# OCTA相关数据的糖肾分类

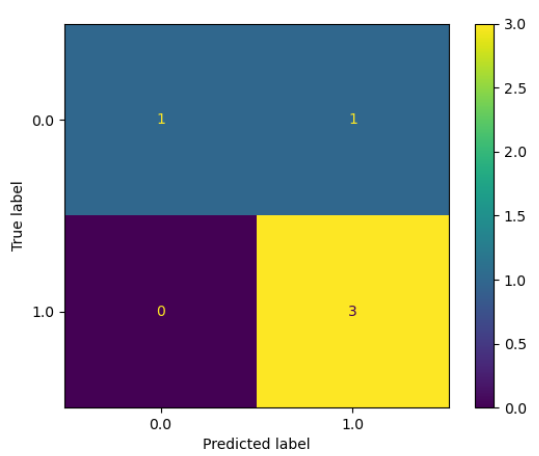
**部分特征分类**

使用全部特征时的分类结果波动很大，所以利用特征选择对特征排序得到了十二个相对重要的特征，分别是G1S、G2S、G2D、G7D、I2D、I4D、G4S、O1S、G1D、G2D、G7D、I4D

。以网格搜索法取最优参数，并以最好结果为例。在糖肾训练集上准确率为0.975，在糖肾测试集准确率为 0.8。得到的分类结果如下表：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 精确度 | 召回率 | F1分数 | 测试样本数量 |
| 无糖肾 | 1.00 | 0.50 | 0.67 | 2 |
| 糖肾 | 0.75 | 1.00 | 0.86 | 3 |

混淆矩阵结果如下图：



多因素方差分析的结果如下表：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | df | sum\_sq | mean\_sq | F | PR(>F) | 特征 |
| A1 | 1.0 | 0.316597 | 0.316597 | 3.951246 | 0.055450 |  |
| A2 | 1.0 | 0.548903 | 0.548903 | 6.850514 | 0.013420 | G2S |
| A3 | 1.0 | 0.272641 | 0.272641 | 3.402656 | 0.074366 |  |
| A4 | 1.0 | 0.044623 | 0.044623 | 0.556909 | 0.460956 |  |
| A5 | 1.0 | 2.365334 | 2.365334 | 29.520242 | 0.000006 | I2D |
| A6 | 1.0 | 0.029987 | 0.029987 | 0.374247 | 0.545021 |  |
| A7 | 1.0 | 0.002072 | 0.002072 | 0.025857 | 0.873261 |  |
| A8 | 1.0 | 2.792648 | 2.792648 | 34.853277 | 0.000001 | O1S |
| A9 | 1.0 | 0.049883 | 0.049883 | 0.622553 | 0.435907 |  |
| A10 | 1.0 | 0.241701 | 0.241701 | 3.016524 | 0.092039 |  |
| A11 | 1.0 | 0.147810 | 0.147810 | 1.844728 | 0.183897 |  |
| A12 | 1.0 | 0.623775 | 0.623775 | 7.784944 | 0.008805 | I4D |

方差分析显示G2S、I2D、O1S、I4D对分类结果有显著影响。然后去除这4个特征，利用剩下8个特征分类时会经常出现比较差的情况，而只用G2S、I2D、O1S、I4D4个特征的分类结果则表现不差。

**总结**

由于数据集是随机划分，所以即使参数相同效果也有时不同，多次实验后基本可以确定G2S、I2D、O1S、I4D等4个特征重要，对分类结果的影响显著。