**实验3：聚类与分类实验指南**

1. **实验目的**
2. 通过实验了解聚类与分类的应用场景；
3. 熟练掌握至少两种聚类算法的原理，并了解不同算法的优劣势；
4. 熟练掌握至少一种分类算法的原理。
5. **实验步骤**

## Python 环境安装

**建议安装python3**

1. 安装原生Python环境

直接在官网下载对应平台（Linux、Windows、Mac OS）最新版本的python，参考文档和教程安装。

Python官网：[https://www.python.org](https://www.python.org/)

Python文档地址：<https://docs.python.org/3/>

1. **安装Anaconda（推荐！）**

Anaconda 提供了很多基础依赖库，并支持自己创建多个不同虚拟环境，且通过conda install命令能自动安装相关依赖库，十分方便。

Anaconda 下载地址<https://www.anaconda.com/distribution/>下载后直接根据提示安装即可。

检验安装成功：在命令行输入 python –-version 能够看到安装的对应版本python。

或者输入conda 可以看到Anaconda相应的提示即可。

## 聚类算法实现

1. clustering文件夹下的todo.py文件部分：

* 在kmeans函数下实现k-means或k-medoids算法
* 在spectral函数下实现谱聚类算法

1. clustering文件夹下的clustering.py文件部分

* 调试knn\_graph的参数，优化谱聚类的效果
* 运行clustering.py文件，保存两种算法的聚类结果，比较分析其不同

## 分类算法实现

1. classification文件夹下todo.py文件部分：

* 在func函数下实现LR、SVM和感知机（Perceptron）三种分类算法的其中一种

1. 在classification文件夹下classification.py文件部分：

* 选择划分训练集和测试集的比例
* 编写计算训练和测试错误率的代码
* 保存分类结果，报告训练和测试错误率

加分项：实现多个分类算法，比较分析其异同

1. **提交内容**

## 实验报告：包含上述第二、第三项实验内容。

## 实验代码：第二、第三项实验的完整代码。

实验3提交方式：“学在浙大”平台