**2019/03/22读书报告**

学号 71117228 姓名 李泓烨

读书进度：统计机器学习第3章

1. **读书报告内容**

**1. 自己提出问题的理解**

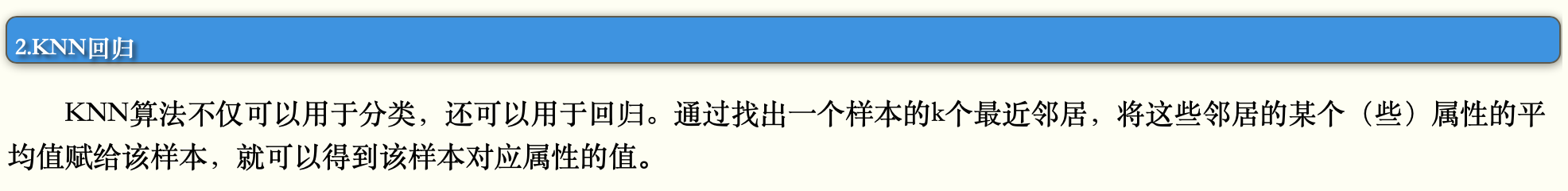
① 怎么理解超矩形区域的意思？

**讨论后的理解**：就是n维欧式空间中的n-1维度的线性子空间；超平面就是二维里面的直线，三维里面的一个面，加上“超”字，我个人认为就是拓展到高维空间里面

**2. 别人提出问题的理解**

1. **提出的问题1**：第三章开头说k近邻法是一种基本的分类与回归的方法，那如何将k近邻用于回归当中呢？是在找到k近邻之后求平均值这样的作法吗？

**讨论后的理解**：



2. **提出的问题2**：在最近邻的搜索算法中，当目标点和“当前最近点”之间的距离为半径构成的圆和另一子节点的区域相交时，会移动到另一个子节点当中递归地进行最近邻搜索。请问是怎么在另一子结点中搜索的？

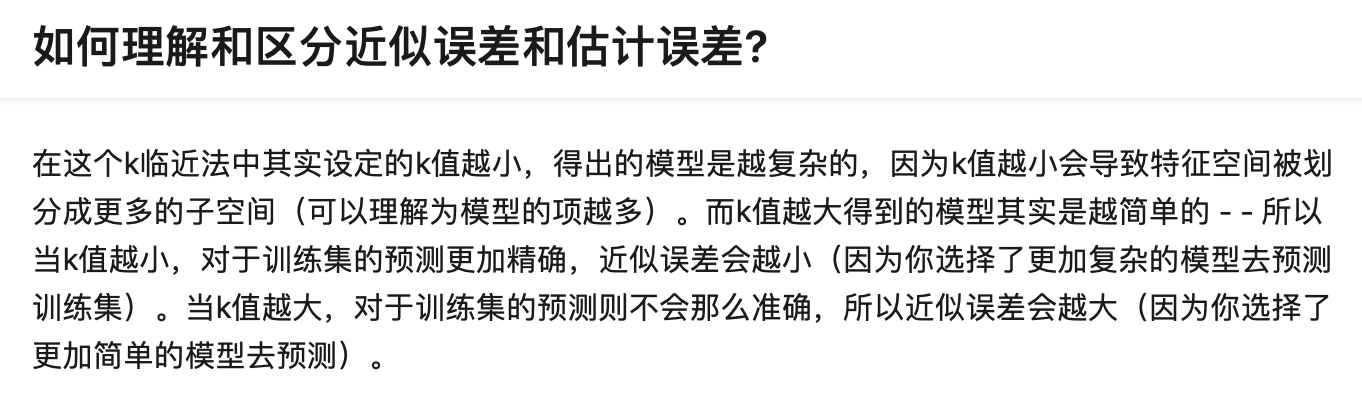
**小组成员的解释**：就是把搜索目标移动为另一子节点；应该是同样的操作直接遍历另一个子节点。

3. **提出的问题3**：k近邻算法要求找到最近的k个最近邻，但书上的kd tree算法貌似只找了最近的一个实例点？请问如何寻找k个近邻呢？我能想到的一种方法是找到最近邻的一个之后将此点去掉再寻找下一个最近邻，请问还有什么效率更高的方法吗？

**自己的理解**：可以用一个队列，边递归边存

4. **提出的问题4**：k值的选择与数据有没有什么联系，除了交叉验证法多次计算选取最优k值之外有无别的方法？

**自己的理解**：有一个经验范围，在那个范围里可以交叉验证



5. **提出的问题5**：如何更好地理解kd树的最近邻搜索（不是很理解里面的递归回退的过程）？

**自己的理解**：从底部向上回退才能搜索完整个数据集

**3. 读书计划**

① 本周完成的内容章节：看完了第3章

② 下周计划：看完第4章（视情况而定，暂时不知道第4章有多少）

**4. 读书摘要、理解&伪代码的具体实现**

1. 读书摘要

这里我以思维导图的形式记录了我的读书笔记（见下一页）

因为这一章公式和它们的推导比较多，我只记了一些很小的部分

