## Lab2实验报告

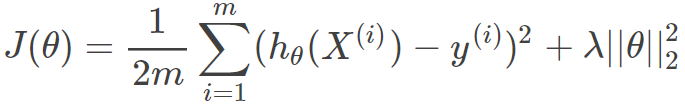
PB19151769 马宇骁

实验目标：

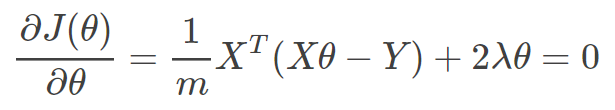
在conda的python3.6上实现线性分类算法，朴素⻉叶斯分类器，SVM算法等。

###### Part1机器学习

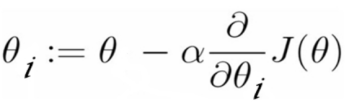
线性分类算法：



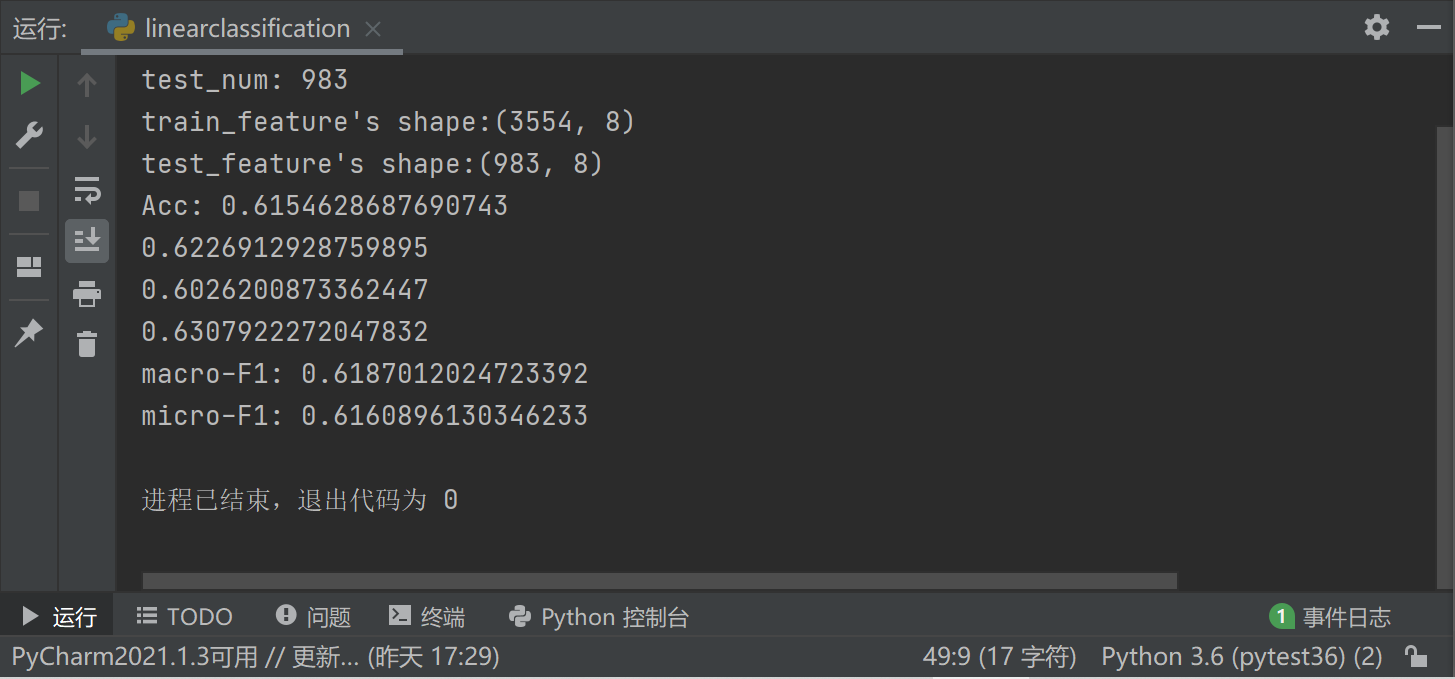
根据最小二乘法，求导：



故，



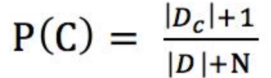
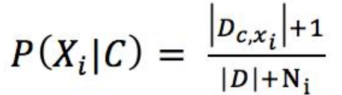
最终得到迭代矩阵，算得结果如下：



分析：

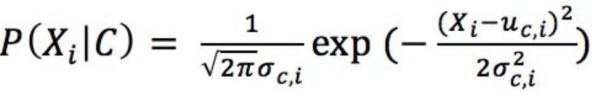
在处理中途不论怎么运行acc都为0，预测准确度也很低，经过print和思考，发现预测结果必须为整型而线性回归结果为浮点数，因此考虑y = np.round(y,0)处理为整数，再次运行无误。

朴素贝叶斯分类器：

根据合适的数据集即可得到贝叶斯分类器的分类结果，但是，若某个属性值在训练集中没有与某个类同时出现过，则直接基于条件概率与先验概率的计算公式进行计算，将出现不论其它属性如何，该类的判别概率都为0的问题。为了避免避免其它属性携带的信息被训练集中未出现的属性值归0，在估计概率值时要使用“平滑”的手段，通常使用的方法为拉普拉斯修正。现在用N表示训练集D中的类别数，Ni表示第i个属性的取值数。

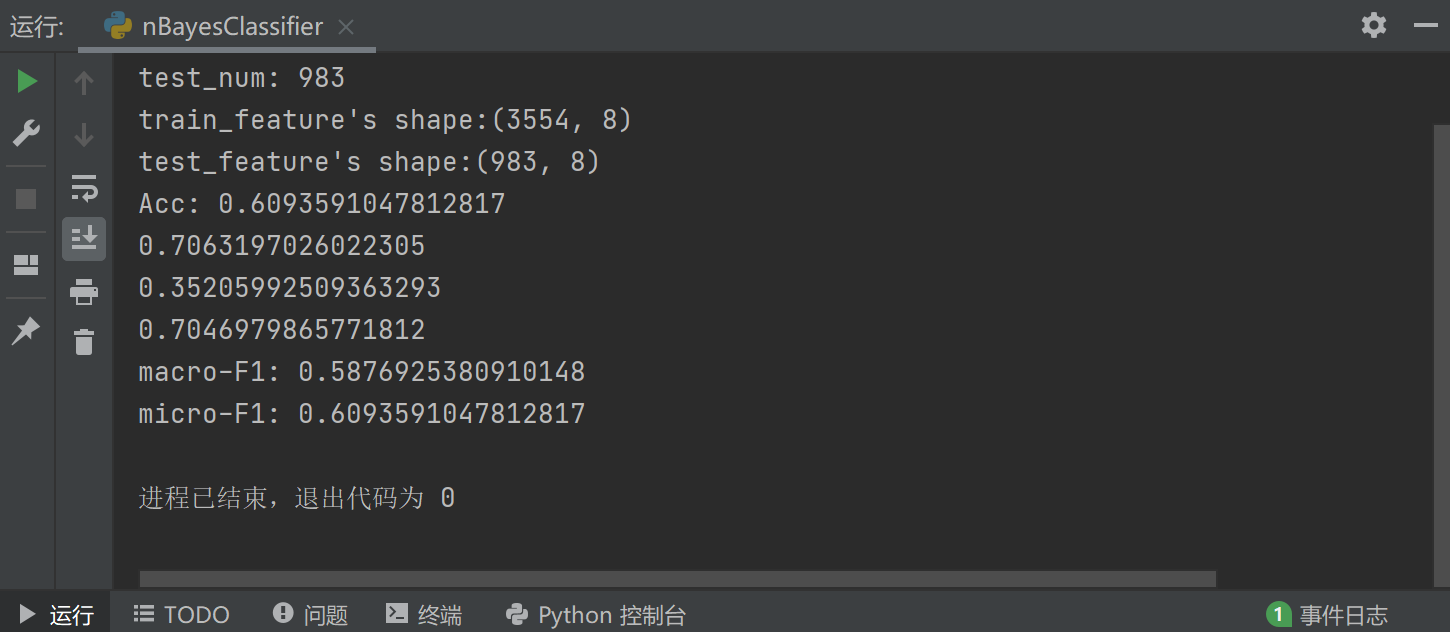
若数据连续，则考虑用高斯分布近似处理：



最终计算



取P(C)最大值的C记为预测值。结果如下：



SVM：

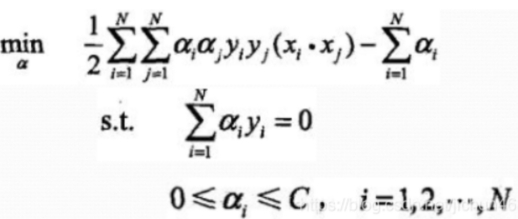
数据集：T={（x1,y1），（x2,y2），…，（xN,yN）}

(xi，yi)为一个训练样本，xi为一个实例，yi为该实例的分类。yi为1时为正例，yi为-1时为负例。

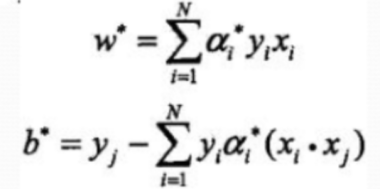
学习目标：在特征空间寻找一个分离超平面（将两类正确分类并且使得两类几何间隔最大）：wx+b=0，该超平面由法向量w和截距b决定。法向量指向的一侧为正类。

相应的分类决策函数：f(x)=sign(wx+b) 称为线性可分支持向量机

求解a：



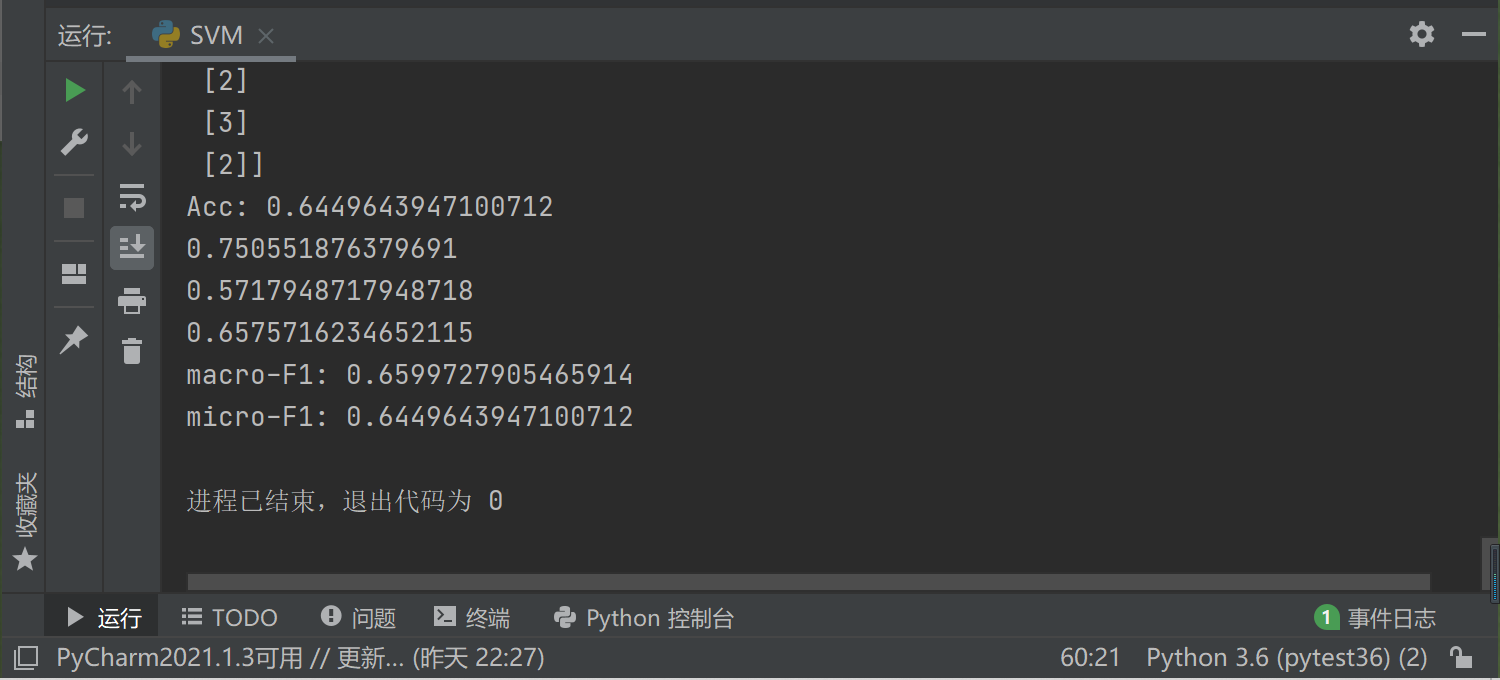
求解w，b：



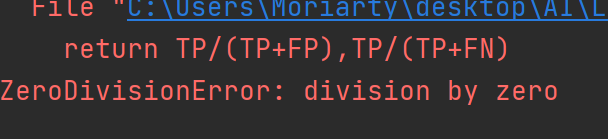


当数据中有一些奇异点，如下图中跑上来的两个红色的点。只是这极个别的点而导致的非完全线性可分。 此时，需要对每个样本点增加一个松弛变量。

结果如下：



问题分析：



实验过程中速度很慢据目测是由于调用核函数引起的。最初代码编译总是报A的矩阵大小不匹配，经过推导，将A最终确定为(np.ones(row) \* y, (1, row))的矩阵。实验过程中总是在进行一半的时候因为出现除以零而中断，此时的准确率只有0.23左右，经过改动参数，直接调用self中的数据最终实现完成svm且准确度正常。

###### Part2深度学习

手写感知机模型并进行反向传播

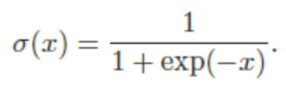
实验目的：

对矩阵链式求导的掌握

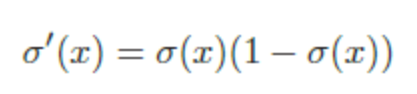
实验内容：

实现一个4层的感知机模型（隐层神经元设置为5，4，4，3，即输入的特征尾为5，输出的类别个数的3，激活函数设置为sigmoid）；实现BP算法；实现梯度下降算法。

sigmod 函数：



导数为：

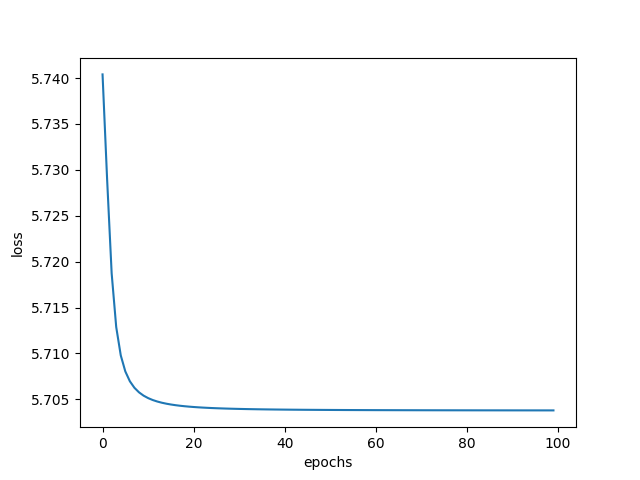


LogisticRegression逻辑回归（softmax回归）

我们要三层的MLP，则只需要HiddenLayer+LogisticRegression，

如果要四层的MLP，则为HiddenLayer+HiddenLayer+LogisticRegression........以此类推。

LOSS图像如下：



问题：

实验中遇到的问题主要集中在对模型的理解上以及最后的loss可视化中间，发现每次运行的时候loss的图像有变化，在少数情况下甚至出现曲线一直增长的情况，应该是由于生成随机数的不确定导致的。

复现MLP-Mixer

实验目的：

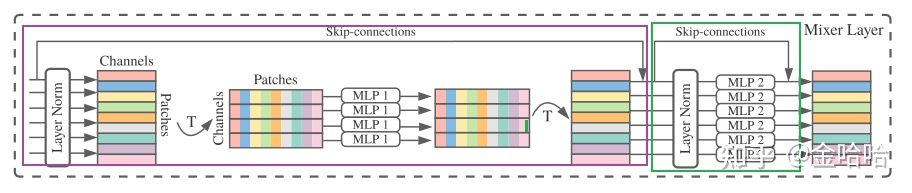
对深度学习的初步掌握，仅使用最基础的多层感知机。考察自行搜索相关资料学习的能力。

实验内容：

复现MLP-Mixer模型，并在MNIST数据集上进行测试。

MLP-Mixer中，通过Mixer Layer使用MLP先后对列、行进行映射，实现空间域和通道域的信息融合。与传统卷积不同的是，Mixer Layer将空间域和通道域分开操作，这种思想与Xception和MobileNet中的深

度可分离卷积相似。



该实验根据csdn的思路跌跌撞撞完成函数代码最终还是未能成功测试数据集，但求给点分吧orz。