

作业

- 1.课上所有的案例都要自己手动完成
- 2.下面以常用的贷款申请样本数据表为样本集，计算信息增益计算过程，并构建ID3决策树。

ID	年龄	有工作	有房子	信贷情况	类别
1	青年	否	否	一般	拒绝
2	青年	否	否	好	拒绝
3	青年	是	否	好	同意
4	青年	是	是	一般	同意
5	青年	否	否	一般	拒绝
6	中年	否	否	一般	拒绝
7	中年	否	否	好	拒绝
8	中年	是	是	好	同意
9	中年	否	是	非常好	同意
10	中年	否	是	非常好	同意
11	老年	否	是	非常好	同意
12	老年	否	是	好	同意
13	老年	是	否	好	同意
14	老年	是	否	非常好	同意
15	老年	否	否	一般	拒绝

思路：

Step1 计算经验熵

Step2 各特征的条件熵

Step3 计算增益

最终构建的决策树

3.说明CART分类树和回归树的特点，并说明构建过程

4.说明决策树剪枝的方法有哪些？ 及各自的特点是什么？