

**《机器学习》课程实验报告**

**学 院 软件学院**

**专 业 软件工程**

**组 员**  占聪聪

**学 号 201530613573**

**邮 箱 [643033786@qq.com](mailto:643033786@qq.com)**

**指导教师**

**提交日期** **2017年 月 日**

## 1. 实验题目: 线性回归，线性分类与梯度下降

## 2. 实验时间：2017年 12 月 2 日

## 3. 报告人: 占聪聪

## 4. 实验目的:

## 1.进一步理解线性回归和梯度下降的原理。

## 2.在小规模数据集上实践。

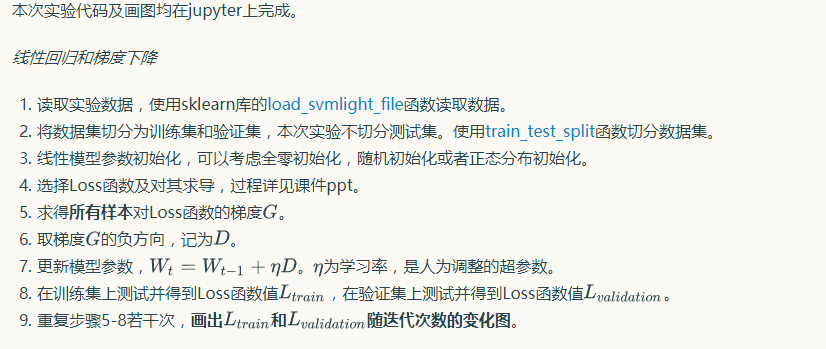
## 3.体会优化和调参的过程。

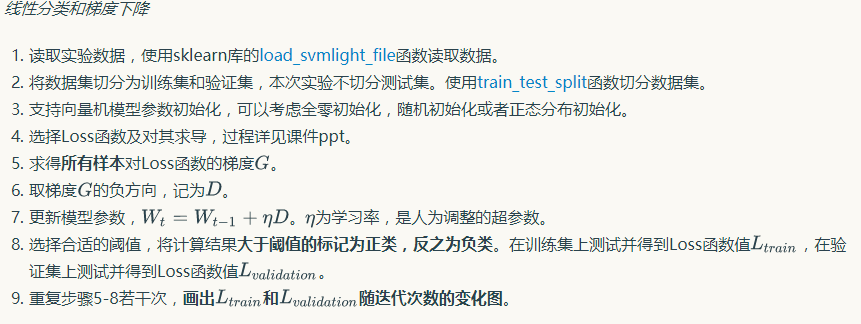
## 5. 数据集以及数据分析：

线性回归使用的是LIBSVM Data中的Housing数据，包含506个样本，每个样本有13个属性。并将其切分为训练集，验证集。

线性分类使用的是LIBSVM Data中的australian数据，包含690个样本，每个样本有14 个属性。并将其切分为训练集，验证集。

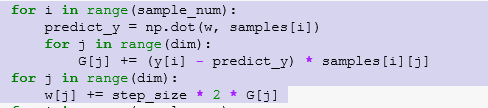
## 6. 实验步骤:



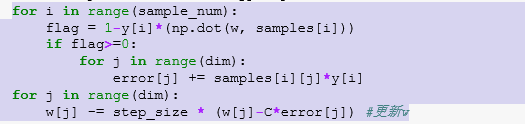


## 7. 代码内容:

线性回归更新w代码：



线性分类更新w代码：



## 线性回归：

## 8. 选择的评估方法（留出法，交叉验证，k折交叉验证等）:

交叉验证

## 9. 模型参数的初始化方法:

全零初始化

## 10.选择的loss函数及其导数:





## 11.实验结果和曲线图:

## 超参数选择（η,epoch等）：

η=0.01 epoch=500(保持不变）

## 评估结果（根据选择的评估方法）：

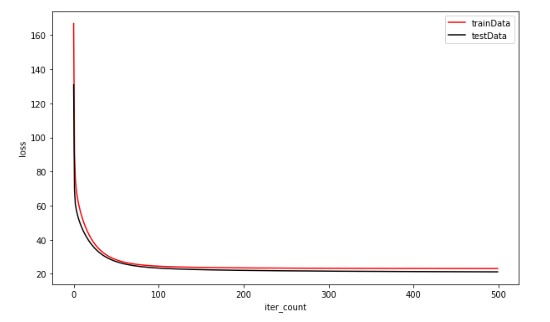
据loss曲线图下降情况情况来选择参数

η=0.001 η=0.0001 η=0.0002 η=0.0003

## 预测结果（最佳结果）：

η=0.0002

## loss曲线图：



## 12.实验结果分析:

训练集的loss和验证集的loss随迭代次数减小至最佳状态趋于平滑，loss值减到很小趋于不变。

## **线性分类：**

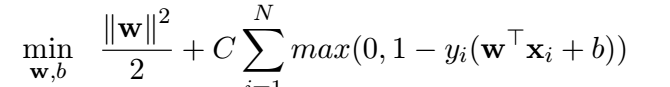
## 8. 选择的评估方法（留出法，交叉验证，k折交叉验证等）:

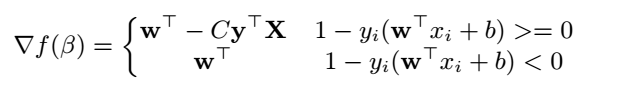
交叉验证

## 9. 模型参数的初始化方法:

全零初始化

## 10.选择的loss函数及其导数：





## 11.实验结果和曲线图:

## 超参数选择（η,epoch等）：

η=0.01 epoch=100(保持不变），C=1.0

## 评估结果（根据选择的评估方法）：

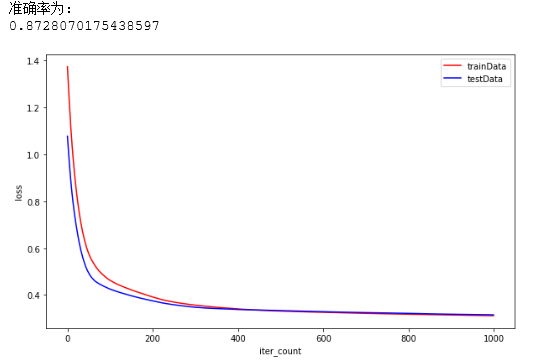
根据loss曲线图下降情况情况来选择参数

η=0.001,C=1.0 η=0.001,C=0.9 η=0.0001,C=0.9 η=0.0002,C=0.9

预测结果（最佳结果）：

η=0.0001,C=0.9

## loss曲线图：



## 12.实验结果分析:

训练集的loss和验证集的loss随迭代次数减小至最佳状态趋于平滑，loss减到很小，趋于不变。

## 13.对比线性回归和线性分类的异同点：

相同之处：都是利用梯度下降方法求得最佳w的近似值。

不同处：线性回归主要运用于处理连续数据问题，预测连续数据的函数值；而线性分类问题则主要用来处理离散数值问题，用于分离离散点，即分类问题。

## 14.实验总结：

通过本次的实验更加深入了解了线性回归以及线性分类问题（支持向量机）的基本原理思想，以及如何确定损失函数和损失函数的求导，到最后的编程实现。