

教务处填写：

年 月 日

考 试 用

# 湖南大学课程考试试卷

课程名称： 机器智能 ； 课程编码： CS06152 ；

试卷编号： A ； 考试形式： 闭卷 ； 考试时间： 120 分钟。

题 号	一	二	三	四	五	六	七	八	九	十	总分
应得分	20	25	20	15	20						100
实得分											
评卷人											

(请在答题纸内作答！)

## 一、(简答题) (20 分)

(a). 请简述什么是贝叶斯网络的近似推理？为什么要引入近似推理？（10 分）

【answer】：近似推理是指根据原有贝叶斯网络的拓扑结构以及条件概率表构造大量的数据样本，再在样本的基础上进行概率统计的方法。(5 分)

精确枚举算法虽然计算正确，但当网络结构复杂层次很深的时候计算量巨大，而近似推理的方法大大地降低了计算的复杂度。(5 分)

(b). 比较有监督学习，无监督学习和半监督之间的差异，并简述三种学习方式的优点；（10 分）

【answer】：有监督学习是在带有标签的数据上进行学习；（1 分）

无监督学习是在没有标签的数据上进行学习；（1 分）

半监督学习是同时使用有标签和无标签的数据进行学习；（2 分）

有监督学习的方式性能优良；（2 分） 无监督的学习方式不需要人工标注的数据，因而节约了标注的成本（2 分） 半监督学习的方式只使用少量的标注数据取得接近有监督学习的性能，在节约标注成本的同时保持了较高的性能；（2 分）

## 二、(概率推理) (25 分)

Pacman 随身携带着两种零食（P）在迷宫中奔跑，奔跑中可能遇到怪兽并与之战斗（W）。Pacman 遇到怪兽时会从口袋里掏出一一种零食吃，如果他吃了巧克力(+p)，那么他在战斗中有 70%的概率获得胜

利(+w)；如果他吃的是豆子(-p)，那么他只有 20%的机会获得胜利。已知他的口袋里面有 5 块巧克力和 15 粒豆子，回答下列问题：

(a). Pacman 在战斗中胜利的概率是多少？Pacman 胜利了是因为吃了巧克力的概率是多少？ (6 分)

【answer】:

$$\begin{aligned}
 P(+w) &= P(+w, +p) + P(+w, -p) = P(+w|+p)P(+p) + P(+w|-p)P(-p) \\
 &= \frac{7}{10} \times \frac{1}{4} + \frac{2}{10} \times \frac{3}{4} = \frac{13}{40} = 0.325
 \end{aligned}$$

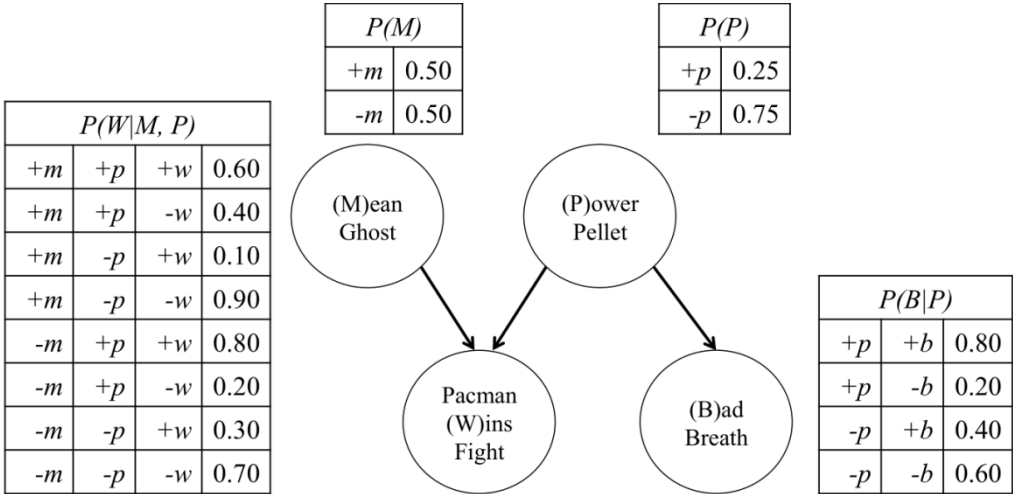
(3 分)

$$\begin{aligned}
 P(+p|+w) &= \frac{P(+w, +p)}{P(+w)} = \frac{P(+w|+p)P(+p)}{P(+w)} \\
 &= \frac{\frac{7}{10} \times \frac{1}{4}}{\frac{13}{40}} = \frac{7}{13} \approx 0.538
 \end{aligned}$$

(3 分)

(b). Pacman 考虑了更多的信息：如 Pacman 吃巧克力会导致有 80%的可能性呼吸困难(+b)，而吃豆子只有 40%的可能性呼吸困难；另外，怪兽分为两种：怪兽一(+m)和怪兽二(-m)，遇见的概率相等。如果吃了巧克力遇见怪兽一胜利的概率有 60%，而吃豆子胜利的概率只有 10%；吃了巧克力遇见怪兽二胜利的概率有 80%，而吃豆子胜利的概率只有 30%，请画出该问题的贝叶斯网络结构图（要求网络中原因在前，结果在后）。(8 分)

【answer】:



(c). 根据 (b) 中的贝叶斯网络指出下列条件独立性是真(T)还是假(F)，其中  $\perp$  表示变量之间独立。

$W \perp B | P$ 、  $M \perp P | W$ 、  $M \perp B | P$ 、  $M \perp B | W$ 、  $W \perp B$ 、  $M \perp P$  (6 分)

【answer】:

.....  $W \perp B | P, (F)$  (1 分)

.....  $M \perp P | W (T)$  (1 分)

.....  $M \perp B | P (T)$  (1 分)

.....  $M \perp B | W (F)$  (1 分)

.....  $W \perp B (T)$  (1 分)

.....  $M \perp P (T)$  (1 分)

(d). 计算给定呼吸困难，遇见怪兽一的条件下 Pacman 胜利的概率 (5 分)

$$P(+w | +b, +m) = a * (0.5 * 0.25 * 0.8 * 0.6 + 0.5 * 0.75 * 0.4 * 0.1) = 0.075a$$

$$P(-w | +b, +m) = a * (0.5 * 0.25 * 0.8 * 0.4 + 0.5 * 0.75 * 0.4 * 0.9) = 0.175a$$

$$\text{归一化 } P(+w | +b, +m) = 0.3$$

### 三、(决策树) (20 分)

假设你正在使用决策树算法学习二进制分类变量 C，您必须决定在当前节点选择哪个属性。已知在这个节点上有 100 个例子；25 个是正的，75 个是负的。如果属性 A 被选中，它的第一个分支得到 5 个正的和 50 个负的例子，它的第二个分支将得到 20 个正的和 25 个负的例子。

(a). 计算熵值  $H(C)$  和选择属性 A 的条件熵值  $H(C|A)$  (8 分)。

$$\text{【answer】: } H(C) = -1/4 \log 1/4 - 3/4 \log 3/4 = 0.81; (4 \text{ 分})$$

$$H(C|A) = 0.55 \cdot (-1/11 \log(1/11) - 10/11 \log(10/11)) + 0.45 \cdot (-4/9 \log(4/9) - 5/9 \log(5/9)) \\ = 0.2417 + 0.4460 = 0.6877 \quad (4 \text{ 分})$$

(b). 判断下列语句的对错并说明原因：当所有的候选属性的信息增益都为 0 的时候当前这个节点应该选为叶子节点(4 分)。

【answer】: 错误 (2 分) 所有属性的信息增益为 0 说明所有属性对于区分样本没有任何帮助，并不表示样本已经区分开来，因而不能作为叶子节点 (2 分)

(c). 在构造完一棵决策树后我们发现该决策树在训练数据上的性能很高但在测试数据上的性能下降，请解释原因并给出解决方案 (8 分)。

【answer】: 产生了过拟合现象 (3 分) 过拟合现象说明当前训练的模型过分的拟合了训练数据，学到了一些无关的信息导致在测试数据上的性能下降 (2 分)；

解决方案：对决策树进行剪枝，剪掉一些层次很深，信息量不大的属性 (3 分)；

### 四、(支持向量机) (15 分)

考虑一个小数据集有四个数据  $(x_1 = 0, x_2 = 0), y_1 = 0$ ;  $(x_1 = 0, x_2 = 1), y_2 = 1$ ;  $(x_1 = 1, x_2 = 0), y_3 = 1$ ;  $(x_1 = 1, x_2 = 1), y_4 = 0$ 。回答下列问题：

(a). 使用线性核函数的支持向量机能分类该数据集吗？ (3 分)

【answer】: 不能, 该数据集不是线性可分的数据集 (3 分)

(b). 如果将数据  $X=(x_1, x_2)$  映射到新的空间  $\phi(X) = (2 \cdot x_1, 2 \cdot x_2, -x_1 - x_2)$  上, 计算上述四个数据映射后的结果是什么? 在新空间中该数据集是线性可分的吗? (6 分)

【answer】:  $\phi(X_1) = (0, 0, 0)$   $y_1=0$

$\phi(X_2) = (0, 2, -1)$   $y_1=1$

$\phi(X_3) = (2, 0, -1)$   $y_1=1$

$\phi(X_4) = (2, 2, -2)$   $y_1=0$  (4 分)

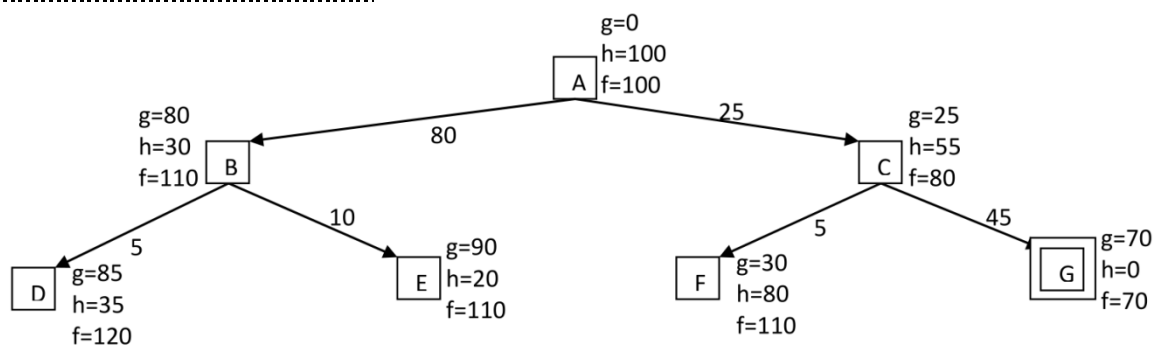
在新空间中该数据集是线性可分的。(2 分)

(c). 核函数的定义是什么? 核函数的主要作用是什么? (6 分)

【answer】: 核函数是将原数据映射到新的空间后在新空间中两个数据的内积; (3 分)

核函数的主要作用是进行数据升维, 从而使得原来线性不可分的数据在新空间中线性可分(3 分)

## 五、(搜索问题) (20 分)



图一: 搜索树

如图一所示的搜索树, 假定 A 是起始结点, G 是目标节点。节点的代价标注在节点附近, 其中 g 表示从起始结点到当前节点的实际代价; h 表示从当前节点到目标节点的估计代价; f 为两者之和; 边上数字表示两个节点之间的实际代价。给出下列搜索算法的结点访问顺序 (到目标节点算法结束)。

(a). 一致代价搜索 (5 分);

【answer】: A,C,F,G

(b). 迭代加深的深度优先搜索 (5 分);

【answer】: A;

A,B,C;

A,B,C,D,E,F,G;

(c). 贪婪最佳优先搜索 (5 分);

【answer】: A,B,E,D,C,G

(d). A\*树搜索 (5 分);

【answer】: ACG