

参数说明：

- 1, 目前采用5折交叉验证
- 2, 把-90°到90°均分成 9/12/15/18 个 bins
- 3, 放缩的层数分成 4/5/6/7
- 4, 在交叉验证时, 目前设定的随机种子是7
- 5, 随机森林和决策树中的随机种子目前也设定为7

分类器说明：

sklearn 中找不到现成的 SMO 和 BP神经网络的代码 我目前没有用这两个分类器
确认一下是不是一定需要这两个分类器 如果需要我再想办法实现

我用的五个分类器的顺序是（与下面的结果对应）：

svm, random forest, decision tree, naive byes, logistic regression

评价标准说明：

根据 Acc & Sensitivity & Specificity & AUC, 我简化为 accuracy, recall, auc , 貌似这三个经常一块儿出现比较科学。

result:

1, Gastritis_Sub 和 Normal_Sub

- bins = 9

(1) 层数 scale = 4

dataset shape: (317, 63)

accuracy:

0.766594662058

0.962147977471

0.958974974398

0.962299987199

0.890023361495

recall:

1.0

0.959013605442

0.979506802721

0.959013605442

0.922193877551

roc_auc:

0.5

0.993513321995

0.935467687075

0.981825396825

0.935504535147

(2) 层数 scale = 5

dataset shape: (317, 81)

accuracy:

0.766594662058

0.962197580645

0.940024961598

0.962299987199

0.915324180748

recall:

```
1.0
0.963180272109
0.971258503401
0.959013605442
0.926360544218
roc_auc:
0.574761904762
0.994642857143
0.904676870748
0.981281179138
0.953032879819
```

(3) 层数 scale = 6

```
accuracy:
0.766594662058
0.955848374296
0.949348758321
0.962299987199
0.886847158218
recall:
1.0
0.967176870748
0.975340136054
0.959013605442
0.914115646259
roc_auc:
0.749387755102
0.993238378685
0.919574829932
0.980192743764
0.940022675737
```

(4) 层数 scale = 7

```
accuracy:
0.766594662058
0.968498783922
0.936698348694
0.962299987199
0.899398361495
recall:
1.0
0.967176870748
0.963010204082
0.959013605442
0.922193877551
roc_auc:
0.862886904762
0.993655045351
0.906267006803
0.987403628118
0.954810090703
```

- bins = 12

(1) 层数 scale = 4

dataset shape: (317, 84)

```
accuracy:
    0.766594662058
    0.968397977471
    0.95595078085
    0.959174987199
    0.902422555044
recall:
    1.0
    0.975425170068
    0.967261904762
    0.959013605442
    0.92193877551
roc_auc:
    0.5
    0.990666099773
    0.943154761905
    0.981825396825
    0.941896258503
```

(2) 层数 scale = 5

```
dataset shape: (317, 108)
accuracy:
    0.766594662058
    0.949598374296
    0.949499167947
    0.959174987199
    0.899398361495
recall:
    1.0
    0.959013605442
    0.967091836735
    0.959013605442
    0.922108843537
roc_auc:
    0.67619047619
    0.991736111111
    0.929260204082
    0.981825396825
    0.955765306122
```

(3) 层数 scale = 6

```
dataset shape: (317, 132)
accuracy:
    0.766594662058
    0.968498783922
    0.959125384025
    0.959174987199
    0.902523361495
recall:
    1.0
    0.971343537415
    0.967176870748
    0.959013605442
    0.918112244898
roc_auc:
    0.828520408163
```

0.99200255102
0.950255102041
0.981255668934
0.94068877551

(4) 层数 scale = 7

dataset shape: (317, 156)

accuracy:

0.766594662058
0.965474590374
0.952574564772
0.962299987199
0.899348758321

recall:

1.0
0.967261904762
0.979421768707
0.959013605442
0.92619047619

roc_auc:

0.88931122449
0.989866780045
0.922091836735
0.981825396825
0.9504138322

- bins = 15

(1) 层数 scale = 4

dataset shape: (317, 105)

accuracy:

0.766594662058
0.971473374296
0.974799987199
0.959174987199
0.91507296467

recall:

1.0
0.975340136054
0.987755102041
0.959013605442
0.930272108844

roc_auc:

0.580952380952
0.993112244898
0.960068027211
0.981553287982
0.950130385488

(2) 层数 scale = 5

dataset shape: (317, 135)

accuracy:

0.766594662058
0.965223374296
0.965276177675
0.959174987199
0.899499167947

recall:

```
1.0
0.971258503401
0.983588435374
0.959013605442
0.930357142857
roc_auc:
0.757142857143
0.993224206349
0.944651360544
0.981825396825
0.960643424036
```

(3) 层数 scale = 6

```
dataset shape: (317, 165)
accuracy:
0.766594662058
0.971522977471
0.962251984127
0.959174987199
0.893298771121
recall:
1.0
0.975340136054
0.975425170068
0.959013605442
0.914115646259
roc_auc:
0.86825255102
0.995426587302
0.947712585034
0.981825396825
0.942083333333
```

(4) 层数 scale = 7

```
dataset shape: (317, 195)
accuracy:
0.766594662058
0.962098374296
0.940024961598
0.962299987199
0.91819796467
recall:
1.0
0.967176870748
0.963010204082
0.959013605442
0.938520408163
roc_auc:
0.909447278912
0.992777777778
0.913886054422
0.981825396825
0.957267573696
```

- bins = 18

(1) 层数 scale = 4

dataset shape: (317, 126)

accuracy:

0.766594662058
0.965373783922
0.978024193548
0.959174987199
0.902372951869

recall:

1.0
0.975425170068
0.983673469388
0.959013605442
0.909778911565

roc_auc:

0.620952380952
0.985673185941
0.971360544218
0.981553287982
0.946366213152

(2) 层数 scale = 5

dataset shape: (317, 162)

accuracy:

0.766594662058
0.952673771121
0.962048771121
0.959174987199
0.896526177675

recall:

1.0
0.954846938776
0.979506802721
0.959013605442
0.909863945578

roc_auc:

0.789039115646
0.994068877551
0.941658163265
0.982097505669
0.95193877551

(3) 层数 scale = 6

dataset shape: (317, 198)

accuracy:

0.766594662058
0.949548771121
0.962149577573
0.962299987199
0.902824180748

recall:

1.0
0.963095238095
0.979506802721
0.959013605442

```
0.910034013605
roc_auc:
0.909472789116
0.992250566893
0.942134353741
0.981553287982
0.94214569161
```

(4) 层数 scale = 7

```
dataset shape: (317, 234)
accuracy:
0.766594662058
0.955848374296
0.939872951869
0.962299987199
0.908723758321
recall:
1.0
0.963095238095
0.975425170068
0.959013605442
0.922108843537
roc_auc:
0.901798469388
0.99310515873
0.898664965986
0.981825396825
0.958738662132
```

结果实在太多了， 这才第一个病例做完， 你先看看并先找周老师商量看是不是最终确定这样跑数据， 不然废了九牛二虎之力跑完后要改个小地方又得全部重新跑， 太浪费时间了。