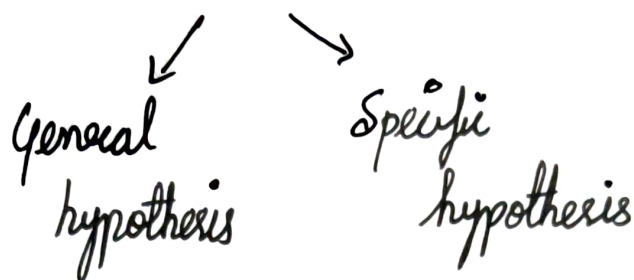


## Concept learning:-

Concept learning is a learning task in which we train our machine to learn some concept by giving some pre-defined example.



### General hypothesis :-

$$G = \{ '?', '?', '?' \}$$

→ Represented by '?'

→ No. of '?' depends on No. of attributes

### Specific hypothesis :-

$$S = \{ '\phi', '\phi', '\phi' \}$$

→ Represented by ' $\phi$ '

→ No. of ' $\phi$ ' depends on No. of attributes.

Let's take a scenario for understanding General & Specific hypothesis.

↳ If you ask your friend to order Pizza without specifying ~~any~~ toppings  $\Rightarrow$  General hypothesis

↳ If you ask your friend to order Pizza with specified toppings  $\Rightarrow$  Specific hypothesis

In, General hypothesis he may choose any toppings.

Find - S algorithm :-

↳ The name itself suggests that algorithm is to find "specific hypothesis".

↳ Also, it will consider only "positive example".

Algorithm :-

1) Initialize hypothesis 'h<sub>0</sub>'

(i.e.,)  $h_0 = \{ \phi, \phi, \phi \}$

2) For each '+ve' example :

For each 'attribute' in example :

if attribute\_value == hypothesis\_values :

Do nothing

else :

Replace hypothesis value with '?'  
(General hypothesis)

Ques:- Days on which person enjoy Sport

	Attribute						Label
	Sky	Temp	Humidity	Wind	Water	Forecast	Enjoy
①	Sunny	Warm	Normal	Strong	Warm	Same	Yes
②	Sunny	Warm	High	Strong	Warm	Same	Yes
③	Rainy	Cold	High	Strong	Warm	Change	No
④	Sunny	Warm	High	Strong	Cool	Same	Yes

Attribute = 6

Label = Enjoy

Step - 1 :-

Initialize hypothesis 'h<sub>0</sub>'

Considering first '+ve' instance as 'h<sub>0</sub>'

h<sub>0</sub> = { 'Sunny', 'Warm', 'Normal', 'Strong',  
'Warm', 'Same' } ①

Step -2 :-

⇒ Comparing with 'tree' instance

→ 'ho' with second instance, (2)

'ho' becomes,

ho = { 'sunny', 'warm', '?', 'strong',

'warm', 'same' }

→ 'ho' can't compare with third instance  
because it is negative (3)

→ 'ho' comparing with fourth instance  
it becomes,

ho = { 'sunny', 'warm', '?', 'strong',

'?', 'same' }

( most general hypothesis )



Disadvantage:-

- consider only 'tve' values
- it is not guarantee that it will match all the data because we ignored '-ve' samples. (most general hypothesis)