

Welcome To: 机器学习及应用  
Machine Learning & Its Application

# 第 3 讲

## 决策树学习

### Decision Tree Learning

主讲: 金连文  
[Lianwen.Jin@gmail.com](mailto:Lianwen.Jin@gmail.com)

1

## Outline

- Story
- 决策树的一些概念及距离
- ID3学习算法
- ID3学习算法举例
- ID3算法的改进及讨论

2

## 一、Story: 生活中的树

3

## Story: 生活中的树

- 古时候,很多人到了弱冠之年, 将面临人生的第一棵树.....

《礼记·曲礼上》“二十曰弱,冠”。



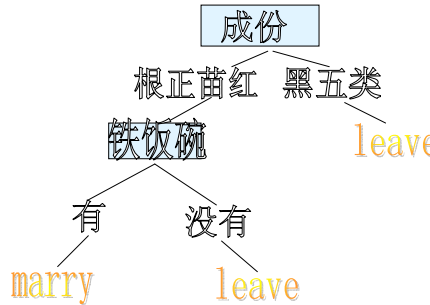
Marry

因为.....那时候,树只有一个叶子,你只能掉死在一片叶子上

4

## Story:生活中的树 (cont..)

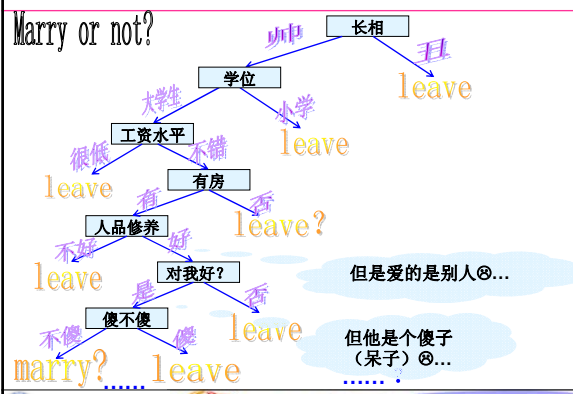
- 60年代



5

## Story :生活中的树 (90年代变得复杂起来)

Marry or not?



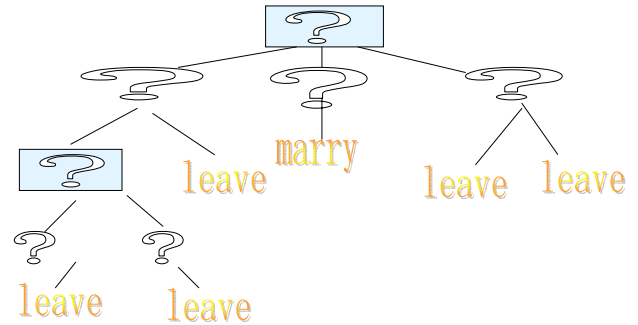
但是爱的是别人@...  
但他是个傻子 (傻子) @...  
.....

6

## Story :生活中的树 (cont..)

- 2000年代到现在,树就越来越复杂,主要表现为:
  - 每个的结点也不一定就是两个的分叉了.慢慢的多了很多个选择条件.比方说文化程度这个结点,就有:
    - 高中,专科,本科,硕士研究生,博士研究生.....
  - 甚至有些结点,是连续值,比如,多少\$ ¥ ?? !!!
  - 或者N多个离散值,比如有多少套房?
  - 还有有些节点,决策的时候不知道是否含有噪声的
    - 例如所陈述的回答与事实真心与否.....?
  - 那么,很多人困惑了.....
- 怎么选最优的树呢?

## Story :生活中的树 (cont..)



## 解决办法

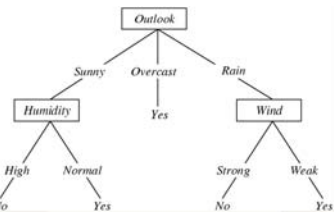
- 机器学习中的决策树方法 (决策树学习算法)
- 机器学习领域前辈及大牛之一Quinlan,J.R,在1983提出ID3决策树算法;
- 针对90后的问题, Quinlan,J.R与时俱进, 1993年正式提出了c4.5算法,并公布了源代码
  - JR Quinlan, C4.5: Programs for Machine Learning - 1993
- 2002年发表C5.0 (See5) 商业版
  - <http://www.rulequest.com/see5-info.html>
- 决策树的另一类家族:
  - CART (1977, 1984, Friedman& Breiman)



## 二、决策树的一些基本概念

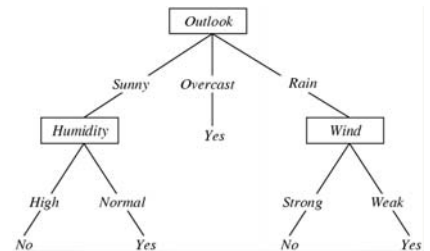
## 1、决策树

- A decision tree is a tree in which
  - each **branch node** represents a choice between a number of alternatives
  - each **leaf node** represents a classification or decision



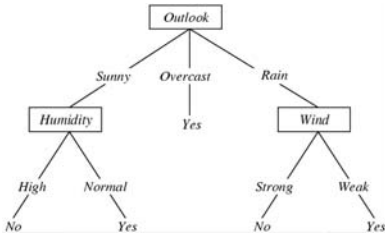
## 2. 节点

- 每一个节点表示对样本实例的某个属性值进行测试, 该节点后相连接的各条路径上的值表示该属性的可能取值 (二叉, 三叉, ...)



### 3. 叶子

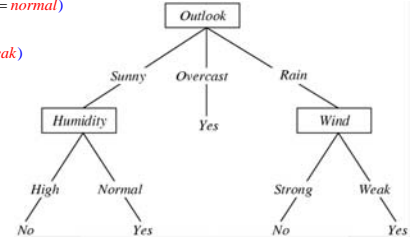
- 每一个叶子产生一条规则，规则由根到该叶子的路径上所有节点的条件，规则的结论是叶子上标注的结论（决策，分类，判断）



### 4. 决策树所产生的规则

- 决策树代表实例属性值约束的合取的析取式。从树根到树叶的每一条路径对应一组属性测试的合取，树本身对应这些合取的析取

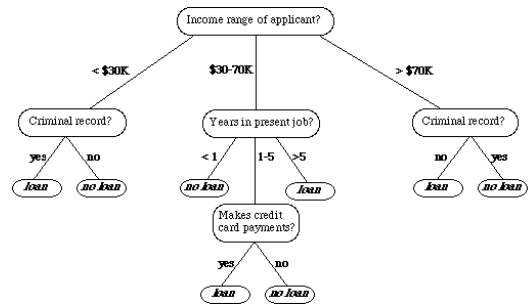
( $Outlook = sunny \wedge Humidity = normal$ )  
 $\vee (Outlook = overcast)$   
 $\vee (Outlook = rain \wedge Wind = weak)$



### 5. 决策树分类的适用性

- 实例由“属性-值”对表示
- 目标函数具有离散的输出值
- 可能需要析取的描述
- 训练数据可以包含错误
- 训练数据可以包含缺少属性值的实例

### 6. 举例1：银行贷款决策



### 6. 举例2：信用风险分析

节点属性可被不同的路径重复使用

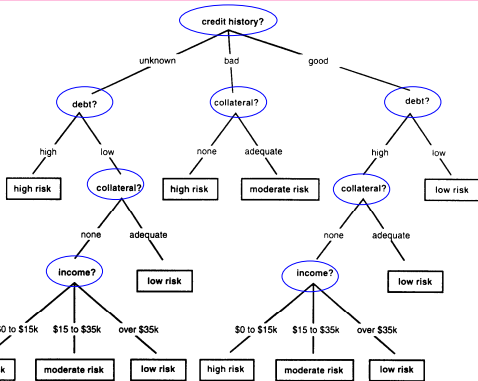


Figure 12.13 A decision tree for credit risk assessment.

### 6. 举例3：

- 根据身高、眼睛、头发判断东西方人



- |                         |                         |
|-------------------------|-------------------------|
| (tall, blond, blue) w   | (short, black, brown) e |
| (short, silver, blue) w | (tall, silver, black) e |
| (short, black, blue) w  | (short, black, brown) e |
| (tall, blond, brown) w  | (tall, black, brown) e  |
| (tall, silver, blue) w  | (tall, black, black) e  |
| (short, blond, blue) w  | (short, blond, black) e |

### 三、ID3学习算法

## 1. ID3基本思想

### ■ ID3的思想

- 自顶向下构造决策树
- 从“哪一个属性将在树的根节点被测试”开始
- 使用统计测试来确定每一个实例属性单独分类训练样例的能力

### ■ ID3的过程

- 分类能力最好的属性被选作树的根节点
- 根节点的每个可能值产生一个分支
- 训练样例排列到适当的分支
- 重复上面的过程直到所有训练样本使用完毕

## 2. 用于概念学习的ID3算法

- ID3(Examples, Target\_attribute, Attributes)
- 创建树的root节点
- 如果Examples都为正,返回label=+的单节点树root
- 如果Examples都为反,返回label=-的单节点树root
- 如果Attributes为空,那么返回单节点root, label=Examples中最普遍的Target\_attribute值
- 否则开始
  - A←Attributes中分类examples能力最好的属性
  - root的决策属性←A
  - 对于A的每个可能值 $v_i$ 
    - 在root下加一个新的分支对应测试A= $v_i$
    - 令Examples $_{v_i}$ 为Examples中满足A属性值为 $v_i$ 的子集
    - 如果Examples $_{v_i}$ 为空
      - 在这个新分支下加一个叶子节点, 节点的label=Examples中最普遍的Target\_attribute值
      - 否则在新分支下加一个子树ID3 (Examples $_{v_i}$ , Target\_attribute, Attributes-{A})
- 结束
- 返回root

## 3. 重要问题：那个属性作为当前测试节点

### ■ 最佳分类属性问题

### ■ 信息增益(Information Gain)

- 用来衡量给定的属性区分训练样例的能力
- ID3算法在增长树的每一步使用信息增益从候选属性中选择属性

### ■ 用熵度量样例的均一性

- 熵刻画了任意样例集的纯度

## 4. 信息论小补

- 一个信息的大小是通过它的不确定性来衡量的。 $p(x)$ 表示随机变量X的概率,那么用 $s(X)$ 表示X的信息量 (stands for surprise)
  - If  $p(x)=1$ , there is no such surprise, thus  $s(X)=0$
  - If  $p(x)=0$ , there is a big surprise and  $s(x)=0$
- 一个随机变量X (假设X有n个可能值) 的平均信息量,也叫熵 (Entropy),是这样计算的:

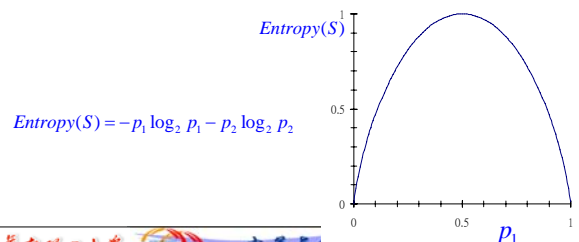
$$H(X) = E[\lg \frac{1}{P(X)}] = -\sum_{i=1}^n P(X^i) \lg P(X^i)$$

例子：书P41

## 4. 信息论小补 (cont)

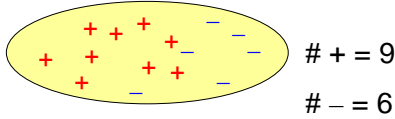
### ■ 熵 (Entropy)的物理意义:

- A formula to calculate the homogeneity of a sample.
- A completely homogeneous sample has entropy of 0.
- An equally divided sample has entropy of 1.



## Entropy (Binary Classification)

$$Entropy(S) = -p_1 \log_2 p_1 - p_2 \log_2 p_2$$



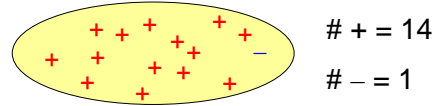
$$p_+ = 9/15 = 0.6$$

$$p_- = 6/15 = 0.4$$

$$Entropy = 0.97095$$

## Entropy (Binary Classification)

$$Entropy(S) = -p_1 \log_2 p_1 - p_2 \log_2 p_2$$



$$p_+ = 14/15$$

$$p_- = 1/15$$

$$Entropy = 0.353359$$

## 5. 信息增益 (Information Gain)

### 信息增益度量期望的熵降低

- > 属性的信息增益，由于使用这个属性分割样例而导致的期望熵降低

$$Gain(S, A) = Entropy(S) - \sum_{v \in \text{Values}(A)} \frac{|S_v|}{|S|} Entropy(S_v)$$

- > **Gain(S,A)**是在知道属性A的值后可以节省的二进制位数

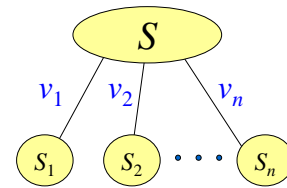
- > 上式中:

- **S**: 目标概念的正负样本的集合
- **S<sub>v</sub>**: 对某个属性v的正负样本的子集
- **Entropy(S<sub>v</sub>)**是将S<sub>v</sub>中的样本划分到c个类的信息熵
- **Value(A)**: 属性A所有可能取值的集合

## 信息增益 (Information Gain)

$$Gain(S, A) = Entropy(S) - \sum_{v \in A} \frac{|S_v|}{|S|} Entropy(S_v)$$

$$\text{Attribute } A = \{v_1, \dots, v_n\}$$



举例P42

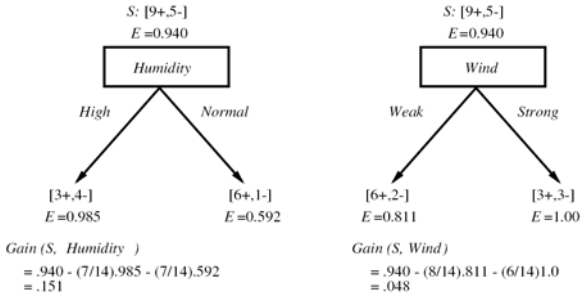
## 四、ID3算法举例

## 1. Textbook的例子

- **Play tennis 观测数据**
- **Target: 什么气候条件下打网球? (Q/A)**

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

# Which attribute is best



- Gain(S, Outlook)=0.246
- Gain(S, Temperature)=0.029

$(D1, D2, \dots, D14)$   
 $[9+,5-]$   
**Outlook**  
 Sunny Overcast Rain  
 $\{D1, D2, D8, D9, D11\}$   $\{D3, D7, D12, D13\}$   $\{D4, D5, D6, D10, D14\}$   
 $[2+,3-]$   $[4+,0-]$   $[3+,2-]$   
 ? Yes ?  
 Which attribute should be tested here?  
 $S_{Sunny} = \{D1, D2, D8, D9, D11\}$   
 $Gain(S_{Sunny}, Humidity) = .970 - (3/5) \cdot 0.0 - (2/5) \cdot 0.0 = .970$   
 $Gain(S_{Sunny}, Temperature) = .970 - (2/5) \cdot 0.0 - (3/5) \cdot 1.0 = .110$   
 $Gain(S_{Sunny}, Wind) = .970 - (2/5) \cdot 1.0 - (3/5) \cdot 0.18 = .818$

## 2. 例子2: 判断东西方人

- 属性: 身高、眼睛、头发

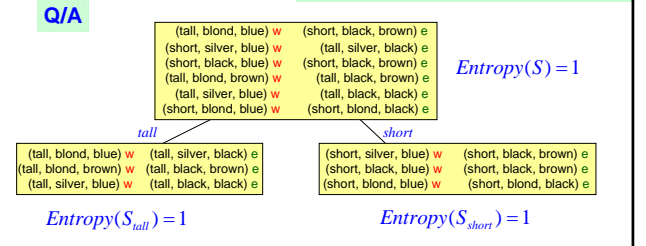


- |                         |                         |
|-------------------------|-------------------------|
| (tall, blond, blue) w   | (short, black, brown) e |
| (short, silver, blue) w | (tall, silver, black) e |
| (short, black, blue) w  | (short, black, brown) e |
| (tall, blond, brown) w  | (tall, black, brown) e  |
| (tall, silver, blue) w  | (tall, black, black) e  |
| (short, blond, blue) w  | (short, blond, black) e |

## 例子2

- 看看身高的信息增益

$$Gain(S, A) = Entropy(S) - \sum_{v \in A} \frac{|S_v|}{|S|} Entropy(S_v)$$



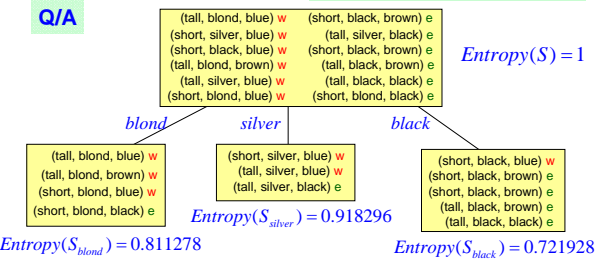
$$Gain(S, Height) = 0$$

## 例子2

- 头发?

Q/A

$$Gain(S, A) = Entropy(S) - \sum_{v \in A} \frac{|S_v|}{|S|} Entropy(S_v)$$



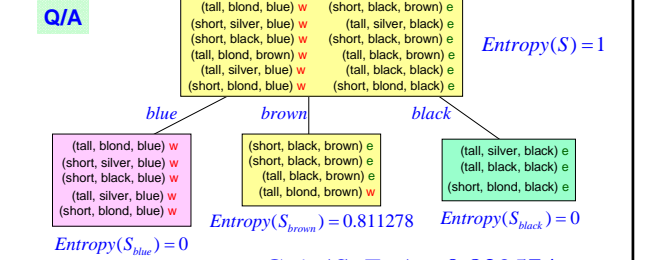
$$Gain(S, Hair) = 0.199197$$

## 例子2

- 眼睛颜色?

Q/A

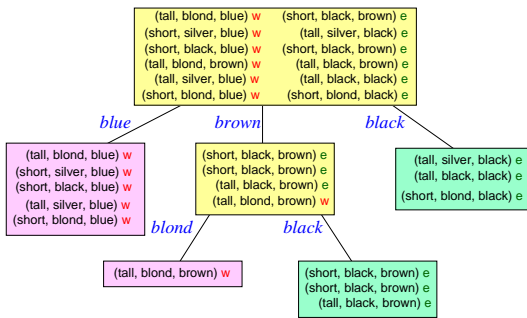
$$Gain(S, A) = Entropy(S) - \sum_{v \in A} \frac{|S_v|}{|S|} Entropy(S_v)$$



$$Gain(S, Eye) = 0.829574$$

- Yes! 眼睛颜色!

## 例子2: ID3得到的结果:



## 五、ID3算法讨论及一些改进措施

### 1. ID3算法讨论

- 观察ID3的搜索空间和搜索策略:
  - 假设空间包含所有的决策树, 它是关于现有属性的有限离散值函数的一个完整空间
  - 维护单一的当前假设 (不同于第二章的变型空间候选消除算法)
  - 不进行回溯, 可能收敛到局部最优
  - 每一步使用所有的训练样例, 不同于基于单独的训练样例递增作出决定, 容错性增强

### 2.ID3 的归纳偏置

- ID3的搜索策略
  - 优先选择较短的树
  - 选择那些信息增益高的属性离根节点较近的树
  - 很难准确刻画ID3的归纳偏置
- 近似的ID3的归纳偏置
  - 较短的树比较长的树优先
  - 近似在于ID3得到局部最优, 而不一定是全局最优
  - 一个精确具有这个归纳偏置的算法, BFS-ID3
- 更贴切近似的归纳偏置
  - 较短的树比较长的树优先, 信息增益高的属性更靠近根节点的树优先

### 3.为什么短的假设优先

- 哲学基础: 奥卡姆剃刀原理 (Occam's Razor):
  - 优先选择拟合数据的最简单的假设
- Story:
  - 奥卡姆剃刀定律, 是由14世纪逻辑学家奥卡姆的威廉 (William of Occam, 约1285年至1349年) 提出。奥卡姆在《箴言书注》2卷15题说“切勿浪费较多东西去做用较少的东西同样可以做好的事情。”
  - 定理的原述: “如无必要, 勿增实体” (Entities should not be multiplied unnecessarily)。
- 例子
  - 物理学家优先选择行星运动的简单假设
  - 简单假设的数量远比复杂假设的数量少
  - 简单假设对训练样例的针对性更小, 更像是泛化的规律, 而不是训练样例的另一种描述

### 4. ID3算法存储的主要问题

- ID3算法存在的主要不足
  - 过度拟合问题
  - 处理连续属性值问题
  - 处理缺少属性值问题
  - 属性选择的度量标准问题
  - 处理不同代价的属性问题
- 针对ID3的这些不足, ID3被扩展成为C4.5

## C4.5

- C4.5 is a algorithm extension of the basic ID3 algorithm to address the following issues not dealt with by ID3:
  - Avoiding overfitting the data
  - Determining how deeply to grow a decision tree.
  - Reduced error pruning.
  - Rule post-pruning.
  - Handling continuous attributes.
    - e.g., temperature
  - Choosing an appropriate attribute selection measure.
  - Handling training data with missing attribute values.
  - Handling attributes with differing costs.
  - Improving computational efficiency.

## 5. 决策树的常见问题

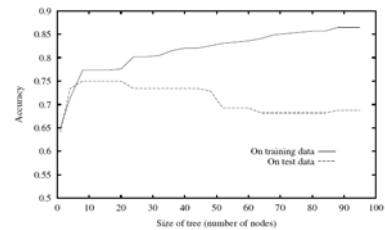
- 决策树学习的实际问题
  - 确定决策树增长的深度
  - 处理连续值的属性
  - 选择一个适当的属性筛选度量标准
  - 处理属性值不完整的训练数据
  - 处理不同代价的属性
  - 提高计算效率

## 6. 过学习、过度拟合 (Overfitting)

- 过学习的概念 (过度拟合, Overfitting)
  - 对于一个假设, 当存在其他的假设对训练样例的拟合比它差, 但事实上在实例的整个分布上表现得却更好时, 我们说这个假设过度拟合训练样例
  - 定义: 给定一个假设空间 $H$ , 一个假设 $h \in H$ , 如果存在其他的假设 $h' \in H$ , 使得在训练样例上 $h$ 的错误率比 $h'$ 小, 但在整个实例分布上 $h'$ 的错误率比 $h$ 小, 那么就称假设 $h$ 过度拟合训练数据。

## 6. 过度拟合 (Overfitting) (cont)

Overfitting in Decision Tree Learning



## 避免过拟合 (过学习)

- 导致过度拟合的原因
  - 一种可能原因是训练样例含有随机错误或噪声
  - 当训练数据没有噪声时, 过度拟合也有可能发生, 特别是当少量的样例被关联到叶子节点时, 很可能出现巧合的规律性, 使得一些属性恰巧可以很好地分割样例, 但却与实际的目标函数并无关系。
- 避免过度拟合的方法
  - 及早停止树增长
  - 后修剪法
- 两种方法的特点
  - 第一种方法更直观
  - 第一种方法中, 精确地估计何时停止树增长很困难
  - 第二种方法被证明在实践中更成功

## 避免过拟合 (过学习)

- 避免过度拟合的关键
  - 使用什么样的准则来确定最终正确树的规模
- 解决方法
  - 使用与训练样例截然不同的一套分离的样例, 来评估通过后修剪方法从树上修建节点的效用。
  - 使用所有可用数据进行训练, 但进行统计测试来估计扩展 (或修剪) 一个特定的节点是否有可能改善在训练集合外的实例上的性能。
  - 使用一个明确的标准来衡量训练样例和决策树的复杂度, 当这个编码的长度最小时停止树增长。



## 避免过拟合

- 方法评述
  - 第一种方法是最普通的，常被称为训练和验证集法。
  - 可用数据分成两个样例集合：
    - 训练集合，形成学习到的假设
    - 验证集合，评估这个假设在后续数据上的精度
  - 方法的动机：即使学习器可能会被训练集合误导，但验证集合不大可能表现出同样的随机波动
  - 验证集合应该足够大，以便它本身可提供具有统计意义的实例样本。
  - 常见的做法是，样例的三分之二作训练集合，三分之一作验证集合。
  - 交叉验证 (Cross-Validation)

## 7. 错误率降低修剪

- 将树上的每一个节点作为修剪得候选对象
- 修剪步骤
  - 删除以此节点为根的子树，使它成为叶结点
  - 把和该节点关联的训练样例的最常见分类赋给它
  - 反复修剪节点，每次总是选取那些删除后可以最大提高决策树在验证集合上的精度的节点
- 继续修剪，直到进一步的修剪是有害的为止
- 数据分成3个子集
  - 训练样例，形成决策树
  - 验证样例，修剪决策树
  - 测试样例，精度的无偏估计
- 如果有大量的数据可供使用，那么使用分离的数据集合来引导修剪

## 8 规则后修剪 (C4.5使用)

- 从训练集合推导出决策树，增长决策树直到尽可能地拟合训练数据，允许过度拟合发生
- 将决策树转化为等价的规则集合，方法是为从根节点到叶节点的每一条路径创建一条规则
- 通过删除任何能导致估计精度提高的前件来修剪每一条规则
- 按照修剪过的规则的估计精度对它们进行排序，并按这样的顺序应用这些规则来分类后来的实例
- 把决策树转化成规则集的好处
  - 可以区分决策节点使用的不同上下文
  - 消除了根节点附近的属性测试和叶节点附近的属性测试的区别
  - 提高了可读性

## 9. 处理连续值

- ID3被限制为取离散值的属性
  - 学习到的决策树要预测的目标属性必须是离散的
  - 树的决策节点的属性也必须是离散的
- 简单删除上面第2个限制的方法
  - 通过动态地定义新的离散值属性来实现，即先把连续值属性的值域分割为离散的区间集合
- 方法的扩展
  - 连续的属性分割成多个区间，而不是单一阈值的两个空间

Create a discrete attribute to test continuous

Temperature = S2 5

PlayTennis = S2 5

Temperature: 40 48 60 72 80 90

PlayTennis: No No Yes Yes Yes No

## 10. 选择其它度量标准

- 信息增益度量存在一个内在偏置，偏向具有较多值的属性
- 避免方法：其他度量函数，比如增益比率
- 增益比率通过加入一个被称作分裂信息的项来惩罚多值属性，分裂信息用来衡量属性分裂数据的广度和均匀性

$$\text{GainRatio}(S, A) = \frac{\text{Gain}(S, A)}{\text{SplitInformation}(S, A)}$$

$$\text{SplitInformation}(S, A) = - \sum_{i=1}^c \frac{|S_i|}{|S|} \log_2 \frac{|S_i|}{|S|}$$

where  $S_i$  is subset of  $S$  for which  $A$  has value  $v_i$

## 11. 缺少属性值的训练样例的处理

- Attribute  $A = \{x, y, z\}$   
 $\text{Gain}(S, A) = ?$
- Possible Approaches:
- Assign most common value of  $A$  to the unknown one.
  - Assign most common value of  $A$  with the same target value to the unknown one.
  - Assign probability to each possible value.

$$S = \left\{ \begin{array}{l} (\dots, x, \dots) + \\ (\dots, x, \dots) - \\ (\dots, x, \dots) + \\ (\dots, y, \dots) - \\ (\dots, y, \dots) + \\ (\dots, y, \dots) + \\ (\dots, y, \dots) + \\ (\dots, z, \dots) + \\ (\dots, z, \dots) + \\ (\dots, z, \dots) - \\ (\dots, ?, \dots) - \\ (\dots, ?, \dots) + \end{array} \right.$$

- 一种策略是赋给它节点n的训练样例中该属性的最常见值
- 另一种策略是赋给它节点n的被分类为c(x)的训练样例中该属性的最常见值
- 更复杂的策略，为A的每个可能值赋予一个概率，而不是简单地将最常见的值赋给A(x)

## 12. More example: 连续值情况

### The Simpsons: Male or Female?

Person	Hair Length	Weight	Age	Class
Homer	0"	250	36	M
Marge	10"	150	34	F
Bart	2"	90	10	M
Lisa	6"	78	8	F
Maggie	4"	20	1	F
Abe	1"	170	70	M
Selma	8"	160	41	F
Otto	10"	180	38	M
Krusty	6"	200	45	M
<b>Comic</b>	<b>8"</b>	<b>290</b>	<b>38</b>	<b>?</b>

Entropy(S) =  $-\frac{p}{p+n} \log_2 \left( \frac{p}{p+n} \right) - \frac{n}{p+n} \log_2 \left( \frac{n}{p+n} \right)$   
 Entropy(4F,5M) =  $-(4/9) \log_2(4/9) - (5/9) \log_2(5/9)$   
 = 0.9911

yes Hair Length <= 5? no

Entropy(1F,3M) =  $-(1/4) \log_2(1/4) - (3/4) \log_2(3/4)$   
 = 0.8113

Entropy(3F,2M) =  $-(3/5) \log_2(3/5) - (2/5) \log_2(2/5)$   
 = 0.9710

Let us try splitting on Hair length

Gain(A) = E(Current set) -  $\sum$  E(all child sets)

Gain(Hair Length <= 5) = 0.9911 - (4/9 \* 0.8113 + 5/9 \* 0.9710) = 0.0911

Entropy(S) =  $-\frac{p}{p+n} \log_2 \left( \frac{p}{p+n} \right) - \frac{n}{p+n} \log_2 \left( \frac{n}{p+n} \right)$   
 Entropy(4F,5M) =  $-(4/9) \log_2(4/9) - (5/9) \log_2(5/9)$   
 = 0.9911

yes Weight <= 160? no

Entropy(4F,1M) =  $-(4/5) \log_2(4/5) - (1/5) \log_2(1/5)$   
 = 0.7219

Entropy(0F,4M) =  $-(0/4) \log_2(0/4) - (4/4) \log_2(4/4)$   
 = 0

Let us try splitting on Weight

Gain(A) = E(Current set) -  $\sum$  E(all child sets)

Gain(Weight <= 160) = 0.9911 - (5/9 \* 0.7219 + 4/9 \* 0) = 0.5900

Entropy(S) =  $-\frac{p}{p+n} \log_2 \left( \frac{p}{p+n} \right) - \frac{n}{p+n} \log_2 \left( \frac{n}{p+n} \right)$   
 Entropy(4F,5M) =  $-(4/9) \log_2(4/9) - (5/9) \log_2(5/9)$   
 = 0.9911

yes age <= 40? no

Entropy(3F,3M) =  $-(3/6) \log_2(3/6) - (3/6) \log_2(3/6)$   
 = 1

Entropy(1F,2M) =  $-(1/3) \log_2(1/3) - (2/3) \log_2(2/3)$   
 = 0.9183

Let us try splitting on Age

Gain(A) = E(Current set) -  $\sum$  E(all child sets)

Gain(Age <= 40) = 0.9911 - (6/9 \* 1 + 3/9 \* 0.9183) = 0.0183

Of the 3 features we had, Weight was best. But while people who weigh over 160 are perfectly classified (as males), the under 160 people are not perfectly classified... So we simply recurse!

This time we find that we can split on Hair length, and we are done!

We don't need to keep the data around, just the test conditions.

How would these people be classified?

```

graph TD
    A[Weight <= 160?] -- yes --> B[Hair Length <= 2?]
    A -- no --> C[Male]
    B -- yes --> D[Male]
    B -- no --> E[Female]
  
```

It is trivial to convert Decision Trees to rules:

**Rules to Classify Males/Females**

If *Weight* greater than 160, classify as **Male**  
 Elseif *Hair Length* less than or equal to 2, classify as **Male**  
 Else classify as **Female**

61

## Further Reading

- **C4. 5: Programs for Machine Learning**  
 JR Quinlan -Morgan Kauffman, Los Altos, CA , 1993 ( [books.google.com](http://books.google.com) )

62

## Homework

- **Textbook**
  - > 习题3.2, 3.4
- **大作业1:**
  - > 在后续的章节学习过程中, 用生动形象的示例阐述某个机器学习算法的原理、应用、物理意义、问题的提出及解决路线 等问题 (选其一)。制作 PPT, or Flash.
  - > **Deadline:** November 30, 2008
  - > **Email to tutor:** [nicholasdk8497@gmail.com](mailto:nicholasdk8497@gmail.com)

63

## 通知

- 《机器学习及其应用》课程网站:
  - > [http://www.hcii-lab.net/lianwen/Course/Machine\\_Learning/](http://www.hcii-lab.net/lianwen/Course/Machine_Learning/)
- 可在上述网页下载课程课件

64

## End of today's lecture

Thanks!

金连文 博士, 教授  
**Email:** [lianwen.jin@gmail.com](mailto:lianwen.jin@gmail.com)  
**Web:** <http://www.hcii-lab.net/lianwen/>

65