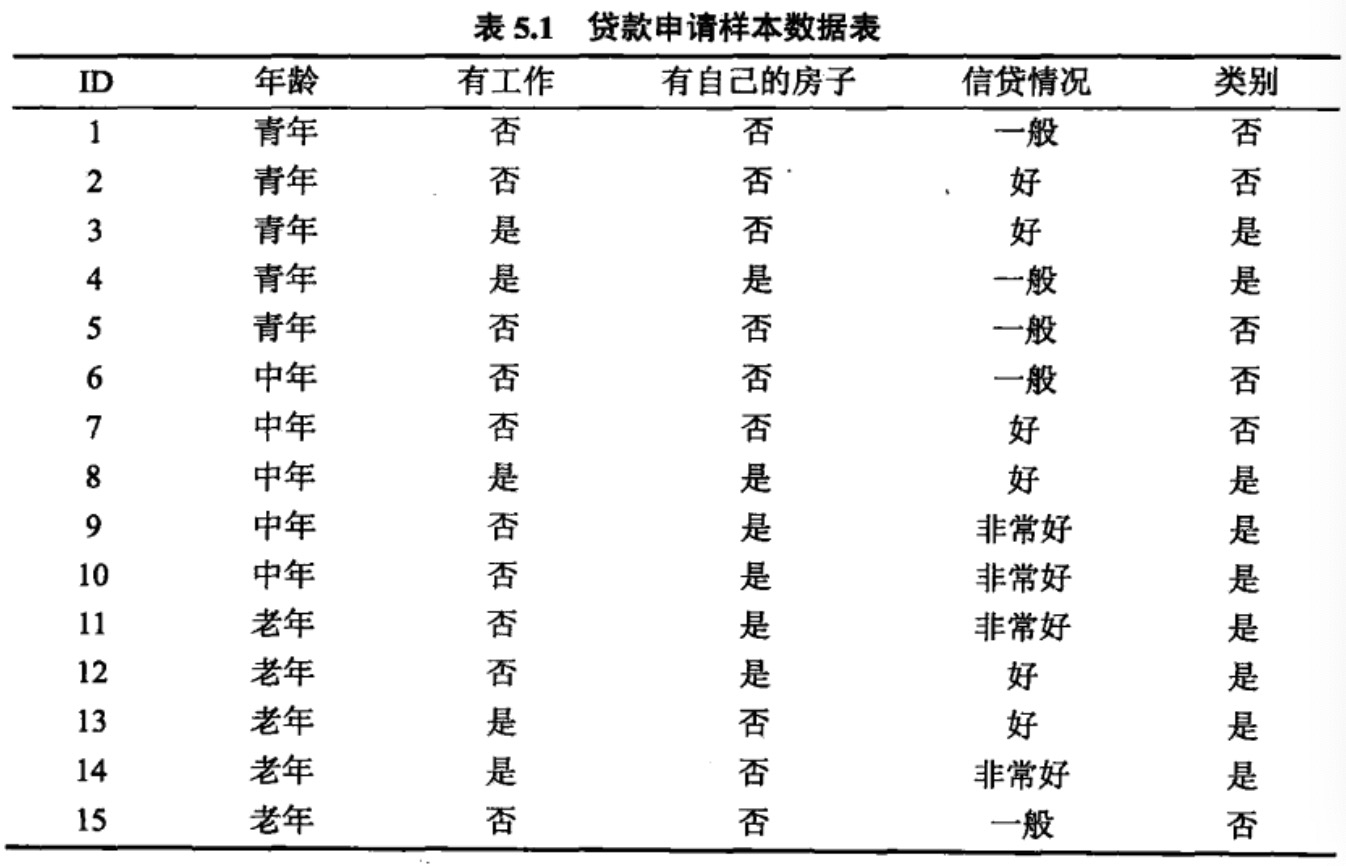
1. 下表是一个由15个样本组成的贷款申请训练数据，包括四个特征（年龄，有无工作，有无房屋，信贷情况），最后一列是类别，表示是否同意其贷款。问题如下：



（1）计算所有特征对上表中数据集的信息增益

（2）用ID3算法建立决策树

2. 用伪代码描述一种决策树剪枝的方法

3. 有N个样本,…, ，每个样本维数D，希望将样本维数降低到K，请给出PCA算法的计算过程