**江西师范大学数字产业学院学生实验报告**

专业 人工智能 姓名 杨志坚 学号 202241600171 日期 2022-12-17

## 实验目的

掌握使用svm和bayes方法进行男女性别的分类

## 实验原理和内容

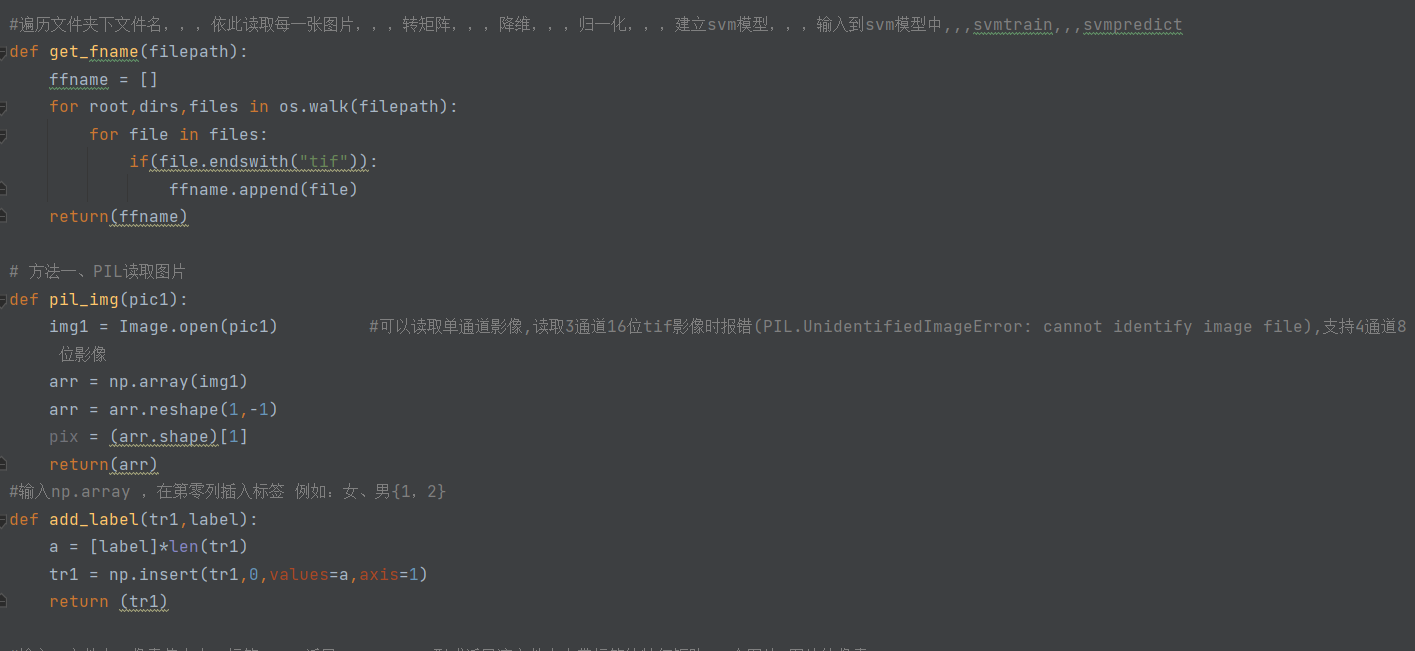
1、Svm和bayes原理

2、使用lisvm库设置进行svm分类，运用bayes原理进行分类

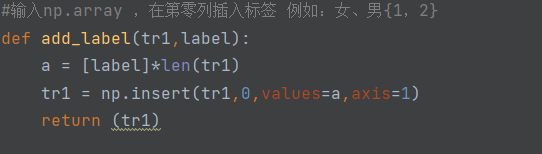
3、设置不同的参数，对比前三十个实验结果

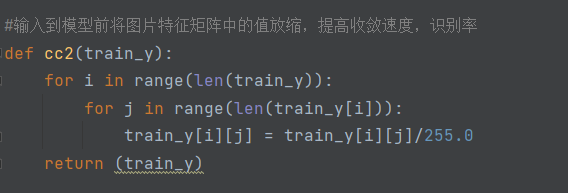
## 实验步骤

### 步骤1：libsvm

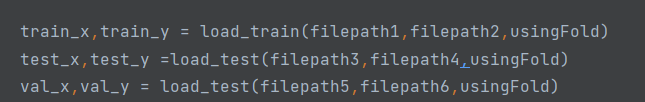
处理图片，将图片转为数组

添加标签

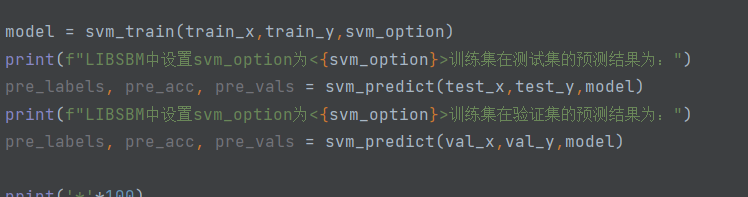




标准化和归一化



建立svm模型

将处理完的数据输入到模型中

使用训练的模型进行预测

预测结果：

在16\_20的图片中，使用中的训练集为tr\_01，使用的测试集为ts\_01，使用的验证集为val\_01

LIBSBM中设置svm\_option为<-t 1 -g 1>训练集在测试集的预测结果为：

Accuracy = 89.4737% (119/133) (classification)

LIBSBM中设置svm\_option为<-t 1 -g 1>训练集在验证集的预测结果为：

Accuracy = 90.9774% (121/133) (classification)

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

LIBSBM中设置svm\_option为<-t 1 -g 2>训练集在测试集的预测结果为：

Accuracy = 89.4737% (119/133) (classification)

LIBSBM中设置svm\_option为<-t 1 -g 2>训练集在验证集的预测结果为：

Accuracy = 90.9774% (121/133) (classification)

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

LIBSBM中设置svm\_option为<-t 1 -g 3>训练集在测试集的预测结果为：

Accuracy = 89.4737% (119/133) (classification)

LIBSBM中设置svm\_option为<-t 1 -g 3>训练集在验证集的预测结果为：

Accuracy = 90.9774% (121/133) (classification)

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

LIBSBM中设置svm\_option为<-t 2 -c 1 -g 1 >训练集在测试集的预测结果为：

Accuracy = 87.9699% (117/133) (classification)

LIBSBM中设置svm\_option为<-t 2 -c 1 -g 1 >训练集在验证集的预测结果为：

Accuracy = 84.2105% (112/133) (classification)

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

LIBSBM中设置svm\_option为<-t 2 -c 10 -g 1 >训练集在测试集的预测结果为：

Accuracy = 85.7143% (114/133) (classification)

LIBSBM中设置svm\_option为<-t 2 -c 10 -g 1 >训练集在验证集的预测结果为：

Accuracy = 81.9549% (109/133) (classification)

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

LIBSBM中设置svm\_option为<-t 2 -c 100 -g 1 >训练集在测试集的预测结果为：

Accuracy = 85.7143% (114/133) (classification)

LIBSBM中设置svm\_option为<-t 2 -c 100 -g 1 >训练集在验证集的预测结果为：

Accuracy = 81.9549% (109/133) (classification)

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

LIBSBM中设置svm\_option为<-t 2 -c 1 -g 0.1 >训练集在测试集的预测结果为：

Accuracy = 88.7218% (118/133) (classification)

LIBSBM中设置svm\_option为<-t 2 -c 1 -g 0.1 >训练集在验证集的预测结果为：

Accuracy = 88.7218% (118/133) (classification)

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

LIBSBM中设置svm\_option为<-t 2 -c 10 -g 0.1 >训练集在测试集的预测结果为：

Accuracy = 90.9774% (121/133) (classification)

LIBSBM中设置svm\_option为<-t 2 -c 10 -g 0.1 >训练集在验证集的预测结果为：

Accuracy = 90.2256% (120/133) (classification)

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

LIBSBM中设置svm\_option为<-t 2 -c 100 -g 0.1 >训练集在测试集的预测结果为：

Accuracy = 92.4812% (123/133) (classification)

LIBSBM中设置svm\_option为<-t 2 -c 100 -g 0.1 >训练集在验证集的预测结果为：

Accuracy = 90.9774% (121/133) (classification)

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

LIBSBM中设置svm\_option为<-t 2 -c 1 -g 0.01 >训练集在测试集的预测结果为：

Accuracy = 85.7143% (114/133) (classification)

LIBSBM中设置svm\_option为<-t 2 -c 1 -g 0.01 >训练集在验证集的预测结果为：

Accuracy = 88.7218% (118/133) (classification)

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

LIBSBM中设置svm\_option为<-t 2 -c 10 -g 0.01 >训练集在测试集的预测结果为：

Accuracy = 88.7218% (118/133) (classification)

LIBSBM中设置svm\_option为<-t 2 -c 10 -g 0.01 >训练集在验证集的预测结果为：

Accuracy = 90.2256% (120/133) (classification)

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

LIBSBM中设置svm\_option为<-t 2 -c 100 -g 0.01 >训练集在测试集的预测结果为：

Accuracy = 89.4737% (119/133) (classification)

LIBSBM中设置svm\_option为<-t 2 -c 100 -g 0.01 >训练集在验证集的预测结果为：

Accuracy = 93.2331% (124/133) (classification)

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

LIBSBM中设置svm\_option为<-t 2 -c 1000 -g 1 >训练集在测试集的预测结果为：

Accuracy = 85.7143% (114/133) (classification)

LIBSBM中设置svm\_option为<-t 2 -c 1000 -g 1 >训练集在验证集的预测结果为：

Accuracy = 81.9549% (109/133) (classification)

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

LIBSBM中设置svm\_option为<-t 2 -c 1000 -g 0.1 >训练集在测试集的预测结果为：

Accuracy = 92.4812% (123/133) (classification)

LIBSBM中设置svm\_option为<-t 2 -c 1000 -g 0.1 >训练集在验证集的预测结果为：

Accuracy = 90.9774% (121/133) (classification)

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

LIBSBM中设置svm\_option为<-t 2 -c 1000 -g 0.01 >训练集在测试集的预测结果为：

Accuracy = 90.2256% (120/133) (classification)

LIBSBM中设置svm\_option为<-t 2 -c 1000 -g 0.01 >训练集在验证集的预测结果为：

Accuracy = 94.7368% (126/133) (classification)

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

**在48\_60的图片中，使用中的训练集为tr\_01，使用的测试集为ts\_01，使用的验证集为val\_01**

**LIBSBM中设置svm\_option为<-t 1 -g 1>训练集在测试集的预测结果为：**

**Accuracy = 90.9774% (121/133) (classification)**

**LIBSBM中设置svm\_option为<-t 1 -g 1>训练集在验证集的预测结果为：**

**Accuracy = 92.4812% (123/133) (classification)**

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**LIBSBM中设置svm\_option为<-t 1 -g 2>训练集在测试集的预测结果为：**

**Accuracy = 90.9774% (121/133) (classification)**

**LIBSBM中设置svm\_option为<-t 1 -g 2>训练集在验证集的预测结果为：**

**Accuracy = 92.4812% (123/133) (classification)**

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**LIBSBM中设置svm\_option为<-t 1 -g 3>训练集在测试集的预测结果为：**

**Accuracy = 90.9774% (121/133) (classification)**

**LIBSBM中设置svm\_option为<-t 1 -g 3>训练集在验证集的预测结果为：**

**Accuracy = 92.4812% (123/133) (classification)**

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**LIBSBM中设置svm\_option为<-t 2 -c 1 -g 1 >训练集在测试集的预测结果为：**

**Accuracy = 43.609% (58/133) (classification)**

**LIBSBM中设置svm\_option为<-t 2 -c 1 -g 1 >训练集在验证集的预测结果为：**

**Accuracy = 54.8872% (73/133) (classification)**

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**LIBSBM中设置svm\_option为<-t 2 -c 10 -g 1 >训练集在测试集的预测结果为：**

**Accuracy = 43.609% (58/133) (classification)**

**LIBSBM中设置svm\_option为<-t 2 -c 10 -g 1 >训练集在验证集的预测结果为：**

**Accuracy = 54.8872% (73/133) (classification)**

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**LIBSBM中设置svm\_option为<-t 2 -c 100 -g 1 >训练集在测试集的预测结果为：**

**Accuracy = 43.609% (58/133) (classification)**

**LIBSBM中设置svm\_option为<-t 2 -c 100 -g 1 >训练集在验证集的预测结果为：**

**Accuracy = 54.8872% (73/133) (classification)**

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**LIBSBM中设置svm\_option为<-t 2 -c 1 -g 0.1 >训练集在测试集的预测结果为：**

**Accuracy = 56.391% (75/133) (classification)**

**LIBSBM中设置svm\_option为<-t 2 -c 1 -g 0.1 >训练集在验证集的预测结果为：**

**Accuracy = 70.6767% (94/133) (classification)**

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**LIBSBM中设置svm\_option为<-t 2 -c 10 -g 0.1 >训练集在测试集的预测结果为：**

**Accuracy = 60.1504% (80/133) (classification)**

**LIBSBM中设置svm\_option为<-t 2 -c 10 -g 0.1 >训练集在验证集的预测结果为：**

**Accuracy = 75.188% (100/133) (classification)**

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**LIBSBM中设置svm\_option为<-t 2 -c 100 -g 0.1 >训练集在测试集的预测结果为：**

**Accuracy = 60.1504% (80/133) (classification)**

**LIBSBM中设置svm\_option为<-t 2 -c 100 -g 0.1 >训练集在验证集的预测结果为：**

**Accuracy = 75.188% (100/133) (classification)**

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**LIBSBM中设置svm\_option为<-t 2 -c 1 -g 0.01 >训练集在测试集的预测结果为：**

**Accuracy = 89.4737% (119/133) (classification)**

**LIBSBM中设置svm\_option为<-t 2 -c 1 -g 0.01 >训练集在验证集的预测结果为：**

**Accuracy = 88.7218% (118/133) (classification)**

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**LIBSBM中设置svm\_option为<-t 2 -c 10 -g 0.01 >训练集在测试集的预测结果为：**

**Accuracy = 90.9774% (121/133) (classification)**

**LIBSBM中设置svm\_option为<-t 2 -c 10 -g 0.01 >训练集在验证集的预测结果为：**

**Accuracy = 88.7218% (118/133) (classification)**

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**LIBSBM中设置svm\_option为<-t 2 -c 100 -g 0.01 >训练集在测试集的预测结果为：**

**Accuracy = 90.9774% (121/133) (classification)**

**LIBSBM中设置svm\_option为<-t 2 -c 100 -g 0.01 >训练集在验证集的预测结果为：**

**Accuracy = 88.7218% (118/133) (classification)**

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**LIBSBM中设置svm\_option为<-t 2 -c 1000 -g 1 >训练集在测试集的预测结果为：**

**Accuracy = 43.609% (58/133) (classification)**

**LIBSBM中设置svm\_option为<-t 2 -c 1000 -g 1 >训练集在验证集的预测结果为：**

**Accuracy = 54.8872% (73/133) (classification)**

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**LIBSBM中设置svm\_option为<-t 2 -c 1000 -g 0.1 >训练集在测试集的预测结果为：**

**Accuracy = 60.1504% (80/133) (classification)**

**LIBSBM中设置svm\_option为<-t 2 -c 1000 -g 0.1 >训练集在验证集的预测结果为：**

**Accuracy = 75.188% (100/133) (classification)**

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**LIBSBM中设置svm\_option为<-t 2 -c 1000 -g 0.01 >训练集在测试集的预测结果为：**

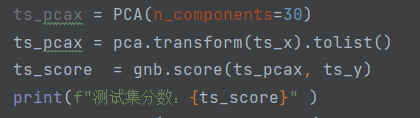
**Accuracy = 90.9774% (121/133) (classification)**

**LIBSBM中设置svm\_option为<-t 2 -c 1000 -g 0.01 >训练集在验证集的预测结果为：**

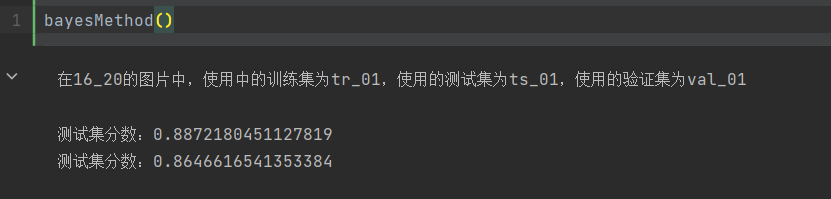
**Accuracy = 88.7218% (118/133) (classification)**

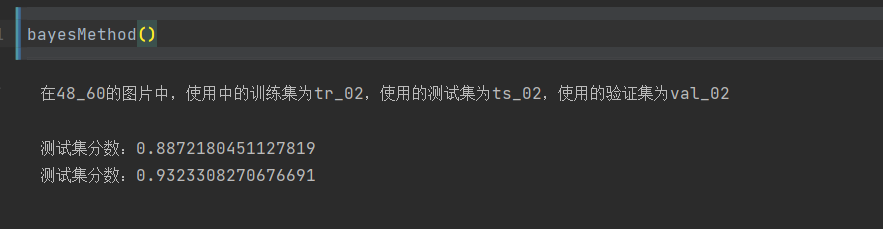
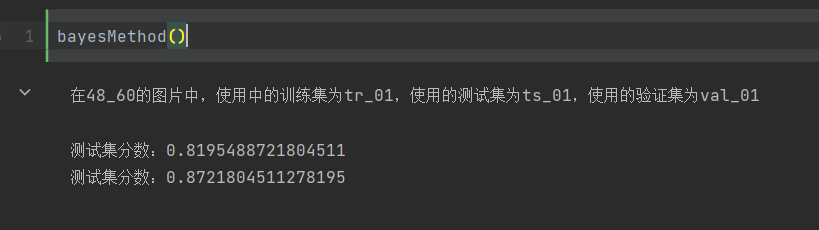
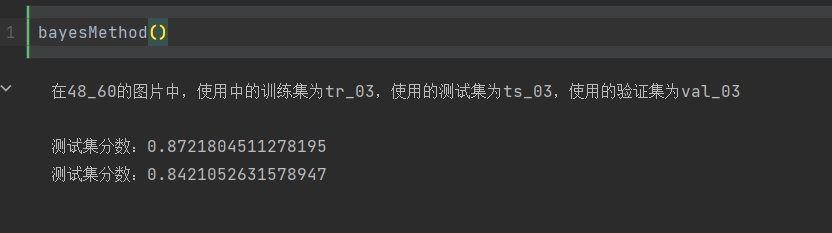
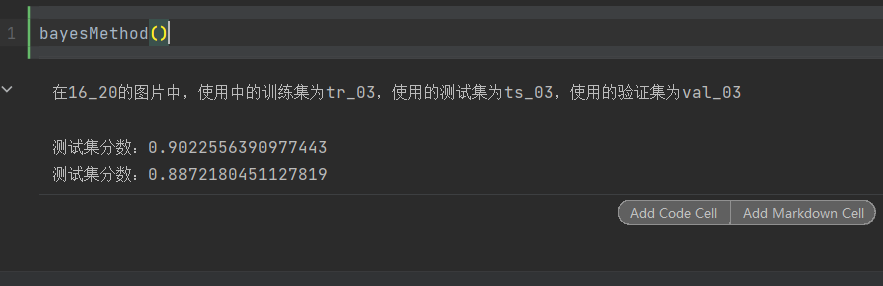
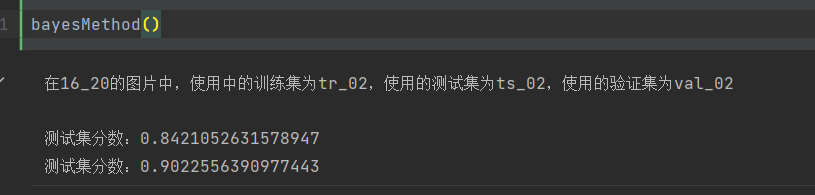
### 步骤2：bayes

使用pca降至30维，并sklearn建立贝叶斯高斯模型进行预测









### 步骤3：对比试验结果

通过对比libsvm和bayes的实验结果，我们发现当使用libsvm进行预测时，svm方法比较依赖选择正确的核函数，并且不同参数配置对模型预测结果影响比较大。当使用bayes进行预测时，需要对数据进行合理处理使其更快更具效率。两种方法的预测结果都能达到80%-90%左右正确率，bayes相对svm更稳定，svm选择正确的核函数并配置合理参数的预测结果也不逊色于bayes。

## 实验总结：

通过本次实验，我熟悉了使用python处理图片数据的方法，学会了使用libsvm和bayes进行预测。

附代码地址

https://github.com/yangzj6677/MLhomework/blob/main/mlhomework.ipynb