



贝叶斯网络 (Bayesian Networks)

作者: Calvin

QQ: 179209347

Mail: 179209347@qq.com

介绍

笔记简介:

- 面向对象: 深度学习初学者
- 依赖课程: **线性代数, 统计概率**, 优化理论, 图论, 离散数学, 微积分, 信息论

知乎专栏:

<https://zhuanlan.zhihu.com/p/693738275>

Github & Gitee 地址:

https://github.com/mymagicpower/AIAS/tree/main/deep_learning

https://gitee.com/mymagicpower/AIAS/tree/main/deep_learning

* 版权声明:

- 仅限用于个人学习
- 禁止用于任何商业用途

贝叶斯网络 (Bayesian Networks)

贝叶斯网络通过**有向图**和**条件概率**分布来表示一组随机变量之间的依赖关系和联合概率分布。

假设我们有一组随机变量 X_1, X_2, \dots, X_n , 它们之间的依赖关系可以用有向无环图 (DAG) 来表示。设节点 X_i 的父节点集合为 $Pa(X_i)$ 。那么贝叶斯网络的联合概率分布可以表示为:

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P(X_i | Pa(X_i))$$

其中, $P(X_i | Pa(X_i))$ 表示节点 X_i 在给定其父节点 $Pa(X_i)$ 的条件下的条件概率分布。

这个公式表明:

整个网络的联合概率分布可以通过每个节点的条件概率分布的乘积来表示。

每个节点的条件概率分布仅依赖于其父节点的取值, 符合贝叶斯网络的条件独立性假设。

如果我们用 x_1, x_2, \dots, x_n 表示对应随机变量的取值, 那么条件概率 $P(X_i = x_i | Pa(X_i))$ 可以通过贝叶斯网络中的参数化形式来表示。通常情况下, 这些参数是由数据估计得到的。

独立性

变量独立性:

随机变量 A 与 B 相互独立, 则:

$$P(AB) = P(A)P(B)$$

$$P(A|B) = P(A)$$

$$P(B|A) = P(B)$$

条件独立性:

随机变量 A 与 B 在给定 C 的条件下相互独立, 则:

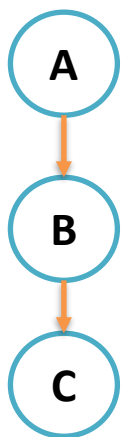
$$P(A, B|C) = P(A|C)P(B|C)$$

$$P(A|B, C) = P(A|C)$$

$$P(B|A, C) = P(B|C)$$

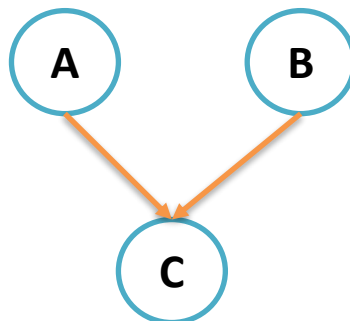
条件独立性

在贝叶斯网络中，如果两个节点是直接连接的，它们是非条件独立的，是直接因果关系。父节点是“因”，子节点是“果”。



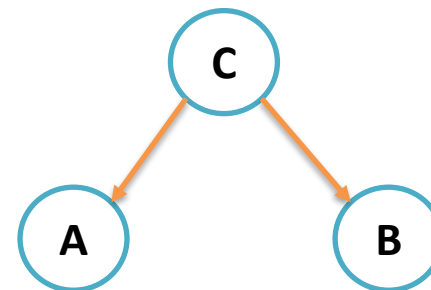
间接因果关系：

在已知 B 时，A 和 C 为条件独立；



共因关系：

在已知 C 时，A 和 B 条件独立；

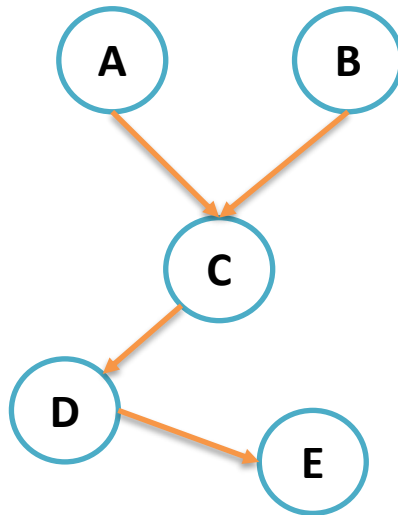


共果关系：

A 和 B 是独立的，在已知 C 时，A 和 B 不独立。

贝叶斯网络中的独立性

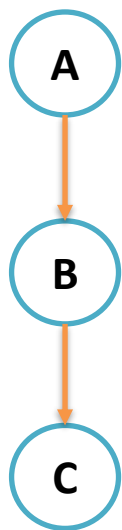
定理：父节点已知时，该节点与其所有非后代的节点（non-descendants）条件独立。



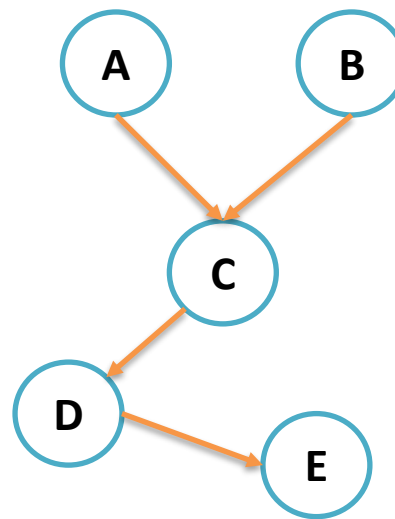
贝叶斯网络例子

联合概率分布:

联合概率可以表示为局部条件概率表的乘积.



$$P(A, B, C) = P(C|B) P(B|A) P(A)$$

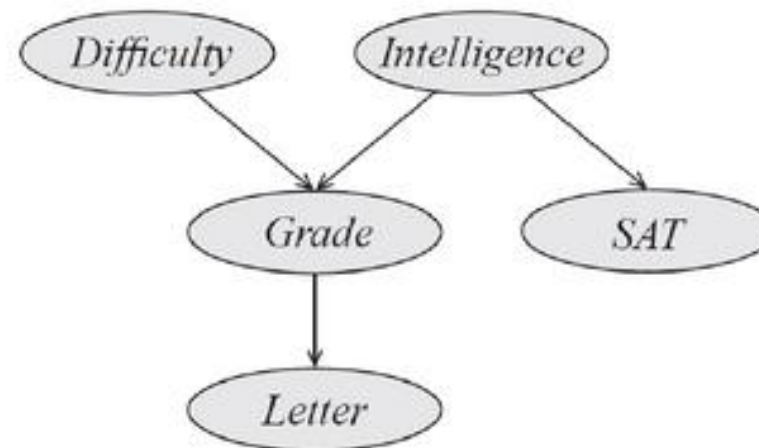


$$P(A, B, C, D, E) = P(E | D)P(D | C)P(C | A, B)P(B)P(A)$$

贝叶斯网络例子 - “学生模型” 的贝叶斯网络图

“学生模型” 的贝叶斯网络图:

学生的成绩不仅取决于他的智力，而且取决于他的学习成绩。
课程的有效性，由随机变量D表示，其域为 {easy, hard}。
我们的学生向他的教授要推荐信。这位教授总是心不在焉，从来不记得学生的名字。她只能看他的成绩，她只根据这些信息给他写信。
她的信的质量是一个随机变量L，其定义域是 {强, 弱}。
信件的实际质量取决于分数。



因此，我们在这一领域有五个随机变量：

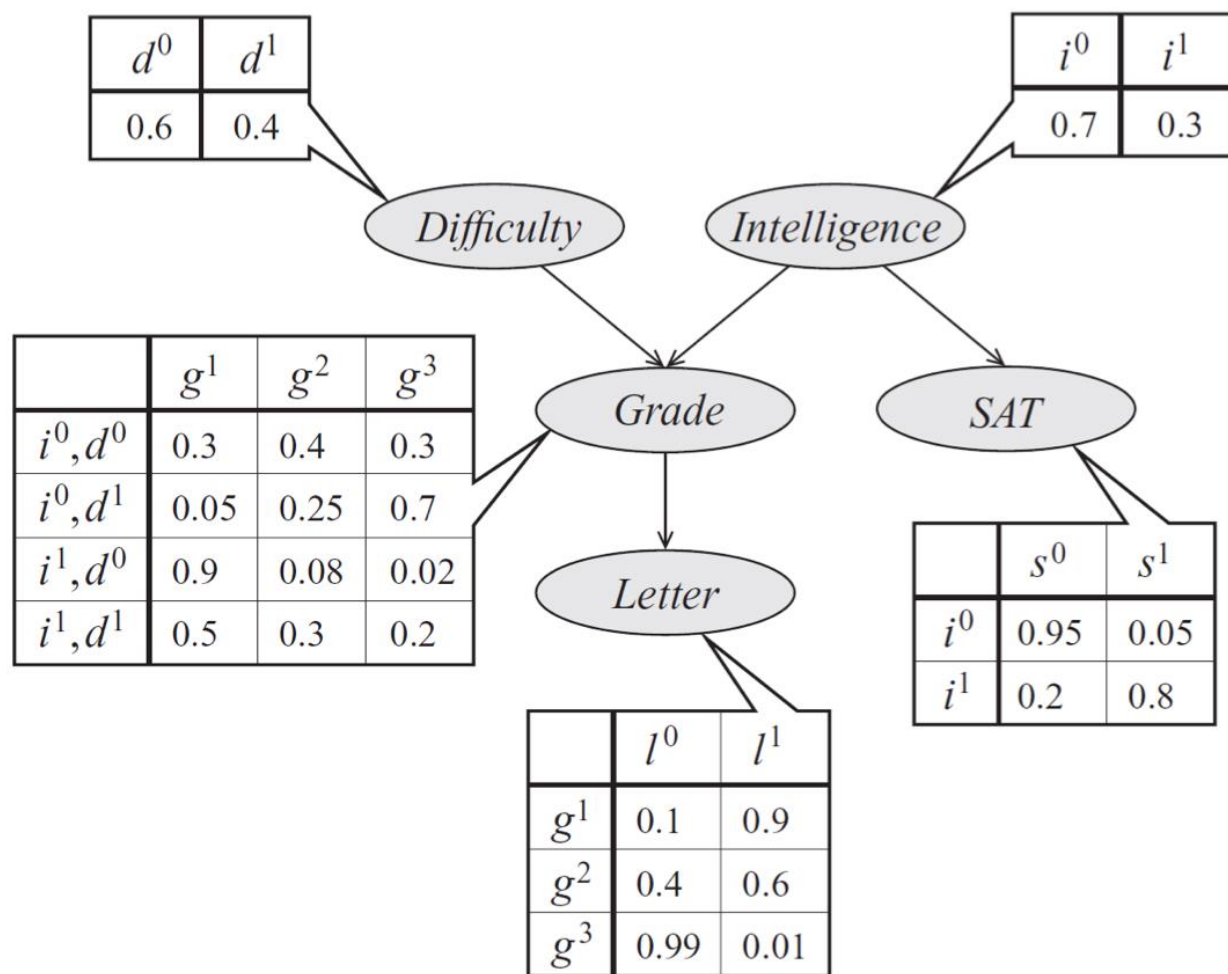
学生智力 (I) ,
课程难度 (D) ,
成绩 (G) ,
学生的SAT成绩 (S) ,
以及推荐信的质量 (L) 。

- $\text{Val}(D) = \langle d^0, d^{01} \rangle = \langle \text{easy}, \text{hard} \rangle$
- $\text{Val}(I) = \langle i^0, i^1 \rangle = \langle \text{non smart}, \text{smart} \rangle$
- $\text{Val}(G) = \langle g^0, g^1, g^2 \rangle = \langle A, B, C \rangle = \langle \text{excellent}, \text{good}, \text{average} \rangle$
- $\text{Val}(S) = \langle s^0, s^1 \rangle = \langle \text{low score}, \text{high score} \rangle$
- $\text{Val}(L) = \langle l^0, l^1 \rangle = \langle \text{weak recomm. letter}, \text{strong recomm. letter} \rangle$

除G外的所有变量都是二值的，G是三值的。因此，联合分布有48个条目。

$$2 \times 2 \times 3 \times 2 \times 2 = 48 \text{ 个参数}$$

贝叶斯网络例子 - “学生模型” 的贝叶斯网络图



联合概率分布:

$$P(I, D, G, S, L) = P(I)P(D)P(G | I, D)P(S | I)P(L | G)$$

$$\begin{aligned}
 P(i^1, d^0, g^2, s^1, l^0) &= P(i^1)P(d^0)P(g^2 | i^1, d^0)P(s^1 | i^1)P(l^0 | g^2) \\
 &= 0.3 \times 0.6 \times 0.08 \times 0.8 \times 0.4 \\
 &= 0.004608.
 \end{aligned}$$



Thank

You