

# Theoretical Concepts and Important Terminologies in ML - I

Dr. Muhammad Wasim

# Example 1: Learning the SICK Concept



# Concept Instance Space (X)

<b>X</b>	<b>T</b>	<b>BP</b>	<b>Sick (SK)</b>
x <sub>1</sub>	L	L	-
x <sub>2</sub>	L	N	-
x <sub>3</sub>	L	H	-
x <sub>4</sub>	N	L	-
x <sub>5</sub>	N	N	-
x <sub>6</sub>	N	H	-
x <sub>7</sub>	H	L	-
x <sub>8</sub>	H	N	-
x <sub>9</sub>	H	H	-

# Concept as a function

- The solution to any problem is a function that converts its inputs to corresponding outputs.
- A concept itself is merely a function, which we don't know yet.

# Concept Space

<b>X</b>	<b>T</b>	<b>BP</b>	<b>Sick (SK)</b>
$x_1$	L	L	0
$x_2$	L	N	0
$x_3$	L	H	1
$x_4$	N	L	0
$x_5$	N	N	0
$x_6$	N	H	1
$x_7$	H	L	1
$x_8$	H	N	1
$x_9$	H	H	1

# Concept Space

- But there are a lot of other possibilities besides this one.
- The question is: how many total concepts can be generated out of this given situation.
- The answer is:  $2^{|X|}$   
Here  $2^9$ , since  $|X| = 9$
- In short the true concept SICK is a function defined over the attributes T and BP, such that it gives a 0 or a 1 as output for each of the 9 instances  $x_i$  belonging to Instance Space X

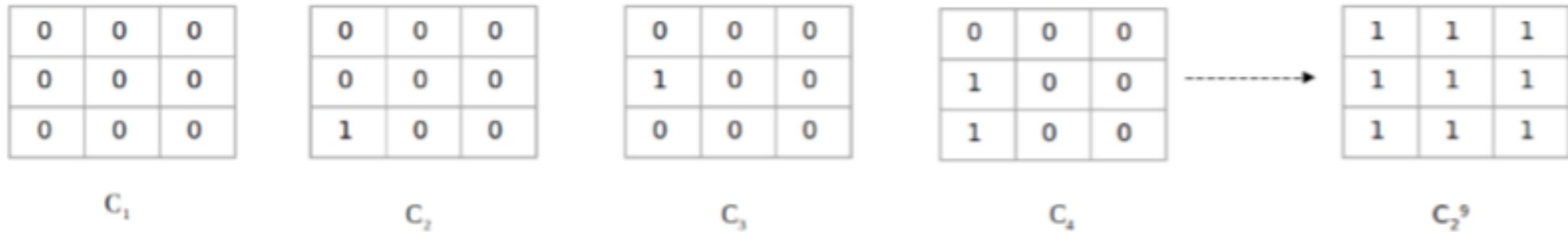
# Concept Space

- In short the true concept SK is a function defined over the attributes T and BP, such that it gives a 0 or a 1 as output for each of the 9 instances  $x_i$  belonging to Instance Space X

$c(x_3)$	$c(x_6)$	$c(x_9)$
$c(x_2)$	$c(x_5)$	$c(x_8)$
$c(x_1)$	$c(x_4)$	$c(x_7)$

# Concept Space (C)

- Since we don't know the true concept yet, so there might be concepts which can produce  $2^9$  different outputs, such as:





# So, What is a Concept?

- Concept is nothing more than a function whose independent variables are the attributes, in this case T and BP
- Maybe the true concept is some complicated arrangement of conjunctions and disjunctions like:

$C = < T = H \text{ AND } BP = H$

OR

$T = N \text{ AND } BP = H$

OR

$T = H \text{ AND } BP = N >$

# Hypothesis Space (H)

- The learner has to apply some hypothesis, that introduces a search bias to reduce the size of the concept space
- This reduced concept space becomes the hypothesis space.  
For example, the most common bias is one that uses the AND relationship between the attributes.
- In other words, the hypothesis space uses the conjunctions (AND) of the attributes T and BP  
i.e.  $h = \langle T, BP \rangle$

# Hypothesis Space

- H denotes the hypothesis space
- Here it is the conjunction of attributes T and BP  
If written in English, it would mean:  
H = < t, bp >: IF “temperature” = t AND “Blood Pressure” = bp  
Then  
H = 1 Otherwise H = 0
- In other words, the function gives a 1 output for all conjunctions of T and BP, e.g., H and H, H and L, H and N, etc.

# Hypothesis Space

- $h = \langle H, H \rangle : \langle \text{temp}, \text{bp} \rangle$

BP				
H	0	0	1	
N	0	0	0	
L	0	0.	0	
	L	N	H	<b>T</b>