UNIT - 5 | Probabilistic Reasoning | ①

* Uncertainty :

Till now, we have learned knowledge representation using First order Logic & propositional Logic with certainty, which mean we sure about predicates.

→ with this knowledge representation, we might write A → B, which means if A is true, then B is true. but consider a situation where we're not sure about whether A is True or not then we can express this statement, this situation is called uncertainty.

→ so to represent uncertain knowledge, where we're not sure about predicates, we need uncertain reasoning (or) Probabilistic reasoning.

Procedure of Handling uncertain knowledge

→ Probabilistic reasoning is a way of knowledge representation where we apply the concept of probability to Indicate the uncertainty

In probabilistic reasoning, we combine probability theory with Logic to handle the uncertainty

In probabilistic reasoning, there are Two ways to solve problems with uncertain knowledge

→ Baye's rule

→ Bayesian statistics

Term:

Probability: It can be defined as a chance than an uncertain Event will occur.

→ It is the numerical measure of likelihood that an event will occur.

value of probability remains always b/w 0 & 1 & That represent ideal uncertainties.

$$\boxed{0 \leq P(A) \leq 1}$$

P(A) = Probability of Event A

P(A) = 0, total uncertainty of Event A

P(A) = 1 = total Certainity of Event A

Formula: probability of uncertain Event

$$\text{probability of occurrence} = \frac{\text{No. of desired outcomes}}{\text{total No. of outcomes}}$$

$P(\neg A)$ = probability of not happening Event

$$P(\neg A) + P(A) = 1$$

Event: Each possible outcome of a variable is called an Event

Sample space: The collection of all possible events is called sample space.

Random variables: Random variables are used to represent the events & objects in the real world

prior probability: prior probability of an event is probability computed before observing new information

Eg: If prior probability, that finger has to be fractured to 0.1, then written as

$p(\text{fracture} = \text{true}) = 0.1$ (or) $P(\text{fracture}) = 0.1$

Condition probability:

Condition probability is a probability of occurring an event when another event has already happened.

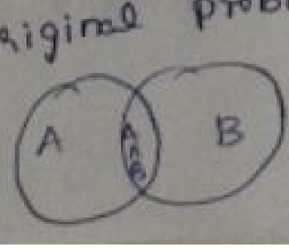
Let's suppose, we want to calculate event A, when event B has already occured.

The probability of A under condition of B

$$\boxed{P(A|B) = \frac{P(A \wedge B)}{P(B)}}$$

$P(A \wedge B)$ = Joint probability of a and B

$P(B)$ = marginal probability of B



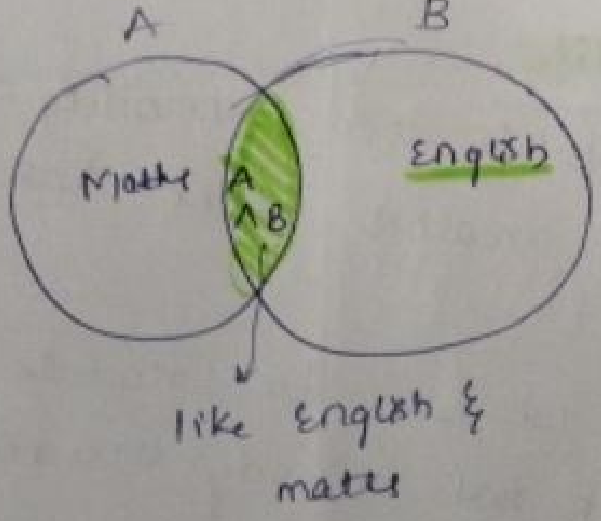
Eg: In a class. 70% of students who like English & 40% of the students who likes English & Maths. what is percent of students who like English also like mathematics.

A → Event that student like maths

B → Event that student like English.

$$P(A|B) = \frac{P(A \wedge B)}{P(B)} = \frac{0.4}{0.7} = 57\%$$

57% are the students who like English & maths



Tossed a Coin :-

Find a possibity of head in 2 events.

Two coins :- A & B

sample space :- { HH, HT, TH, TT }

possibility of head in Each Coin = 0.5 = 50%

Condition probability example

$P(A/B) = ?$

$S = \{1, 2, 3, 4, 5, 6\}$

$A = \{1, 3, 5\}$ $B = \{3, 4, 5\}$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A \text{ and } B)}{P(B)} = \frac{2/6}{3/6} = \frac{2}{3}$$

Baye's Rule:

→ Bayes' Theorem is also known as Bayes' rule.

→ Bayes' Law (or) Bayesian reasoning, which determines the probability of an event with uncertain knowledge.

In probability theory, it relates the conditional probability & marginal probabilities of two random events.

* It is a way to calculate the value of P(B|A) with knowledge of P(A|B)

→ Bayes' theorem allows updating the probability prediction of event by observing new information of real world.

Eg: If cancer corresponds to one's age then by using Bayes' theorem, we can determine probability of cancer more accurately with help of age.

→ Bayes theorem can be derived using product rule and conditional probability of event A with known event B.

Product rule;

$P(A \wedge B) = P(A|B)\, P(B)$ or

⑥

Similiarly, probability of Event B with known A,

$$P(A \wedge B) = P(B|A)\, P(A) \longrightarrow ②$$

[∵ $P(A|B)$ = Prob of Event A when Event B happens similarly $P(B|A)$]

Equate ① & ②

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

$$\boxed{P(A|B) = \frac{P(B|A)\, P(A)}{P(B)}} = \text{Bayes rule}$$

$P(A|B)$ = known as posterior, which need to calculate.

$P(B|A)$ = called likelihood,

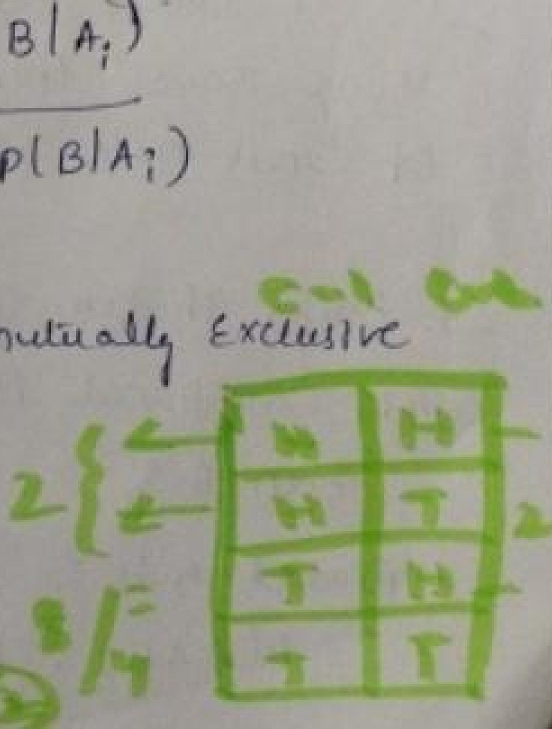
$P(A)$ = called prior probability. Probability of hypothesis before considering Evidence.

$P(B)$ = Marginal probability. Pure probability of an Evidence

General Equation. $P(B)$ written with ~~respect A~~, $P(B) = P(A) \cdot P(B|A)$

$$P(A_i|B) = \frac{P(A_i) * P(B|A_i)}{\sum_{i=1}^{k} P(A_i) * P(B|A_i)}$$

$A_1, A_2, A_3 \ldots$ An are set of mutually Exclusive and Exhaustive Events



Applying Baye's Rule:

Bayes rule allow us to Compute single-term P(B|A) in terms of P(A|B), P(B) & P(A).

→ This is useful in cases where we have good probability of these 3 terms & want to determine 4th one

→ suppose we want to perceive the Effect of some unknown cause, & want to compute that cause,

$$P(\text{cause} \mid \text{Effect}) = \frac{P(\text{effect} \mid \text{cause})\, P(\text{cause})}{P(\text{effect})}$$

eg: What is the probability that a patient has diseases ~~meaningitis~~ Cancer with a stiff neck?

Data: A Doctor is aware that Disease cancer causes a patient to have a stiff neck, & it occurs 80% of time. He is also aware of some more facts.

→ known probability that a patient has cancer disease is $\frac{1}{30,000}$

→ known probability that a patient has stiff neck is 2%

$P(a) = 0.2$ $\qquad$ $P(a/b) = 0.8$

$P(b) = 1/30000$

$$P(b|a) = \frac{P(a|b)\,P(b)}{P(a)}$$

$$= \frac{0.8 * \left(\frac{1}{30000}\right)}{0.02} = 0.0013$$

Eg:2: From a deck of playing cards, a single card is taken. Probability that the card is king is 4/52. Calculate posterior probability P(king/face) which means the drawn face card is king card.

$$P(\text{king}|\text{face}) = \frac{P(\text{face}|\text{king}) * P(\text{king})}{P(\text{face})}$$

$P(\text{king}) = 4/52$

$P(\text{face}) = 3/13 \rightarrow$ (Jack, Queen, king)

$P(\text{face}|\text{king}) = 1$

$$P(\text{king}|\text{face}) = \frac{1 * (1/13)}{(3/13)} = 1/3$$

Bayesian Networks

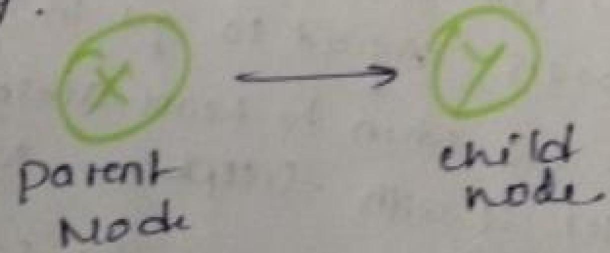
→ Bayesian Networks are probabilistic, because these Networks are Built from a probability distribution

→ Bayesian Network also known as Belief network, probabilistic Network, Causal Network or a knowledge map is directed graph in which Each node has a quantitive probability Information.

Properties:

(i) A node in a Bayesian Network represents a discrete or a Continuous random variable.

(ii) A set of Directed Arrows/ links is used for Connecting the nodes within a Network i.e, an arrow from node x to y represents that x is a parent of y.



(iii) It is DAG (Directed Acyclic graph) & it doesn't Contain any Directed Cycle.

(iv) Each Individual node has an associated Conditional probability disturbution $P(X_i | Parents(X_i))$ that determines the Effect of parents on a particular node.

Bayesian Network can be used for Building models from Data & expert opinions. & It Consists of two parts.

1. Directed Acyclic graph
2. Table of Conditional probabilities

→ Generalized form of Bayesian Network that represents & solve Decision problems under uncertain knowledge is Called Influence Diagram.

Explanation of Bayesian Network.

Let's understand the Bayesian Network through an Example by creating the ~~Directed Acyclic graph~~ Joint probability distribution

You have installed a burglar alarm at home. The Alarm not only Detects but also responds to minor Earthquakes. You have 2 neighbours, Agreed to get in touch when its rings

chris → call you when he hears alarm but sometimes Confuses it with telephone ringing & calls.

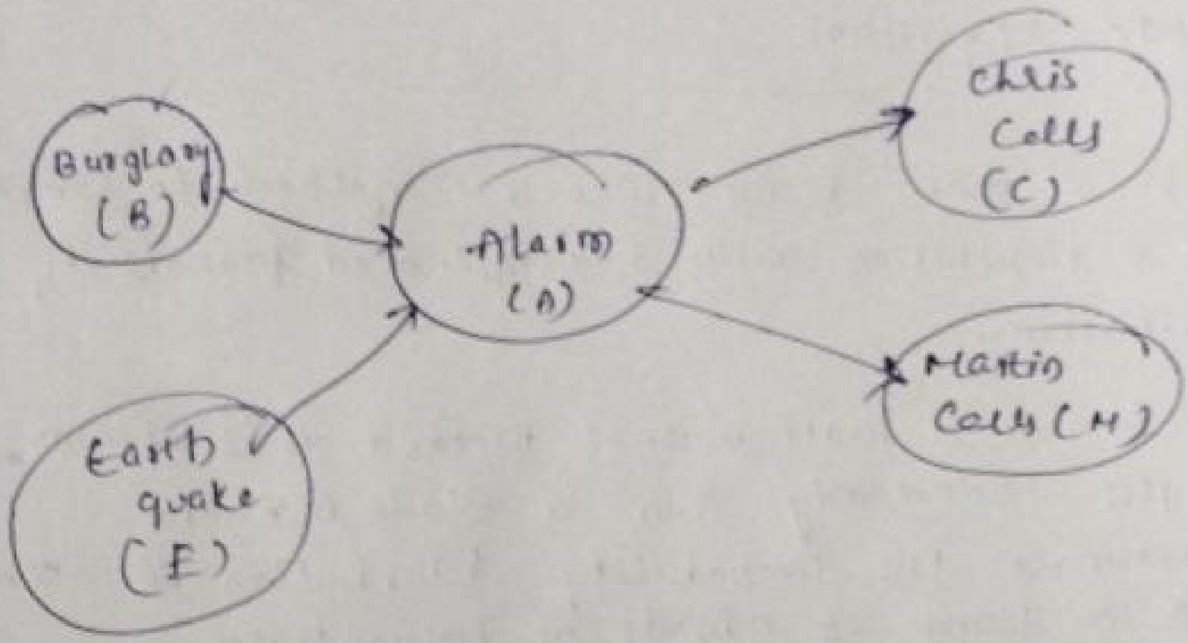
Martin → is a music lover who sometimes misses alarm due to Loud noise he plays.

Problem: In a Bayesian Network, we consider nodes as random variables.

There are 5 nodes.

Links acts as Casual dependencies that define the Relationship Between the nodes.

Both chris & Martin Call when there is a alarm.



probability of 5 nodes,

$$P[B, E, A, C, M] = P[C|A]\ P[M|A]\ P[A|B,E]\ P[B|E]\ P[E]$$

NODE - B

True	0.015
False	0.985

NODE E

True	0.010
False	0.990

NODE A:

B	E	P(A=T)	P(A=F)
T	T	0.90	0.10
T	F	0.95	0.05
F	F	0.85	0.75
F	T	0.99	0.99

NODE C

Alarm	P(C=T	P(C=F)
T	0.90	0.10
F	0.10	0.90

NODE M:

Alarm	P(M=True	P(M=F
T	0.85	0.15
F	0.05	0.95

MCMC Algorithm

Markov chain Monte Carlo is a method to sample from a population with a complicated probability disturbution.

→ MCMC is a mathematical method that draws samples randomly from a Black box to approximate the probability disturbution of attributes over a Range of objects or future states.

→ MCMC : The solution to sampling probability disturbutions in high dimensions is to use "MCMC"