1. 自己提出的问题的理解（罗列全部）：
2. 提出的问题1：5.1.1中种子的类别标识表示了聚类结果中各个簇的类别标识是什么意思？

讨论后的理解：先对已有label的数据进行聚类，再利用这些粗来对其余无标注的数据进行聚类，而之前已有标注的数据就是种子。

1. 提出的问题2：co-training可否将属性分为三个或者多个集合来训练三个或者多个分类器进行学习呢？

讨论后的理解：分成多个集合的话，会使得第一个假设很难被满足，这样的情况下是没有讨论多个集合的意义的。

1. 别人提出的问题的理解（选择几个问题罗列，并给出理解）：
2. 问题3：最小割方法的代价函数|vi-vj|代表的是什么？是两点之前的距离吗？

自己的理解：不是距离，整个函数是为了保证相似度高的两个数据点的类别标识一致。因为相似度越高，则|vi-vj|的系数越大，则需要他们的值相同使其变为零才能达到最小。

1. 问题4：co-training的第二个假设，能否举一个通俗易懂的例子？

自己的理解：向身高和体重，他们实际上隐含有一定的联系的，身高越高，体重越大。

1. 问题5：self-training是不是会导致泛化能力不够强？因为它是用自己训练的模型来教自己，会不会产生很大的误差？  
   自己的理解：若出现这样的情况，可以加一些调整，若要改进的话只能看怎么选择无标签数据。
2. （必填）读书计划
3. 本周完成的内容章节：第五章
4. 下周计划：进行统计学习方法的阅读

四、读书摘要及理解或伪代码的具体实现（读书摘要、伪代码的具体实现代码等可以写到这个部分）

读书摘要及理解：