

一、概述

第五章将搭建一个推荐引擎。我们将分为两节，第一节学习一些机器学习中的基本方法。第二节构建推荐引擎。

二、内容

1、构造函数组合

这里用到python中的高级函数，就是一个简单的用法，可以在网上看下reduce和map函数的用法。

2、构建机器学习流水线

一个机器学习模型中，有很多可移动的组件需要被组合在一起，模型才能被执行并成功的得到结果。把机器学习过程的各个部分结合在一起的过程就被称作一条 pipeline（管道，流水线）。

代码中，我们生成了20维的向量特征，搭建特征选择器，随机森林分类器，最后构建机器学习流水线。

3.寻找最邻近

在数据集中找出所给一点，举例它最近的n个点即为寻找最邻近。使用NearestNeighbors模型。

4.构建KNN分类器和回归器

1、分类器

就是对数据进行一个分类，学习代码即可，没有特殊知识点。

2、回归器

这里我们生成0-1中随机100个数据作为X，添加噪点以及使用np.sinc函数进行x映射到y。进行回归训练后可看到回归结果。