

**《机器学习》课程实验报告**

**学 院 软件学院**

**专 业 软件工程**

**组 员**  **陈世鸿**

**学 号 201530611234**

**邮 箱 yingzi827165351@vip.qq.com**

**指导教师**  **吴庆耀**

**提交日期** **2017年12月7日**

## 1. 实验题目: 逻辑回归、线性分类与随机梯度下降

## 2. 实验时间：2017年12月2日

## 3. 报告人:陈世鸿

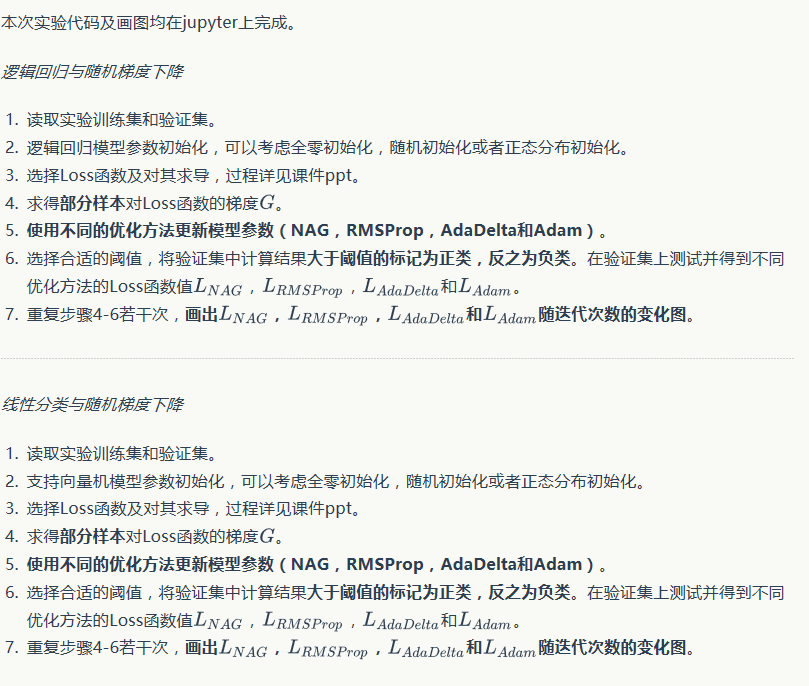
## 4. 实验目的:

1. 对比理解梯度下降和随机梯度下降的区别与联系。
2. 对比理解逻辑回归和线性分类的区别与联系。
3. 进一步理解SVM的原理并在较大数据上实践。

## 5. 数据集以及数据分析：

实验使用的是LIBSVM Data的中的a9a数据，包含32561 / 16281(testing)个样本，每个样本有123/123 (testing)个属性。请自行下载训练集和验证集。

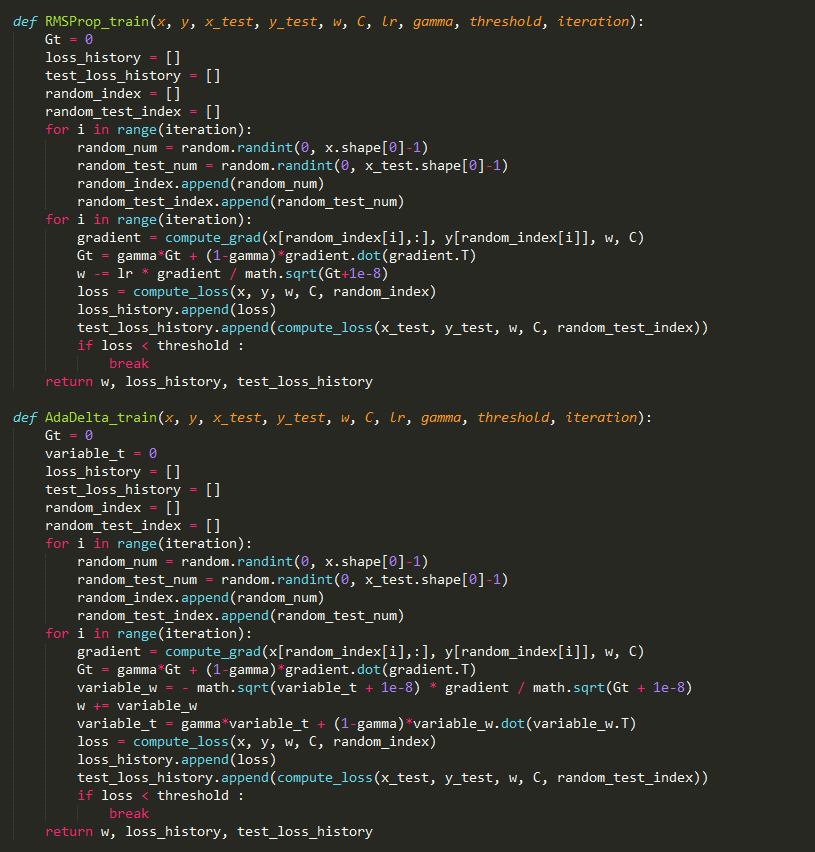
## 6. 实验步骤:

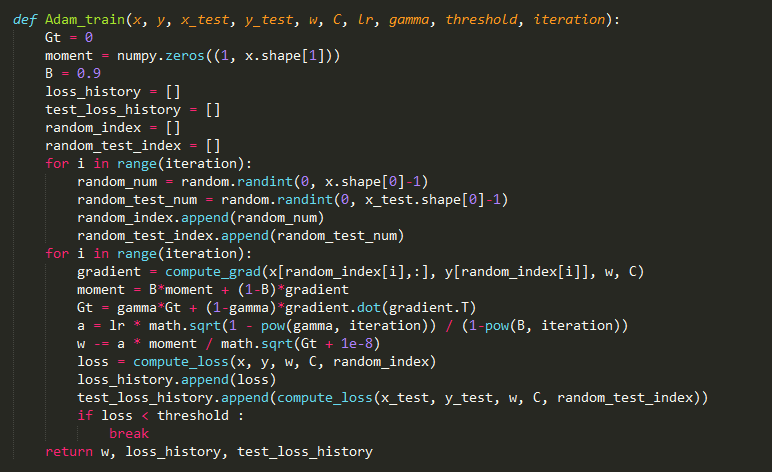


## 7. 代码内容:

RegressionExperiment:

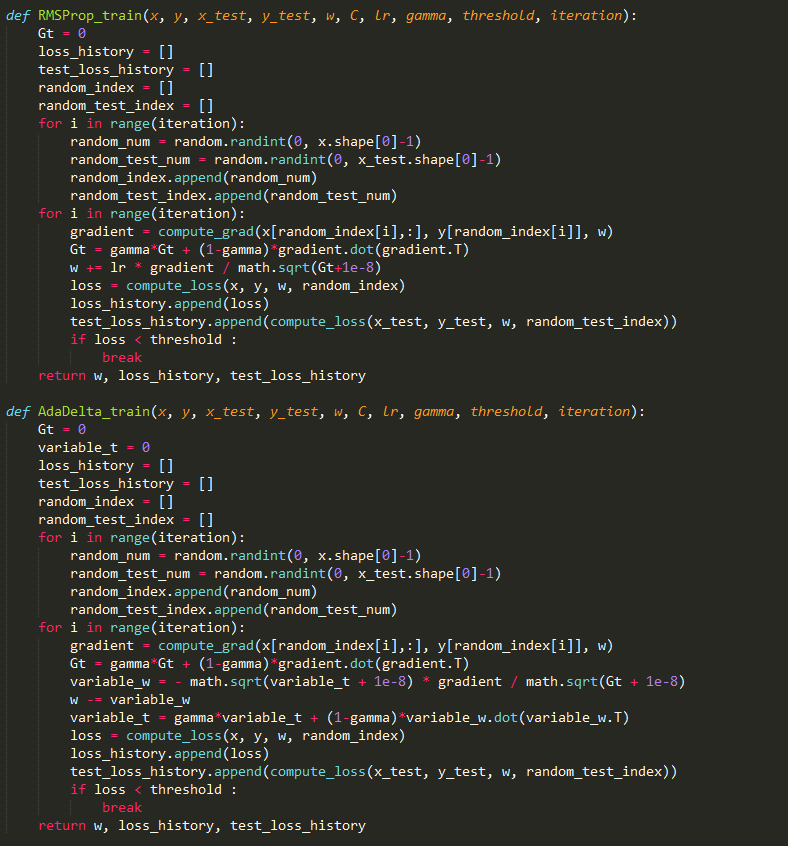


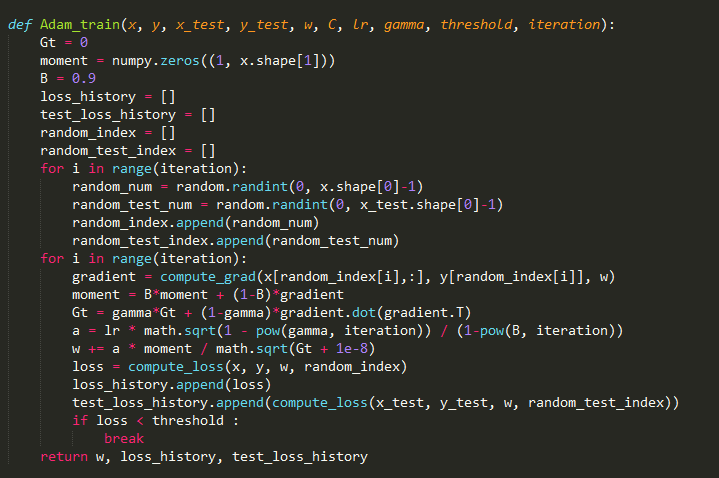




ClassificationExperiment:









## 8. 模型参数的初始化方法:

**Regression Experiment：**全零初始化。

**Classification Experiment：**全零初始化。

## 9.选择的loss函数及其导数:

**Regression Experiment：**

Loss函数：

导数：+ λw

**Classification Experiment：**

Loss函数：

其中，

导数：

## 10.实验结果和曲线图:

## 超参数选择：

**Regression Experiment：**

NAG：λ=0.3 learning rate=0.001

gamma = 0.9 threshold=0.001 iteration=3000

RMSProp：λ=0.3 learning rate=0.001

gamma = 0.9 threshold=0.001 iteration=3000

AdaDelta：λ=0.3 learning rate=0.001

gamma = 0.9 threshold=0.001 iteration=3000

Adam：λ=0.3 learning rate=0.001

gamma = 0.9 threshold=0.001 iteration=3000

**Classification Experiment:**

NAG：λ=0.3 learning rate=0.001

gamma = 0.9 threshold=0.001 iteration=3000

RMSProp：λ=0.3 learning rate=0.001

gamma = 0.9 threshold=0.001 iteration=3000

AdaDelta：λ=0.3 learning rate=0.001

gamma = 0.9 threshold=0.001 iteration=3000

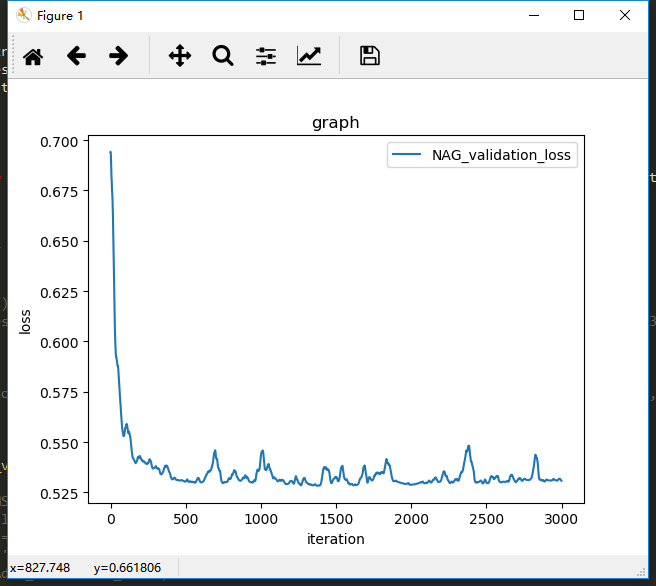
Adam：λ=0.3 learning rate=0.001

gamma = 0.9 threshold=0.001 iteration=3000

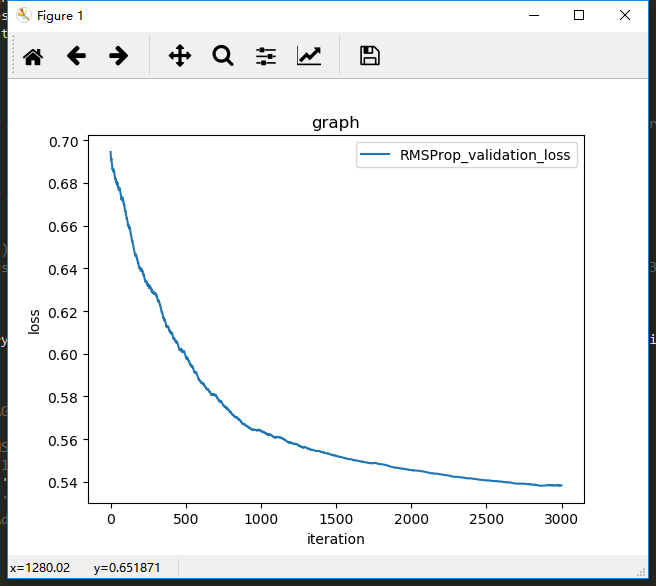
## 预测结果（最佳结果）：

**Regression Experiment：**

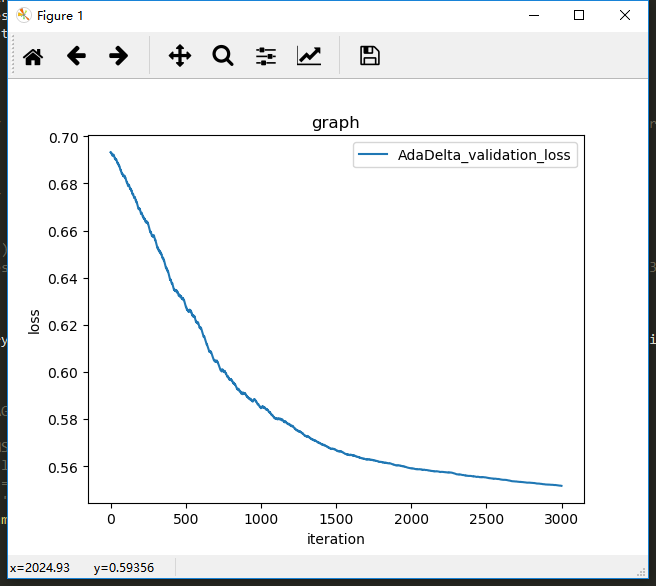
NAG：



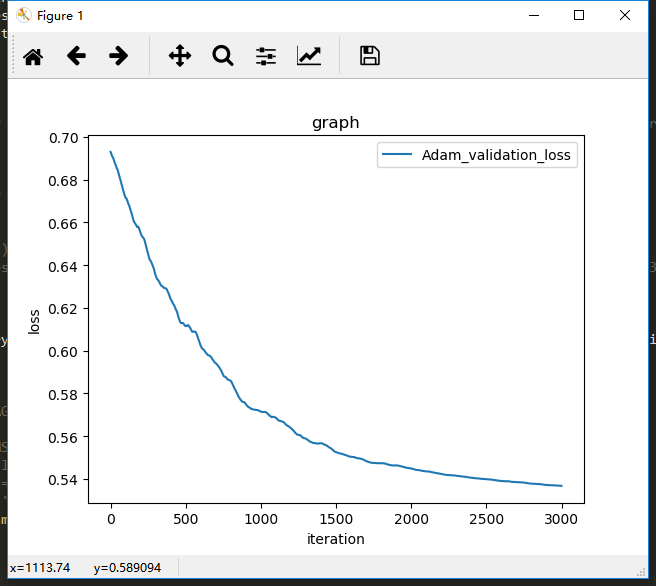
RMSProp：



AdaDelta：

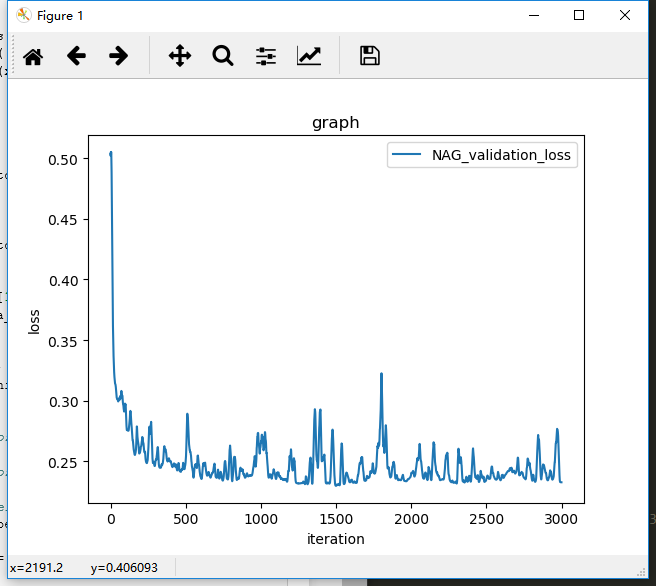


Adam：

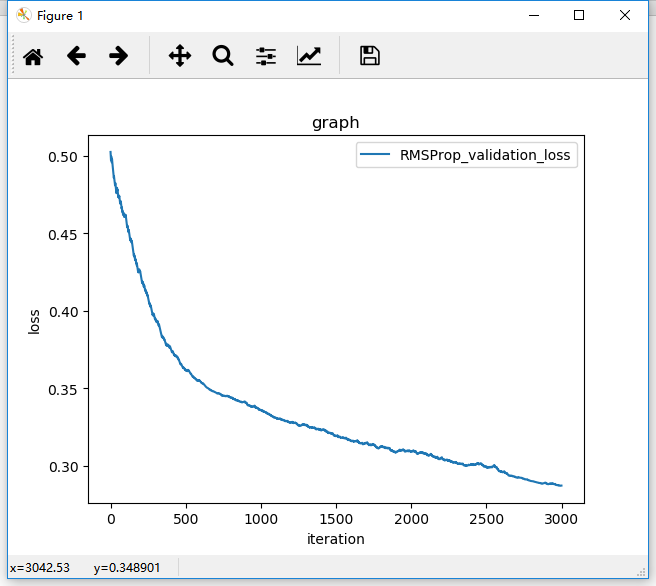


**Classification Experiment：**

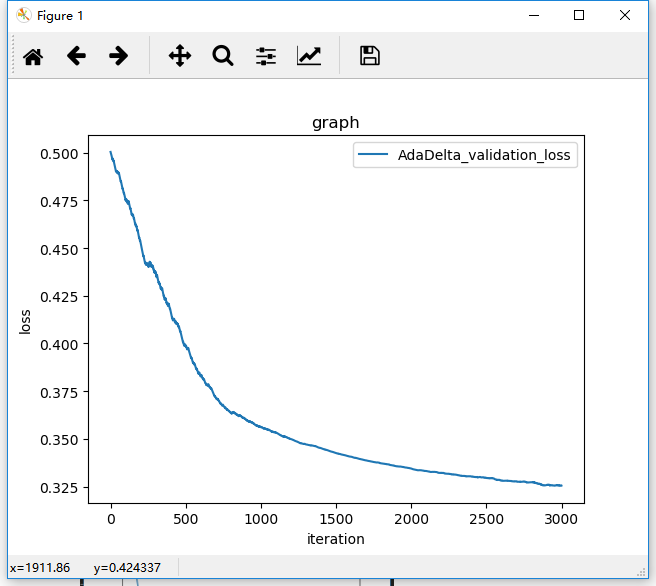
NAG:



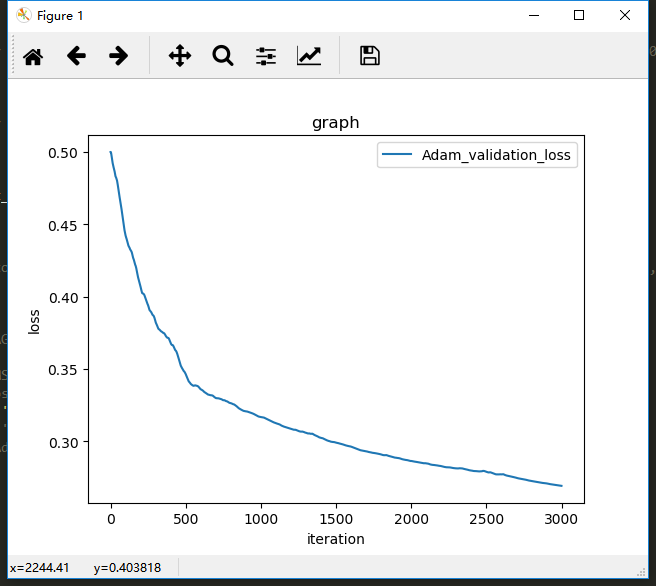
RMSProp:



AdaDelta:

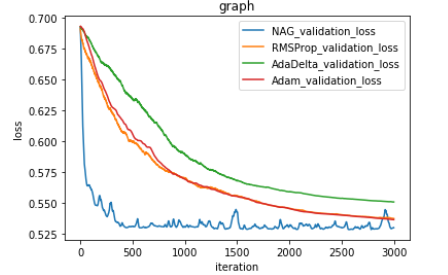


Adam:

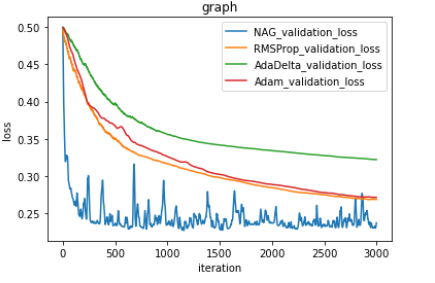


## loss曲线图：

**Regression Experiment：**



**Classification Experiment:**



## 11.实验结果分析:

NAG收敛速度最快，但是震荡明显；；RMSProp 与 Adam的收敛速度几乎持平，震荡不明显；AdaDelta收敛较慢但是比较稳定。

## 12.对比逻辑回归和线性分类的异同点：

**同：**

都属于分类问题，都用于预测。

**异：**

找最优超平面的方法不同，，形象点说，logistic模型找的那个超平面，是尽量让所有点都远离它，而SVM线性分类寻找的那个超平面，是只让最靠近中间分割线的那些点尽量远离，即只用到那些“支持向量”的样本。

逻辑回归只可以处理线性可分情况；SVM则二者皆可。

## 13.实验总结：

在本次实验当中，我学习到了很多SGD在实践运用的经验，将课程中学习到的知识运用在实际问题上。但是由于自身对于知识把握懂得程度不高，在实现的过程中遇到了诸如无法正确实现SGD优化算法，调参不够灵活的问题，在总结反思之后解决了问题并顺利完成了实验。、

在对模型的训练过程中，我体会到了灵活调参的重要性。在一开始，因为超参数C, learning rate等设置得不合理，导致loss图像与预期相差甚远，模型参数无法收敛或者收敛过慢，跑数据集时间过长等问题，导致无法拟合数据；或者是因为迭代次数过少，参数还未收敛便停止了训练。在进行数次的不同的调参后，模型往预期方向改变，我也从中学得一些经验。