



SNS COLLEGE OF TECHNOLOGY

Coimbatore-35
An Autonomous Institution

Accredited by NBA – AICTE and Accredited by NAAC – UGC with 'A++' Grade
Approved by AICTE, New Delhi & Affiliated to Anna University, Chennai

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

23AMB201 - MACHINE LEARNING

II YEAR IV SEM

UNIT V – Reinforcement Learning
Bayesian network

Temporal Model/S.Rajarajeswari/AP/AIML/SNSCT/

Redesigning Common Mind & Business Towards Excellence



Build an Entrepreneurial Mindset Through Our Design Thinking Frame/Work



WHAT IS BAYESIAN NETWORK

Redesigning Common Mind & Business Towards Excellence



Build an Entrepreneurial Mindset Through Our Design Thinking FrameWork

- Bayesian belief network is key computer technology for dealing with probabilistic events and to solve a problem which has uncertainty. We can define a Bayesian network as:
- "A Bayesian network is a probabilistic graphical model which represents a set of variables and their conditional dependencies using a directed acyclic graph."
- It is also called a **Bayes network**, **belief network**, **decision network**, or **Bayesian model**.
- Bayesian networks are probabilistic, because these networks are built from a **probability distribution**, and also use probability theory for prediction and anomaly detection.





- Real world applications are probabilistic in nature, and to represent the relationship between multiple events, we need a Bayesian network. It can also be used in various tasks including **prediction, anomaly detection, diagnostics, automated insight, reasoning, time series prediction, and decision making under uncertainty.**
- Note: The Bayesian network graph does not contain any cyclic graph. Hence, it is known as a **directed acyclic graph or DAG.**





Bayesian Network can be used for building models from data and experts opinions, and it consists of two parts:

- **Directed Acyclic Graph**
- **Table of conditional probabilities.**

The generalized form of Bayesian network that represents and solve decision problems under uncertain knowledge is known as an **Influence diagram**





Bayesian Network can be used for building models from data and experts opinions, and it consists of two parts:

- **Directed Acyclic Graph**
- **Table of conditional probabilities.**

The generalized form of Bayesian network that represents and solve decision problems under uncertain knowledge is known as an **Influence diagram**



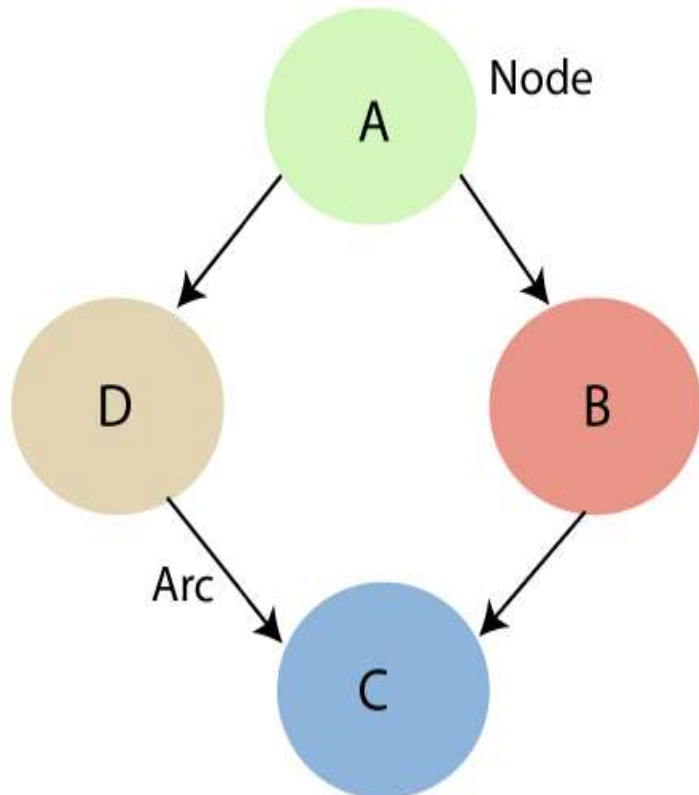


Bayesian Network can be used for building models from data and opinions, and it consists of two parts:

Redesigning Common Mind & Business Towards Excellence



Build an Entrepreneurial Mindset Through Our Design Thinking Framework



Each **node** corresponds to the random variables, and a variable can be **continuous** or **discrete**. **Arc or directed arrows** represent the causal relationship or conditional probabilities between random variables. These directed arrows connect the pair of nodes in the graph.

These links represent that one node directly influence the other node, and if there is no directed link that means that nodes are independent with each other

In the above diagram, A, B, C, and D are random variables represented by the nodes of the network graph.

If we are considering node B, which is connected with node A by a directed arrow, then node A is called the **parent** of Node B.

Node C is independent of node A.



○ COMPONENTS

The Bayesian network has mainly two components:

Causal Component

Actual numbers

Each node in the Bayesian network has condition probability distribution $P(X_i | \text{Parent}(X_i))$, which determines the effect of the parent on that node

Bayesian network is based on Joint probability distribution and conditional probability. So let's first understand the joint probability distribution:.





- Joint probability distribution:
- If we have variables $x_1, x_2, x_3, \dots, x_n$, then the probabilities of a different combination of $x_1, x_2, x_3 \dots x_n$, are known as Joint probability distribution.
- $P[x_1, x_2, x_3, \dots, x_n]$, it can be written as the following way in terms of the joint probability distribution.
- $= P[x_1 | x_2, x_3, \dots, x_n] P[x_2, x_3, \dots, x_n]$
- $= P[x_1 | x_2, x_3, \dots, x_n] P[x_2 | x_3, \dots, x_n] \dots P[x_{n-1} | x_n] P[x_n]$.





- let's understand the Bayesian network through an example by creating a directed acyclic graph:

- **Example:** Harry installed a new burglar alarm at his home to detect burglary. The alarm reliably responds at detecting a burglary but also responds for minor earthquakes. Harry has two neighbors David and Sophia, who have taken a responsibility to inform Harry at work when they hear the alarm. David always calls Harry when he hears the alarm, but sometimes he got confused with the phone ringing and calls at that time too. On the other hand, Sophia likes to listen to high music, so sometimes she misses to hear the alarm. Here we would like to compute the probability of Burglary Alarm.





- Calculate the probability that alarm has sounded, but neither a burglary, nor an earthquake occurred, and David and Sophia both called the Harry.

Solution:

- The Bayesian network for the above problem is given below. The network structure is showing that burglary and earthquake is the parent node of the alarm and directly affecting the probability of alarm's going off, but David and Sophia's calls depend on alarm probability.
- The network is representing that our assumptions do not directly perceive the burglary and also do not notice the minor earthquake, and they also not confer before calling.
- The conditional distributions for each node are given as conditional probabilities table or CPT.
- Each row in the CPT must be sum to 1 because all the entries in the table represent an exhaustive set of cases for the variable.
- In CPT, a boolean variable with k boolean parents contains 2^k probabilities. Hence, if there are two parents, then CPT will contain 4 probability values





Redesigning Common Mind & Business Towards Excellence



Build an Entrepreneurial Mindset Through Our Design Thinking Framework

- **List of all events occurring in this network:**
- **Burglary (B)**
- **Earthquake(E)**
- **Alarm(A)**
- **David Calls(D)**
- **Sophia calls(S)**



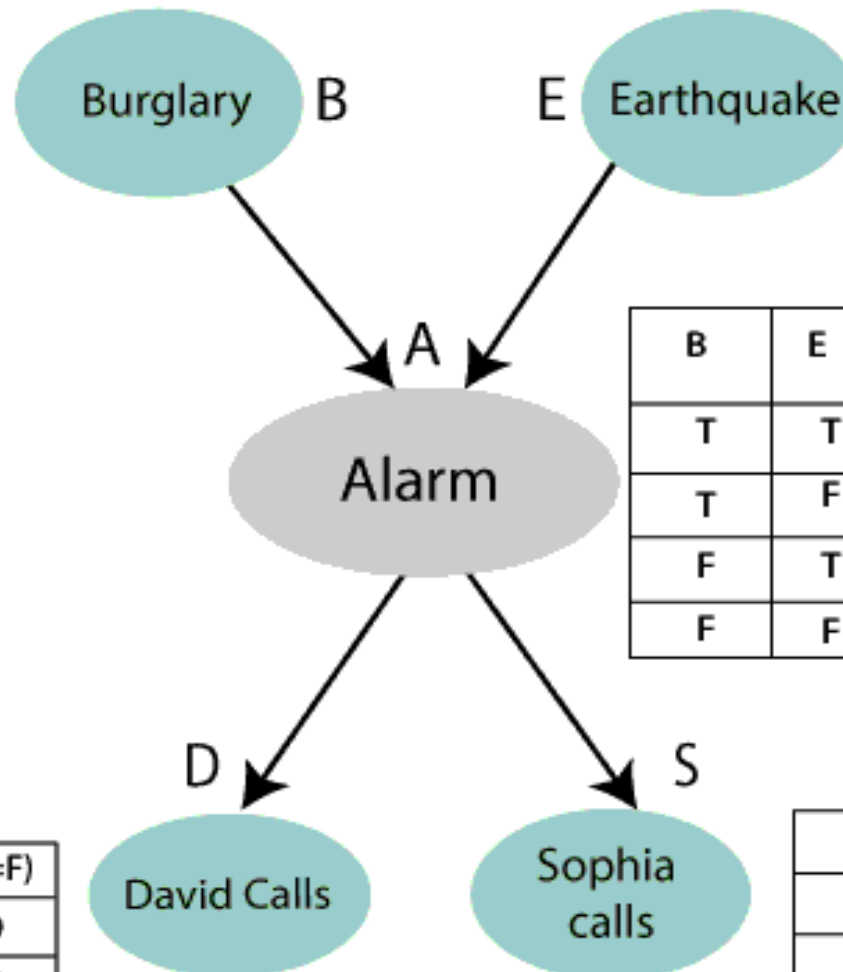


- We can write the events of problem statement in the form of probability: $P[D, S, A, B, E]$, can rewrite the above probability statement using joint probability distribution:
- $P[D, S, A, B, E] = P[D \mid S, A, B, E] \cdot P[S, A, B, E]$
- $= P[D \mid S, A, B, E] \cdot P[S \mid A, B, E] \cdot P[A, B, E]$
- $= P[D \mid A] \cdot P[S \mid A, B, E] \cdot P[A, B, E]$
- $= P[D \mid A] \cdot P[S \mid A] \cdot P[A \mid B, E] \cdot P[B, E]$
- $= P[D \mid A] \cdot P[S \mid A] \cdot P[A \mid B, E] \cdot P[B \mid E] \cdot P[E]$





T	0.002
F	0.998



T	0.001
F	0.999

B	E	P(A=T)	P(A=F)
T	T	0.94	0.06
T	F	0.95	0.04
F	T	0.69	0.69
F	F	0.999	0.999

A	P(D=T)	P(D=F)
T	0.91	0.09
F	0.05	0.95

A	P(S=T)	P(S=F)
T	0.75	0.25
F	0.02	0.98



- let's take the observed probability for the Burglary and earthquake component:
- $P(B = \text{True}) = 0.002$, which is the probability of burglary.
- $P(B = \text{False}) = 0.998$, which is the probability of no burglary.
- $P(E = \text{True}) = 0.001$, which is the probability of a minor earthquake
- $P(E = \text{False}) = 0.999$, Which is the probability that an earthquake not occurred.





- conditional probability table for Alarm A:
- The Conditional probability of Alarm A depends

Conditional probability table for Alarm A:

The Conditional probability of Alarm A depends on Burglar and earthquake:

B	E	P(A= True)	P(A= False)
True	True	0.94	0.06
True	False	0.95	0.04
False	True	0.31	0.69
False	False	0.001	0.999





Conditional probability table for David Calls:

The Conditional probability of David that he will call depends on the probability of Alarm.

A	P(D= True)	P(D= False)
True	0.91	0.09
False	0.05	0.95





Conditional probability table for Sophia Calls:

The Conditional probability of Sophia that she calls is depending on its Parent Node "Alarm."

A	P(S= True)	P(S= False)
True	0.75	0.25
False	0.02	0.98

From the formula of joint distribution, we can write the problem statement in the form of probability distribution:

$$P(S, D, A, \neg B, \neg E) = P(S|A) * P(D|A) * P(A|\neg B \wedge \neg E) * P(\neg B) * P(\neg E).$$

$$= 0.75 * 0.91 * 0.001 * 0.998 * 0.999$$

$$= 0.00068045.$$





- Hence, a Bayesian network can answer any query about the domain by using Joint distribution.
- The semantics of Bayesian Network:
- There are two ways to understand the semantics of the Bayesian network, which is given below:
- 1. To understand the network as the representation of the Joint probability distribution.
- It is helpful to understand how to construct the network.
- 2. To understand the network as an encoding of a collection of conditional independence statements.
- It is helpful in designing inference procedure.

