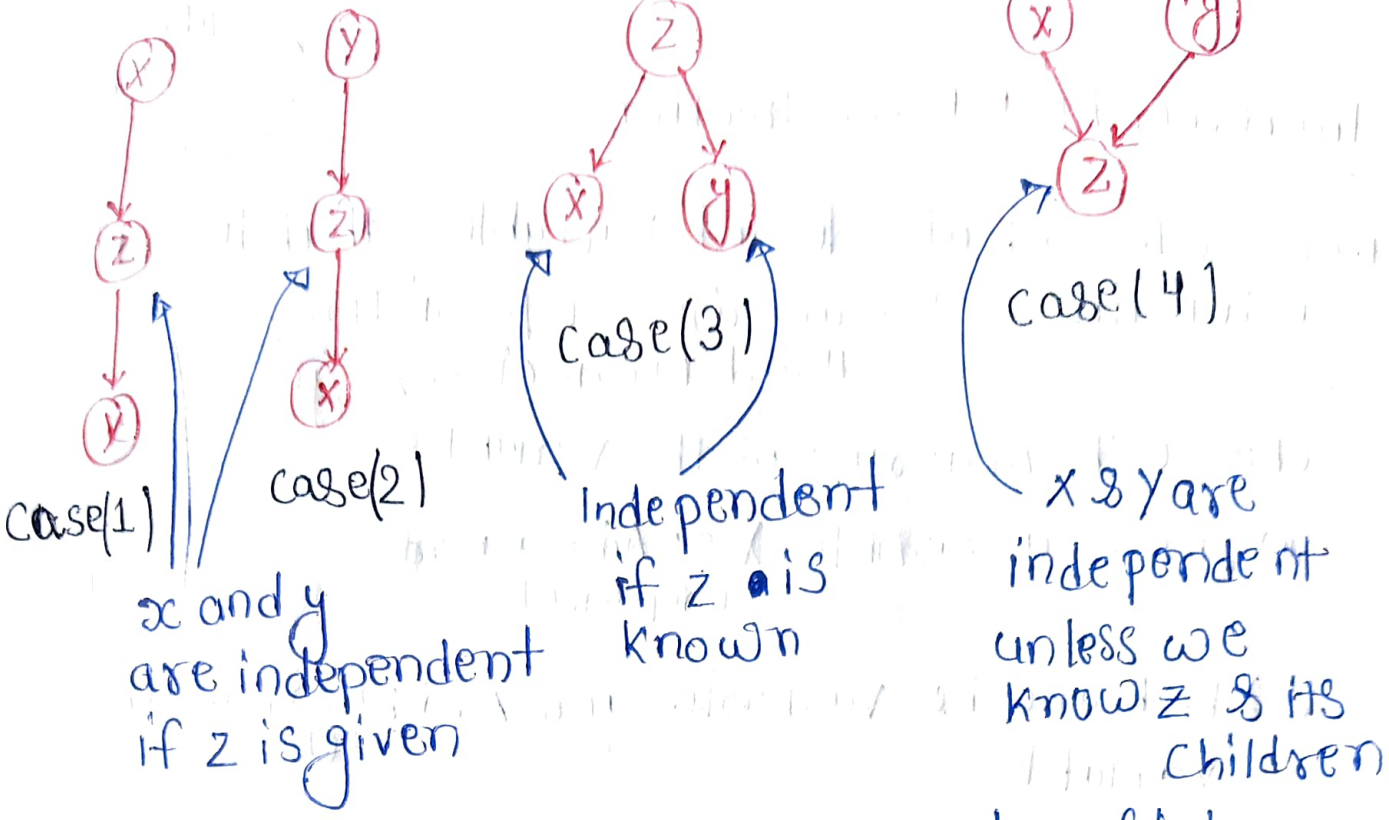
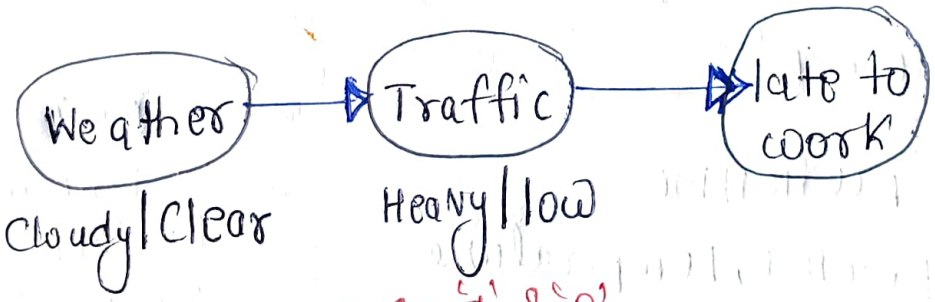
Lecture-17



Dependency mean: if we know the value of 'x' then we can determine the likelihood of 'y'

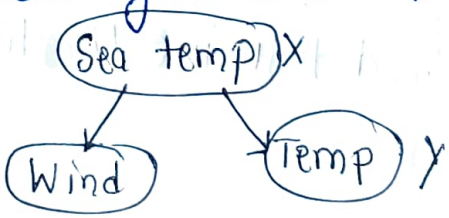


① Bayesian Network
② Fuzzy reasoning

Joint → JP (Probability)
→ Structure → B N
→ Quantative Analysis
→ CP table (Probability)
↑ Conditional

Incremental B.N Construction

1. Choose the set of relevant variables that describes a domain ⟶ heart disease prediction, Health Parameters, sym
2. Choose the ordering of the variables
3. (a) Choose a variable X ~ Add a node
   Independent
   (b) Pick another variable and Add parent (set) of the variable
   (c) Conditional Probability table (CPT)



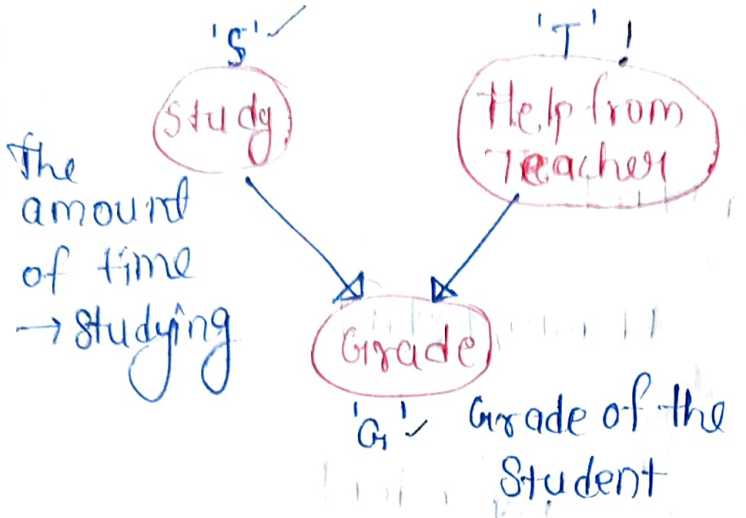
Problem statement:

Variable: Cyclone predicition
→ Sea temp, wind, temperature
→ likelihood of cyclone

Dataset
1. Domain
2. ML
Data → Pattern
Modelity

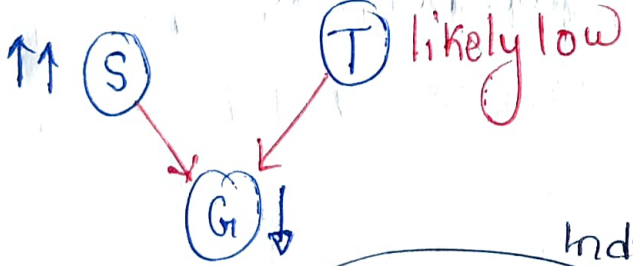
$$P(Y|X) \leftarrow \text{Statistics} = \frac{P(X\ Y)}{P(X)}$$

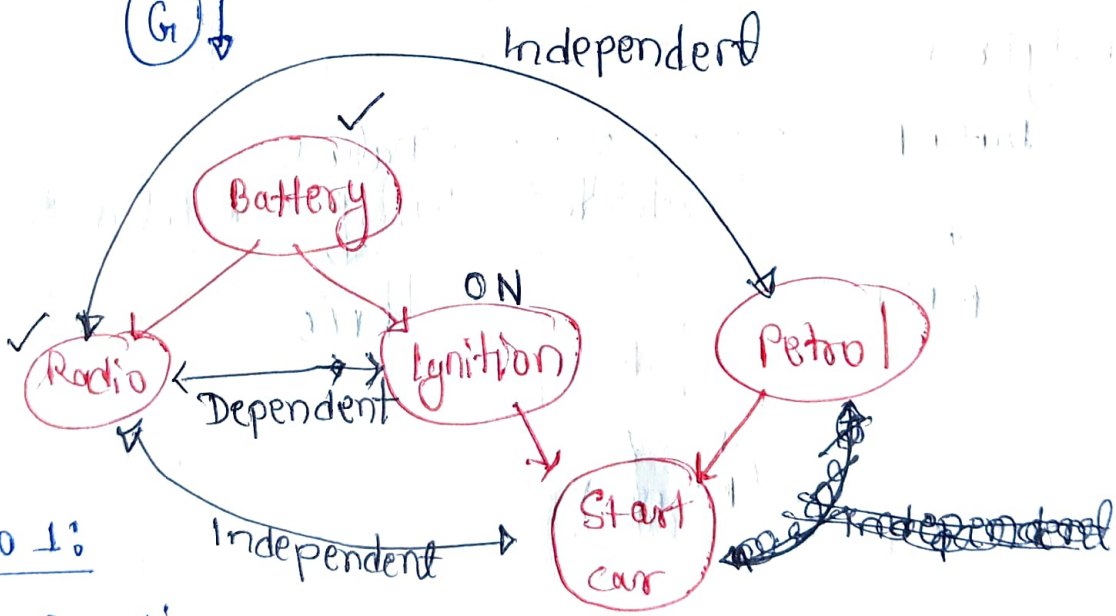
$$\frac{\#\text{ sample } X \text{ \& } Y}{\#\text{ sample}}$$



'S' and 'T' remains independent unless we know 'G'

We know that Grade is 'High', spent less amount of time studying, then the likelihood that help from teacher is 'high'.





Scenario 1:

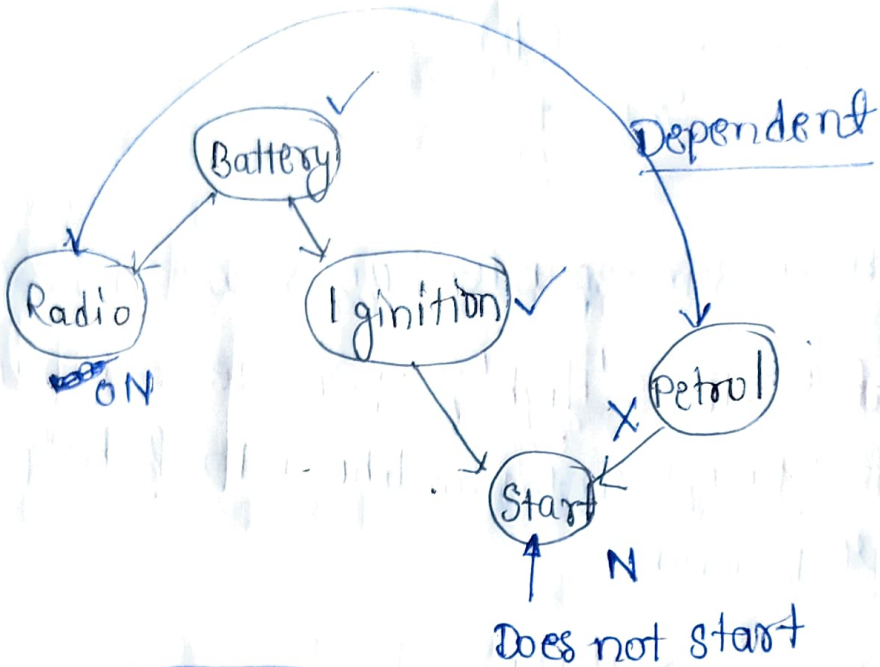
Weather

Petrol & Radio are independent or not given about ignition takes place or not.

Scenario 2:

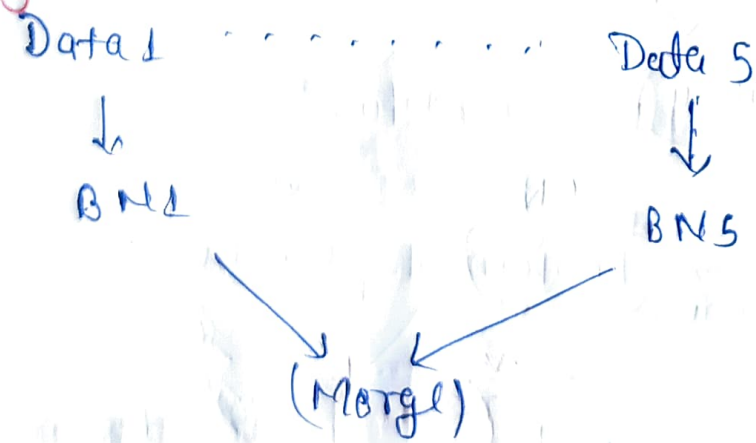
Evidence: Car Starts

↳ Radio, Petrol



Case 4

Polytree



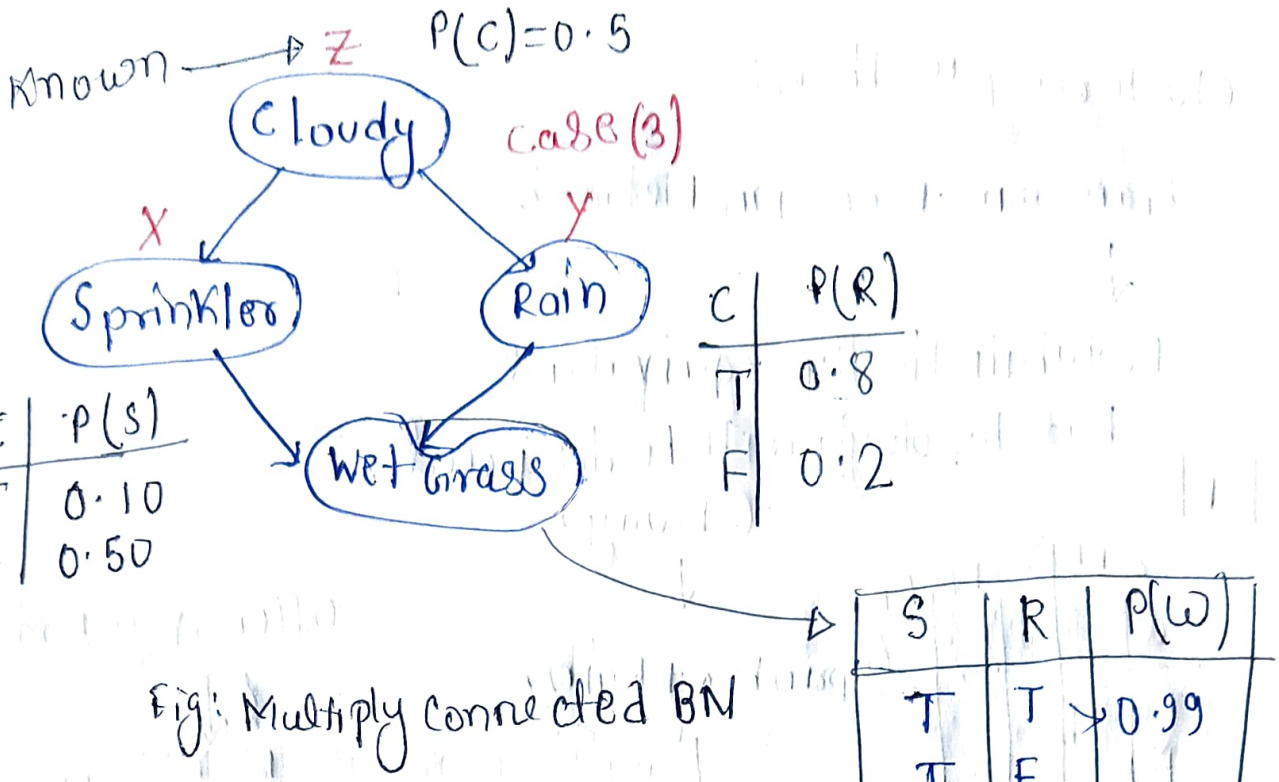
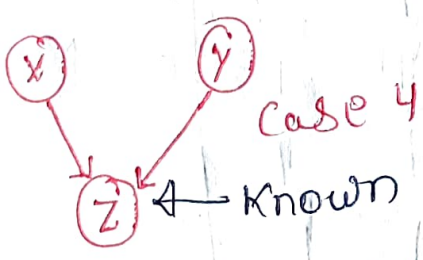


Fig: Multiply connected BN

1. S, R are independent



2. S, R are not independent.

Methods to resolve the above problem

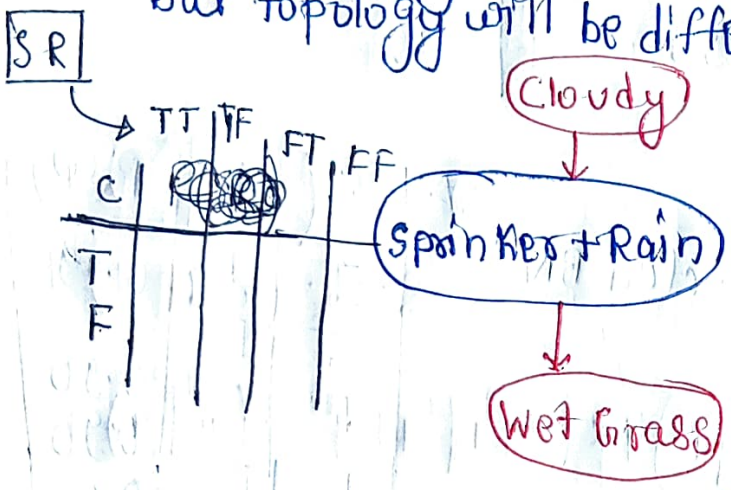
1. → Clustering method
2. → cutset conditioning method

Clustering method:

Transform Bayesian Network

↓

Probabilistically equivalent but topology will be different



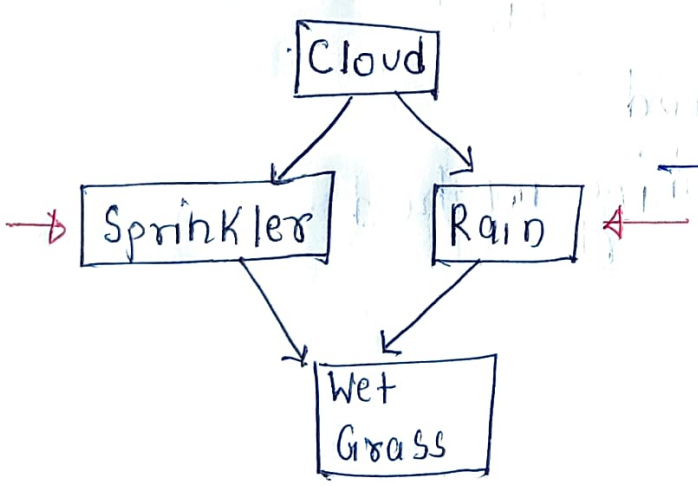
Offending node

↓

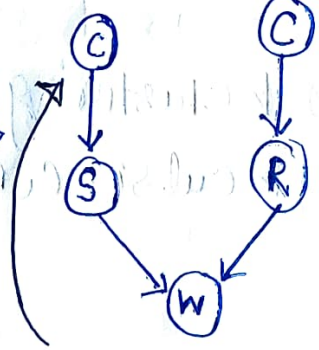
Merge

	S	R	P(W)
	T	T	
	T	F	
	F	T	
	F	F	

Cutset Conditioning Method:

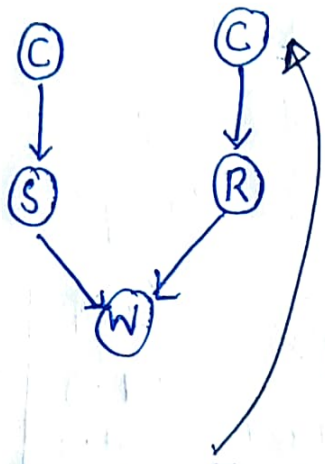


1.



For C = True

2.



For C = True

Simpson's Paradox

Males	Recovered	Not recovered	Recovery rate
Given	18	12	60%
Not Given	7	3	70%

Females	Recovered	Not Recovered	Recovery rate
Given	2	8	20%
Not Given	9	21	30%

Combined	Recover	Not Recover	Recovery Rate
Given	20	20	50%
Not Given	16	24	40%

Exp 1

Male
40

Exp 2

Female
40

# of Sample size
# of patients

$$P(R|MG) = 0.6$$
$$P(R|FG) = 0.2$$

$$P(R|G) = P(R|MG) \times P(G|M) + P(R|FG) \times P(G|F)$$

$$= 0.6 \times \frac{30}{40} + 0.2 \times \frac{10}{40}$$

$$= 0.6 \times 0.75 + 0.2 \times 0.25$$

$$= 0.45 + 0.05$$

$$= \boxed{0.5} \text{ Ans.}$$

Bias is coming from here

Statement

40% Indians 'Tall'

$P(\text{Tall}) = 0.4$

P(Event)

$0 \leq P \leq 1$

→ How Tall the person is ? ← Fuzzy Logic

We want to find out!

▷ What is the truth value of an event ?

How well an object satisfy an abstract description ?

Fuzzy rules

Graph Conversion

1. Diet is low and exercise is high → Balanced
2. -11- high or -11- low → Unbalanced
3. Balanced → Risk is low
4. Unbalanced → Risk is high

[Cardiac Health Management]

Diet = 3000 cal/day

Exercise = Burning 1000 cal/day

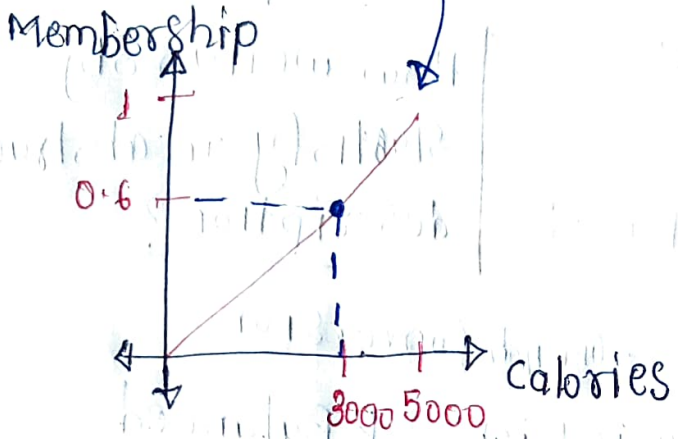
Q: What is the risk of heart disease ?

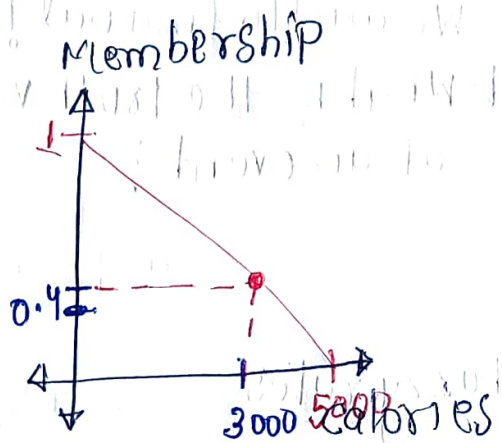
A: (Unbalanced) → (High Risk) ?

Membership function

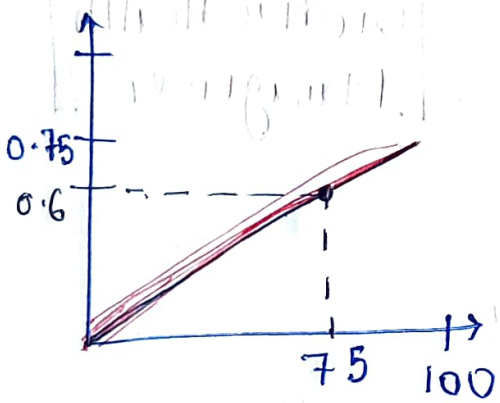
↳ connects rules $\{1, 2, 3, 4\}$ to some values

$$f_{diethigh}(x) = \frac{1}{5000}x$$





$$f_{dietlow}(x) = 1 - \frac{1}{5000}x$$



$$f_{riskhigh}(x) = \frac{1}{125}x$$

$$0.6 = \frac{1}{125}x$$

$$\Rightarrow \boxed{x = 75}$$