《模式识别与深度学习》

实验2——PCA降维SVM分类算法实验指导书

实验时间：2022年9月22日-2022年10月07日

1. 实验目的
2. 熟练掌握PCA降维算法原理及代码实现，能够对已实现算法进行灵活调用获取降维重建结果。
3. 熟练掌握PCA降维以及线性和非线性SVM算法，并且根据实验数据的特性调节PCA降维数以及SVM线性函数或核函数。
4. 实验环境
5. Python3
6. Numpy （科学计算，数据处理）
7. Scikit-learn（开源机器学习库）
8. Pillow （图像存取库）
9. 实验要求

本次实验包含两部分：

1. 在理解并掌握PCA人脸降维重建算法原理的基础上，更改给出的重建代码，实现对本班级采集人脸数据进行PCA降维重建实验，调节参数n\_components的不同取值，汇报不同取值对实验结果的影响（不同取值的重建结果截图展示），并尝试对结果进行分析，哪些取值对实验结果影响比较大，原因是什么。
2. 补全已给出的代码（pca\_svm.py），使用本班级人脸采集数据进行PCA降维SVM分类实验，汇报不同参数（PCA降维参数n\_components以及SVM分类参数）下训练集以及测试集的准确率。调节参数，汇报不同的参数对实验结果的影响，并尝试对结果进行分析，哪些参数对实验结果影响比较大，原因是什么。
3. 实验内容
4. 准备工作

安装好对应的实验环境，简单了解Numpy、Scikit-learn，Pillow等函数库的使用。

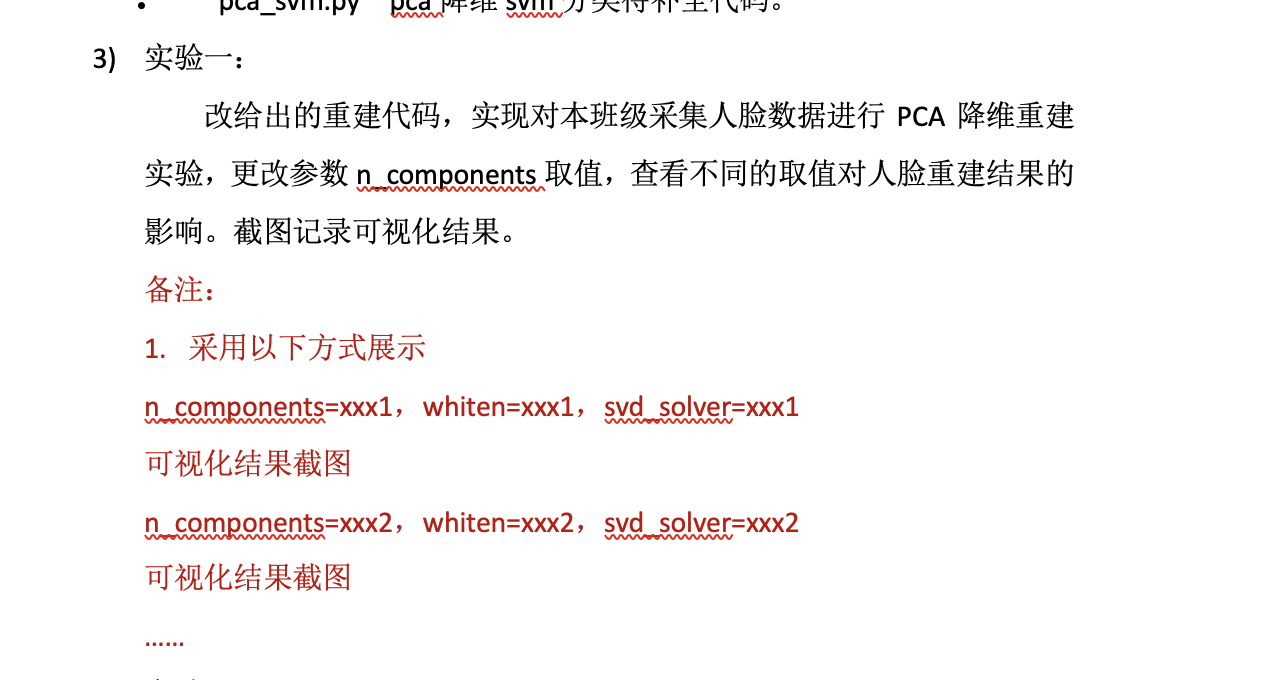
1. 实验材料（已上传到智慧树作业考试实验2-PCA）

* “pca\_recon/lfw-100\_ori”lfw人脸数据。
* “pca\_recon/lfw-100\_recon”演示pca降维重建实验结果。
* “pca\_svm/class4-facedata-rgb”本班人脸rgb数据。
* “pca\_recon/pca-reconstraction.py”pca降维重建实验完整代码。
* “pca\_svm/class4-facedata”本班人脸灰度数据。
* “pca\_recon/lfw-100\_recon”演示pca降维重建实验结果。
* “pca\_svm.py”pca降维svm分类待补全代码。

1. 实验一：

改给出的重建代码，实现对本班级采集人脸数据进行PCA降维重建实验，更改参数n\_components取值，查看不同的取值对人脸重建结果的影响。截图记录可视化结果。

备注：

1. 采用pca实验ppt中p12页的形式挑选一些样本展示即可
2. ****
3. 实验二：

补全pca\_svm.py代码，使用本班级人脸采集数据进行PCA降维，然后使用降维后的数据进行SVM分类实验。汇报不同参数（PCA降维参数以及SVM分类参数）下训练集以及测试集的准确率。调节参数，汇报不同的参数对实验结果的影响，并尝试对结果进行分析，哪些参数对实验结果影响比较大，原因是什么。

1. 实验方式

每位同学独立上机编程实验，由助教进行实验指导。

1. 参考内容

Numpy相关：<https://www.runoob.com/numpy/numpy-tutorial.html>

Scikit-learn包：<https://scikit-learn.org/stable/>

SVM函数：[https://scikit-learn.org/stable/modules/svm.html#svm-classification](https://scikit-learn.org/stable/modules/svm.html%23svm-classification)

PCA函数：[https://scikit-learn.org/stable/modules/decomposition.html#principal-component-analysis-pca](https://scikit-learn.org/stable/modules/decomposition.html%23principal-component-analysis-pca)

Scikit-learn函数库安装方法：使用命令 pip install scikit-learn

Numpy包安装方法：使用命令 pip install numpy

Pillow函数库安装方法：使用命令 pip install pillow

1. 实验报告要求
2. 现已给出PCA人脸降维重建算法实现，要求对给出的算法做到理解，更改给出的重建代码，实现对本班级采集人脸数据进行PCA降维重建实验，并调试参数n\_components的不同取值，查看不同取值对实验结果的影响，将不同取值对应的重建结果通过截图记录，并尝试对结果进行分析，哪些参数对实验结果影响比较大，原因是什么。
3. 补全已给出的pca\_svm.py代码，使用本班级人脸采集数据进行PCA降维，然后使用降维后的数据进行SVM分类实验，并提交代码，调试不同参数，查看不同参数对实验结果的影响，将不同参数对应的训练集和测试集准确率通过表格记录，并尝试对结果进行分析，哪些参数对实验结果影响比较大，原因是什么。
4. 实验结果提交
   1. 将代码和实验报告打包为“实验2-学号-姓名.zip”，其中实验报告的命名是“实验2-实验报告-学号-姓名.docx”；
   2. Zip文件通过智慧树作业页面进行提交。