机器学习报告

姓名:牛晋宇 班级:计算机1604 学号:1611640414

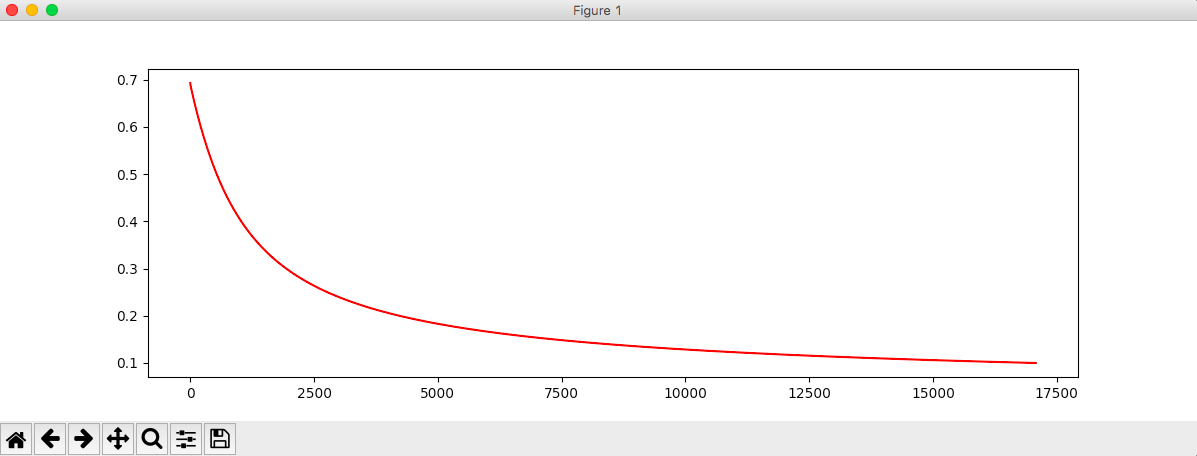
题目：使用python实现对数几率回归模型

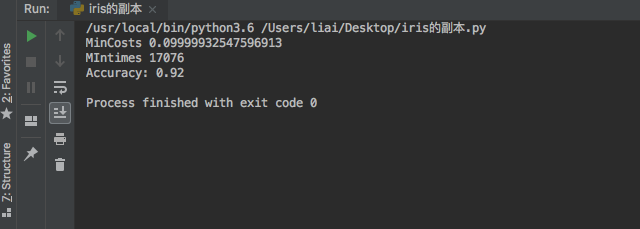
问题描述：用python构造一个回归函数来对iris数据集进行分类

数据集描述：总共150个数据，其中50个setonsa数据，50个versicolor数据，50个virginica，每个数据包含4个属性，可以通过花萼长度，花萼宽度，花瓣长度，花瓣宽度4个属性来判断出是哪一种花。进行二分类 我选取了versicolor和virginica来进行实现对数几率回归模型，并将损失值小于等于0.1来作为迭代的终止条件

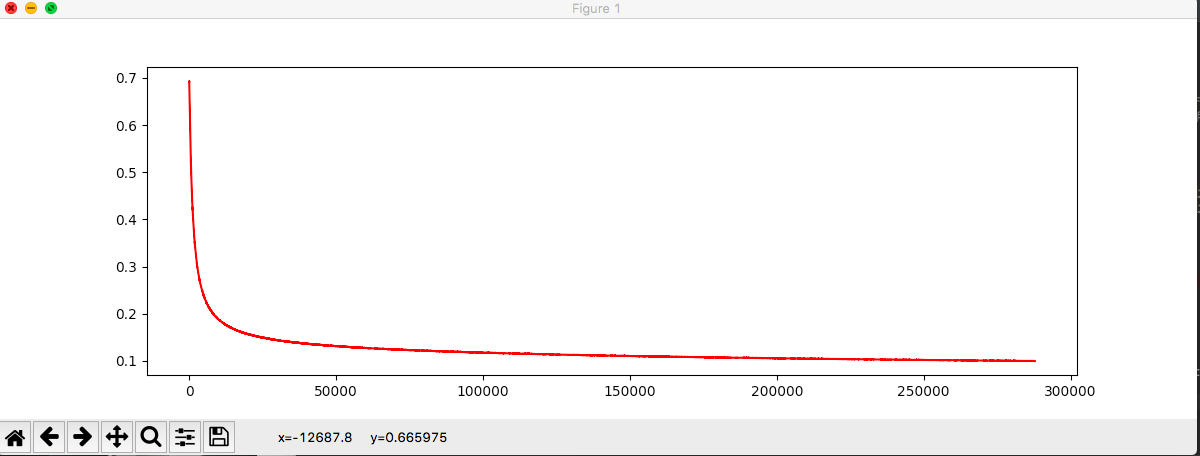
实验结果图：

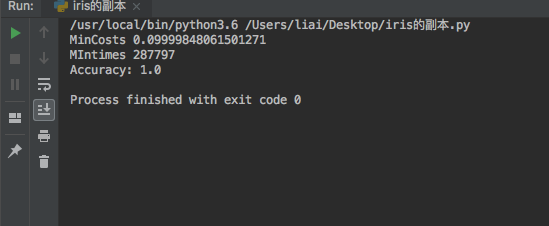
1.将数据集的50%作为训练集，50%作为测试集，

损失值和迭代次数的关系图：

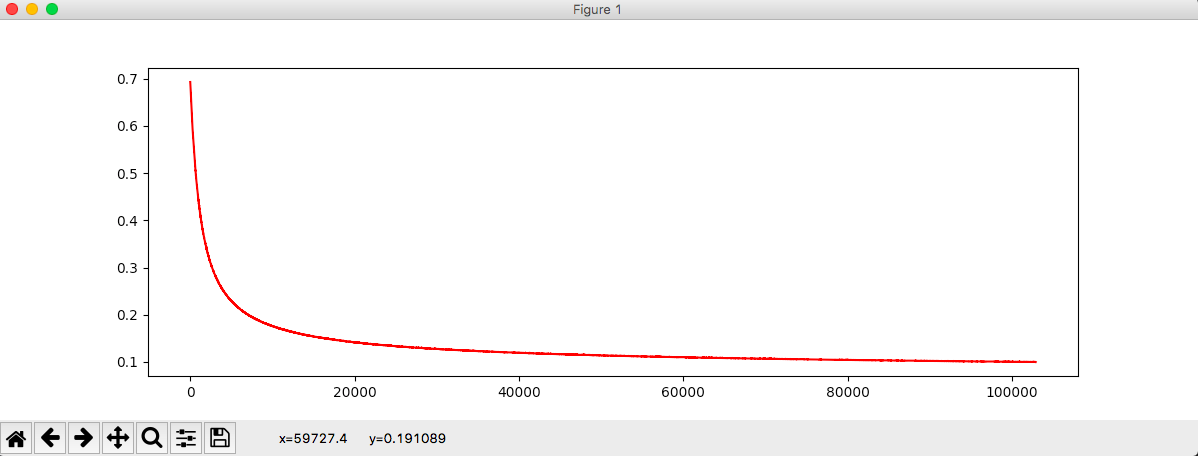
准确率，最小迭代次数，和最小损失值：

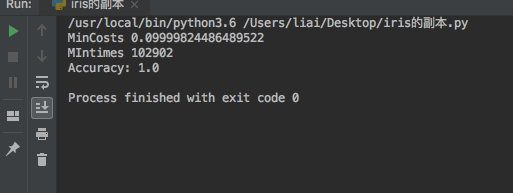
2.将数据集的70%作为训练集，30%作为测试集

损失值和迭代次数的关系图：

准确率，最小迭代次数，和最小损失值：

3.将数据集的90%作为训练集，10%作为测试集

损失值和迭代次数的关系图：

准确率，最小迭代次数，和最小损失值：

实验结果分析：当数据集用了50%来作为训练集来训练的时候，损失值只迭代了17076次就小于0.1了，但精度不行才0.92，当数据集用大约70%的训练集训练的，迭代次数达到需要287797次损失值才能小于0.1 ，精度可以达到1.0. 但精度达到1.0 可能是因为测试的数据太少了，同理90%的数据集来训练 ，精度同样可以达到1.0。所以感觉数据太少了，并没有可以发现确切的结论