**

**HUNAN UNIVERSITY**

《DM&BID》

Report 8

|  |  |
| --- | --- |
| **报告名称：** | K近邻算法 |
| **学生姓名：** | 杨超然 |
| **学生学号：** | 202106060220 |
| **专业班级：** | 电商2102班 |
| **学 院：** | 工商管理学院 |
| **指导老师：** | 江资斌 |
| **日 期：** | 2023.3.20 |

目录

[一、K近邻算法 2](#_Toc1572861201)

[（一）K近邻算法代码实现 2](#_Toc1701994284)

[1.1模型搭建与训练 2](#_Toc935753148)

[1.2模型预测 3](#_Toc1180411455)

[1.3K近邻算法回归模型 3](#_Toc721393199)

[二、 数据归一化 4](#_Toc1910308278)

[（一）数据归一化代码实现 4](#_Toc1670705696)

[1.1min-max标准化 4](#_Toc1201948147)

[1.2Z-score标准化 4](#_Toc1911322947)

[三、 手写数字识别模型 5](#_Toc1544378403)

[（一）模型搭建 5](#_Toc1880461579)

[（二）模型预测及评估 5](#_Toc400925354)

[四、 图像识别原理 6](#_Toc1696224039)

[（一） 图片大小调整及显示 6](#_Toc592009548)

[（二） 图片灰度处理 6](#_Toc612736685)

[（三） 图片二值化处理 7](#_Toc1081377430)

[（四） 二维数组转一维数组 7](#_Toc556361449)

[五、 学习小记 8](#_Toc623074305)

[（一） K近邻算法原理 8](#_Toc879581363)

# 一、K近邻算法

## （一）K近邻算法代码实现

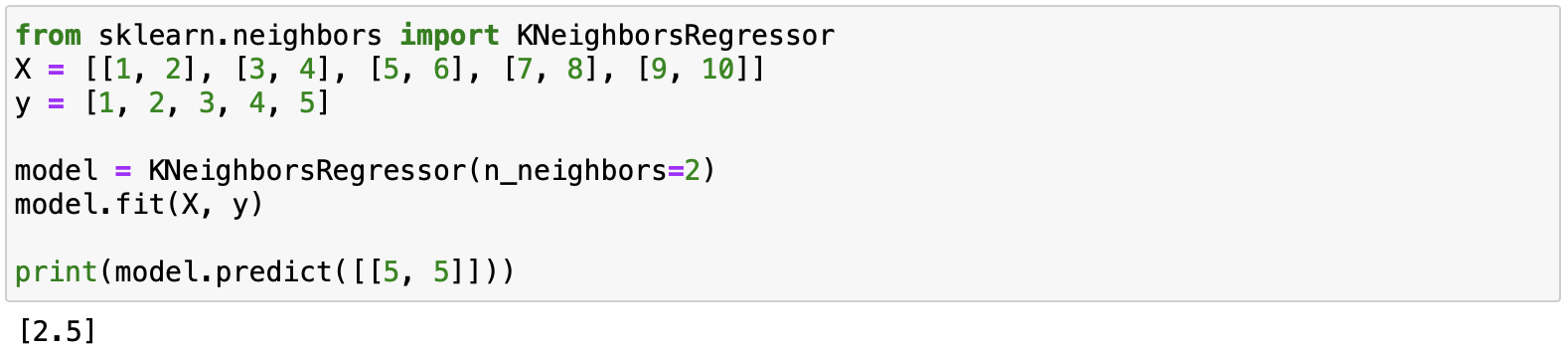
1.1模型搭建与训练



1.2模型预测



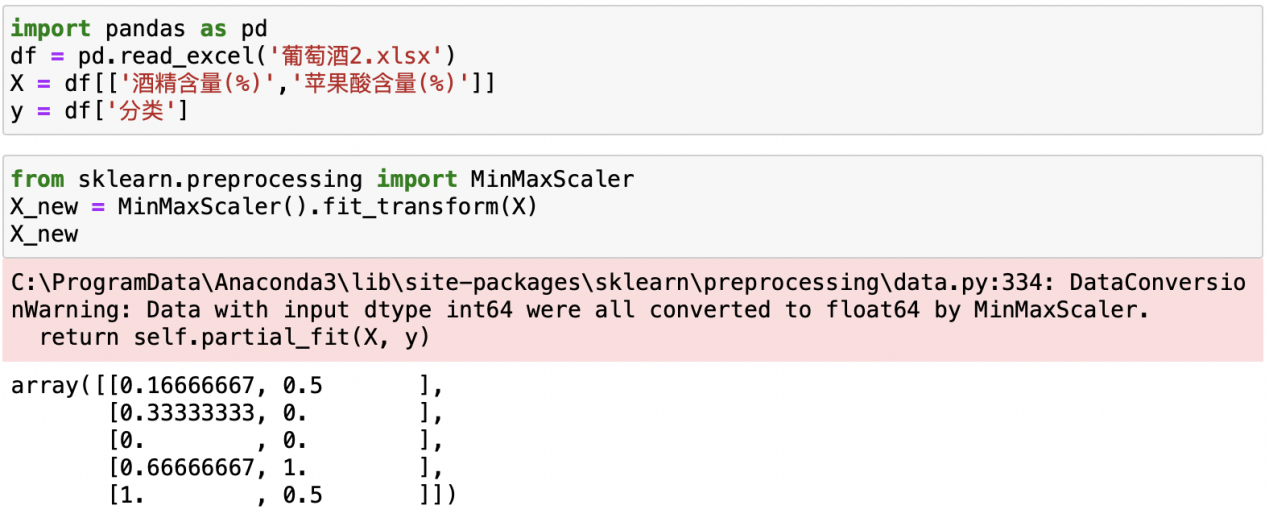
1.3K近邻算法回归模型



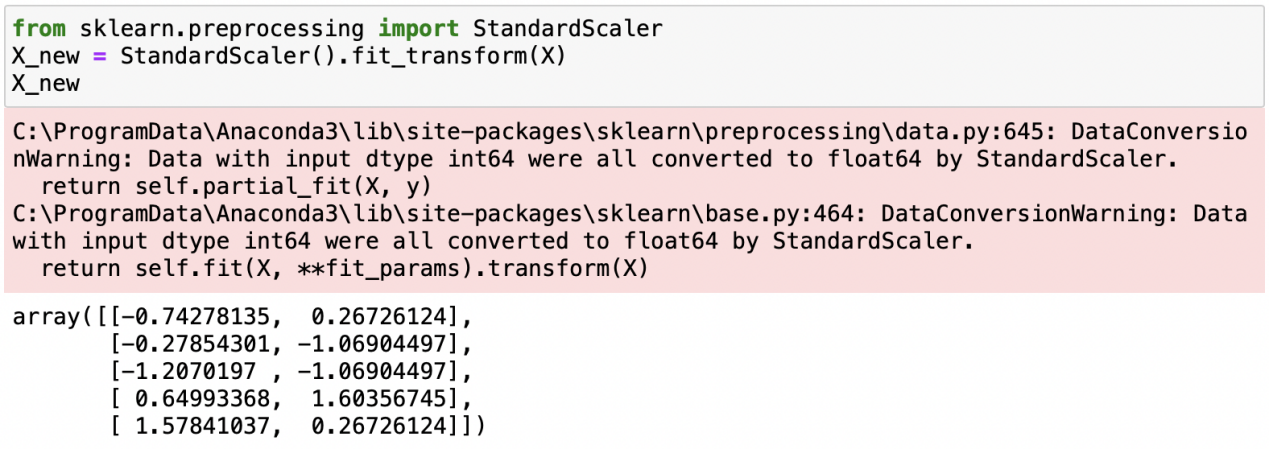
# 数据归一化

## （一）数据归一化代码实现

1.1min-max标准化



1.2Z-score标准化



# 手写数字识别模型

## （一）模型搭建

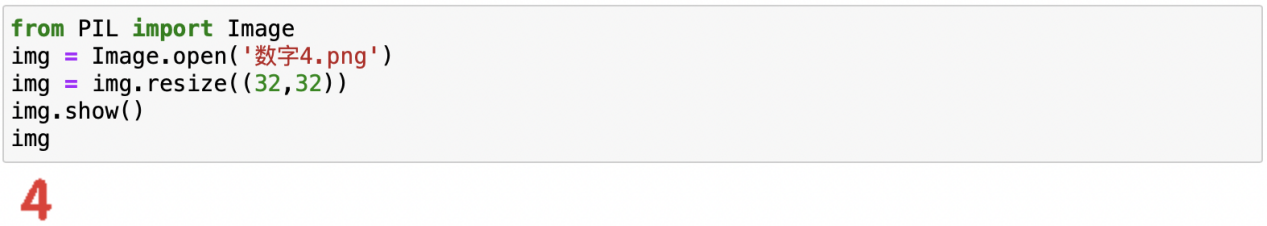


## （二）模型预测及评估

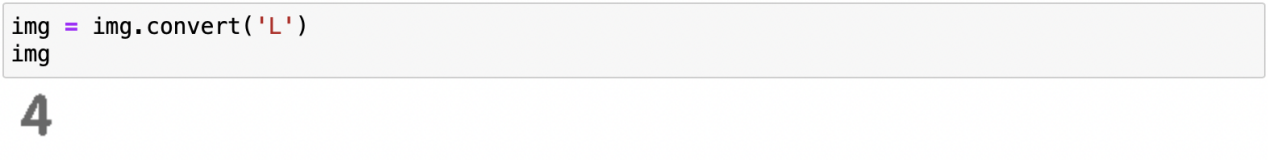


# 图像识别原理

## 图片大小调整及显示

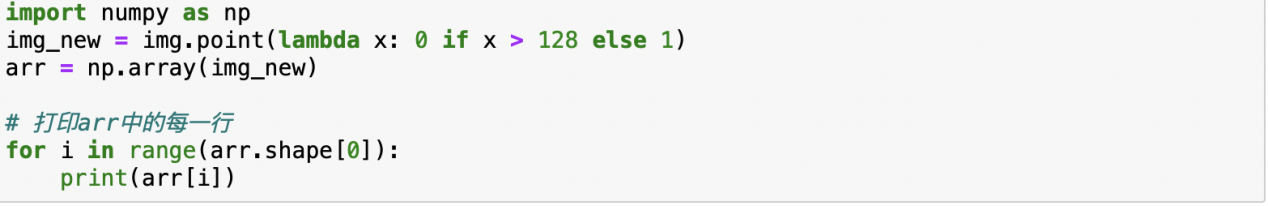


## 图片灰度处理

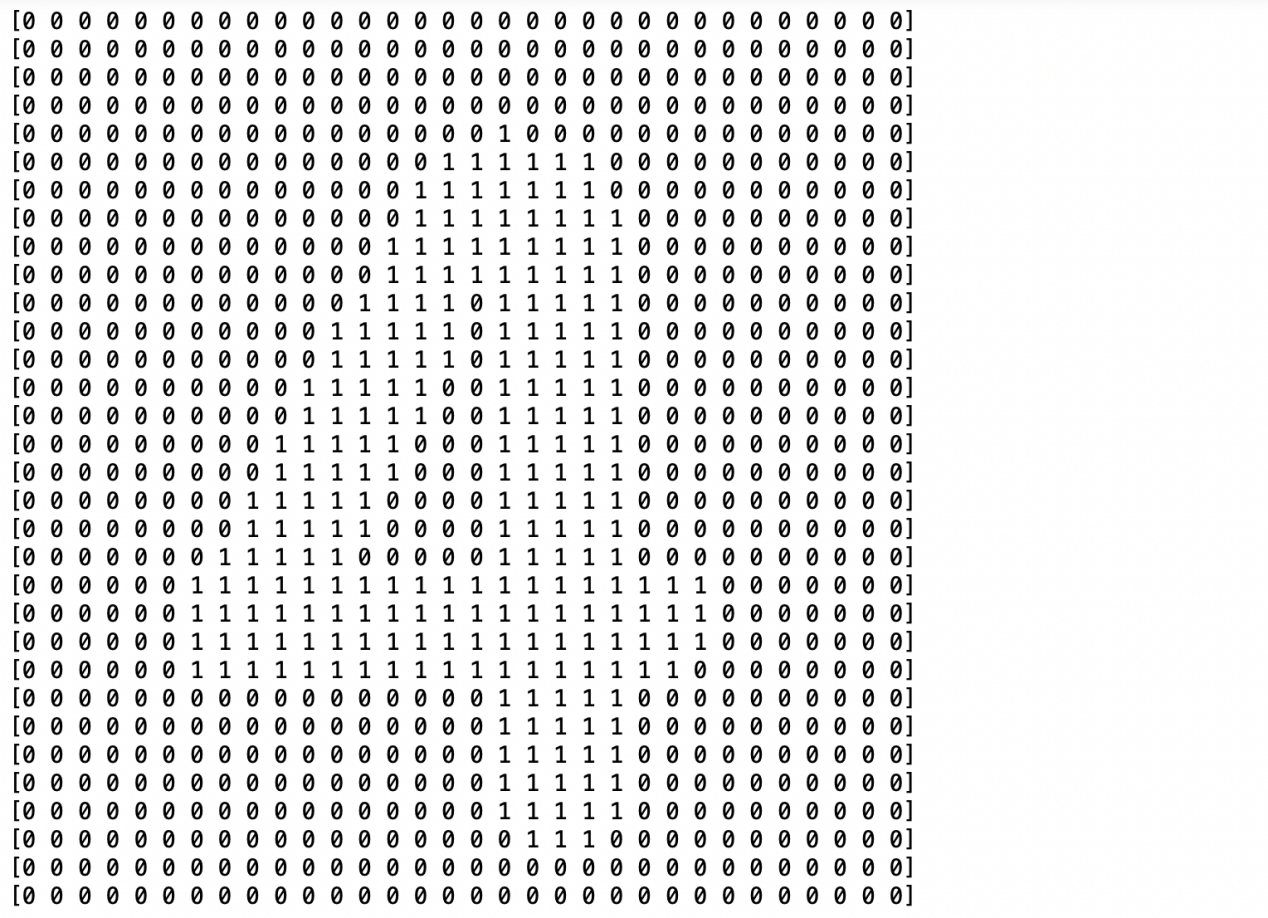


## 图片二值化处理

即通过如下代码将图片转换为计算机能够识别的内容：数字0和1，并将二值化后的图片并打印

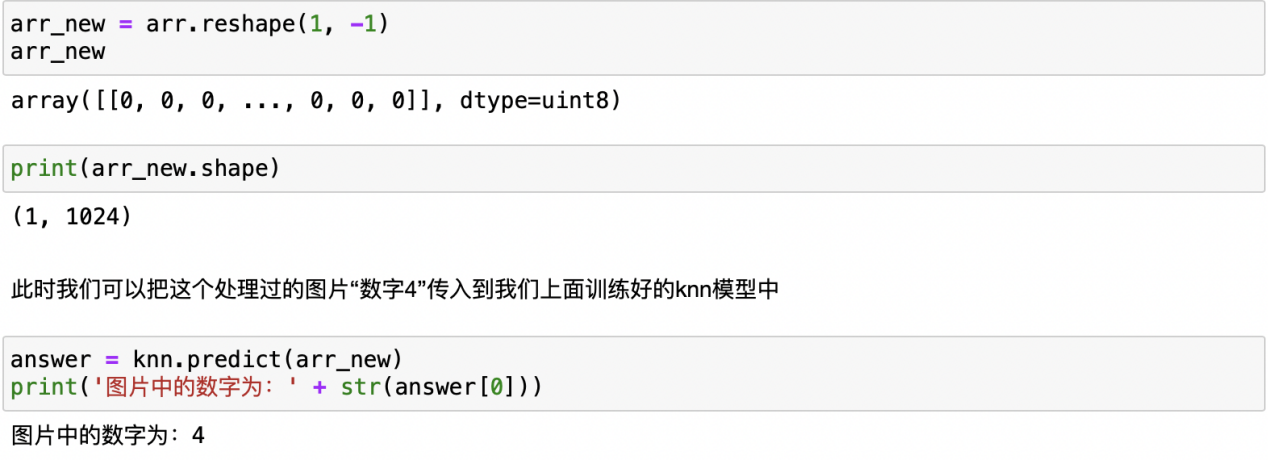


结果如下：



## 二维数组转一维数组

经过图像二值化的处理之后，我们获得了一个二维的数组(32x32的0-1矩 阵)，为了方便进行机器学习建模，还需要对这个二维数组进行一个简单处理:在上图二维数组中，第一行之后依次拼接第二行的32个数字，第三行的32个数字……直至第32行的32个数字，这时候便得到一个1x1024的一维数组。



# 学习小记

## K近邻算法原理

对于一个新的数据而言，K近邻算法的目的就是在已有数据中**寻找与它最相似的K个数据**，或者说“离它最近”的K个数据， **如果这K个点大多数属于某一个类别，则该样本也属于这个类别。**

