python与数据挖掘第一次作业

姓名：佘浩楠

学号：2019102965

专业：软件工程

导师：王文永

方向：图像处理、模式识别（具体未定）

我们在研一上学期并没有确定研究的方向，大概方向是图像处理，未来细致方向或许与教育、医疗相关。所以谈及python在此方向的作用并没有具体进行实践及应用过。

就我前期查找的相关资料来讲，图像处理用到的主要编程语言有Matlab、c/c++ 、python等。原因很简单，他们都有很多第三方库，不用我们从零开始编程。

在进行数字图像处理时，我们经常需要对图像进行读取、保存、缩放、裁剪、旋转、颜色转换等基本操作。在使用python进行编程时，涉及到多个不同的图像处理库的选择，opencv、scikit-image、scipy、pillow、matplotlib等等。

以下内容是我查找的一些之后可能会用到的一些常用库：

1. PIL(Python Imaging Library)

PIL（Python Imaging Library）是Python常用的图像处理库，而Pillow是PIL的一个友好Fork，提供了了广泛的文件格式支持，强大的图像处理能力，主要包括图像储存、图像显示、格式转换以及基本的图像处理操作等。

2. scipy.misc

python在科学计算领域有三个非常受欢迎库，numpy、SciPy、matplotlib。numpy是一个高性能的多维数组的计算库，SciPy是构建在numpy的基础之上的，它提供了许多的操作numpy的数组的函数。SciPy是一款方便、易于使用、专为科学和工程设计的python工具包，它包括了统计、优化、整合以及线性代数模块、傅里叶变换、信号和图像图例，常微分方差的求解等。

3. Opencv

OpenCV中已经包含如下应用领域功能：二维和三维特征工具箱、运动估算、人脸识别系统、姿势识别、人机交互、移动机器人、运动理解、对象鉴别、分割与识别、立体视觉、运动跟踪、增强现实（AR技术）。基于上述功能实现需要，OpenCV中还包括以下基于统计学机器学习库：Boosting算法、Decision Tree(决策树)学习、Gradient Boosting算法、EM算法(期望最大化)、KNN算法、朴素贝叶斯分类、人工神经网络、随机森林、支掌向量机。

4. matplotlib

我们处理数学运算、绘制图表，或者在图像上绘制点、直线和曲线时，Matplotlib 是个很好的类库，具有比 PIL 更强大的绘图功能。Matplotlib 可以绘制出高质量的图表，就像书中的许多插图一样。Matplotlib 中的 PyLab 接口包含很多方便用户创建图像的函数。

5. skimage

scikit-image是一组用于图像处理和计算机视觉的算法。“skimage”的主要包只提供了一些用于转换图像数据类型的实用程序;大多数功能程序存在其子包中。读取功能包含在io模块中。