使用Python实现对数几率回归模型

1. 问题描述：

基于Python编程语言，读取鸢尾花数据，并对其数据进行训练集和测试集的划分，通过对数几率回归模型进行训练，并观察模型在测试集上分类的正确率。

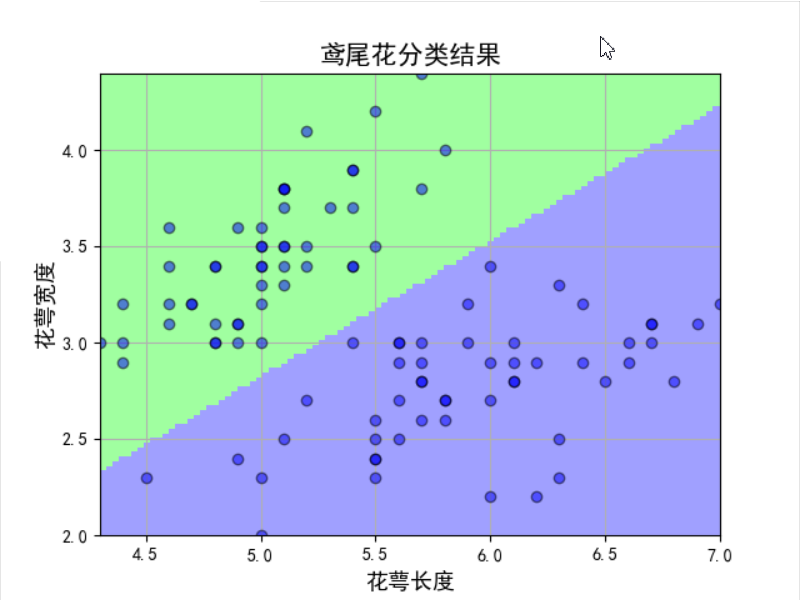
1. 数据集描述：

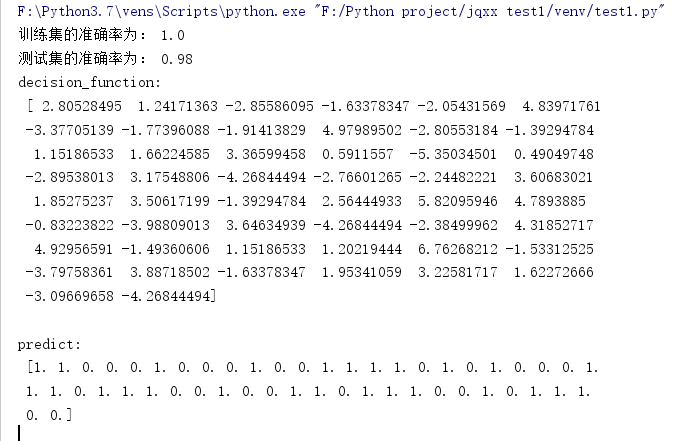
IRIS数据集以鸢尾花的特征作为数据来源，数据集包含150个数据集，分为3类，每类50个数据，每个数据包含4个属性，是在数据挖掘、数据分类中非常常用的测试集、训练集。

本次实验，取用了setosa和versicolor这两类中的花萼长度和花萼宽度这两个属性，进行训练及测试。

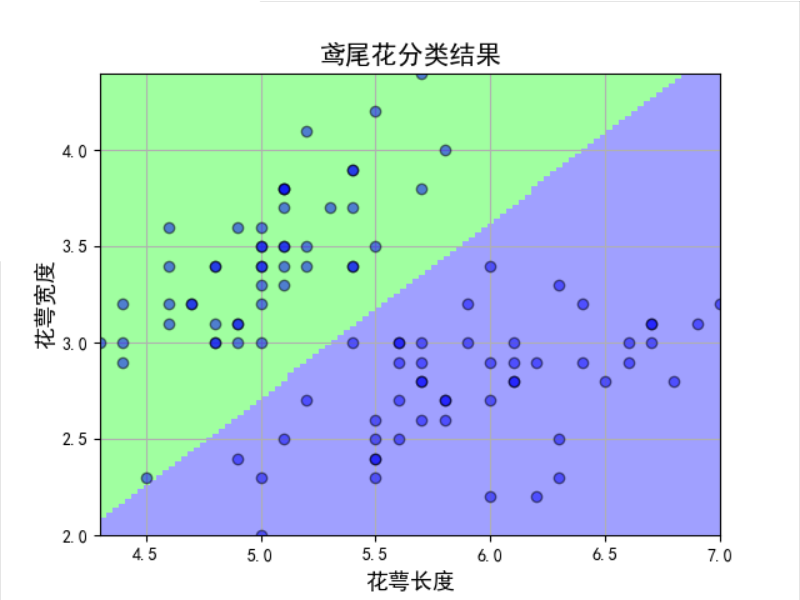
1. 实验结果图：

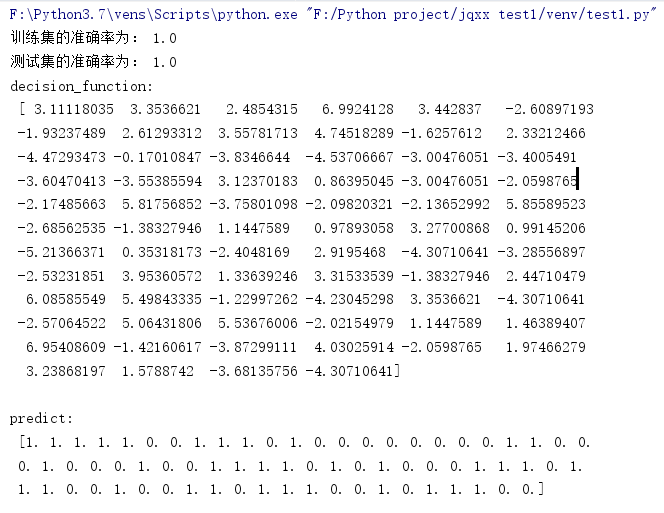
50%训练集，50%测试集：



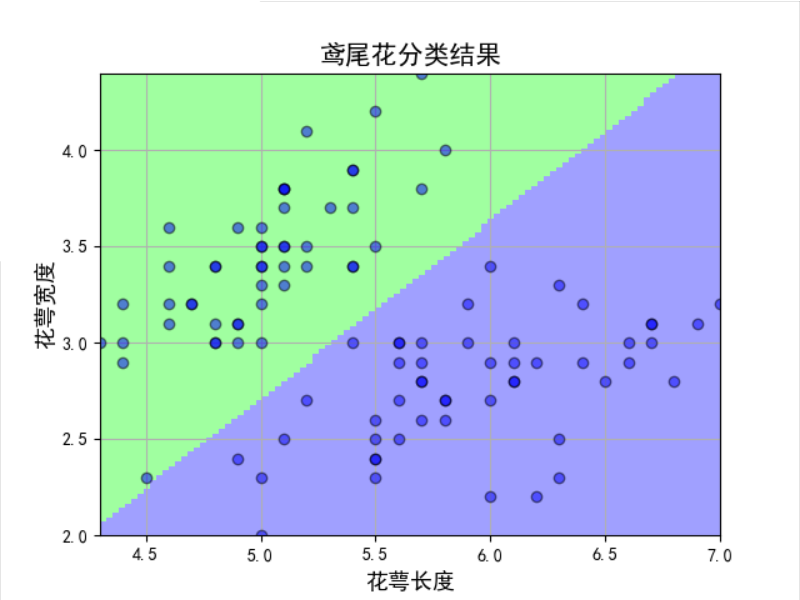


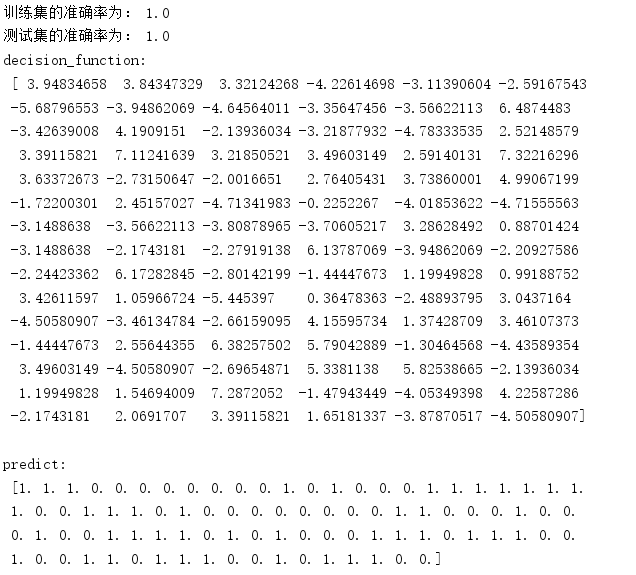
70%训练集，30%测试集：





90%训练集，10%测试集：





1. 实验结果分析：

当测试集占比越小，准确率越大。