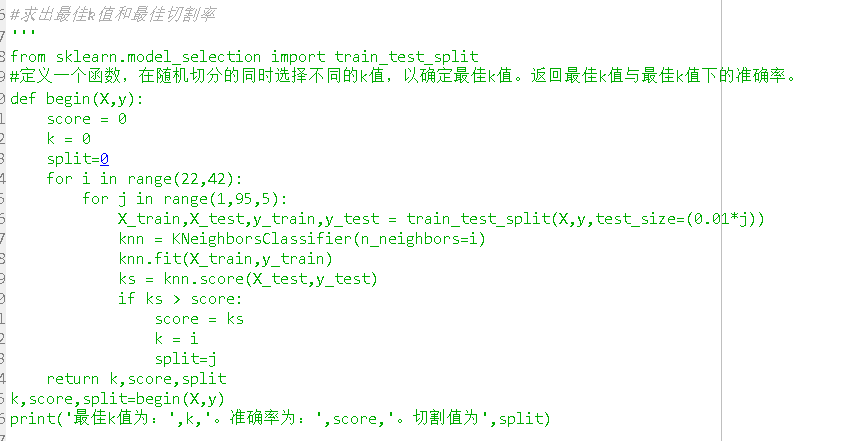
1. 本次任务我用的是knn算法。但是发现训练的效果不是很理想，正确率一直提升不上去。
2. 在这个过程中，我对一下遇到问题进行了解决
3. 由于数据集中存在了大量的“?”，所以必须把它们替换掉才能做操作，但由于它是字符串类型，所以我把它替换成了数据集中不存在的整形100，然后进行替换操作，将它们替换到本列属性的均值。



1. 在写代码期间为了提高正确率，我使用了对数据集的预处理，分别是标准化、归一化、正则化，但是发现效果并不是很显著。于是我就不再进行预处理



1. 然后我写了一个函数用来返回最佳k值、最佳切分点和正确率



（4）以上是我遇到的比较主要的问题。