信息基础2 作业

徐瑞 20180012045

Logistic Regression 分类散点

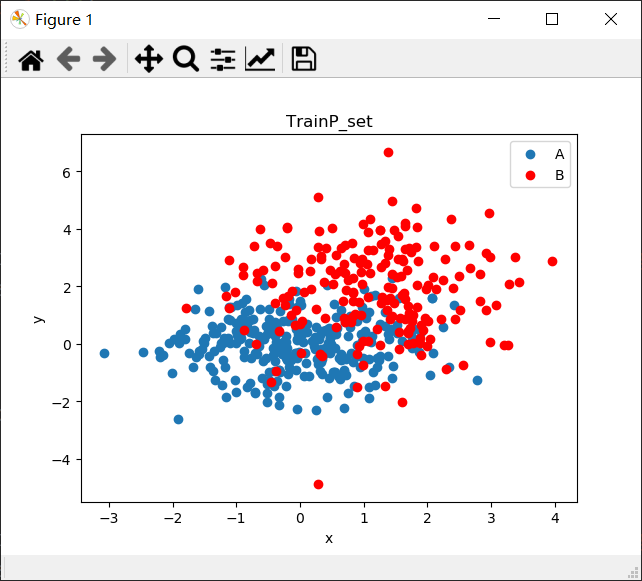
**问题**：有两类空间点，A类和B类。A 类点以(0, 0)为中心、(1, 0; 0, 1)为协方差矩阵的二维高斯分布；B 类点以(1, 2) 为中心、(1, 0; 0, 2)为协方差矩阵的二维高斯分布；随机生成300个A类点，200个B类点，并用Logistic Regression的方法进行分类(GD 和SGD)。画出不同 epoch等参数下的结果。

**答案：**我先产生了一组符合题目要求的散点作为训练集，然后我又产生了一组相同分布不同点的点集作为测试集。我使用的学习率为0.01；A类点的标签我用数值1来代替，B类点的标签我用数值0来代替；对于epoch我取了5个不同的值，分别为：100，300，500，800，1000；对于每一个epoch我都是分别用GD(梯度下降)和SGD（随机梯度下降）的方法用训练集训练模型，然后再分别对训练集和测试集进行分类。

梯度下降（GD）：每次梯度下降都使用全部的训练样本；

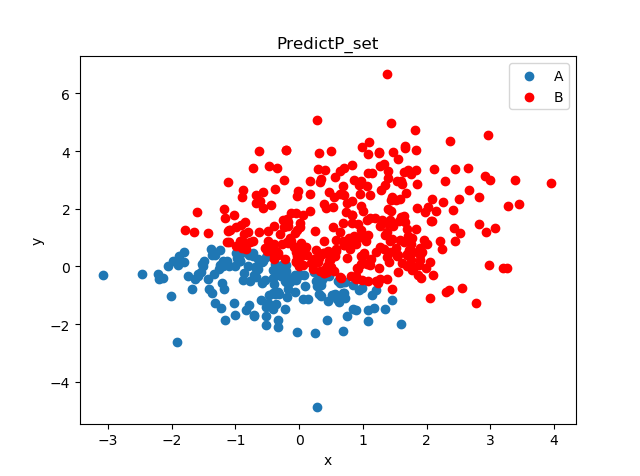
随机梯度下降（SGD）：每次梯度下降只使用一个训练样本；

**训练集分布如下：**

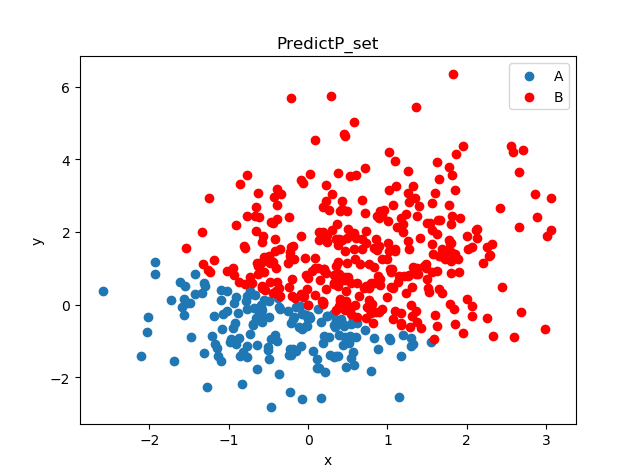


epoch = 100时：

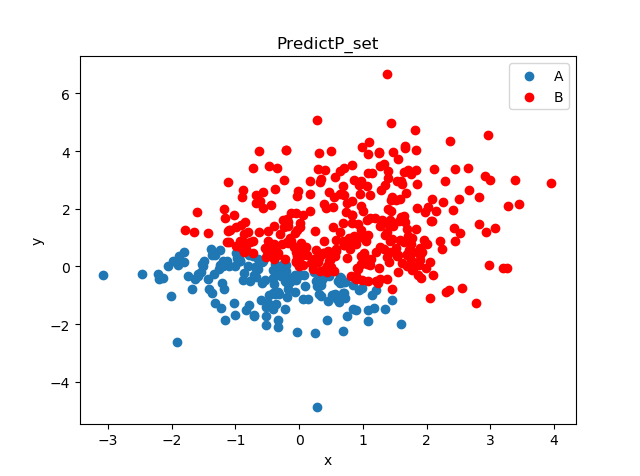
GD训练集分类结果如下：



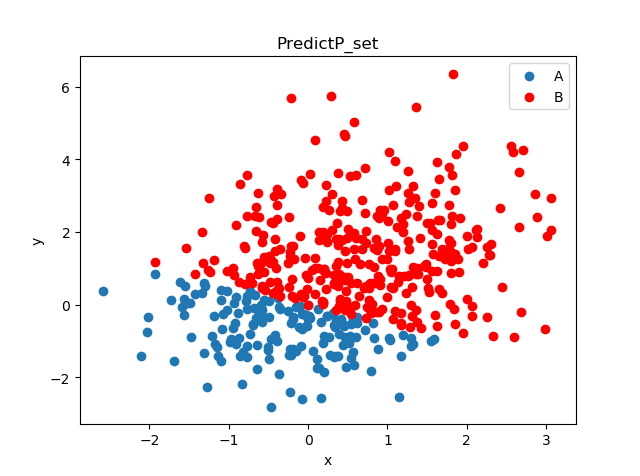
GD测试集分类结果如下：



SGD训练集分类结果如下：

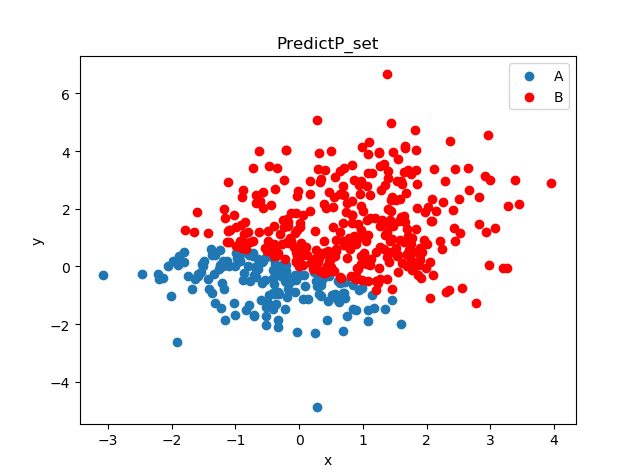


SGD测试集分类结果如下：

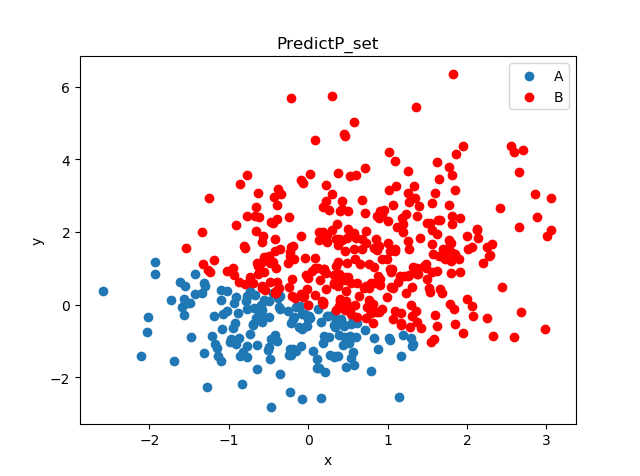


epoch = 300时：

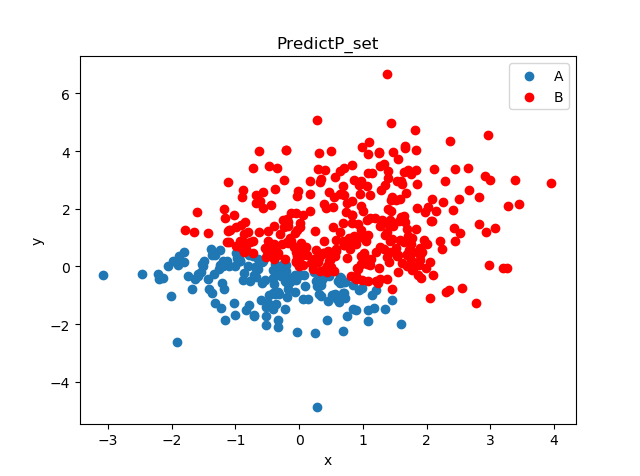
GD训练集分类结果如下：



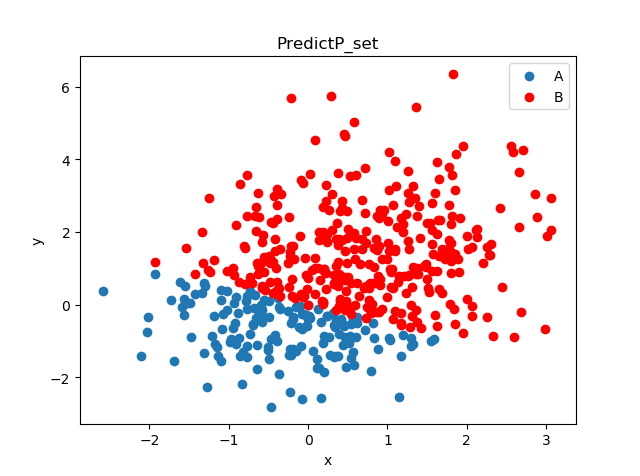
GD测试集分类结果如下：



SGD训练集分类结果如下：

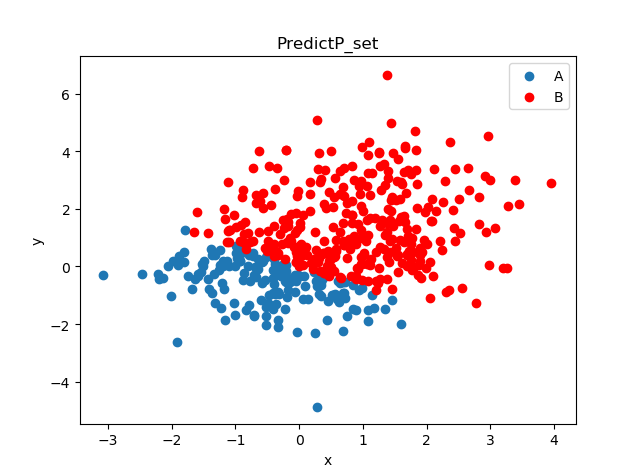


SGD测试集分类结果如下：

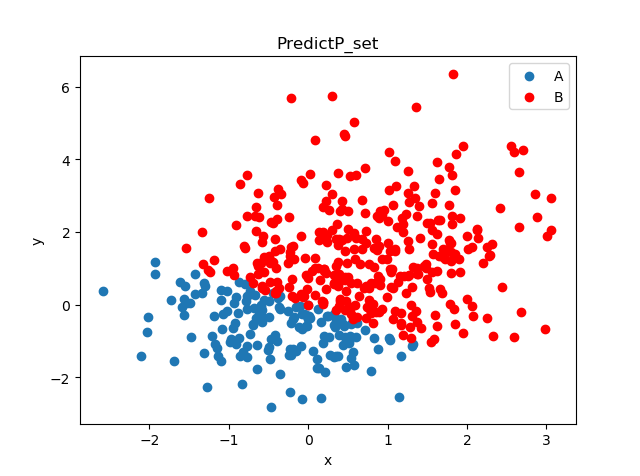


epoch = 500时：

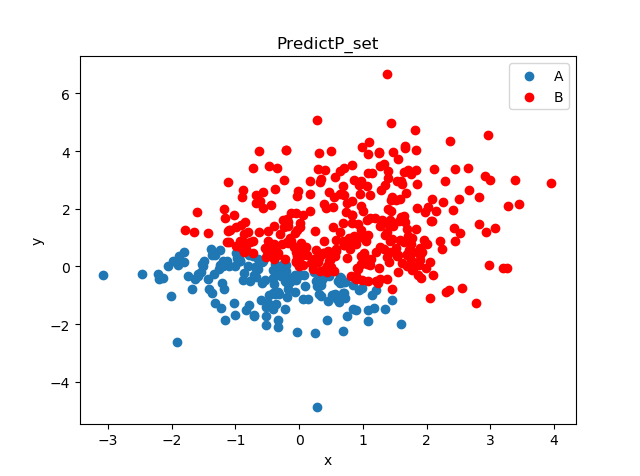
GD训练集分类结果如下：



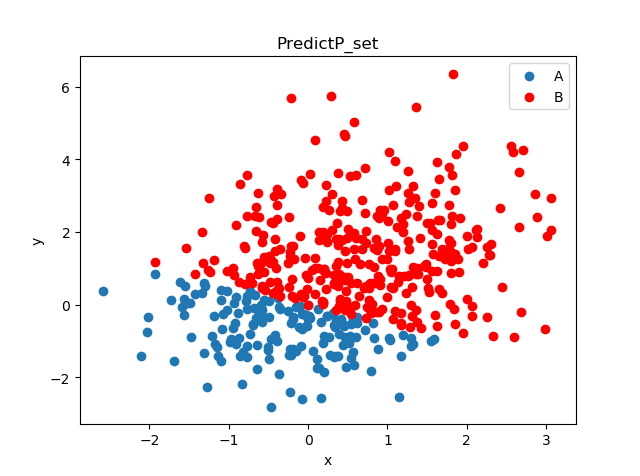
GD测试集分类结果如下：



SGD训练集分类结果如下：

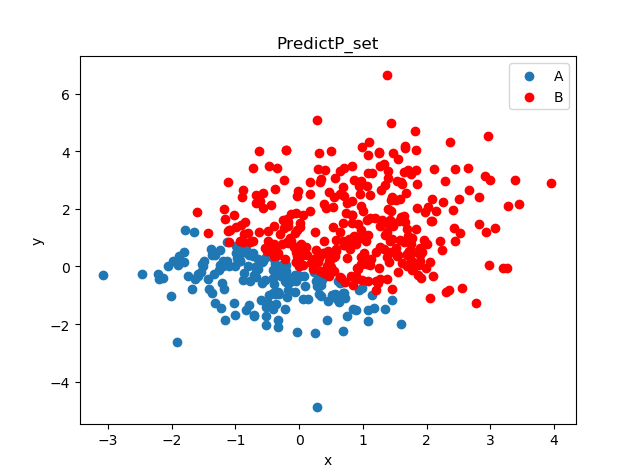


SGD测试集分类结果如下：

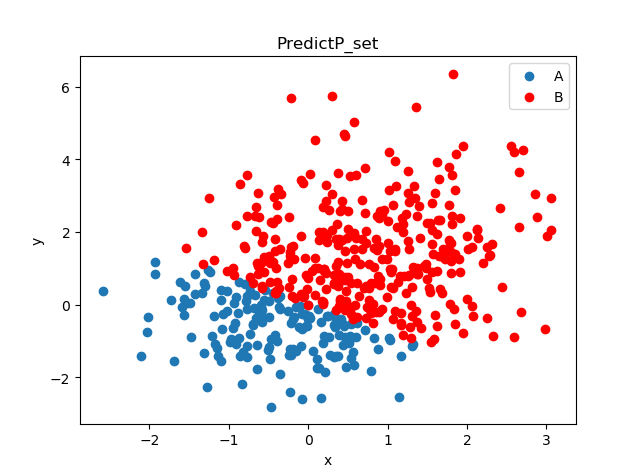


epoch = 1000时：

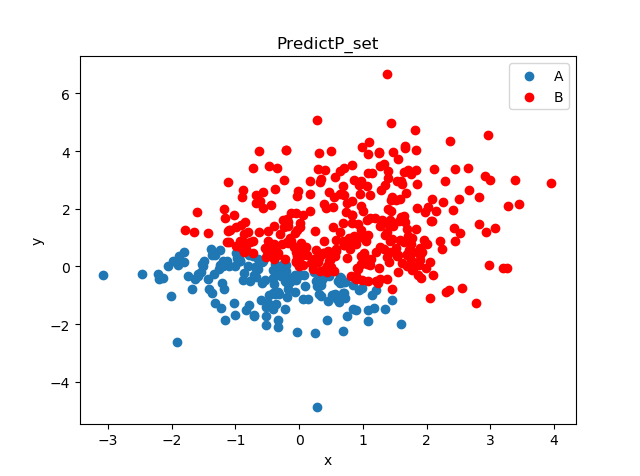
GD训练集分类结果如下：



GD测试集分类结果如下：



SGD训练集分类结果如下：



SGD测试集分类结果如下：

