

**《机器学习》课程实验报告**

**学 院 软件学院**

**专 业 软件工程**

**组 员**  **麦栋铖**

**学 号 201536612525**

**邮 箱** [379857334@qq.com](mailto:379857334@qq.com)

**指导教师**  **谭明奎**

**提交日期** **2017年12月 8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8日

## 3. 报告人:麦栋铖

## 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。

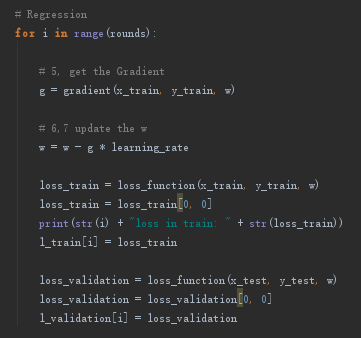
## 数据集以及数据分析：

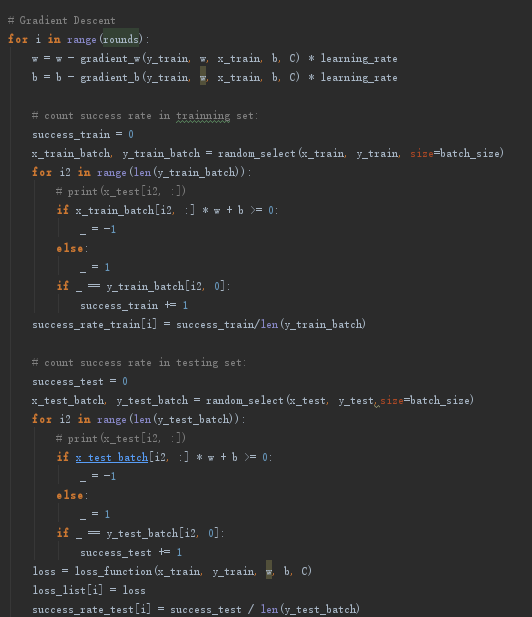
线性回归使用的是[LIBSVM Data](https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/)中的[Housing](https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/regression.html#housing)数据，包含506个样本，每个样本有13个属性。请自行下载scaled版本，并将其切分为训练集，验证集。   
线性分类使用的是[LIBSVM Data](https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/)中的[australian](https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/binary.html#australian)数据，包含690个样本，每个样本有14 个属性。请自行下载scaled版本，并将其切分为训练集，验证集。

## 实验步骤:



## 代码内容:





## 8. 选择的评估方法（留出法，交叉验证，k折交叉验证等）:

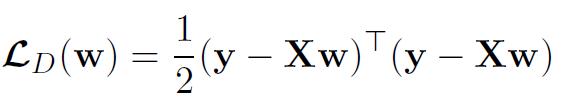
留出法

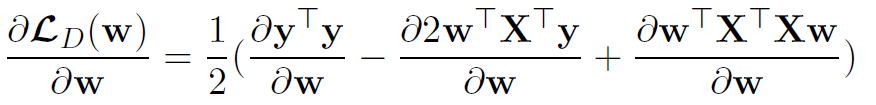
## 9. 模型参数的初始化方法:

使用全零初始化

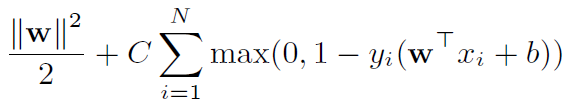
## 10.选择的loss函数及其导数:

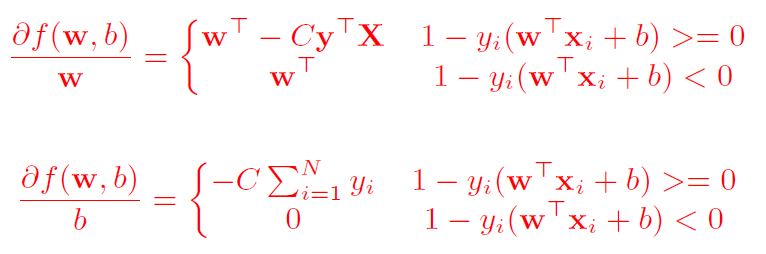
回归：





分类：





## 11.实验结果和曲线图:

## 超参数选择（η,epoch等）：

## 评估结果（根据选择的评估方法）：

## 预测结果（最佳结果）：

## loss曲线图：

## 12.实验结果分析:

梯度下降会随着迭代次数增加而收敛到某个值或者震荡。

调参非常的重要

## 13.对比线性回归和线性分类的异同点：

相同点：目的都是拟合一个超平面

不同点：回归是使点尽量靠近超平面，而分类是使点尽量远离超平面并划分为两类，在超平面的不同位置。

## 14.实验总结：

机器学习是一种很有效的拟合数据的方法

调参和优化需要非常有耐心的尝试。