参数与结果记录

初始参数结果：

*{'num\_thread': 4, 'num\_leaves': 12, 'metric': 'binary', 'objective': 'binary', 'num\_round': 2000, 'learning\_rate': 0.02, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}*

*The minimum is attained in round 421*

*The minimum is attained in round 428*

*The minimum is attained in round 399*

*The minimum is attained in round 610*

*The minimum is attained in round 297*

*8.008000000000001 %*

参数记录：

考虑算力情况，num\_round 定为2000，因此降低学习率，此时num\_round大约在1000左右；第四个，1500才到最低，与其它差的比较多；

*{'num\_thread': 12, 'num\_leaves': 12, 'metric': 'binary', 'objective': 'binary', 'num\_round': 2000, 'learning\_rate': 0.01, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}*

*The minimum is attained in round 879*

*The minimum is attained in round 850*

*The minimum is attained in round 944*

*The minimum is attained in round 1521*

*The minimum is attained in round 648*

*8.005999999999998 %*

尝试一些极端值；

首先，增加树的复杂度num\_leaves=50，同时需要适当降低num\_round=1000，发现效果变差了0.1%。

*{'num\_thread': 12, 'num\_leaves': 50, 'metric': 'binary', 'objective': 'binary', 'num\_round': 1000, 'learning\_rate': 0.01, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}*

*The minimum is attained in round 693*

*The minimum is attained in round 509*

*The minimum is attained in round 565*

*The minimum is attained in round 503*

*The minimum is attained in round 512*

*8.128000000000002 %*

降低学习率learning\_rate=0.005，num\_round=2000，减慢收敛速度，效果更差了！

**本来想的是，降低lr，能够找更好的找到局部最优值！！**

*{'num\_thread': 12, 'num\_leaves': 50, 'metric': 'binary', 'objective': 'binary', 'num\_round': 2000, 'learning\_rate': 0.005, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}*

*The minimum is attained in round 1445*

*The minimum is attained in round 965*

*The minimum is attained in round 1114*

*The minimum is attained in round 1075*

*The minimum is attained in round 963*

*8.167999999999997 %*

**本来想的是，降低lr，能够找更好的找到局部最优值！！==>> 增加lr，**learning\_rate=0.02，效果依然不好

*{'num\_thread': 12, 'num\_leaves': 50, 'metric': 'binary', 'objective': 'binary', 'num\_round': 1000, 'learning\_rate': 0.02, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}*

*The minimum is attained in round 330*

*The minimum is attained in round 241*

*The minimum is attained in round 313*

*The minimum is attained in round 262*

*The minimum is attained in round 253*

*8.166000000000002 %*

考虑，降低树的复杂度num\_leaves=6，同时需要适当增加num\_round=3000，相比修改前，还是差0.01%。

*{'num\_thread': 12, 'num\_leaves': 6, 'metric': 'binary', 'objective': 'binary', 'num\_round': 3000, 'learning\_rate': 0.01, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}*

*The minimum is attained in round 1279*

*The minimum is attained in round 1607*

*The minimum is attained in round 1263*

*The minimum is attained in round 1427*

*The minimum is attained in round 985*

*8.018 %*

使用循环，观察不同num\_leaves下的效果；

learning\_rate=0.01，num\_leaves = [6, 12, 20, 32, 48, 64, 80, 96, 127],比较好的结果主要集中在**num\_leaves=6-20之间，因此深度较大，可能没有帮助。**

*{'num\_thread': 12, 'num\_leaves': 6, 'metric': 'binary', 'objective': 'binary', 'num\_round': 2000, 'learning\_rate': 0.01, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}*

*The minimum is attained in round 1279*

*The minimum is attained in round 1607*

*The minimum is attained in round 1263*

*The minimum is attained in round 1427*

*The minimum is attained in round 985*

*8.018 %*

*{'num\_thread': 12, 'num\_leaves': 12, 'metric': 'binary', 'objective': 'binary', 'num\_round': 2000, 'learning\_rate': 0.01, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}*

*The minimum is attained in round 879*

*The minimum is attained in round 850*

*The minimum is attained in round 944*

*The minimum is attained in round 1521*

*The minimum is attained in round 648*

*8.005999999999998 %*

*{'num\_thread': 12, 'num\_leaves': 20, 'metric': 'binary', 'objective': 'binary', 'num\_round': 2000, 'learning\_rate': 0.01, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}*

*The minimum is attained in round 820*

*The minimum is attained in round 671*

*The minimum is attained in round 713*

*The minimum is attained in round 840*

*The minimum is attained in round 542*

*8.078000000000001 %*

*{'num\_thread': 12, 'num\_leaves': 32, 'metric': 'binary', 'objective': 'binary', 'num\_round': 2000, 'learning\_rate': 0.01, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}*

*The minimum is attained in round 748*

*The minimum is attained in round 558*

*The minimum is attained in round 645*

*The minimum is attained in round 518*

*The minimum is attained in round 511*

*8.069999999999997 %*

*{'num\_thread': 12, 'num\_leaves': 48, 'metric': 'binary', 'objective': 'binary', 'num\_round': 2000, 'learning\_rate': 0.01, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}*

*The minimum is attained in round 680*

*The minimum is attained in round 502*

*The minimum is attained in round 559*

*The minimum is attained in round 510*

*The minimum is attained in round 519*

*8.131999999999998 %*

根据上面结论，研究固定num\_leaves下，不同learning\_rate的影响。num\_leaves=12，

learning\_rate=[0.001, 0.005, 0.01, 0.015, 0.02, 0.05, 0.1, 0.2, 0.4]

**learning\_rate=0.004 情况下，最低！！**

*{'num\_thread': 12, 'num\_leaves': 12, 'metric': 'binary', 'objective': 'binary', 'num\_round': 4000, 'learning\_rate': 0.002, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}*

*The minimum is attained in round 3963*

*Finished loading model, total used 4000 iterations*

*The minimum is attained in round 3999*

*Finished loading model, total used 4000 iterations*

*The minimum is attained in round 3996*

*Finished loading model, total used 4000 iterations*

*The minimum is attained in round 3572*

*Finished loading model, total used 4000 iterations*

*The minimum is attained in round 3150*

*Finished loading model, total used 4000 iterations*

*8.031999999999998 %*

*{'num\_thread': 12, 'num\_leaves': 12, 'metric': 'binary', 'objective': 'binary', 'num\_round': 4000, 'learning\_rate': 0.004, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}*

*The minimum is attained in round 2143*

*Finished loading model, total used 4000 iterations*

*The minimum is attained in round 2194*

*Finished loading model, total used 4000 iterations*

*The minimum is attained in round 2226*

*Finished loading model, total used 4000 iterations*

*The minimum is attained in round 3538*

*Finished loading model, total used 4000 iterations*

*The minimum is attained in round 1695*

*Finished loading model, total used 4000 iterations*

*7.983999999999999 %*

*{'num\_thread': 12, 'num\_leaves': 12, 'metric': 'binary', 'objective': 'binary', 'num\_round': 4000, 'learning\_rate': 0.005, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}*

*The minimum is attained in round 1738*

*The minimum is attained in round 1701*

*The minimum is attained in round 1552*

*The minimum is attained in round 2554*

*The minimum is attained in round 1297*

*8.012 %*

*{'num\_thread': 12, 'num\_leaves': 12, 'metric': 'binary', 'objective': 'binary', 'num\_round': 4000, 'learning\_rate': 0.008, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}*

*The minimum is attained in round 1058*

*Finished loading model, total used 4000 iterations*

*The minimum is attained in round 1120*

*Finished loading model, total used 4000 iterations*

*The minimum is attained in round 1105*

*Finished loading model, total used 4000 iterations*

*The minimum is attained in round 1588*

*Finished loading model, total used 4000 iterations*

*The minimum is attained in round 811*

*Finished loading model, total used 4000 iterations*

*8.016000000000004 %*

*{'num\_thread': 12, 'num\_leaves': 12, 'metric': 'binary', 'objective': 'binary', 'num\_round': 4000, 'learning\_rate': 0.01, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}*

*The minimum is attained in round 879*

*The minimum is attained in round 850*

*The minimum is attained in round 944*

*The minimum is attained in round 1521*

*The minimum is attained in round 648*

*8.005999999999998 %*

*{'num\_thread': 12, 'num\_leaves': 12, 'metric': 'binary', 'objective': 'binary', 'num\_round': 4000, 'learning\_rate': 0.015, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}*

*The minimum is attained in round 601*

*The minimum is attained in round 606*

*The minimum is attained in round 588*

*The minimum is attained in round 794*

*The minimum is attained in round 444*

*8.022 %*

*{'num\_thread': 12, 'num\_leaves': 12, 'metric': 'binary', 'objective': 'binary', 'num\_round': 4000, 'learning\_rate': 0.02, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}*

*The minimum is attained in round 421*

*The minimum is attained in round 428*

*The minimum is attained in round 399*

*The minimum is attained in round 610*

*The minimum is attained in round 297*

*8.008000000000001 %*

*{'num\_thread': 12, 'num\_leaves': 12, 'metric': 'binary', 'objective': 'binary', 'num\_round': 4000, 'learning\_rate': 0.05, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}*

*The minimum is attained in round 168*

*The minimum is attained in round 147*

*The minimum is attained in round 149*

*The minimum is attained in round 238*

*The minimum is attained in round 121*

*8.024000000000001 %*

**根据上面结论，固定learning\_rate=0.004，num\_leaves=range(6,16,2)，num\_leaves=8**

考虑参数feature\_fraction = bagging\_fraction, fractions = [0.6, 0.7, 0.8, 0.9, 1.0],

**Feature\_fraction = 0.8 最好**

{'num\_thread': 12, 'num\_leaves': 12, 'metric': 'binary', 'objective': 'binary', 'num\_round': 4000, 'learning\_rate': 0.004, 'feature\_fraction': 0.6, 'bagging\_fraction': 0.6}

The minimum is attained in round 2385

The minimum is attained in round 2577

The minimum is attained in round 2404

The minimum is attained in round 3486

The minimum is attained in round 1759

8.033999999999999 %

{'num\_thread': 12, 'num\_leaves': 12, 'metric': 'binary', 'objective': 'binary', 'num\_round': 4000, 'learning\_rate': 0.004, 'feature\_fraction': 0.7, 'bagging\_fraction': 0.7}

The minimum is attained in round 2288

The minimum is attained in round 2173

The minimum is attained in round 2059

The minimum is attained in round 3284

The minimum is attained in round 1730

8.012 %

{'num\_thread': 12, 'num\_leaves': 12, 'metric': 'binary', 'objective': 'binary', 'num\_round': 4000, 'learning\_rate': 0.004, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}

The minimum is attained in round 2143

The minimum is attained in round 2194

The minimum is attained in round 2226

The minimum is attained in round 3538

The minimum is attained in round 1695

7.983999999999999 %

{'num\_thread': 12, 'num\_leaves': 12, 'metric': 'binary', 'objective': 'binary', 'num\_round': 4000, 'learning\_rate': 0.004, 'feature\_fraction': 0.9, 'bagging\_fraction': 0.9}

The minimum is attained in round 1976

The minimum is attained in round 2137

The minimum is attained in round 2011

The minimum is attained in round 3195

The minimum is attained in round 1577

8.015999999999998 %

{'num\_thread': 12, 'num\_leaves': 12, 'metric': 'binary', 'objective': 'binary', 'num\_round': 4000, 'learning\_rate': 0.004, 'feature\_fraction': 1.0, 'bagging\_fraction': 1.0}

The minimum is attained in round 2240

The minimum is attained in round 2133

The minimum is attained in round 1830

The minimum is attained in round 2742

The minimum is attained in round 1576

8.018000000000002 %

GBDT

测试集测试：

五折结果，7.983999999999999 %

params = {'num\_thread': 12, 'num\_leaves': 12, 'metric': 'binary', 'objective': 'binary', 'num\_round': 4000, 'learning\_rate': 0.004, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}

第一种方式：# Vote

五折后的五个模型，分别得到预测结果求平均， 测试集8.204 %

第二种方式：# mean, Predict

五折后的probability 求平均后，得到预测结果测试集8.206000000000003%

正则化（测试效果不好）

L2（先粗调，0到1之间，再精调）L1（定下L2，先粗调，0到1之间，再精调）（测试）

{'num\_thread': 12, 'num\_leaves': 12, 'metric': 'binary', 'objective': 'binary', 'num\_round': 200, 'learning\_rate': 0.1, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8, 'lambda\_l1': 0.6063157894736843, 'lambda\_l2': 0.26}

The minimum is attained in round 115

The minimum is attained in round 93

The minimum is attained in round 68

The minimum is attained in round 119

The minimum is attained in round 66

7.962 %

**Hyopt**

训练记录

{'bagging\_fraction': 0.936883991179611, 'boosting': 'goss', 'device\_type': 'cpu', 'drop\_rate': 0.11852930241041167, 'extra\_trees': True, 'feature\_fraction': 0.5227806161720856, 'lambda\_l1': 4.3254509530619, 'lambda\_l2': 2.5119994536346746, 'learning\_rate': 0.0608188253086666, 'metric': 'binary\_error', 'min\_gain\_to\_split': 0.2888168506543208, 'num\_leaves': 6, 'num\_round': 2000, 'num\_threads': 12, 'objective': 'binary', 'uniform\_drop': False}

The minimum is attained in round 730

The minimum is attained in round 208

The minimum is attained in round 256

The minimum is attained in round 700

The minimum is attained in round 1749

7.847999999999999

100%|███████████████████████████████████████████████████████| 100/100 [1:06:26<00:00, 39.86s/trial, best loss: 0.07848]

测试结果

# Vote 8.177999999999997%

# mean, Predict, 8.189999999999998%

Goss

{'num\_thread': 12, 'num\_leaves': 8, 'boosting': 'goss', 'metric': 'binary', 'objective': 'binary', 'num\_round': 4000, 'learning\_rate': 0.005, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}

The minimum is attained in round 3183

The minimum is attained in round 2095

The minimum is attained in round 2771

The minimum is attained in round 3827

The minimum is attained in round 2231

7.928000000000002 %

测试结果

# Vote 8.187999999999995%

# mean, Predict, 8.206000000000003%

Dart

{'num\_thread': 12, 'num\_leaves': 8, 'boosting': 'goss', 'metric': 'binary', 'objective': 'binary', 'num\_round': 4000, 'learning\_rate': 0.005, 'feature\_fraction': 0.8, 'bagging\_fraction': 0.8}

The minimum is attained in round 3183

The minimum is attained in round 2095

The minimum is attained in round 2771

The minimum is attained in round 3827

The minimum is attained in round 2231

7.928000000000002 %

测试结果

# Vote 8.260000000000002%

# mean, Predict, 8.206000000000003%

Merge（测试结果）

模型类别

1. gbdt
2. hyperopt
3. gbdt加正则化（没啥用）
4. Goss
5. Dart

测试结果

124

# Vote 8.199999999999996%

# mean, Predict, 8.164000000000005%

1234

# Vote 8.186000000000003%

# mean, Predict, 8.179999999999998%

14

# Vote 8.192%

# mean, Predict, 8.196000000000003%

145

# Vote 8.179999999999998%

# mean, Predict, 8.153999999999995%

1245

# Vote 8.16%

# mean, Predict, 8.17%