

**《机器学习》课程实验报告**

**学 院 软件学院**

**专 业 软件工程**

**组 员**  **李嘉灏**

**学 号 20153061951**

**邮 箱 786699266@qq.com**

**指导教师**  **吴庆耀**

**提交日期** **2017年12月13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13日

## 3. 报告人:李嘉灏

## 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](https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/)的中的[a9a](https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/binary.html#a9a)数据，包含32561 / 16281(testing)个样本，每个样本有123/123 (testing)个属性。

其中label表示该记录的类别，+1为正类，-1为负类。其中正类有11687个样本，负类有37155个样本。

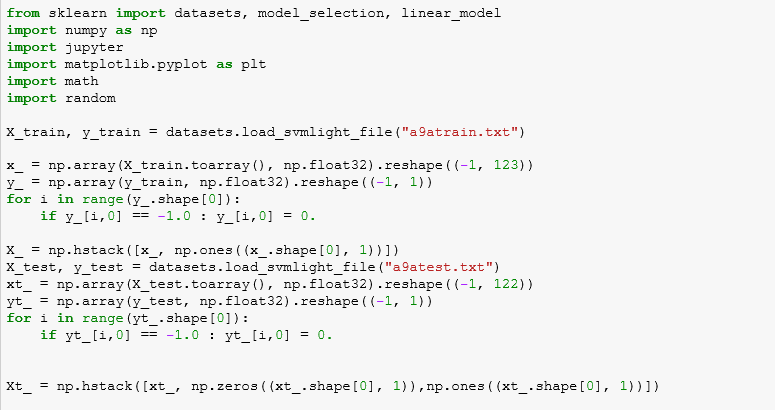
## 6. 实验步骤:

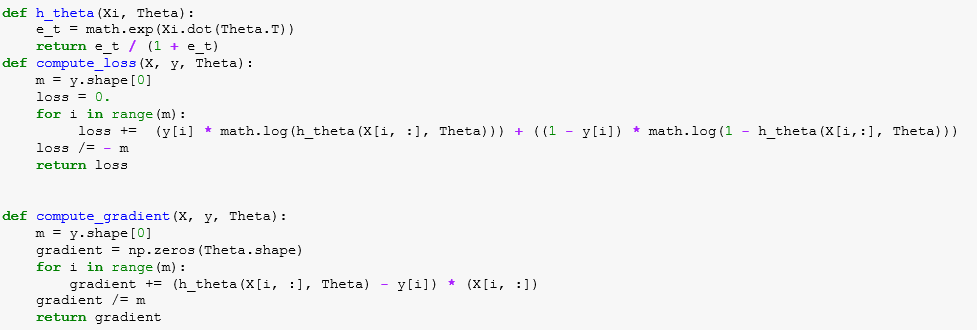


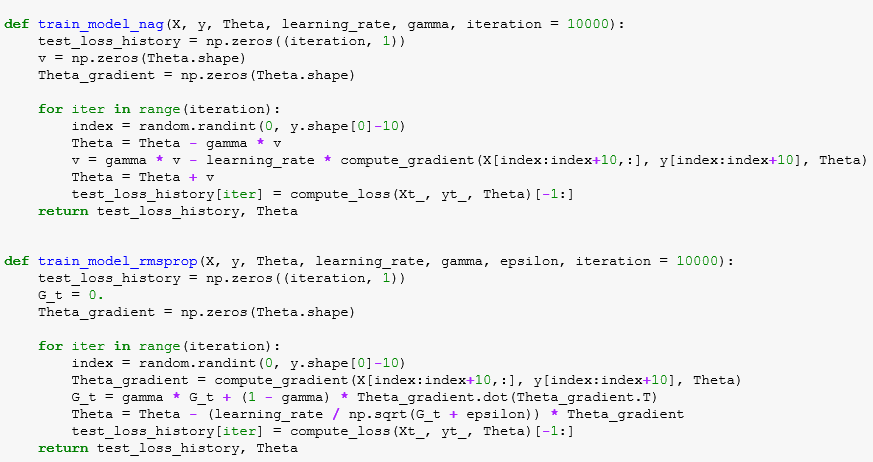


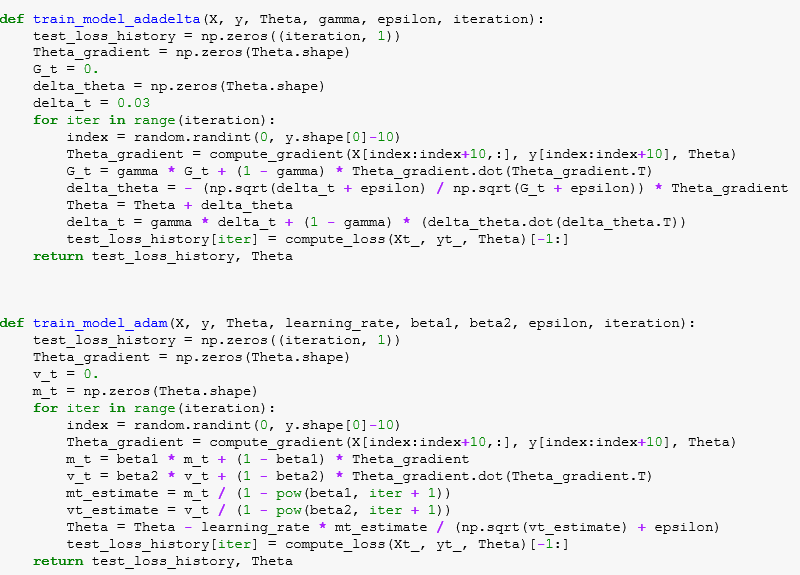
## 7. 代码:

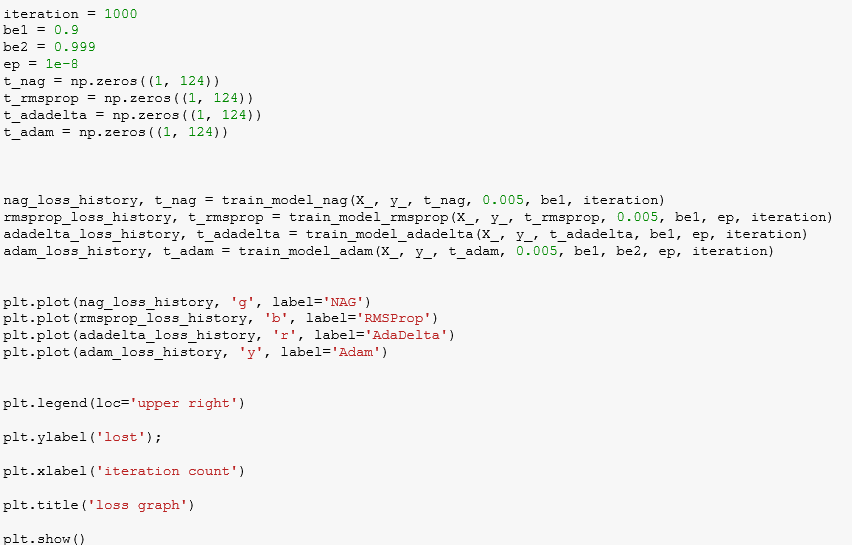
**逻辑回归**：



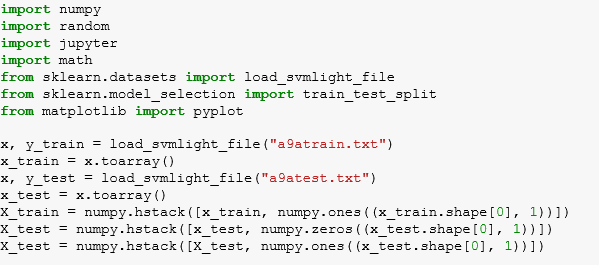


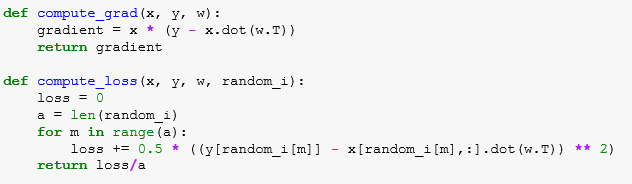


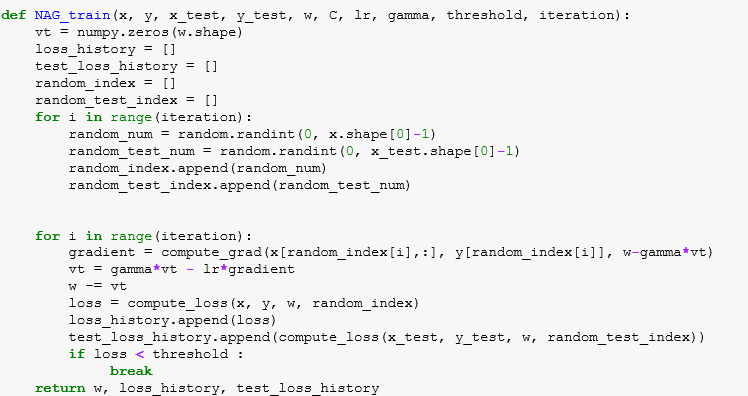


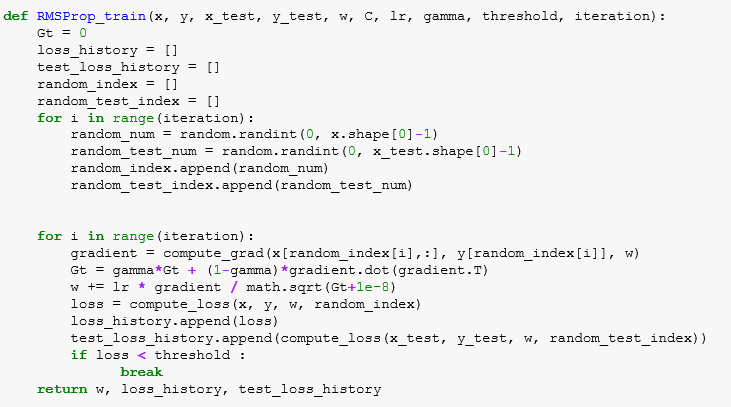


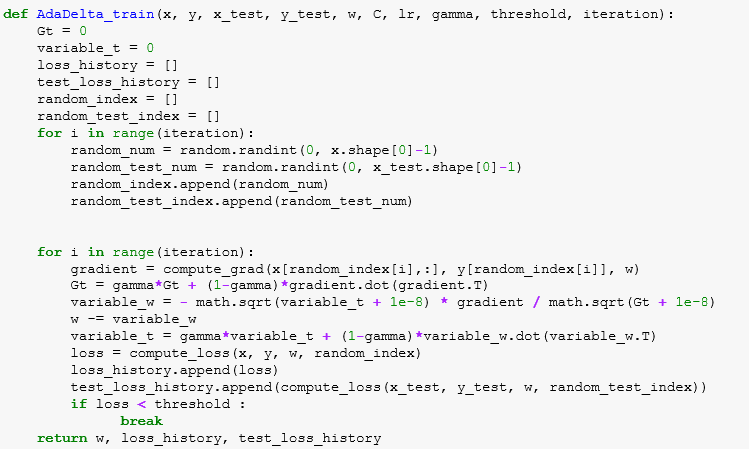
**线性分类：**

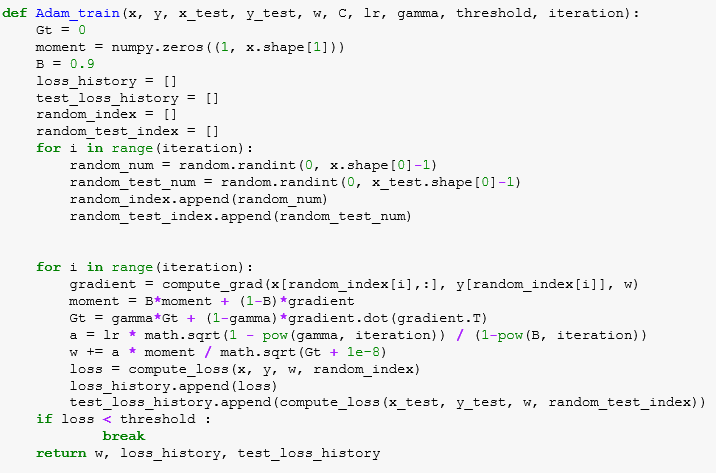


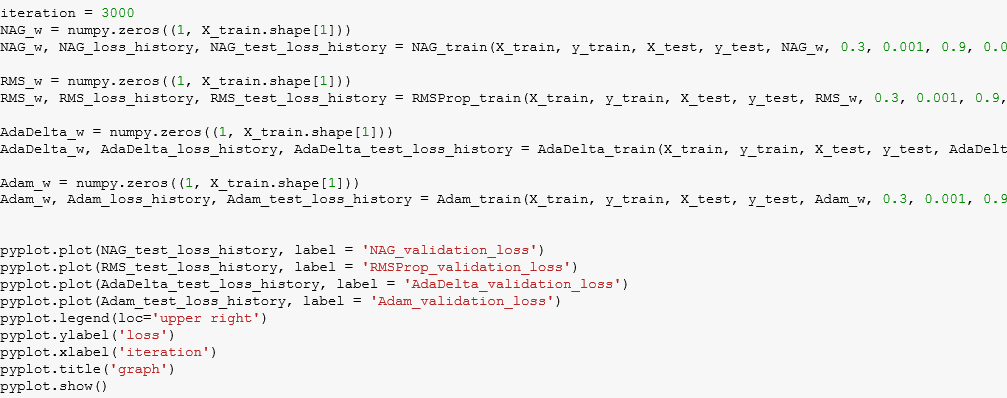












**逻辑回归：**

## 8. 模型参数的初始化方法:

全零初始化。

## 9.选择的loss函数及其导数:

loss function:

PS：

Gradient:

## 10.实验结果和曲线图:

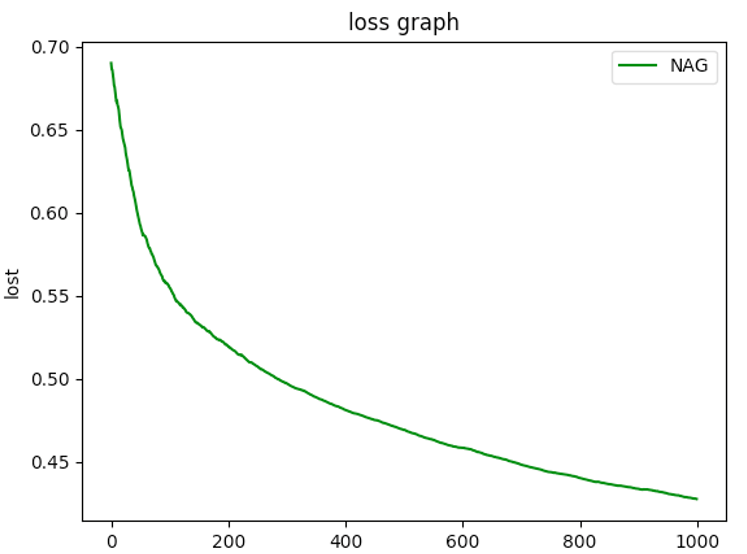
## 1.NAG:

## 超参数选择：η=0.005

## =0.9

## epoch = 1000

## 预测结果（最佳结果）：



**2.RMSProp:**

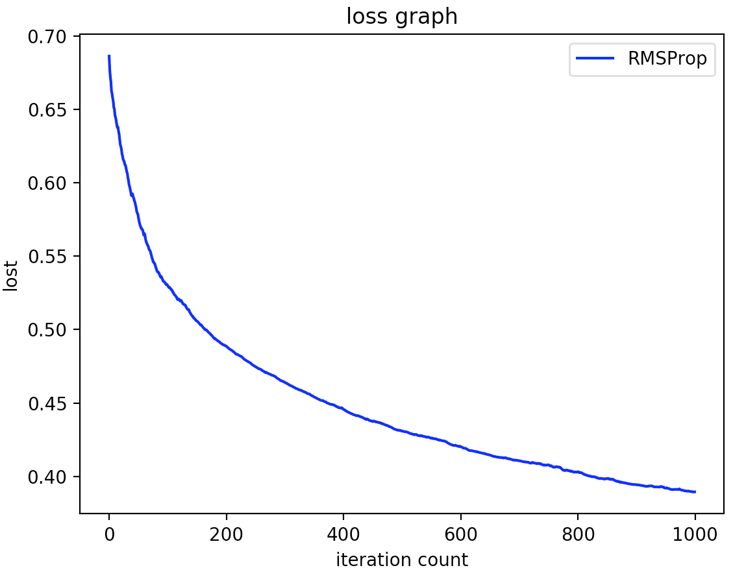
## 超参数选择：η=0.005

## = 0.9

## = 1e-8

## epoch = 1000

## 预测结果（最佳结果）：



**3.Adam:**

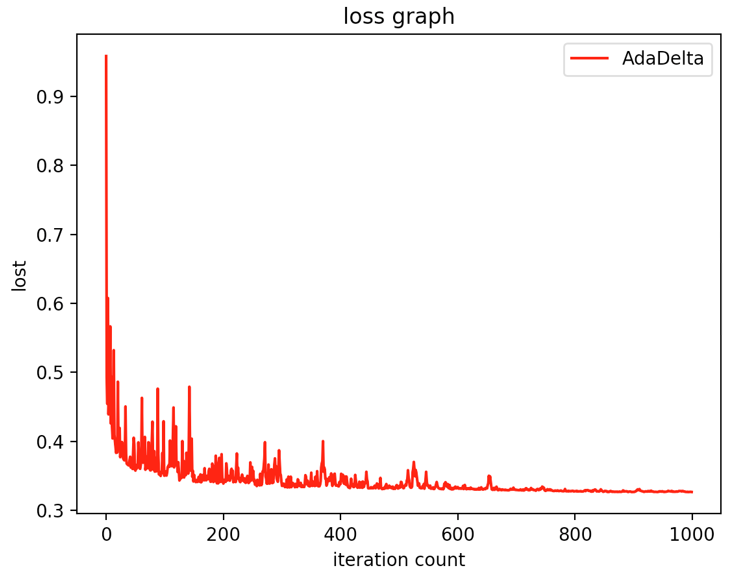
## 超参数选择：η=0.005

## = 0.9

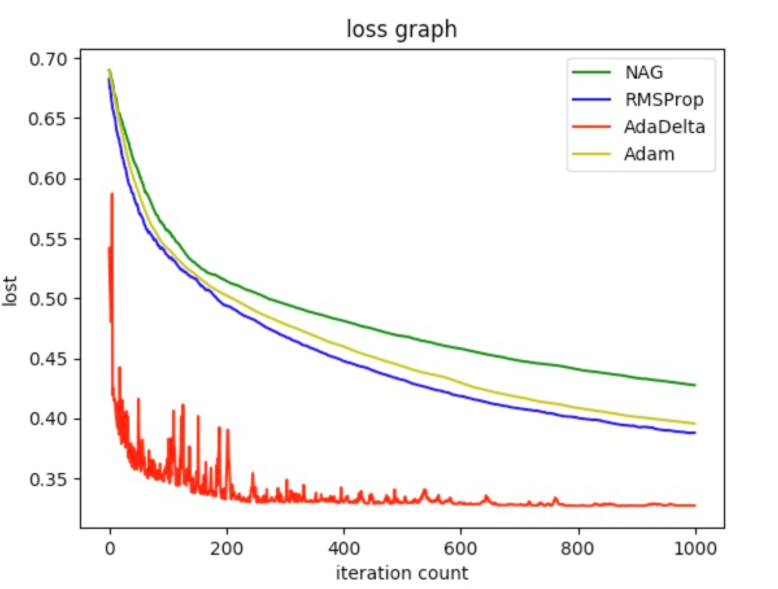
## = 0.999

## = 1e-8 epoch=1000

## 预测结果（最佳结果）：



## loss曲线图：



## 11.实验结果分析:

1.AdaDelta优点是较快的达到了局部最优解，缺点是震荡明显；

2.NAG缺点是收敛速度和收敛幅度上都较小，优点则是曲线较为平滑；

3.RMSProp和Adam的曲线基本是一致的，收敛速度比NAG快，但是比AdaDelta慢。

结论：四种方法都可以较为快速收敛到局部最优解，都比普通梯度下降法更好

**线性分类：**

## 8. 模型参数的初始化方法:

全零初始化

## 9.选择的loss函数及其导数:

loss function:

PS:

gradient:

## 10.实验结果和曲线图:

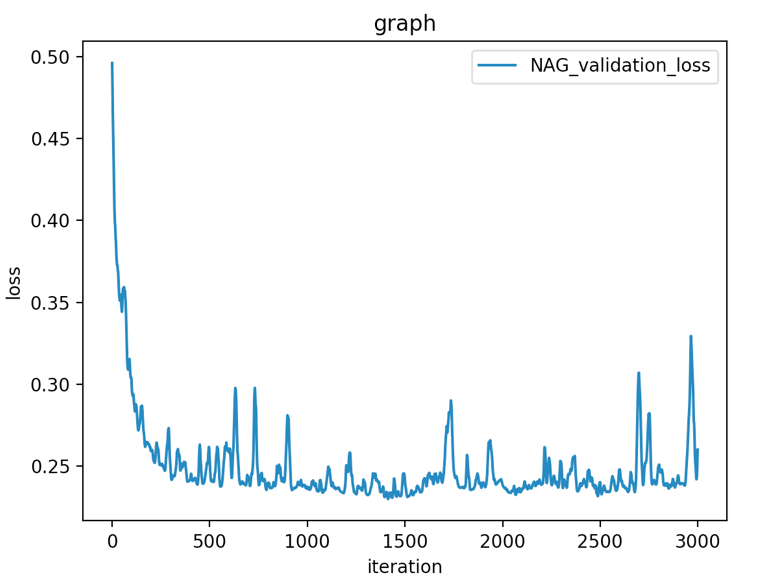
## 1.NAG:

## 超参数选择：η=0.001

## =0.9

## epoch = 3000

## 预测结果（最佳结果）：



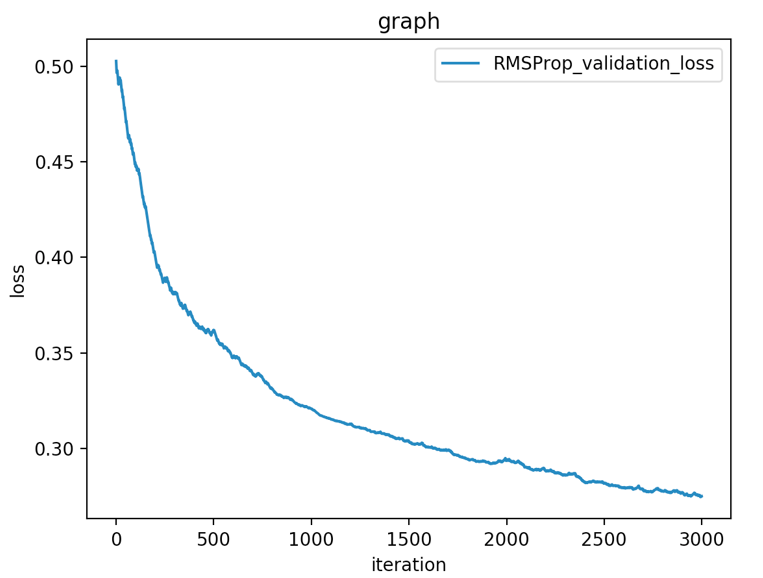
**2.RMSProp:**

## 超参数选择：η=0.001

## = 0.9

## = 1e-8 epoch = 3000

## 预测结果（最佳结果）：



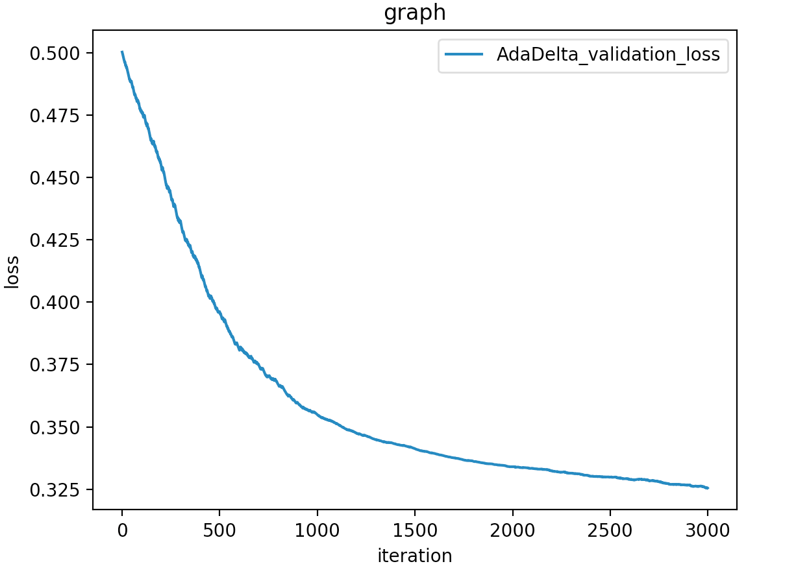
**3.AdaDelta:**

## 超参数选择： = 0.9

## = 1e-8

## epoch=3000

## 预测结果（最佳结果）：



**4.Adam:**

## 超参数选择：η=0.005

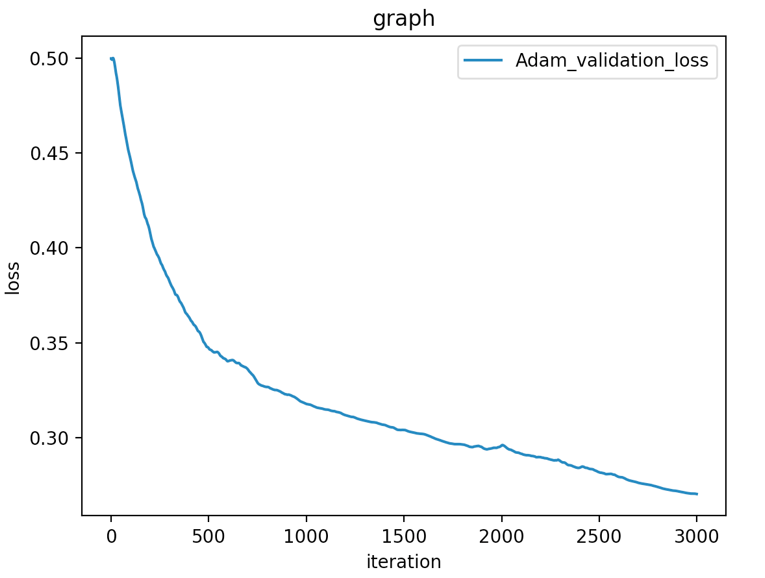
## = 0.9

## = 0.999

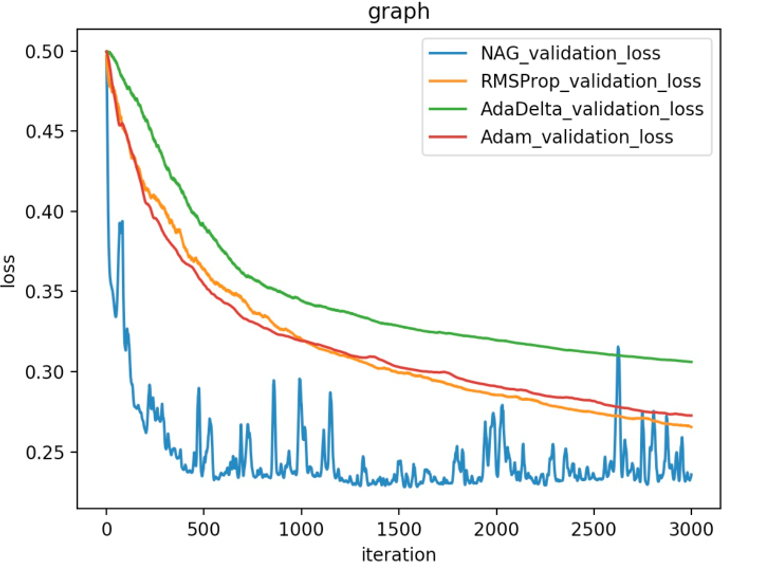
## = 1e-8

## epoch=3000

## 预测结果（最佳结果）：



**loss曲线图：**



## 11.实验结果分析:

1.NAG优点是收敛速度最快，缺点是震荡明显；

2.RMSProp 与 Adam的收敛速度几乎是一样的，优点是震荡不明显；

3.AdaDelta缺点是收敛较慢，优点是是曲线比较平滑。

结论：相比普通梯度下降算法，这四种算法在不同程度上做了改进，更加优秀。

## 12.对比逻辑回归和线性分类的异同点：

相同点都属于分类问题并且都用于预测。

不同点为逻辑回归只可以处理线性的情况，但是SVM都可以。

## 13.实验总结：

本次实验中，我学习到了很多回归、线性和随机梯度等在实践运用的经验，将课程中学习到的知识运用在实际问题上。同时在实验中通过漫长的调试最终还是幸运的完成了实验。

在对模型的训练过程中，我体会到了调整参数的重要性。最初由于超参数学习率等参数设置得不合理，导致loss图像与最佳的预期相差很多。好在我在进行数次的不同的调参后，模型往预计方向改变，我也从中学到了很多。