In [1]:
```
# 导入操作系统库
import os
# 更改工作目录
os.chdir(r"D:\softwares\applied statistics\pythoncodelearning\chap1\sourcecode")
# 导入绘图库
import matplotlib.pyplot as plt
# 导入数据集划分函数
from sklearn.model_selection import train_test_split
# 导入KNN分类器
from sklearn.neighbors import KNeighborsClassifier
# 导入计算分类准确率的工具
from sklearn.metrics import accuracy_score
# 导入breast_cancer数据集
from sklearn.datasets import load_breast_cancer
# 导入绘图库中的字体管理包
from matplotlib import font_manager
# 实现中文字符正常显示
font = font_manager.FontProperties(fname=r"C:\Windows\Fonts\SimKai.ttf")
# 使用seaborn风格绘图
plt.style.use("seaborn-v0_8")
# 生成数据
X, y = load_breast_cancer(return_X_y=True, as_frame=True)
x_train, x_test, y_train, y_test = train_test_split(
    X, y, # 数据
    test_size=0.3, # 测试集的数量
    random_state=1 # 随机数种子
)
# 建立分类器
knn = KNeighborsClassifier(
    n_neighbors=2 # 最近邻的数量，分几类就几个
)
# 模型拟合
res = knn.fit(x_train, y_train)
# 模型预测
y_pred = res.predict(x_test)
# 预测准确率
acc = accuracy_score(y_test, y_pred)
print("预测准确率为：", acc, sep="\n")
```

```
预测准确率为：
0.9005847953216374
```