

习题 6

6.1 什么是学习和机器学习？为什么要研究机器学习？

解：略。

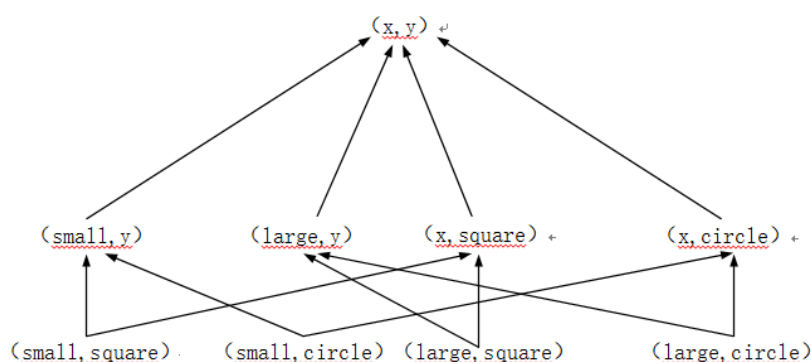
6.2 简单的学习模型是由哪几部分组成的？各部分的功能是什么？

解：略。

6.3 到目前为止，机器学习的方法有哪些？如何对它们进行分类？

解：略。

6.4 现假设有一个物体，用两个属性来描述：大小和形状。大小只有两个值：大和小；形状也有两个值：圆（Circle）、方（Square）。每个物体可以用一个向量表示 (x,y) ： x 表示物体的大小， y 表示物体的形状。初始变型空间可以用下图进行描述。用候选消除算法学习“圆”概念，即 $(x \text{ circle})$ ，给出学习过程。



解：首先将集合 H 初始化为整个规则空间， $\langle G, S \rangle$;

$G = \{(x,y)\}$

$S = \{(small, square), (small, circle), (large, square), (large, circle)\}$

提供一个正例 $(small, circle)$ ，对 $\langle G, S \rangle$ 进行更新：

$G = \{(x,y)\}$

$S = \{(small, circle)\}$

新的描述空间如图 1 所示。

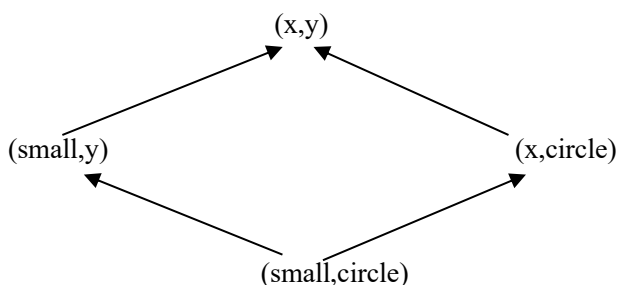


图 1 第一个正例后的变型空间

第二个实例为反例 $(large, square)$ ，对 G 集合进行特殊化： $G = \{(small, y), (x, circle)\}$

S 不变，新的描述空间如图 2 所示。

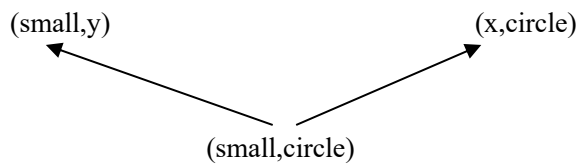


图2 第二个反例后的变型空间

最后一个正例为(large,circle)，从 S 集合中删去不覆盖该例的元素(small,y)，并对 S 集合进行一般化，得到： $G=S=\{(x,circle)\}$

从而算法停止，并给出学习得到的概念(x,circle)。

6.5 设训练例子集如下表所示：

| 序号 | 属性 | | 分类 |
|----|----------------|----------------|----|
| | x ₁ | x ₂ | |
| 1 | T | T | + |
| 2 | T | T | + |
| 3 | T | F | - |
| 4 | F | F | + |
| 5 | F | T | - |
| 6 | F | T | - |

请用 ID3 算法完成其学习过程。

解：设根节点为 S，具有最大的信息熵：

$$H(S) = -(P(+) \log_2 P(+) + P(-) \log_2 P(-)) = -((3/6) \log_2 (3/6) + (3/6) \log_2 (3/6)) = 1$$

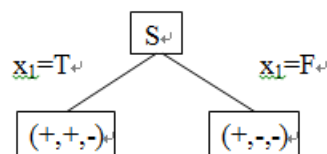
S 关于属性 x₁ 的条件熵：

$$H(S|x_1) = (|S_T|/|S|)H(S_T) + (|S_F|/|S|)H(S_F) = 0.9183$$

S 关于属性 x₂ 的条件熵：

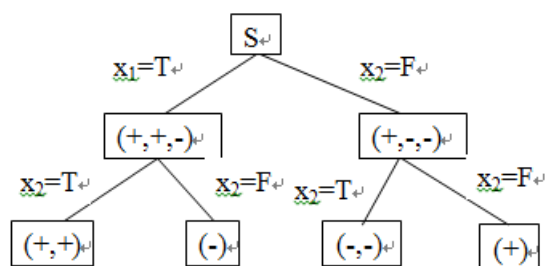
$$H(S|x_2) = (|S_T|/|S|)H(S_T) + (|S_F|/|S|)H(S_F) = 1$$

选择属性 x₁ 对根节点进行扩展。用 x₁ 对 S 扩展后所得到的部分决策树如下图所示。



扩展 x₁ 后的部分决策树

在该决策树中，其 2 个叶节点均不是最终决策方案，因此还需要继续扩展。而要继续扩展，只有属性 x₂ 可选择，因此不需要再进行条件熵的计算，可直接对属性 x₂ 进行扩展。对 x₂ 扩展后所得到的决策树如下图所示：



扩展 x_2 后得到的完整决策树

6.6 下图给出了一个可能带有噪音的数据集合。它有四个属性，Outlook、Temperature、Humidity、Windy。它被分为两类，P 与 N，分别为正例与反例。试构造决策树将数据进行分类。

| 属性 | Outlook | Temperature | Humidity | Windy | 类 |
|----|----------|-------------|----------|--------|---|
| 1 | Overcast | Hot | High | Not | N |
| 2 | Overcast | Hot | High | Very | N |
| 3 | Overcast | Hot | High | Medium | N |
| 4 | Sunny | Hot | High | Not | P |
| 5 | Sunny | Hot | High | Medium | P |
| 6 | Rain | Mild | High | Not | N |
| 7 | Rain | Mild | High | Medium | N |
| 8 | Rain | Hot | Normal | Not | P |
| 9 | Rain | Cool | Normal | Medium | N |
| 10 | Rain | Hot | Normal | Very | N |
| 11 | Sunny | Cool | Normal | Very | P |
| 12 | Sunny | Cool | Normal | Medium | P |
| 13 | Overcast | Mild | High | Not | N |
| 14 | Overcast | Mild | High | Medium | N |
| 15 | Overcast | Cool | Normal | Not | P |
| 16 | Overcast | Cool | Normal | Medium | P |
| 17 | Rain | Mild | Normal | Not | N |
| 18 | Rain | Mild | Normal | Medium | N |
| 19 | Overcast | Mild | Normal | Medium | P |
| 20 | Overcast | Mild | Normal | Very | P |
| 21 | Sunny | Mild | High | Very | P |
| 22 | Sunny | Mild | High | Medium | P |
| 23 | Sunny | Hot | Normal | Not | P |
| 24 | Rain | Mild | High | Very | N |

解：初始时刻的熵值为： $H(X) = -\frac{12}{24} \log \frac{12}{24} - \frac{12}{24} \log \frac{12}{24} = 1$

选取 Outlook 作为测试属性： $H(X/Outlook) = 0.5528$

选取 Temperature 作为测试属性： $H(X/Temp) = 0.6739$

选取 Humidity 作为测试属性： $H(X/Humid) = 0.9183$

选取 Windy 作为测试属性： $H(X/Windy) = 1$

$H(X/Outlook)$ 最小，选择 Outlook 作为测试属性之后将训练实例集分为三个子集，生成三个叶结点，对每个叶结点依次利用上面过程则生成图 3 的决策树。

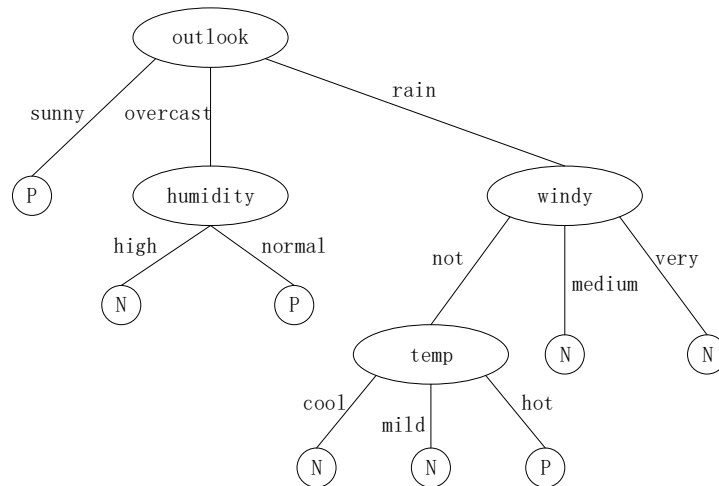
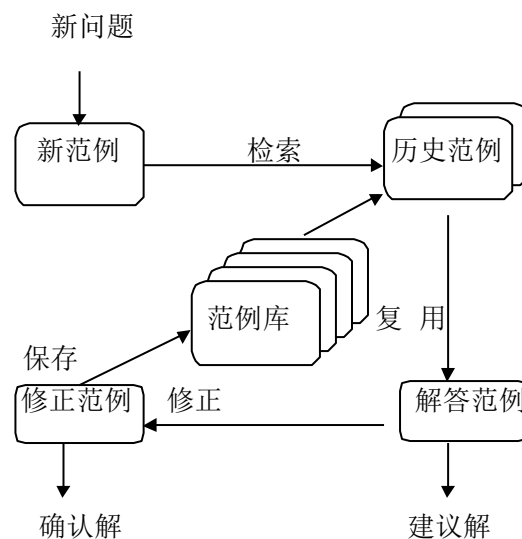


图 3 生成的决策树

6.7 给出基于范例的学习的原理和过程模型。

- (1) 原理略。
- (2) 过程模型：



6.8 试解释强化学习模型及其与其它机器学习方法的异同。

解：略。