说明文档

def init\_data() #读取和处理数据

def sigmoid(z) #sigmoid函数

def grad\_descent(dataMatIn, classLabels) #梯度下降

def plotBestFIt(weights) #可视化

def pulsarTest () #得出测试集的结果

def csv\_save(answer) # 将测试集的结果保存到csv文件里

在if \_\_name\_\_ == '\_\_main\_\_'中， 先通过init\_data()读取和处理数据，再通过grad\_descent () 用梯度下降算法对训练集进行训练，用plotBestFIt () 函数将训练集训练的结果可视化，通过pulsarTest() 得出测试集的结果，用csv\_save() 函数将得到的结果保存到csv文件里。核心代码是用梯度下降训练训练集和得出测试集的结果。

第一次测试的准确率是0.78，这时的学习率定义的是alpha = 0.01，接着修改测试率由0.02到0.06，通过可视化可以明显看到二分类的效果有进步。学习率为0.06时测试的准确率是0.80，学习率为0.02时测试的准确率是0.82，最后定学习率为0.02 。