

**课 程 实 验 报 告**

**课程名称： 数字图像处理课程设计**

**专业班级： 种子1901**

**组 员： 李瑞堃、李 可**

**胡玉洁、董浣羽**

**学 号： U201913780、U201914150、**

**U201915157、U201815671**

**指导教师： 王兴刚**

**报告日期： 2022年3月21日**

**电子信息与通信学院**

**目 录**

[实验1 线性模型 1](#_Toc98884919)

[1.1实验内容 1](#_Toc98884920)

[1.2实验原理 1](#_Toc98884921)

[1.2.1线性回归方程 1](#_Toc98884922)

[1.2.2 损失函数设计 1](#_Toc98884923)

[1.2.3 梯度下降优化 1](#_Toc98884924)

[1.3实验步骤 2](#_Toc98884925)

[1.4训练&参数讨论 3](#_Toc98884926)

[1.5 实验结果 3](#_Toc98884927)

[1.6 组员介绍&分工 4](#_Toc98884928)

[1.7 仓库链接 5](#_Toc98884929)

# 实验1 线性模型

## 1.1实验内容

设计并训练线性感知机，对两簇正态分布样本点进行二分类，样本点数量为500×2=1000个，分布如下：

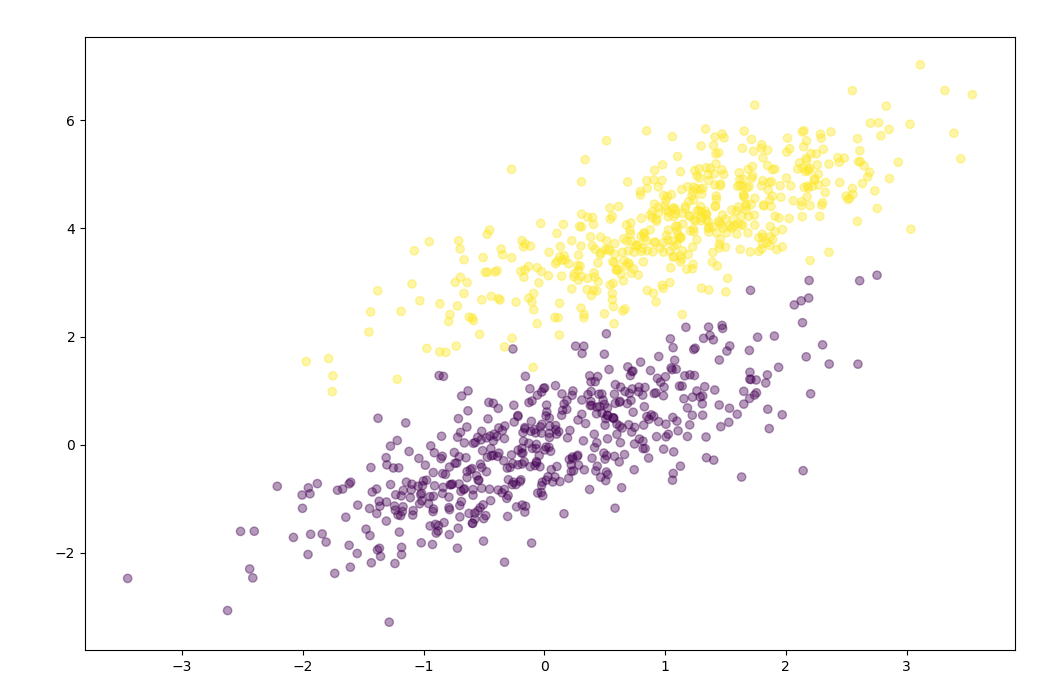
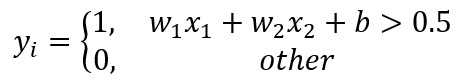


图1.1 样本分布

## 1.2实验原理

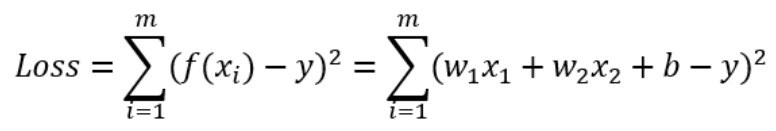
### 1.2.1线性回归方程

数据集 D={(x1,y1),(x2,y2),……,(xm,ym)}，标签 y∈[0, 1]，根据线性回归方程结果预测y：



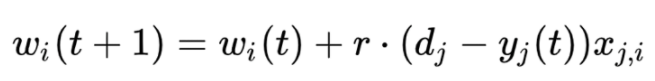
### 1.2.2 损失函数设计

使用回归问题常用的均方损失函数：



### 1.2.3 梯度下降优化

使用最小二乘法思想，使损失值最小。对每个样本或每批量样本进行预测，然后使损失函数Loss依次对参量 w1、w2求导，并依据梯度反向更新参量以减小后续误差：



循环此过程至最大轮数或者至损失值符合要求。

## 1.3实验步骤

**（1）数据预处理**

对1000样本的数据集进行乱序处理，考虑到样本属性值已满足0~1间的正态分布，故不做其他处理，按照9:1比例划分训练集和测试集。

**（2）批量训练**

使用逐样本更新的方式，每对一个样本完成预测后计算损失梯度并依据学习率进行参数更新，然后对下一个样本重复此过程，以此为一轮，共进行 MAX\_STEP=100轮。

**（3）测试集验证**

用训练后的最终权重对测试集进行预测，统计测准率P、召回率R、正确率acc。

**（4）鲁棒性测试**

1. lr 对 acc、P、R的影响

改变lr取值为：1e-5~1e-3，取相同数据集，相同步数观察模型性能变化

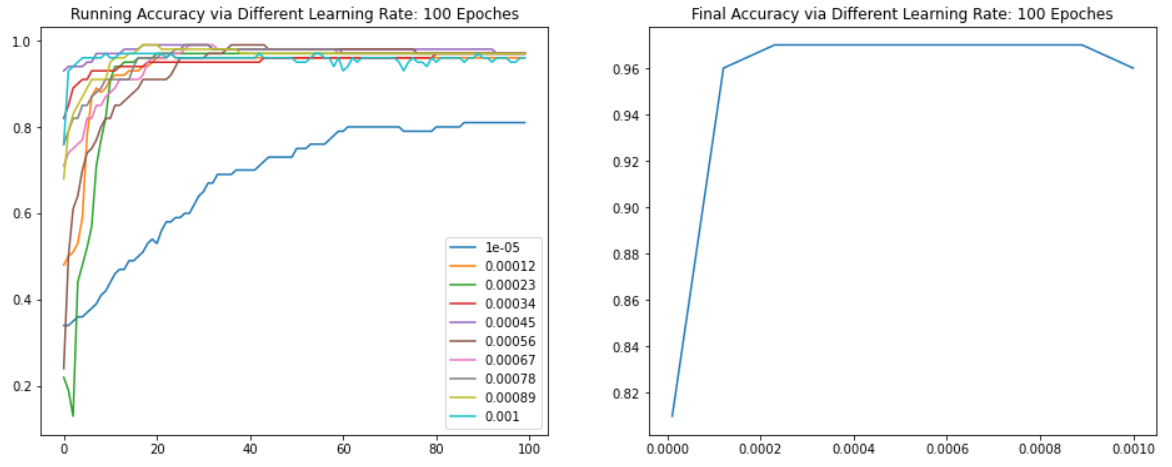


图 1.2 不同学习率下的准确率

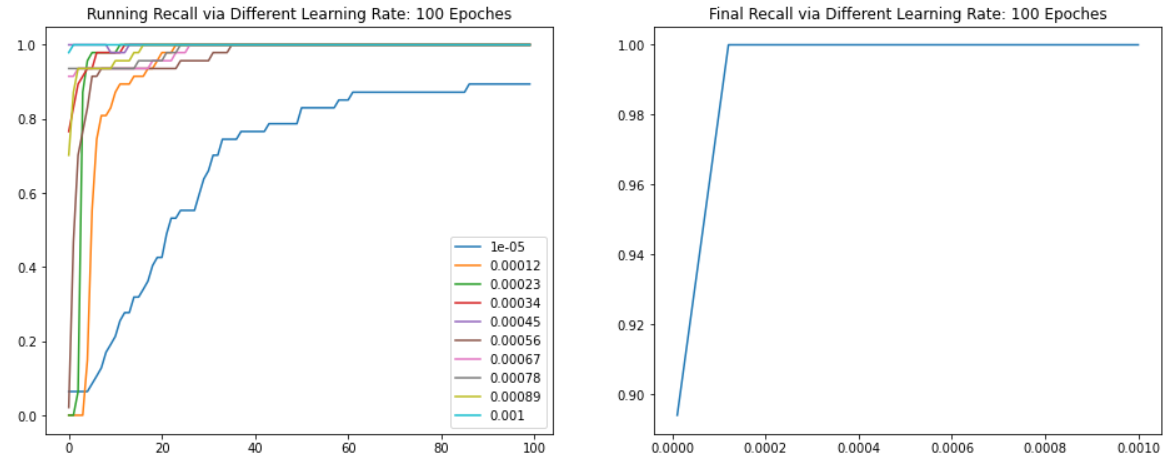


图 1.3 不同学习率下的召回率

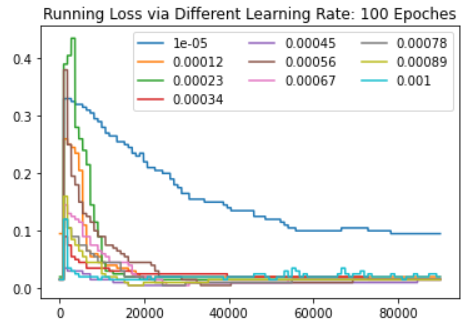


图 1.4 不同学习率下的损失函数

## 1.4训练&参数讨论

**batch size 与 learning rate**

逐样本更新时，batch size=1，此时如果学习率太小会导致收敛过慢，需要增加训练轮数

全样本更新时，如果loss值取累加值而不是平均，则需要调低学习率，否则单次更新步长太大，难以收敛到最优解

**loss function 与 threshold**

使用均方误差作为损失函数，本质上是使用了欧式距离的度量思想，应当保证分类阈值到各个样本点的距离相同。

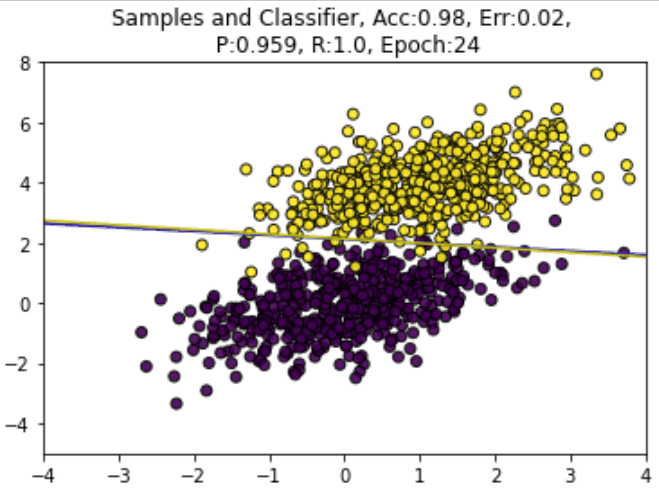
问题建模中，我们用0、1来表示两种样本标签，虽然类别之间无大小和次序之分，但是在计算时不得不考虑到数值0、1的大小之差，因此采用0.5而不是0作为预测分类的阈值

**参数明细**

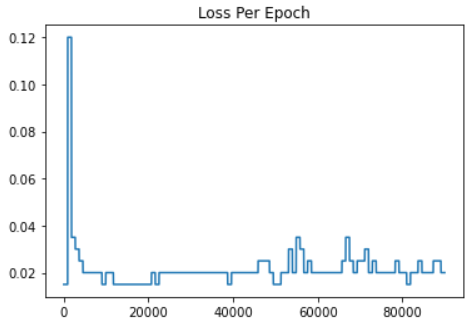
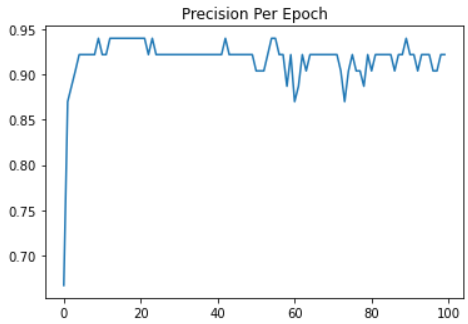
|  |  |
| --- | --- |
| 随机种子 | 666 |
| 训练集、测试集样本数 | 900、100 |
| 学习率 | 1e-3 |
| 训练轮数 | <100，或者模型收敛 |

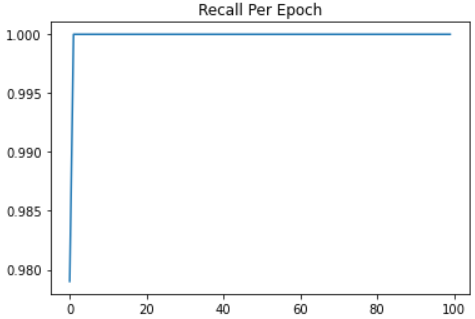
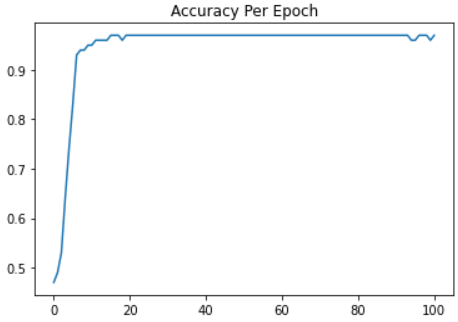
## 1.5 实验结果

训练过程可视化：



性能指标变化图：

## 1.6 组员介绍&分工

**组员项目组背景**

董浣羽：Intel组BM方向，DRL求解波束赋形优化问题，擅长语言：Python

李瑞堃：Intel组NS方向，Multi-Agent DRL求解频谱资源调度问题，擅长语言：Python

李可：Intel组RAN方向， DU与CU间F1AP通信，擅长语言：C

胡玉洁：洛凯组客户端方向，安卓，客户端开发，擅长语言：Java

**分工情况：**

|  |  |
| --- | --- |
| **姓名** | **工作** |
| 李瑞堃 # | 实现代码、流程设计 |
| 董浣羽 # | 实现代码、学习率实验 |
| 胡玉洁 ## | 数据集划分、总结分析、语言学习 |
| 李可 ## | 评价指标、总结分析、语言学习 |

## 1.7 仓库链接

<https://gitee.com/huanyuuu/dip_hw/>

本实验代码在lrk和dhy分支各有一个版本，分别是按照不同风格写的。