging_temperature	1

classes_count	0
auto_class_weights	None
sparse_features_conflict_fraction	0
leaf_estimation_backtracking	AnyImprovement
best_model_min_trees	1
model_shrink_rate	0
min_data_in_leaf	1
loss_function	MultiClass
learning_rate	0.5
score_function	Cosine
leaf_estimation_iterations	1
bootstrap_type	Bayesian
max_leaves	64

OpenFE is a framework for automated feature generation for tabular data. It can discover effective candidate features for improving the learning performance of both GBDT and neural networks. The experiment use LightGBM as classifier, following the configuration specified in the OpenFE paper. The hyperparameter settings of OpenFE in the experiments are shown in Table S11.

Table S11 Parameter settings for OpenFE

OpenFE Parameter	Settings
candidate_features_list	None
categorical_features	None
n_data_blocks	8
min_candidate_features	2000
metric	multi_logloss
stage1_metric	predictive
stage2_metric	gain_importance
stage2_params	None
feature_boosting	FALSE
is_stage1	TRUE
drop_columns	None
boosting_type	gbdt
num_leaves	31
max_depth	-1
learning_rate	0.1
n_estimators	50
subsample_for_bin	200000
objective	None
class_weight	None
min_split_gain	0
min_child_weight	1.00E-03
min_child_samples	20
subsample	1

subsample_freq	0
colsample_bytree	1
reg_alpha	0
reg_lambda	0
importance_type	split

DeepForest is a deep learning algorithm implemented based on the Decision Forest featuring independence of gradient computation, with cascade adaptive growth, a small number of hyperparameters. Except for the key parameters specified in the main text, all other parameters of MDGP-Forest maintain identical configurations to those in DeepForest. The hyperparameter settings of DeepForest in the experiments are shown in Table S12.

Table S12 Parameter settings for DeepForest

DeepForest Parameter	Settings
Number of Decision Forests	4
Number of Decision Trees per Forest	50
early_stop_rounds	1
if_stacking	FALSE

For the decision forest configuration, both Deep Forest and MDGP-Forest incorporate two RandomForest and two completely-RandomForest. The hyperparameter settings of the random forests align with those specified in Table S4. The hyperparameter settings of completely-RandomForest in the experiments are shown in Table S13.

Table S13 Parameter settings for completely-RandomForest

completely-RandomForest Parameter	Settings
estimator	ExtraTree
n_estimators	50
criterion	gini
splitter	random
max_depth	None
min_samples_split	2
min_samples_leaf	1
min_weight_fraction_leaf	0
max_features	sqrt
max_leaf_nodes	None
min_impurity_decrease	0
bootstrap	FALSE
class_weight	None
ccp_alpha	0
max_samples	None

Imbalance-DF was developed by integrating SMOTE and AdaBoost with Deep Forest architecture to address class imbalance problems. We reproduced Imbalance-DF and extended its functionality to support multi-class classification. The hyperparameter settings of Imbalance-DF in the experiments are shown in Table S14.

Table S14 Parameter settings for Imbalance-DF

Imbalance-DF Parameter	Settings
Number of Decision Forests	4
Number of Decision Trees per Forest	50
early_stop_rounds	1
if_stacking	FALSE
k_neighbors	3
learning_rate	1
algorithm	SAMME
early_termination	FALSE

2. Supplementary data for comparative experiments

This section presents supplementary data for the comparative experiments. We first detail the comparative experimental results based on the MCC metric. Subsequently, we provide definitions for supplementary metrics (Accuracy, Balanced Accuracy, Precision, macro-AUC, and OVO-AUC), along with their comparative experimental results and nonparametric statistical test outcomes. Finally, we present the Friedman test statistics for all metrics, accompanied by significance heatmaps derived from p-values.

2.1 Comparative experimental results on MCC metric

The MCC metric results from the comparative experiment are presented in Table S15, with the top-performing and second-best algorithms highlighted in bold and underlined text.

2.2 Comparative experimental results on other metrics

This section presents the comparative experimental results of MDGP-Forest on other metrics. The metrics used include Accuracy, Balanced Accuracy, Precision, macro-AUC, and ovo-AUC. Accuracy is defined in formula (7).

$$\text{Accuracy} = \frac{\sum_{i=1}^{c} TP_i}{N} \tag{7}$$

where, N represent the total number of instances in a dataset. Although this metric can be dominated by the majority class in imbalanced datasets, it still provides a valuable reference for algorithm performance evaluation.

Balanced Accuracy, calculated as the average of per-class Recall using the same method as formula

(2), measures the balance of the model's recognition capability across different classes. Precision is computed using macro-averaging as defined in formula (3), which measures the average reliability of the model’s positive predictions.

The macro-AUC first computes the AUC for each class individually, then averages the AUC values across all classes as specified in formula (8).

$$\mathrm{macro-AUC} = \frac{1}{C}\sum_{i=1}^{C} \mathrm{AUC}_i \tag{8}$$

where, AUC_i represents the AUC when treating class i as positive and all other classes as negative. The macro-AUC reflects the algorithm performance to separate one class from all others.

The ovo-AUC is obtained by computing the AUC for each pair of classes and then taking the average, as shown in formula (9).

$$\mathrm{ovo-AUC} = \frac{1}{C(C-1)}\sum_{i=1}^{C-1} \sum_{j=i+1}^{C} \mathrm{AUC}_{i,j} \tag{9}$$

where, $\mathrm{AUC}_{i,j}$ represents the AUC when class i is treated as positive and class j as negative. This metric evaluates the algorithm's separability between class pairs.

The evaluation results of MDGP-Forest and the algorithms used for comparison on various datasets are presented in Table S16, Table S17, and Table S18, respectively. From the results, it is evident that MDGP-Forest is ranked first in terms of Accuracy, Balanced Accuracy, Precision, macro-AUC, ovo-AUC on 23, 17, 19, 19, and 19 datasets, respectively, demonstrating superior predictive performance compared to other algorithms.

Next, the Nonparametric Statistical Test is employed to verify whether the performance of MDGP-Forest is significantly different from that of other algorithms. This step ignores high-dimensional datasets where OpenFE cannot be run, and the Friedman test statistic for all metrics are presented in Table S21. All p-values in the Friedman test were significantly below the significance level $\alpha = 0.05$, indicating statistically significant differences among the algorithms.

The Nemenyi post-hoc test can provide a more intuitive representation of the overall performance of each algorithm. The results are shown in Fig S1. From the experimental results, it is evident that the classification performance of MDGP-Forest is significantly better than other algorithms across all metrics.

Table S15 The MCC results of the performance comparison experiment

Dataset	MDGP-Forest	DeepForest	imbalance-DF	RandomForest	XGBoost	lightGBM	CatBoost	SelfPacedEnsemble	BalancedRandomForest	EasyEnsemble	BalanceCascade	RUSBoost	SMOTEBoost	OpenFE
D1	**0.9388** ±0.0177	0.9362 ±0.0135	0.9354 ±0.0107	0.9141 ±0.0133	0.9386 ±0.0112	0.9364 ±0.0119	0.9291 ±0.012	0.8849 ±0.0147	0.9074 ±0.0124	0.9077 ±0.0136	0.8912 ±0.0154	0.797 ±0.0249	0.8963 ±0.0339	-
D2	**0.8631** ±0.0334	0.8546 ±0.0349	0.8319 ±0.0404	0.8413 ±0.0422	0.8506 ±0.0347	0.8506 ±0.0294	0.8381 ±0.0382	0.8325 ±0.0441	0.8256 ±0.0381	0.8281 ±0.0437	0.819 ±0.0453	0.7902 ±0.0406	0.7734 ±0.0488	**0.8631** ±0.0323
D3	**0.9099** ±0.0095	0.9078 ±0.0085	0.8862 ±0.0081	0.8938 ±0.0082	0.8968 ±0.0074	0.8984 ±0.0083	0.8758 ±0.0085	0.8825 ±0.0093	0.8752 ±0.0083	0.8779 ±0.0078	0.882 ±0.01	0.8262 ±0.0143	0.8227 ±0.0144	0.8926 ±0.0096
D4	**0.7691** ±0.085	0.751 ±0.064	0.7518 ±0.0611	0.7522 ±0.0602	0.725 ±0.0627	0.7407 ±0.0702	0.7383 ±0.0571	0.7332 ±0.0667	0.7569 ±0.0698	0.7517 ±0.0665	0.7336 ±0.0733	0.6909 ±0.0763	0.6821 ±0.0746	0.7295 ±0.0714
D5	0.6643 ±0.0131	**0.6657** ±0.0129	0.6091 ±0.0144	0.6352 ±0.0099	0.6449 ±0.014	0.6419 ±0.0114	0.5843 ±0.0133	0.6229 ±0.0122	0.592 ±0.0115	0.613 ±0.0132	0.6243 ±0.0108	0.2951 ±0.0209	0.579 ±0.012	-
D6	**0.9472** ±0.0388	0.9176 ±0.0604	0.9338 ±0.049	0.9038 ±0.0742	0.8893 ±0.0621	0.8934 ±0.0611	0.9126 ±0.0608	0.888 ±0.0629	0.9154 ±0.0503	0.8752 ±0.0766	0.8911 ±0.0628	0.8988 ±0.0577	0.8446 ±0.0836	0.8917 ±0.06
D7	0.945 ±0.0229	0.8744 ±0.0385	0.9395 ±0.0298	0.6997 ±0.0356	0.7471 ±0.0412	0.7678 ±0.0418	0.8027 ±0.0314	0.6297 ±0.0475	0.5588 ±0.0474	0.6145 ±0.0433	0.5892 ±0.0539	0.6517 ±0.0469	0.6137 ±0.0517	**0.9594** ±0.027
D8	**0.5765** ±0.0907	0.5272 ±0.0695	0.4649 ±0.0641	0.5288 ±0.069	0.5345 ±0.0663	0.5296 ±0.067	0.5352 ±0.0591	0.4769 ±0.0601	0.4163 ±0.074	0.4986 ±0.0669	0.4298 ±0.0659	0.3364 ±0.068	0.4089 ±0.0756	0.5723 ±0.0691
D9	0.9741 ±0.0389	0.971 ±0.0403	0.9709 ±0.0427	0.9641 ±0.0384	0.959 ±0.0397	0.9403 ±0.0463	0.9655 ±0.0405	**0.9776** ±0.0343	0.9307 ±0.0564	0.9419 ±0.0472	0.9632 ±0.0424	0.947 ±0.0478	0.9633 ±0.0478	0.9532 ±0.0491
D10	**0.7411** ±0.0442	0.722 ±0.0765	0.5497 ±0.0841	0.7085 ±0.0849	0.6854 ±0.0822	0.6845 ±0.087	0.6873 ±0.0769	0.5371 ±0.0714	0.5469 ±0.0675	0.5432 ±0.0712	0.4996 ±0.1106	0.4846 ±0.1048	0.5764 ±0.0992	0.6543 ±0.0827
D11	**0.9709** ±0.0107	0.969 ±0.0108	0.9615 ±0.0148	0.9468 ±0.0189	0.967 ±0.0129	0.9661 ±0.0134	0.9675 ±0.0142	0.9651 ±0.0143	0.9499 ±0.0172	0.9591 ±0.0175	0.9612 ±0.0175	0.8982 ±0.0393	0.9621 ±0.0169	0.966 ±0.0135
D12	0.7388 ±0.0342	0.7316 ±0.0263	0.6425 ±0.0262	0.7128 ±0.0277	0.7436 ±0.0273	0.7362 ±0.0274	0.7135 ±0.031	0.6424 ±0.0257	0.626 ±0.0251	0.6488 ±0.0265	0.6129 ±0.0252	0.4363 ±0.0551	0.6185 ±0.0347	**0.7503** ±0.0278
D13	**0.3483** ±0.0699	0.3298 ±0.063	0.3202 ±0.0843	0.3316 ±0.0828	0.323 ±0.0867	0.3211 ±0.0829	0.3385 ±0.077	0.2915 ±0.0627	0.3187 ±0.0737	0.3152 ±0.0793	0.2102 ±0.0716	0.2602 ±0.0896	0.2523 ±0.0706	0.325 ±0.096
D14	**0.8408** ±0.0329	0.8361 ±0.049	0.8194 ±0.053	0.8299 ±0.0473	0.7966 ±0.0501	0.7981 ±0.052	0.8249 ±0.0515	0.7996 ±0.0525	0.7739 ±0.0589	0.7979 ±0.0474	0.7336 ±0.0773	0.6988 ±0.08	0.7135 ±0.062	0.7992 ±0.0492

D15	0.9972 ±0.002	0.9928 ±0.003	0.9754 ±0.0052	0.9754 ±0.0042	<u>0.9996</u> ±0.0006	0.9993 ±0.0007	0.9995 ±0.0008	0.9911 ±0.0025	0.9069 ±0.0091	0.951 ±0.0062	0.9658 ±0.0148	0.8793 ±0.0193	0.9902 ±0.0033	**0.9998** ±0.0005
D16	0.501 ±0.0613	**0.5112** ±0.0416	0.215 ±0.0363	<u>0.5031</u> ±0.0415	0.4938 ±0.0363	0.4924 ±0.0403	0.4262 ±0.0434	0.1801 ±0.0331	0.1917 ±0.0284	0.2296 ±0.028	0.07 ±0.0294	0.1614 ±0.0519	0.382 ±0.0362	0.4844 ±0.0332
D17	**0.9151** ±0.0052	<u>0.9127</u> ±0.0089	0.7882 ±0.0135	0.8657 ±0.0088	0.9039 ±0.0087	0.9004 ±0.0091	0.8781 ±0.0096	0.7688 ±0.0136	0.7466 ±0.0127	0.7706 ±0.0138	0.7863 ±0.0136	0.5274 ±0.0508	0.7784 ±0.0118	0.8997 ±0.0088
D18	<u>0.4998</u> ±0.0318	**0.5045** ±0.0318	0.3996 ±0.0298	0.4952 ±0.0329	0.4698 ±0.0349	0.4608 ±0.0312	0.4716 ±0.0315	0.3854 ±0.0377	0.3706 ±0.0352	0.4159 ±0.0316	0.2196 ±0.0373	0.262 ±0.0443	0.3749 ±0.0408	0.4674 ±0.0338
D19	**0.8772** ±0.0239	<u>0.863</u> ±0.0216	0.7059 ±0.0232	0.8574 ±0.0247	0.8571 ±0.0221	0.8504 ±0.022	0.8565 ±0.0206	0.7677 ±0.0256	0.6687 ±0.023	0.6881 ±0.019	0.6124 ±0.0712	0.6602 ±0.0702	0.8388 ±0.0239	0.8583 ±0.0207
D20	<u>0.5146</u> ±0.02	**0.5223** ±0.0225	0.1612 ±0.0236	0.5117 ±0.0256	0.4749 ±0.0232	0.4591 ±0.0242	0.3734 ±0.0254	0.1325 ±0.027	0.1261 ±0.0203	0.1578 ±0.0225	0.0455 ±0.0174	0.0931 ±0.0273	0.4041 ±0.0239	0.4787 ±0.0237
D21	**0.9351** ±0.0301	0.9264 ±0.0588	0.926 ±0.0558	0.9267 ±0.0521	0.9258 ±0.0551	0.9228 ±0.056	<u>0.932</u> ±0.0515	0.9217 ±0.0588	0.9237 ±0.0565	0.9307 ±0.0537	0.9199 ±0.0632	0.926 ±0.0523	0.9152 ±0.0601	0.9285 ±0.0527
D22	<u>0.9833</u> ±0.0126	**0.984** ±0.0106	0.9642 ±0.0176	0.9614 ±0.0143	0.9005 ±0.0216	0.9242 ±0.0168	0.9293 ±0.0214	0.8039 ±0.0304	0.9589 ±0.0135	0.9007 ±0.0234	0.8059 ±0.0301	0.9366 ±0.0203	0.799 ±0.0291	0.9142 ±0.0255
D23	**0.9916** ±0.0025	0.9851 ±0.0037	0.9776 ±0.0046	0.9734 ±0.0049	0.9789 ±0.0049	0.9816 ±0.0048	0.963 ±0.0071	0.9248 ±0.0093	0.9734 ±0.0047	0.9631 ±0.0059	0.9256 ±0.0088	0.9748 ±0.0054	0.9239 ±0.009	<u>0.9871</u> ±0.0043
D24	<u>0.9758</u> ±0.0035	**0.9759** ±0.0054	0.9726 ±0.0063	0.971 ±0.0074	0.9573 ±0.0081	0.9618 ±0.0081	0.9546 ±0.0091	0.8806 ±0.0144	0.9713 ±0.0062	0.9398 ±0.0096	0.8791 ±0.015	0.9511 ±0.0112	0.8763 ±0.016	-
D25	**0.9712** ±0.0077	<u>0.968</u> ±0.0075	0.9643 ±0.007	0.9603 ±0.0091	0.9585 ±0.0087	0.9618 ±0.0075	0.9406 ±0.0118	0.8882 ±0.0158	0.9614 ±0.009	0.9407 ±0.0112	0.8882 ±0.0159	0.9522 ±0.0111	0.8858 ±0.016	-
D26	<u>0.9283</u> ±0.011	**0.9296** ±0.014	0.9133 ±0.0163	0.908 ±0.0159	0.902 ±0.0186	0.8228 ±0.0231	0.9028 ±0.0187	0.8518 ±0.0245	0.9086 ±0.0163	0.8677 ±0.0207	0.8513 ±0.0268	0.8758 ±0.0216	0.8502 ±0.0238	-
D27	<u>0.7943</u> ±0.0142	**0.7965** ±0.0149	0.7855 ±0.0165	0.7733 ±0.0173	0.7782 ±0.0161	0.78 ±0.0145	0.7762 ±0.0148	0.6553 ±0.0228	0.772 ±0.0172	0.75 ±0.0181	0.6454 ±0.0182	0.7145 ±0.0196	0.6213 ±0.0214	0.7789 ±0.0191
D28	**0.7919** ±0.0161	<u>0.791</u> ±0.015	0.7849 ±0.0164	0.7697 ±0.0153	0.7697 ±0.0164	0.7733 ±0.0157	0.7741 ±0.0146	0.673 ±0.0227	0.7714 ±0.0188	0.7517 ±0.0171	0.6609 ±0.0251	0.7271 ±0.0163	0.6355 ±0.0221	0.7748 ±0.018
D29	**0.9852** ±0.0042	<u>0.985</u> ±0.0035	0.9816 ±0.0038	0.9777 ±0.0044	0.9748 ±0.0047	0.9774 ±0.0044	0.9663 ±0.0059	0.9157 ±0.0108	0.9786 ±0.0042	0.9549 ±0.007	0.9144 ±0.0112	0.9679 ±0.0052	0.9013 ±0.0107	0.9766 ±0.0056

D30	**0.9447** ±0.0116	0.9414 ±0.0147	0.9315 ±0.0153	0.9232 ±0.0182	0.9105 ±0.0187	0.918 ±0.0167	0.9011 ±0.0181	0.7807 ±0.028	0.9227 ±0.0163	0.858 ±0.0213	0.7632 ±0.0261	0.8657 ±0.0207	0.7458 ±0.0297	-
D31	**0.7499** ±0.0356	0.7099 ±0.0384	0.6794 ±0.0396	0.6697 ±0.0368	0.6829 ±0.0349	0.6865 ±0.0381	0.6883 ±0.0394	0.6304 ±0.0397	0.6635 ±0.0329	0.6601 ±0.032	0.6223 ±0.0435	0.65 ±0.045	0.6102 ±0.0478	0.7203 ±0.041
D32	**0.3289** ±0.0154	0.3267 ±0.0252	0.3234 ±0.0208	0.3161 ±0.0234	0.3169 ±0.0214	0.3235 ±0.0254	0.3257 ±0.0248	0.2822 ±0.0242	0.3196 ±0.0219	0.3103 ±0.0227	0.2788 ±0.0214	0.2697 ±0.024	0.2362 ±0.0271	0.3173 ±0.0204
D33	**0.9907** ±0.0023	0.99 ±0.0023	0.9841 ±0.0024	0.9725 ±0.0039	0.9873 ±0.0028	0.986 ±0.0027	0.9681 ±0.0042	0.9568 ±0.0064	0.9701 ±0.0034	0.9588 ±0.0049	0.9529 ±0.0063	0.9291 ±0.0074	0.9245 ±0.0063	-
D34	**0.9983** ±0.0027	0.9966 ±0.004	0.9891 ±0.0085	0.9871 ±0.0084	0.9675 ±0.0105	0.9746 ±0.0125	0.9735 ±0.0111	0.9392 ±0.0209	0.9779 ±0.0102	0.9476 ±0.0184	0.9168 ±0.0266	0.9482 ±0.0193	0.8524 ±0.0292	0.9757 ±0.0131
D35	0.2921 ±0.0287	0.2799 ±0.0397	**0.318** ±0.042	0.254 ±0.0402	0.2746 ±0.0366	0.2872 ±0.0373	0.2821 ±0.0325	0.2468 ±0.0318	0.2775 ±0.0357	0.2664 ±0.0358	0.2275 ±0.0391	0.1864 ±0.0408	0.2191 ±0.0393	0.275 ±0.0404

Table S16 The Accuracy results of the performance comparison experiment

Dataset	MDGP-Forest	DeepForest	imbalance-DF	RandomForest	XGBoost	lightGBM	CatBoost	SelfPacedEnsemble	BalancedRandomForest	EasyEnsemble	BalanceCascade	RUSBoost	SMOTEBoost	OpenFE
D1	**0.9622** ±0.0109	0.9606 ±0.0083	0.9598 ±0.0067	0.9472 ±0.0082	0.9619 ±0.007	0.9606 ±0.0075	0.9558 ±0.0076	0.9289 ±0.0091	0.9417 ±0.008	0.9423 ±0.0086	0.9324 ±0.0097	0.876 ±0.015	0.9357 ±0.021	-
D2	**0.9017** ±0.0237	0.8954 ±0.025	0.8782 ±0.0294	0.8858 ±0.0303	0.8925 ±0.0249	0.8925 ±0.021	0.8835 ±0.0273	0.8791 ±0.0316	0.8738 ±0.0273	0.8755 ±0.0316	0.8692 ±0.0324	0.8491 ±0.029	0.8359 ±0.035	0.9014 ±0.0233
D3	**0.9269** ±0.0076	0.9252 ±0.0069	0.9071 ±0.0067	0.9138 ±0.0066	0.9163 ±0.006	0.9176 ±0.0067	0.8994 ±0.0069	0.9044 ±0.0075	0.8977 ±0.0069	0.9005 ±0.0064	0.9041 ±0.0081	0.8591 ±0.0116	0.856 ±0.0117	0.913 ±0.0077
D4	**0.8552** ±0.0533	0.8423 ±0.0406	0.8394 ±0.0395	0.8439 ±0.0374	0.8265 ±0.0384	0.8365 ±0.0435	0.8355 ±0.0357	0.829 ±0.0426	0.8435 ±0.0444	0.8413 ±0.0424	0.8297 ±0.0476	0.8052 ±0.0482	0.7994 ±0.0463	0.8299 ±0.0437
D5	0.7197 ±0.0109	**0.7207** ±0.0107	0.6695 ±0.0124	0.6955 ±0.0081	0.7034 ±0.0117	0.7013 ±0.0094	0.6503 ±0.0111	0.6808 ±0.0107	0.6552 ±0.01	0.6741 ±0.0112	0.6833 ±0.0092	0.4164 ±0.0173	0.6472 ±0.0103	-
D6	**0.9749** ±0.0181	0.9609 ±0.0284	0.9674 ±0.024	0.9549 ±0.0343	0.9479 ±0.0288	0.9498 ±0.0283	0.9591 ±0.028	0.9465 ±0.0294	0.9572 ±0.0263	0.9391 ±0.0367	0.947 ±0.0308	0.9521 ±0.0272	0.9256 ±0.0404	0.9489 ±0.0279
D7	0.9685 ±0.0131	0.9259 ±0.0225	0.9632 ±0.0188	0.8298 ±0.0217	0.8573 ±0.0242	0.8698 ±0.0233	0.889 ±0.0178	0.7459 ±0.0374	0.6798 ±0.0411	0.7474 ±0.0342	0.7336 ±0.0391	0.7968 ±0.0302	0.7746 ±0.0333	**0.9768** ±0.0154
D8	**0.7822**	0.7579	0.685	0.7587	0.7589	0.7571	0.7606	0.6769	0.6556	0.7125	0.6399	0.6506	0.6727	0.7781

	±0.0428	±0.0331	±0.0374	±0.0329	±0.0323	±0.0334	±0.0289	±0.0348	±0.0441	±0.037	±0.0414	±0.038	±0.0405	±0.0336
D9	0.9798	0.9774	0.9775	0.9722	0.9682	0.9537	0.9732	**0.9826**	0.9455	0.9547	0.9713	0.9588	0.9713	0.9639
	±0.0305	±0.0313	±0.033	±0.0298	±0.0309	±0.0359	±0.0317	±0.0267	±0.0449	±0.0371	±0.033	±0.0375	±0.0375	±0.0377
D10	**0.8083**	0.7949	0.6343	0.7856	0.7678	0.7678	0.7706	0.6242	0.6294	0.6379	0.5872	0.6228	0.6804	0.7453
	±0.0311	±0.0551	±0.0718	±0.0606	±0.0597	±0.0631	±0.0553	±0.0616	±0.0571	±0.0603	±0.0943	±0.0737	±0.0737	±0.0597
D11	**0.9893**	0.9886	0.9859	0.9806	0.9879	0.9876	0.9881	0.9872	0.9815	0.9848	0.9857	0.9606	0.986	0.9875
	±0.0039	±0.0039	±0.0054	±0.0068	±0.0047	±0.0049	±0.0052	±0.0052	±0.0063	±0.0065	±0.0064	±0.0175	±0.0063	±0.0049
D12	0.7973	0.7917	0.7014	0.778	0.8014	0.7956	0.7779	0.6914	0.6837	0.7084	0.6603	0.5684	0.6996	**0.8065**
	±0.0259	±0.0203	±0.0218	±0.021	±0.0209	±0.0213	±0.0237	±0.023	±0.0218	±0.0227	±0.0221	±0.0405	±0.0271	±0.0213
D13	**0.6759**	0.6653	0.5956	0.6619	0.6456	0.652	0.6619	0.5524	0.5772	0.5789	0.4657	0.6033	0.5612	0.6529
	±0.0273	±0.027	±0.0582	±0.0373	±0.0442	±0.0424	±0.0381	±0.0504	±0.0617	±0.0582	±0.0656	±0.0702	±0.0508	±0.0472
D14	**0.8847**	0.8816	0.8656	0.8771	0.8527	0.8539	0.8735	0.85	0.8295	0.85	0.7981	0.7722	0.7894	0.8545
	±0.0242	±0.035	±0.0404	±0.0338	±0.0359	±0.0373	±0.0368	±0.0397	±0.0446	±0.0363	±0.0617	±0.0663	±0.0458	±0.0352
D15	0.9981	0.9951	0.9831	0.9832	0.9997	0.9995	0.9997	0.9939	0.935	0.9661	0.9764	0.9161	0.9933	**0.9998**
	±0.0013	±0.0021	±0.0036	±0.0029	±0.0004	±0.0005	±0.0005	±0.0017	±0.0064	±0.0043	±0.0105	±0.0143	±0.0022	±0.0003
D16	0.6909	**0.6974**	0.3996	0.6921	0.6839	0.6835	0.6435	0.3482	0.3598	0.4153	0.1791	0.4262	0.5889	0.6791
	±0.0365	±0.0252	±0.0333	±0.0249	±0.0224	±0.0245	±0.027	±0.0304	±0.0333	±0.0258	±0.0332	±0.0446	±0.0254	±0.0205
D17	**0.9587**	0.9575	0.8857	0.936	0.9536	0.9519	0.9411	0.8734	0.8579	0.877	0.8827	0.7234	0.8886	0.9516
	±0.0025	±0.0043	±0.0084	±0.0041	±0.0041	±0.0043	±0.0046	±0.0087	±0.0084	±0.0082	±0.0091	±0.0495	±0.0062	±0.0042
D18	0.615	**0.6187**	0.4917	0.6117	0.5918	0.5861	0.5927	0.4854	0.4792	0.5232	0.3115	0.4109	0.5058	0.5907
	±0.0244	±0.0245	±0.0295	±0.0251	±0.0264	±0.024	±0.0238	±0.0386	±0.0325	±0.0281	±0.0411	±0.0377	±0.0319	±0.0262
D19	**0.977**	0.9744	0.9159	0.9735	0.9732	0.972	0.9732	0.9413	0.9001	0.911	0.8608	0.9111	0.9681	0.9735
	±0.0045	±0.004	±0.0104	±0.0045	±0.0041	±0.0041	±0.0037	±0.009	±0.0106	±0.0088	±0.048	±0.0311	±0.005	±0.0038
D20	0.6839	**0.6888**	0.3247	0.6822	0.6569	0.6472	0.595	0.2952	0.2876	0.3224	0.1324	0.3066	0.5819	0.6599
	±0.0127	±0.0143	±0.0305	±0.0163	±0.015	±0.0155	±0.0157	±0.0343	±0.0258	±0.0241	±0.0319	±0.0365	±0.0173	±0.0153
D21	**0.9553**	0.9493	0.9493	0.95	0.9493	0.9473	0.9533	0.9467	0.948	0.9527	0.9453	0.9493	0.942	0.9513
	±0.0209	±0.0405	±0.0382	±0.0352	±0.0376	±0.0381	±0.0356	±0.0398	±0.0382	±0.0363	±0.043	±0.0358	±0.0409	±0.0358
D22	0.9846	**0.9854**	0.9673	0.9646	0.909	0.9307	0.9354	0.8209	0.9624	0.9092	0.8227	0.942	0.8164	0.9216
	±0.0116	±0.0097	±0.0161	±0.0131	±0.0197	±0.0153	±0.0196	±0.0277	±0.0123	±0.0214	±0.0274	±0.0187	±0.0266	±0.0234
D23	**0.9924**	0.9865	0.9796	0.9758	0.9808	0.9833	0.9663	0.9316	0.9758	0.9664	0.9323	0.9771	0.9308	0.9883

	±0.0023	±0.0034	±0.0042	±0.0044	±0.0044	±0.0044	±0.0064	±0.0085	±0.0042	±0.0053	±0.008	±0.0049	±0.0082	±0.0039
D24	0.9781	**0.9782**	0.9753	0.9738	0.9615	0.9655	0.959	0.8923	0.974	0.9457	0.8908	0.9558	0.8884	-
	±0.0032	±0.0049	±0.0057	±0.0067	±0.0073	±0.0073	±0.0082	±0.0131	±0.0056	±0.0087	±0.0136	±0.0101	±0.0144	
D25	**0.9739**	0.9712	0.9678	0.9642	0.9625	0.9656	0.9464	0.8991	0.9651	0.9465	0.8991	0.9568	0.8969	-
	±0.007	±0.0068	±0.0064	±0.0082	±0.0078	±0.0068	±0.0107	±0.0142	±0.0082	±0.0101	±0.0143	±0.0101	±0.0145	
D26	0.9358	**0.937**	0.9224	0.9177	0.9124	0.8411	0.913	0.8672	0.9182	0.8816	0.8669	0.8889	0.8658	-
	±0.0099	±0.0125	±0.0146	±0.0143	±0.0166	±0.0209	±0.0169	±0.022	±0.0147	±0.0186	±0.024	±0.0194	±0.0213	
D27	0.8617	**0.8632**	0.856	0.8481	0.8517	0.8529	0.8505	0.77	0.8472	0.8327	0.7634	0.8094	0.7474	0.8522
	±0.0098	±0.0102	±0.0111	±0.0116	±0.0108	±0.0097	±0.0098	±0.0152	±0.0116	±0.0122	±0.0122	±0.0131	±0.0143	±0.0127
D28	**0.8605**	0.8598	0.8562	0.8461	0.8462	0.8487	0.8493	0.7818	0.8472	0.8342	0.7738	0.8179	0.7569	0.8497
	±0.011	±0.0101	±0.0109	±0.0102	±0.0109	±0.0104	±0.0097	±0.0151	±0.0125	±0.0114	±0.0167	±0.0109	±0.0147	±0.012
D29	**0.9867**	0.9865	0.9835	0.9799	0.9773	0.9797	0.9696	0.924	0.9807	0.9594	0.9229	0.971	0.9111	0.9789
	±0.0038	±0.0032	±0.0034	±0.004	±0.0042	±0.0039	±0.0053	±0.0097	±0.0038	±0.0063	±0.01	±0.0047	±0.0096	±0.0051
D30	**0.95**	0.9471	0.9381	0.9306	0.9191	0.9259	0.9107	0.802	0.9302	0.8718	0.7862	0.8787	0.7706	-
	±0.0105	±0.0132	±0.0138	±0.0165	±0.0169	±0.0151	±0.0164	±0.0252	±0.0147	±0.0193	±0.0236	±0.0187	±0.0268	
D31	**0.8115**	0.7811	0.7579	0.7508	0.7613	0.7638	0.7653	0.722	0.7463	0.7436	0.7159	0.7364	0.7068	0.7894
	±0.0264	±0.0289	±0.0299	±0.0277	±0.0261	±0.0287	±0.0295	±0.0297	±0.0249	±0.0241	±0.0325	±0.0337	±0.0359	±0.0308
D32	**0.5539**	0.5524	0.5452	0.5448	0.5459	0.5503	0.5516	0.5202	0.5435	0.5397	0.5183	0.5147	0.4915	0.5461
	±0.0099	±0.0166	±0.0137	±0.0156	±0.0141	±0.0167	±0.0165	±0.0162	±0.0147	±0.0151	±0.0143	±0.0159	±0.018	±0.0136
D33	**0.9923**	0.9917	0.9868	0.9771	0.9894	0.9883	0.9735	0.964	0.9751	0.9657	0.9608	0.941	0.9372	-
	±0.0019	±0.0019	±0.002	±0.0033	±0.0023	±0.0022	±0.0035	±0.0054	±0.0029	±0.0041	±0.0052	±0.0062	±0.0052	
D34	**0.9985**	0.997	0.9904	0.9887	0.9714	0.9777	0.9767	0.9465	0.9806	0.9539	0.9269	0.9544	0.8704	0.9786
	±0.0024	±0.0035	±0.0075	±0.0074	±0.0092	±0.0109	±0.0098	±0.0184	±0.009	±0.0161	±0.0233	±0.017	±0.0256	±0.0115
D35	**0.5449**	0.539	0.5386	0.5218	0.5316	0.5415	0.5377	0.5037	0.5193	0.5187	0.4946	0.4795	0.495	0.5337
	±0.0194	±0.0247	±0.0271	±0.0257	±0.0234	±0.0238	±0.0208	±0.021	±0.0245	±0.0233	±0.0265	±0.0264	±0.0258	±0.0251

Table S17 The Balanced Accuracy results of the performance comparison experiment

Dataset	MDGP-Forest	DeepForest	imbalance-DF	RandomForest	XGBoost	lightGBM	CatBoost	SelfPacedEnsemble	BalancedRandomForest	EasyEnsemble	BalanceCascade	RUSBoost	SMOTEBoost	OpenFE
D1	0.9583	0.9558	0.9598	0.9385	**0.9606**	0.9592	0.9563	0.9199	0.9452	0.941	0.9271	0.8563	0.9314	-
	±0.0139	±0.0104	±0.0079	±0.0103	±0.0078	±0.0084	±0.0081	±0.0107	±0.0085	±0.0098	±0.0109	±0.02	±0.0234	

D2	**0.8919** ±0.0248	0.8839 ±0.0273	0.8799 ±0.0292	0.876 ±0.0331	0.8806 ±0.028	0.8775 ±0.0259	0.8715 ±0.031	0.871 ±0.0362	0.8751 ±0.0293	0.8735 ±0.033	0.8643 ±0.0355	0.8296 ±0.0356	0.8313 ±0.0395	0.8881 ±0.0266
D3	**0.9049** ±0.0095	0.903 ±0.0094	0.8992 ±0.008	0.8878 ±0.0092	0.8951 ±0.0081	0.8962 ±0.0094	0.8731 ±0.0098	0.8909 ±0.0094	0.8927 ±0.0077	0.8897 ±0.0079	0.8897 ±0.0103	0.8277 ±0.0153	0.8368 ±0.0137	0.8911 ±0.0101
D4	0.8146 ±0.0721	0.8 ±0.055	0.8159 ±0.0486	0.7946 ±0.0561	0.7728 ±0.0574	0.7836 ±0.0589	0.7816 ±0.0492	0.795 ±0.0553	**0.8163** ±0.0575	0.808 ±0.053	0.7947 ±0.06	0.7449 ±0.0639	0.7474 ±0.065	0.7757 ±0.0632
D5	**0.6951** ±0.0101	0.6949 ±0.012	0.6684 ±0.0128	0.6628 ±0.0093	0.6764 ±0.0124	0.6722 ±0.0111	0.602 ±0.0137	0.6816 ±0.0108	0.6553 ±0.01	0.6699 ±0.012	0.6789 ±0.0099	0.3569 ±0.0202	0.6364 ±0.0099	-
D6	0.9627 ±0.0402	0.9321 ±0.0607	**0.963** ±0.0443	0.9244 ±0.0694	0.9112 ±0.0614	0.9197 ±0.0578	0.9255 ±0.0635	0.9235 ±0.0585	0.9614 ±0.0373	0.9298 ±0.0626	0.9325 ±0.0508	0.9167 ±0.0593	0.9007 ±0.0615	0.9181 ±0.0564
D7	0.9385 ±0.0335	0.8964 ±0.0525	**0.9717** ±0.0158	0.6002 ±0.0155	0.6411 ±0.0312	0.6403 ±0.0269	0.6684 ±0.027	0.7095 ±0.0612	0.6556 ±0.058	0.6527 ±0.0464	0.6296 ±0.0646	0.6095 ±0.0396	0.5692 ±0.0303	0.947 ±0.0362
D8	0.6588 ±0.0963	0.6279 ±0.0631	0.6897 ±0.0504	0.6294 ±0.0592	0.6493 ±0.0667	0.6406 ±0.0565	0.6388 ±0.0537	**0.7107** ±0.0512	0.66 ±0.0583	0.7028 ±0.0609	0.6838 ±0.054	0.5625 ±0.0554	0.623 ±0.0609	0.6692 ±0.0639
D9	0.9589 ±0.0622	0.9492 ±0.0753	0.9467 ±0.0821	0.9267 ±0.0819	0.9167 ±0.0842	0.8917 ±0.0883	0.9381 ±0.0781	**0.9633** ±0.0611	0.9006 ±0.0923	0.9059 ±0.0887	0.935 ±0.0847	0.9137 ±0.0839	0.9376 ±0.0815	0.9158 ±0.0909
D10	**0.7553** ±0.0671	0.745 ±0.0865	0.7263 ±0.0594	0.7243 ±0.0963	0.7121 ±0.0776	0.706 ±0.0869	0.6837 ±0.0967	0.7287 ±0.0638	0.743 ±0.0513	0.7206 ±0.0561	0.7123 ±0.0782	0.5658 ±0.113	0.7086 ±0.0887	0.7015 ±0.0841
D11	0.9776 ±0.0096	0.9768 ±0.01	0.9758 ±0.011	0.9537 ±0.0158	0.9755 ±0.0106	0.9748 ±0.0115	0.9749 ±0.0113	0.9764 ±0.0105	0.9708 ±0.0128	**0.9795** ±0.0113	0.9758 ±0.0126	0.9513 ±0.0156	0.9754 ±0.011	0.9751 ±0.0113
D12	0.8047 ±0.0437	0.8039 ±0.029	0.7943 ±0.0231	0.7829 ±0.0283	0.8071 ±0.0304	0.7986 ±0.0293	0.7884 ±0.0298	**0.8114** ±0.0219	0.7883 ±0.0226	0.7953 ±0.025	0.7977 ±0.0214	0.5345 ±0.0498	0.7299 ±0.0363	0.8092 ±0.0307
D13	0.3435 ±0.0711	0.3149 ±0.0504	0.3444 ±0.0808	0.3304 ±0.0619	0.3452 ±0.0781	0.3279 ±0.0672	0.3415 ±0.069	0.3429 ±0.0722	**0.3683** ±0.0719	0.3554 ±0.0819	0.323 ±0.0799	0.3006 ±0.0624	0.3192 ±0.0703	0.331 ±0.0822
D14	0.8563 ±0.0406	0.8493 ±0.0503	**0.8675** ±0.0386	0.8287 ±0.0668	0.7919 ±0.0767	0.7257 ±0.0935	0.8255 ±0.0851	0.854 ±0.0434	0.8384 ±0.046	0.848 ±0.0436	0.809 ±0.0564	0.6883 ±0.1005	0.775 ±0.0532	0.7574 ±0.0876
D15	0.9962 ±0.0036	0.9894 ±0.0065	0.9867 ±0.003	0.9399 ±0.0124	**0.9997** ±0.0008	0.9981 ±0.0034	0.9991 ±0.0019	0.9953 ±0.0013	0.95 ±0.0049	0.974 ±0.0033	0.9808 ±0.0084	0.8937 ±0.0277	0.9873 ±0.0062	**0.9997** ±0.0011
D16	0.3414 ±0.0262	0.3395 ±0.0197	**0.3978** ±0.0639	0.3523 ±0.0291	0.3555 ±0.0249	0.3564 ±0.0272	0.3213 ±0.0223	0.3699 ±0.072	0.3742 ±0.0698	0.3856 ±0.0649	0.3507 ±0.06	0.2631 ±0.0332	0.3667 ±0.0505	0.3453 ±0.0224

D17	0.8505 ±0.0145	0.8478 ±0.0227	0.8538 ±0.0165	0.7271 ±0.0189	0.8173 ±0.0222	0.8015 ±0.0208	0.8062 ±0.0204	0.847 ±0.0198	0.8398 ±0.0188	0.8271 ±0.0226	**0.8624** ±0.0176	0.5677 ±0.0354	0.7629 ±0.0229	0.8029 ±0.0201
D18	0.5744 ±0.0619	0.5638 ±0.0583	0.5509 ±0.0488	0.5266 ±0.062	0.5267 ±0.0666	0.4206 ±0.0303	0.4889 ±0.06	0.5643 ±0.0398	0.5458 ±0.0474	**0.5795** ±0.0351	0.4606 ±0.0483	0.3314 ±0.0387	0.4977 ±0.0523	0.4512 ±0.0353
D19	0.877 ±0.0376	0.8482 ±0.046	0.9406 ±0.0186	0.8373 ±0.0462	0.8492 ±0.0473	0.8346 ±0.0428	0.8544 ±0.0421	**0.9479** ±0.0158	0.9315 ±0.018	0.9384 ±0.0153	0.9388 ±0.0178	0.8416 ±0.0714	0.8737 ±0.0435	0.8358 ±0.0455
D20	0.3834 ±0.0141	0.3794 ±0.0154	0.3562 ±0.0724	0.3749 ±0.0156	0.3717 ±0.0178	0.3667 ±0.0196	0.2979 ±0.0144	0.3156 ±0.0738	0.3071 ±0.0676	0.352 ±0.0822	0.2829 ±0.0728	0.2185 ±0.0256	**0.3852** ±0.0266	0.3758 ±0.0212
D21	**0.9553** ±0.0209	0.9493 ±0.0405	0.9493 ±0.0382	0.95 ±0.0352	0.9493 ±0.0376	0.9473 ±0.0381	0.9533 ±0.0356	0.9467 ±0.0398	0.948 ±0.0382	0.9527 ±0.0363	0.9453 ±0.043	0.9493 ±0.0358	0.942 ±0.0409	0.9513 ±0.0358
D22	0.9846 ±0.0116	**0.9854** ±0.0097	0.9673 ±0.0161	0.9646 ±0.0131	0.909 ±0.0197	0.9307 ±0.0153	0.9354 ±0.0196	0.8209 ±0.0277	0.9624 ±0.0123	0.9092 ±0.0214	0.8227 ±0.0274	0.942 ±0.0187	0.8164 ±0.0266	0.9216 ±0.0234
D23	**0.9924** ±0.0023	0.9865 ±0.0034	0.9796 ±0.0042	0.9758 ±0.0044	0.9808 ±0.0044	0.9833 ±0.0044	0.9663 ±0.0064	0.9316 ±0.0085	0.9758 ±0.0042	0.9664 ±0.0053	0.9323 ±0.008	0.9771 ±0.0049	0.9308 ±0.0082	0.9883 ±0.0039
D24	0.9781 ±0.0032	**0.9782** ±0.0049	0.9753 ±0.0057	0.9738 ±0.0067	0.9615 ±0.0073	0.9655 ±0.0073	0.959 ±0.0082	0.8923 ±0.0131	0.974 ±0.0056	0.9457 ±0.0087	0.8908 ±0.0136	0.9558 ±0.0101	0.8884 ±0.0144	-
D25	**0.9739** ±0.007	0.9712 ±0.0068	0.9678 ±0.0064	0.9642 ±0.0082	0.9625 ±0.0078	0.9656 ±0.0068	0.9464 ±0.0107	0.8991 ±0.0142	0.9651 ±0.0082	0.9465 ±0.0101	0.8991 ±0.0143	0.9568 ±0.0101	0.8969 ±0.0145	-
D26	0.9358 ±0.0099	**0.937** ±0.0125	0.9224 ±0.0146	0.9177 ±0.0143	0.9124 ±0.0166	0.8411 ±0.0209	0.913 ±0.0169	0.8672 ±0.022	0.9182 ±0.0147	0.8816 ±0.0186	0.8669 ±0.024	0.8889 ±0.0194	0.8658 ±0.0213	-
D27	0.8622 ±0.0096	**0.8637** ±0.0101	0.8565 ±0.011	0.8486 ±0.0116	0.8521 ±0.0108	0.8532 ±0.0097	0.8508 ±0.0098	0.7702 ±0.0152	0.8477 ±0.0116	0.8331 ±0.0121	0.7635 ±0.0122	0.8097 ±0.0131	0.7476 ±0.0143	0.8525 ±0.0127
D28	**0.8602** ±0.011	0.8595 ±0.0101	0.8559 ±0.0109	0.8458 ±0.0102	0.846 ±0.0109	0.8485 ±0.0104	0.8491 ±0.0097	0.7816 ±0.0151	0.847 ±0.0126	0.8339 ±0.0114	0.7736 ±0.0167	0.8177 ±0.0109	0.7567 ±0.0147	0.8495 ±0.0119
D29	**0.9867** ±0.0038	0.9865 ±0.0032	0.9835 ±0.0034	0.9799 ±0.0039	0.9773 ±0.0042	0.9796 ±0.0039	0.9696 ±0.0053	0.9241 ±0.0097	0.9807 ±0.0038	0.9594 ±0.0062	0.923 ±0.0101	0.9711 ±0.0047	0.9112 ±0.0096	0.9789 ±0.0051
D30	**0.9499** ±0.0105	0.947 ±0.0132	0.938 ±0.0138	0.9303 ±0.0165	0.9188 ±0.0169	0.9257 ±0.0152	0.9104 ±0.0164	0.8013 ±0.0253	0.9299 ±0.0148	0.8713 ±0.0194	0.7855 ±0.0236	0.8784 ±0.0188	0.7699 ±0.0269	-
D31	**0.8138** ±0.0259	0.7835 ±0.0285	0.7608 ±0.0293	0.7537 ±0.0272	0.7638 ±0.0257	0.7664 ±0.0283	0.7679 ±0.0291	0.7241 ±0.0295	0.7494 ±0.0243	0.7465 ±0.0236	0.7181 ±0.032	0.7386 ±0.0332	0.7087 ±0.0357	0.7917 ±0.0303

D32	**0.5538** ±0.0101	0.5524 ±0.0165	0.5531 ±0.0138	0.5473 ±0.0153	0.547 ±0.014	0.5512 ±0.0168	0.5526 ±0.0161	0.5255 ±0.016	0.5508 ±0.0145	0.5442 ±0.0151	0.5232 ±0.0142	0.5153 ±0.0161	0.494 ±0.018	0.5475 ±0.0132
D33	**0.9925** ±0.0017	0.992 ±0.0019	0.9871 ±0.0019	0.977 ±0.0034	0.9895 ±0.0023	0.9885 ±0.0023	0.9736 ±0.0035	0.963 ±0.0057	0.9752 ±0.0029	0.9656 ±0.0041	0.9598 ±0.0055	0.9426 ±0.0059	0.936 ±0.0054	-
D34	**0.9986** ±0.0023	0.9971 ±0.0035	0.9909 ±0.0072	0.9886 ±0.0077	0.9718 ±0.0092	0.9778 ±0.0106	0.9769 ±0.0097	0.9463 ±0.019	0.9811 ±0.0087	0.9546 ±0.016	0.9273 ±0.0237	0.9536 ±0.018	0.8697 ±0.0264	0.9788 ±0.0115
D35	0.5214 ±0.0205	0.5088 ±0.027	**0.5466** ±0.0305	0.4941 ±0.0274	0.5085 ±0.0249	0.5158 ±0.0255	0.5146 ±0.0221	0.4953 ±0.023	0.517 ±0.025	0.5064 ±0.026	0.4824 ±0.0267	0.4488 ±0.0267	0.4736 ±0.0262	0.5087 ±0.0285

Table S18 The Precision results of the performance comparison experiment

Dataset	MDGP-Forest	DeepForest	imbalance-DF	RandomForest	XGBoost	lightGBM	CatBoost	SelfPacedEnsemble	BalancedRandomForest	EasyEnsemble	BalanceCascade	RUSBoost	SMOTEBoost	OpenFE
D1	**0.9566** ±0.012	0.9551 ±0.0101	0.9509 ±0.0084	0.9425 ±0.0094	0.9546 ±0.0088	0.9532 ±0.0093	0.9464 ±0.0095	0.9182 ±0.011	0.9287 ±0.0096	0.9318 ±0.0107	0.9201 ±0.0121	0.8758 ±0.0155	0.9252 ±0.0238	-
D2	0.9043 ±0.0241	0.9017 ±0.0263	0.8756 ±0.0316	0.8938 ±0.0284	0.8984 ±0.0266	0.9003 ±0.0219	0.8897 ±0.029	0.881 ±0.0327	0.8731 ±0.0271	0.8758 ±0.0315	0.8686 ±0.033	0.864 ±0.032	0.8322 ±0.0378	**0.9045** ±0.0252
D3	**0.9203** ±0.0092	0.9178 ±0.0085	0.8893 ±0.0076	0.9048 ±0.009	0.905 ±0.0073	0.9069 ±0.0083	0.886 ±0.0088	0.8873 ±0.0087	0.8804 ±0.0079	0.8831 ±0.0079	0.8869 ±0.0099	0.8406 ±0.0144	0.8348 ±0.0135	0.9019 ±0.0095
D4	**0.8142** ±0.0692	0.8058 ±0.0518	0.8045 ±0.047	0.804 ±0.0486	0.7814 ±0.052	0.7952 ±0.0581	0.7928 ±0.0496	0.7911 ±0.0511	0.8091 ±0.055	0.8043 ±0.0517	0.7904 ±0.0561	0.7582 ±0.0628	0.7475 ±0.0627	0.7859 ±0.0599
D5	0.7205 ±0.0094	**0.7234** ±0.0132	0.6513 ±0.0125	0.6963 ±0.0109	0.7107 ±0.0133	0.696 ±0.0111	0.6844 ±0.0141	0.6637 ±0.0107	0.6428 ±0.0104	0.6615 ±0.011	0.6657 ±0.0098	0.4755 ±0.0201	0.6363 ±0.0107	-
D6	**0.9725** ±0.0236	0.9656 ±0.0346	0.9548 ±0.0365	0.9567 ±0.0392	0.9535 ±0.0368	0.9488 ±0.0363	0.9663 ±0.028	0.9386 ±0.0412	0.936 ±0.042	0.9176 ±0.0481	0.9323 ±0.0492	0.9585 ±0.0346	0.9057 ±0.0585	0.9485 ±0.0366
D7	0.9433 ±0.0336	0.8457 ±0.0416	0.8997 ±0.0394	0.5917 ±0.0089	0.6366 ±0.0486	0.6368 ±0.0754	0.6803 ±0.0699	0.691 ±0.0277	0.6735 ±0.0244	0.663 ±0.022	0.6451 ±0.0324	0.6157 ±0.035	0.5847 ±0.0232	**0.9704** ±0.0255
D8	**0.7971** ±0.0499	0.7888 ±0.056	0.6205 ±0.0387	0.7869 ±0.0601	0.7515 ±0.0592	0.7618 ±0.0629	0.7649 ±0.0599	0.6133 ±0.034	0.5882 ±0.0447	0.6427 ±0.04	0.5973 ±0.0419	0.6379 ±0.0692	0.5938 ±0.0478	0.771 ±0.0583
D9	0.9542 ±0.075	0.9401 ±0.0897	0.9414 ±0.0952	0.9167 ±0.0985	0.9038 ±0.1062	0.8767 ±0.1181	0.9367 ±0.095	**0.9644** ±0.0684	0.8793 ±0.1083	0.888 ±0.1082	0.9333 ±0.1	0.8992 ±0.101	0.9277 ±0.1017	0.9048 ±0.1163
D10	**0.8281**	0.7958	0.6657	0.7943	0.7551	0.7539	0.7722	0.6311	0.654	0.6417	0.664	0.6345	0.7082	0.746

	±0.061	±0.0959	±0.0752	±0.1037	±0.1091	±0.1013	±0.1136	±0.0676	±0.0659	±0.0705	±0.0856	±0.1206	±0.088	±0.0905
D11	0.985	0.9835	0.9788	0.9815	0.9867	0.9865	**0.9877**	0.9839	0.9659	0.9748	0.9805	0.9389	0.9811	0.9859
	±0.0079	±0.0088	±0.0103	±0.0092	±0.0079	±0.0077	±0.0078	±0.0097	±0.0127	±0.0127	±0.0113	±0.0298	±0.0122	±0.0084
D12	0.8303	0.8329	0.7139	0.8257	0.8449	0.8396	0.8157	0.6734	0.6743	0.7006	0.6851	0.7091	0.6986	**0.8482**
	±0.0287	±0.0259	±0.0278	±0.0246	±0.0235	±0.0235	±0.0287	±0.0252	±0.0254	±0.0232	±0.0203	±0.034	±0.0335	±0.0245
D13	0.3667	0.3324	0.3351	0.3568	0.3629	0.3512	**0.369**	0.3234	0.345	0.342	0.3034	0.3145	0.3142	0.3416
	±0.1072	±0.0794	±0.0813	±0.1095	±0.108	±0.1038	±0.1072	±0.0641	±0.067	±0.0749	±0.0612	±0.0935	±0.0765	±0.1092
D14	**0.8811**	0.88	0.8535	0.8624	0.8086	0.7443	0.8441	0.8223	0.7932	0.8194	0.7686	0.6967	0.7727	0.7683
	±0.0505	±0.0482	±0.0572	±0.0668	±0.0877	±0.09	±0.0912	±0.0703	±0.0686	±0.0666	±0.0844	±0.1131	±0.0597	±0.0979
D15	0.9981	0.9954	0.9782	0.987	0.9998	0.9996	0.9997	0.9907	0.8562	0.94	0.9387	0.9219	0.9864	**0.9999**
	±0.0016	±0.0023	±0.0086	±0.0024	±0.0003	±0.0004	±0.0004	±0.0046	±0.0115	±0.0145	±0.0414	±0.0148	±0.0072	±0.0003
D16	0.3878	0.3617	0.2831	**0.3969**	0.3819	0.394	0.3516	0.2795	0.2804	0.2897	0.2282	0.238	0.3417	0.3875
	±0.068	±0.04	±0.0255	±0.086	±0.0606	±0.0694	±0.0625	±0.0218	±0.0215	±0.0162	±0.0277	±0.0244	±0.0334	±0.0795
D17	0.927	0.9244	0.7205	**0.9281**	0.9257	0.9172	0.897	0.6907	0.6952	0.7035	0.7122	0.5807	0.7067	0.9219
	±0.0146	±0.0176	±0.0133	±0.0454	±0.0164	±0.0258	±0.0197	±0.0138	±0.011	±0.0125	±0.0129	±0.0545	±0.0172	±0.0198
D18	**0.6259**	0.5944	0.4772	0.57	0.5613	0.4397	0.5327	0.4449	0.424	0.4562	0.3229	0.3425	0.4683	0.4875
	±0.0574	±0.0671	±0.0496	±0.081	±0.0675	±0.0447	±0.0812	±0.0401	±0.0337	±0.0439	±0.0416	±0.0541	±0.0465	±0.0618
D19	**0.8974**	0.8856	0.6603	0.8758	0.871	0.8659	0.8735	0.6949	0.5667	0.6042	0.62	0.688	0.8154	0.8761
	±0.0408	±0.0371	±0.037	±0.0346	±0.0325	±0.0345	±0.0334	±0.0345	±0.0201	±0.0256	±0.0388	±0.0604	±0.0323	±0.0368
D20	0.5232	**0.5326**	0.2337	0.5254	0.4833	0.4781	0.417	0.2227	0.2133	0.2315	0.1883	0.1925	0.3422	0.4949
	±0.0214	±0.0276	±0.0118	±0.0275	±0.0321	±0.0281	±0.0372	±0.0139	±0.0126	±0.0135	±0.0181	±0.0123	±0.0151	±0.035
D21	**0.9593**	0.9536	0.953	0.953	0.9526	0.9505	0.9568	0.9498	0.9512	0.9557	0.9487	0.953	0.9458	0.9541
	±0.0191	±0.0375	±0.0359	±0.0342	±0.0356	±0.0364	±0.0327	±0.0385	±0.0371	±0.0351	±0.0411	±0.0338	±0.0393	±0.0344
D22	0.986	**0.9864**	0.9692	0.9668	0.9141	0.9341	0.9387	0.8289	0.9644	0.9145	0.8304	0.9456	0.8243	0.9253
	±0.0103	±0.0088	±0.0154	±0.0124	±0.0189	±0.0152	±0.0188	±0.028	±0.0116	±0.0196	±0.0281	±0.0172	±0.027	±0.0224
D23	**0.9925**	0.9867	0.9799	0.9761	0.981	0.9835	0.9668	0.9324	0.9761	0.9669	0.9331	0.9776	0.9316	0.9884
	±0.0023	±0.0033	±0.0041	±0.0044	±0.0044	±0.0043	±0.0064	±0.0084	±0.0042	±0.0053	±0.008	±0.0047	±0.0081	±0.0038
D24	0.9787	**0.9788**	0.976	0.9746	0.9626	0.9665	0.9601	0.8953	0.9748	0.9474	0.8939	0.9577	0.8916	-
	±0.0032	±0.0047	±0.0055	±0.0064	±0.0071	±0.007	±0.0081	±0.0124	±0.0054	±0.0085	±0.0128	±0.0093	±0.014	
D25	**0.9748**	0.972	0.9688	0.9652	0.9635	0.9666	0.9476	0.9015	0.9663	0.9477	0.9014	0.9584	0.8994	-

	±0.0066	±0.0065	±0.0062	±0.0078	±0.0076	±0.0066	±0.0104	±0.014	±0.0079	±0.0099	±0.0142	±0.0095	±0.0143	
D26	0.9394 ±0.009	**0.9407** ±0.012	0.927 ±0.0138	0.9221 ±0.0133	0.9172 ±0.0153	0.8558 ±0.0185	0.92 ±0.0146	0.8777 ±0.0198	0.9226 ±0.0135	0.8903 ±0.0163	0.8771 ±0.022	0.8969 ±0.0174	0.8762 ±0.0194	-
D27	0.8638 ±0.0093	**0.8653** ±0.0099	0.8577 ±0.011	0.8492 ±0.0116	0.8523 ±0.0107	0.8535 ±0.0097	0.851 ±0.0098	0.7705 ±0.0153	0.8485 ±0.0114	0.8336 ±0.012	0.7641 ±0.0122	0.81 ±0.013	0.748 ±0.0143	0.8527 ±0.0128
D28	**0.8616** ±0.0103	0.8611 ±0.0101	0.8568 ±0.011	0.8464 ±0.0103	0.8463 ±0.011	0.8488 ±0.0105	0.8494 ±0.0098	0.7819 ±0.0153	0.8475 ±0.0126	0.8343 ±0.0115	0.774 ±0.0168	0.818 ±0.0108	0.757 ±0.0147	0.8498 ±0.012
D29	**0.9869** ±0.0036	0.9867 ±0.0031	0.9837 ±0.0033	0.9801 ±0.0039	0.9776 ±0.0042	0.9799 ±0.0039	0.97 ±0.0052	0.925 ±0.0096	0.981 ±0.0037	0.9598 ±0.0062	0.9239 ±0.0098	0.9718 ±0.0045	0.9121 ±0.0094	0.9792 ±0.005
D30	**0.9523** ±0.0097	0.9488 ±0.0131	0.9403 ±0.0134	0.9329 ±0.0161	0.9217 ±0.0164	0.9284 ±0.0149	0.9135 ±0.0159	0.8062 ±0.0253	0.9322 ±0.0145	0.8754 ±0.0185	0.7903 ±0.0238	0.8833 ±0.0177	0.7753 ±0.0264	-
D31	**0.8108** ±0.0273	0.78 ±0.0309	0.7542 ±0.0332	0.7426 ±0.031	0.7591 ±0.0273	0.7606 ±0.0309	0.7601 ±0.0316	0.7222 ±0.0313	0.7374 ±0.0282	0.7367 ±0.0271	0.7166 ±0.0342	0.7414 ±0.0358	0.7086 ±0.0367	0.7875 ±0.0328
D32	**0.5493** ±0.0094	0.5464 ±0.0166	0.5411 ±0.0143	0.5405 ±0.0152	0.5418 ±0.0141	0.545 ±0.0163	0.5453 ±0.0173	0.5221 ±0.016	0.5395 ±0.0151	0.5381 ±0.0151	0.5211 ±0.0141	0.5199 ±0.0164	0.4946 ±0.0171	0.5412 ±0.014
D33	**0.9925** ±0.0019	0.992 ±0.0019	0.987 ±0.0019	0.9771 ±0.0033	0.9895 ±0.0023	0.9886 ±0.0023	0.9737 ±0.0035	0.9633 ±0.0055	0.975 ±0.0028	0.9657 ±0.0041	0.9599 ±0.0055	0.9431 ±0.0058	0.9362 ±0.0054	-
D34	**0.9986** ±0.0022	0.9972 ±0.0033	0.9911 ±0.0071	0.9894 ±0.007	0.973 ±0.009	0.9789 ±0.0103	0.978 ±0.0093	0.9486 ±0.0179	0.9814 ±0.0085	0.9562 ±0.0156	0.9302 ±0.0228	0.9581 ±0.0152	0.8752 ±0.0233	0.9798 ±0.0109
D35	0.5277 ±0.0232	0.5223 ±0.0292	**0.5457** ±0.0289	0.4993 ±0.0284	0.5125 ±0.0244	0.5219 ±0.0257	0.5194 ±0.0214	0.4922 ±0.0213	0.514 ±0.0231	0.5049 ±0.0244	0.4793 ±0.0258	0.4551 ±0.0312	0.4745 ±0.0267	0.514 ±0.0282

Table S19 The macro-AUC results of the performance comparison experiment

Dataset	MDGP-Forest	DeepForest	imbalance-DF	RandomForest	XGBoost	lightGBM	CatBoost	SelfPacedEnsemble	BalancedRandomForest	EasyEnsemble	BalanceCascade	RUSBoost	SMOTEBoost	OpenFE
D1	0.992 ±0.0041	0.9941 ±0.002	0.9942 ±0.0018	0.992 ±0.002	**0.9945** ±0.0018	0.9942 ±0.0018	0.9936 ±0.0017	0.979 ±0.0056	0.9921 ±0.0019	0.9893 ±0.0028	0.9807 ±0.0054	0.9697 ±0.0058	0.9774 ±0.0199	-
D2	0.9725 ±0.0053	**0.9764** ±0.0097	0.9713 ±0.0115	0.9702 ±0.0133	0.9714 ±0.0136	0.9709 ±0.0136	0.9725 ±0.0108	0.9612 ±0.0168	0.9724 ±0.0114	0.9671 ±0.0147	0.9604 ±0.0161	0.9593 ±0.0109	0.9137 ±0.024	0.9709 ±0.0135
D3	0.9919 ±0.0014	**0.9921** ±0.0012	0.9888 ±0.0016	0.9888 ±0.0016	0.9903 ±0.0013	0.9904 ±0.0013	0.9869 ±0.0014	0.988 ±0.0017	0.9886 ±0.0014	0.9876 ±0.0018	0.9876 ±0.0017	0.9754 ±0.0033	0.9289 ±0.0068	0.9898 ±0.0016

D4	**0.9554** ±0.0247	0.9486 ±0.0187	0.9478 ±0.0177	0.9466 ±0.0172	0.9401 ±0.0188	0.9428 ±0.0187	0.9472 ±0.0164	0.9337 ±0.0204	0.9439 ±0.0185	0.9404 ±0.0193	0.9305 ±0.0207	0.9253 ±0.0299	0.8497 ±0.0398	0.9421 ±0.0183
D5	0.9357 ±0.0047	**0.9386** ±0.004	0.9181 ±0.005	0.9215 ±0.0042	0.9213 ±0.0047	0.9263 ±0.0042	0.9031 ±0.0044	0.9176 ±0.0045	0.9072 ±0.0045	0.9149 ±0.0046	0.9118 ±0.0041	0.7607 ±0.0101	0.8218 ±0.0063	-
D6	**0.9992** ±0.0014	0.998 ±0.0034	0.998 ±0.0035	0.9967 ±0.0049	0.9949 ±0.0051	0.9959 ±0.0044	0.9964 ±0.0054	0.9948 ±0.0056	0.9979 ±0.0028	0.9938 ±0.0072	0.9947 ±0.0061	0.995 ±0.0059	0.9266 ±0.0456	0.9959 ±0.0043
D7	**0.9963** ±0.0023	0.9837 ±0.008	0.9951 ±0.0045	0.8258 ±0.023	0.9269 ±0.0154	0.9218 ±0.0167	0.9474 ±0.0131	0.8781 ±0.0185	0.8556 ±0.0222	0.8699 ±0.0196	0.8713 ±0.0195	0.8751 ±0.0214	0.7474 ±0.0263	0.9919 ±0.0069
D8	0.8673 ±0.0439	0.8335 ±0.0383	0.8395 ±0.0323	0.8277 ±0.0389	0.8514 ±0.0343	0.8555 ±0.0333	0.8483 ±0.0386	0.844 ±0.0316	0.8081 ±0.0404	0.8496 ±0.0354	0.8177 ±0.0304	0.7484 ±0.0405	0.7345 ±0.046	**0.8759** ±0.0369
D9	0.9995 ±0.0017	0.9995 ±0.0017	0.9996 ±0.0016	0.9994 ±0.0016	0.9994 ±0.0025	0.9815 ±0.0214	0.9981 ±0.0056	**0.9998** ±0.001	0.9932 ±0.0138	0.9985 ±0.0044	0.9969 ±0.0086	0.9982 ±0.0041	0.9758 ±0.0388	0.989 ±0.0176
D10	**0.959** ±0.0151	0.9568 ±0.0187	0.909 ±0.0297	0.95 ±0.02	0.9365 ±0.0236	0.9343 ±0.0259	0.949 ±0.0211	0.9212 ±0.0207	0.9133 ±0.0216	0.914 ±0.0251	0.9026 ±0.0343	0.8836 ±0.03	0.8398 ±0.0429	0.9291 ±0.0259
D11	**0.9976** ±0.0022	0.9975 ±0.0025	0.9968 ±0.0022	0.9965 ±0.0028	0.9975 ±0.002	0.9962 ±0.0034	0.9971 ±0.0022	0.9954 ±0.0044	0.9962 ±0.0024	0.9959 ±0.0044	0.9946 ±0.0052	0.9882 ±0.0067	0.9926 ±0.0083	0.9962 ±0.0032
D12	0.9662 ±0.0091	0.9634 ±0.0083	0.944 ±0.0079	0.9577 ±0.0077	0.9642 ±0.007	0.9641 ±0.0068	0.9592 ±0.0071	0.9508 ±0.0072	0.9403 ±0.0089	0.9444 ±0.0092	0.941 ±0.0069	0.8814 ±0.0151	0.8686 ±0.0207	**0.9672** ±0.0067
D13	**0.7944** ±0.0338	0.7875 ±0.0379	0.7443 ±0.0408	0.7656 ±0.0423	0.776 ±0.0381	0.7734 ±0.0409	0.7937 ±0.0414	0.703 ±0.0516	0.7275 ±0.045	0.7286 ±0.0456	0.6498 ±0.05	0.7181 ±0.0561	0.6251 ±0.0445	0.7703 ±0.0434
D14	0.9742 ±0.0109	**0.9763** ±0.0091	0.9711 ±0.011	0.9732 ±0.0104	0.9683 ±0.0138	0.9611 ±0.0224	0.974 ±0.0114	0.9721 ±0.013	0.9701 ±0.0118	0.9731 ±0.0105	0.9577 ±0.0173	0.9559 ±0.0148	0.885 ±0.0292	0.9666 ±0.0171
D15	0.9999 ±0.0004	0.9998 ±0.0007	0.9993 ±0.0003	0.9994 ±0.0002	**1** ±0.0	**1** ±0.0	**1** ±0.0	0.9999 ±0.0001	0.9938 ±0.001	0.9983 ±0.0004	0.9991 ±0.0006	0.9897 ±0.0028	0.9946 ±0.0028	**1** ±0.0
D16	0.8777 ±0.0205	**0.8783** ±0.0182	0.7774 ±0.0332	0.8099 ±0.0379	0.8354 ±0.03	0.8274 ±0.0329	0.8296 ±0.024	0.7509 ±0.0392	0.7537 ±0.0337	0.7722 ±0.0323	0.7128 ±0.0346	0.692 ±0.0474	0.6745 ±0.0378	0.8379 ±0.0329
D17	**0.9903** ±0.0024	0.9894 ±0.0046	0.9746 ±0.0071	0.9774 ±0.0068	0.9902 ±0.0031	0.9886 ±0.0037	0.9865 ±0.0036	0.9699 ±0.0078	0.9649 ±0.0079	0.9663 ±0.0087	0.9653 ±0.0079	0.9046 ±0.0151	0.9057 ±0.0112	0.9887 ±0.0032
D18	**0.8895** ±0.0152	0.8826 ±0.0217	0.8532 ±0.0237	0.868 ±0.0222	0.8743 ±0.0238	0.828 ±0.0419	0.8773 ±0.0187	0.8435 ±0.022	0.8389 ±0.0229	0.851 ±0.0231	0.7987 ±0.0221	0.7947 ±0.0244	0.7413 ±0.0306	0.8384 ±0.0522

D19	0.9921 ±0.0039	0.992 ±0.0042	0.9845 ±0.0071	0.988 ±0.0055	**0.9938** ±0.0023	**0.9938** ±0.0022	0.9936 ±0.0029	0.9853 ±0.006	0.9851 ±0.0048	0.9866 ±0.006	0.9783 ±0.0072	0.9752 ±0.0116	0.9878 ±0.0051	0.9936 ±0.0031
D20	**0.871** ±0.0336	0.8625 ±0.0369	0.7328 ±0.0356	0.811 ±0.0414	0.8302 ±0.0387	0.8257 ±0.0432	0.8087 ±0.0256	0.6878 ±0.0434	0.6936 ±0.0407	0.7178 ±0.0404	0.6588 ±0.038	0.6285 ±0.049	0.6972 ±0.041	0.8083 ±0.0483
D21	**0.9981** ±0.0037	0.9954 ±0.0059	0.9946 ±0.0069	0.9919 ±0.0118	0.9856 ±0.0174	0.9901 ±0.0106	0.9938 ±0.0091	0.9683 ±0.0275	0.9909 ±0.0122	0.9897 ±0.0151	0.9682 ±0.0275	0.9839 ±0.0198	0.9643 ±0.0278	0.9902 ±0.0106
D22	**0.9997** ±0.0005	**0.9997** ±0.0004	0.9981 ±0.0019	0.9986 ±0.001	0.9954 ±0.0019	0.9965 ±0.0019	0.9975 ±0.0013	0.9195 ±0.0138	0.9987 ±0.0007	0.9955 ±0.0019	0.9196 ±0.0137	0.9979 ±0.0013	0.9219 ±0.0138	0.9958 ±0.0022
D23	**0.9999** ±0.0001	0.9998 ±0.0002	0.9997 ±0.0002	0.9995 ±0.0002	0.9997 ±0.0002	0.9997 ±0.0001	0.9991 ±0.0003	0.973 ±0.0038	0.9995 ±0.0001	0.9989 ±0.0004	0.973 ±0.0039	0.9995 ±0.0002	0.9774 ±0.0041	**0.9999** ±0.0001
D24	0.9992 ±0.0005	**0.9993** ±0.0004	0.9985 ±0.0009	0.9983 ±0.0009	0.9985 ±0.0007	0.9986 ±0.0006	0.9986 ±0.0007	0.9568 ±0.0063	0.9981 ±0.0012	0.9963 ±0.0017	0.9569 ±0.0063	0.9982 ±0.0007	0.968 ±0.0059	-
D25	**0.999** ±0.0006	0.9988 ±0.0006	0.998 ±0.001	0.9978 ±0.001	0.9984 ±0.0008	0.9985 ±0.0007	0.9975 ±0.0009	0.9624 ±0.0071	0.9979 ±0.001	0.9967 ±0.0016	0.9623 ±0.0073	0.9981 ±0.0008	0.972 ±0.0065	-
D26	0.996 ±0.0016	**0.9963** ±0.0014	0.994 ±0.0027	0.9937 ±0.0021	0.994 ±0.0018	0.9807 ±0.0042	0.9929 ±0.0024	0.9441 ±0.0121	0.9938 ±0.002	0.9864 ±0.0051	0.9441 ±0.0121	0.988 ±0.0035	0.9504 ±0.0123	-
D27	0.9719 ±0.0034	**0.9722** ±0.0038	0.969 ±0.0045	0.9643 ±0.0043	0.9674 ±0.004	0.9673 ±0.004	0.9652 ±0.0042	0.9167 ±0.0108	0.9643 ±0.004	0.957 ±0.005	0.909 ±0.0115	0.947 ±0.0047	0.8405 ±0.0109	0.9688 ±0.0041
D28	**0.9705** ±0.0027	0.9698 ±0.0032	0.9687 ±0.0036	0.964 ±0.0038	0.9662 ±0.0031	0.9668 ±0.0034	0.9655 ±0.0033	0.9264 ±0.01	0.964 ±0.0039	0.9585 ±0.0046	0.9193 ±0.0112	0.9525 ±0.0043	0.8448 ±0.0107	0.9677 ±0.0032
D29	**0.9998** ±0.0002	0.9997 ±0.0002	0.9989 ±0.0007	0.9995 ±0.0002	0.9995 ±0.0002	0.9995 ±0.0002	0.9992 ±0.0003	0.9879 ±0.0038	0.9995 ±0.0002	0.9979 ±0.0008	0.9862 ±0.0045	0.9991 ±0.0003	0.9708 ±0.004	0.9995 ±0.0002
D30	**0.9976** ±0.001	**0.9976** ±0.0013	0.9952 ±0.0022	0.9944 ±0.0023	0.9952 ±0.0019	0.9959 ±0.0018	0.9948 ±0.0019	0.9592 ±0.0101	0.9948 ±0.0018	0.9861 ±0.0042	0.9526 ±0.0109	0.9901 ±0.0029	0.9218 ±0.0137	-
D31	**0.9548** ±0.0104	0.9435 ±0.0094	0.9335 ±0.0092	0.9289 ±0.0094	0.9318 ±0.0086	0.9318 ±0.0082	0.9317 ±0.0103	0.9064 ±0.0166	0.9279 ±0.0099	0.9249 ±0.0104	0.9023 ±0.016	0.9082 ±0.0149	0.8309 ±0.0232	0.9456 ±0.0114
D32	**0.7593** ±0.0081	0.7554 ±0.0098	0.7548 ±0.0087	0.745 ±0.0101	0.7484 ±0.0106	0.7507 ±0.0116	0.754 ±0.01	0.7196 ±0.0112	0.7472 ±0.011	0.7417 ±0.011	0.7153 ±0.0109	0.7039 ±0.0117	0.637 ±0.0143	0.7483 ±0.0102
D33	0.9998 ±0.0001	0.9998 ±0.0001	0.9996 ±0.0001	0.9994 ±0.0001	**0.9999** ±0.0001	0.9998 ±0.0001	0.9992 ±0.0002	0.9981 ±0.0006	0.9993 ±0.0002	0.9985 ±0.0003	0.9975 ±0.0008	0.9955 ±0.001	0.9749 ±0.003	-

D34	**1** ±0.0	**1** ±0.0	0.9998 ±0.0003	0.9998 ±0.0001	0.9992 ±0.0006	0.9996 ±0.0004	0.9996 ±0.0003	0.9967 ±0.002	0.9996 ±0.0003	0.998 ±0.0013	0.9948 ±0.0028	0.9984 ±0.0011	0.9505 ±0.0125	0.9995 ±0.0006
D35	0.7262 ±0.0127	0.7201 ±0.0227	**0.731** ±0.0227	0.6952 ±0.0205	0.7155 ±0.0216	0.7188 ±0.0213	0.7213 ±0.0206	0.6815 ±0.0193	0.7061 ±0.0213	0.7061 ±0.0193	0.6564 ±0.0209	0.6436 ±0.0271	0.6538 ±0.0234	0.7169 ±0.02

Table S20 The ovo-AUC results of the performance comparison experiment

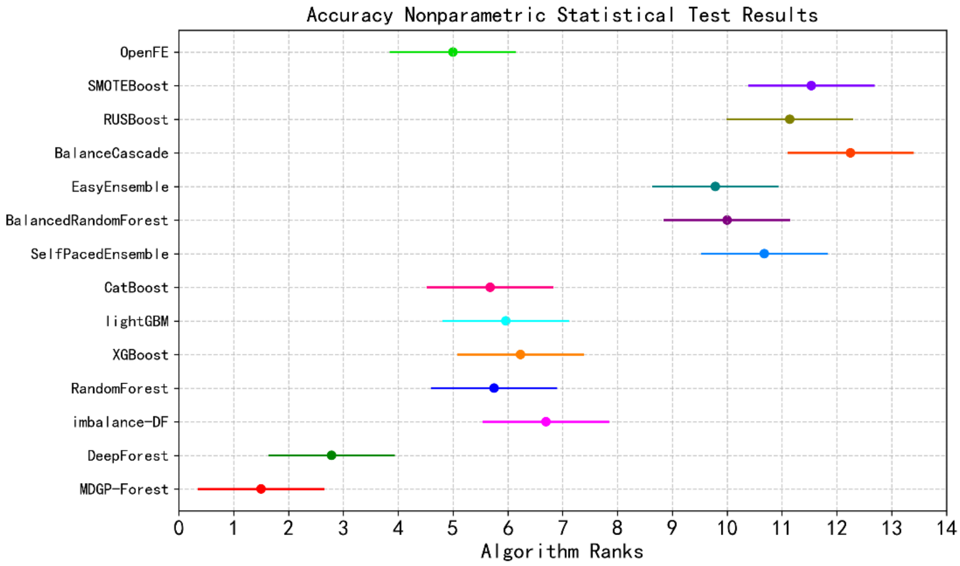
Dataset	MDGP-Forest	DeepForest	imbalance-DF	RandomForest	XGBoost	lightGBM	CatBoost	SelfPacedEnsemble	BalancedRandomForest	EasyEnsemble	BalanceCascade	RUSBoost	SMOTEBoost	OpenFE
D1	0.9917 ±0.0044	0.9938 ±0.0021	0.9944 ±0.0018	0.9916 ±0.0021	**0.9945** ±0.0018	0.9942 ±0.0019	0.9935 ±0.0018	0.9785 ±0.0058	0.9921 ±0.002	0.9888 ±0.0031	0.9802 ±0.0057	0.9695 ±0.0058	0.9767 ±0.0204	-
D2	0.9727 ±0.0053	**0.9763** ±0.01	0.9729 ±0.0112	0.9701 ±0.0137	0.9707 ±0.0146	0.9708 ±0.0141	0.9727 ±0.0107	0.9616 ±0.0168	0.9733 ±0.0115	0.9675 ±0.0149	0.9609 ±0.0161	0.9592 ±0.0108	0.9135 ±0.0245	0.9702 ±0.0142
D3	0.9915 ±0.0014	**0.9918** ±0.0012	0.9886 ±0.0016	0.9884 ±0.0016	0.9899 ±0.0013	0.99 ±0.0014	0.9863 ±0.0015	0.9878 ±0.0017	0.9884 ±0.0014	0.9873 ±0.0018	0.9874 ±0.0017	0.9748 ±0.0033	0.927 ±0.007	0.9894 ±0.0016
D4	**0.9398** ±0.0331	0.9306 ±0.0245	0.9315 ±0.0232	0.9282 ±0.0229	0.9202 ±0.0257	0.9239 ±0.0243	0.9293 ±0.0219	0.9153 ±0.0262	0.9277 ±0.0237	0.9237 ±0.0242	0.9121 ±0.0265	0.9049 ±0.0346	0.8326 ±0.0456	0.9228 ±0.0237
D5	0.9349 ±0.0042	**0.9378** ±0.004	0.9199 ±0.0049	0.9202 ±0.0043	0.9205 ±0.0048	0.925 ±0.0044	0.9012 ±0.0046	0.92 ±0.0045	0.9098 ±0.0044	0.9166 ±0.0045	0.9153 ±0.004	0.7557 ±0.01	0.8215 ±0.0063	-
D6	**0.9993** ±0.0011	0.9982 ±0.0031	0.9983 ±0.0029	0.9973 ±0.004	0.9956 ±0.0048	0.9965 ±0.0037	0.9966 ±0.0061	0.9956 ±0.0049	0.9982 ±0.0024	0.9948 ±0.006	0.9955 ±0.0051	0.9958 ±0.0051	0.9289 ±0.0451	0.9965 ±0.0036
D7	**0.9948** ±0.0032	0.9772 ±0.0108	0.9938 ±0.0051	0.7993 ±0.0253	0.9097 ±0.0194	0.9067 ±0.0203	0.9388 ±0.0152	0.8503 ±0.0243	0.8248 ±0.0285	0.8422 ±0.0245	0.8408 ±0.0256	0.8459 ±0.0242	0.6987 ±0.0294	0.9912 ±0.0072
D8	0.8649 ±0.0528	0.8355 ±0.0398	0.8585 ±0.0337	0.8267 ±0.0419	0.8514 ±0.0382	0.8551 ±0.0366	0.8476 ±0.0428	0.8606 ±0.0343	0.8269 ±0.0427	0.8633 ±0.0388	0.8437 ±0.0332	0.76 ±0.0449	0.7391 ±0.0494	**0.8744** ±0.0403
D9	0.999 ±0.004	0.999 ±0.0036	0.9993 ±0.0028	0.9988 ±0.0036	0.9987 ±0.0054	0.9699 ±0.0341	0.9972 ±0.0079	**0.9994** ±0.0027	0.9948 ±0.01	0.9984 ±0.0047	0.9952 ±0.0106	0.9975 ±0.0064	0.9726 ±0.0436	0.9817 ±0.0283
D10	**0.9628** ±0.0142	0.9586 ±0.0207	0.9315 ±0.0252	0.953 ±0.0202	0.9363 ±0.0269	0.9376 ±0.0294	0.9485 ±0.0217	0.9425 ±0.0178	0.9377 ±0.0187	0.9351 ±0.0232	0.9326 ±0.0274	0.8962 ±0.0279	0.8491 ±0.0452	0.9365 ±0.0267
D11	**0.9971** ±0.0028	0.997 ±0.0027	0.997 ±0.0023	0.9952 ±0.0033	0.9959 ±0.0033	0.995 ±0.0046	0.9958 ±0.0033	0.9956 ±0.0048	0.9966 ±0.0023	0.9962 ±0.0048	0.995 ±0.0054	0.9901 ±0.0062	0.9929 ±0.0081	0.9947 ±0.0047
D12	**0.9726**	0.9704	0.9605	0.9646	0.9704	0.9699	0.9655	0.9652	0.9564	0.9592	0.9599	0.8881	0.8758	0.9723

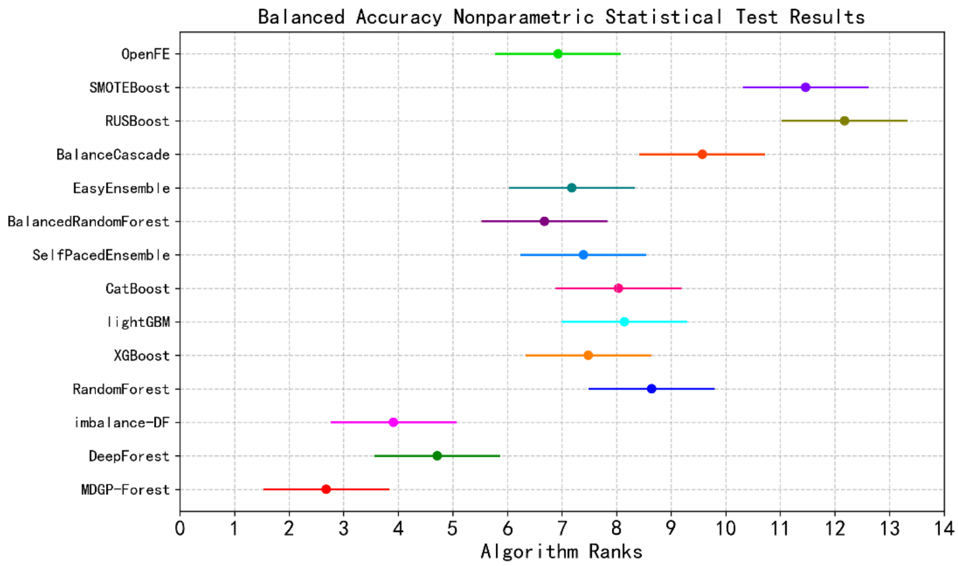
	±0.009	±0.0081	±0.007	±0.0074	±0.0069	±0.0066	±0.007	±0.0066	±0.0084	±0.0088	±0.0067	±0.0141	±0.0212	±0.0064
D13	<u>0.676</u>	0.6678	0.6633	0.6593	0.6672	0.6673	**0.6799**	0.6407	0.667	0.652	0.622	0.638	0.5881	0.6634
	±0.051	±0.0424	±0.0486	±0.0447	±0.0409	±0.0443	±0.0484	±0.0621	±0.0521	±0.0532	±0.0539	±0.0523	±0.0439	±0.0476
D14	0.9697	0.9712	0.9722	0.969	0.961	0.9438	0.9659	**0.9732**	0.9712	<u>0.9727</u>	0.9621	0.9469	0.8854	0.9545
	±0.0126	±0.0111	±0.0093	±0.0123	±0.0167	±0.0344	±0.0144	±0.013	±0.0107	±0.0115	±0.0156	±0.0184	±0.03	±0.0244
D15	<u>0.9999</u>	0.9996	0.9994	0.9979	**1**	**1**	**1**	<u>0.9999</u>	0.9954	0.9988	0.9993	0.9898	0.9938	**1**
	±0.0004	±0.0009	±0.0005	±0.0006	±0.0	±0.0	±0.0	±0.0	±0.0007	±0.0003	±0.0005	±0.0027	±0.0034	±0.0
D16	<u>0.8416</u>	**0.8437**	0.7929	0.7783	0.8001	0.7953	0.7978	0.7668	0.7684	0.7883	0.747	0.6952	0.6476	0.7988
	±0.0193	±0.0174	±0.0332	±0.0388	±0.0281	±0.0307	±0.0249	±0.0475	±0.036	±0.0342	±0.0404	±0.0428	±0.0358	±0.0345
D17	**0.9821**	0.9815	0.9738	0.969	<u>0.9819</u>	0.9799	0.9775	0.973	0.968	0.968	0.971	0.8954	0.8933	0.98
	±0.0039	±0.0059	±0.006	±0.0078	±0.0048	±0.0054	±0.005	±0.007	±0.007	±0.0081	±0.0068	±0.014	±0.0127	±0.0048
D18	0.8767	0.8667	<u>0.8796</u>	0.8561	0.8666	0.8035	0.8545	0.8786	0.8698	**0.8841**	0.8446	0.79	0.7459	0.8188
	±0.0183	±0.0271	±0.0201	±0.0235	±0.0266	±0.0347	±0.0217	±0.0189	±0.0201	±0.0202	±0.0218	±0.0251	±0.0336	±0.0448
D19	0.9847	0.9852	0.9873	0.9826	0.984	0.985	0.9854	<u>0.9887</u>	0.988	**0.9889**	0.987	0.9747	0.9813	0.9832
	±0.0075	±0.0061	±0.0049	±0.0068	±0.0057	±0.0071	±0.0059	±0.0054	±0.0051	±0.006	±0.0056	±0.0122	±0.0072	±0.0091
D20	**0.8458**	<u>0.8362</u>	0.7561	0.7875	0.8053	0.806	0.7741	0.7085	0.713	0.7322	0.6873	0.6313	0.677	0.7909
	±0.034	±0.0384	±0.046	±0.0413	±0.04	±0.0389	±0.0291	±0.0534	±0.0528	±0.0495	±0.05	±0.048	±0.0299	±0.0436
D21	**0.9981**	<u>0.9954</u>	0.9946	0.9919	0.9856	0.9901	0.9938	0.9683	0.9909	0.9897	0.9682	0.9839	0.9643	0.9902
	±0.0037	±0.0059	±0.0069	±0.0118	±0.0174	±0.0106	±0.0091	±0.0275	±0.0122	±0.0151	±0.0275	±0.0198	±0.0278	±0.0106
D22	**0.9997**	**0.9997**	0.9981	0.9986	0.9954	0.9965	0.9975	0.9195	<u>0.9987</u>	0.9955	0.9196	0.9979	0.9219	0.9958
	±0.0005	±0.0004	±0.0019	±0.001	±0.0019	±0.0019	±0.0013	±0.0138	±0.0007	±0.0019	±0.0137	±0.0013	±0.0138	±0.0022
D23	**0.9999**	<u>0.9998</u>	0.9997	0.9995	0.9997	0.9997	0.9991	0.973	0.9995	0.9989	0.973	0.9995	0.9774	**0.9999**
	±0.0001	±0.0002	±0.0002	±0.0002	±0.0002	±0.0001	±0.0003	±0.0038	±0.0001	±0.0004	±0.0039	±0.0002	±0.0041	±0.0001
D24	<u>0.9992</u>	**0.9993**	0.9985	0.9983	0.9985	0.9986	0.9986	0.9568	0.9981	0.9963	0.9569	0.9982	0.968	-
	±0.0005	±0.0004	±0.0009	±0.0009	±0.0007	±0.0006	±0.0007	±0.0063	±0.0012	±0.0017	±0.0063	±0.0007	±0.0059	
D25	**0.999**	<u>0.9988</u>	0.998	0.9978	0.9984	0.9985	0.9975	0.9624	0.9979	0.9967	0.9623	0.9981	0.972	-
	±0.0006	±0.0006	±0.001	±0.001	±0.0008	±0.0007	±0.0009	±0.0071	±0.001	±0.0016	±0.0073	±0.0008	±0.0065	
D26	<u>0.996</u>	**0.9963**	0.994	0.9937	0.994	0.9807	0.9929	0.9441	0.9938	0.9864	0.9441	0.988	0.9504	-
	±0.0016	±0.0014	±0.0027	±0.0021	±0.0018	±0.0042	±0.0024	±0.0121	±0.002	±0.0051	±0.0121	±0.0035	±0.0123	
D27	<u>0.972</u>	**0.9722**	0.9691	0.9644	0.9674	0.9674	0.9653	0.9167	0.9644	0.9571	0.909	0.947	0.8406	0.9689

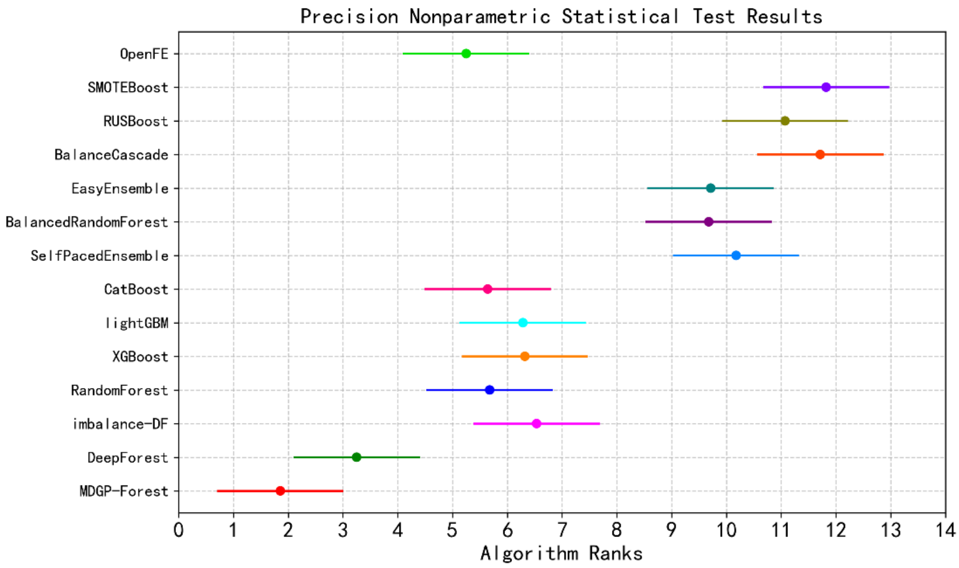
	±0.0034	±0.0037	±0.0045	±0.0043	±0.004	±0.004	±0.0042	±0.0108	±0.004	±0.005	±0.0115	±0.0047	±0.0109	±0.0041
D28	**0.9704**	0.9697	0.9687	0.9639	0.9661	0.9668	0.9655	0.9263	0.964	0.9584	0.9192	0.9525	0.8447	0.9677
	±0.0027	±0.0032	±0.0036	±0.0038	±0.0031	±0.0034	±0.0033	±0.01	±0.0039	±0.0046	±0.0112	±0.0043	±0.0107	±0.0032
D29	**0.9998**	0.9997	0.9989	0.9995	0.9995	0.9995	0.9992	0.9879	0.9995	0.9979	0.9862	0.9991	0.9708	0.9995
	±0.0002	±0.0002	±0.0007	±0.0002	±0.0002	±0.0002	±0.0003	±0.0038	±0.0002	±0.0008	±0.0045	±0.0003	±0.004	±0.0002
D30	**0.9976**	**0.9976**	0.9952	0.9944	0.9952	0.9959	0.9948	0.9591	0.9948	0.986	0.9526	0.9901	0.9217	-
	±0.001	±0.0013	±0.0022	±0.0023	±0.0019	±0.0018	±0.0019	±0.0101	±0.0018	±0.0042	±0.0109	±0.0029	±0.0137	
D31	**0.9555**	0.9444	0.9346	0.9299	0.9327	0.9328	0.9327	0.9072	0.9289	0.9258	0.9031	0.9091	0.8314	0.9464
	±0.0101	±0.0091	±0.0089	±0.0092	±0.0084	±0.008	±0.0101	±0.0166	±0.0097	±0.0102	±0.0159	±0.0147	±0.0231	±0.0112
D32	**0.7625**	0.7586	0.7581	0.7481	0.7515	0.7538	0.7571	0.7224	0.7503	0.7447	0.718	0.7066	0.6385	0.7515
	±0.0082	±0.0098	±0.0087	±0.01	±0.0106	±0.0116	±0.01	±0.0112	±0.011	±0.0109	±0.0109	±0.0116	±0.0143	±0.0102
D33	0.9998	0.9998	0.9996	0.9994	**0.9999**	0.9998	0.9992	0.9981	0.9993	0.9985	0.9975	0.9958	0.9748	-
	±0.0001	±0.0001	±0.0001	±0.0001	±0.0	±0.0001	±0.0002	±0.0007	±0.0002	±0.0003	±0.0009	±0.0009	±0.003	
D34	**1**	**1**	0.9998	0.9998	0.9992	0.9996	0.9996	0.9967	0.9996	0.998	0.9948	0.9984	0.9504	0.9995
	±0.0	±0.0	±0.0003	±0.0001	±0.0006	±0.0004	±0.0003	±0.002	±0.0003	±0.0013	±0.0028	±0.0011	±0.0125	±0.0006
D35	0.7212	0.7158	**0.7284**	0.69	0.709	0.7126	0.7154	0.6779	0.7033	0.7015	0.6554	0.6392	0.648	0.7105
	±0.0133	±0.0234	±0.0236	±0.0205	±0.0216	±0.0215	±0.0207	±0.0198	±0.0219	±0.0198	±0.0217	±0.0269	±0.0232	±0.0202

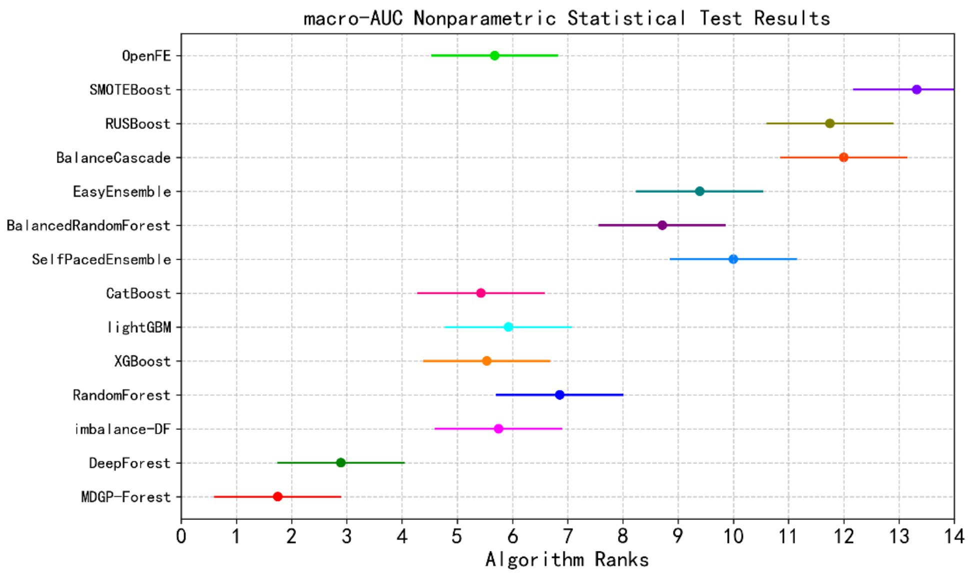
Table S21 The Friedman test statistic of the comparison experiment on all metrics

Evaluation Metrics	MDGP-Forest	DeepForest	imbalance-DF	RandomForest	XGBoost	lightGBM	CatBoost	SelfPaced Ensemble	Balanced RandomForest	EasyEnsemble	Balance Cascade	RUSBoost	SMOTEBoost	OpenFE	Friedman test statistic	p-value
F1	1.714	3.750	6.143	6.429	5.786	6.464	5.857	10.000	9.357	9.250	11.571	11.679	11.500	5.500	198.457	2.83E-35
G-mean	2.393	4.571	4.179	8.179	7.214	7.750	7.893	7.786	7.143	7.250	10.179	12.286	11.464	6.714	149.012	3.26E-25
MCC	1.536	2.964	6.250	5.964	6.464	6.000	5.643	10.179	9.464	9.750	12.000	11.714	11.893	5.179	233.008	2.13E-42
Accuracy	1.500	2.786	6.696	5.750	6.232	5.964	5.679	10.679	10.000	9.786	12.250	11.143	11.536	5.000	238.681	1.43E-43
Recall	2.679	4.714	3.911	8.643	7.482	8.143	8.036	7.393	6.679	7.179	9.571	12.179	11.464	6.929	142.263	7.40E-24
Precision	1.857	3.250	6.536	5.679	6.321	6.286	5.643	10.179	9.679	9.714	11.714	11.071	11.821	5.250	210.465	9.63E-38
macro-AUC	1.750	2.893	5.750	6.857	5.536	5.929	5.429	10.000	8.714	9.393	12.000	11.750	13.321	5.679	248.335	1.42E-45
ovo-AUC	2.196	3.196	4.750	7.893	6.500	6.536	6.179	8.893	7.357	8.643	11.250	11.714	13.643	6.250	211.901	4.88E-38


Accuracy Nonparametric Statistical Test Results
OpenFE
SMOTEBoost
RUSBoost
BalanceCascade
EasyEnsemble
BalancedRandomForest
SelfPacedEnsemble
CatBoost
lightGBM
XGBoost
RandomForest
imbalance-DF
DeepForest
MDGP-Forest
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14
Algorithm Ranks


Balanced Accuracy Nonparametric Statistical Test Results
OpenFE
SMOTEBoost
RUSBoost
BalanceCascade
EasyEnsemble
BalancedRandomForest
SelfPacedEnsemble
CatBoost
lightGBM
XGBoost
RandomForest
imbalance-DF
DeepForest
MDGP-Forest
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14
Algorithm Ranks


Precision Nonparametric Statistical Test Results
OpenFE
SMOTEBoost
RUSBoost
BalanceCascade
EasyEnsemble
BalancedRandomForest
SelfPacedEnsemble
CatBoost
lightGBM
XGBoost
RandomForest
imbalance-DF
DeepForest
MDGP-Forest
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14
Algorithm Ranks



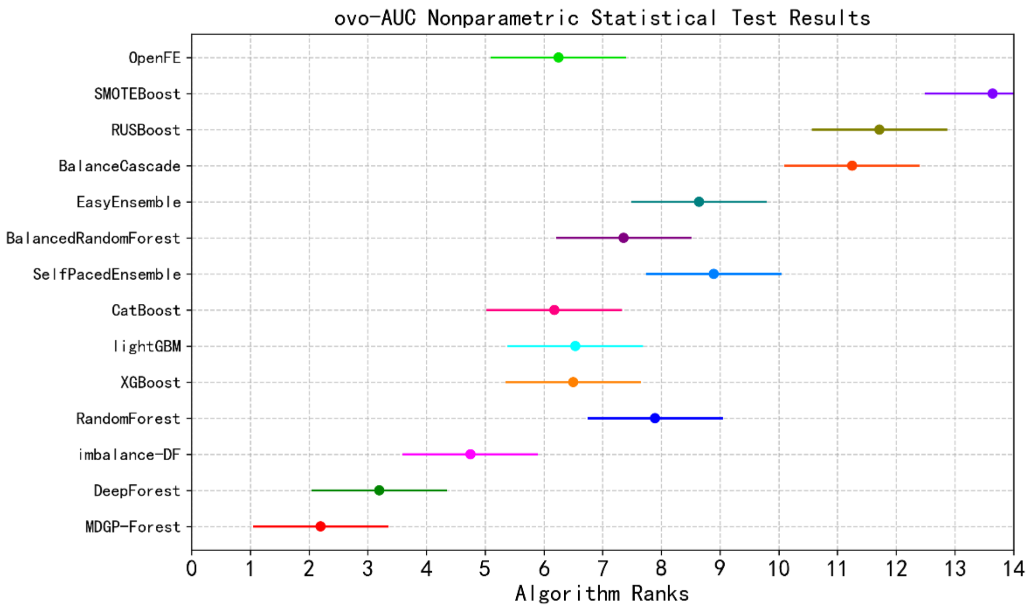
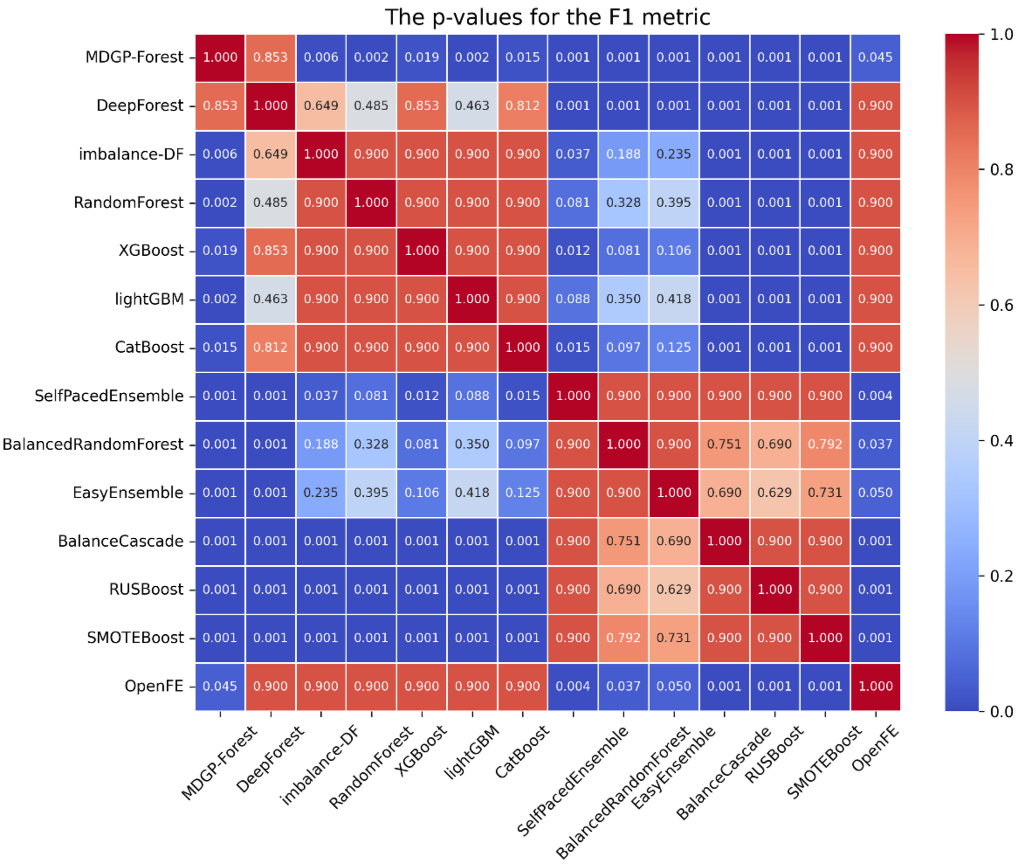
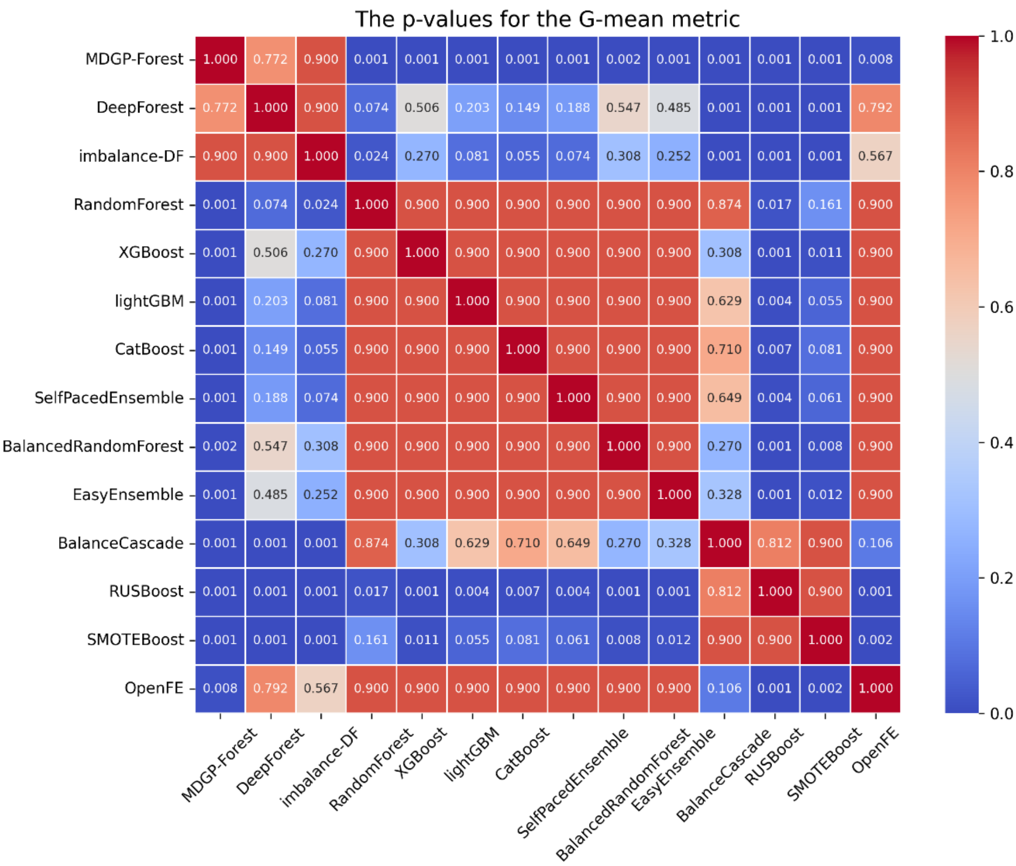


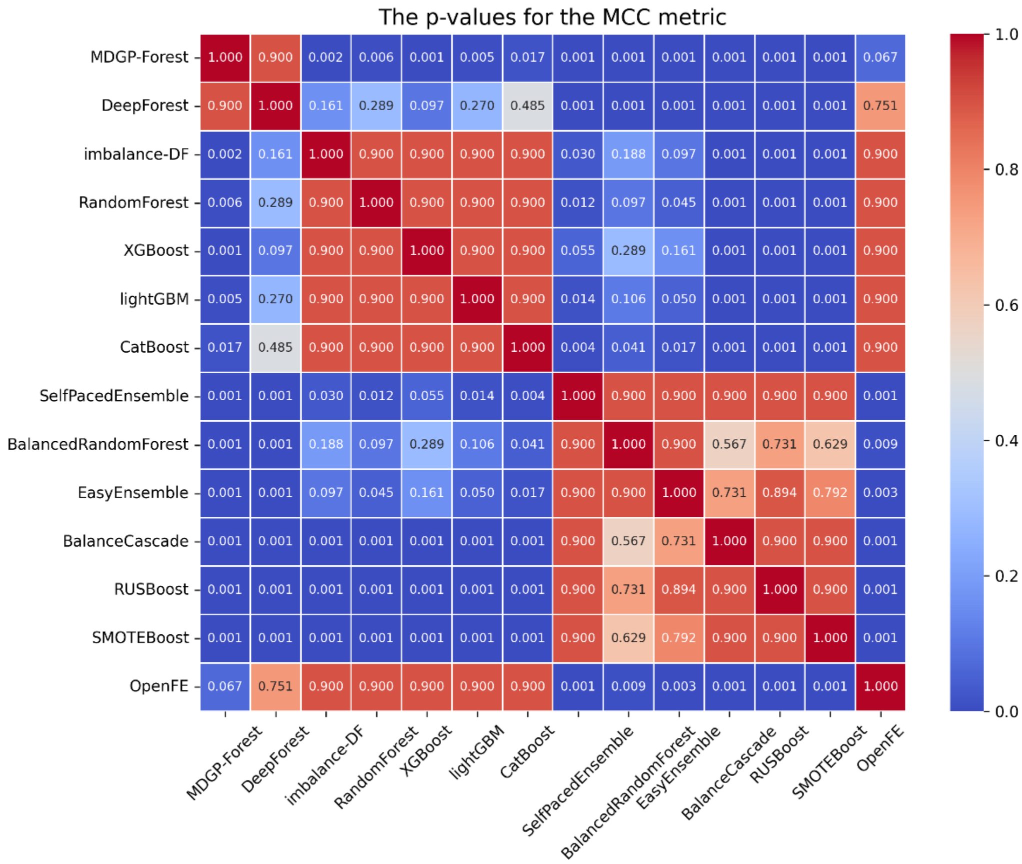
Fig S1 Nonparametric Statistical Test for the comparison experiment on Balanced Accuracy, Precision, macro-AUC, and ovo-AUC. The average rank of each algorithm is represented by a point, and the horizontal bar atop each point displays the range of the Nemenyi test value.

2.3 Statistical significance test results of the algorithms

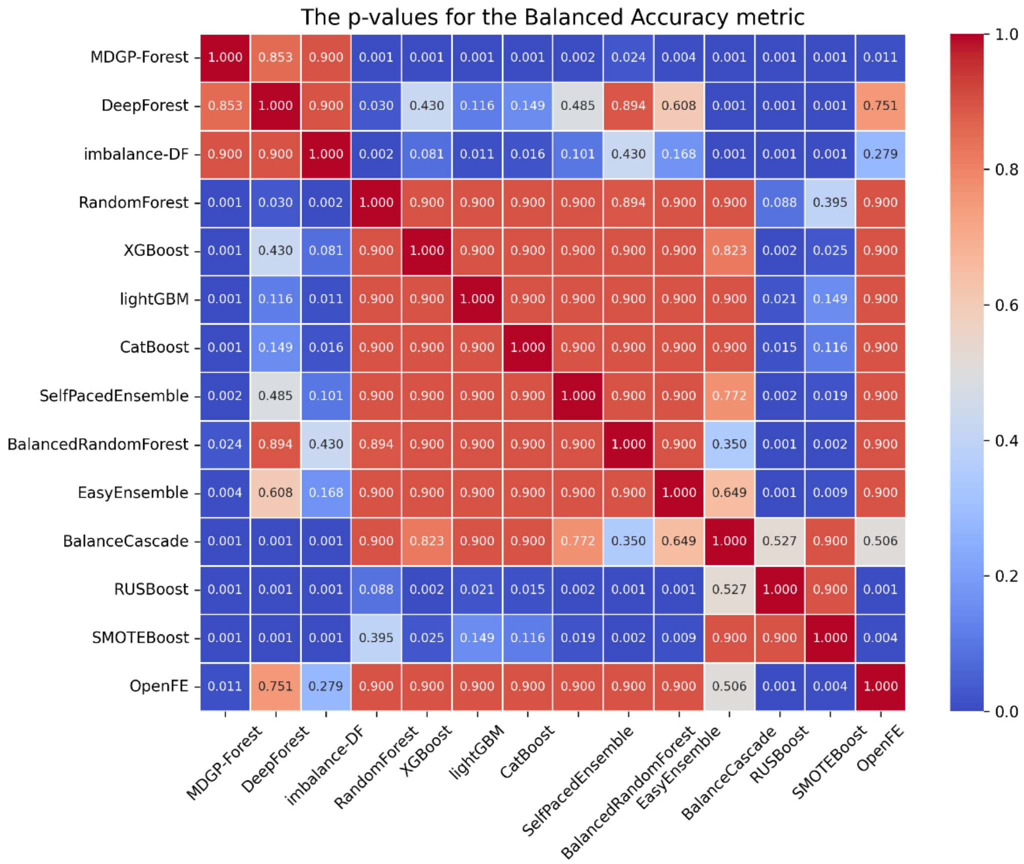
For the p-value matrix from the Nemenyi post-hoc test, we generated a significance heatmap to visualize inter-algorithm significant differences, as illustrated in Fig S2. Each row and column in the matrix represent an algorithm, with individual cells indicating the p-values between algorithm pairs. Color intensity corresponds to p-value magnitude, where cool colors denote smaller p-values (statistically significant differences), and warm colors indicate larger p-values (non-significant differences).

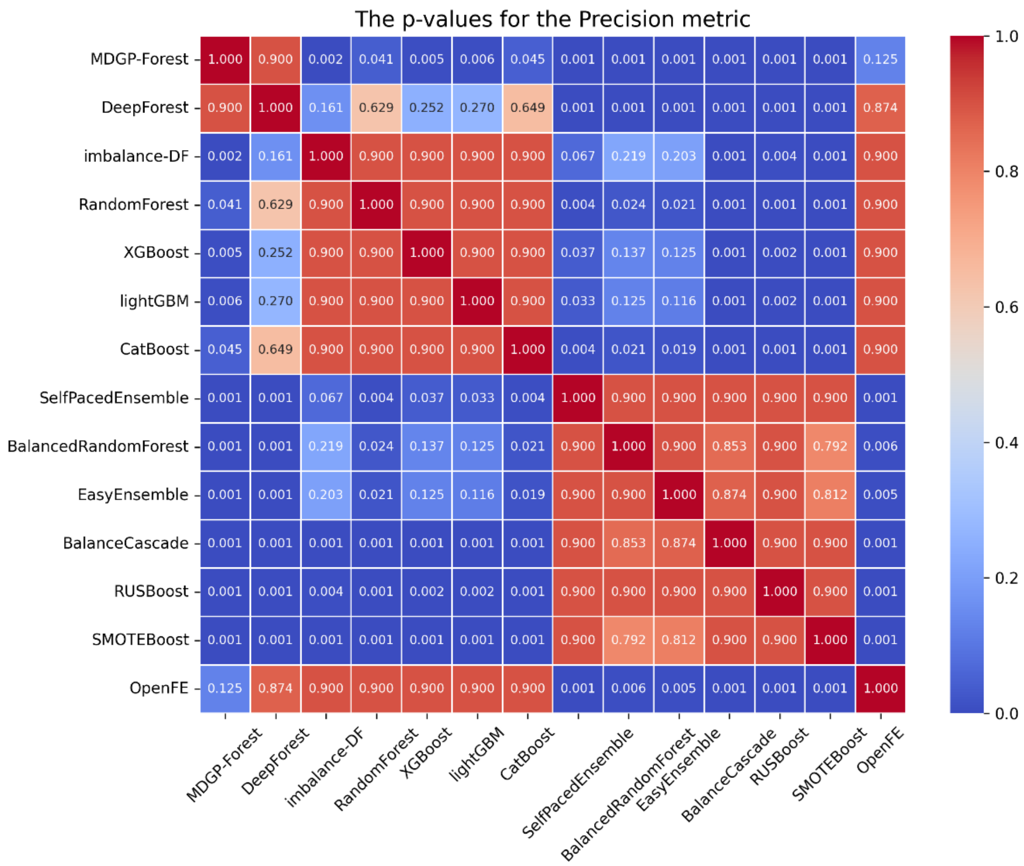

The p-values for the F1 metric
MDGP-Forest 1.000 0.853 0.006 0.002 0.019 0.002 0.015 0.001 0.001 0.001 0.001 0.001 0.001 0.045
DeepForest 0.853 1.000 0.649 0.485 0.853 0.463 0.812 0.001 0.001 0.001 0.001 0.001 0.001 0.900
imbalance-DF 0.006 0.649 1.000 0.900 0.900 0.900 0.900 0.037 0.188 0.235 0.001 0.001 0.001 0.900
RandomForest 0.002 0.485 0.900 1.000 0.900 0.900 0.900 0.081 0.328 0.395 0.001 0.001 0.001 0.900
XGBoost 0.019 0.853 0.900 0.900 1.000 0.900 0.900 0.012 0.081 0.106 0.001 0.001 0.001 0.900
lightGBM 0.002 0.463 0.900 0.900 0.900 1.000 0.900 0.088 0.350 0.418 0.001 0.001 0.001 0.900
CatBoost 0.015 0.812 0.900 0.900 0.900 0.900 1.000 0.015 0.097 0.125 0.001 0.001 0.001 0.900
SelfPacedEnsemble 0.001 0.001 0.037 0.081 0.012 0.088 0.015 1.000 0.900 0.900 0.900 0.900 0.900 0.004
BalancedRandomForest 0.001 0.001 0.188 0.328 0.081 0.350 0.097 0.900 1.000 0.900 0.751 0.690 0.792 0.037
EasyEnsemble 0.001 0.001 0.235 0.395 0.106 0.418 0.125 0.900 0.900 1.000 0.690 0.629 0.731 0.050
BalanceCascade 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.900 0.751 0.690 1.000 0.900 0.900 0.001
RUSBoost 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.900 0.690 0.629 0.900 1.000 0.900 0.001
SMOTEBoost 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.900 0.792 0.731 0.900 0.900 1.000 0.001
OpenFE 0.045 0.900 0.900 0.900 0.900 0.900 0.900 0.004 0.037 0.050 0.001 0.001 0.001 1.000
MDGP-Forest DeepForest imbalance-DF RandomForest XGBoost lightGBM CatBoost SelfPacedEnsemble BalancedRandomForest EasyEnsemble BalanceCascade RUSBoost SMOTEBoost OpenFE
1.0 0.8 0.6 0.4 0.2 0.0


The p-values for the G-mean metric
MDGP-Forest 1.000 0.772 0.900 0.001 0.001 0.001 0.001 0.001 0.002 0.001 0.001 0.001 0.001 0.008
DeepForest 0.772 1.000 0.900 0.074 0.506 0.203 0.149 0.188 0.547 0.485 0.001 0.001 0.001 0.792
imbalance-DF 0.900 0.900 1.000 0.024 0.270 0.081 0.055 0.074 0.308 0.252 0.001 0.001 0.001 0.567
RandomForest 0.001 0.074 0.024 1.000 0.900 0.900 0.900 0.900 0.900 0.900 0.874 0.017 0.161 0.900
XGBoost 0.001 0.506 0.270 0.900 1.000 0.900 0.900 0.900 0.900 0.900 0.308 0.001 0.011 0.900
lightGBM 0.001 0.203 0.081 0.900 0.900 1.000 0.900 0.900 0.900 0.900 0.629 0.004 0.055 0.900
CatBoost 0.001 0.149 0.055 0.900 0.900 0.900 1.000 0.900 0.900 0.900 0.710 0.007 0.081 0.900
SelfPacedEnsemble 0.001 0.188 0.074 0.900 0.900 0.900 0.900 1.000 0.900 0.900 0.649 0.004 0.061 0.900
BalancedRandomForest 0.002 0.547 0.308 0.900 0.900 0.900 0.900 0.900 1.000 0.900 0.270 0.001 0.008 0.900
EasyEnsemble 0.001 0.485 0.252 0.900 0.900 0.900 0.900 0.900 0.900 1.000 0.328 0.001 0.012 0.900
BalanceCascade 0.001 0.001 0.001 0.874 0.308 0.629 0.710 0.649 0.270 0.328 1.000 0.812 0.900 0.106
RUSBoost 0.001 0.001 0.001 0.017 0.001 0.004 0.007 0.004 0.001 0.001 0.812 1.000 0.900 0.001
SMOTEBoost 0.001 0.001 0.001 0.161 0.011 0.055 0.081 0.061 0.008 0.012 0.900 0.900 1.000 0.002
OpenFE 0.008 0.792 0.567 0.900 0.900 0.900 0.900 0.900 0.900 0.900 0.106 0.001 0.002 1.000
MDGP-Forest DeepForest imbalance-DF RandomForest XGBoost lightGBM CatBoost SelfPacedEnsemble BalancedRandomForest EasyEnsemble BalanceCascade RUSBoost SMOTEBoost OpenFE
1.0 0.8 0.6 0.4 0.2 0.0


The p-values for the MCC metric
MDGP-Forest 1.000 0.900 0.002 0.006 0.001 0.005 0.017 0.001 0.001 0.001 0.001 0.001 0.001 0.067
DeepForest 0.900 1.000 0.161 0.289 0.097 0.270 0.485 0.001 0.001 0.001 0.001 0.001 0.001 0.751
imbalance-DF 0.002 0.161 1.000 0.900 0.900 0.900 0.900 0.030 0.188 0.097 0.001 0.001 0.001 0.900
RandomForest 0.006 0.289 0.900 1.000 0.900 0.900 0.900 0.012 0.097 0.045 0.001 0.001 0.001 0.900
XGBoost 0.001 0.097 0.900 0.900 1.000 0.900 0.900 0.055 0.289 0.161 0.001 0.001 0.001 0.900
lightGBM 0.005 0.270 0.900 0.900 0.900 1.000 0.900 0.014 0.106 0.050 0.001 0.001 0.001 0.900
CatBoost 0.017 0.485 0.900 0.900 0.900 0.900 1.000 0.004 0.041 0.017 0.001 0.001 0.001 0.900
SelfPacedEnsemble 0.001 0.001 0.030 0.012 0.055 0.014 0.004 1.000 0.900 0.900 0.900 0.900 0.900 0.001
BalancedRandomForest 0.001 0.001 0.188 0.097 0.289 0.106 0.041 0.900 1.000 0.900 0.567 0.731 0.629 0.009
EasyEnsemble 0.001 0.001 0.097 0.045 0.161 0.050 0.017 0.900 0.900 1.000 0.731 0.894 0.792 0.003
BalanceCascade 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.900 0.567 0.731 1.000 0.900 0.900 0.001
RUSBoost 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.900 0.731 0.894 0.900 1.000 0.900 0.001
SMOTEBoost 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.900 0.629 0.792 0.900 0.900 1.000 0.001
OpenFE 0.067 0.751 0.900 0.900 0.900 0.900 0.900 0.001 0.009 0.003 0.001 0.001 0.001 1.000
MDGP-Forest DeepForest imbalance-DF RandomForest XGBoost lightGBM CatBoost SelfPacedEnsemble BalancedRandomForest EasyEnsemble BalanceCascade RUSBoost SMOTEBoost OpenFE
1.0
0.8
0.6
0.4
0.2
0.0


The p-values for the Accuracy metric
MDGP-Forest 1.000 0.900 0.001 0.011 0.002 0.005 0.014 0.001 0.001 0.001 0.001 0.001 0.001 0.097
DeepForest 0.900 1.000 0.031 0.308 0.111 0.203 0.350 0.001 0.001 0.001 0.001 0.001 0.001 0.751
imbalance-DF 0.001 0.031 1.000 0.900 0.900 0.900 0.900 0.025 0.155 0.243 0.001 0.005 0.001 0.900
RandomForest 0.011 0.308 0.900 1.000 0.900 0.900 0.900 0.001 0.011 0.021 0.001 0.001 0.001 0.900
XGBoost 0.002 0.111 0.900 0.900 1.000 0.900 0.900 0.005 0.048 0.085 0.001 0.001 0.001 0.900
lightGBM 0.005 0.203 0.900 0.900 0.900 1.000 0.900 0.002 0.021 0.041 0.001 0.001 0.001 0.900
CatBoost 0.014 0.350 0.900 0.900 0.900 0.900 1.000 0.001 0.008 0.017 0.001 0.001 0.001 0.900
SelfPacedEnsemble 0.001 0.001 0.025 0.001 0.005 0.002 0.001 1.000 0.900 0.900 0.900 0.900 0.900 0.001
BalancedRandomForest 0.001 0.001 0.155 0.011 0.048 0.021 0.008 0.900 1.000 0.900 0.731 0.900 0.900 0.001
EasyEnsemble 0.001 0.001 0.243 0.021 0.085 0.041 0.017 0.900 0.900 1.000 0.608 0.900 0.900 0.002
BalanceCascade 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.900 0.731 0.608 1.000 0.900 0.900 0.001
RUSBoost 0.001 0.001 0.005 0.001 0.001 0.001 0.001 0.900 0.900 0.900 0.900 1.000 0.900 0.001
SMOTEBoost 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.900 0.900 0.900 0.900 0.900 1.000 0.001
OpenFE 0.097 0.751 0.900 0.900 0.900 0.900 0.900 0.001 0.001 0.002 0.001 0.001 0.001 1.000
MDGP-Forest DeepForest imbalance-DF RandomForest XGBoost lightGBM CatBoost SelfPacedEnsemble BalancedRandomForest EasyEnsemble BalanceCascade RUSBoost SMOTEBoost OpenFE
1.0
0.8
0.6
0.4
0.2
0.0


The p-values for the Balanced Accuracy metric
MDGP-Forest
DeepForest
imbalance-DF
RandomForest
XGBoost
lightGBM
CatBoost
SelfPacedEnsemble
BalancedRandomForest
EasyEnsemble
BalanceCascade
RUSBoost
SMOTEBoost
OpenFE
1.0
0.8
0.6
0.4
0.2
0.0


The p-values for the Precision metric
MDGP-Forest
DeepForest
imbalance-DF
RandomForest
XGBoost
lightGBM
CatBoost
SelfPacedEnsemble
BalancedRandomForest
EasyEnsemble
BalanceCascade
RUSBoost
SMOTEBoost
OpenFE
1.0
0.8
0.6
0.4
0.2
0.0

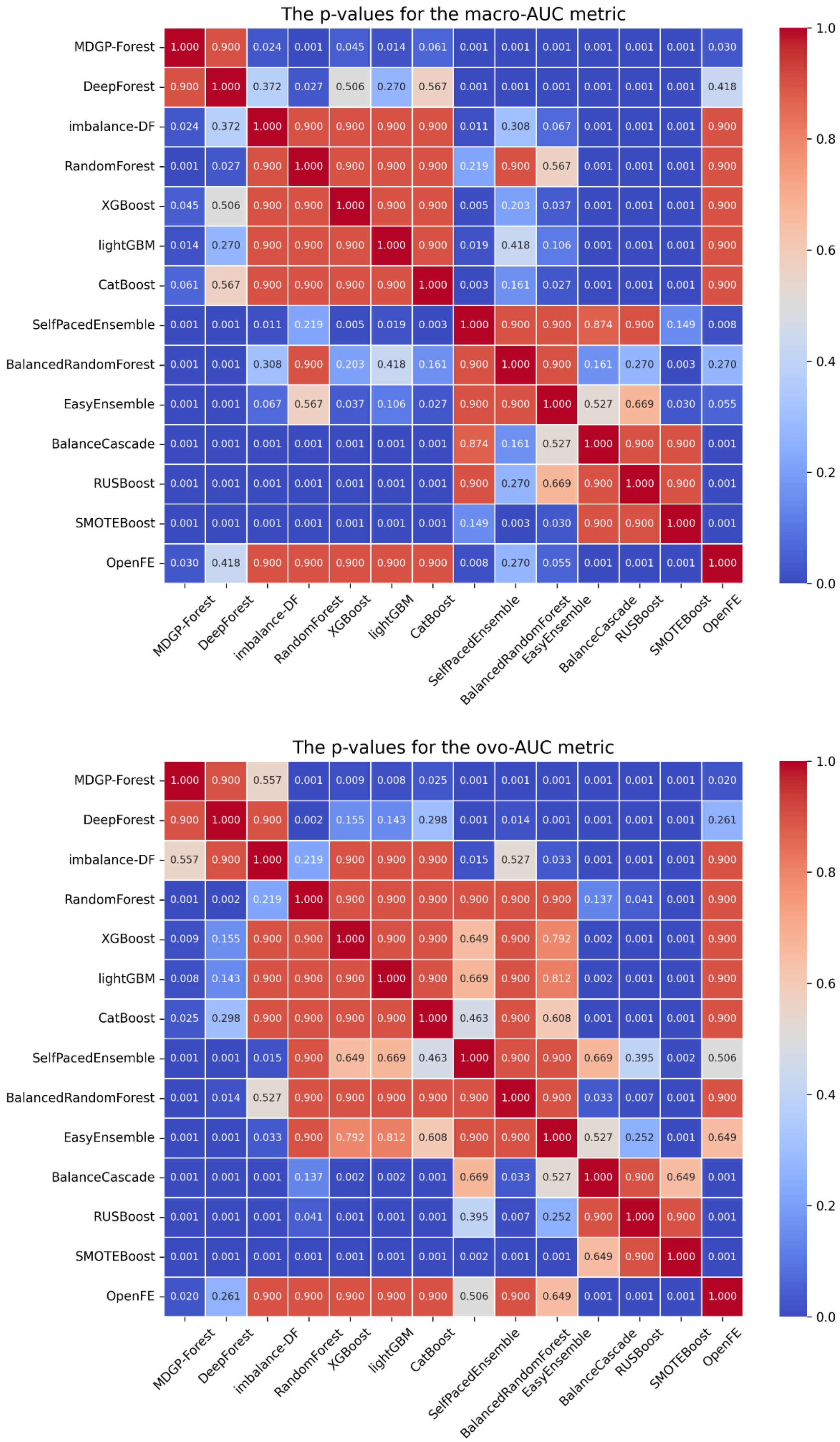


Fig S2 Significance heatmaps for all metrics.

3. Detailed data of ablation study

Due to space limitations in the main text, only the result figure of the Nonparametric Statistical Test from the ablation study was presented. In this section, we showcase the details of this part. The results of the ablation study were presented in Table S22, Table S23, and Table S24.

Table S22 The F1 results of the ablation study

Dataset	MDGP-Forest	No_hardness_resampling	Disassembly_with_RUS	No_multiclass_disassembly	No_original_features
D1	<u>0.9573</u> ±0.0128	**0.9578** ±0.0123	0.9491 ±0.0125	0.9547 ±0.0118	0.9496 ±0.0136
D2	0.8964 ±0.0242	**0.8987** ±0.0279	0.8923 ±0.0239	0.8927 ±0.026	<u>0.8983</u> ±0.0271
D3	**0.9104** ±0.0091	<u>0.9102</u> ±0.0083	0.8984 ±0.0078	0.9091 ±0.0085	0.9005 ±0.0074
D4	<u>0.8128</u> ±0.0707	0.8097 ±0.0656	**0.8154** ±0.0549	0.8013 ±0.0704	0.8081 ±0.0572
D5	**0.7033** ±0.0098	<u>0.7029</u> ±0.009	0.6242 ±0.0112	0.7003 ±0.008	0.6454 ±0.0102
D6	<u>0.9654</u> ±0.0265	**0.967** ±0.0254	0.9613 ±0.0257	0.9566 ±0.0326	0.9634 ±0.0262
D7	0.9392 ±0.0282	0.9416 ±0.029	<u>0.9475</u> ±0.0366	0.8742 ±0.0311	**0.9484** ±0.0314
D8	**0.6823** ±0.1001	<u>0.6808</u> ±0.1073	0.6701 ±0.1018	0.6769 ±0.0827	0.6678 ±0.0991
D9	<u>0.9505</u> ±0.074	0.9297 ±0.0827	**0.9677** ±0.0589	0.9378 ±0.0798	0.9443 ±0.0746
D10	**0.7636** ±0.0552	<u>0.7554</u> ±0.059	0.7004 ±0.0676	0.7524 ±0.0349	0.7025 ±0.0859
D11	0.9812 ±0.0072	0.9818 ±0.0058	<u>0.9825</u> ±0.0058	0.9789 ±0.0094	**0.9832** ±0.0073
D12	0.8149 ±0.0355	<u>0.8156</u> ±0.0366	0.812 ±0.032	0.8092 ±0.0343	**0.8219** ±0.0277
D13	0.341 ±0.0778	<u>0.3573</u> ±0.0869	0.3446 ±0.0703	0.3442 ±0.0909	**0.3596** ±0.0867
D14	**0.8631** ±0.0472	0.8483 ±0.0618	0.8374 ±0.073	<u>0.8593</u> ±0.051	0.838 ±0.073
D15	0.9972 ±0.0022	0.9969 ±0.0022	<u>0.9973</u> ±0.0025	0.9949 ±0.0037	**0.9988** ±0.0015
D16	0.3457 ±0.0272	0.3401 ±0.0227	**0.3536** ±0.03	0.3454 ±0.0268	<u>0.3521</u> ±0.0361
D17	<u>0.8826</u> ±0.0129	**0.8832** ±0.0135	0.8578 ±0.0132	0.8813 ±0.0129	0.859 ±0.0163
D18	**0.5835**	0.5804	0.5764	<u>0.5822</u>	0.5818

	±0.0614	±0.0599	±0.0666	±0.0585	±0.0648
D19	0.8834	**0.8849**	0.8804	0.8785	0.8833
	±0.0368	±0.0378	±0.0399	±0.0402	±0.0389
D20	0.4221	0.4234	0.4152	**0.4249**	0.4207
	±0.0159	±0.0147	±0.0208	±0.012	±0.0218
D21	**0.9551**	0.9538	0.9538	0.9475	0.9531
	±0.0211	±0.0203	±0.0243	±0.0289	±0.0225
D22	**0.9846**	0.9843	0.9626	0.9835	0.9629
	±0.0118	±0.0112	±0.0169	±0.0122	±0.016
D23	0.9924	0.9923	0.9945	0.9889	**0.9951**
	±0.0023	±0.0029	±0.0022	±0.003	±0.0024
D24	**0.9781**	0.9778	0.973	0.9777	0.974
	±0.0032	±0.0028	±0.0047	±0.0028	±0.0037
D25	0.9739	**0.974**	0.9723	0.9724	0.9736
	±0.007	±0.0075	±0.0078	±0.0069	±0.0068
D26	**0.9358**	0.9356	0.8898	0.9324	0.8988
	±0.0099	±0.0114	±0.0186	±0.0102	±0.0185
D27	0.8612	**0.8624**	0.8579	0.8614	0.8596
	±0.01	±0.0095	±0.0101	±0.0089	±0.012
D28	**0.8596**	0.8565	0.8559	0.8586	0.8575
	±0.0112	±0.0087	±0.0118	±0.0108	±0.0096
D29	**0.9867**	**0.9867**	0.9796	0.9866	0.9812
	±0.0038	±0.0036	±0.0047	±0.0035	±0.0045
D30	**0.95**	0.9493	0.9408	0.9492	0.9395
	±0.0105	±0.0117	±0.0137	±0.0104	±0.016
D31	**0.8103**	0.8049	0.791	0.7826	0.7918
	±0.0261	±0.0314	±0.0279	±0.0276	±0.0246
D32	0.5489	0.5485	0.5461	**0.5503**	0.545
	±0.009	±0.0137	±0.0138	±0.013	±0.0141
D33	**0.9925**	**0.9925**	0.9903	0.9911	0.9915
	±0.0018	±0.0019	±0.0022	±0.0024	±0.0019
D34	**0.9986**	0.9982	0.9957	0.9983	0.9958
	±0.0023	±0.0024	±0.0044	±0.0022	±0.0052
D35	**0.5227**	0.5215	0.5181	0.5064	0.5174
	±0.0214	±0.0276	±0.0253	±0.0292	±0.0268

Table S23 The G-mean results of the ablation study

Dataset	MDGP-Forest	No_hardness_resampling	Disassembly_with_RUS	No_multiclass_disassembly	No_original_features
D1	0.9693	**0.9696**	0.963	0.9674	0.9633
	±0.0099	±0.0094	±0.0108	±0.0095	±0.0114
D2	0.9277	0.9294	0.9244	0.9245	**0.9296**
	±0.0168	±0.0199	±0.0165	±0.0191	±0.0178
D3	**0.9441**	0.9438	0.937	0.9432	0.9383

	±0.0057	±0.0051	±0.0048	±0.0052	±0.0048
D4	0.8706 ±0.0502	0.8693 ±0.0473	**0.8734** ±0.0399	0.8624 ±0.0482	0.8677 ±0.0407
D5	**0.8134** ±0.0067	0.8133 ±0.0062	0.763 ±0.0081	0.8116 ±0.006	0.777 ±0.0068
D6	0.9712 ±0.0299	0.9717 ±0.0282	0.9734 ±0.023	0.9578 ±0.0357	**0.9742** ±0.0222
D7	0.9605 ±0.0203	0.9591 ±0.0201	0.9681 ±0.0232	0.9234 ±0.0273	**0.9689** ±0.02
D8	**0.7465** ±0.0694	0.7462 ±0.0738	0.7404 ±0.0646	0.7315 ±0.0553	0.7425 ±0.0614
D9	0.9769 ±0.0347	0.9665 ±0.0396	**0.9842** ±0.0274	0.9698 ±0.0384	0.9731 ±0.0349
D10	**0.8482** ±0.0397	0.844 ±0.0414	0.8098 ±0.0449	0.8438 ±0.0274	0.8091 ±0.0552
D11	0.9833 ±0.0066	0.9837 ±0.0052	0.9833 ±0.0052	0.9814 ±0.0081	**0.9848** ±0.0055
D12	0.879 ±0.0256	0.8802 ±0.0255	0.8769 ±0.0231	0.8754 ±0.0244	**0.8823** ±0.0207
D13	0.5459 ±0.0572	0.5599 ±0.0634	0.5508 ±0.055	0.5464 ±0.0646	**0.5602** ±0.0636
D14	**0.9134** ±0.0231	0.9056 ±0.0311	0.9034 ±0.0365	0.9121 ±0.0235	0.9023 ±0.0371
D15	0.9978 ±0.002	0.9976 ±0.002	0.998 ±0.0019	0.9961 ±0.0032	**0.999** ±0.0013
D16	0.5587 ±0.0247	0.5557 ±0.0218	**0.5619** ±0.0224	0.5595 ±0.0245	0.5615 ±0.0275
D17	0.9164 ±0.0081	**0.9167** ±0.0087	0.897 ±0.0088	0.9153 ±0.0082	0.8986 ±0.0121
D18	0.7369 ±0.0412	0.7351 ±0.0413	0.7321 ±0.0448	**0.7371** ±0.0397	0.7339 ±0.045
D19	0.924 ±0.0214	**0.9249** ±0.0232	0.9229 ±0.0234	0.9165 ±0.0259	0.924 ±0.023
D20	0.596 ±0.0116	0.5972 ±0.0109	0.5902 ±0.0151	**0.5977** ±0.0091	0.5948 ±0.0161
D21	**0.9664** ±0.0157	0.9654 ±0.0151	0.9654 ±0.0182	0.9609 ±0.0211	0.9649 ±0.0168
D22	**0.9915** ±0.0064	0.9914 ±0.0062	0.9793 ±0.0094	0.991 ±0.0066	0.9796 ±0.0088
D23	0.9958 ±0.0013	0.9958 ±0.0016	0.997 ±0.0012	0.9938 ±0.0017	**0.9973** ±0.0013
D24	**0.9878** ±0.0018	0.9876 ±0.0016	0.985 ±0.0026	0.9876 ±0.0016	0.9855 ±0.002

D25	**0.9855** ±0.0039	**0.9855** ±0.0042	0.9846 ±0.0044	0.9846 ±0.0039	<u>0.9853</u> ±0.0038
D26	**0.9635** ±0.0057	<u>0.9634</u> ±0.0064	0.9362 ±0.0113	0.9616 ±0.0057	0.9418 ±0.0109
D27	0.8959 ±0.0073	**0.8968** ±0.007	0.8933 ±0.0075	<u>0.896</u> ±0.0065	0.8945 ±0.009
D28	**0.8945** ±0.0084	0.8922 ±0.0066	0.8916 ±0.0089	<u>0.8939</u> ±0.0081	0.8928 ±0.0073
D29	**0.9926** ±0.0021	**0.9926** ±0.002	0.9886 ±0.0026	<u>0.9925</u> ±0.002	0.9895 ±0.0025
D30	**0.9719** ±0.0059	<u>0.9716</u> ±0.0066	0.9668 ±0.0077	0.9715 ±0.0059	0.966 ±0.009
D31	**0.8732** ±0.0181	<u>0.8693</u> ±0.0215	0.8603 ±0.0183	0.8545 ±0.0193	0.8606 ±0.0163
D32	<u>0.6549</u> ±0.0081	0.6546 ±0.0113	0.6524 ±0.0115	**0.6566** ±0.0107	0.6516 ±0.0124
D33	**0.9955** ±0.0011	**0.9955** ±0.0011	0.9942 ±0.0014	0.9946 ±0.0014	<u>0.9948</u> ±0.0012
D34	**0.9992** ±0.0013	0.9989 ±0.0013	0.9976 ±0.0024	<u>0.999</u> ±0.0012	0.9976 ±0.003
D35	**0.6313** ±0.0157	<u>0.6303</u> ±0.0206	0.6282 ±0.0202	0.617 ±0.0229	0.6275 ±0.0213

Table S24 The MCC results of the ablation study

Dataset	MDGP-Forest	No_hardness_resampling	Disassembly_with_RUS	No_multiclass_disassembly	No_original_features
D1	<u>0.9388</u> ±0.0177	**0.9393** ±0.0169	0.9265 ±0.0184	0.9352 ±0.0165	0.9273 ±0.0196
D2	0.8631 ±0.0334	<u>0.8667</u> ±0.0381	0.8598 ±0.033	0.8581 ±0.0375	**0.8672** ±0.0354
D3	**0.9099** ±0.0095	<u>0.9095</u> ±0.0088	0.8978 ±0.008	0.9088 ±0.009	0.8995 ±0.008
D4	<u>0.7691</u> ±0.085	0.7651 ±0.0801	**0.7703** ±0.0686	0.7539 ±0.0806	0.7627 ±0.0721
D5	<u>0.6643</u> ±0.0131	**0.6652** ±0.012	0.5864 ±0.0139	0.6621 ±0.0117	0.6057 ±0.0123
D6	<u>0.9472</u> ±0.0388	**0.95** ±0.0372	0.9378 ±0.0412	0.9354 ±0.047	0.9406 ±0.0426
D7	0.945 ±0.0229	0.9456 ±0.0246	**0.9555** ±0.0303	0.892 ±0.0271	<u>0.9532</u> ±0.0275
D8	0.5765 ±0.0907	0.5738 ±0.0971	<u>0.5775</u> ±0.0861	0.543 ±0.0712	**0.582** ±0.0747
D9	<u>0.9741</u> ±0.0389	0.9687 ±0.0354	**0.9815** ±0.0299	0.9703 ±0.0372	0.9725 ±0.0327

D10	**0.7411** ±0.0442	<u>0.7367</u> ±0.041	0.6652 ±0.0586	0.7365 ±0.0344	0.6621 ±0.0613
D11	0.9709 ±0.0107	**0.9713** ±0.009	0.9699 ±0.0095	0.9681 ±0.0138	<u>0.9712</u> ±0.0135
D12	0.7388 ±0.0342	<u>0.7403</u> ±0.035	0.7383 ±0.032	0.73 ±0.0277	**0.7511** ±0.0324
D13	0.3483 ±0.0699	**0.3677** ±0.0789	0.3448 ±0.0768	<u>0.3607</u> ±0.0828	0.3503 ±0.0746
D14	<u>0.8408</u> ±0.0329	0.8377 ±0.0278	0.8369 ±0.0311	**0.8423** ±0.0325	0.8306 ±0.0342
D15	0.9972 ±0.002	0.997 ±0.0021	<u>0.9973</u> ±0.0021	0.9951 ±0.0031	**0.9986** ±0.0014
D16	<u>0.501</u> ±0.0613	0.5006 ±0.0508	0.4928 ±0.0624	**0.5067** ±0.0617	0.493 ±0.0629
D17	**0.9151** ±0.0052	<u>0.9146</u> ±0.0051	0.8905 ±0.01	0.9133 ±0.0053	0.8918 ±0.0086
D18	0.4998 ±0.0318	**0.5017** ±0.0305	0.4973 ±0.0362	<u>0.5015</u> ±0.0304	0.4988 ±0.0305
D19	<u>0.8772</u> ±0.0239	**0.8787** ±0.0262	0.8725 ±0.029	0.8737 ±0.0277	0.8769 ±0.0285
D20	0.5146 ±0.02	<u>0.5149</u> ±0.0206	0.5104 ±0.0289	**0.521** ±0.0206	0.5128 ±0.0245
D21	**0.9351** ±0.0301	<u>0.9332</u> ±0.0289	0.933 ±0.0357	0.9254 ±0.0378	0.9321 ±0.0329
D22	**0.9833** ±0.0126	<u>0.9829</u> ±0.0121	0.9594 ±0.0184	0.9822 ±0.013	0.9598 ±0.0171
D23	0.9916 ±0.0025	0.9915 ±0.0032	<u>0.9939</u> ±0.0024	0.9877 ±0.0034	**0.9946** ±0.0026
D24	**0.9758** ±0.0035	<u>0.9754</u> ±0.0031	0.9702 ±0.0051	0.9753 ±0.0031	0.9713 ±0.004
D25	<u>0.9712</u> ±0.0077	**0.9713** ±0.0082	0.9694 ±0.0087	0.9695 ±0.0076	0.9708 ±0.0075
D26	**0.9283** ±0.011	<u>0.9281</u> ±0.0124	0.8759 ±0.0214	0.9246 ±0.0108	0.8866 ±0.0208
D27	0.7943 ±0.0142	**0.7961** ±0.0135	0.7889 ±0.0147	<u>0.7946</u> ±0.0126	0.7911 ±0.0176
D28	**0.7919** ±0.0161	0.7872 ±0.0127	0.7858 ±0.0174	<u>0.7909</u> ±0.0154	0.7881 ±0.0143
D29	**0.9852** ±0.0042	**0.9852** ±0.004	0.9773 ±0.0052	<u>0.9851</u> ±0.0039	0.9791 ±0.005
D30	**0.9447** ±0.0116	<u>0.944</u> ±0.0129	0.9346 ±0.015	0.9438 ±0.0115	0.9332 ±0.0177
D31	**0.7499**	<u>0.7424</u>	0.7253	0.7133	0.7259

	±0.0356	±0.0417	±0.0357	±0.0377	±0.0315
D32	0.3289 ±0.0154	0.3284 ±0.0213	0.3243 ±0.022	**0.3317** ±0.0202	0.3225 ±0.024
D33	0.9907 ±0.0023	**0.9908** ±0.0023	0.9882 ±0.0028	0.9889 ±0.003	0.9895 ±0.0023
D34	**0.9983** ±0.0027	0.9978 ±0.0028	0.995 ±0.0049	0.998 ±0.0026	0.995 ±0.0062
D35	**0.2921** ±0.0287	0.2914 ±0.0368	0.2893 ±0.0375	0.2669 ±0.0431	0.2876 ±0.04

4. Supplementary data of parameter studies

4.1 Detailed data for parameter studies in the main text

This section provides further details for the analysis of the impact of different hyperparameters on the performance of MDGP-Forest.

The MDGP-Forest is configured with four different θ values $\{0\%, 5\%, 10\%, 15\%\}$, and the experimental results are presented in Table S25, Table S26, and Table S27.

Table S25 F1 results of the different θ parameter values

Dataset	θ=0%	θ=5%	θ=10%	θ=15%
D1	0.9578 ±0.0123	0.9573 ±0.0128	**0.9585** ±0.0119	0.957 ±0.012
D2	**0.8987** ±0.0279	0.8964 ±0.0242	0.8943 ±0.0261	0.8966 ±0.0244
D3	0.9102 ±0.0083	**0.9104** ±0.0091	0.9093 ±0.0078	0.9094 ±0.0084
D4	0.8097 ±0.0656	**0.8128** ±0.0707	0.8094 ±0.0624	0.7999 ±0.0618
D5	0.7029 ±0.009	**0.7033** ±0.0098	0.7026 ±0.0087	0.703 ±0.009
D6	**0.967** ±0.0254	0.9654 ±0.0265	0.965 ±0.0335	0.9638 ±0.0292
D7	0.9416 ±0.029	0.9392 ±0.0282	**0.9434** ±0.0333	0.9344 ±0.0367
D8	0.6808 ±0.1073	0.6823 ±0.1001	0.6809 ±0.1047	**0.6878** ±0.1002
D9	0.9297 ±0.0827	**0.9505** ±0.074	0.9419 ±0.0781	0.9335 ±0.082
D10	0.7554 ±0.059	**0.7636** ±0.0552	0.7549 ±0.0478	0.7578 ±0.0607
D11	**0.9818** ±0.0058	0.9812 ±0.0072	0.9799 ±0.0081	0.9801 ±0.0073

D12	**0.8156** ±0.0366	0.8149 ±0.0355	<u>0.8155</u> ±0.0359	0.8147 ±0.0346
D13	**0.3573** ±0.0869	0.341 ±0.0778	0.3462 ±0.0846	<u>0.3572</u> ±0.0975
D14	0.8483 ±0.0618	**0.8631** ±0.0472	<u>0.8538</u> ±0.0666	0.8413 ±0.0685
D15	0.9969 ±0.0022	**0.9972** ±0.0022	0.9969 ±0.0024	<u>0.997</u> ±0.0027
D16	0.3401 ±0.0227	**0.3457** ±0.0272	0.3384 ±0.0295	<u>0.3415</u> ±0.0256
D17	**0.8832** ±0.0135	0.8826 ±0.0129	<u>0.883</u> ±0.0133	0.8824 ±0.014
D18	0.5804 ±0.0599	**0.5835** ±0.0614	0.5779 ±0.0671	<u>0.5813</u> ±0.0616
D19	**0.8849** ±0.0378	<u>0.8834</u> ±0.0368	0.8786 ±0.0403	0.8809 ±0.0358
D20	0.4234 ±0.0147	0.4221 ±0.0159	**0.4274** ±0.0157	<u>0.4239</u> ±0.0154
D21	0.9538 ±0.0203	<u>0.9551</u> ±0.0211	**0.9572** ±0.0215	0.9518 ±0.0195
D22	<u>0.9843</u> ±0.0112	**0.9846** ±0.0118	0.9837 ±0.0127	0.9835 ±0.013
D23	<u>0.9923</u> ±0.0029	**0.9924** ±0.0023	**0.9924** ±0.0022	0.992 ±0.0025
D24	0.9778 ±0.0028	<u>0.9781</u> ±0.0032	**0.9783** ±0.0029	0.9779 ±0.0027
D25	**0.974** ±0.0075	<u>0.9739</u> ±0.007	0.9736 ±0.007	<u>0.9739</u> ±0.0073
D26	<u>0.9356</u> ±0.0114	**0.9358** ±0.0099	0.9331 ±0.0093	0.9337 ±0.0116
D27	<u>0.8624</u> ±0.0095	0.8612 ±0.01	**0.8625** ±0.0105	0.8615 ±0.01
D28	0.8565 ±0.0087	**0.8596** ±0.0112	<u>0.859</u> ±0.0101	0.8583 ±0.0101
D29	**0.9867** ±0.0036	**0.9867** ±0.0038	<u>0.9866</u> ±0.0034	0.9865 ±0.0036
D30	<u>0.9493</u> ±0.0117	**0.95** ±0.0105	0.9488 ±0.0123	0.9485 ±0.0111
D31	0.8049 ±0.0314	<u>0.8103</u> ±0.0261	0.802 ±0.0295	**0.8121** ±0.0285
D32	0.5485 ±0.0137	<u>0.5489</u> ±0.009	0.5487 ±0.0113	**0.5493** ±0.0124
D33	**0.9925**	**0.9925**	<u>0.9924</u>	**0.9925**

	±0.0019	±0.0018	±0.002	±0.002
D34	0.9982	**0.9986**	**0.9986**	0.9981
	±0.0024	±0.0023	±0.0023	±0.0024
D35	0.5215	**0.5227**	0.5162	0.5208
	±0.0276	±0.0214	±0.025	±0.0285

Table S26 G-mean results of the different θ parameter values

Dataset	θ=0%	θ=5%	θ=10%	θ=15%
D1	0.9696	0.9693	**0.9701**	0.969
	±0.0094	±0.0099	±0.0092	±0.0091
D2	**0.9294**	0.9277	0.9261	0.9279
	±0.0199	±0.0168	±0.0182	±0.0162
D3	0.9438	**0.9441**	0.9436	0.9434
	±0.0051	±0.0057	±0.005	±0.0051
D4	0.8693	**0.8706**	0.8688	0.8624
	±0.0473	±0.0502	±0.0446	±0.0437
D5	0.8133	**0.8134**	0.813	0.8131
	±0.0062	±0.0067	±0.0062	±0.0063
D6	**0.9717**	0.9712	0.969	0.9701
	±0.0282	±0.0299	±0.0329	±0.0286
D7	0.9591	0.9605	**0.9607**	0.9559
	±0.0201	±0.0203	±0.022	±0.0269
D8	0.7462	0.7465	0.7461	**0.7494**
	±0.0738	±0.0694	±0.0714	±0.0696
D9	0.9665	**0.9769**	0.972	0.9683
	±0.0396	±0.0347	±0.0374	±0.0389
D10	0.844	**0.8482**	0.8443	0.8434
	±0.0414	±0.0397	±0.0346	±0.0422
D11	**0.9837**	0.9833	0.9823	0.9828
	±0.0052	±0.0066	±0.0075	±0.0064
D12	**0.8802**	0.879	**0.8802**	0.8794
	±0.0255	±0.0256	±0.0256	±0.0255
D13	**0.5599**	0.5459	0.5508	0.5566
	±0.0634	±0.0572	±0.0628	±0.073
D14	0.9056	**0.9134**	0.9093	0.9028
	±0.0311	±0.0231	±0.0337	±0.0336
D15	0.9976	**0.9978**	0.9976	0.9975
	±0.002	±0.002	±0.0021	±0.0025
D16	0.5557	**0.5587**	0.5544	0.5559
	±0.0218	±0.0247	±0.027	±0.0241
D17	**0.9167**	0.9164	0.9165	0.9161
	±0.0087	±0.0081	±0.0087	±0.0088
D18	0.7351	**0.7369**	0.7331	0.7355

	±0.0413	±0.0412	±0.0458	±0.0427
D19	**0.9249** ±0.0232	0.924 ±0.0214	0.9206 ±0.0264	0.923 ±0.0209
D20	0.5972 ±0.0109	0.596 ±0.0116	**0.5995** ±0.0116	0.5972 ±0.0116
D21	0.9654 ±0.0151	0.9664 ±0.0157	**0.9679** ±0.0161	0.9639 ±0.0145
D22	0.9914 ±0.0062	**0.9915** ±0.0064	0.9911 ±0.0069	0.991 ±0.007
D23	**0.9958** ±0.0016	**0.9958** ±0.0013	**0.9958** ±0.0012	0.9956 ±0.0014
D24	0.9876 ±0.0016	0.9878 ±0.0018	**0.9879** ±0.0016	0.9877 ±0.0015
D25	**0.9855** ±0.0042	**0.9855** ±0.0039	0.9853 ±0.0039	0.9854 ±0.0041
D26	0.9634 ±0.0064	**0.9635** ±0.0057	0.9619 ±0.0054	0.9623 ±0.0065
D27	**0.8968** ±0.007	0.8959 ±0.0073	**0.8968** ±0.0077	0.8961 ±0.0073
D28	0.8922 ±0.0066	**0.8945** ±0.0084	0.8941 ±0.0076	0.8935 ±0.0076
D29	**0.9926** ±0.002	**0.9926** ±0.0021	0.9925 ±0.0019	0.9925 ±0.0021
D30	0.9716 ±0.0066	**0.9719** ±0.0059	0.9713 ±0.0069	0.9711 ±0.0062
D31	0.8693 ±0.0215	0.8732 ±0.0181	0.8673 ±0.0202	**0.8742** ±0.0194
D32	0.6546 ±0.0113	0.6549 ±0.0081	0.655 ±0.0101	**0.6553** ±0.0112
D33	**0.9955** ±0.0011	**0.9955** ±0.0011	0.9954 ±0.0012	**0.9955** ±0.0012
D34	0.9989 ±0.0013	**0.9992** ±0.0013	**0.9992** ±0.0013	0.9989 ±0.0013
D35	0.6303 ±0.0206	**0.6313** ±0.0157	0.6266 ±0.0189	0.63 ±0.0217

Table S27 MCC results of the different θ parameter values

Dataset	θ=0%	θ=5%	θ=10%	θ=15%
D1	0.9393 ±0.0169	0.9388 ±0.0177	**0.9403** ±0.0165	0.938 ±0.0165
D2	**0.8667** ±0.0381	0.8631 ±0.0334	0.8604 ±0.0354	0.8635 ±0.032
D3	0.9095	**0.9099**	0.9085	0.9089

	±0.0088	±0.0095	±0.0081	±0.0086
D4	0.7651 ±0.0801	**0.7691** ±0.085	0.7641 ±0.0752	0.7529 ±0.0733
D5	**0.6652** ±0.012	0.6643 ±0.0131	0.6642 ±0.0122	0.6644 ±0.0123
D6	**0.95** ±0.0372	0.9472 ±0.0388	0.9475 ±0.0491	0.9449 ±0.0427
D7	0.9456 ±0.0246	0.945 ±0.0229	**0.9476** ±0.0286	0.94 ±0.0285
D8	0.5738 ±0.0971	0.5765 ±0.0907	0.5756 ±0.0967	**0.5768** ±0.0943
D9	0.9687 ±0.0354	**0.9741** ±0.0389	0.9725 ±0.0351	0.97 ±0.0353
D10	0.7367 ±0.041	**0.7411** ±0.0442	0.7329 ±0.0347	0.7337 ±0.0479
D11	**0.9713** ±0.009	0.9709 ±0.0107	0.9694 ±0.0126	0.9691 ±0.0109
D12	0.7403 ±0.035	0.7388 ±0.0342	**0.7404** ±0.0347	0.7399 ±0.0339
D13	**0.3677** ±0.0789	0.3483 ±0.0699	0.3609 ±0.0833	0.367 ±0.0863
D14	0.8377 ±0.0278	0.8408 ±0.0329	**0.8414** ±0.0352	0.8386 ±0.0321
D15	0.997 ±0.0021	**0.9972** ±0.002	0.9971 ±0.0019	**0.9972** ±0.0022
D16	0.5006 ±0.0508	**0.501** ±0.0613	0.4981 ±0.0634	0.4993 ±0.0586
D17	0.9146 ±0.0051	**0.9151** ±0.0052	0.915 ±0.0056	0.9145 ±0.0058
D18	0.5017 ±0.0305	0.4998 ±0.0318	**0.5018** ±0.0316	0.5001 ±0.0331
D19	**0.8787** ±0.0262	0.8772 ±0.0239	0.8764 ±0.0273	0.8768 ±0.0235
D20	0.5149 ±0.0206	0.5146 ±0.02	**0.5193** ±0.0192	0.515 ±0.0222
D21	0.9332 ±0.0289	0.9351 ±0.0301	**0.9379** ±0.0312	0.9305 ±0.0276
D22	0.9829 ±0.0121	**0.9833** ±0.0126	0.9824 ±0.0135	0.9822 ±0.0138
D23	0.9915 ±0.0032	**0.9916** ±0.0025	**0.9916** ±0.0025	0.9911 ±0.0027
D24	0.9754 ±0.0031	0.9758 ±0.0035	**0.976** ±0.0032	0.9755 ±0.003

D25	**0.9713** ±0.0082	0.9712 ±0.0077	0.9708 ±0.0077	0.9711 ±0.0081
D26	0.9281 ±0.0124	**0.9283** ±0.011	0.9253 ±0.0103	0.926 ±0.0124
D27	**0.7961** ±0.0135	0.7943 ±0.0142	0.7959 ±0.0148	0.7947 ±0.0141
D28	0.7872 ±0.0127	**0.7919** ±0.0161	0.7911 ±0.0148	0.79 ±0.0147
D29	**0.9852** ±0.004	**0.9852** ±0.0042	0.9851 ±0.0038	0.985 ±0.0041
D30	0.944 ±0.0129	**0.9447** ±0.0116	0.9435 ±0.0135	0.9431 ±0.0122
D31	0.7424 ±0.0417	0.7499 ±0.0356	0.7384 ±0.0395	**0.7516** ±0.0382
D32	0.3284 ±0.0213	0.3289 ±0.0154	0.3293 ±0.0191	**0.3298** ±0.0213
D33	**0.9908** ±0.0023	0.9907 ±0.0023	0.9905 ±0.0026	0.9907 ±0.0026
D34	0.9978 ±0.0028	**0.9983** ±0.0027	**0.9983** ±0.0027	0.9977 ±0.0029
D35	0.2914 ±0.0368	**0.2921** ±0.0287	0.2855 ±0.0345	0.2913 ±0.0396

MDGP-Forest is configured with four different p values: {5,10,20,50}, and the experimental results are presented in Table S28, Table S29, and Table S30.

Table S28 F1 results of the different p parameter values

Dataset	p=5	p=10	p=20	p=50
D2	0.8962 ±0.022	0.8964 ±0.0242	**0.9** ±0.0288	0.8971 ±0.0293
D3	0.91 ±0.0079	**0.9104** ±0.0091	0.9092 ±0.0088	0.9084 ±0.0094
D4	0.8035 ±0.0562	**0.8128** ±0.0707	0.808 ±0.0704	0.8075 ±0.068
D6	0.963 ±0.0287	**0.9654** ±0.0265	0.9606 ±0.0306	0.9616 ±0.029
D7	0.9329 ±0.0309	**0.9392** ±0.0282	0.9376 ±0.0356	0.9321 ±0.0439
D8	0.6757 ±0.1022	0.6823 ±0.1001	0.6835 ±0.1029	**0.6843** ±0.1003
D9	0.9329 ±0.0826	0.9505 ±0.074	0.9471 ±0.0752	**0.9507** ±0.0718
D10	0.7635 ±0.0583	**0.7636** ±0.0552	0.7512 ±0.0484	0.7307 ±0.0724

D11	0.981 ±0.0065	<u>0.9812</u> ±0.0072	**0.9823** ±0.0059	0.9809 ±0.0071
D12	0.8127 ±0.0364	0.8149 ±0.0355	<u>0.8157</u> ±0.0349	**0.8174** ±0.0349
D13	0.3384 ±0.0893	0.341 ±0.0778	**0.3672** ±0.0962	<u>0.3634</u> ±0.0917
D14	0.8487 ±0.0587	**0.8631** ±0.0472	<u>0.856</u> ±0.0618	0.8548 ±0.0658
D15	0.9963 ±0.0027	<u>0.9972</u> ±0.0022	**0.9973** ±0.0022	**0.9973** ±0.0022
D16	0.3421 ±0.0264	<u>0.3457</u> ±0.0272	0.3405 ±0.0279	**0.3478** ±0.0238
D17	0.8825 ±0.0126	0.8826 ±0.0129	**0.8834** ±0.0128	<u>0.883</u> ±0.0133
D18	0.5774 ±0.0621	**0.5835** ±0.0614	0.5796 ±0.0623	<u>0.5797</u> ±0.0637
D19	**0.8843** ±0.0354	<u>0.8834</u> ±0.0368	0.8783 ±0.0413	0.8792 ±0.0354
D20	**0.4261** ±0.015	0.4221 ±0.0159	0.4223 ±0.0122	<u>0.4227</u> ±0.0109
D21	0.9501 ±0.0334	**0.9551** ±0.0211	<u>0.9531</u> ±0.018	<u>0.9531</u> ±0.0166
D22	<u>0.984</u> ±0.0128	**0.9846** ±0.0118	0.9839 ±0.0125	0.9833 ±0.0123
D23	0.9918 ±0.0023	0.9924 ±0.0023	<u>0.9931</u> ±0.0024	**0.9934** ±0.0025
D27	0.8602 ±0.0097	<u>0.8612</u> ±0.01	0.8608 ±0.0105	**0.8622** ±0.009
D28	<u>0.859</u> ±0.0104	**0.8596** ±0.0112	0.858 ±0.0104	0.8586 ±0.0101
D29	0.9865 ±0.0034	0.9867 ±0.0038	<u>0.987</u> ±0.0039	**0.9875** ±0.0038
D31	0.8039 ±0.0253	**0.8103** ±0.0261	<u>0.8048</u> ±0.0304	0.8008 ±0.0197
D32	0.5479 ±0.0119	<u>0.5489</u> ±0.009	**0.5498** ±0.0103	0.5488 ±0.0116
D34	0.9983 ±0.0022	**0.9986** ±0.0023	<u>0.9985</u> ±0.0023	0.9984 ±0.0025
D35	0.52 ±0.0275	<u>0.5227</u> ±0.0214	**0.5234** ±0.0191	0.5155 ±0.0166

Table S29 G-mean results of the different p parameter values

Dataset	*p*=5	*p*=10	*p*=20	*p*=50
D2	<u>0.9281</u> ±0.0149	0.9277 ±0.0168	**0.9303** ±0.0202	0.928 ±0.0199
D3	<u>0.9438</u> ±0.0049	**0.9441** ±0.0057	0.9433 ±0.0054	0.9427 ±0.0058
D4	0.8648 ±0.0399	**0.8706** ±0.0502	0.8672 ±0.0504	<u>0.8675</u> ±0.0489
D6	<u>0.9676</u> ±0.03	**0.9712** ±0.0299	0.967 ±0.0322	0.965 ±0.0316
D7	<u>0.9573</u> ±0.0203	**0.9605** ±0.0203	0.9567 ±0.0241	0.9507 ±0.0318
D8	0.7415 ±0.0703	0.7465 ±0.0694	**0.7499** ±0.0699	<u>0.7493</u> ±0.0671
D9	0.9683 ±0.0389	**0.9769** ±0.0347	0.9757 ±0.0353	<u>0.9759</u> ±0.0352
D10	<u>0.8481</u> ±0.0392	**0.8482** ±0.0397	0.8409 ±0.0326	0.8288 ±0.0478
D11	<u>0.9836</u> ±0.0058	0.9833 ±0.0066	**0.9841** ±0.0055	0.9827 ±0.0068
D12	0.8786 ±0.0271	0.879 ±0.0256	<u>0.8791</u> ±0.0245	**0.8793** ±0.0245
D13	0.5413 ±0.0629	0.5459 ±0.0572	**0.5626** ±0.0677	<u>0.5613</u> ±0.0656
D14	0.9074 ±0.0276	**0.9134** ±0.0231	<u>0.909</u> ±0.0316	0.908 ±0.0344
D15	0.9971 ±0.0023	<u>0.9978</u> ±0.002	0.9977 ±0.002	**0.9979** ±0.0019
D16	0.5569 ±0.0248	<u>0.5587</u> ±0.0247	0.5564 ±0.0266	**0.5614** ±0.0218
D17	0.916 ±0.0081	**0.9164** ±0.0081	**0.9164** ±0.0084	<u>0.9162</u> ±0.0082
D18	0.7331 ±0.0431	**0.7369** ±0.0412	0.7347 ±0.0428	<u>0.7348</u> ±0.0431
D19	<u>0.9235</u> ±0.022	**0.924** ±0.0214	0.9205 ±0.0252	0.9213 ±0.0213
D20	**0.5994** ±0.011	<u>0.596</u> ±0.0116	0.5959 ±0.0081	<u>0.596</u> ±0.0074
D21	0.9629 ±0.0245	**0.9664** ±0.0157	<u>0.9649</u> ±0.0134	<u>0.9649</u> ±0.0124
D22	<u>0.9912</u> ±0.0069	**0.9915** ±0.0064	<u>0.9912</u> ±0.0068	0.9908 ±0.0067
D23	0.9955	0.9958	<u>0.9962</u>	**0.9964**

	±0.0013	±0.0013	±0.0013	±0.0014
D27	0.8952 ±0.0072	0.8959 ±0.0073	0.8956 ±0.0077	**0.8966** ±0.0066
D28	0.8941 ±0.0078	**0.8945** ±0.0084	0.8934 ±0.0078	0.8938 ±0.0076
D29	0.9925 ±0.0019	0.9926 ±0.0021	0.9927 ±0.0022	**0.993** ±0.0022
D31	0.8688 ±0.0174	**0.8732** ±0.0181	0.8696 ±0.0206	0.8669 ±0.0135
D32	0.6546 ±0.01	0.6549 ±0.0081	**0.6557** ±0.0091	0.6547 ±0.0101
D34	0.9991 ±0.0012	**0.9992** ±0.0013	0.9991 ±0.0013	0.9991 ±0.0014
D35	0.6287 ±0.0204	0.6313 ±0.0157	**0.6315** ±0.015	0.6247 ±0.012

Table S30 MCC results of the different p parameter values

Dataset	p=5	p=10	p=20	p=50
D2	0.863 ±0.0308	0.8631 ±0.0334	**0.8677** ±0.0392	0.8644 ±0.0386
D3	0.9094 ±0.0084	**0.9099** ±0.0095	0.9085 ±0.0089	0.9079 ±0.0095
D4	0.7564 ±0.0682	**0.7691** ±0.085	0.7624 ±0.0844	0.7635 ±0.0814
D6	0.9439 ±0.0417	**0.9472** ±0.0388	0.9399 ±0.0447	0.9417 ±0.0418
D7	0.9399 ±0.025	**0.945** ±0.0229	0.9449 ±0.0277	0.9438 ±0.0307
D8	0.5611 ±0.0935	0.5765 ±0.0907	0.5855 ±0.0927	**0.5867** ±0.0905
D9	0.9701 ±0.0351	0.9741 ±0.0389	0.9753 ±0.0344	**0.9763** ±0.032
D10	**0.7463** ±0.0405	0.7411 ±0.0442	0.7335 ±0.0389	0.7189 ±0.0521
D11	0.9708 ±0.0096	0.9709 ±0.0107	**0.972** ±0.0086	0.9706 ±0.0108
D12	0.7361 ±0.0361	0.7388 ±0.0342	0.7405 ±0.0348	**0.7438** ±0.0347
D13	0.347 ±0.0751	0.3483 ±0.0699	0.3712 ±0.0828	**0.3716** ±0.078
D14	0.8327 ±0.0303	0.8408 ±0.0329	0.8414 ±0.0264	**0.8415** ±0.0314
D15	0.9964	0.9972	0.9975	**0.9979**

	±0.0025	±0.002	±0.0018	±0.0016
D16	0.5027 ±0.0594	0.501 ±0.0613	0.5009 ±0.0613	**0.506** ±0.0565
D17	0.9132 ±0.0051	0.9151 ±0.0052	0.916 ±0.0048	**0.9171** ±0.0055
D18	0.5003 ±0.0318	0.4998 ±0.0318	**0.5016** ±0.0331	0.4992 ±0.0361
D19	**0.8777** ±0.0251	0.8772 ±0.0239	0.8745 ±0.0254	0.8746 ±0.0216
D20	0.5169 ±0.0223	0.5146 ±0.02	0.5173 ±0.0203	**0.5186** ±0.0218
D21	0.9294 ±0.0447	**0.9351** ±0.0301	0.9322 ±0.0257	0.9322 ±0.0241
D22	0.9827 ±0.0136	**0.9833** ±0.0126	0.9826 ±0.0133	0.982 ±0.0131
D23	0.991 ±0.0025	0.9916 ±0.0025	0.9924 ±0.0026	**0.9928** ±0.0028
D27	0.7929 ±0.0139	0.7943 ±0.0142	0.7936 ±0.015	**0.7956** ±0.0127
D28	0.7912 ±0.0149	**0.7919** ±0.0161	0.7898 ±0.0149	0.7904 ±0.0147
D29	0.985 ±0.0039	0.9852 ±0.0042	0.9855 ±0.0044	**0.9861** ±0.0043
D31	0.7413 ±0.0342	**0.7499** ±0.0356	0.7434 ±0.0401	0.7381 ±0.0264
D32	0.3285 ±0.0185	0.3289 ±0.0154	**0.3302** ±0.0173	0.3282 ±0.0194
D34	0.998 ±0.0026	**0.9983** ±0.0027	0.9982 ±0.0027	0.9981 ±0.003
D35	0.2881 ±0.0358	0.2921 ±0.0287	**0.2924** ±0.0295	0.2813 ±0.0222

MDGP-Forest is configured with six different gen values $\{5,10,20,30,40,50\}$, and the experimental results are presented in Table S31, Table S32, and Table S33.

Table S31 F1 results of the different gen parameter values

Dataset	*gen*=5	*gen*=10	*gen*=20	*gen*=30	*gen*=40	*gen*=50
D2	**0.897** ±0.0243	0.8946 ±0.0275	0.8964 ±0.0242	0.8959 ±0.0251	**0.897** ±0.0246	0.8962 ±0.0245
D3	0.9086 ±0.0089	0.9098 ±0.0089	**0.9104** ±0.0091	0.91 ±0.0086	0.9093 ±0.0081	0.9091 ±0.0092
D4	0.8093 ±0.0629	0.8089 ±0.0652	**0.8128** ±0.0707	0.8079 ±0.0689	0.805 ±0.0647	0.8094 ±0.0611
D6	0.9573	0.9608	0.9654	0.9649	0.967	**0.9716**

	±0.0278	±0.0301	±0.0265	±0.0276	±0.022	±0.0213
D7	0.9147	0.9237	0.9392	0.9438	0.9577	**0.9618**
	±0.0411	±0.0335	±0.0282	±0.0474	±0.0271	±0.0278
D8	0.6799	**0.6877**	0.6823	0.6777	0.6842	0.6806
	±0.0946	±0.0939	±0.1001	±0.1034	±0.1014	±0.1022
D9	0.9307	0.9292	0.9505	**0.9519**	0.9397	0.9506
	±0.084	±0.083	±0.074	±0.0731	±0.0781	±0.0724
D10	0.7595	0.7492	**0.7636**	0.7556	0.7587	0.7472
	±0.0498	±0.0676	±0.0552	±0.0522	±0.052	±0.0569
D11	0.9805	0.9803	0.9812	0.9813	0.9809	**0.9815**
	±0.0072	±0.0066	±0.0072	±0.0072	±0.0075	±0.0069
D12	0.8147	0.8166	0.8149	0.8167	**0.8177**	0.817
	±0.0334	±0.035	±0.0355	±0.0335	±0.0381	±0.0341
D13	0.3479	0.3571	0.341	0.3547	**0.3579**	0.3556
	±0.0741	±0.0936	±0.0778	±0.0894	±0.0842	±0.085
D14	0.8428	0.8558	**0.8631**	0.8529	0.861	0.856
	±0.0639	±0.0541	±0.0472	±0.0602	±0.0476	±0.0508
D15	0.9971	0.9969	**0.9972**	0.997	0.9969	**0.9972**
	±0.0023	±0.0023	±0.0022	±0.0025	±0.0023	±0.0023
D16	**0.3488**	0.34	0.3457	0.3406	0.3424	0.3472
	±0.026	±0.0265	±0.0272	±0.0271	±0.0263	±0.0235
D17	0.8823	0.883	0.8826	**0.8832**	0.8811	0.8821
	±0.0136	±0.0125	±0.0129	±0.014	±0.0123	±0.0117
D18	0.58	0.5833	**0.5835**	0.582	0.5809	0.581
	±0.0632	±0.0611	±0.0614	±0.0608	±0.0629	±0.0645
D19	0.878	0.8807	0.8834	**0.8852**	0.8836	0.8844
	±0.0371	±0.0391	±0.0368	±0.0382	±0.0366	±0.0366
D20	0.4229	0.4256	0.4221	0.4255	0.4236	**0.4267**
	±0.0142	±0.0158	±0.0159	±0.0147	±0.0157	±0.0121
D21	0.9524	0.9518	0.9551	**0.9558**	0.9531	0.9545
	±0.0182	±0.0183	±0.0211	±0.0219	±0.0204	±0.0252
D22	0.9826	**0.9846**	**0.9846**	0.9844	0.9845	0.9837
	±0.012	±0.0112	±0.0118	±0.012	±0.0122	±0.012
D23	0.991	0.9918	0.9924	0.9924	**0.9932**	**0.9932**
	±0.0031	±0.0026	±0.0023	±0.0029	±0.0024	±0.0023
D27	0.8616	0.8614	0.8612	**0.8624**	0.8612	0.8612
	±0.0097	±0.0093	±0.01	±0.0099	±0.0095	±0.0098
D28	0.858	0.8594	**0.8596**	0.8585	0.858	0.8576
	±0.0109	±0.0113	±0.0112	±0.0112	±0.0109	±0.0113
D29	**0.9868**	**0.9868**	0.9867	0.9866	0.9865	**0.9868**
	±0.0032	±0.0035	±0.0038	±0.0037	±0.0034	±0.0037
D31	0.8003	0.8061	0.8103	0.8094	0.8108	**0.8149**
	±0.0298	±0.0263	±0.0261	±0.0288	±0.032	±0.03

D32	**0.552** ±0.0104	0.5517 ±0.0123	0.5489 ±0.009	0.5502 ±0.0124	0.5462 ±0.0122	0.5463 ±0.0121
D34	0.9982 ±0.0022	0.9979 ±0.0025	0.9986 ±0.0023	0.9988 ±0.002	0.9981 ±0.0027	**0.9989** ±0.0019
D35	0.5192 ±0.0243	0.515 ±0.0262	0.5227 ±0.0214	0.5144 ±0.0224	**0.5228** ±0.0235	0.52 ±0.0211

Table S32 G-mean results of the different gen parameter values

Dataset	*gen*=5	*gen*=10	*gen*=20	*gen*=30	*gen*=40	*gen*=50
D2	0.9285 ±0.0162	0.9262 ±0.0187	0.9277 ±0.0168	0.9275 ±0.0175	**0.9291** ±0.0162	0.9278 ±0.0163
D3	0.943 ±0.0055	0.9437 ±0.0056	**0.9441** ±0.0057	0.9437 ±0.0053	0.9432 ±0.0051	0.9433 ±0.0057
D4	0.8686 ±0.044	0.8685 ±0.0462	**0.8706** ±0.0502	0.8674 ±0.0488	0.8652 ±0.0458	0.8685 ±0.0433
D6	0.9601 ±0.0325	0.9641 ±0.0326	0.9712 ±0.0299	0.9721 ±0.0291	0.9752 ±0.0247	**0.9796** ±0.0232
D7	0.9455 ±0.0283	0.9491 ±0.0236	0.9605 ±0.0203	0.9602 ±0.0315	0.9681 ±0.0203	**0.9738** ±0.0189
D8	0.7421 ±0.0655	0.7469 ±0.0648	0.7465 ±0.0694	0.7429 ±0.072	**0.75** ±0.0718	0.7459 ±0.0717
D9	0.9674 ±0.0398	0.9665 ±0.0396	0.9769 ±0.0347	**0.9775** ±0.0341	0.9713 ±0.037	0.9759 ±0.0351
D10	0.846 ±0.0355	0.8388 ±0.0454	**0.8482** ±0.0397	0.8437 ±0.0367	0.8432 ±0.0352	0.8398 ±0.0398
D11	0.9831 ±0.0061	0.9828 ±0.0061	0.9833 ±0.0066	0.9834 ±0.0064	0.983 ±0.0066	**0.9836** ±0.0064
D12	0.8794 ±0.0244	0.8802 ±0.0249	0.879 ±0.0256	0.8807 ±0.0237	**0.8812** ±0.0272	0.8809 ±0.0245
D13	0.5495 ±0.0538	0.5563 ±0.0698	0.5459 ±0.0572	0.5516 ±0.063	**0.5572** ±0.0596	0.554 ±0.0598
D14	0.9023 ±0.0319	0.9084 ±0.0279	**0.9134** ±0.0231	0.9106 ±0.0299	0.9127 ±0.0231	0.9107 ±0.0248
D15	0.9975 ±0.0021	0.9974 ±0.0021	**0.9978** ±0.002	0.9976 ±0.0023	0.9976 ±0.0021	**0.9978** ±0.002
D16	**0.5613** ±0.0234	0.555 ±0.0256	0.5587 ±0.0247	0.5559 ±0.025	0.5573 ±0.0247	0.5601 ±0.0215
D17	0.9162 ±0.0087	0.9164 ±0.008	0.9164 ±0.0081	**0.9174** ±0.0092	0.9153 ±0.0086	0.9166 ±0.0078
D18	0.7355 ±0.0432	0.7366 ±0.0422	**0.7369** ±0.0412	0.7357 ±0.042	0.734 ±0.0429	0.7344 ±0.0451
D19	0.9204 ±0.0218	0.9198 ±0.0257	0.924 ±0.0214	0.9245 ±0.0235	0.9229 ±0.0209	**0.9252** ±0.0211

D20	0.5965 ±0.0103	0.5984 ±0.0111	0.596 ±0.0116	0.5988 ±0.0105	0.5973 ±0.0114	**0.5994** ±0.009
D21	0.9644 ±0.0135	0.9639 ±0.0136	0.9664 ±0.0157	**0.9669** ±0.0164	0.9649 ±0.0152	0.9659 ±0.0188
D22	0.9905 ±0.0065	**0.9915** ±0.0061	**0.9915** ±0.0064	**0.9915** ±0.0065	**0.9915** ±0.0066	0.9911 ±0.0065
D23	0.9951 ±0.0017	0.9955 ±0.0014	0.9958 ±0.0013	0.9958 ±0.0016	**0.9963** ±0.0013	**0.9963** ±0.0013
D27	0.8963 ±0.0071	0.896 ±0.0069	0.8959 ±0.0073	**0.8967** ±0.0072	0.8959 ±0.007	0.8958 ±0.0072
D28	0.8933 ±0.0082	0.8944 ±0.0084	**0.8945** ±0.0084	0.8937 ±0.0085	0.8933 ±0.0083	0.893 ±0.0085
D29	**0.9926** ±0.0018	**0.9926** ±0.002	**0.9926** ±0.0021	0.9925 ±0.0021	0.9925 ±0.0019	**0.9926** ±0.0021
D31	0.8664 ±0.0204	0.8704 ±0.018	0.8732 ±0.0181	0.8726 ±0.0197	0.8732 ±0.022	**0.8758** ±0.0207
D32	**0.6578** ±0.0091	0.6575 ±0.0108	0.6549 ±0.0081	0.6558 ±0.0102	0.6528 ±0.0111	0.6527 ±0.0104
D34	0.999 ±0.0012	0.9989 ±0.0014	0.9992 ±0.0013	**0.9993** ±0.0011	0.9989 ±0.0015	**0.9993** ±0.0011
D35	0.6283 ±0.0186	0.6249 ±0.0198	**0.6313** ±0.0157	0.625 ±0.0175	**0.6313** ±0.0181	0.6292 ±0.0161

Table S33 MCC results of the different gen parameter values

Dataset	*gen*=5	*gen*=10	*gen*=20	*gen*=30	*gen*=40	*gen*=50
D2	0.864 ±0.0341	0.8615 ±0.0367	0.8631 ±0.0334	0.8626 ±0.0336	**0.8641** ±0.0333	0.8629 ±0.0319
D3	0.9082 ±0.0093	0.9091 ±0.0088	**0.9099** ±0.0095	0.9092 ±0.0088	0.9086 ±0.0084	0.9085 ±0.0093
D4	0.7642 ±0.0749	0.7644 ±0.0759	**0.7691** ±0.085	0.7648 ±0.0819	0.7595 ±0.0781	0.765 ±0.0741
D6	0.9357 ±0.04	0.9418 ±0.0427	0.9472 ±0.0388	0.946 ±0.0411	0.9484 ±0.034	**0.9554** ±0.0329
D7	0.9258 ±0.0312	0.9319 ±0.0269	0.945 ±0.0229	0.9493 ±0.035	0.9586 ±0.024	**0.9635** ±0.0273
D8	0.5673 ±0.0872	0.5737 ±0.0867	0.5765 ±0.0907	0.5689 ±0.0998	**0.5785** ±0.096	0.5742 ±0.0969
D9	0.9698 ±0.0354	0.9686 ±0.0355	0.9741 ±0.0389	**0.9765** ±0.0342	0.9725 ±0.0326	**0.9765** ±0.0317
D10	0.7383 ±0.0385	0.7292 ±0.0543	**0.7411** ±0.0442	0.728 ±0.0461	0.7287 ±0.0348	0.7278 ±0.0407
D11	0.9703 ±0.0105	0.9698 ±0.0097	0.9709 ±0.0107	**0.9711** ±0.0113	0.9704 ±0.0115	0.9709 ±0.0103

D12	0.7371 ±0.03	0.7399 ±0.0326	0.7388 ±0.0342	<u>0.7414</u> ±0.0315	**0.7445** ±0.0396	0.7413 ±0.0347
D13	0.36 ±0.0734	**0.3699** ±0.0921	0.3483 ±0.0699	0.3566 ±0.0758	<u>0.3644</u> ±0.0824	0.357 ±0.0697
D14	0.8346 ±0.0252	0.8387 ±0.0273	**0.8408** ±0.0329	0.8391 ±0.032	<u>0.84</u> ±0.0299	0.8329 ±0.0282
D15	0.997 ±0.002	0.997 ±0.0021	<u>0.9972</u> ±0.002	<u>0.9972</u> ±0.0021	**0.9973** ±0.002	<u>0.9972</u> ±0.002
D16	**0.5062** ±0.0598	0.4971 ±0.0578	0.501 ±0.0613	0.4988 ±0.0566	0.5018 ±0.0583	<u>0.5034</u> ±0.0551
D17	0.9147 ±0.0046	0.9148 ±0.0059	**0.9151** ±0.0052	<u>0.915</u> ±0.0058	0.9137 ±0.0054	0.9144 ±0.0049
D18	0.5009 ±0.0302	0.4998 ±0.0328	0.4998 ±0.0318	<u>0.5015</u> ±0.034	**0.5033** ±0.0299	0.4984 ±0.037
D19	0.8726 ±0.0226	0.8754 ±0.0275	0.8772 ±0.0239	<u>0.8779</u> ±0.0255	0.8761 ±0.0231	**0.8787** ±0.0254
D20	<u>0.5195</u> ±0.0206	0.5194 ±0.0221	0.5146 ±0.02	**0.52** ±0.0216	0.517 ±0.0213	0.5188 ±0.0185
D21	0.9313 ±0.0258	0.9304 ±0.0258	<u>0.9351</u> ±0.0301	**0.9362** ±0.0315	0.9323 ±0.0293	0.9341 ±0.0365
D22	0.9812 ±0.0128	**0.9833** ±0.0121	**0.9833** ±0.0126	<u>0.9832</u> ±0.0129	**0.9833** ±0.0129	0.9824 ±0.0128
D23	0.9901 ±0.0034	0.991 ±0.0028	<u>0.9916</u> ±0.0025	<u>0.9916</u> ±0.0032	**0.9925** ±0.0026	**0.9925** ±0.0025
D27	<u>0.7951</u> ±0.0135	0.7943 ±0.0135	0.7943 ±0.0142	**0.7959** ±0.0139	0.7942 ±0.0134	0.7941 ±0.014
D28	0.7896 ±0.0157	<u>0.7918</u> ±0.0161	**0.7919** ±0.0161	0.7902 ±0.0166	0.7895 ±0.016	0.7891 ±0.0164
D29	**0.9853** ±0.0035	**0.9853** ±0.0039	<u>0.9852</u> ±0.0042	0.9851 ±0.0041	0.985 ±0.0038	**0.9853** ±0.0041
D31	0.7369 ±0.0395	0.7446 ±0.0352	<u>0.7499</u> ±0.0356	0.7488 ±0.0386	0.7498 ±0.0429	**0.7553** ±0.0406
D32	**0.3345** ±0.0179	<u>0.3339</u> ±0.0207	0.3289 ±0.0154	0.3304 ±0.0196	0.3247 ±0.0212	0.3245 ±0.0203
D34	0.9979 ±0.0026	0.9976 ±0.0029	0.9983 ±0.0027	<u>0.9985</u> ±0.0024	0.9977 ±0.0032	**0.9986** ±0.0023
D35	0.2884 ±0.0349	0.2817 ±0.0366	**0.2921** ±0.0287	0.2807 ±0.0321	<u>0.2916</u> ±0.0349	0.2882 ±0.0292

4.2 Supplementary parameter studies

This section supplements six parameter studies for MDGP-Forest, focusing on crossover rate, mutation rate, population size, max depth, min samples split, and min samples leaf. To ensure consistency in the experimental setup, results from high-dimensional datasets will be excluded from these experiments.

The crossover rate cr determines the probability that two parent individuals will exchange subtrees to generate offspring during each generation of category-related GP feature construction in MDGP-Forest. Crossover primarily facilitates local improvements by utilizing effective features from the current population. A higher crossover rate indicates the algorithm's stronger tendency to optimize existing features rather than explore new ones. MDGP-Forest is configured with five different cr values $\{0.3,0.4,0.5,0.6,0.7\}$, and the experimental results are presented in Table S34, Table S35, and Table S36.

Table S34 F1 results of the different cr parameter values

Dataset	cr=0.3	cr =0.4	cr =0.5	cr =0.6	cr =0.7
D2	<u>0.8962</u> ±0.0249	0.8952 ±0.0273	**0.8964** ±0.0242	0.8949 ±0.0273	0.8952 ±0.0269
D3	0.9092 ±0.008	<u>0.9097</u> ±0.0085	**0.9104** ±0.0091	0.9082 ±0.0083	0.9096 ±0.0086
D4	0.8038 ±0.0614	0.8086 ±0.0646	**0.8128** ±0.0707	<u>0.8109</u> ±0.0673	0.8069 ±0.0646
D6	<u>0.9673</u> ±0.0238	0.9669 ±0.0277	0.9654 ±0.0265	**0.9684** ±0.028	0.9648 ±0.0296
D7	0.9371 ±0.0364	0.9326 ±0.0338	<u>0.9392</u> ±0.0282	**0.9436** ±0.0304	0.9391 ±0.0367
D8	0.6825 ±0.1019	0.6814 ±0.0997	0.6823 ±0.1001	**0.6857** ±0.1009	<u>0.6838</u> ±0.1043
D9	0.949 ±0.0743	0.941 ±0.0786	**0.9505** ±0.074	<u>0.9499</u> ±0.0737	0.9368 ±0.0792
D10	0.7588 ±0.0608	0.7549 ±0.0475	**0.7636** ±0.0552	0.7481 ±0.0513	<u>0.7601</u> ±0.0506
D11	<u>0.9804</u> ±0.0063	0.9799 ±0.0086	**0.9812** ±0.0072	0.9798 ±0.0076	**0.9812** ±0.0068
D12	0.8146 ±0.036	<u>0.8153</u> ±0.0346	0.8149 ±0.0355	0.815 ±0.0352	**0.8157** ±0.0337
D13	0.3594 ±0.0908	0.348 ±0.0803	0.341 ±0.0778	**0.3668** ±0.0857	<u>0.3602</u> ±0.0878
D14	0.8528 ±0.0654	0.8508 ±0.0634	**0.8631** ±0.0472	0.848 ±0.0598	<u>0.8603</u> ±0.0576
D15	<u>0.9971</u> ±0.0024	0.9967 ±0.0025	**0.9972** ±0.0022	0.997 ±0.0023	0.997 ±0.0023
D16	<u>0.3445</u> ±0.0246	0.3419 ±0.0234	**0.3457** ±0.0272	0.3418 ±0.0261	0.3399 ±0.0304
D17	0.8821 ±0.0118	**0.8834** ±0.0127	<u>0.8826</u> ±0.0129	0.8818 ±0.0131	0.8799 ±0.0136
D18	0.5831	<u>0.5855</u>	0.5835	0.5829	**0.588**

	±0.0644	±0.0614	±0.0614	±0.0614	±0.0606
D19	0.8813	0.8827	**0.8834**	0.8823	0.8826
	±0.0381	±0.038	±0.0368	±0.037	±0.0336
D20	**0.4341**	0.4281	0.4221	0.4249	0.4242
	±0.0132	±0.0188	±0.0159	±0.0143	±0.0164
D21	0.9497	0.9531	0.9551	**0.9558**	**0.9558**
	±0.0219	±0.0204	±0.0211	±0.0197	±0.0208
D22	**0.9852**	0.9843	0.9846	0.9842	0.9835
	±0.011	±0.0116	±0.0118	±0.0111	±0.0117
D23	0.9921	0.9922	0.9924	**0.9926**	0.9924
	±0.0025	±0.003	±0.0023	±0.0026	±0.0028
D27	0.8615	0.8614	0.8612	**0.8622**	0.862
	±0.0097	±0.0088	±0.01	±0.0098	±0.0094
D28	0.8585	0.8582	**0.8596**	0.8584	0.8576
	±0.01	±0.0095	±0.0112	±0.0108	±0.0079
D29	0.986	**0.9867**	**0.9867**	**0.9867**	0.9865
	±0.0038	±0.0035	±0.0038	±0.0034	±0.0036
D31	0.8051	0.8045	**0.8103**	0.8058	0.8083
	±0.0289	±0.0291	±0.0261	±0.0248	±0.0295
D32	0.5498	0.5491	0.5489	**0.5507**	0.5478
	±0.0114	±0.0128	±0.009	±0.0101	±0.0113
D34	0.998	0.9986	0.9986	**0.9987**	0.9983
	±0.0024	±0.0021	±0.0023	±0.002	±0.0023
D35	**0.5231**	0.5217	0.5227	0.523	0.5206
	±0.025	±0.0259	±0.0214	±0.0248	±0.0253

Table S35 G-mean results of the different *cr* parameter values

Dataset	*cr*=0.3	*cr* =0.4	*cr* =0.5	*cr* =0.6	*cr* =0.7
D2	**0.928**	0.9273	0.9277	0.926	0.9267
	±0.0166	±0.0185	±0.0168	±0.0189	±0.0187
D3	0.9433	0.9436	**0.9441**	0.9426	0.9435
	±0.005	±0.0054	±0.0057	±0.0053	±0.0053
D4	0.8655	0.8676	**0.8706**	0.8704	0.8675
	±0.0438	±0.0459	±0.0502	±0.048	±0.0458
D6	**0.9736**	0.9735	0.9712	0.9722	0.9712
	±0.0278	±0.0265	±0.0299	±0.0306	±0.0294
D7	0.9573	0.9571	0.9605	**0.9619**	0.9562
	±0.0236	±0.021	±0.0203	±0.0207	±0.0273
D8	0.7461	0.7447	0.7465	0.7478	**0.7484**
	±0.071	±0.0696	±0.0694	±0.07	±0.0715
D9	0.9757	0.972	**0.9769**	0.9759	0.9695
	±0.0353	±0.0374	±0.0347	±0.0351	±0.0379
D10	0.8452	0.8431	**0.8482**	0.8401	0.846

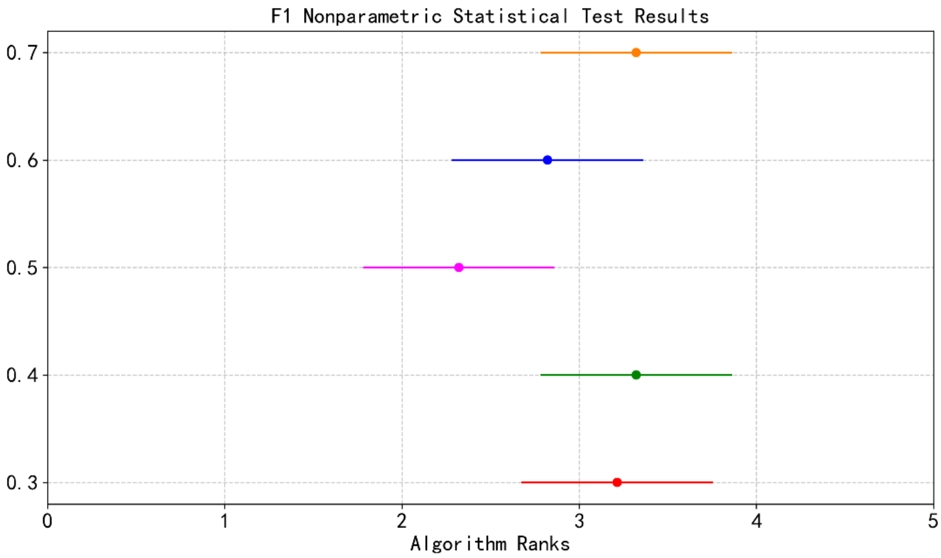
	±0.0427	±0.0337	±0.0397	±0.0358	±0.0372
D11	<u>0.9833</u>	0.9826	<u>0.9833</u>	0.9824	**0.9836**
	±0.006	±0.0079	±0.0066	±0.0065	±0.0061
D12	0.8789	<u>0.8796</u>	0.879	0.8792	**0.8799**
	±0.0254	±0.025	±0.0256	±0.0254	±0.025
D13	0.5585	0.5517	0.5459	**0.5647**	<u>0.5589</u>
	±0.0636	±0.0574	±0.0572	±0.0654	±0.06
D14	0.9087	0.9088	**0.9134**	0.9066	<u>0.9109</u>
	±0.0326	±0.0314	±0.0231	±0.0298	±0.0307
D15	**0.9978**	0.9974	**0.9978**	**0.9978**	<u>0.9977</u>
	±0.0021	±0.0021	±0.002	±0.0021	±0.0021
D16	<u>0.5575</u>	<u>0.5575</u>	**0.5587**	0.5566	0.555
	±0.0222	±0.0233	±0.0247	±0.024	±0.0271
D17	0.9158	**0.9166**	<u>0.9164</u>	0.9156	0.9146
	±0.0074	±0.0082	±0.0081	±0.0082	±0.0091
D18	0.7366	**0.7379**	<u>0.7369</u>	0.7363	<u>0.7369</u>
	±0.0456	±0.0421	±0.0412	±0.0426	±0.0401
D19	0.9234	<u>0.9235</u>	**0.924**	0.9219	0.923
	±0.0237	±0.0222	±0.0214	±0.0231	±0.0199
D20	**0.6056**	<u>0.6001</u>	0.596	0.5978	0.5977
	±0.0114	±0.0136	±0.0116	±0.0108	±0.012
D21	0.9624	0.9649	<u>0.9664</u>	**0.9669**	**0.9669**
	±0.0163	±0.0152	±0.0157	±0.0147	±0.0156
D22	**0.9919**	0.9914	<u>0.9915</u>	0.9913	0.991
	±0.006	±0.0064	±0.0064	±0.0061	±0.0064
D23	0.9957	0.9957	<u>0.9958</u>	**0.9959**	<u>0.9958</u>
	±0.0014	±0.0017	±0.0013	±0.0014	±0.0015
D27	0.8961	0.8961	0.8959	**0.8966**	<u>0.8965</u>
	±0.0071	±0.0064	±0.0073	±0.0071	±0.0069
D28	<u>0.8937</u>	0.8935	**0.8945**	0.8936	0.8929
	±0.0075	±0.0071	±0.0084	±0.0081	±0.0059
D29	0.9922	**0.9926**	**0.9926**	**0.9926**	<u>0.9925</u>
	±0.0021	±0.002	±0.0021	±0.0019	±0.002
D31	0.8694	0.8691	**0.8732**	0.8701	<u>0.8716</u>
	±0.0199	±0.0201	±0.0181	±0.0172	±0.0202
D32	<u>0.656</u>	0.6554	0.6549	**0.6574**	0.6543
	±0.0104	±0.0108	±0.0081	±0.0083	±0.0096
D34	0.9989	**0.9992**	**0.9992**	**0.9992**	<u>0.999</u>
	±0.0013	±0.0012	±0.0013	±0.0012	±0.0013
D35	**0.6317**	0.6305	<u>0.6313</u>	**0.6317**	0.6297
	±0.0187	±0.0197	±0.0157	±0.0179	±0.019

Table S36 MCC results of the different *cr* parameter values

Dataset	*cr*=0.3	*cr* =0.4	*cr* =0.5	*cr* =0.6	*cr* =0.7
D2	<u>0.863</u> ±0.0336	0.8613 ±0.0363	**0.8631** ±0.0334	0.8603 ±0.036	0.862 ±0.0353
D3	0.9087 ±0.008	0.9089 ±0.0087	**0.9099** ±0.0095	0.9073 ±0.0085	<u>0.9091</u> ±0.0088
D4	0.7584 ±0.0732	0.7641 ±0.078	**0.7691** ±0.085	<u>0.7672</u> ±0.0794	0.7626 ±0.0766
D6	<u>0.9492</u> ±0.0365	0.9491 ±0.0414	0.9472 ±0.0388	**0.9518** ±0.0412	0.946 ±0.0442
D7	0.9409 ±0.0307	0.939 ±0.0267	<u>0.945</u> ±0.0229	**0.9474** ±0.0251	0.9444 ±0.0295
D8	0.5738 ±0.095	0.5692 ±0.0937	<u>0.5765</u> ±0.0907	**0.5773** ±0.0946	0.5763 ±0.0962
D9	<u>0.9754</u> ±0.0341	0.9724 ±0.0352	0.9741 ±0.0389	**0.9763** ±0.0319	0.9712 ±0.0329
D10	<u>0.7371</u> ±0.0463	0.7341 ±0.0381	**0.7411** ±0.0442	0.7258 ±0.0446	0.7324 ±0.0378
D11	0.9702 ±0.0095	0.9693 ±0.0134	<u>0.9709</u> ±0.0107	0.969 ±0.0117	**0.9711** ±0.0102
D12	0.7386 ±0.0329	<u>0.7398</u> ±0.0329	0.7388 ±0.0342	0.7385 ±0.0336	**0.7406** ±0.0329
D13	**0.373** ±0.0829	0.3572 ±0.0708	0.3483 ±0.0699	0.3683 ±0.0895	<u>0.3698</u> ±0.0725
D14	0.8406 ±0.0308	0.8394 ±0.033	<u>0.8408</u> ±0.0329	0.8351 ±0.0273	**0.8417** ±0.0302
D15	<u>0.9974</u> ±0.0019	0.9971 ±0.0021	0.9972 ±0.002	**0.9975** ±0.002	0.9972 ±0.0021
D16	0.4987 ±0.0543	**0.506** ±0.0593	<u>0.501</u> ±0.0613	0.5005 ±0.0579	0.4973 ±0.0599
D17	0.9139 ±0.0052	<u>0.9149</u> ±0.0055	**0.9151** ±0.0052	0.9145 ±0.0056	0.913 ±0.005
D18	0.5004 ±0.0358	**0.5032** ±0.0314	0.4998 ±0.0318	<u>0.5011</u> ±0.0338	0.4974 ±0.0301
D19	0.8759 ±0.0249	<u>0.8768</u> ±0.0258	**0.8772** ±0.0239	0.8763 ±0.0261	0.8753 ±0.0214
D20	**0.5246** ±0.0154	<u>0.5187</u> ±0.0209	0.5146 ±0.02	0.5165 ±0.0222	0.5165 ±0.0214
D21	0.9277 ±0.0305	0.9323 ±0.029	<u>0.9351</u> ±0.0301	**0.936** ±0.0286	**0.936** ±0.0302
D22	**0.9839** ±0.0119	0.9829 ±0.0126	<u>0.9833</u> ±0.0126	0.9828 ±0.0119	0.9822 ±0.0125
D23	0.9913	0.9915	<u>0.9916</u>	**0.9919**	<u>0.9916</u>

	±0.0028	±0.0033	±0.0025	±0.0029	±0.003
D27	0.7948	0.7945	0.7943	**0.7957**	0.7954
	±0.0137	±0.0124	±0.0142	±0.0137	±0.0133
D28	0.7903	0.79	**0.7919**	0.7902	0.7887
	±0.0144	±0.0136	±0.0161	±0.0157	±0.0111
D29	0.9845	**0.9852**	**0.9852**	**0.9852**	0.985
	±0.0042	±0.004	±0.0042	±0.0038	±0.004
D31	0.7424	0.7421	**0.7499**	0.7441	0.7467
	±0.0391	±0.0394	±0.0356	±0.0336	±0.0394
D32	0.331	0.33	0.3289	**0.3337**	0.3277
	±0.0199	±0.0204	±0.0154	±0.0152	±0.0182
D34	0.9976	0.9983	0.9983	**0.9984**	0.998
	±0.0029	±0.0025	±0.0027	±0.0024	±0.0028
D35	**0.2945**	0.2916	0.2921	0.2933	0.2896
	±0.0344	±0.036	±0.0287	±0.0312	±0.0343

The Nonparametric Statistical Test was performed on the experimental data, and the results are shown in Fig S3. Optimal performance occurs at $cr = 0.5$, with a slight decline at 0.6. Values beyond this range (both higher and lower) result in significantly poorer performance. Setting the cr too high may cause rapid convergence of individuals in the algorithm population, leading to loss of diversity and entrapment in local optima. Conversely, setting it too low may result in inefficient evolution, causing the algorithm to degenerate into approximate random search with reduced convergence speed. Therefore, it is recommended to set the crossover rate of MDGP-Forest to a moderate value, preferably within the range of 0.4 to 0.6.



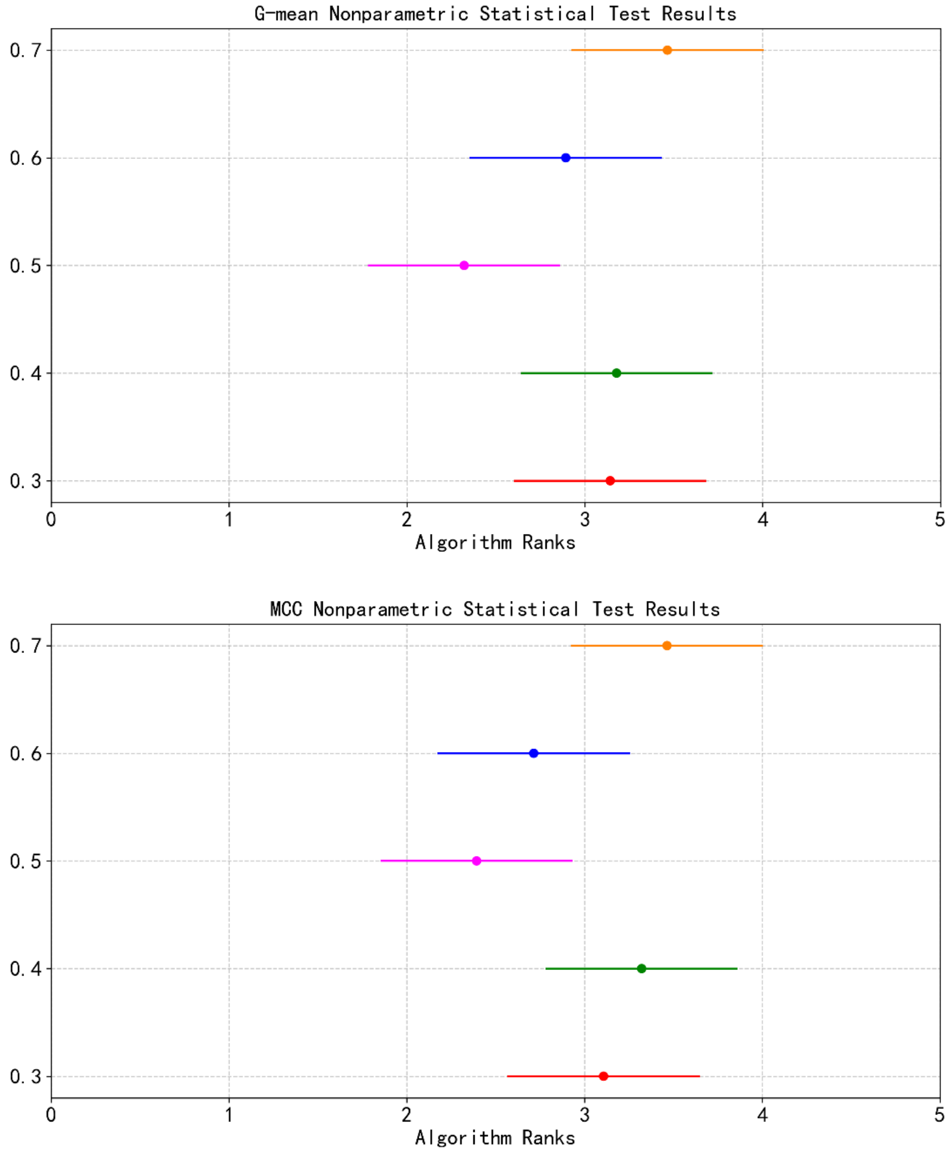


Fig S3 Nonparametric statistical test figure for results of different cr values.

The mutation rate mr determines the probability that individuals will be modified through random mutation operations in each generation of the category-related GP feature construction population in MDGP-Forest. This parameter controls the trade-off between exploiting existing high-quality features and globally exploring new feature possibilities in the algorithm. MDGP-Forest is configured with five different mr values {0.1,0.2,0.3,0.4,0.5}, and the experimental results are presented in Table S37, Table S38, and Table S39.

Table S37 F1 results of the different mr parameter values

Dataset	mr=0.1	mr =0.2	mr =0.3	mr=0.4	mr=0.5
D2	0.8943 ±0.0256	<u>0.8964</u> ±0.0242	0.8952 ±0.021	0.8949 ±0.0231	**0.8965** ±0.0266
D3	<u>0.9089</u>	**0.9104**	<u>0.9089</u>	<u>0.9089</u>	0.9085

	±0.0089	±0.0091	±0.0092	±0.0092	±0.0095
D4	0.8044 ±0.0638	<u>0.8128</u> ±0.0707	0.8043 ±0.066	**0.8144** ±0.0665	0.8082 ±0.0667
D6	0.9624 ±0.0276	0.9654 ±0.0265	**0.9707** ±0.0269	0.9652 ±0.0277	<u>0.967</u> ±0.0274
D7	0.9422 ±0.0378	0.9392 ±0.0282	<u>0.9434</u> ±0.0308	**0.9468** ±0.0285	0.942 ±0.0354
D8	0.6807 ±0.103	0.6823 ±0.1001	<u>0.685</u> ±0.1051	0.6826 ±0.1106	**0.6858** ±0.1046
D9	<u>0.9403</u> ±0.0771	**0.9505** ±0.074	0.938 ±0.0806	0.9386 ±0.0796	0.938 ±0.0801
D10	<u>0.7539</u> ±0.0515	**0.7636** ±0.0552	0.7524 ±0.049	0.7526 ±0.0606	0.746 ±0.0586
D11	0.9797 ±0.0075	**0.9812** ±0.0072	<u>0.9808</u> ±0.0072	0.98 ±0.0086	0.98 ±0.0073
D12	**0.8191** ±0.0352	0.8149 ±0.0355	0.814 ±0.035	<u>0.8158</u> ±0.035	0.813 ±0.0422
D13	<u>0.3676</u> ±0.092	0.341 ±0.0778	**0.3724** ±0.0788	0.3577 ±0.0738	0.3521 ±0.0835
D14	0.8444 ±0.0656	**0.8631** ±0.0472	0.8458 ±0.0655	<u>0.8623</u> ±0.0506	0.8618 ±0.0502
D15	0.997 ±0.0022	<u>0.9972</u> ±0.0022	**0.9973** ±0.0023	0.9971 ±0.0023	<u>0.9972</u> ±0.0021
D16	0.3406 ±0.0267	<u>0.3457</u> ±0.0272	0.3418 ±0.0313	0.3432 ±0.0249	**0.3458** ±0.029
D17	0.8821 ±0.0109	<u>0.8826</u> ±0.0129	0.8818 ±0.0134	**0.8836** ±0.0142	0.8821 ±0.0116
D18	0.5828 ±0.0619	<u>0.5835</u> ±0.0614	0.5822 ±0.0625	**0.6035** ±0.0621	0.5812 ±0.0672
D19	0.8802 ±0.0397	<u>0.8834</u> ±0.0368	0.8815 ±0.0394	**0.885** ±0.04	0.8812 ±0.0399
D20	0.4238 ±0.0162	0.4221 ±0.0159	**0.4271** ±0.0117	0.422 ±0.0165	<u>0.4252</u> ±0.0141
D21	0.9531 ±0.0216	**0.9551** ±0.0211	0.9538 ±0.0202	<u>0.9545</u> ±0.0232	0.953 ±0.0249
D22	0.9836 ±0.0128	0.9846 ±0.0118	0.9831 ±0.0124	<u>0.9847</u> ±0.0117	**0.9849** ±0.0105
D23	**0.9926** ±0.002	0.9924 ±0.0023	<u>0.9925</u> ±0.0028	**0.9926** ±0.0024	0.9924 ±0.0023
D27	<u>0.862</u> ±0.0103	0.8612 ±0.01	**0.8621** ±0.0098	**0.8621** ±0.0091	0.8614 ±0.0102
D28	<u>0.8587</u> ±0.0111	**0.8596** ±0.0112	0.8579 ±0.0107	0.8585 ±0.011	0.8582 ±0.0106

D29	**0.9867** ±0.0037	**0.9867** ±0.0038	<u>0.9866</u> ±0.0035	<u>0.9866</u> ±0.0034	<u>0.9866</u> ±0.0037
D31	0.8071 ±0.0311	**0.8103** ±0.0261	0.8025 ±0.0296	<u>0.8072</u> ±0.0323	0.8042 ±0.0308
D32	0.5484 ±0.0127	0.5489 ±0.009	<u>0.5516</u> ±0.0116	0.5468 ±0.013	**0.5527** ±0.0132
D34	0.9983 ±0.0025	<u>0.9986</u> ±0.0023	**0.9987** ±0.002	<u>0.9986</u> ±0.0024	0.9984 ±0.0025
D35	0.5222 ±0.0295	<u>0.5227</u> ±0.0214	0.5218 ±0.0188	**0.5248** ±0.0225	0.5208 ±0.0249

Table S38 G-mean results of the different mr parameter values

Dataset	*mr*=0.1	*mr* =0.2	*mr* =0.3	*mr*=0.4	*mr*=0.5
D2	0.926 ±0.0173	<u>0.9277</u> ±0.0168	0.9269 ±0.0138	0.9269 ±0.015	**0.9278** ±0.0181
D3	0.9432 ±0.0056	**0.9441** ±0.0057	<u>0.9433</u> ±0.0056	0.9432 ±0.0057	0.943 ±0.0058
D4	0.8654 ±0.0454	<u>0.8706</u> ±0.0502	0.8652 ±0.0473	**0.8725** ±0.0473	0.8684 ±0.0477
D6	0.9651 ±0.032	0.9712 ±0.0299	**0.9741** ±0.0282	0.9718 ±0.0289	<u>0.9731</u> ±0.0288
D7	0.9611 ±0.0236	0.9605 ±0.0203	<u>0.9619</u> ±0.0223	**0.9634** ±0.0203	0.9592 ±0.0236
D8	0.7463 ±0.0717	0.7465 ±0.0694	<u>0.7475</u> ±0.0725	0.7472 ±0.076	**0.7497** ±0.0728
D9	<u>0.9713</u> ±0.0371	**0.9769** ±0.0347	0.9705 ±0.038	0.9704 ±0.038	0.9704 ±0.038
D10	<u>0.8418</u> ±0.0371	**0.8482** ±0.0397	0.8408 ±0.0363	0.8409 ±0.0414	0.8385 ±0.04
D11	0.9822 ±0.0069	**0.9833** ±0.0066	<u>0.983</u> ±0.0069	0.9824 ±0.0074	0.9828 ±0.0063
D12	**0.8816** ±0.025	0.879 ±0.0256	0.8789 ±0.0251	<u>0.8797</u> ±0.0249	0.8785 ±0.0271
D13	<u>0.5658</u> ±0.0691	0.5459 ±0.0572	**0.5699** ±0.0609	0.56 ±0.0592	0.5562 ±0.0648
D14	0.903 ±0.0336	**0.9134** ±0.0231	0.9049 ±0.0336	<u>0.913</u> ±0.0265	0.9123 ±0.0256
D15	<u>0.9975</u> ±0.002	**0.9978** ±0.002	**0.9978** ±0.0021	**0.9978** ±0.0021	**0.9978** ±0.0019
D16	0.5554 ±0.0242	<u>0.5587</u> ±0.0247	0.5563 ±0.0281	0.5577 ±0.0221	**0.5592** ±0.0253
D17	0.9158 ±0.0076	<u>0.9164</u> ±0.0081	0.916 ±0.0089	**0.9168** ±0.0087	0.9162 ±0.0076

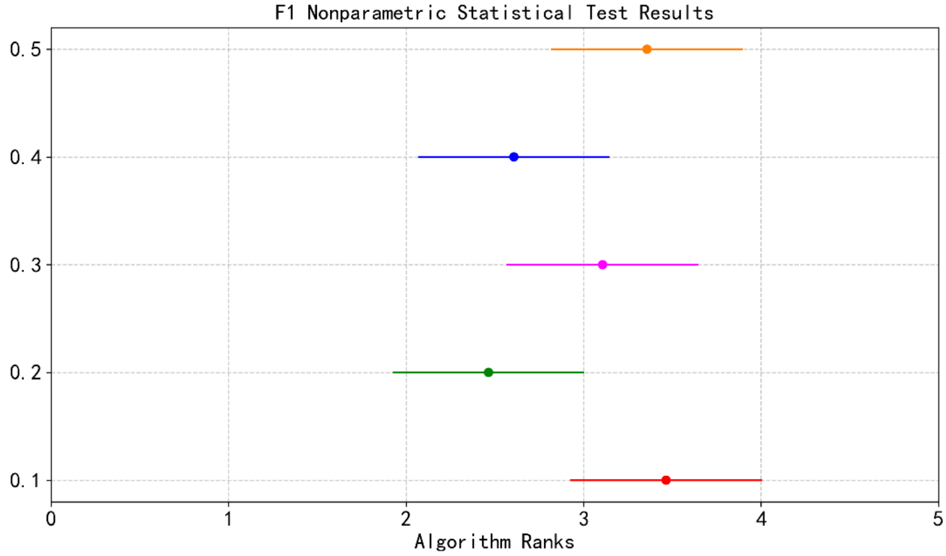
D18	<u>0.7373</u> ±0.0425	0.7369 ±0.0412	0.7361 ±0.0432	**0.7478** ±0.0426	0.7338 ±0.0446
D19	0.9235 ±0.0237	<u>0.924</u> ±0.0214	0.9235 ±0.0254	**0.9242** ±0.0243	0.9228 ±0.0244
D20	0.5968 ±0.0116	0.596 ±0.0116	**0.5989** ±0.0093	0.5956 ±0.0115	<u>0.5985</u> ±0.0103
D21	0.9649 ±0.016	**0.9664** ±0.0157	0.9654 ±0.0151	<u>0.9659</u> ±0.0174	0.9649 ±0.0183
D22	0.991 ±0.0069	<u>0.9915</u> ±0.0064	0.9907 ±0.0067	**0.9916** ±0.0063	**0.9916** ±0.0059
D23	**0.9959** ±0.0011	<u>0.9958</u> ±0.0013	**0.9959** ±0.0015	**0.9959** ±0.0013	<u>0.9958</u> ±0.0013
D27	<u>0.8965</u> ±0.0077	0.8959 ±0.0073	**0.8966** ±0.0072	<u>0.8965</u> ±0.0067	0.896 ±0.0076
D28	<u>0.8938</u> ±0.0083	**0.8945** ±0.0084	0.8933 ±0.008	0.8937 ±0.0083	0.8935 ±0.0079
D29	<u>0.9925</u> ±0.0021	**0.9926** ±0.0021	<u>0.9925</u> ±0.002	<u>0.9925</u> ±0.0019	<u>0.9925</u> ±0.0021
D31	<u>0.8711</u> ±0.0211	**0.8732** ±0.0181	0.8677 ±0.0207	0.871 ±0.0221	0.869 ±0.0213
D32	0.6546 ±0.0107	0.6549 ±0.0081	<u>0.6574</u> ±0.0101	0.6538 ±0.0112	**0.6579** ±0.0114
D34	0.999 ±0.0014	<u>0.9992</u> ±0.0013	**0.9993** ±0.0011	<u>0.9992</u> ±0.0014	0.9991 ±0.0014
D35	0.6309 ±0.022	<u>0.6313</u> ±0.0157	0.6307 ±0.0144	**0.6327** ±0.0168	0.6296 ±0.0193

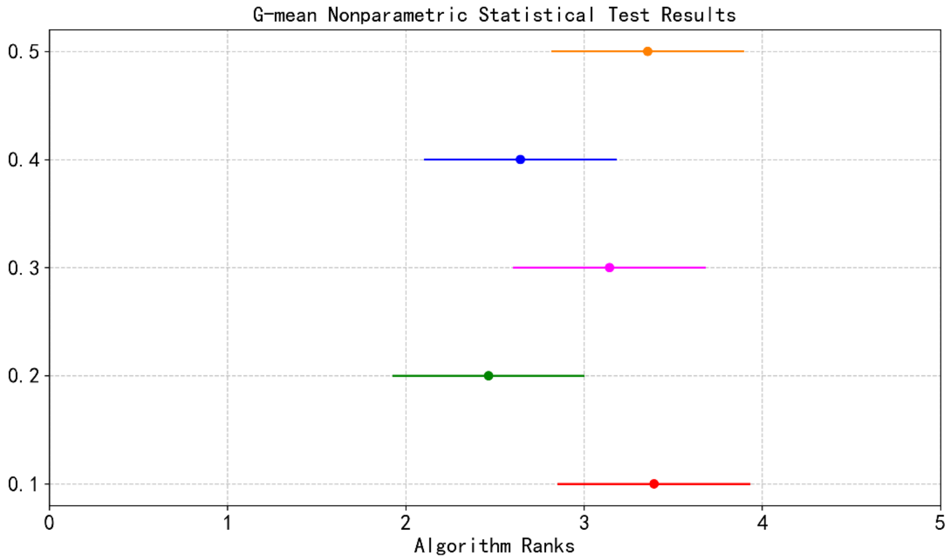
Table S39 MCC results of the different mr parameter values

Dataset	mr=0.1	mr =0.2	mr =0.3	mr=0.4	mr=0.5
D2	0.8613 ±0.035	<u>0.8631</u> ±0.0334	0.862 ±0.0294	0.8612 ±0.0305	**0.8633** ±0.0342
D3	0.9083 ±0.009	**0.9099** ±0.0095	<u>0.9085</u> ±0.0092	<u>0.9085</u> ±0.0095	0.9079 ±0.0097
D4	0.757 ±0.0781	<u>0.7691</u> ±0.085	0.7574 ±0.0791	**0.7714** ±0.0814	0.7635 ±0.0818
D6	0.9438 ±0.0394	0.9472 ±0.0388	**0.9557** ±0.0394	0.9461 ±0.042	<u>0.949</u> ±0.0405
D7	0.9465 ±0.0305	0.945 ±0.0229	<u>0.9467</u> ±0.0236	**0.9499** ±0.0254	0.9457 ±0.0286
D8	0.5725 ±0.0944	0.5765 ±0.0907	0.5775 ±0.0983	<u>0.5787</u> ±0.1057	**0.5804** ±0.1005
D9	<u>0.9725</u> ±0.0326	**0.9741** ±0.0389	<u>0.9725</u> ±0.0326	0.9724 ±0.0328	0.9724 ±0.0328

D10	0.7318 ±0.0409	**0.7411** ±0.0442	0.7291 ±0.0402	<u>0.7333</u> ±0.0399	0.725 ±0.045
D11	0.9689 ±0.011	**0.9709** ±0.0107	<u>0.9703</u> ±0.0115	0.9691 ±0.0135	0.97 ±0.0102
D12	**0.7471** ±0.0289	0.7388 ±0.0342	0.7389 ±0.0334	0.7395 ±0.0346	<u>0.7457</u> ±0.0362
D13	<u>0.3739</u> ±0.0935	0.3483 ±0.0699	**0.3747** ±0.079	0.3641 ±0.0802	0.3612 ±0.0842
D14	0.8354 ±0.0319	<u>0.8408</u> ±0.0329	0.8393 ±0.0292	0.8406 ±0.0292	**0.8411** ±0.0301
D15	0.9971 ±0.002	0.9972 ±0.002	**0.9974** ±0.0021	<u>0.9973</u> ±0.0019	<u>0.9973</u> ±0.0019
D16	0.4975 ±0.0558	<u>0.501</u> ±0.0613	0.4994 ±0.0631	0.5005 ±0.0532	**0.5033** ±0.0559
D17	0.9137 ±0.005	**0.9151** ±0.0052	0.9142 ±0.0055	<u>0.9146</u> ±0.0056	<u>0.9146</u> ±0.0045
D18	<u>0.5019</u> ±0.0333	0.4998 ±0.0318	<u>0.5019</u> ±0.0339	**0.5101** ±0.0356	0.4954 ±0.0329
D19	0.8756 ±0.0253	<u>0.8772</u> ±0.0239	0.8759 ±0.0249	**0.8779** ±0.0273	0.8739 ±0.0243
D20	<u>0.5163</u> ±0.0208	0.5146 ±0.02	0.5116 ±0.015	0.5154 ±0.0248	**0.518** ±0.0203
D21	0.9324 ±0.0303	**0.9351** ±0.0301	0.9332 ±0.0292	<u>0.9342</u> ±0.0336	0.9325 ±0.034
D22	0.9823 ±0.0136	0.9833 ±0.0126	0.9817 ±0.0132	**0.9835** ±0.0124	<u>0.9834</u> ±0.0116
D23	**0.9918** ±0.0022	0.9916 ±0.0025	**0.9918** ±0.0031	**0.9918** ±0.0027	<u>0.9917</u> ±0.0026
D27	0.7952 ±0.015	0.7943 ±0.0142	**0.7956** ±0.0138	<u>0.7954</u> ±0.0129	0.7945 ±0.0147
D28	<u>0.7907</u> ±0.0158	**0.7919** ±0.0161	0.7896 ±0.0153	0.7905 ±0.0159	0.7899 ±0.0153
D29	**0.9852** ±0.0042	**0.9852** ±0.0042	<u>0.9851</u> ±0.0039	<u>0.9851</u> ±0.0038	<u>0.9851</u> ±0.0042
D31	<u>0.7461</u> ±0.0411	**0.7499** ±0.0356	0.7394 ±0.0404	0.7458 ±0.0431	0.7418 ±0.0416
D32	0.3283 ±0.0204	0.3289 ±0.0154	<u>0.3338</u> ±0.0199	0.3271 ±0.0214	**0.3345** ±0.0221
D34	0.998 ±0.003	<u>0.9983</u> ±0.0027	**0.9984** ±0.0024	<u>0.9983</u> ±0.0029	0.9981 ±0.003
D35	<u>0.2928</u> ±0.0385	0.2921 ±0.0287	0.2917 ±0.0269	**0.2946** ±0.0295	0.2899 ±0.036

The Nonparametric Statistical Test was performed on the experimental data, and the results are shown in Fig S4. The results demonstrate that superior performance is achieved when mr is set to either 0.2 or 0.4, indicating that the optimal mr setting should be determined based on dataset characteristics. An excessively high setting may disrupt existing effective features, causing the algorithm to degenerate into random search behavior with impaired convergence. Conversely, an insufficiently low setting may fail to generate sufficient novel features, potentially trapping the algorithm in local optima. For easier feature construction scenarios, higher mutation rates (0.3-0.5) help explore wider solution spaces, while more challenging cases require lower rates (0.1-0.3) to focus on refining existing features.





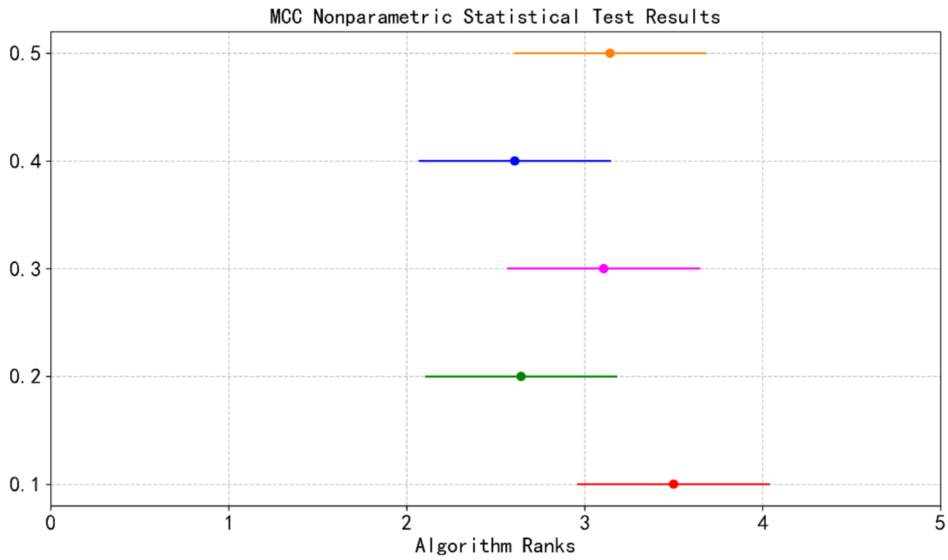


Fig S4 Nonparametric statistical test figure for results of different mr values.

The population size pop determines the number of candidate feature constructions (i.e., individuals) in each generation of category-related GP feature construction in MDGP-Forest. A larger population enables simultaneous exploration of more regions in the search space, thereby increasing the probability of discovering globally optimal feature constructions. MDGP-Forest is configured with five different pop values $\{10,30,50,70,90\}$, and the experimental results are presented in Table S40, Table S41, and Table S42.

Table S40 F1 results of the different pop parameter values

Dataset	*pop*=10	*pop* =30	*pop* =50	*pop* =70	*pop* =90
D2	0.8944 ±0.0256	**0.8964** ±0.0239	**0.8964** ±0.0242	0.8952 ±0.0258	<u>0.8953</u> ±0.024
D3	0.9085 ±0.0088	0.9094 ±0.0078	**0.9104** ±0.0091	0.9095 ±0.0083	<u>0.91</u> ±0.0083
D4	0.8079 ±0.0628	0.8078 ±0.0648	**0.8128** ±0.0707	0.8059 ±0.0658	<u>0.8095</u> ±0.0603
D6	0.9604 ±0.0311	0.963 ±0.0298	<u>0.9654</u> ±0.0265	**0.9656** ±0.0277	0.9605 ±0.0313
D7	0.8894 ±0.0477	0.9274 ±0.0335	0.9392 ±0.0282	<u>0.9518</u> ±0.03	**0.9602** ±0.0272
D8	0.6804 ±0.0924	<u>0.6872</u> ±0.1017	0.6823 ±0.1001	0.6778 ±0.107	**0.6896** ±0.1058
D9	0.9255 ±0.0839	0.9225 ±0.0839	**0.9505** ±0.074	0.9367 ±0.0789	<u>0.9458</u> ±0.0761
D10	0.7456 ±0.0522	0.7481 ±0.0546	**0.7636** ±0.0552	0.7507 ±0.0558	<u>0.7519</u> ±0.0565
D11	0.9802	<u>0.9812</u>	<u>0.9812</u>	**0.9816**	<u>0.9812</u>

	±0.0075	±0.0067	±0.0072	±0.0071	±0.0072
D12	0.8123	**0.8163**	0.8149	0.8153	0.8158
	±0.0368	±0.0347	±0.0355	±0.034	±0.0356
D13	0.3494	0.3502	0.341	**0.3519**	0.3503
	±0.0695	±0.0841	±0.0778	±0.0874	±0.0773
D14	0.8552	0.8624	**0.8631**	0.8549	0.8594
	±0.0584	±0.0497	±0.0472	±0.0558	±0.0531
D15	0.9971	0.9965	**0.9972**	0.9969	0.9968
	±0.0024	±0.0025	±0.0022	±0.0024	±0.0022
D16	0.3446	0.3416	**0.3457**	0.3382	0.3407
	±0.0212	±0.0261	±0.0272	±0.0248	±0.0249
D17	0.8837	0.8828	0.8826	**0.8844**	0.8822
	±0.0113	±0.014	±0.0129	±0.0131	±0.0135
D18	0.58	0.5798	0.5835	**0.5887**	0.5817
	±0.0615	±0.0639	±0.0614	±0.0566	±0.0632
D19	0.8778	0.881	0.8834	0.8826	**0.8845**
	±0.0408	±0.0369	±0.0368	±0.0371	±0.0381
D20	0.4226	**0.426**	0.4221	0.424	0.4253
	±0.014	±0.014	±0.0159	±0.0135	±0.0168
D21	0.9496	0.9522	**0.9551**	0.9524	0.9524
	±0.028	±0.0282	±0.0211	±0.0205	±0.0228
D22	0.9843	0.9836	**0.9846**	0.984	0.9843
	±0.0113	±0.0128	±0.0118	±0.0119	±0.0113
D23	0.9906	0.9916	0.9924	0.9928	**0.9931**
	±0.003	±0.0029	±0.0023	±0.0026	±0.0021
D27	0.8616	0.8605	0.8612	0.8608	**0.862**
	±0.0094	±0.0096	±0.01	±0.01	±0.0094
D28	0.8583	0.8583	**0.8596**	0.8585	0.8584
	±0.0098	±0.0091	±0.0112	±0.0106	±0.0102
D29	**0.9867**	**0.9867**	**0.9867**	0.9866	**0.9867**
	±0.0039	±0.0034	±0.0038	±0.0033	±0.0037
D31	0.7954	0.8028	0.8103	0.8113	**0.8118**
	±0.0239	±0.0295	±0.0261	±0.0312	±0.0299
D32	**0.5504**	0.5492	0.5489	0.5477	0.5462
	±0.0123	±0.0122	±0.009	±0.016	±0.014
D34	0.9984	0.9983	0.9986	**0.9987**	0.998
	±0.0023	±0.0023	±0.0023	±0.002	±0.0026
D35	0.5182	0.5225	**0.5227**	0.5202	0.5217
	±0.0257	±0.0217	±0.0214	±0.0223	±0.0237

Table S41 G-mean results of the different *pop* parameter values

Dataset	*pop*=10	*pop* =30	*pop* =50	*pop* =70	*pop* =90
D2	0.926 ±0.0177	**0.9283** ±0.0158	0.9277 ±0.0168	0.9269 ±0.0179	0.9268 ±0.0155
D3	0.9428 ±0.0054	0.9435 ±0.005	**0.9441** ±0.0057	0.9436 ±0.0051	0.944 ±0.0053
D4	0.8676 ±0.0434	0.867 ±0.046	**0.8706** ±0.0502	0.8663 ±0.047	0.8689 ±0.0428
D6	0.9635 ±0.0334	0.965 ±0.0319	0.9712 ±0.0299	**0.9718** ±0.0281	0.967 ±0.0324
D7	0.9288 ±0.0327	0.9532 ±0.022	0.9605 ±0.0203	0.9654 ±0.0208	**0.9716** ±0.0188
D8	0.7401 ±0.0624	0.7478 ±0.0689	0.7465 ±0.0694	0.7432 ±0.074	**0.7512** ±0.0756
D9	0.964 ±0.0405	0.9631 ±0.0401	**0.9769** ±0.0347	0.9695 ±0.0379	0.9741 ±0.0362
D10	0.8391 ±0.0359	0.8397 ±0.0369	**0.8482** ±0.0397	0.839 ±0.0382	0.8416 ±0.0388
D11	0.9828 ±0.0063	0.9832 ±0.0057	0.9833 ±0.0066	**0.984** ±0.0062	0.9835 ±0.0063
D12	0.8772 ±0.0263	0.8799 ±0.0247	0.879 ±0.0256	0.8799 ±0.024	**0.88** ±0.0249
D13	0.5505 ±0.0537	0.5517 ±0.0595	0.5459 ±0.0572	0.5533 ±0.0638	**0.5542** ±0.058
D14	0.9114 ±0.0266	**0.9137** ±0.0241	0.9134 ±0.0231	0.9101 ±0.0278	0.9125 ±0.0253
D15	0.9977 ±0.0022	0.9973 ±0.0023	**0.9978** ±0.002	0.9976 ±0.0021	0.9975 ±0.002
D16	**0.5588** ±0.0206	0.5559 ±0.0232	0.5587 ±0.0247	0.5541 ±0.0245	0.5557 ±0.0236
D17	0.9169 ±0.0071	0.9163 ±0.0091	0.9164 ±0.0081	**0.9173** ±0.0081	0.9163 ±0.0089
D18	0.7357 ±0.042	0.734 ±0.0441	0.7369 ±0.0412	**0.7396** ±0.0392	0.7352 ±0.0432
D19	0.9196 ±0.0259	0.9227 ±0.0229	**0.924** ±0.0214	0.923 ±0.0225	0.9233 ±0.0237
D20	0.5957 ±0.01	**0.5985** ±0.0105	0.596 ±0.0116	0.5973 ±0.01	0.5984 ±0.0124
D21	0.9624 ±0.0205	0.9644 ±0.0204	**0.9664** ±0.0157	0.9644 ±0.0153	0.9644 ±0.0169
D22	0.9914 ±0.0061	0.991 ±0.0069	**0.9915** ±0.0064	0.9912 ±0.0065	0.9914 ±0.0062
D23	0.9948	0.9954	0.9958	0.996	**0.9962**

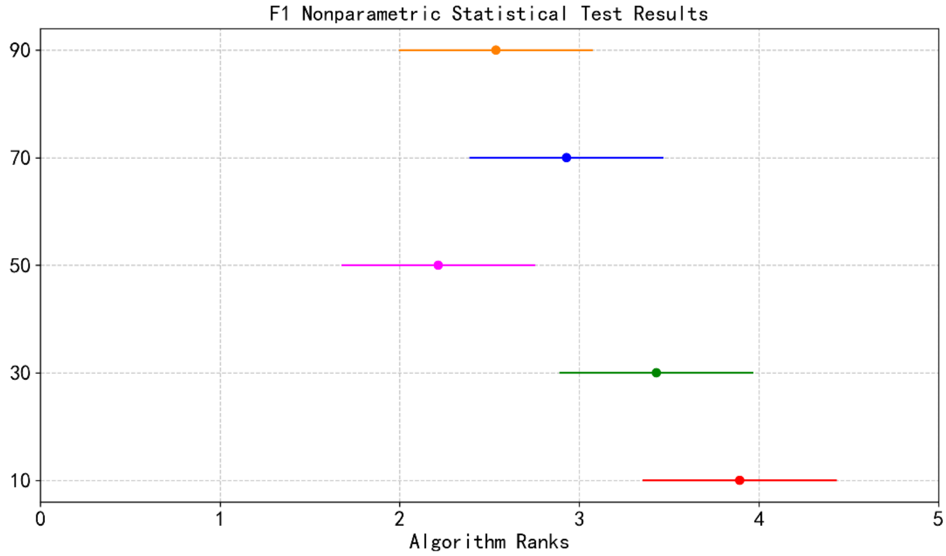
	±0.0017	±0.0016	±0.0013	±0.0015	±0.0012
D27	<u>0.8963</u> ±0.0068	0.8954 ±0.007	0.8959 ±0.0073	0.8956 ±0.0074	**0.8964** ±0.0069
D28	0.8936 ±0.0073	0.8935 ±0.0068	**0.8945** ±0.0084	<u>0.8937</u> ±0.008	0.8936 ±0.0077
D29	**0.9926** ±0.0022	**0.9926** ±0.0019	**0.9926** ±0.0021	<u>0.9925</u> ±0.0019	**0.9926** ±0.0021
D31	0.863 ±0.0165	0.8678 ±0.0203	0.8732 ±0.0181	<u>0.8736</u> ±0.0213	**0.874** ±0.0207
D32	**0.6565** ±0.0103	<u>0.655</u> ±0.0108	0.6549 ±0.0081	0.6545 ±0.014	0.6528 ±0.0121
D34	<u>0.9991</u> ±0.0013	<u>0.9991</u> ±0.0013	**0.9992** ±0.0013	**0.9992** ±0.0012	0.9989 ±0.0014
D35	0.6277 ±0.02	<u>0.6311</u> ±0.017	**0.6313** ±0.0157	0.6295 ±0.0169	0.6307 ±0.0176

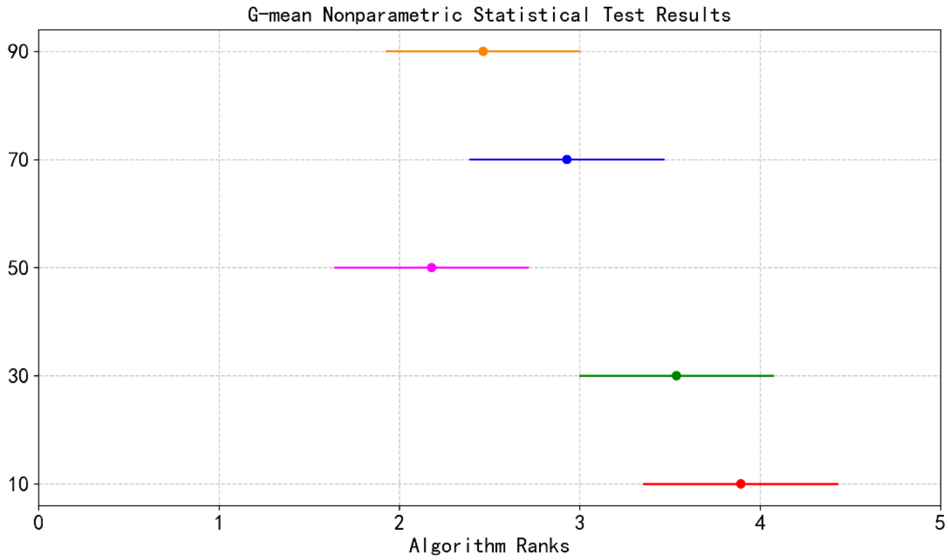
Table S42 MCC results of the different *pop* parameter values

Dataset	*pop*=10	*pop* =30	*pop* =50	*pop* =70	*pop* =90
D2	0.8608 ±0.0357	**0.864** ±0.0323	<u>0.8631</u> ±0.0334	0.8617 ±0.0333	0.8619 ±0.0313
D3	0.9082 ±0.009	0.909 ±0.008	**0.9099** ±0.0095	0.9087 ±0.0087	<u>0.9093</u> ±0.0084
D4	0.7616 ±0.0735	0.7612 ±0.0775	**0.7691** ±0.085	0.7612 ±0.0784	<u>0.7655</u> ±0.0733
D6	0.9406 ±0.0444	0.9446 ±0.0426	<u>0.9472</u> ±0.0388	**0.9481** ±0.0396	0.9399 ±0.045
D7	0.9079 ±0.035	0.934 ±0.0286	0.945 ±0.0229	<u>0.9533</u> ±0.0268	**0.9587** ±0.0255
D8	0.5638 ±0.0799	<u>0.5776</u> ±0.0903	0.5765 ±0.0907	0.5688 ±0.1031	**0.5809** ±0.1079
D9	0.965 ±0.0393	0.9636 ±0.0393	<u>0.9741</u> ±0.0389	0.9711 ±0.033	**0.9754** ±0.0318
D10	0.7317 ±0.0368	<u>0.7344</u> ±0.0389	**0.7411** ±0.0442	0.7332 ±0.0451	0.7292 ±0.0475
D11	0.97 ±0.0102	0.9707 ±0.0098	<u>0.9709</u> ±0.0107	**0.9716** ±0.0103	<u>0.9709</u> ±0.0107
D12	0.7362 ±0.0341	**0.741** ±0.0349	0.7388 ±0.0342	0.7397 ±0.0335	<u>0.7408</u> ±0.0346
D13	0.3579 ±0.0775	**0.3629** ±0.0726	0.3483 ±0.0699	<u>0.3613</u> ±0.0737	0.3605 ±0.0773
D14	0.84 ±0.0328	**0.8414** ±0.03	<u>0.8408</u> ±0.0329	0.8377 ±0.0295	0.8398 ±0.032
D15	**0.9972**	<u>0.997</u>	**0.9972**	**0.9972**	**0.9972**

	±0.0022	±0.0021	±0.002	±0.0021	±0.0021
D16	**0.5029**	0.4976	<u>0.501</u>	0.4982	0.5004
	±0.054	±0.0557	±0.0613	±0.0556	±0.0586
D17	0.9147	0.9148	**0.9151**	<u>0.915</u>	0.9142
	±0.0052	±0.0064	±0.0052	±0.0055	±0.0057
D18	<u>0.5016</u>	0.4984	0.4998	**0.502**	0.501
	±0.0322	±0.0327	±0.0318	±0.0302	±0.0319
D19	0.8732	0.8746	**0.8772**	0.8756	<u>0.8771</u>
	±0.0257	±0.0254	±0.0239	±0.0233	±0.0251
D20	<u>0.5178</u>	0.5165	0.5146	0.5163	**0.5181**
	±0.0213	±0.0224	±0.02	±0.0201	±0.022
D21	0.928	<u>0.9317</u>	**0.9351**	0.9313	0.9316
	±0.0375	±0.0367	±0.0301	±0.0294	±0.0318
D22	0.9829	0.9823	**0.9833**	0.9827	<u>0.983</u>
	±0.012	±0.0136	±0.0126	±0.0127	±0.0121
D23	0.9896	0.9908	0.9916	<u>0.992</u>	**0.9924**
	±0.0033	±0.0032	±0.0025	±0.0029	±0.0023
D27	<u>0.795</u>	0.7932	0.7943	0.7936	**0.7952**
	±0.013	±0.0135	±0.0142	±0.0143	±0.0134
D28	<u>0.7903</u>	0.7901	**0.7919**	0.7902	0.7902
	±0.014	±0.013	±0.0161	±0.0156	±0.0148
D29	**0.9852**	**0.9852**	**0.9852**	<u>0.9851</u>	**0.9852**
	±0.0043	±0.0039	±0.0042	±0.0038	±0.0042
D31	0.7301	0.7392	0.7499	<u>0.7509</u>	**0.7516**
	±0.032	±0.0396	±0.0356	±0.0416	±0.0403
D32	**0.3318**	0.3288	0.3289	<u>0.329</u>	0.325
	±0.0197	±0.0202	±0.0154	±0.0266	±0.0233
D34	0.9981	0.998	<u>0.9983</u>	**0.9984**	0.9976
	±0.0028	±0.0028	±0.0027	±0.0024	±0.0031
D35	0.2872	<u>0.2925</u>	0.2921	0.2896	**0.2926**
	±0.0369	±0.0323	±0.0287	±0.0316	±0.0313

The Nonparametric Statistical Test was performed on the experimental data, and the results are shown in Fig S5. Results show the best overall performance occurs when population size exceeds 50. When pop is set too low, the population tends to converge rapidly to local optima, making it difficult to generate high-quality features. Moreover, using smaller populations, mutations or crossover operations may excessively influence the evolutionary trajectory, resulting in unstable outcomes. When pop is set too high, each generation requires evaluating a large number of individuals, resulting in excessively long training times. Therefore, it is generally recommended to set pop to a necessary yet moderate value.





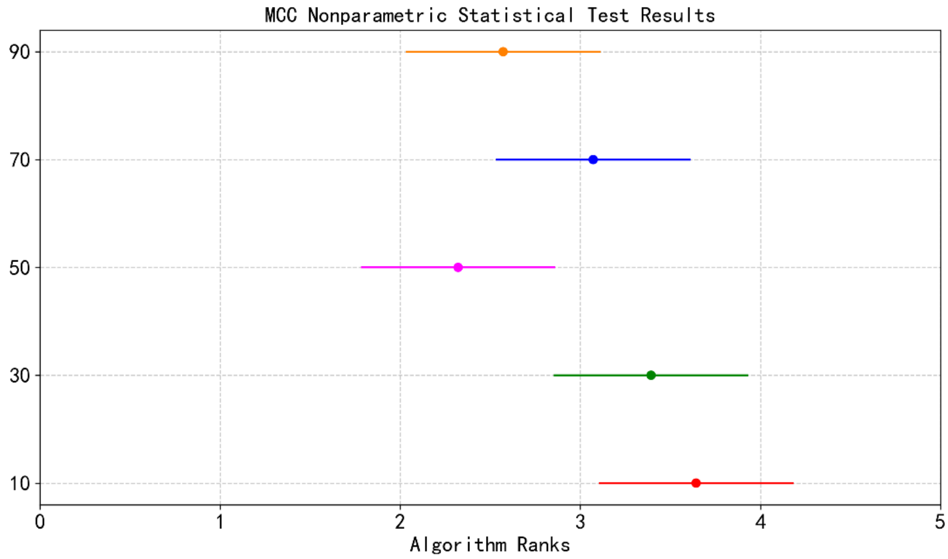


Fig S5 Nonparametric statistical test figure for results of different *pop* values.

The parameter max_depth controls the maximum depth of individual decision trees within the MDGP-Forest's decision forest module, defined as the number of nodes along the longest path from the root node to a leaf node. When set to None, the trees will grow recursively until all leaves become pure. MDGP-Forest is configured with five different max_depth values $\{\text{None}, 5, 10, 20, 30\}$, and the experimental results are presented in Table S43, Table S44, and Table S45.

Table S43 F1 results of the different *max_depth* parameter values

Dataset	*max_depth* =None	*max_depth* =5	*max_depth* =10	*max_depth* =20	*max_depth* =30
D2	0.8964 ±0.0242	0.8949 ±0.0275	0.8953 ±0.0255	**0.8968** ±0.0249	**0.8968** ±0.0256
D3	**0.9104** ±0.0091	0.8619 ±0.0108	0.9021 ±0.0096	0.9089 ±0.009	0.9092 ±0.0084
D4	**0.8128** ±0.0707	0.8102 ±0.0609	0.8031 ±0.0624	0.8095 ±0.0607	0.8076 ±0.0664
D6	0.9654 ±0.0265	0.9642 ±0.0281	**0.9664** ±0.0255	0.9657 ±0.0255	0.9647 ±0.0262
D7	0.9392 ±0.0282	0.9291 ±0.0387	0.9328 ±0.0414	**0.9416** ±0.029	0.9365 ±0.0292
D8	0.6823 ±0.1001	0.6553 ±0.0921	**0.6864** ±0.0988	0.681 ±0.1043	0.6851 ±0.1043
D9	**0.9505** ±0.074	0.9293 ±0.0827	0.9447 ±0.0768	0.9479 ±0.0755	0.9391 ±0.0794
D10	**0.7636** ±0.0552	0.6591 ±0.0782	0.731 ±0.0769	0.7606 ±0.0604	0.7534 ±0.0609
D11	0.9812 ±0.0072	0.9784 ±0.0093	**0.9817** ±0.0083	0.9813 ±0.0063	0.98 ±0.0073
D12	0.8149 ±0.0355	0.7229 ±0.0351	0.8074 ±0.0369	0.8149 ±0.035	**0.8172** ±0.0365
D13	0.341 ±0.0778	0.3401 ±0.0779	**0.3625** ±0.0743	0.3531 ±0.0884	0.3499 ±0.0765
D14	0.8631 ±0.0472	0.8388 ±0.0752	**0.8694** ±0.0466	0.8605 ±0.0521	0.8424 ±0.0724
D15	0.9972 ±0.0022	0.9609 ±0.0092	**0.9973** ±0.0021	0.9969 ±0.0024	0.9967 ±0.0025
D16	**0.3457** ±0.0272	0.2893 ±0.0175	0.3242 ±0.0245	0.3369 ±0.0263	0.342 ±0.0259
D17	0.8826 ±0.0129	0.7623 ±0.0247	0.8666 ±0.0138	0.882 ±0.0137	**0.8837** ±0.0118
D18	0.5835 ±0.0614	0.551 ±0.0391	**0.5837** ±0.0636	0.583 ±0.0641	0.583 ±0.0611
D19	**0.8834**	0.822	0.8825	0.8804	0.8825

	±0.0368	±0.0551	±0.0386	±0.0429	±0.037
D20	0.4221 ±0.0159	0.2613 ±0.0155	0.3606 ±0.0185	<u>0.4222</u> ±0.015	**0.4248** ±0.0167
D21	**0.9551** ±0.0211	0.9542 ±0.0222	**0.9551** ±0.0211	0.9531 ±0.0204	<u>0.9544</u> ±0.0243
D22	**0.9846** ±0.0118	0.879 ±0.0387	0.9785 ±0.0123	0.9833 ±0.0117	<u>0.9836</u> ±0.0118
D23	**0.9924** ±0.0023	0.9659 ±0.006	0.9848 ±0.0044	**0.9924** ±0.0023	<u>0.9923</u> ±0.002
D27	0.8612 ±0.01	0.8612 ±0.0067	**0.8631** ±0.0078	<u>0.8625</u> ±0.0089	0.8604 ±0.0096
D28	<u>0.8596</u> ±0.0112	0.858 ±0.0111	**0.8597** ±0.0113	0.8583 ±0.0101	0.8581 ±0.0106
D29	**0.9867** ±0.0038	0.9592 ±0.0047	0.9832 ±0.0035	<u>0.9866</u> ±0.0036	<u>0.9866</u> ±0.0037
D31	**0.8103** ±0.0261	0.7777 ±0.0324	0.7984 ±0.0282	<u>0.8055</u> ±0.0309	0.8037 ±0.0267
D32	0.5489 ±0.009	<u>0.5529</u> ±0.013	**0.5532** ±0.0143	0.5505 ±0.0128	0.5504 ±0.0115
D34	<u>0.9986</u> ±0.0023	0.9677 ±0.0121	0.996 ±0.0034	**0.9989** ±0.0019	0.9983 ±0.0023
D35	0.5227 ±0.0214	**0.5278** ±0.0235	<u>0.5271</u> ±0.0217	0.5198 ±0.0228	0.5238 ±0.0232

Table S44 G-mean results of the different *max_depth* parameter values

Dataset	*max_depth* =None	*max_depth* =5	*max_depth* =10	*max_depth* =20	*max_depth* =30
D2	0.9277 ±0.0168	0.9258 ±0.0197	0.9268 ±0.0178	<u>0.9279</u> ±0.0172	**0.928** ±0.0174
D3	**0.9441** ±0.0057	0.9151 ±0.0068	0.9391 ±0.0059	0.9432 ±0.0057	<u>0.9433</u> ±0.0051
D4	**0.8706** ±0.0502	<u>0.869</u> ±0.0437	0.8651 ±0.0445	0.8684 ±0.0431	0.868 ±0.0475
D6	<u>0.9712</u> ±0.0299	0.97 ±0.028	**0.9734** ±0.0267	0.9708 ±0.0303	0.9695 ±0.0288
D7	<u>0.9605</u> ±0.0203	0.951 ±0.0287	0.9565 ±0.0263	**0.961** ±0.0196	0.9579 ±0.0176
D8	0.7465 ±0.0694	0.725 ±0.0572	<u>0.7474</u> ±0.0691	0.7451 ±0.0717	**0.7479** ±0.072
D9	**0.9769** ±0.0347	0.9674 ±0.0387	0.9735 ±0.0367	<u>0.9757</u> ±0.0353	0.9704 ±0.038
D10	**0.8482** ±0.0397	0.7891 ±0.0555	0.8295 ±0.0506	<u>0.8456</u> ±0.0421	0.8433 ±0.0436

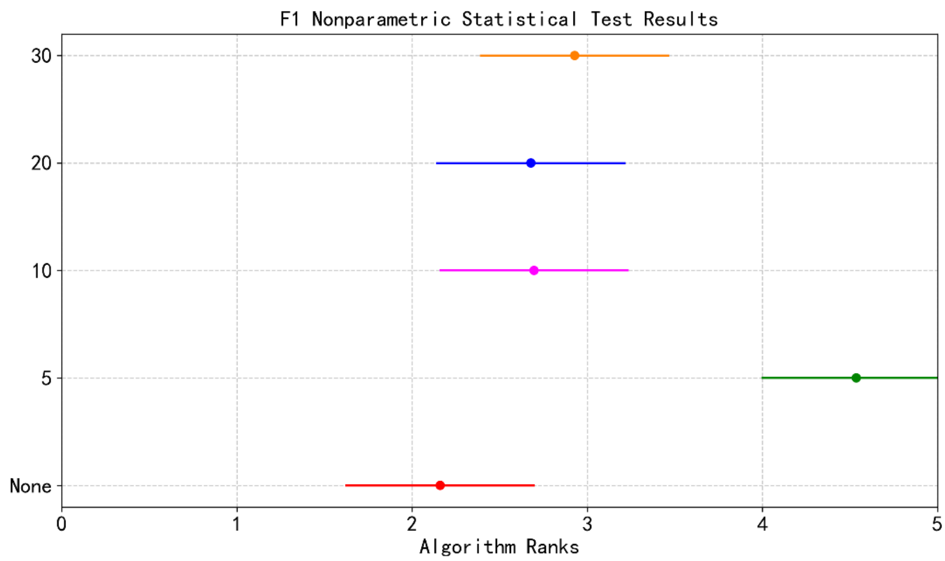
D11	0.9833 ±0.0066	0.9793 ±0.0091	**0.984** ±0.0072	0.9835 ±0.0054	0.9824 ±0.0064
D12	0.879 ±0.0256	0.7994 ±0.0235	0.8704 ±0.0275	0.8791 ±0.0252	**0.8812** ±0.0261
D13	0.5459 ±0.0572	0.5467 ±0.0585	**0.561** ±0.0602	0.5561 ±0.0649	0.5524 ±0.0563
D14	0.9134 ±0.0231	0.899 ±0.0413	**0.9185** ±0.0228	0.9125 ±0.0268	0.9031 ±0.0365
D15	0.9978 ±0.002	0.9717 ±0.0081	**0.9981** ±0.0016	0.9977 ±0.0021	0.9973 ±0.0022
D16	**0.5587** ±0.0247	0.5092 ±0.0166	0.5432 ±0.0236	0.5546 ±0.026	0.5559 ±0.0233
D17	0.9164 ±0.0081	0.833 ±0.0113	0.9028 ±0.0081	0.9162 ±0.0088	**0.9172** ±0.0074
D18	**0.7369** ±0.0412	0.7073 ±0.028	0.7349 ±0.0424	0.7363 ±0.0446	0.736 ±0.0421
D19	0.924 ±0.0214	0.8728 ±0.031	**0.9253** ±0.022	0.9219 ±0.0258	0.9239 ±0.0223
D20	0.596 ±0.0116	0.4789 ±0.0102	0.55 ±0.0134	0.5957 ±0.0107	**0.5977** ±0.0118
D21	**0.9664** ±0.0157	0.9658 ±0.0165	**0.9664** ±0.0157	0.9649 ±0.0152	0.9659 ±0.0181
D22	**0.9915** ±0.0064	0.9334 ±0.0212	0.9882 ±0.0067	0.9908 ±0.0064	0.991 ±0.0064
D23	**0.9958** ±0.0013	0.981 ±0.0033	0.9916 ±0.0025	**0.9958** ±0.0013	**0.9958** ±0.0011
D27	0.8959 ±0.0073	0.8959 ±0.005	**0.8972** ±0.0058	0.8968 ±0.0065	0.8952 ±0.0071
D28	0.8945 ±0.0084	0.8933 ±0.0084	**0.8946** ±0.0085	0.8936 ±0.0076	0.8933 ±0.008
D29	**0.9926** ±0.0021	0.977 ±0.0027	0.9906 ±0.002	0.9925 ±0.0021	0.9925 ±0.0021
D31	**0.8732** ±0.0181	0.8519 ±0.0229	0.8656 ±0.0195	0.8698 ±0.0211	0.8684 ±0.0184
D32	0.6549 ±0.0081	**0.6632** ±0.0114	0.6607 ±0.0121	0.6566 ±0.0106	0.6559 ±0.0098
D34	0.9992 ±0.0013	0.9812 ±0.0069	0.9977 ±0.002	**0.9994** ±0.0011	0.999 ±0.0013
D35	0.6313 ±0.0157	**0.6366** ±0.0202	0.6353 ±0.0179	0.629 ±0.0172	0.6325 ±0.0181

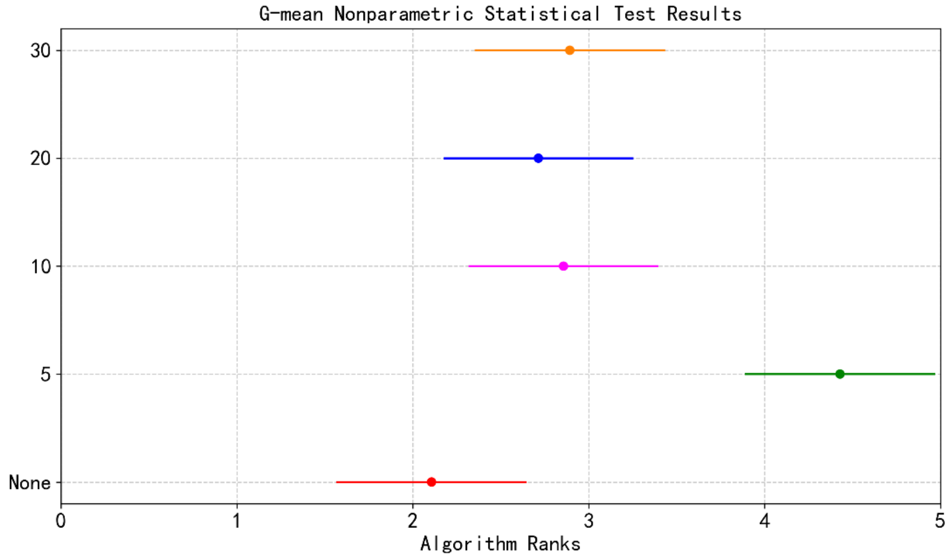
Table S45 MCC results of the different *max_depth* parameter values

Dataset	*max_depth* =None	*max_depth* =5	*max_depth* =10	*max_depth* =20	*max_depth* =30
D2	0.8631 ±0.0334	0.8602 ±0.0393	0.8619 ±0.0343	**0.8642** ±0.034	<u>0.864</u> ±0.0346
D3	**0.9099** ±0.0095	0.8638 ±0.0114	0.9017 ±0.0096	0.9083 ±0.0092	<u>0.9085</u> ±0.0087
D4	**0.7691** ±0.085	<u>0.7665</u> ±0.0749	0.7588 ±0.0741	0.7638 ±0.074	0.7613 ±0.0798
D6	<u>0.9472</u> ±0.0388	0.9449 ±0.0422	**0.9482** ±0.0381	0.9471 ±0.0376	0.9459 ±0.0383
D7	<u>0.945</u> ±0.0229	0.9436 ±0.0267	0.9389 ±0.0351	**0.9453** ±0.0252	0.9413 ±0.0246
D8	0.5765 ±0.0907	0.5678 ±0.0917	**0.5785** ±0.0965	0.5733 ±0.0968	<u>0.5769</u> ±0.0973
D9	<u>0.9741</u> ±0.0389	0.9686 ±0.0353	0.9727 ±0.0372	**0.9752** ±0.0344	0.9724 ±0.0328
D10	<u>0.7411</u> ±0.0442	0.6558 ±0.0718	0.7233 ±0.0533	**0.7418** ±0.0469	0.7291 ±0.0459
D11	0.9709 ±0.0107	0.965 ±0.0156	**0.9721** ±0.0122	<u>0.9712</u> ±0.0091	0.9686 ±0.011
D12	0.7388 ±0.0342	0.6539 ±0.0303	0.7273 ±0.036	<u>0.7398</u> ±0.0343	**0.7425** ±0.0345
D13	0.3483 ±0.0699	<u>0.3668</u> ±0.0788	**0.3691** ±0.0775	0.3647 ±0.0867	0.361 ±0.0706
D14	0.8408 ±0.0329	<u>0.8413</u> ±0.037	**0.8492** ±0.0331	0.8405 ±0.0289	0.8359 ±0.0312
D15	0.9972 ±0.002	0.9748 ±0.0063	**0.9975** ±0.0017	<u>0.9973</u> ±0.0018	0.9969 ±0.0022
D16	<u>0.501</u> ±0.0613	0.3857 ±0.0447	0.4808 ±0.0533	**0.5047** ±0.0576	0.4951 ±0.057
D17	**0.9151** ±0.0052	0.8277 ±0.0074	0.8963 ±0.005	<u>0.9138</u> ±0.0054	**0.9151** ±0.0054
D18	0.4998 ±0.0318	0.4792 ±0.0278	<u>0.5033</u> ±0.0273	**0.5045** ±0.0307	0.5023 ±0.0327
D19	<u>0.8772</u> ±0.0239	0.864 ±0.0349	**0.8781** ±0.0265	0.875 ±0.027	0.8761 ±0.0243
D20	0.5146 ±0.02	0.3442 ±0.0203	0.448 ±0.0221	**0.5196** ±0.0188	<u>0.5161</u> ±0.0221
D21	**0.9351** ±0.0301	0.9341 ±0.0313	**0.9351** ±0.0301	0.9322 ±0.0293	<u>0.9343</u> ±0.0348
D22	**0.9833** ±0.0126	0.8718 ±0.0397	0.9768 ±0.013	0.9819 ±0.0125	<u>0.9823</u> ±0.0126

D23	**0.9916** ±0.0025	0.9624 ±0.0065	0.9833 ±0.0049	**0.9916** ±0.0025	<u>0.9915</u> ±0.0022
D27	0.7943 ±0.0142	0.7943 ±0.0101	**0.7967** ±0.0113	<u>0.796</u> ±0.0125	0.793 ±0.0137
D28	<u>0.7919</u> ±0.0161	0.7894 ±0.0165	**0.7921** ±0.0164	0.7901 ±0.0145	0.7895 ±0.0155
D29	**0.9852** ±0.0042	0.9544 ±0.0053	0.9813 ±0.0039	<u>0.9851</u> ±0.0041	<u>0.9851</u> ±0.0042
D31	**0.7499** ±0.0356	0.712 ±0.0452	0.7356 ±0.0381	<u>0.7437</u> ±0.041	0.7404 ±0.0361
D32	0.3289 ±0.0154	**0.3544** ±0.0227	<u>0.3439</u> ±0.0235	0.3328 ±0.0202	0.3308 ±0.0191
D34	<u>0.9983</u> ±0.0027	0.9624 ±0.0146	0.9951 ±0.0041	**0.9986** ±0.0023	0.998 ±0.0028
D35	0.2921 ±0.0287	**0.3087** ±0.0371	<u>0.3017</u> ±0.0337	0.2884 ±0.0325	0.2957 ±0.0363

The Nonparametric Statistical Test was performed on the experimental data, and the results are shown in Fig S6. It can be observed that when max_depth is set to 5, the overall performance is significantly lower than in other cases, while setting it to None yields the best results. When max_depth is set to a smaller value, the growth of the trees is strictly constrained, leading to an overly simplistic model structure. In such cases, the decision forest module may under-fitting, failing to learn sufficiently complex patterns, which results in a notable decline in performance. Therefore, it is recommended setting $max\_depth = None$ for optimal MDGP-Forest performance.





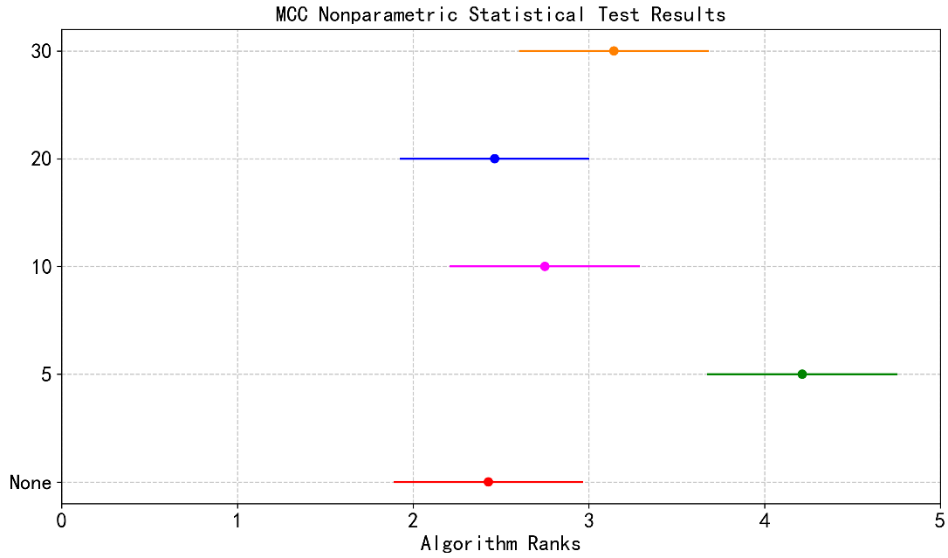


Fig S6 Nonparametric statistical test figure for results of different max_depth values.

The parameter $min\_samples\_split$ determines the minimum number of samples required to split an internal node in the decision trees of MDGP-Forest's decision forest module. When set to relatively small values, this hyperparameter imposes more lenient splitting conditions, permitting node division even when only a limited number of samples are present. MDGP-Forest is configured with five different $min\_samples\_split$ values {2,5,10,15,20}, and the experimental results are presented in Table S46, Table S47, and Table S48.

Table S46 F1 results of the different *min_samles_split* parameter values

Dataset	*min_samles_split* =2	*min_samles_split* =5	*min_samles_split* =10	*min_samles_split* =15	*min_samles_split* =20
D2	**0.8964** ±0.0242	0.8959 ±0.0289	0.8957 ±0.0259	<u>0.8963</u> ±0.0269	0.8948 ±0.0255

D3	**0.9104** ±0.0091	<u>0.9081</u> ±0.009	0.9048 ±0.0097	0.9031 ±0.009	0.9009 ±0.0095
D4	0.8128 ±0.0707	0.8058 ±0.0665	<u>0.814</u> ±0.0625	0.8089 ±0.0647	**0.816** ±0.0611
D6	<u>0.9654</u> ±0.0265	**0.9679** ±0.0254	0.9631 ±0.0226	0.9602 ±0.0245	0.9555 ±0.0273
D7	**0.9392** ±0.0282	0.9233 ±0.0392	0.9332 ±0.0301	<u>0.9382</u> ±0.0303	0.9294 ±0.0361
D8	0.6823 ±0.1001	**0.6882** ±0.1026	<u>0.6866</u> ±0.1091	0.6794 ±0.1	0.6744 ±0.0953
D9	<u>0.9505</u> ±0.074	**0.9581** ±0.0663	0.897 ±0.0812	0.7779 ±0.097	0.7192 ±0.0907
D10	**0.7636** ±0.0552	<u>0.7318</u> ±0.0812	0.688 ±0.0966	0.6482 ±0.1021	0.5759 ±0.1213
D11	0.9812 ±0.0072	0.9815 ±0.008	0.9821 ±0.0085	**0.9825** ±0.0084	<u>0.9822</u> ±0.0084
D12	<u>0.8149</u> ±0.0355	**0.8154** ±0.0395	0.813 ±0.0405	0.8088 ±0.0407	0.8058 ±0.0425
D13	0.341 ±0.0778	**0.3498** ±0.0786	<u>0.3424</u> ±0.0666	0.3277 ±0.066	0.3131 ±0.0691
D14	**0.8631** ±0.0472	<u>0.8569</u> ±0.0609	0.8426 ±0.0736	0.8453 ±0.0673	0.8007 ±0.0733
D15	<u>0.9972</u> ±0.0022	**0.9977** ±0.0019	0.9971 ±0.0022	0.9964 ±0.0024	0.9956 ±0.0027
D16	**0.3457** ±0.0272	<u>0.3385</u> ±0.0255	0.3268 ±0.0254	0.3171 ±0.0217	0.3137 ±0.0233
D17	**0.8826** ±0.0129	<u>0.8807</u> ±0.0143	0.8765 ±0.0133	0.8737 ±0.0171	0.8682 ±0.0172
D18	**0.5835** ±0.0614	<u>0.5821</u> ±0.0615	0.5809 ±0.0613	0.573 ±0.0618	0.5686 ±0.0551
D19	**0.8834** ±0.0368	<u>0.8811</u> ±0.0403	0.8764 ±0.04	0.8734 ±0.0343	0.8715 ±0.031
D20	**0.4221** ±0.0159	<u>0.412</u> ±0.014	0.3982 ±0.0164	0.3786 ±0.0142	0.3626 ±0.0147
D21	**0.9551** ±0.0211	0.9538 ±0.0165	0.9538 ±0.0165	<u>0.9545</u> ±0.019	0.951 ±0.0219
D22	**0.9846** ±0.0118	<u>0.9802</u> ±0.0108	0.9712 ±0.0183	0.9623 ±0.0187	0.9498 ±0.0202
D23	**0.9924** ±0.0023	<u>0.9916</u> ±0.0024	0.9891 ±0.0034	0.9884 ±0.0036	0.9871 ±0.0035
D27	0.8612 ±0.01	0.8626 ±0.0092	<u>0.8632</u> ±0.0093	0.8629 ±0.009	**0.8633** ±0.0091
D28	0.8596	0.8581	0.858	<u>0.8598</u>	**0.8602**

	±0.0112	±0.0101	±0.0111	±0.0108	±0.0112
D29	**0.9867** ±0.0038	0.9859 ±0.0037	0.9846 ±0.0033	0.9835 ±0.0037	0.9827 ±0.0035
D31	**0.8103** ±0.0261	0.8001 ±0.0275	0.7968 ±0.0286	0.7907 ±0.0289	0.7859 ±0.0305
D32	0.5489 ±0.009	0.5493 ±0.0147	0.5523 ±0.0131	0.5543 ±0.0175	**0.5556** ±0.0145
D34	**0.9986** ±0.0023	0.9973 ±0.0028	0.9952 ±0.0041	0.9927 ±0.0045	0.9896 ±0.0057
D35	0.5227 ±0.0214	0.5211 ±0.0231	0.5237 ±0.0211	0.5284 ±0.0215	**0.5345** ±0.0175

Table S47 G-mean results of the different *min_samles_split* parameter values

Dataset	*min_samles_split* =2	*min_samles_split* =5	*min_samles_split* =10	*min_samles_split* =15	*min_samles_split* =20
D2	**0.9277** ±0.0168	0.9273 ±0.0199	0.9274 ±0.018	0.9275 ±0.0186	0.9266 ±0.0182
D3	**0.9441** ±0.0057	0.9427 ±0.0056	0.9408 ±0.0058	0.9399 ±0.0055	0.9385 ±0.0057
D4	0.8706 ±0.0502	0.8664 ±0.048	0.8717 ±0.0447	0.868 ±0.0467	**0.8731** ±0.044
D6	0.9712 ±0.0299	**0.975** ±0.025	0.9701 ±0.0236	0.9701 ±0.0207	0.9675 ±0.0242
D7	**0.9605** ±0.0203	0.9466 ±0.0269	0.9507 ±0.024	0.9544 ±0.0217	0.9493 ±0.0253
D8	0.7465 ±0.0694	0.7498 ±0.0709	**0.7506** ±0.0736	0.7442 ±0.067	0.7405 ±0.0636
D9	0.9769 ±0.0347	**0.9805** ±0.031	0.9508 ±0.0377	0.8917 ±0.05	0.8582 ±0.0472
D10	**0.8482** ±0.0397	0.8302 ±0.0528	0.7995 ±0.0601	0.7758 ±0.0683	0.7299 ±0.084
D11	0.9833 ±0.0066	**0.9842** ±0.0067	0.9838 ±0.0073	0.9839 ±0.0069	0.9838 ±0.007
D12	**0.879** ±0.0256	0.8788 ±0.0281	0.8762 ±0.0294	0.8723 ±0.03	0.8703 ±0.0309
D13	0.5459 ±0.0572	**0.5498** ±0.0548	0.5486 ±0.0502	0.5384 ±0.0499	0.5294 ±0.0555
D14	**0.9134** ±0.0231	0.9102 ±0.031	0.9036 ±0.0384	0.9031 ±0.0358	0.8777 ±0.0375
D15	0.9978 ±0.002	**0.9985** ±0.0015	0.9982 ±0.0016	0.9982 ±0.0013	0.9976 ±0.0019
D16	**0.5587** ±0.0247	0.5547 ±0.0242	0.5441 ±0.0233	0.537 ±0.0213	0.5334 ±0.0218

D17	**0.9164** ±0.0081	<u>0.9144</u> ±0.0091	0.9115 ±0.0079	0.9093 ±0.0105	0.9058 ±0.0104
D18	**0.7369** ±0.0412	0.7352 ±0.0418	<u>0.736</u> ±0.0419	0.7324 ±0.0419	0.7304 ±0.0376
D19	<u>0.924</u> ±0.0214	0.9211 ±0.0253	0.9188 ±0.0264	0.921 ±0.0225	**0.9242** ±0.0197
D20	**0.596** ±0.0116	<u>0.5889</u> ±0.0099	0.5783 ±0.0124	0.5642 ±0.011	0.5532 ±0.0118
D21	**0.9664** ±0.0157	0.9654 ±0.0123	0.9654 ±0.0123	<u>0.9659</u> ±0.0141	0.9634 ±0.0162
D22	**0.9915** ±0.0064	<u>0.9891</u> ±0.0059	0.984 ±0.0102	0.9792 ±0.0104	0.9722 ±0.0113
D23	**0.9958** ±0.0013	<u>0.9954</u> ±0.0013	0.994 ±0.0019	0.9936 ±0.002	0.9929 ±0.002
D27	0.8959 ±0.0073	0.8969 ±0.0068	**0.8974** ±0.0068	<u>0.8971</u> ±0.0066	**0.8974** ±0.0067
D28	0.8945 ±0.0084	0.8934 ±0.0076	0.8934 ±0.0084	<u>0.8947</u> ±0.008	**0.895** ±0.0084
D29	**0.9926** ±0.0021	<u>0.9922</u> ±0.0021	0.9914 ±0.0019	0.9908 ±0.0021	0.9903 ±0.002
D31	**0.8732** ±0.0181	<u>0.8665</u> ±0.0189	0.8647 ±0.0197	0.861 ±0.0197	0.858 ±0.0208
D32	0.6549 ±0.0081	0.6556 ±0.0127	0.6579 ±0.0113	<u>0.6598</u> ±0.0148	**0.661** ±0.0126
D34	**0.9992** ±0.0013	<u>0.9984</u> ±0.0016	0.9973 ±0.0023	0.9958 ±0.0025	0.9941 ±0.0032
D35	0.6313 ±0.0157	0.6298 ±0.019	0.632 ±0.0172	<u>0.6362</u> ±0.0174	**0.6409** ±0.0144

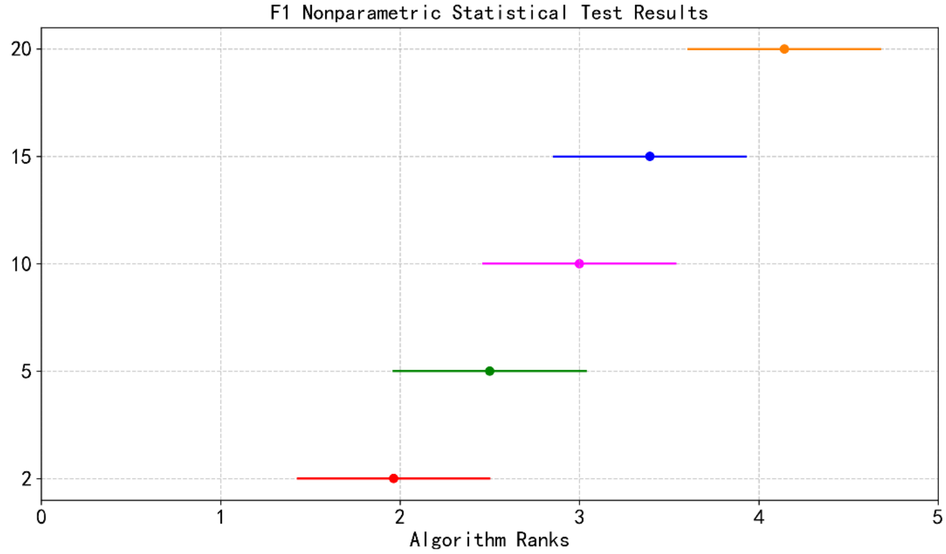
Table S48 MCC results of the different *min_samles_split* parameter values

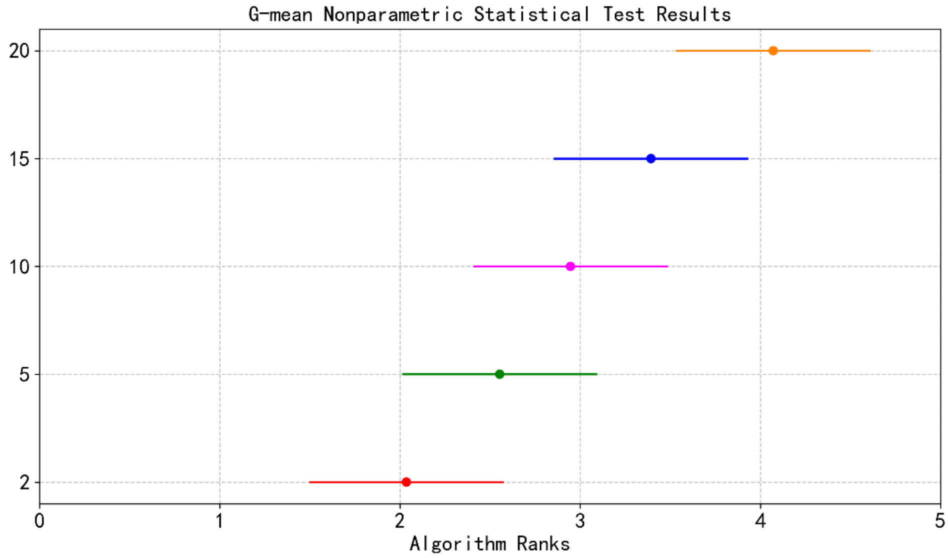
Dataset	*min_samles_split* =2	*min_samles_split* =5	*min_samles_split* =10	*min_samles_split* =15	*min_samles_split* =20
D2	**0.8631** ±0.0334	<u>0.8627</u> ±0.0392	0.862 ±0.0365	0.8623 ±0.0384	0.8604 ±0.0375
D3	**0.9099** ±0.0095	<u>0.9076</u> ±0.0094	0.9045 ±0.01	0.9026 ±0.0092	0.9005 ±0.0092
D4	0.7691 ±0.085	0.7622 ±0.0818	<u>0.7718</u> ±0.0762	0.7643 ±0.0803	**0.772** ±0.0761
D6	<u>0.9472</u> ±0.0388	**0.9501** ±0.0382	0.9432 ±0.0339	0.9378 ±0.0381	0.9301 ±0.0432
D7	<u>0.945</u> ±0.0229	0.9355 ±0.0279	0.9435 ±0.0203	**0.9467** ±0.0227	0.944 ±0.0238
D8	0.5765	<u>0.5796</u>	**0.5879**	0.5777	0.5751

	±0.0907	±0.0949	±0.0994	±0.0904	±0.0911
D9	<u>0.9741</u>	**0.9791**	0.9512	0.9037	0.8687
	±0.0389	±0.0308	±0.0345	±0.0549	±0.0607
D10	**0.7411**	<u>0.725</u>	0.6846	0.6616	0.6181
	±0.0442	±0.0572	±0.0674	±0.0668	±0.0864
D11	0.9709	0.9721	<u>0.9729</u>	**0.973**	0.9727
	±0.0107	±0.0118	±0.0124	±0.0128	±0.013
D12	<u>0.7388</u>	**0.7396**	0.7357	0.7296	0.7277
	±0.0342	±0.0376	±0.0382	±0.039	±0.0425
D13	0.3483	**0.3581**	<u>0.3579</u>	0.3549	0.3467
	±0.0699	±0.0662	±0.0636	±0.0744	±0.0834
D14	0.8408	0.8442	<u>0.8452</u>	**0.8519**	0.8345
	±0.0329	±0.0297	±0.0366	±0.0338	±0.027
D15	0.9972	**0.998**	<u>0.9976</u>	<u>0.9976</u>	0.9969
	±0.002	±0.0017	±0.0016	±0.0014	±0.0018
D16	**0.501**	<u>0.4985</u>	0.4755	0.4594	0.4505
	±0.0613	±0.0561	±0.0514	±0.0473	±0.0493
D17	**0.9151**	<u>0.9118</u>	0.9065	0.9033	0.8975
	±0.0052	±0.0056	±0.0048	±0.0058	±0.0065
D18	0.4998	0.5016	<u>0.5037</u>	**0.5045**	0.5031
	±0.0318	±0.0283	±0.0258	±0.0289	±0.0305
D19	0.8772	**0.8783**	<u>0.8777</u>	0.8769	0.8739
	±0.0239	±0.0258	±0.0256	±0.0261	±0.0259
D20	**0.5146**	<u>0.5055</u>	0.4848	0.4627	0.4473
	±0.02	±0.0223	±0.0241	±0.0216	±0.0229
D21	**0.9351**	0.9331	0.9331	<u>0.9341</u>	0.9296
	±0.0301	±0.024	±0.024	±0.0271	±0.0303
D22	**0.9833**	<u>0.9785</u>	0.9687	0.9591	0.9456
	±0.0126	±0.0115	±0.0198	±0.0201	±0.0216
D23	**0.9916**	<u>0.9907</u>	0.9881	0.9873	0.9858
	±0.0025	±0.0027	±0.0037	±0.0039	±0.0039
D27	0.7943	0.7961	**0.7972**	<u>0.7964</u>	**0.7972**
	±0.0142	±0.0132	±0.0134	±0.0129	±0.0132
D28	0.7919	0.7899	0.7897	<u>0.7924</u>	**0.7928**
	±0.0161	±0.0145	±0.0162	±0.0153	±0.0161
D29	**0.9852**	<u>0.9844</u>	0.9828	0.9816	0.9807
	±0.0042	±0.0041	±0.0037	±0.0042	±0.0039
D31	**0.7499**	<u>0.7371</u>	0.7337	0.727	0.721
	±0.0356	±0.0368	±0.0388	±0.039	±0.0406
D32	0.3289	0.3301	0.3351	<u>0.3391</u>	**0.3416**
	±0.0154	±0.0244	±0.021	±0.0277	±0.0237
D34	**0.9983**	<u>0.9967</u>	0.9944	0.9913	0.9877
	±0.0027	±0.0034	±0.0047	±0.0052	±0.0067

D35	0.2921 ±0.0287	0.2902 ±0.0385	0.295 ±0.0337	<u>0.3034</u> ±0.0328	**0.3112** ±0.0279

The Nonparametric Statistical Test was performed on the experimental data, and the results are shown in Fig S7. The experimental results reveal a progressive degradation in MDGP-Forest's overall performance across all three metrics as the $min\_samples\_split$ value increases. When configured with larger values, this hyperparameter enforces more conservative splitting behavior, which ultimately restricts tree growth and oversimplifies the model, consequently inducing underfitting. We therefore recommend setting $min\_samples\_split$ =2 for optimal MDGP-Forest performance.





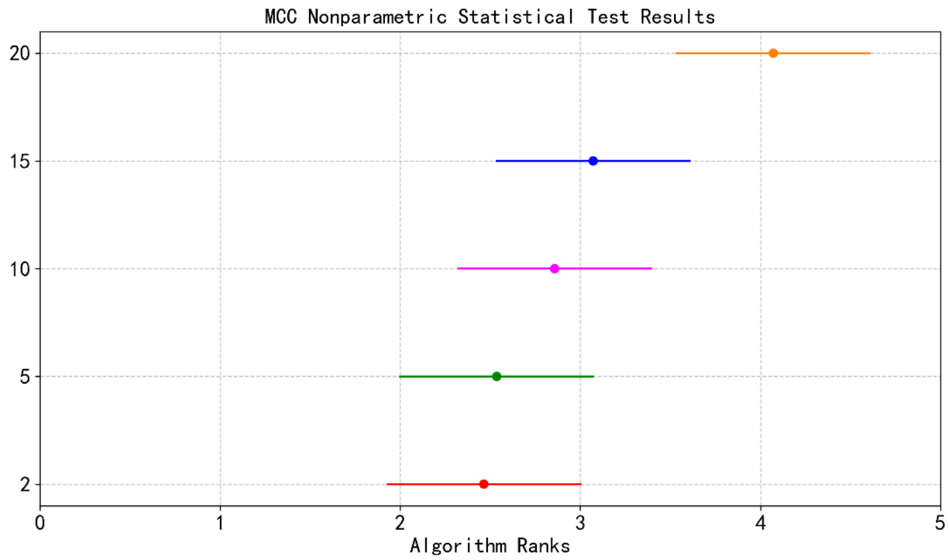


Fig S7 Nonparametric statistical test figure for results of different $min\_samples\_split$ values.

The parameter $min\_samples\_leaf$ regulates the minimum number of samples required at each leaf node in the decision trees comprising MDGP-Forest's decision forest module. This hyperparameter determines the stopping criterion during tree growth by specifying the least number of training instances that must reside in any terminal node, thereby constraining the final topology of the developed trees. MDGP-Forest is configured with five different $min\_samples\_leaf$ values $\{1,2,5,10\}$, and the experimental results are presented in Table S49, Table S50, and Table S51**Table S48**.

Table S49 F1 results of the different *min_samles_leaf* parameter values

Dataset	*min_samles_leaf*=1	*min_samles_leaf* =2	*min_samles_leaf* =5	*min_samles_leaf* =10
D2	0.8964 ±0.0242	0.8976 ±0.0257	**0.9003** ±0.023	<u>0.8984</u> ±0.0216
D3	**0.9104** ±0.0091	<u>0.9065</u> ±0.0087	0.9016 ±0.0093	0.8948 ±0.0097
D4	**0.8128** ±0.0707	<u>0.8126</u> ±0.0639	0.8084 ±0.0653	0.811 ±0.0636
D6	<u>0.9654</u> ±0.0265	0.9574 ±0.0279	**0.9657** ±0.0287	0.9652 ±0.0287
D7	**0.9392** ±0.0282	<u>0.9351</u> ±0.0323	0.9313 ±0.0428	0.9103 ±0.0474
D8	<u>0.6823</u> ±0.1001	**0.6836** ±0.1032	0.6695 ±0.0991	0.6359 ±0.0887
D9	**0.9505** ±0.074	<u>0.9327</u> ±0.0844	0.8107 ±0.1146	0.6806 ±0.1253
D10	**0.7636** ±0.0552	<u>0.7052</u> ±0.0828	0.6095 ±0.1148	0.4049 ±0.0635
D11	0.9812	<u>0.9817</u>	**0.9823**	0.9786

	±0.0072	±0.0085	±0.0079	±0.0079
D12	**0.8149**	0.8125	0.7986	0.7808
	±0.0355	±0.0378	±0.0336	±0.0309
D13	0.341	**0.3419**	0.3231	0.2933
	±0.0778	±0.0772	±0.0615	±0.0448
D14	**0.8631**	0.8384	0.7268	0.6881
	±0.0472	±0.0782	±0.0683	±0.0302
D15	0.9972	**0.9975**	0.9963	0.9939
	±0.0022	±0.002	±0.0025	±0.0033
D16	**0.3457**	0.3303	0.3108	0.3003
	±0.0272	±0.0276	±0.022	±0.0204
D17	**0.8826**	0.8789	0.8674	0.8498
	±0.0129	±0.0119	±0.0127	±0.0133
D18	**0.5835**	0.5759	0.4808	0.4302
	±0.0614	±0.0611	±0.0338	±0.0239
D19	**0.8834**	0.8813	0.8666	0.8412
	±0.0368	±0.0417	±0.0378	±0.0601
D20	**0.4221**	0.4079	0.353	0.2901
	±0.0159	±0.0172	±0.0227	±0.0147
D21	**0.9551**	0.9524	0.9483	0.9475
	±0.0211	±0.0227	±0.026	±0.0313
D22	**0.9846**	0.9776	0.951	0.901
	±0.0118	±0.0131	±0.0209	±0.0367
D23	**0.9924**	0.9908	0.9876	0.9827
	±0.0023	±0.0032	±0.0038	±0.0043
D27	0.8612	0.8626	**0.8648**	0.8641
	±0.01	±0.0107	±0.0089	±0.0079
D28	0.8596	0.859	0.8594	**0.8606**
	±0.0112	±0.0112	±0.0105	±0.011
D29	**0.9867**	0.9855	0.9815	0.9765
	±0.0038	±0.0038	±0.0037	±0.0044
D31	**0.8103**	0.8001	0.7856	0.7723
	±0.0261	±0.0279	±0.0272	±0.032
D32	0.5489	0.5526	0.556	**0.5599**
	±0.009	±0.0119	±0.0125	±0.0107
D34	**0.9986**	0.9961	0.9894	0.9706
	±0.0023	±0.0032	±0.0061	±0.0125
D35	0.5227	0.5231	0.5293	**0.5335**
	±0.0214	±0.0202	±0.0196	±0.02

Table S50 G-mean results of the different *min_samles_leaf* parameter values

Dataset	*min_samles_leaf*=1	*min_samles_leaf* =2	*min_samles_leaf* =5	*min_samles_leaf* =10
D2	0.9277 ±0.0168	0.9288 ±0.0174	**0.9308** ±0.0165	0.9287 ±0.0153
D3	**0.9441** ±0.0057	0.9419 ±0.0054	0.9388 ±0.0056	0.935 ±0.0058
D4	0.8706 ±0.0502	**0.8711** ±0.046	0.8679 ±0.047	0.8694 ±0.0453
D6	0.9712 ±0.0299	0.9599 ±0.0324	**0.9725** ±0.0229	0.9677 ±0.027
D7	**0.9605** ±0.0203	0.9555 ±0.0228	0.9519 ±0.0282	0.94 ±0.0317
D8	0.7465 ±0.0694	**0.748** ±0.0687	0.7378 ±0.0657	0.7153 ±0.0525
D9	**0.9769** ±0.0347	0.9685 ±0.0389	0.9096 ±0.0535	0.8428 ±0.0657
D10	**0.8482** ±0.0397	0.8139 ±0.0511	0.7548 ±0.0775	0.628 ±0.0491
D11	0.9833 ±0.0066	**0.9839** ±0.0068	0.9838 ±0.0068	0.9806 ±0.0069
D12	**0.879** ±0.0256	0.8772 ±0.0274	0.8653 ±0.0257	0.8496 ±0.0251
D13	0.5459 ±0.0572	**0.5501** ±0.0594	0.5344 ±0.0464	0.521 ±0.0385
D14	**0.9134** ±0.0231	0.9001 ±0.0417	0.8407 ±0.0381	0.8211 ±0.019
D15	0.9978 ±0.002	**0.9984** ±0.0017	0.9978 ±0.0017	0.9964 ±0.0024
D16	**0.5587** ±0.0247	0.5486 ±0.0272	0.5314 ±0.0216	0.5208 ±0.0193
D17	**0.9164** ±0.0081	0.9135 ±0.0071	0.9048 ±0.0079	0.8931 ±0.0091
D18	**0.7369** ±0.0412	0.7316 ±0.041	0.6718 ±0.0247	0.6458 ±0.0212
D19	**0.924** ±0.0214	0.921 ±0.0258	0.916 ±0.0251	0.8982 ±0.0374
D20	**0.596** ±0.0116	0.5859 ±0.0126	0.5466 ±0.0167	0.507 ±0.0104
D21	**0.9664** ±0.0157	0.9644 ±0.0169	0.9614 ±0.0191	0.9609 ±0.0228
D22	**0.9915** ±0.0064	0.9877 ±0.0072	0.9729 ±0.0116	0.9456 ±0.02
D23	**0.9958**	0.9949	0.9932	0.9905

	±0.0013	±0.0017	±0.0021	±0.0024
D27	0.8959 ±0.0073	0.8969 ±0.0078	**0.8985** ±0.0066	0.898 ±0.0059
D28	0.8945 ±0.0084	0.8941 ±0.0084	0.8944 ±0.0079	**0.8952** ±0.0083
D29	**0.9926** ±0.0021	0.9919 ±0.0021	0.9897 ±0.0021	0.9869 ±0.0025
D31	**0.8732** ±0.0181	0.8666 ±0.0189	0.8577 ±0.0183	0.849 ±0.0223
D32	0.6549 ±0.0081	0.6582 ±0.0102	0.6618 ±0.0111	**0.6647** ±0.0101
D34	**0.9992** ±0.0013	0.9978 ±0.0018	0.9939 ±0.0035	0.983 ±0.0072
D35	0.6313 ±0.0157	0.6314 ±0.0163	0.6366 ±0.016	**0.6406** ±0.0169

Table S51 MCC results of the different *min_samles_leaf* parameter values

Dataset	*min_samles_leaf*=1	*min_samles_leaf* =2	*min_samles_leaf* =5	*min_samles_leaf* =10
D2	0.8631 ±0.0334	0.8649 ±0.0358	**0.8685** ±0.0341	0.8663 ±0.0314
D3	**0.9099** ±0.0095	0.9057 ±0.009	0.9011 ±0.0095	0.894 ±0.0097
D4	**0.7691** ±0.085	0.7686 ±0.0778	0.7654 ±0.0794	0.7689 ±0.0787
D6	**0.9472** ±0.0388	0.937 ±0.038	0.9453 ±0.0457	0.9457 ±0.0448
D7	**0.945** ±0.0229	0.9434 ±0.0254	0.9435 ±0.0293	0.9315 ±0.031
D8	0.5765 ±0.0907	**0.5852** ±0.0871	0.5761 ±0.0963	0.5515 ±0.0825
D9	**0.9741** ±0.0389	0.97 ±0.0353	0.9162 ±0.0518	0.8512 ±0.0809
D10	**0.7411** ±0.0442	0.6934 ±0.0518	0.6445 ±0.0664	0.5439 ±0.0665
D11	0.9709 ±0.0107	**0.9726** ±0.0123	0.9722 ±0.012	0.9665 ±0.0127
D12	**0.7388** ±0.0342	0.7357 ±0.0351	0.7202 ±0.0368	0.7003 ±0.0355
D13	0.3483 ±0.0699	**0.3594** ±0.0723	0.3491 ±0.0681	0.345 ±0.0712
D14	**0.8408** ±0.0329	0.8403 ±0.0369	0.8185 ±0.0363	0.8078 ±0.0337
D15	0.9972	**0.998**	0.9973	0.9955

	±0.002	±0.0016	±0.0018	±0.0026
D16	**0.501** ±0.0613	0.494 ±0.0623	0.447 ±0.0491	0.4186 ±0.0433
D17	**0.9151** ±0.0052	0.9104 ±0.006	0.8987 ±0.006	0.884 ±0.0063
D18	0.4998 ±0.0318	**0.5034** ±0.0269	0.4995 ±0.0269	0.4904 ±0.0314
D19	0.8772 ±0.0239	**0.8785** ±0.0267	0.8735 ±0.0289	0.8662 ±0.0356
D20	**0.5146** ±0.02	0.5017 ±0.0236	0.4433 ±0.0266	0.4099 ±0.0219
D21	**0.9351** ±0.0301	0.9314 ±0.0321	0.9261 ±0.0351	0.9252 ±0.0416
D22	**0.9833** ±0.0126	0.9758 ±0.0141	0.9469 ±0.0222	0.8947 ±0.0376
D23	**0.9916** ±0.0025	0.9898 ±0.0035	0.9864 ±0.0042	0.981 ±0.0048
D27	0.7943 ±0.0142	0.7962 ±0.0152	**0.7991** ±0.013	0.7981 ±0.0117
D28	0.7919 ±0.0161	0.7911 ±0.0163	0.7919 ±0.0154	**0.7929** ±0.0163
D29	**0.9852** ±0.0042	0.9839 ±0.0042	0.9794 ±0.0041	0.9738 ±0.0049
D31	**0.7499** ±0.0356	0.7374 ±0.0367	0.7211 ±0.0358	0.7043 ±0.0436
D32	0.3289 ±0.0154	0.3355 ±0.0192	0.3437 ±0.0208	**0.3497** ±0.0197
D34	**0.9983** ±0.0027	0.9953 ±0.0038	0.9875 ±0.0072	0.966 ±0.0149
D35	0.2921 ±0.0287	0.2922 ±0.0315	0.3032 ±0.0297	**0.3123** ±0.0324

The Nonparametric Statistical Test was performed on the experimental data, and the results are shown in Fig S8. The experimental results demonstrate a consistent decline in MDGP-Forest's overall performance across all three metrics as the $min\_samples\_leaf$ value increases. When configured with excessively large values, this hyperparameter forces each terminal node to contain more samples, thereby reducing the tree's sensitivity to fluctuations in the training data. This renders the model overly conservative, impairing its ability to capture complex data patterns and ultimately leading to underfitting. We therefore recommend setting $min\_samples\_leaf = 1$ for optimal MDGP-Forest performance.

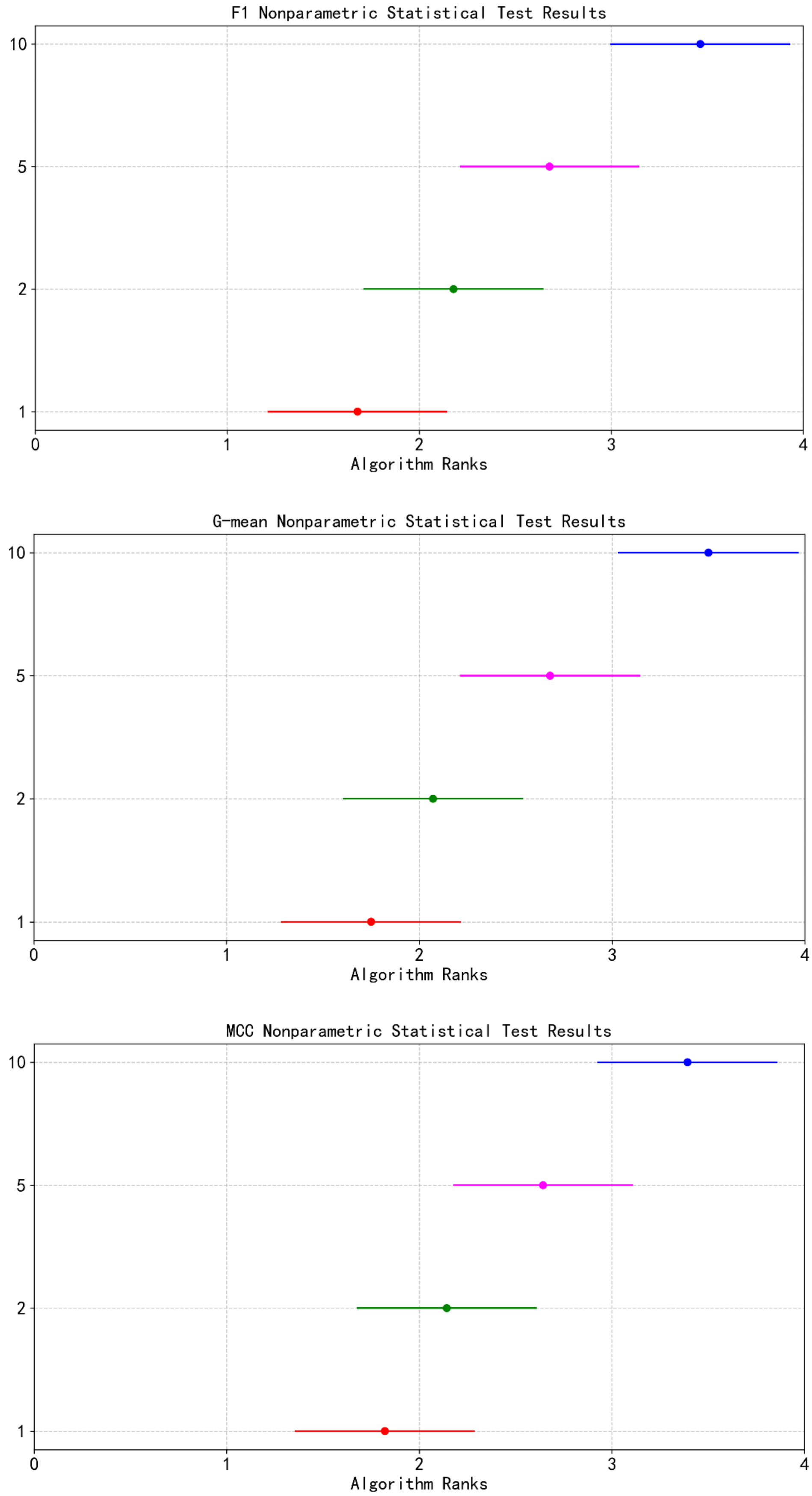


Fig S8 Nonparametric statistical test figure for results of different $min\_samples\_leaf$ values.

4.3 Parameter Setting for High-Dimensional Datasets

In high-dimensional datasets, instances possess rich feature information. If the population size in MDGP-Forest is set too small, the initialized candidate features may only contain a limited subset of the original features, resulting in significant information loss. If the number of constructed features per class is insufficient, the algorithm fails to fully utilize class information, leading to poor inter-class discriminability. Therefore, it is advisable to set both parameters at relatively higher values for high-dimensional datasets, with recommended settings as shown in Table S52.

Table S52 Hyperparameter setting for High-Dimensional dataset

Parameter	High-dimensional setting	Default setting
GP Population Size	100	50
Number of Features Constructed per Population	50	10

Table S53, Table S54, and Table S55 present the comparative results of MDGP-Forest's two parameter settings on high-dimensional datasets. The high-dimensional setting demonstrates superior performance to the default setting in all metrics across all tested high-dimensional datasets. The results demonstrate that the High-dimensional setting is more suitable for high-dimensional datasets and can more effectively mitigate information loss issues.

Table S53 F1 results of the High-dimensional setting

Dataset	High-dimensional setting	Default setting
D1	**0.9573** ±0.0128	<u>0.9547</u> ±0.0136
D5	**0.7033** ±0.0098	<u>0.7012</u> ±0.0099
D24	**0.9781** ±0.0032	<u>0.9779</u> ±0.0027
D25	**0.9739** ±0.007	<u>0.9734</u> ±0.006
D26	**0.9358** ±0.0099	<u>0.9294</u> ±0.0097
D30	**0.95** ±0.0105	<u>0.9477</u> ±0.011
D33	**0.9925** ±0.0018	<u>0.9915</u> ±0.0022

Table S54 G-mean results of the High-dimensional setting

Dataset	High-dimensional setting	Default setting
D1	**0.9693** ±0.0099	<u>0.9673</u> ±0.0107
D5	**0.8134** ±0.0067	<u>0.8119</u> ±0.007
D24	**0.9878** ±0.0018	<u>0.9877</u> ±0.0015
D25	**0.9855** ±0.0039	<u>0.9852</u> ±0.0033
D26	**0.9635** ±0.0057	<u>0.9599</u> ±0.0052
D30	**0.9719** ±0.0059	<u>0.9707</u> ±0.0062
D33	**0.9955** ±0.0011	<u>0.9949</u> ±0.0013

Table S55 MCC results of the High-dimensional setting

Dataset	High-dimensional setting	Default setting
D1	**0.9388** ±0.0177	<u>0.9348</u> ±0.0189
D5	**0.6643** ±0.0131	<u>0.6617</u> ±0.014
D24	**0.9758** ±0.0035	<u>0.9756</u> ±0.003
D25	**0.9712** ±0.0077	<u>0.9706</u> ±0.0066
D26	**0.9283** ±0.011	<u>0.9214</u> ±0.0098
D30	**0.9447** ±0.0116	<u>0.9422</u> ±0.0122
D33	**0.9907** ±0.0023	<u>0.9895</u> ±0.0027

5. Detailed data of runtime analysis

This section conducts a detailed analysis of the training time for the four datasets presented in the main text, and reporting of both train and test time data in all datasets.

5.1 Training time analysis for the four datasets in the main text

In the main text, the Runtime analysis section only presented the training time for four datasets: thyroid, page_blocks, optdigits, and mfeat_pixel. This section presents a detailed analysis of the training time for the four datasets in the main text. The characteristics of these four datasets are shown in Table S56.

Table S56 Characteristics of Typical Datasets for training time Analysis.

Dataset ID	Dataset	Number of Instances	Number of Features	Number of Classes	Imbalance Ratio
D6	thyroid	small	small	small	secondary
D19	page_blocks	high	small	secondary	high
D29	optdigits	high	secondary	high	small
D24	mfeat_pixel	secondary	high	high	small

In the thyroid dataset, which has a relatively small number of instances, features, and classes, the training time of MDGP-Forest is close to that of DeepForest. This indicates that MDGP-Forest does not experience a significant increase in training time on small-scale datasets. Imbalance-DF has the shortest training time, suggesting that the algorithm completes training with a small number of cascade layers. OpenFE also has a very short training time on this dataset. This is because when the number of features is small, the acceleration effect of OpenFE's SuccessivePruning algorithm is particularly evident.

The shorter training time of MDGP-Forest compared to Deep Forest on the thyroid dataset represents a special case observed with small-scale datasets. This phenomenon occurs because the computational overhead of class-related GP feature construction becomes negligible for small-scale data. Furthermore, MDGP-Forest employs Python's multiprocessing toolkit to optimize this module, effectively avoiding significant increases in training time. Simultaneously, MDGP-Forest's utilization of higher-order feature information may enable faster convergence compared to Deep Forest. These factors contribute to this special situation. While in most cases, MDGP-Forest requires longer training time than Deep Forest due to its additional components specifically designed for multi-class imbalanced learning.

The page blocks dataset has a large number of instances and a high imbalance ratio. Imbalance-DF has the longest training time on this dataset. This is because Imbalance-DF uses the SOMTE-based imbalanced learning algorithm as the forest module for cascade layers, which employs oversampling to generate a large number of synthetic instances when the imbalance ratio is high, thereby increasing its training time. MDGP-Forest includes steps such as Multi-class disassembly and Sampling, and Class-related GP feature construction, so its training time is longer than that of DeepForest. Since the number of features in page blocks is relatively small, the number of constructed features generated by OpenFE is limited, resulting in the shortest training time.

The optdigits dataset has a relatively large number of classes and instances, which results in MDGP-Forest requiring more training time. Due to the very small imbalance ratio of the optdigits dataset, the resampling of imbalance-DF does not play a significant role, leading to a training time close to that of Deep Forest. The optdigits dataset has a large number of features, which causes OpenFE to generate a massive number of constructed features before performing feature selection, resulting in the longest training time.

The mfeat_pixel dataset is high-dimensional, and therefore OpenFE generates a larger scale of features, resulting in a significantly longer training time compared to other algorithms. While generating massive features, OpenFE also occupies a large amount of memory space, potentially leading to insufficient memory issues. In contrast, MDGP-Forest, which implements feature construction based on GP, does not experience excessive computational expansion, and achieves an acceptable training time.

5.2 Detailed data for training time

The training times of MDGP-Forest, DeepForest, imbalance-DF, and OpenFE on 35 datasets under 5-fold cross-validation are shown in Table S57. Specifically, the training time of OpenFE on the mfeat pixel dataset was obtained by running each fold of the 5-fold cross-validation separately in sequence. For the other datasets with more than 100 features, OpenFE was unable to complete the 5-fold cross-validation under the same experimental settings, and these cases are marked with "-" in the table.

Table S57 Training time of each algorithm (seconds)

Dataset	MDGP-Forest	DeepForest	imbalance-DF	OpenFE
D1	277.89	26.94	28.95	-
D2	35.57	17.86	4.82	12.31
D3	111.74	22.98	30.98	115.08
D4	21.39	20.13	3.34	5.10
D5	1107.33	28.18	173.65	-
D6	15.45	18.32	3.41	5.12
D7	35.56	27.38	5.41	5.81
D8	30.27	21.58	4.99	7.15
D9	15.47	19.34	2.48	6.74
D10	15.58	17.69	3.34	4.59
D11	27.28	24.58	5.59	19.71
D12	49.59	22.16	16.54	25.54
D13	20.76	16.81	4.21	5.93
D14	16.09	15.73	3.65	5.82
D15	258.02	30.33	32.91	19.37
D16	37.33	22.01	7.12	10.66

D17	146.82	28.05	193.16	46.11
D18	38.33	17.46	6.61	8.55
D19	28.82	20.27	54.08	9.90
D20	179.94	25.79	24.96	16.07
D21	10.53	16.59	2.28	4.96
D22	27.66	21.85	4.49	9.67
D23	135.29	24.76	27.96	64.88
D24	136.06	21.12	18.95	2905.41
D25	137.75	18.58	30.25	-
D26	151.02	20.47	15.26	-
D27	63.17	19.34	18.49	23.88
D28	53.50	15.35	16.03	14.09
D29	126.53	24.53	24.15	270.36
D30	93.85	24.18	12.91	-
D31	23.37	22.56	4.12	13.78
D32	89.65	21.40	10.14	7.65
D33	1482.22	50.38	927.26	-
D34	22.26	20.93	14.41	44.94
D35	36.98	26.36	6.29	7.41

5.3 Detailed data for testing time

The testing time of MDGP-Forest, DeepForest, imbalance-DF, and OpenFE on 35 datasets under 5-fold cross-validation is shown in Table S58. Since OpenFE employs LightGBM as its downstream classifier, which has been meticulously optimized for parallel computing, it achieves the shortest testing time. However, OpenFE’s excessively long training time on high-dimensional datasets significantly limits its broader application in complex scenarios. MDGP-Forest exhibits longer testing time than Deep Forest due to the additional computational overhead from GP-based feature transformations between cascade layers.

Table S58 Testing time of each algorithm (seconds)

Dataset	MDGP-Forest	DeepForest	imbalance-DF	OpenFE
D1	13.08792	6.51403	0.30895	-
D2	12.96983	3.49055	0.07974	0.00093
D3	6.68611	5.29139	0.30541	0.00202
D4	7.93642	4.25758	0.06548	0.00082
D5	25.41330	5.45902	2.23956	-
D6	5.89452	3.76026	0.07126	0.00078
D7	20.67550	6.66827	0.13806	0.00095
D8	11.05255	4.70833	0.09821	0.00110
D9	6.36225	4.07422	0.04407	0.00085
D10	5.19761	3.46015	0.06040	0.00073

D11	7.02177	5.72818	0.05019	0.00143
D12	6.24201	4.94263	0.14464	0.00180
D13	6.94348	3.17873	0.06850	0.00094
D14	5.00895	2.80713	0.06154	0.00104
D15	10.49603	7.65634	0.62359	0.00226
D16	5.29106	4.80114	0.07065	0.00124
D17	14.66319	6.85858	0.52137	0.00455
D18	4.84697	3.35159	0.04715	0.00136
D19	5.97900	4.33715	0.27484	0.00239
D20	14.85740	6.03467	0.13553	0.00196
D21	4.55888	3.11818	0.04366	0.00076
D22	4.94700	4.75981	0.07315	0.00167
D23	10.11967	5.89055	0.30479	0.01091
D24	7.42685	4.54527	0.28821	-
D25	10.31474	3.80163	0.25366	-
D26	8.00850	4.34969	0.27762	-
D27	4.42248	4.10056	0.16233	0.00146
D28	6.66332	2.76316	0.19496	0.00192
D29	7.00825	5.79953	0.33866	0.00310
D30	7.85699	5.56619	0.19690	0.00434
D31	5.56014	5.01463	0.06860	0.00124
D32	5.84066	4.61077	0.18431	0.00173
D33	45.99570	11.49115	4.85346	-
D34	3.86982	4.45870	0.19027	0.00157
D35	4.62015	6.20511	0.14922	0.00133

To address the computational time limitations of MDGP-Forest, the following potential solutions could be considered in future work: (1) developing a dedicated parallel and distributed computing framework specifically for MDGP-Forest to enhance its workflow computational efficiency; (2) optimizing the fitness evaluation function of GP in MDGP-Forest to enable assessment of all individuals in the population with reduced computational overhead.

6. Visualization of constructed features in MDGP-Forest

This section demonstrates the effectiveness of MDGP-Forest's class-related GP feature construction using the cardiotocography dataset as an example. The dataset contains 2,126 instances with 22 features across 3 classes. The scatter plot visualization after PCA-based dimensionality reduction is shown in Fig S9, in which purple points represent the majority class (most instances) in the cardiotocography dataset, while the other two classes are relatively minority. Class overlap is observed between the purple class and the other two classes. The yellow class separates into two distinct groups, with one group demonstrating complex overlap patterns with the teal class.

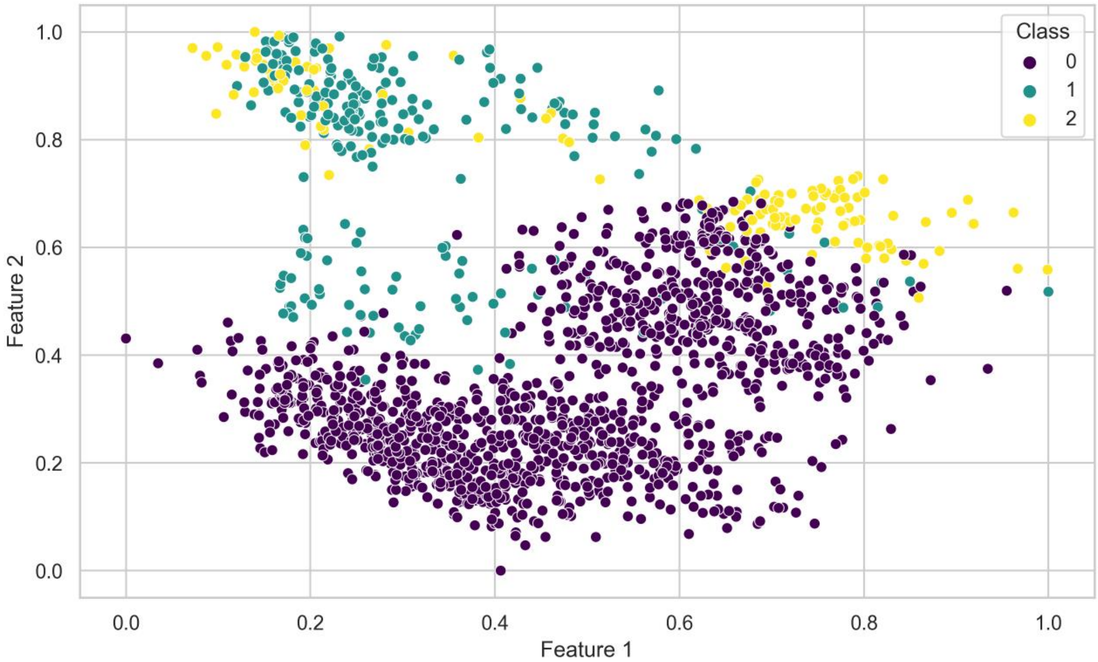


Fig S9 PCA-reduced scatter plot of the cardiotocography dataset

The visualization first extracts constructed features from each cascade layer of MDGP-Forest trained on the cardiotocography dataset. Subsequently, multi-class disassembly is employed to construct binary classification datasets and match them with corresponding class's constructed features. Finally, PCA is applied for dimensionality reduction followed by scatter plot visualization.

The trained MDGP-Forest comprises a 4-layer cascade structure. Visualization results of the constructed features in the subsequent three layers, categorized by class, are presented in Fig S10, Fig S11, and Fig S12. Fig S10 demonstrates that the features constructed by MDGP-Forest in the second cascade layer have improved inter-class separation to some extent. As the MDGP-Forest training progresses, the quality of constructed features continually improves. In the final cascade layer shown in Fig S12, these features effectively discriminate most instances of each class from other classes.

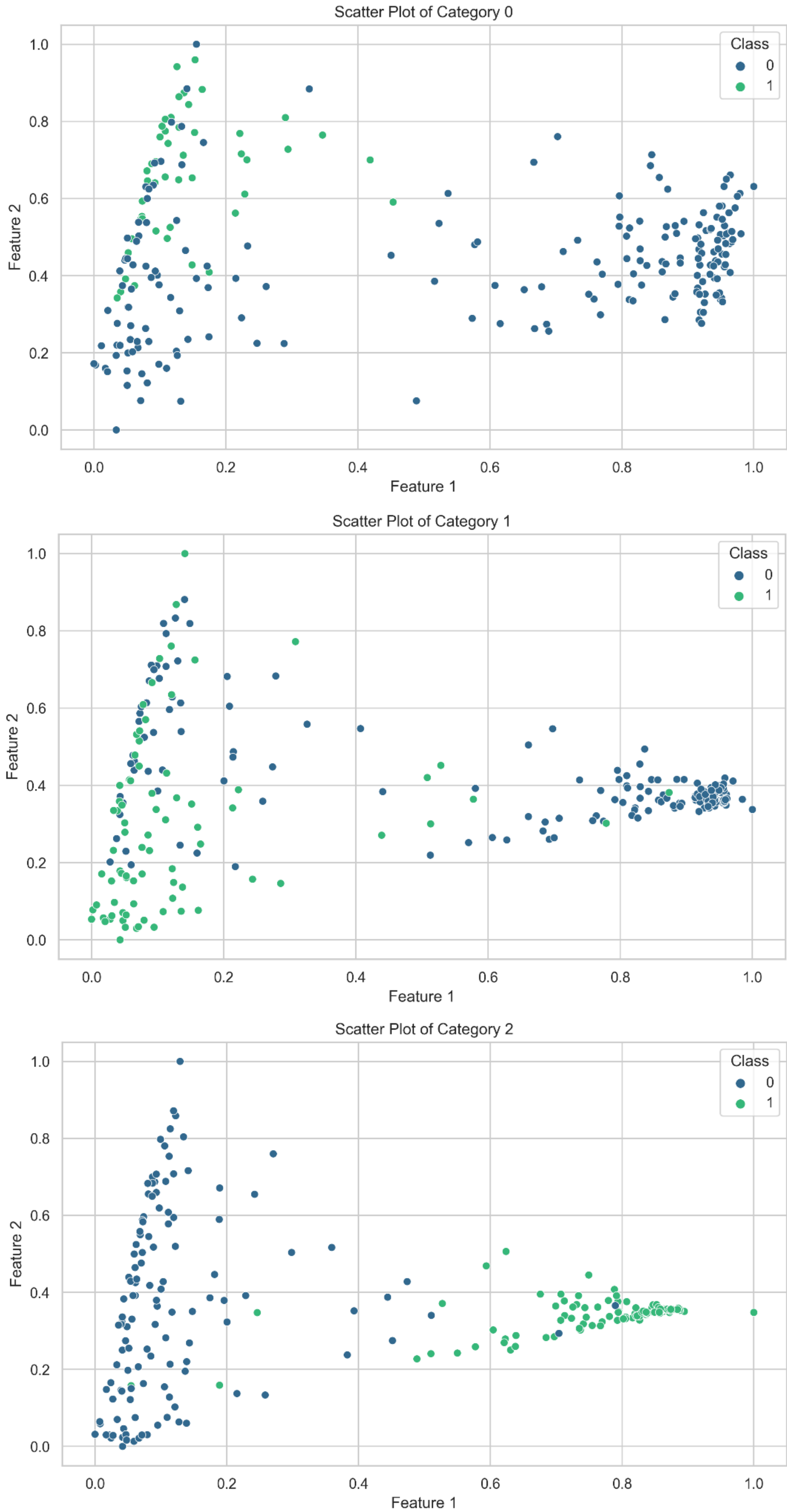


Fig S10 Visualization of Constructed Features in MDGP-Forest’s Second Cascade Layer

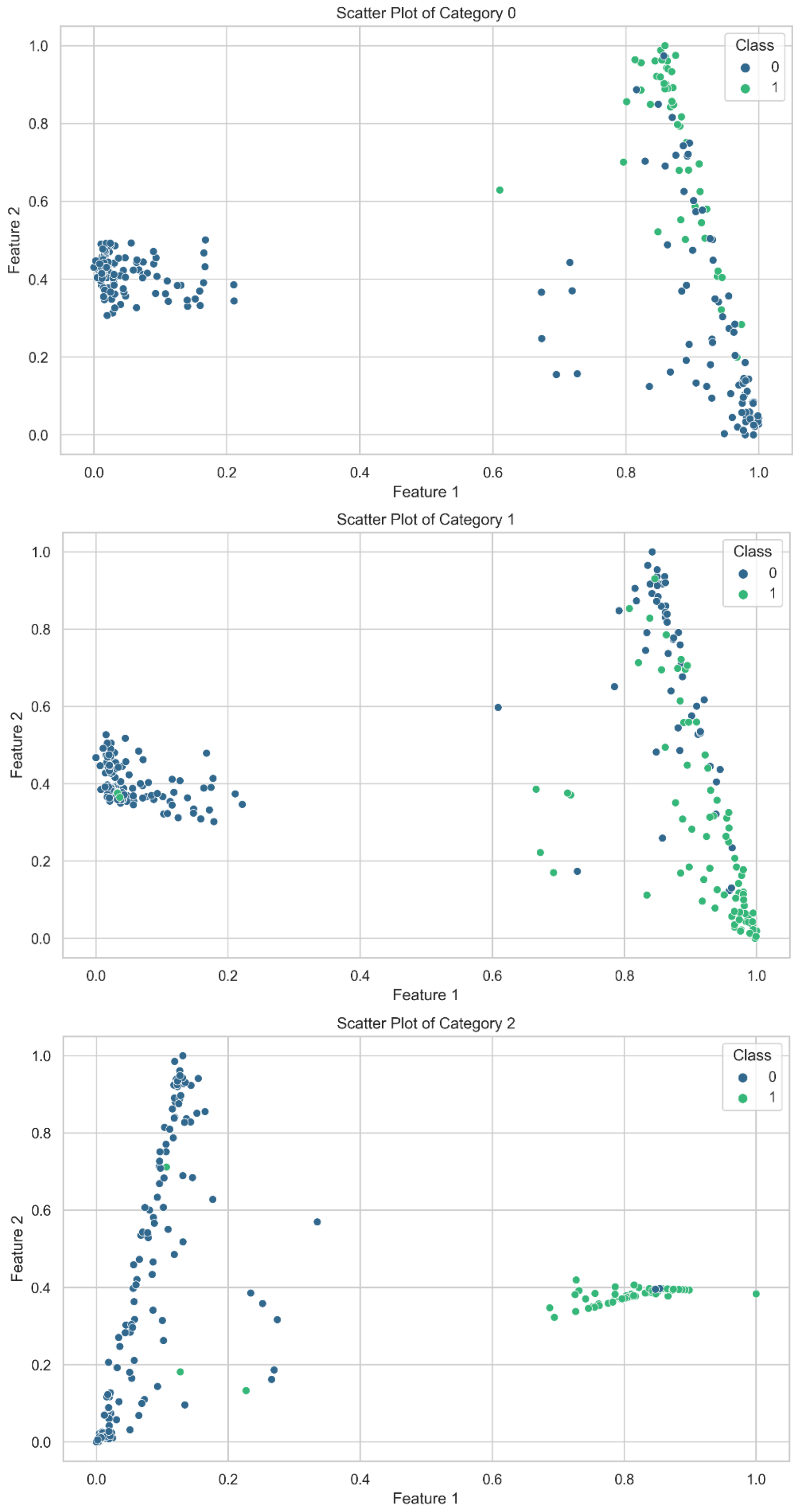


Fig S11 Visualization of Constructed Features in MDGP-Forest's Third Cascade Layer

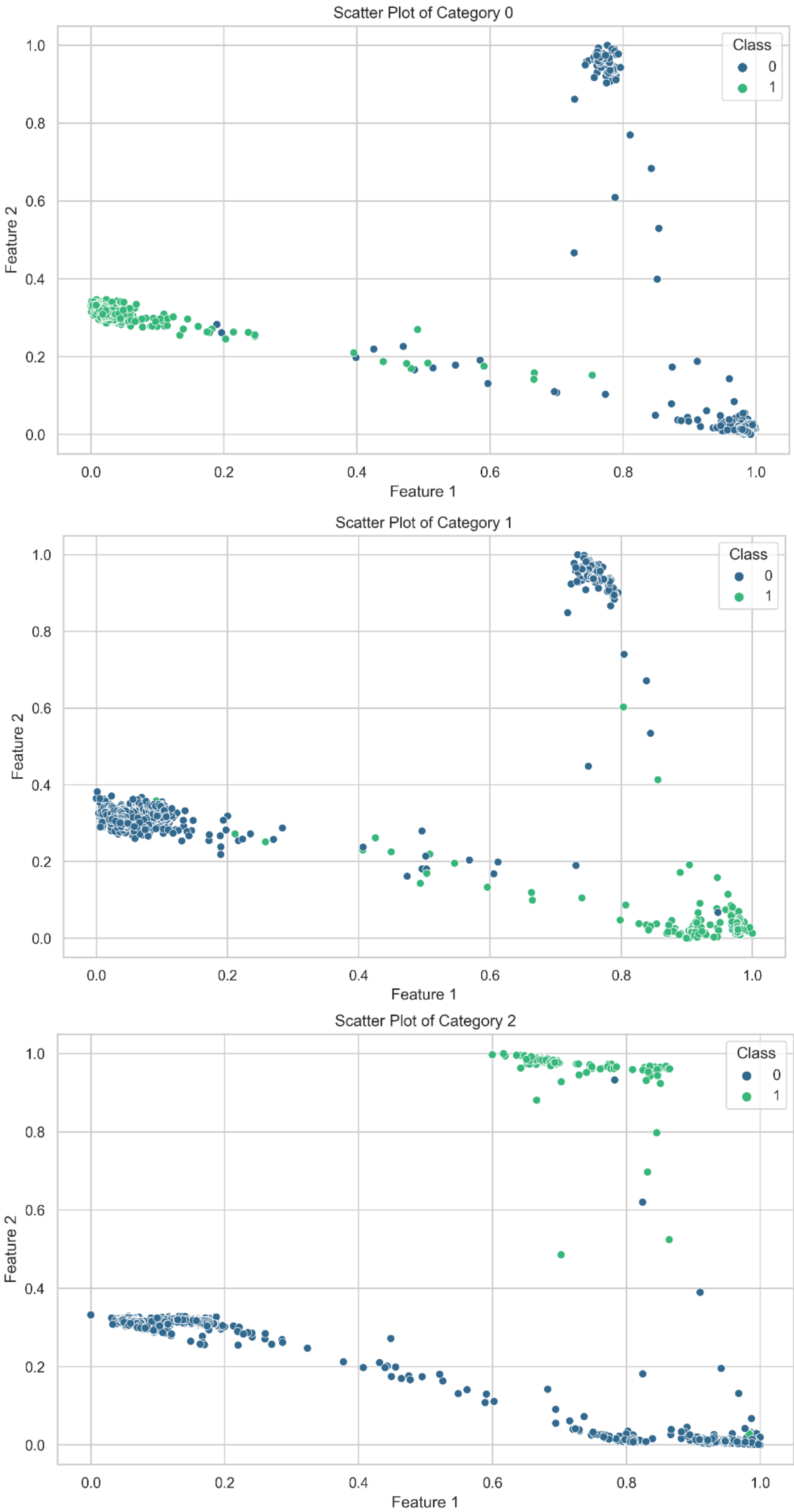


Fig S12 Visualization of Constructed Features in MDGP-Forest's Fourth Cascade Layer