

《人工智能及应用》

实验报告

姓 名： 王浩

学 号： 2020204246

专业班级： 计算机科学技术学院图灵班

指导教师： 杨环

提交日期： 2022. 10. 26

实验题目：Fisher 线性分类器的设计与实现

1.实验目的及要求

- 1) 掌握 Fisher 线性判别的基本原理
- 2) 利用 Fisher 线性判别解决基本的两类线性分类问题

2.实验内容

实验预处理

实验一：

只需要把数据进行转置，就可以进行 fisher 线性判别

实验二, 三：

因为每个文件夹中是一个人的数据，每个人有十张照片，所以我们采用前 8 个照片作为训练集，后 2 张照片作为预测集，然后将图片数据导入，进行转置，标准化，处理。然后采用 PCA 进行降维，将 10000 多维的数据转化为 2 维处理。

预处理之后的结果如下图所示：

```
In [8]: new_train_data

Out[8]: array([[ 39.5392097, -48.23393704],
 [ 253.41650477, 335.02507101],
 [ 489.35489913, 552.50156784],
 [ 100.18039223, 48.99575378],
 [ 711.94301814, -442.45815878],
 [ 596.4836973, 341.77506308],
 [ 527.40509308, -461.26078368],
 [ 327.49970493, -138.96335549],
 [-348.39481593, -132.09781853],
 [-331.08573095, -72.78278395],
 [-428.58042125, 44.70650403],
 [-408.2191111, -54.48853215],
 [-438.30268631, 56.14758017],
 [-303.63545787, -59.88205287],
 [-455.53256683, 120.0678003 ],
 [-332.07172904, -89.05191771]])
```

```
In [9]: new_test_data

Out[9]: array([[ 427.71529601, 327.79844478],
 [ 692.25358885, -285.11725081],
 [-456.86803202, 139.97333227],
 [-663.10085283, -182.65452624]])
```

数据预处理结果

实验过程

实验 1

参数 A	参数 B	是否合格
2.95	6.63	合格
2.53	7.79	合格
3.57	5.65	合格
3.16	5.47	合格
2.58	4.46	不合格
2.16	6.22	不合格
3.27	3.52	不合格

实验 1 数据图

分析：

这是利用 LDA 进行一个分类的问题：

假设一个产品有两个参数柔软性 A 和钢性 B 来衡量它是否合格。

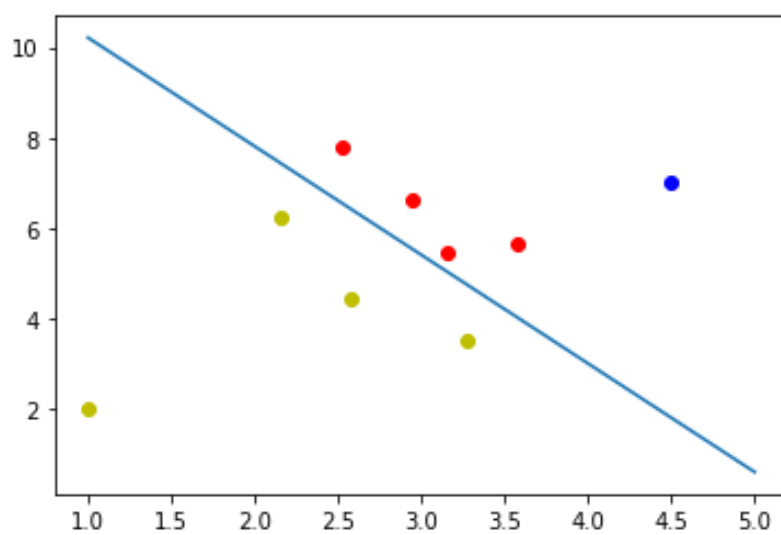
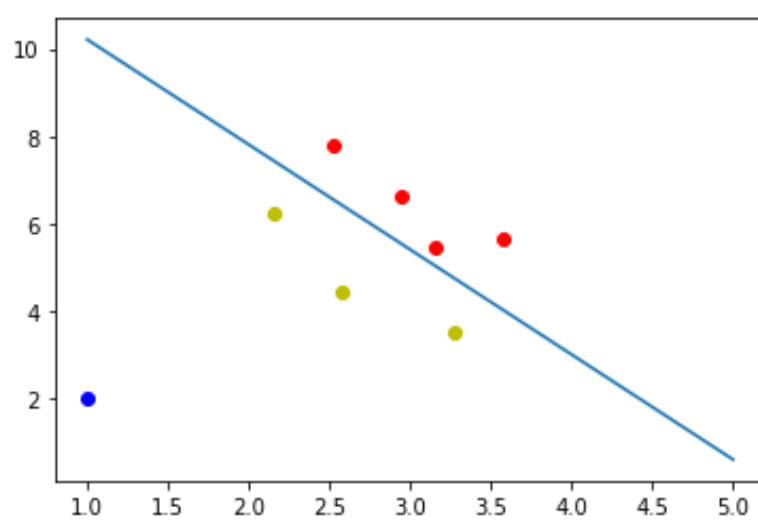
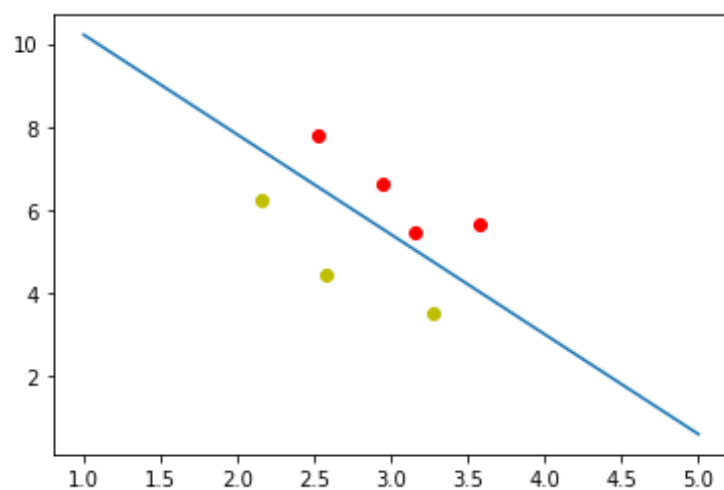
根据上图中的数据，我们可以把样本分为两类，一类是合格的产品，一类是不合格的产品。

首先通过 LDA 算法对训练样本的投影获得判别函数，

然后判断测试样本的类别，即输入一个样本的两个参数，

判断该产品是否合格。

最后实验结果可视化如下图所示：

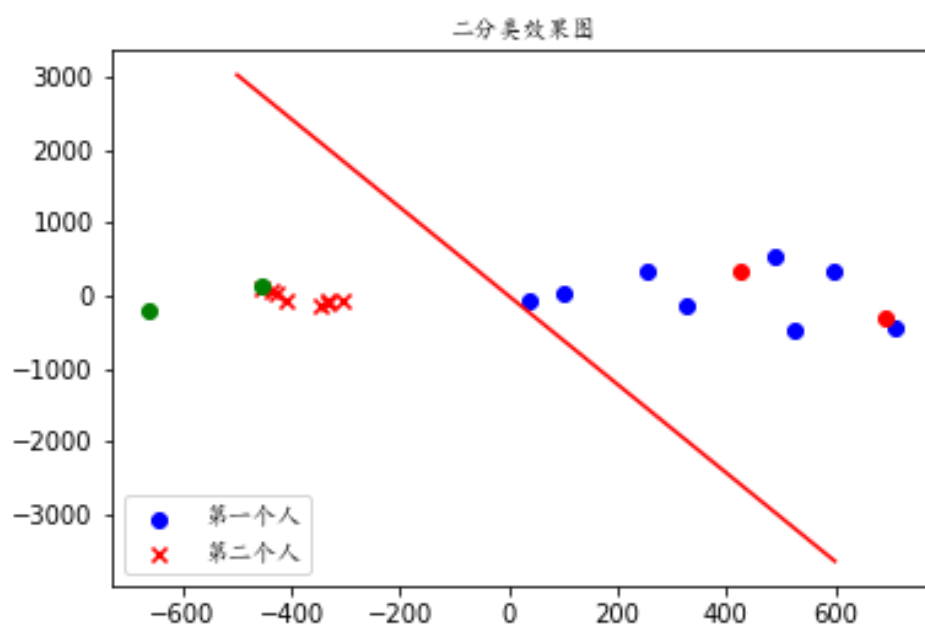


实验一演示图

实验 2

基于 ORL 人脸库，实验样本主要来自于两个人，每人 45 张图片，共有 90 个样本，其中的 80 个样本作为训练样本，10 个作为测试样本。通过 LDA 实现两类问题的线性判别。

初始化之后进行，通过 LDA 线性判别，获得线性判别函数的权向量，最后进行可视化如下图所示



实验二演示图

实验 3

(选做) 基于 ORL 人脸库，基于 Fisher 线性分类器实现多类人脸的识别问题。

实验原理

通过将多分类问题转化为多个 2 分类问题

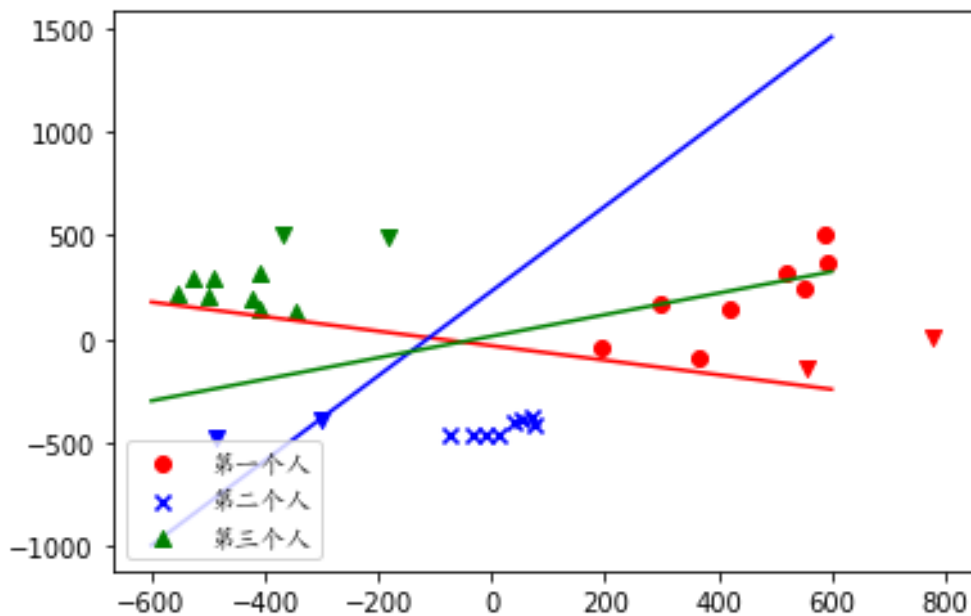
每类转化为 $c-1$ 个两类问题 (pairwise):

$$\omega_i \text{ 与 } \omega_j, \quad j \neq i$$

C 类则 $c(c-1)$ 个两类问题, 但其中半数相同, 故

C 类转化为 $\frac{c(c-1)}{2}$ 个两类问题

$$g_{ij}(x) \begin{cases} > 0 \rightarrow \text{当 } x \in \omega_i \\ < 0 \rightarrow \text{当 } x \in \omega_j \end{cases} \quad i \neq j$$



实验三演示图

实验结果及分析

通过 fisher 线性判别, 经过求数学公式的推导, 最后求出权向量的值, 求出判别函数后, 使其等于 0, 画图, 直观观察分类效果, 从结果上看, 分类效果比较好。

实验中存在的问题及解决方法

一开始没有采用 PCA 降维, 直接用 10000 多维的数据进行运算。结果, 代码运行时间

长，权向量的值非常不合理，预测结果也不合理，经过查阅相关资料显示，出现这种情况的原因是数据噪声太严重了，模型根本没有办法对训练数据集进行拟合。因此采用归一化，缩小维数，提高可操作性。