# 特征预处理

|  |
| --- |
|  |

## 1.什么是特征预处理

|  |
| --- |
|  |

## 2.包含内容

|  |
| --- |
|  |

## 3.预处理模块

|  |
| --- |
|  |

### 为什么要进行标准化/归一化?

|  |
| --- |
|  |

## 4.归一化

|  |
| --- |
|  |
|  |
|  |

### 应用举例

|  |
| --- |
|  |

### 这个例子需要用到pandas,我们已经把这些数据放到一个叫做dating.txt文件中

### 代码如下

|  |
| --- |
| # 归一化实例 import pandas as pd from sklearn.preprocessing import MinMaxScaler  """  数据需要从dating.txt中加载,使用pandas来加载 """ # 1.加载数据 data = pd.read\_csv("./dating.txt") # 2.获取数据的前三列,因为第四列是目标值不能作为特征值 data = data.iloc[:,:3] # print(data) # 3.创建转换器 transfer = MinMaxScaler() # 4.归一化处理 ret\_data = transfer.fit\_transform(data) print(ret\_data) # 显示结果 |

### 效果

|  |
| --- |
| [[0.47073382 0.58819286 0.53237967]  [0.06187353 0.48794044 1. ]  [0.24074991 0. 0.43571351]  [1. 1. 0.19139157]  [0.27620344 0.01947089 0. ]  [0.31604999 0.58819286 0.53237967]  [0.21655736 0.48794044 1. ]  [0.14793961 0. 0.43571351]  [0.53594852 1. 0.19139157]  [0.12151961 0.01947089 0. ]  [0.16136617 0.58819286 0.53237967]  [0. 0.48794044 1. ]  [0.14793961 0. 0.43571351]  [0.69063235 1. 0.19139157]  [0.89493875 0.01947089 0. ]  [0.31604999 0.58819286 0.53237967]  [0.21655736 0.48794044 1. ]  [0.28715506 0. 0.43571351]  [0.53594852 1. 0.19139157]  [0.74025492 0.01947089 0. ]] |