

《模式识别与机器学习》

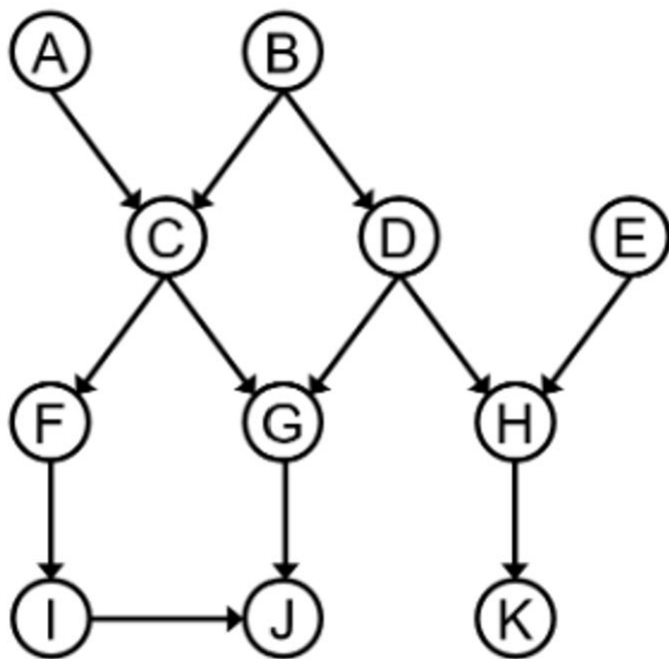
第2次习题课

助教：许修为、黄原辉

2023年12月05日

一、选择题

1. (多选题) 对于下图中的贝叶斯网络, 下列独立性关系成立的有 (AB) 2 分



A. $E \perp F$ B. $A \perp D$ C. $A \perp K | C$ D. $D \perp I | J$ E. $C \perp E | K$ F. $A \perp E | J, K$

一般步骤:

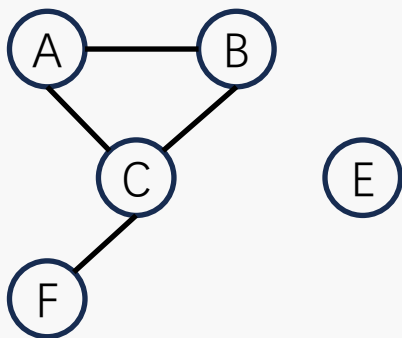
1. 根据原始概率图, 构建包括表达式中包含的变量以及这些变量的ancestor节点 (父节点、父节点的父节点...) 的图
2. 道德化该子图
3. 从图中删除需要判断的概率表达式中作为条件的变量, 以及和他们相连的路径
4. 根据待观测变量在图上的连通性判断独立性

参考: [d-separation.pdf \(mit.edu\)](https://www.meritt.berkeley.edu/doc/publications/d-separation.pdf)

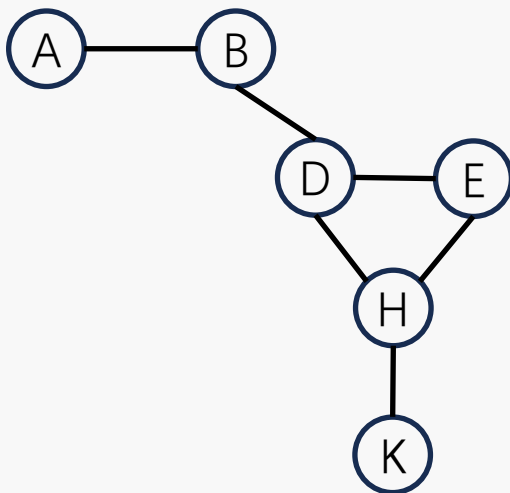
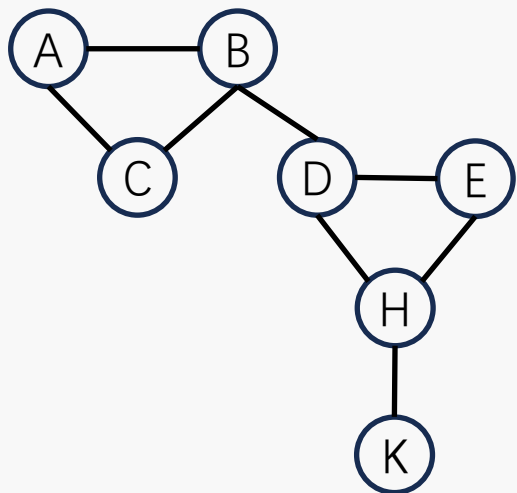
一、选择题

以 A 选项和 C 选项为例：

A. $E \perp F$



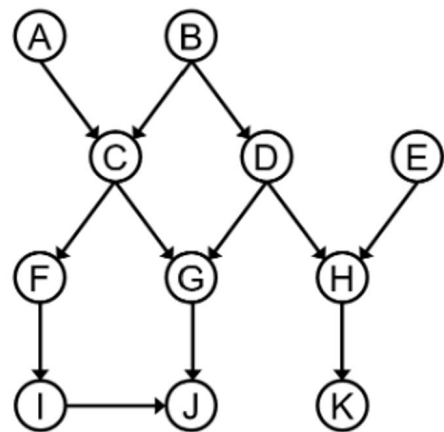
C. $A \perp K | C$



一般步骤：

1. 根据原始概率图，构建包括表达式中包含的变量以及这些变量的ancestor节点（父节点、父节点的父节点...）的图
2. 道德化该子图
3. 从图中删除需要判断的概率表达式中作为条件的变量，以及和他们相连的路径
4. 根据待观测变量在图上的连通性判断独立性

1. （多选题）对于下图中的贝叶斯网络，下列独立性关系成立的有（ AB ） 2分



A. $E \perp F$ B. $A \perp D$ C. $A \perp K | C$ D. $D \perp I | J$ E. $C \perp E | K$ F. $A \perp E | J, K$

二、计算题

1. （贝叶斯网络）写出如下贝叶斯网络所有变量的联合概率分布：4分

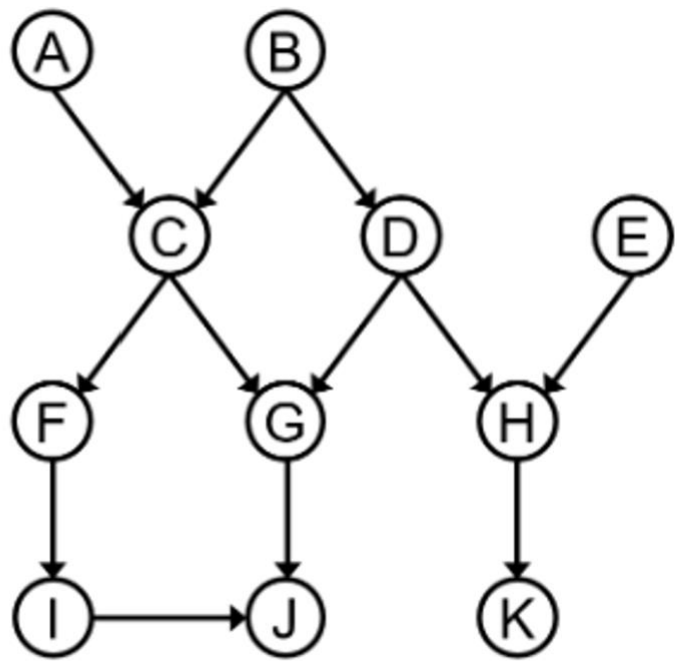


图 1 贝叶斯网各个变量的关联关系图

$$P(ABCDEFGHIJK) = P(A)P(B)P(C|AB)P(D|B)P(E)P(F|C)P(G|CD)P(H|DE)P(I|F)P(J|G)P(K|H)$$

二、计算题

2. (贝叶斯网的学习) 已知西瓜数据集如图 2 所示, 计算图 3 中贝叶斯网的 BIC 评分。

编号	色泽	根蒂	敲声	纹理	脐部	触感	好瓜
1	青绿	蜷缩	浊响	清晰	凹陷	硬滑	是
2	乌黑	蜷缩	沉闷	清晰	凹陷	硬滑	是
3	乌黑	蜷缩	浊响	清晰	凹陷	硬滑	是
4	青绿	蜷缩	沉闷	清晰	凹陷	硬滑	是
5	浅白	蜷缩	浊响	清晰	凹陷	硬滑	是
6	青绿	稍蜷	浊响	清晰	稍凹	软粘	是
7	乌黑	稍蜷	浊响	稍糊	稍凹	软粘	是
8	乌黑	稍蜷	浊响	清晰	稍凹	硬滑	是
9	乌黑	稍蜷	沉闷	稍糊	稍凹	硬滑	否
10	青绿	硬挺	清脆	清晰	平坦	软粘	否
11	浅白	硬挺	清脆	模糊	平坦	硬滑	否
12	浅白	蜷缩	浊响	模糊	平坦	软粘	否
13	青绿	稍蜷	浊响	稍糊	凹陷	硬滑	否
14	浅白	稍蜷	沉闷	稍糊	凹陷	硬滑	否
15	乌黑	稍蜷	浊响	清晰	稍凹	软粘	否
16	浅白	蜷缩	浊响	模糊	平坦	硬滑	否
17	青绿	蜷缩	沉闷	稍糊	稍凹	硬滑	否

图 2 西瓜数据集

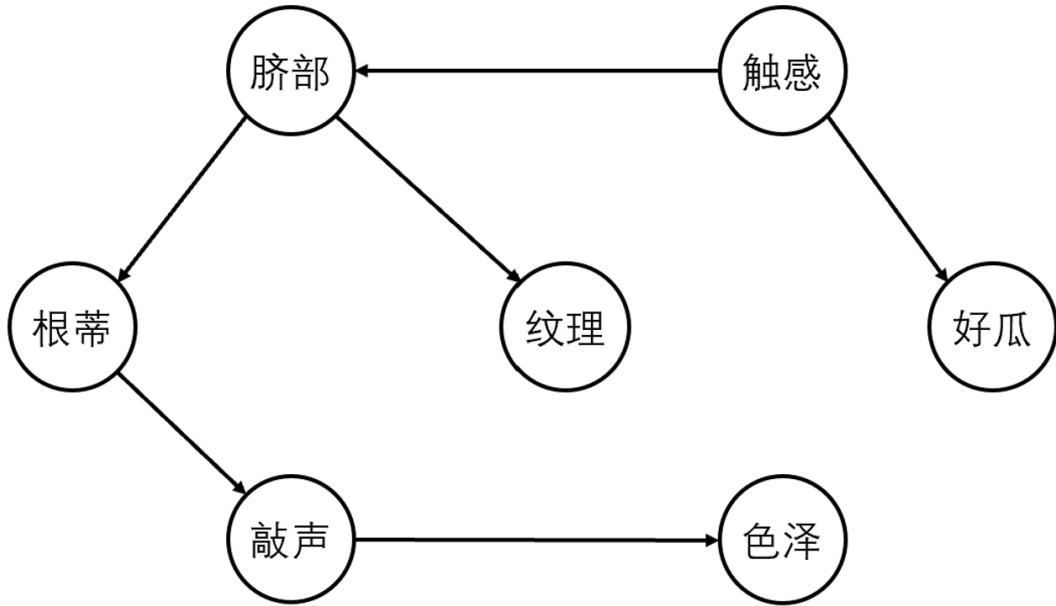


图 3 贝叶斯网

$$BIC = \frac{\log m}{2} |B| - LL(B|D)$$

(1) $m = 17$

二、计算题

$$BIC = \frac{\log m}{2} |B| - LL(B|D)$$

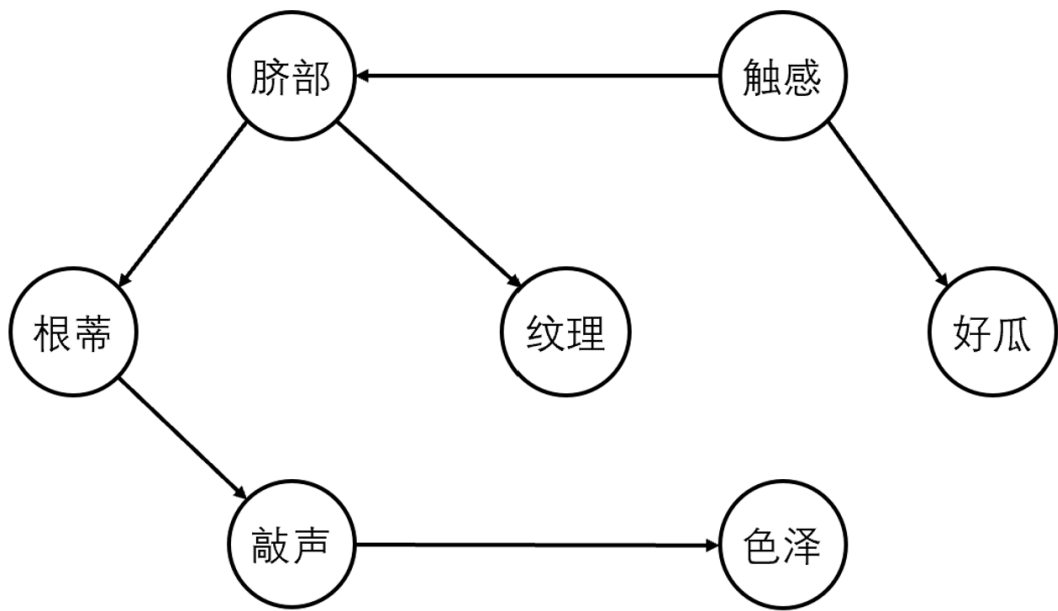


图 3 贝叶斯网

(2) 计算 $|B| = 4 + 1 + 6 \times 4 + 2 = 31$

脐部：3 种取值， $2 \times (3 - 1) = 4$

触感：2 种取值，1

根蒂：3 种取值， $3 \times (3 - 1) = 6$

纹理：3 种取值， $3 \times (3 - 1) = 6$

敲声：3 种取值， $3 \times (3 - 1) = 6$

色泽：3 种取值， $3 \times (3 - 1) = 6$

好瓜：2 种取值， $2 \times (2 - 1) = 2$

二、计算题

$$BIC = \frac{\log m}{2} |B| - LL(B|D)$$

编号	色泽	根蒂	敲声	纹理	脐部	触感	好瓜
1	青绿	蜷缩	浊响	清晰	凹陷	硬滑	是
2	乌黑	蜷缩	沉闷	清晰	凹陷	硬滑	是
3	乌黑	蜷缩	浊响	清晰	凹陷	硬滑	是
4	青绿	蜷缩	沉闷	清晰	凹陷	硬滑	是
5	浅白	蜷缩	浊响	清晰	凹陷	硬滑	是
6	青绿	稍蜷	浊响	清晰	稍凹	软粘	是
7	乌黑	稍蜷	浊响	稍糊	稍凹	软粘	是
8	乌黑	稍蜷	浊响	清晰	稍凹	硬滑	是
9	乌黑	稍蜷	沉闷	稍糊	稍凹	硬滑	否
10	青绿	硬挺	清脆	清晰	平坦	软粘	否
11	浅白	硬挺	清脆	模糊	平坦	硬滑	否
12	浅白	蜷缩	浊响	模糊	平坦	软粘	否
13	青绿	稍蜷	浊响	稍糊	凹陷	硬滑	否
14	浅白	稍蜷	沉闷	稍糊	凹陷	硬滑	否
15	乌黑	稍蜷	浊响	清晰	稍凹	软粘	否
16	浅白	蜷缩	浊响	模糊	平坦	硬滑	否
17	青绿	蜷缩	沉闷	稍糊	稍凹	硬滑	否

图 2 西瓜数据集

(3) 使用极大似然法估计网络参数

P(脐部 触感)	硬滑	软粘
凹陷	7/12	0
稍凹	3/12	3/5
平坦	2/12	2/5

P(触感)	
硬滑	12/17
软粘	5/17

P(根蒂 脐部)	凹陷	稍凹	平坦
蜷缩	5/7	1/6	2/4
稍蜷	2/7	5/6	0
硬挺	0	0	2/4

□ 二、计算题

$$BIC = \frac{\log m}{2} |B| - LL(B|D)$$

(4) 计算各样本的似然概率 \leftarrow

$$LL(\text{脐部}|\text{触感}) = 7 \log \frac{7}{12} + 3 \log \frac{3}{12} + 2 \log \frac{2}{12} + 3 \log \frac{3}{5} + 2 \log \frac{2}{5} = -14.88 \leftarrow$$

$$LL(\text{触感}) = 12 \log \frac{12}{17} + 5 \log \frac{5}{17} = -10.30 \leftarrow$$

$$LL(\text{根蒂}|\text{脐部}) = 5 \log \frac{5}{7} + 2 \log \frac{2}{7} + \log \frac{1}{6} + 5 \log \frac{5}{6} + 2 \log \frac{2}{4} + 2 \log \frac{2}{4} = -9.66 \leftarrow$$

$$LL(\text{纹理}|\text{脐部}) = 5 \log \frac{5}{7} + 2 \log \frac{2}{7} + 3 \log \frac{3}{6} + 3 \log \frac{3}{6} + \log \frac{1}{4} + 3 \log \frac{3}{4} = -10.60 \leftarrow$$

$$LL(\text{好瓜}|\text{触感}) = 6 \log \frac{6}{12} + 6 \log \frac{6}{12} + 2 \log \frac{2}{5} + 3 \log \frac{3}{5} = -11.68 \leftarrow$$

$$LL(\text{敲声}|\text{根蒂}) = 5 \log \frac{5}{8} + 3 \log \frac{3}{8} + 5 \log \frac{5}{7} + 2 \log \frac{2}{7} + 2 \log \frac{2}{2} = -9.48 \leftarrow$$

$$LL(\text{色泽}|\text{敲声}) = 3 \log \frac{3}{10} + 4 \log \frac{4}{10} + 3 \log \frac{3}{10} + 2 \log \frac{2}{5} + 2 \log \frac{2}{5} + \log \frac{1}{5} + \log \frac{1}{2} + \log \frac{1}{2} = -17.55 \leftarrow$$

$$LL(B|D) = -84.15 \leftarrow$$

(5) $BIC = \frac{\log m}{2} |B| - LL(B|D) = \mathbf{128.06}$ 或者 $\mathbf{184.76}$ (以 2 为底) \leftarrow

□ 选择题

- B: 弱分类器的错误率小于0.5是集成学习的前提, 否则该分类器参与集成将会产生负面影响
- C: Adaboost是串行算法

1. (多选题) 下列说法错误的是 (BC) ←

- A. 集成学习需要各个弱分类器之间具备一定的差异性 ←
- B. 弱分类器的错误率可以大于 0.5 ←
- C. Adaboost 算法使用并行式策略训练基学习器 ←
- D. Adaboost 算法结合弱学习器以避免欠拟合 ←

□ 选择题

- A: boosting是组合弱学习器避免欠拟合, bagging是组合强学习器避免过拟合
- D: Boosting是串行算法

2. (多选题) 下列说法正确的是 (BC) ↵

- A. Bagging 和 Boosting 都是对弱学习器的组合↵
- B. Bagging 算法能够降低模型预测结果的方差↵
- C. 随机森林能进一步减小 Bagging 算法中基学习器之间的相关性↵
- D. Boosting 能通过并行计算求解↵



第八章作业

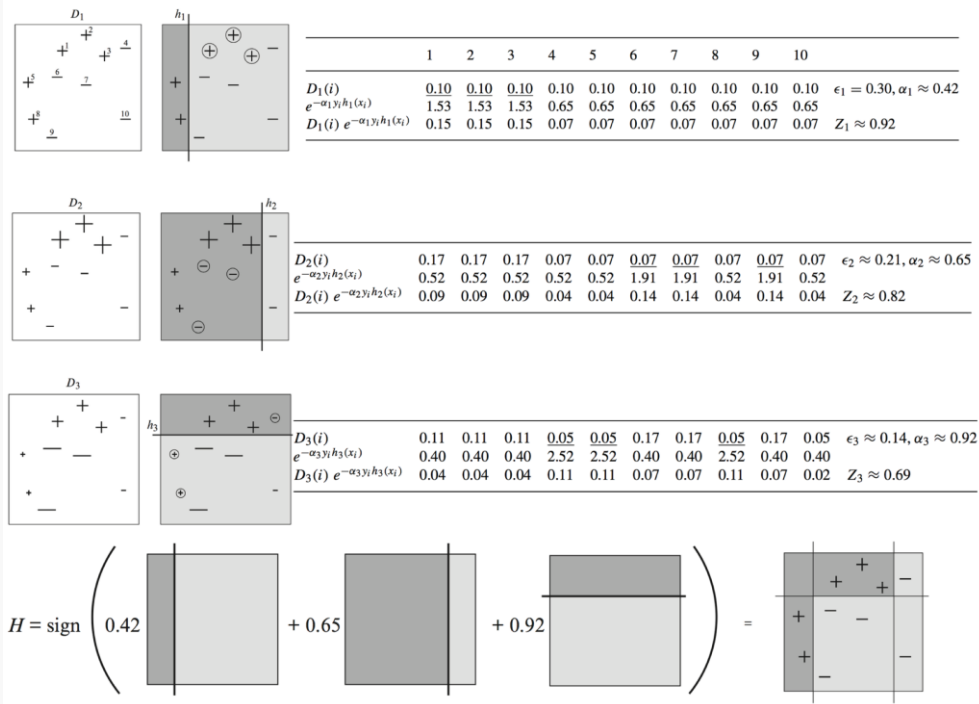
□ 计算题

■ **注意：** 答案并不唯一，和选择弱分类器的顺序有关

(AdaBoost 算法) 训练数据如下表所示，假设弱分类器由竖直或水平分类面 ($x_1 < v_1$ 或 $x_1 > v_1$ 或 $x_2 < v_2$ 或 $x_2 > v_2$) 产生，其阈值 v_1 或 v_2 使该分类器在训练数据集上分类错误率最低。试用 AdaBoost 算法学习一个强分类器。

序号	1	2	3	4	5	6	7	8	9	10
(x_1, x_2)	(2,2)	(3,1)	(4,2)	(5,2)	(1,3)	(2,3)	(3,3)	(1,4)	(2,5)	(5,4)
y	1	1	1	-1	1	-1	-1	1	-1	-1

$$H = \text{sign}(0.42(x_1 < 1.5) + 0.65(x_1 < 4.5) + 0.92(x_2 < 2.5))$$



□ 一、选择题

1. (单选题) 下列哪个神经网络结构会发生权重共享 (D)
A. 卷积神经网络
B. 循环神经网络
C. 全连接神经网络
D. 卷积神经网络和循环神经网络
2. (单选题) 与传统机器学习方法相比, 深度学习的优势在于 (A)
A. 深度学习可以自动学习特征
B. 深度学习不需要做数据预处理
C. 深度学习完全不提取底层特征, 如图像边缘、纹理等
D. 深度学习不需要调参
3. (单选题) LSTM 总共有三个门机制, 不包括下面哪项? (B)
A. 输入门
B. 更新门
C. 遗忘门
D. 输出门

二、计算题

1. (卷积神经网络) 假设如下左图是二维卷积神经网络某层某通道的特征图，如下右图为下一层的一个 3×3 卷积核：↵

(a) 特征图						(b) 卷积核		
3	2	5	0	5	3	0	1	0
9	4	0	1	2	6	-1	0	1
3	0	2	6	7	7	0	-1	0
1	1	3	7	1	0			
7	9	8	6	6	9			
3	9	0	4	1	7			

- (1) 卷积操作是卷积神经网络的必要步骤。请写出上述卷积核滤波后的特征图，其中边界延拓 (padding) 参数为 0，卷积步长 (stride) 参数为 1。↵
- (2) 请写出卷积之后的特征图再经过一个最大值池化 (max-pooling) 层之后的特征图，其中 kernel 大小为 2，stride 参数为 2。↵
- (3) 请写出 (1) 中卷积核滤波后的特征图以 ReLU 函数为激活函数的输出特征图。↵



第九章作业

二、计算题

(a) 特征图

3	2	5	0	5	3
9	4	0	1	2	6
3	0	2	6	7	7
1	1	3	7	1	0
7	9	8	6	6	9
3	9	0	4	1	7

(b) 卷积核

0	1	0
-1	0	1
0	-1	0

(1) 卷积运算

-7	0	-4	3
2	3	-1	2
-7	0	-2	-6
-7	0	1	3

(2) 最大值池化

3	3
0	3

(2) ReLU激活

3	0	0	3
2	3	0	2
0	0	0	0
0	0	1	3