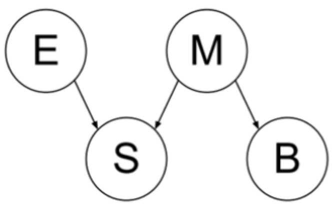
贝叶斯网络、机器学习(1) 课外作业

1. 假设，空气中弥漫着硫磺(S)的气味既可能是鸡蛋(E)发臭所散发的，也可能是玛雅启示(M)所引发的厄运征兆，且玛雅启示还会导致海水沸腾(B)。对应的贝叶斯网络和部分条件概率分布如图所示。



E	P(E)
T	0.4
F	0.6

M	P(M)
T	0.1
F	0.9

M	B	P(B|M)
T	T	1.0
T	F	0.0
F	T	0.1
F	F	0.9

E	M	S	P(S|E,M)
T	T	T	1.0
T	T	F	0.0
T	F	T	0.8
T	F	F	0.2
F	T	T	0.3
F	T	F	0.7
F	F	T	0.1
F	F	F	0.9

(1) 计算联合概率分布 $P(E=F, S=F, M=F, B=F)$;
(2) 海水沸腾的概率为多少？
(3) 假设海洋正在沸腾，那么玛雅启示发生的概率是多少？
(4) 假设空气中有硫磺的气味、海水正在沸腾、鸡蛋已经发臭，那么玛雅启示发生的概率是多少？
(5) 假设玛雅启示正在发生，那么鸡蛋发臭的概率是多少？

解答：

(1) $P(E=F, S=F, M=F, B=F)=P(E=F)P(M=F)P(S=F|E=F, M=F)P(B=F|M=F)= 0.6\times0.9\times0.9\times0.9 = 0.44$.

(2) $P(B=T) = P(B=T|M=T)P(M=T) + P(B=T|M=F)P(M=F) = 1\times0.1+0.1\times0.9 = 0.19$.

(3) $P(M=T|B=T) = P(B=T|M=T)P(M=T) / P(B=T) = (1\times0.1) / 0.19 = 0.53$.

(4) 所求即 $P(M=T|S=T, B=T, E=T)$.

首先求 $P(E=T, S=T, M=T, B=T) = P(E=T)P(M=T)P(S=T|E=T, M=T)P(B=T|M=T) = 0.4\times0.1\times1.0\times1.0 = 0.04$,

再求 $P(S=T, B=T, E=T) = P(E=T)P(M=T)P(S=T|E=T, M=T)P(B=T|M=T) + P(E=T)P(M=F)P(S=T|E=T, M=F)P(B=T|M=F) = 0.4\times0.1\times1.0\times1.0 + 0.4\times0.9\times0.8\times0.1 = 0.0688$,

因此：$P(M=T|S=T, B=T, E=T) = P(E=T, S=T, M=T, B=T) / P(S=T, B=T, E=T) = 0.04/0.0688 = 0.58$.

(5) $P(E=T|M=T)=P(E=T)=0.4$.

2. 根据以下数据集，使用 ID3 算法计算属性“天气”和“湿度”的信息增益，并说明哪个属性更适合作为根节点。

天气	湿度	活动（是否进行）
晴	高	否
晴	高	否
多云	高	是
雨	中	是
雨	中	是

解答：

解答过程

Step 1：计算总熵

总样本数：5

- 正类（活动=是）：3
- 负类（活动=否）：2

总熵公式：

$$Entropy(D) = -\frac{3}{5}\log_2\frac{3}{5} - \frac{2}{5}\log_2\frac{2}{5}$$

计算：

$$Entropy(D) = -0.6 \times (-0.737) - 0.4 \times (-1.322) = 0.971\,\text{bit}$$

（注：$\log_2\frac{3}{5} \approx -0.737$，$\log_2\frac{2}{5} \approx -1.322$）

Step 2：计算“天气”属性的信息增益

天气的取值：晴（2个）、多云（1个）、雨（2个）

- **晴（2个样本）**：活动均为“否”

$$Entropy(\text{晴}) = -0 \times \log_2 0 - 1 \times \log_2 1 = 0\,\text{bit}$$

- **多云（1个样本）**：活动为“是”

$$Entropy(\text{多云}) = -1 \times \log_2 1 = 0\,\text{bit}$$

- **雨（2个样本）**：活动均为“是”

$$Entropy(\text{雨}) = -1 \times \log_2 1 = 0\,\text{bit}$$

条件熵：

$$Entropy(\text{天气}) = \frac{2}{5} \times 0 + \frac{1}{5} \times 0 + \frac{2}{5} \times 0 = 0\,\text{bit}$$

信息增益：

$$Gain(\text{天气}) = Entropy(D) - Entropy(\text{天气}) = 0.971 - 0 = 0.971\,\text{bit}$$

Step 3：计算“湿度”属性的信息增益

湿度的取值：高（3个）、中（2个）

- **高（3个样本）**：活动为“否”（2个）、“是”（1个）

$$Entropy(\text{高}) = -\frac{2}{3}\log_2\frac{2}{3} - \frac{1}{3}\log_2\frac{1}{3} \approx 0.918\ \text{bit}$$

- **中（2个样本）**：活动均为“是”

$$Entropy(\text{中}) = -1 \times \log_2 1 = 0\ \text{bit}$$

条件熵：

$$Entropy(\text{湿度}) = \frac{3}{5} \times 0.918 + \frac{2}{5} \times 0 \approx 0.551\ \text{bit}$$

信息增益：

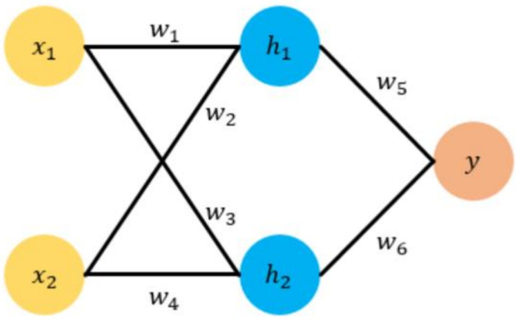
$$Gain(\text{湿度}) = Entropy(D) - Entropy(\text{湿度}) = 0.971 - 0.551 = 0.420\ \text{bit}$$

Step 4：选择根节点

- 天气的信息增益：0.971 bit
- 湿度的信息增益：0.420 bit

结论：选择“天气”作为根节点（信息增益更大）。

3. 如图是一个 MLP 模型。现有一个仅包含一个数据的数据集，该数据输入$x_1 = 1$，$x_2 = 0.5$，输出的目标值$t = 4$。如果随机初始化后，$w_1 = 0.5$，$w_2 = 1.5$，$w_3 = 2.3$，$w_4 = 3$，$w_5 = 1$，$w_6 = 1$，若学习率$\eta = 0.1$，激活函数均为 ReLU，求经过1轮反向传播后，权重的更新值w_5^+和w_1^+.



解答：

(1)首先前向传播：

$$h_{1,in} = w_1 \cdot x_1 + w_2 \cdot x_2 = 1.25$$
$$h_1 = ReLU\left(h_{1,in}\right) = 1.25$$
$$h_{2,in} = w_3 \cdot x_1 + w_4 \cdot x_2 = 3.8$$

$$h_2 = ReLU(h_{2,in}) = 3.8$$
$$y_{in} = w_5 \cdot h_1 + w_6 \cdot h_2 = 5.05$$
$$y = ReLU(y_{in})$$
$$E = \frac{1}{2}(y - t)^2 = 0.55125$$

(2)反向传播：

根据链式法则有：$\frac{\partial E}{\partial w_5} = \frac{\partial E}{\partial y} \cdot \frac{\partial y}{\partial y_{in}} \cdot \frac{\partial y_{in}}{\partial w_5}$。

其中：$\frac{\partial E}{\partial y} = 2 \cdot \frac{1}{2} \cdot (t - y) \cdot (-1) = y - t = 5.05 - 4 = 1.05$；

由于$y_{in} > 0$，故$\frac{\partial y}{\partial y_{in}} = 1$；

而$y = w_5 \cdot h_1 + w_6 \cdot h_2$，所以$\frac{\partial y}{\partial w_5} = h_1 + 0 = h_1 = 1.25$；

因此$\frac{\partial E}{\partial w_5} = \frac{\partial E}{\partial w_5} = \frac{\partial E}{\partial y} \cdot \frac{\partial y}{\partial y_{in}} \cdot \frac{\partial y_{in}}{\partial w_5} = (y - t) \cdot 1 \cdot h_1 = 1.05 \times 1 \times 1.25 = 1.3125$;

更新参数：$w_5^+ = w_5 - \eta \cdot \frac{\partial E}{\partial w_5} = 1 - 0.1 \times 1.3125 = 0.86875.$

根据链式法则：$\frac{\partial E}{\partial w_1} = \frac{\partial E}{\partial y} \cdot \frac{\partial y}{\partial y_{in}} \cdot \frac{\partial y_{in}}{\partial h_1} \cdot \frac{\partial h_1}{\partial h_{1,in}} \cdot \frac{\partial h_{1,in}}{\partial w_1}$。

其中：$\frac{\partial y_{in}}{\partial h_1} = w_5 + 0 = w_5$；

因为$h_{1,in} > 0$，因此$\frac{\partial h_1}{\partial h_{1,in}} = 1$；

由$h_{1,in} = w_1 \cdot x_1 + w_2 \cdot x_2$，故$\frac{\partial h_{1,in}}{\partial w_1} = x_1 + 0 = x_1$。

因此，$\frac{\partial E}{\partial w_1} = \frac{\partial E}{\partial y} \cdot \frac{\partial y}{\partial y_{in}} \cdot \frac{\partial y_{in}}{\partial h_1} \cdot \frac{\partial h_1}{\partial h_{1,in}} \cdot \frac{\partial h_{1,in}}{\partial w_1} = (y - t) \cdot 1 \cdot w_5 \cdot 1 \cdot x_1$

故$w_1^+ = w_1 - \frac{\eta \partial E}{\partial w_1} = w_1 - \eta \cdot (y - t) \cdot w_5 \cdot x_1 = 0.5 - 0.1 \times 1.05 \times 1 \times 1 = 0.395$。

4.给定输入和卷积核如下：

输入：2 通道，每通道为 1 个 3*3 矩阵：$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}$和$\begin{bmatrix} 3 & 2 & 1 \\ 2 & 1 & 3 \\ 1 & 3 & 2 \end{bmatrix}$；

卷积核 1：$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$和$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$，偏置值为 1；

卷积核 2：$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$和$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$，偏置值为 2。

若卷积计算步长为 1，填充为 1：

(1)计算经过卷积计算后的结果，并用 CNN 特征图维度计算公式验证输出的宽度和高度；

(2)分别使用 2*2 平均池化和 2*2 最大池化处理(1)中卷积计算结果。

解答：

(1)卷积结果第一个通道：$\begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 4 & 3 & 3 \\ 3 & 3 & 5 & 1 \\ 0 & 3 & 1 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 3 & 2 & 1 \\ 3 & 4 & 2 & 3 \\ 2 & 2 & 6 & 2 \\ 1 & 3 & 2 & 0 \end{bmatrix} + 1 = \begin{bmatrix} 2 & 6 & 6 & 2 \\ 6 & 9 & 6 & 7 \\ 6 & 6 & 12 & 4 \\ 2 & 7 & 4 & 3 \end{bmatrix}$

卷积结果第二个通道：$\begin{bmatrix} -1 & -2 & -3 & 0 \\ -2 & -2 & 1 & 3 \\ -3 & 1 & 1 & 1 \\ 0 & 3 & 1 & 2 \end{bmatrix} + \begin{bmatrix} 0 & -3 & -2 & -1 \\ 3 & 0 & 0 & -3 \\ 2 & 0 & 0 & -2 \\ 1 & 3 & 2 & 0 \end{bmatrix} + 2 = \begin{bmatrix} 1 & -3 & -3 & 1 \\ 3 & 0 & 3 & 2 \\ 1 & 3 & 3 & 1 \\ 3 & 8 & 5 & 4 \end{bmatrix}$

(2)平均池化后：$\begin{bmatrix} 5.75 & 5.25 \\ 5.25 & 5.75 \end{bmatrix}$和$\begin{bmatrix} 0.25 & 0.75 \\ 3.75 & 3.25 \end{bmatrix}$

最大池化后：$\begin{bmatrix} 9 & 7 \\ 7 & 12 \end{bmatrix}$和$\begin{bmatrix} 3 & 3 \\ 8 & 5 \end{bmatrix}$