

课堂练习3

姓名：骆禹松

班级：2018211129

学号：2018210071

1 实验目的

- 设计用于三分类任务的基于核的模型并编程实现。
- 使用sklearn中提供的鸢尾花数据集作为训练和测试数据。
- 通过可视化方式呈现分类的决策边界。

2 实验原理

注：本实验采用了svm库中提供的rbf核模型。

相比线性模型，基于核的非线性模型具有更灵活的非线性决策边界，其形状取决于核的类型及设置的参数。

下面对本次实验选择的sklearn.svm.SVC()函数中的参数进行说明：

```
sklearn.svm.SVC(kernel="rbf", gamma=0.7, C=1.0)
```

- 参数C是惩罚参数，C越大，对误分类的惩罚越大，趋向于对训练集全分对的情况，这样对训练集测试时准确率很高，但泛化能力弱，可能会导致模型过拟合。C越小，则对误分类的惩罚减小，泛化能力较强。
- 参数kernel是核函数，默认是rbf，还可以选择linear, poly, sigmoid和precomputed等。
- gamma是rbf、poly和sigmoid的核系数，当前设置为0.7，可以根据实际情况进行调整。

3 实验步骤

3.1 实验数据

本次实验采用鸢尾花数据集作为训练和测试数据(只取数据集的前两列，即花萼长度和花萼宽度)。

```
import numpy as np
import matplotlib.pyplot as plt
from sklearn import svm
from sklearn import datasets

iris = datasets.load_iris()
# 取前两个特征值
X = iris.data[:, :2]
Y = iris.target
```

如上代码块中，X表示提取的鸢尾花数据集，Y表示鸢尾花的种类（分别用0，1，2表示三种鸢尾花）。

3.2 创建模型

本次实验采用svm库中提供的rbf核模型。

```
# 创建分类器
clf = svm.SVC(kernel="rbf", gamma=0.7, C=1.0)
# 创建模型
model = clf.fit(X, Y)
```

3.3 生成测试数据

为检验模型效果，这里需要先生成一组测试数据。分别以鸢尾花数据集中每列数据的最大最小值作为边界生成一组等间隔数组。

```
[X_test1, X_test2] = np.meshgrid(np.arange(X0.min()-1, \
X0.max()+1, 0.02), np.arange(X1.min()-1, X1.max()+1, 0.02))
```

3.4 预测分类

使用章节3.2中创建的模型来对测试数据进行分类预测。

```
X_test_predict = clf.predict(np.c_[X_test1.ravel(), X_test2.ravel()])
X_test_predict = X_test_predict.reshape(X_test1.shape)
```

3.5 绘制图像

首先绘制提取出的鸢尾花特征数据图。

```
X0 = X[:,0]
X1 = X[:,1]
plt.scatter(X0, X1, c=Y, cmap=plt.cm.coolwarm, s=20, edgecolors="k")
plt.xlabel("Sepal Length")
plt.ylabel("Sepal Width")
```

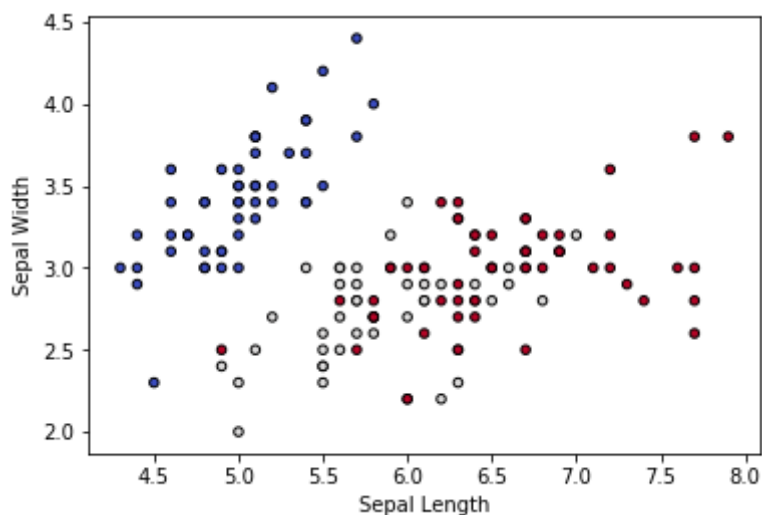
然后绘制预测的测试数据分类图，即通过可视化的形式呈现分类的决策边界。

```
plt.contourf(X_test1, X_test2, X_test_predict, cmap=plt.cm.coolwarm, alpha=0.8)
plt.xlim(X_test1.min(), X_test1.max())
plt.ylim(X_test2.min(), X_test2.max())
```

4 实验结果

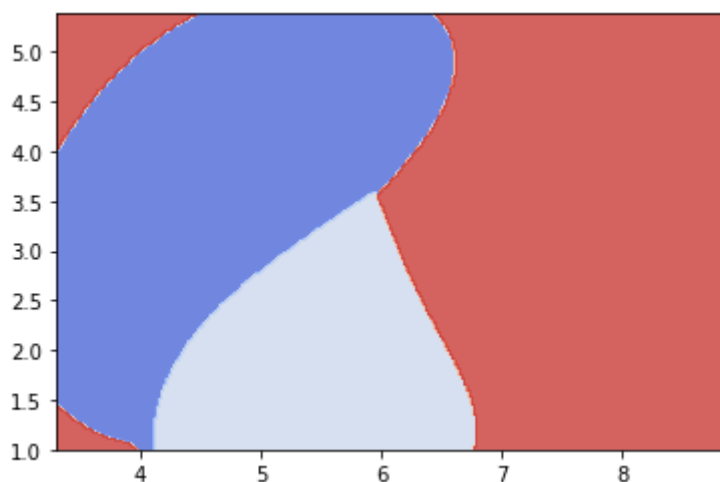
4.1 鸢尾花特征数据图

利用章节3.5中的代码可以绘制提取出的鸢尾花特征数据图，结果如下图所示：



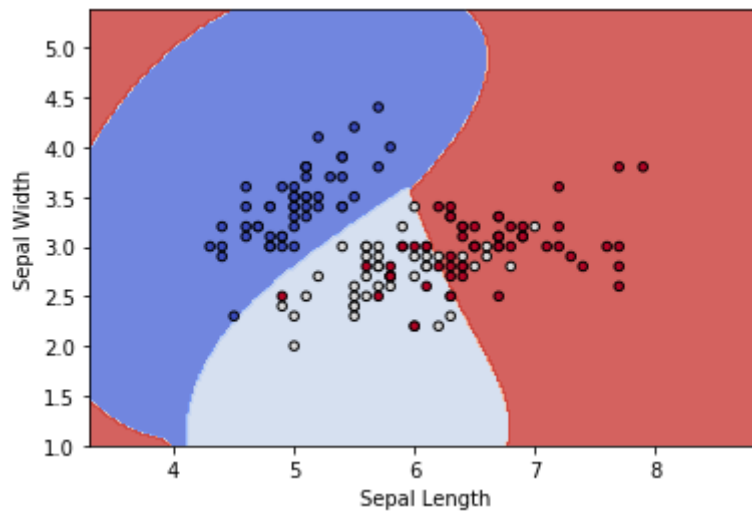
4.2 决策边界绘制

利用章节3.5中的代码可以绘制出分类的决策边界，结果如下图所示：



4.3 图像合成

将章节4.1和4.2中的图像组合在一起展示，结果如下图所示：



可以看出相比第二次作业采用的线性模型，对于鸢尾花数据集分类来说，本次实验采用的基于rbf核的非线性模型效果更好。

5 实验感悟

通过本次实验，我更深入地理解了非线性核模型，学会了使用非线性核模型进行分类决策，学会了使用Python的sklearn库中的svm模型。