姓名和学号：李胜志 220180232

章节名称：朴素贝叶斯算法的鸢尾花数据集分类应用

知识目标：使用贝叶斯算法对鸢尾花数据集进行分类。

能力目标：掌握朴素贝叶斯算法的sklearn方法。

素质目标：养成分析问题、事前规划的良好习惯。

知识重点：根据具体情境选择合适的朴素贝叶斯算法对鸢尾花数据集进行分类。

知识难点：根据特征变量的不同选择不同的朴素贝叶斯算法。

朴素贝叶斯算法的鸢尾花数据集分类应用：

数据分析：鸢尾花数据集中有4个特征变量，且均为连续变量，因此使用高斯朴素贝叶斯算法。

1. 导入高斯朴素贝叶斯算法GaussianNB
2. 输入训练集数据，划分特征和标签，划分训练集和测试集
3. 创建模型并训练
4. 求解模型的准确率

代码：

# 导入所需的库

from sklearn.naive\_bayes import GaussianNB

from sklearn.datasets import load\_iris

from sklearn.model\_selection import train\_test\_split

# 创建数据集

mydata = load\_iris()

# 划分特征和变量

x = mydata.data

y = mydata.target

# 划分训练集和测试集

x\_train, x\_test, y\_train, y\_test = train\_test\_split(x, y, test\_size=0.25)

# 创建模型并训练模型

gs = GaussianNB()

gs.fit(x\_train, y\_train)

print("高斯朴素贝叶斯算法的准确率是: ", gs.score(x\_test, y\_test))

运行结果：

高斯朴素贝叶斯算法的准确率是: 0.9736842105263158