Vaidehi Vatsaraj
2018130060
Data Science, 2022
# Tut 7: Machine Learning 2

Bayesian networks provide useful benefits as a probabilistic model.

For example:

- **Visualization**. The model provides a direct way to visualize the structure of the model and motivate the design of new models.
- **Relationships**. Provides insights into the presence and absence of the relationships between random variables.
- **Computations**. Provides a way to structure complex probability calculations.

## How to Develop and Use a Bayesian Network

Designing a Bayesian Network requires defining at least three things:

- **Random Variables**. What are the random variables in the problem?
- **Conditional Relationships**. What are the conditional relationships between the variables?
- **Probability Distributions**. What are the probability distributions for each variable?

It may be possible for an expert in the problem domain to specify some or all of these aspects in the design of the model.

In many cases, the architecture or topology of the graphical model can be specified by an expert, but the probability distributions must be estimated from data from the domain.

Both the probability distributions and the graph structure itself can be estimated from data, although it can be a challenging process. As such, it is common to use learning algorithms for this purpose; for example, assuming a Gaussian distribution for continuous random variables gradient ascent for estimating the distribution parameters.

Once a Bayesian Network has been prepared for a domain, it can be used for reasoning, e.g. making decisions.

Reasoning is achieved via inference with the model for a given situation. For example, the outcome for some events is known and plugged into the random variables. The model can be used to estimate the probability of causes for the events or possible further outcomes.

*"Reasoning (inference) is then performed by introducing evidence that sets variables in known states, and subsequently computing probabilities of interest, conditioned on this evidence."*

Practical examples of using Bayesian Networks in practice include medicine (symptoms and diseases), bioinformatics (traits and genes), and speech recognition (utterances and time).

- ## Example of a Bayesian Network

We can make Bayesian Networks concrete with a small example.

Consider a problem with three random variables: A, B, and C. A is dependent upon B, and C is dependent upon B.

We can state the conditional dependencies as follows:

- A is conditionally dependent upon B, e.g. P(A|B)
- C is conditionally dependent upon B, e.g. P(C|B)

We know that C and A have no effect on each other.


- A is conditionally dependent upon B, e.g. P(A|B)
- C is conditionally dependent upon B, e.g. P(C|B)

We know that C and A have no effect on each other.


We can also state the conditional independencies as follows:

- A is conditionally independent from C: P(A|B, C)
- C is conditionally independent from A: P(C|B, A)

Notice that the conditional dependence is stated in the presence of the conditional independence. That is, A is conditionally independent of C, or A is conditionally dependent upon B in the presence of C.

We might also state the conditional independence of A given C as the conditional dependence of A given B, as A is unaffected by C and can be calculated from A given B alone.

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

We can see that B is unaffected by A and C and has no parents; we can simply state the conditional independence of B from A and C as P(B, P(A|B), P(C|B)) or P(B).

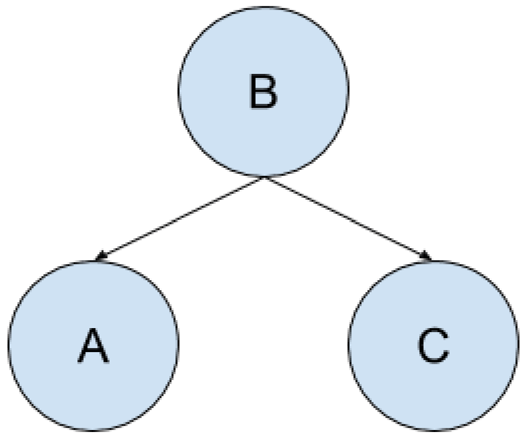
We can also write the joint probability of A and C given B or conditioned on B as the product of two conditional probabilities; for example:

- P(A, C | B) = P(A|B) * P(C|B)

The model summarizes the joint probability of P(A, B, C), calculated as:

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

We can draw the graph as follows:




Bayesian Network

Notice that the random variables are each assigned a node, and the conditional probabilities are stated as directed connections between the nodes. Also notice that it is not possible to navigate the graph in a cycle, e.g. no loops are possible when navigating from node to node via the edges.

Also notice that the graph is useful even at this point where we don't know the probability distributions for the variables.

**<u>Exercise</u>**

Suppose you are working for a financial institution and you are asked to build a fraud detection system. You plan to use the following information:

– When the card holder is traveling abroad, fraudulent transaction are more likely since tourists are prime targets for thieves. More precisely, 1% of transactions are fraudulent when the card holder is traveling, whereas only 0.2% of the transactions are fraudulent when he is not traveling. On average, 5% of all transactions happen while the card holder is traveling. If a transaction is fraudulent, then the likelihood of a foreign purchase increases, unless the card holder happens to be traveling. More precisely, when the card holder is not traveling, 10% of the fraudulent transactions are foreign purchases, whereas only 1% of the legitimate transactions are foreign purchases. On the other hand, when the card holder is traveling, 90% of the transactions are foreign purchases regardless of the legitimacy of the transactions

Bayesian network

Travel ——— Forn purchase

Fraud

↑ prob of foreign purchase ; Fraud likely

travel, fraud ———→ Forn purchase

travel explains forn purchase ∴ Fraud less likely

| T | F |
|---|---|
| 0.05 | 0.95 |

| Travel | T | F |
|---|---|---|
| T | 0.01 | 0.99 |
| F | 0.002 | 0.998 |

| Travel | Fraud | T | F |
|---|---|---|---|
| T | T | 0.9 | 0.1 |
| F | T | 0.1 | 0.9 |
| T | F | 0.9 | 0.1 |
| F | F | 0.01 | 0.99 |

1) System has detected the foreign purchases. What is the probability of a fraud if we don't know whether the card holder is traveling or not?

1) [ travel = ?, forn purchase = T,
fraud = ? ]

$P(\text{fraud} = T \mid \text{forn purchase} = T)$

$= \alpha * [P(\text{fraud} = T \mid \text{travel} = T) * \ldots$
$\ldots * P(\text{travel} = F)]$

$= \alpha * [0.01 * 0.9 * 0.05 \underset{+}{*} 0.002 * 0.1 * 0.95]$

$= \alpha * [0.00045 + 0.00019] = 0.00064\alpha$

Similarly,

$P(\text{fraud} = F \mid \text{forn purchase} = T)$
$= 0.054081\,\alpha$

$$\alpha = \frac{1}{0.00064 + 0.054081} = 18.291$$

$P(\text{fraud} = T \mid \text{forn purchase} = T)$
$= 0.00064 \times 18.291$
$= 0.0117 = 1.17\%$

2) Suppose that probability more than 1% causes an agent to call the client to confirm the transaction. An agent calls but the card holder is not at home. Her spouse confirms that she is out of town on a business trip. How does the probability of the fraud changes based on this new piece of information?

2) $P(\text{fraud} = T \mid \text{forn purchase} = T \mid \text{travel} = T)$
$= 0.00045\,\alpha$

$P(\text{fraud} = F \mid \text{forn purchase} = T \mid \text{travel} = T)$
$= 0.04455\,\alpha$

$$\alpha = \frac{1}{0.00045 + 0.04455} = 22.222$$

$\therefore$ $P(\text{fraud} = T \mid \text{forn purchase} = T \mid \text{travel} = T)$
$= 0.00045 \times 22.222$
$= 0.01 = 1\%$