**Supplementary Table 1. The detailed parameters of the 5 Algorithms.**

|  |  |
| --- | --- |
| Algorithms | Parameters |
| Logistics Regression | class\_weight={0:0.1, 1:0.9}, penalty='l1', solver='liblinear', random\_state=2024 |
| Random Forest | n\_estimators=100, max\_depth=3, min\_samples\_split=10, min\_samples\_leaf=4, max\_features=1, oob\_score=True, random\_state=2023 |
| LGBM | boosting\_type='gbdt' ,objective='binary' ,learning\_rate=0.02 ,n\_estimators=100 ,max\_depth=4 ,num\_leaves=30 ,min\_child\_samples=6 ,min\_child\_weight=1 ,bagging\_fraction=1 ,feature\_fraction=0.9 ,subsample=1 ,colsample\_bytree=0.8 ,reg\_alpha=5 ,reg\_lambda=10 ,random\_state=2024 |
| XGBoost | objective='binary:logistic' ,learning\_rate=0.02 ,n\_estimators=230 ,max\_depth=4 ,min\_child\_weight=5 ,gamma=0 ,subsample=0.8 ,colsample\_bytree=1 ,scale\_pos\_weight=1 ,reg\_alpha=2 ,reg\_lambda=1 ,n\_jobs=-1, eval\_metric=['logloss','auc','error'], seed=2024 |
| CatBoost | loss\_function="Logloss", eval\_metric="AUC", learning\_rate=0.01, iterations=1000, random\_seed=42, od\_type="Iter", depth=4, early\_stopping\_rounds=800, colsample\_bylevel=0.1, l2\_leaf\_reg=40, random\_strength=800, scale\_pos\_weight=1, silent=True |

**Supplementary Figure 1. Feature importance ranking by RFE.**

