

# 什么是贝叶斯网络

简单地说，贝叶斯网络是描述事件之间概率关系的计算模型，是机器学习中最重要方法之一。很多重要的模型，包括高斯混合模型（GMM）、隐马尔可夫模型（HMM）、主成份分析（PCA）、线性区分性分析（LDA）等都可以归结为贝叶斯网络的特殊形式。

我们知道，机器要拥有智能，自身需要拥有一定的“知识”。早期研究者是通过人来将知识梳理出来并教给计算机的，这些知识通常以“如果...那么”这种带条件的判断句子来表示的。这种方法简单直观，但知识之间可能发生冲突，而且知识本身也具有不确定性，当系统中的这样的知识多达几万、几十万以后，知识的维护和使用都会出现问题。后来人们就想到让机器自动去学习知识，并在学习过程中解决知识的不确定性和自洽性问题。贝叶斯网络正是这样一种学习结构。

## 1. 贝叶斯概率模型

让我们从一个天气预报系统讲起。我们知道下雨和云量、气压有很强的相关性，当下雨时云量会增加，气压会降低。将这些知识表示成概率形式，记为  $P(\text{云量}, \text{气压} | \text{下雨})$ ，意思是如果下雨的话，云量和气压的分布范围是多少。这一概率称为“条件概率”。有了这两条知识，依概率中的贝叶斯公式，就可以推理出当观察到云量为  $c$ ，气压为  $p$  时，下雨的可能性，即：

$$P(\text{下雨} | \text{云量}=c, \text{气压}=p) \propto P(\text{下雨}) P(\text{云量}=c, \text{气压}=p | \text{下雨})$$

其中  $P(\text{下雨})$  称为“下雨”这一事件的“先验概率”，意思是当没有观察到云量和气压这些“证据”时下雨的可能性。这一概率可以认为是一种经验知识，例如可以通过地理位置和季节等对是否可能会发生下雨有个先期预判。

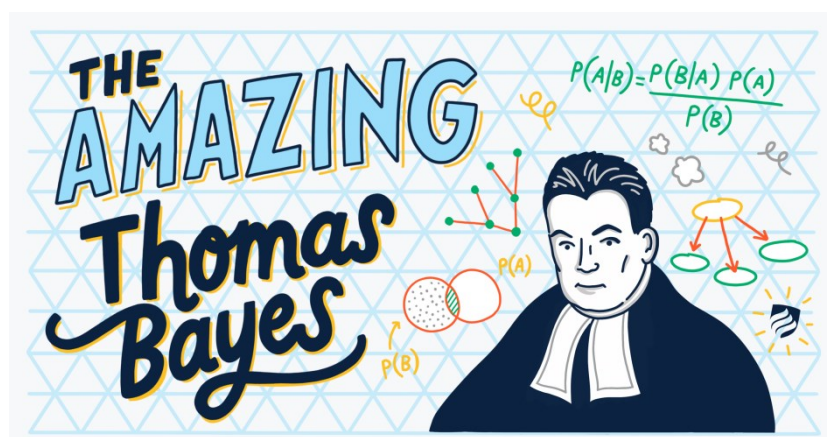


图 1: Thomas Bayes 和他的贝叶斯公式[1]

上述模型称为贝叶斯概率模型。值得强调的是，这一模型是可学习的，特别是其中的条件概率部分，如  $P(\text{云量}=c, \text{气压}=p | \text{下雨})$ ，是可以通过实验观测到的数据进行估计的，这一估计过程称为“模型训练”或“模型学习”。

除了可学习性，贝叶斯概率模型还具有如下特点：

- 第一，所有知识都是以概率形式出现的，因此可以天然描述现实世界的不确定性；
- 第二，模型将先验概率和条件概率结合起来，前者可以认为是一种经验知识，后者是由数据中得来的证据，因此兼顾了知识和数据两种信息源；
- 第三，推理是一个反向过程，即我们模型是如果下雨，云量和气压会如何变化，推理时利用观察到的云量和气压，反推下雨的可能性，这和我们日常生活中的经验是一致的。
- 第四，模型具有严格的数学基础，推理过程是可靠的；
- 第五，模型具有很强的可扩展性，如果还有其他观察变量（如气温、风速等），可以统一到同样的推理框架中。

可以看到，贝叶斯概率模型构建了一个完整的、可学习的建模和推理框架，非常适合机器学习任务。

## 2. 贝叶斯网络

原则上，贝叶斯概率模型可以描述非常复杂的系统，特别是引入隐藏变量以后，模型的表达能力可得到显著提高。然而，当系统中的变量过多是，模型的描述变得不那么直观。一个办法是将整个系统一幅有向图表示出来，图中每个节点代表一个变量，节点间的边代表变量之间的概率关系，这样表示出来的有向图通常称为贝叶斯网络（Bayes Network），也称为信任网络（Belief Network）[2]。基于这一图形化表示，可以清楚看到变量之间的相关性，推理出当某一变量发生变化时对其他变量的影响。图 2 给出前面天气预报模型的贝叶斯网络，图 3 给出一个更复杂天气系统的贝叶斯网络。

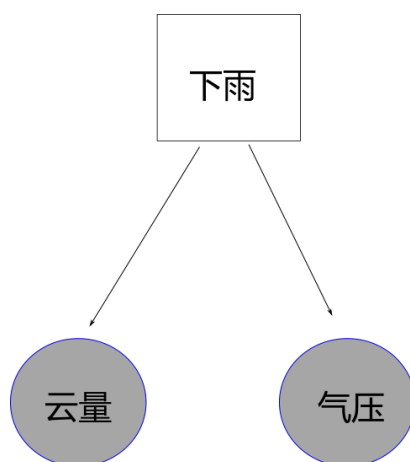


图 2:用于天气预报的简单贝叶斯网络

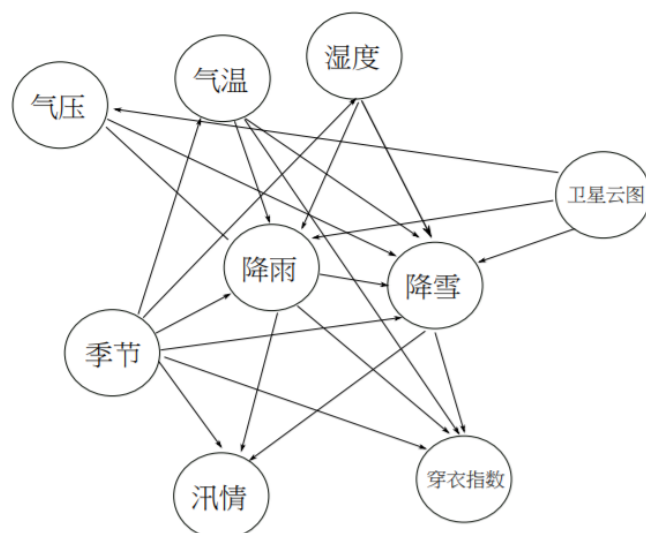


图 3: 一个复杂天气系统的贝叶斯网络[3]

贝叶斯网络为贝叶斯概率模型提供了一个强大的表示工具。基于贝叶斯网络，研究者总结出了一套通用的训练方法和推理算法，极大简化了建模过程，也促进了这一方法的推广使用。

和神经网络模型相比，贝叶斯网络具有更强的先验假设，因此对模型的约束也要严格一些。这些约束使得模型学习更可靠，可解释性也更高，但同时也限制了对复杂数据的学习能力。因此，深度学习的兴起以后，贝叶斯网络的发展相对平缓。最近一段时间，深度学习在可解释性方面的不足引起研究者的反思，人们提出将贝叶斯模型和神经网络模型相结合的新思路，如贝叶斯神经网络、深度贝叶斯模型等。从历史发展来看，概率模型和神经模型从来没有截然分开过，今天这两种源自不同思路的学习方法互相融合应该也是历史的必然。

1. Jim Kulich , The Amazing Thomas Bayes, <https://www.elmhurst.edu/blog/thomas-bayes/>
2. Pearl, Judea (1988). *Probabilistic Reasoning in Intelligent Systems* (2nd revised ed.). San Mateo
3. 王东,《机器学习导论》，清华大学出版社，2021.2.