TRIBHUVAN UNIVERSITY

INSTITUTE OF ENGINEERING
PULCHOWK CAMPUS

A
REPORT
ON

BAYESIAN LEARNING & BAYESIAN BELIEF NETWORKS

SUBMITTED BY:

Diwas Adhikari **(076BEI014)**

SUBMITTED TO:
Prof. Basanta Joshi
Department of Electronics & Computer Engineering
Pulchowk Campus, IOE

SUBMISSION DATE

August 11, 2023

(For course completion of BEI – Artificial Intelligence)

Joint Probability Distribution

A joint probability distribution shows a probability distribution for two or more than two random variables. For two events X and Y, the joint probability distribution function is given by:

$$\mathbf{f}_{X,Y}\,(x, y) = P\,(X = x,\ Y = y)$$

The whole point of the joint distribution is to look for a relationship between two variables. The following table shows some probabilities for X and Y happening at the same time. For example: if we want to evaluate the probability for an event with Y = 2 and X = 3, we find the intersection of Y = 2 and X = 3.

		X		
		1	2	3
	1	0	1/6	1/6
Y	2	1/6	0	1/6
	3	1/6	1/6	0

		X		
		1	2	3
	1	0	1/6	1/6
Y	2	1/6	0	1/6
	3	1/6	1/6	0

Conditional Probability & Independence

Conditional probability is a measure of the probaility of an event occurring, given that another event has already occurred. If E and F be any two events such that P(F) ≠ 0. Then the conditional probability of event E given that event F has occurred, is denoted as P(E|F).

$$P(E/F) = \frac{P(E \cap F)}{P(F)}$$

Events turn out to be mutually exclusive if they have no common probabilities in their sample space. Hence, two events E and F are independent if and only if their joint probaiblity equals the product of their probabilities such that:

$$P(E \cap F) = P(E).P(F)$$

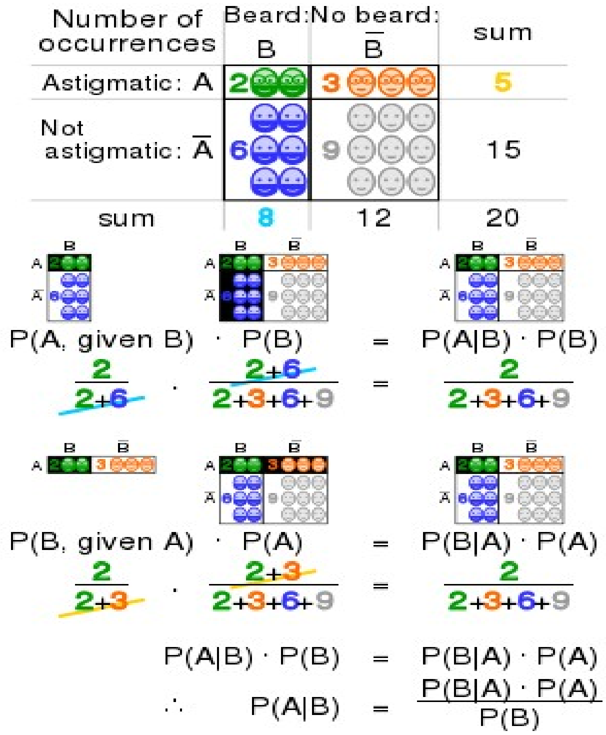
This further implies that if these two events are conditionally independent, the occurrence of E does not affect the probability of event F and vice versa. In other words, these events are totally independent to each other such that their conditional probabilities are as follows:

$$P(E/F) = \frac{P(E \cap F)}{P(F)} = P(E) \qquad P(F/E) = \frac{P(E \cap F)}{P(E)} = P(F)$$

Bayes' Theorem

Bayes' theorem describes the probability of occurrence of an event related to any condition. For any event E to occur based on event F's occurrence, we determine that posterior probability if we have an idea of the likelihood of event F based on event E's occurrence.

$$P(E/F) = \frac{P(F/E).P(E)}{P(F)}$$



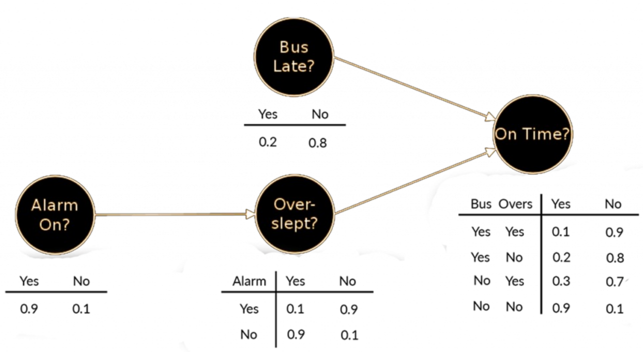
Bayesian Learning / Bayesian Inference

Bayesian inference is a learning technique that uses probabilities to define and reason about our beliefs. In particular, this method gives us a way to properly update our beliefs when new observations are made. Bayesian Learning is not about identifying a sample space and determining probabilities of elementary events and random variables to compute values in joint probability distributions. Instead, we identify random variables directly, and we determine probabilistic relationships among the random variables. All of the nodes are Boolean variables such that Bayes' theorem of conditional probability P(E|F) will work for P(¬E|F) too to predict the conditional probability of an event E not occuring due to event F where P(¬E) = 1 – P(E).

$$P(E/F)=\frac{P(F/E).P(E)}{P(F)} \qquad\qquad P(\neg E/F)=\frac{P(F/\neg E).P(\neg E)}{P(F)}$$

Bayesian Belief Networks

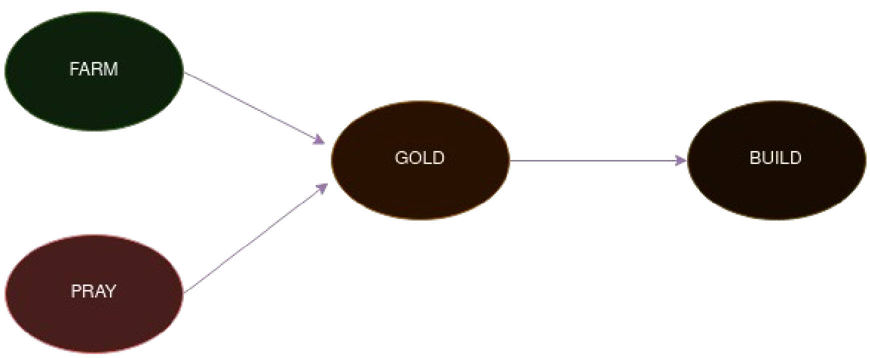
Bayesian networks are probabilistic graph models which are represented by directed acyclic graphs (DAG) whose nodes represent variables associated to their probabilistic functions and their edges represent conditional dependencies. Hence, if there are nodes which aren't connected to any other node, it represents variables that are conditionally independent of each other.



Construction of Bayesian Networks

To construct a Bayesian network, we need to first evaluate a scenario with all the possible events and make sure if there is any conditional dependence between any two events. Thus, each event is associated to a particular node with a certain probabilistic function. If an event with high likelihood can trigger another event to occur, it is directed by an edge (arrow) towards that other event's node ensuring the dependency. We should know that if event E is conditionally dependent on event F, it doesn't guarantee that event F will happen due to the occurrence of event E. Hence, this is why belief networks turn out to be directed graphs. For example: the event of an adult getting married can result to an event of a baby being born with a certain probability but, that doesn't suggest that the dependency holds true in the reverse order.

Let's create our own Bayesian network based on a hypothetical scenario of Ancient Egypt. Assume that we have four different events: "farm crops" denoted by node 'Farm', "pray to Isis" denoted by node 'Pray', "earn gold" denoted by node 'Gold' and "build pyramid" denoted by node 'Build'. If an Egyptian nomad farms various crops, he can earn lots of gold but, he can also make a fortune with a small certainty if he prays to his deity – Isis. If he manages to earn a lot, he can build his own pyramid. Here, the event 'Gold' is dependent on two events 'Farm' and 'Pray' with different probabilities so, two arrows are directed respectively from those events to node 'Gold'. Similarly, the event of building a pyramid depends on him earning wealth so that arrow from node 'Gold' is directed towards node 'Build'.



The joint probabiility of this network based on the conditional dependencies on the graph is given by:

$$P(F,P,G,B)=P(Farm).P(Pray).P(Gold/Farm,Pray).P(Build/Gold)$$

Thus, by analysis or research, we might obtain the probabilities of these events occuring with certainty. Those prior probabilities are crucial to use the model above for evaluating predictions for this scenario. The probability of a nomad being engaged in farming is low at about 10% whereas a nomad being religious and a devotee to Isis is very high at about 80%. The chances of being wealthy and earning gold is dependent on those other events. Similarly, the chances of someone building a pyramid is based on the event of earning gold. Thus, we need to develop a Boolean table with probabilities for those nodes.

FARM	PRAY	GOLD
F	F	0.01
F	T	0.25
T	F	0.5
T	T	0.75

GOLD	BUILD
F	0.001
T	0.05

Prediction using Bayesian Networks

Using the above joint probability model, we can eventually predict the likelihood of a chain of events occuring together due to their dependencies.

If we want to find out the probability of someone not being able to build a pyramid despite earning gold by praying to Isis, we can use the above equation.

$$P(\neg F, P, G, \neg B) = P(\neg Farm) . P(Pray) . P(Gold / \neg Farm, Pray) . P(\neg Build / Gold)$$

Since, P(Farm) = 10% = 0.1 ;

P(¬Farm) = 1 - P(Farm) = 1 - 0.1 = 0.9

P(Pray) = 80% = 0.8

P(Gold / ¬Farm, Pray) = P(Gold) < Farm = F, Pray = T >

= 0.25

P(¬Build / Gold) = 1 - P(Build) <Gold = T>

= 1 – 0.05 = 0.95

$$P(\neg F, P, G, \neg B) = 0.9 * 0.8 * 0.25 * 0.95 = 0.171 = 17.1\%$$

Hence, the probabiility of someone not being able to build a pyramid despite earning gold by praying to Isis is about 17.1%.

CONCLUSION

Thus, in this way, Bayesian belief networks can be constructed easily and predictions can be made based on inferences which are condition-based on the prior information. I was able to make a Bayesian network by making up a random problem on my own using this simple approach. I would say that Bayesian inference is a very convenient method which is heavily influenced by the prior information fed to it.

SUBMITTED BY:

Diwas Adhikari
076BEI014 – (Roll Number: 14)
Bachelor in Electronics & Communication Engineering (BEI)
076bei014.diwas@pcampus.edu.np
lunaticoda123@gmail.com