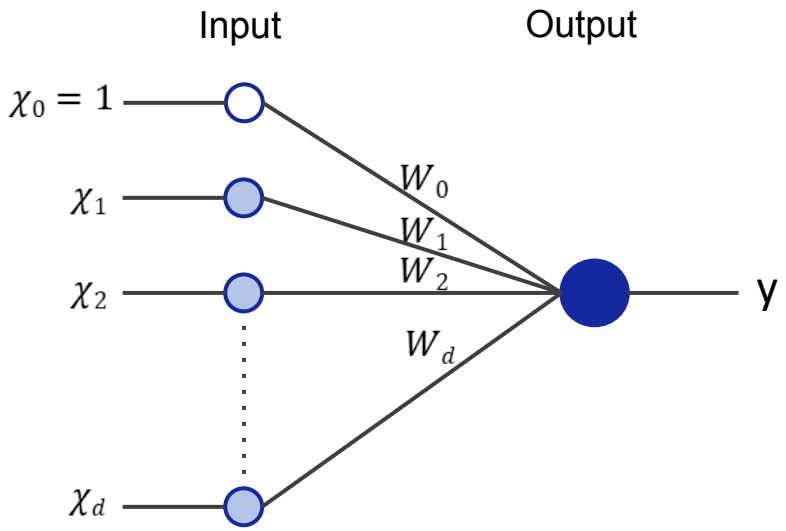
Perceptron

The Structure of a Perceptron



Structure of a perceptron

- Comprehend the structure of a perceptron by solving the OR function of the truth table.

Perceptron

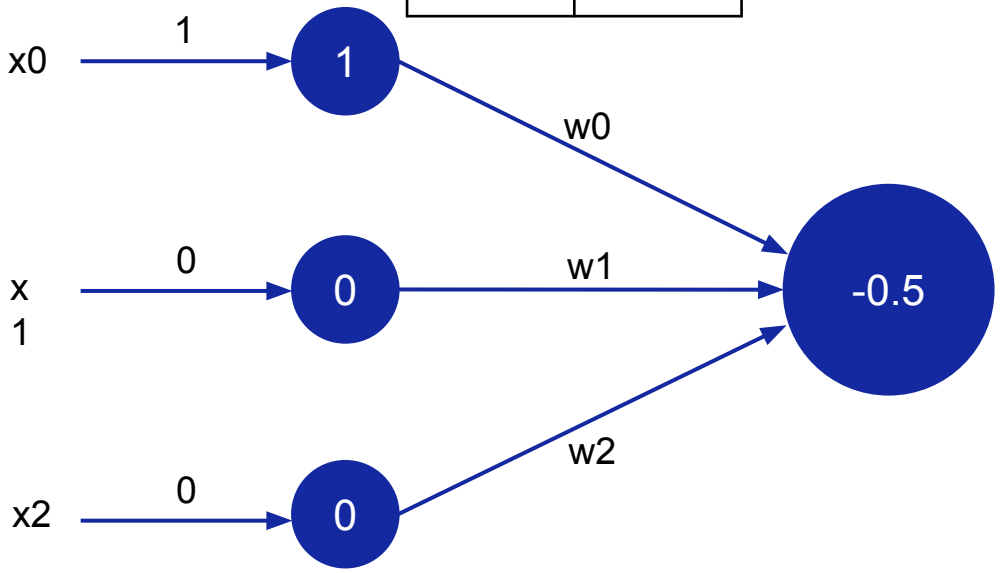
- First, you need to understand the OR operation of the truth table. The results of the OR operation are below.

X1	X2	OR operation
False	False	F
True	False	T
False	True	T
True	True	T

- What are the weight values to solve the OR operation problem?

Perceptron

- Suppose that we know the ‘w’ values to solve a problem. w0 is -0,5, w1 is 1, w2 is 1.
  (w: a weight vector)
- Execute Excel function after substituting x1 with 0, and x2 with 0.



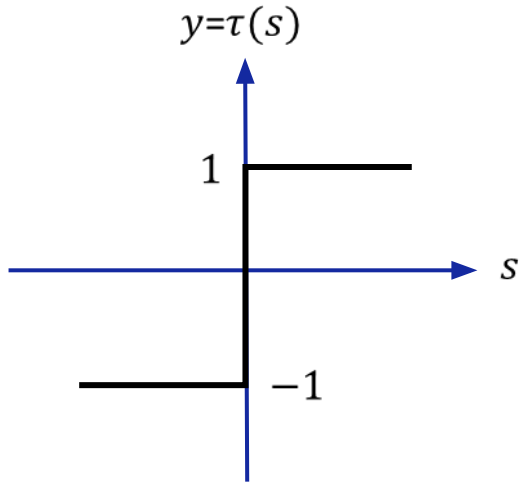
x1	X2
0	0

w0	w1	w2
-0.5	1	1

x0 * w0	-0.5
+	
x0 * w1	0
+	
x0 * w2	0
Sum	-0.5

Perceptron

- Put the value of the sum into a function.
- Here, that function is the step function.



Use threshold function as an activation function τ(s)

Perceptron

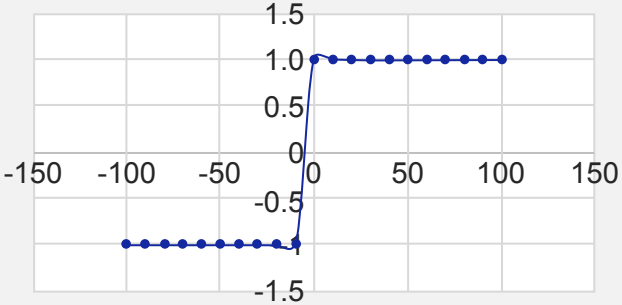
- A threshold function can be defined as follows, and we can create a graph by putting random values in the Excel table.

A threshold can be defined as follows,

$$y = \tau(s)$$

here $s = w_0 + \sum_{i=1}^{d} w_i x_i$, $\tau(s) = \begin{cases} 1 & s \geq 0 \\ -1 & s < 0 \end{cases}$

-100	-1
-90	-1
-80	-1
-70	-1
-60	-1
-50	-1
-40	-1
-30	-1
-20	-1
-10	-1
0	1
10	1
20	1
30	1
40	1
50	1
60	1
70	1
80	1
90	1
100	1



Perceptron

- Put the result form the previous slide, -0.5 as s, and run it through the threshold function, the return is -1.
- Let's learn about the truth table.

A truth table displays truth or false for all results of the propositions or the combination of their Boolean functions.

For example, in case of the conjunction of two statements P and Q, $P \wedge Q$, the truth table can be constructed as below.

In addition, true·false is also notated as T·F or 1·0.

Proposition P	Proposition Q	
True	True	True
True	False	False
False	True	False
False	False	False

Perceptron

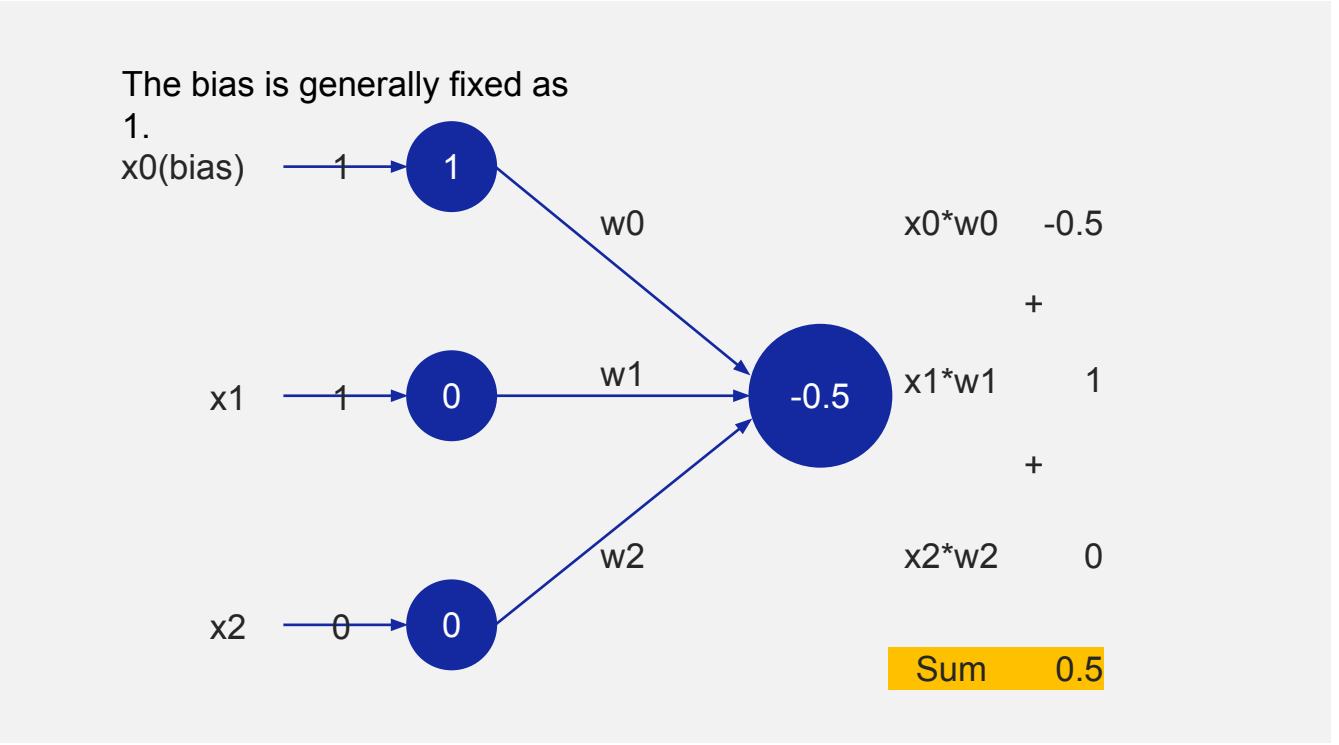
- A truth table can be expressed with 1 and 0 as follows.

x1	x2
0	0
1	0
0	1
1	1

- The figure from Slide 9 put 0 for the first variable x1, and 0 for x2 as well.
- x0 is called a bias, and it initializes as 1.

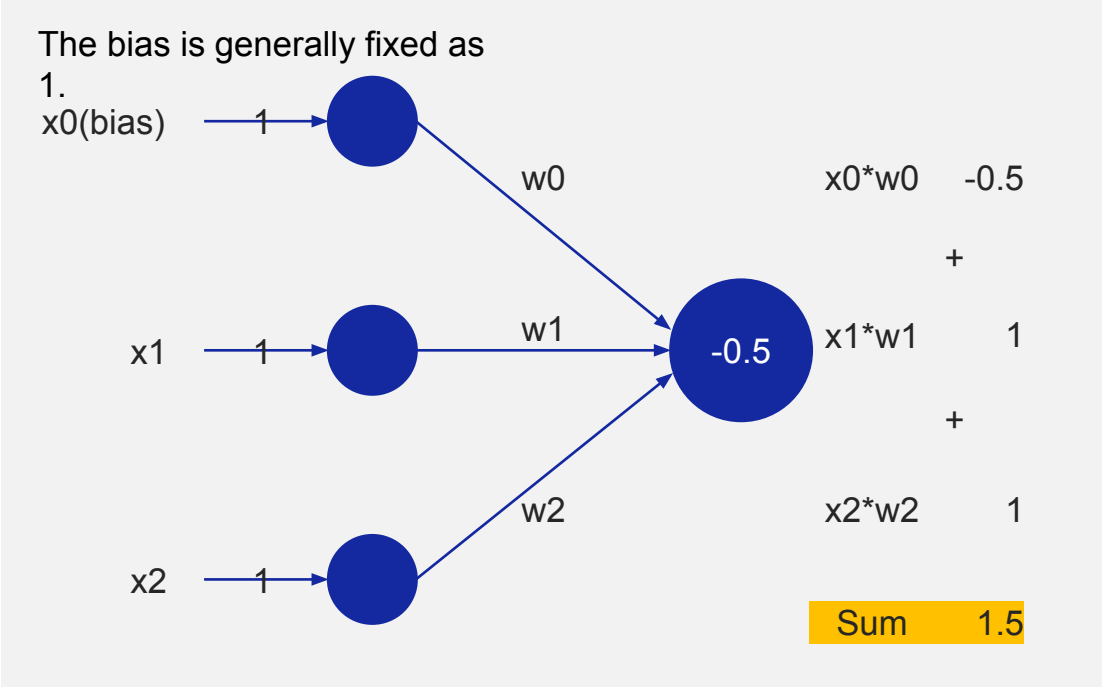
Perceptron

- Substitute x1 with 1, and x2 with 0.



Perceptron

- The sum is 0.5. Run this through the threshold function, and the result is 1.
- Likewise, if we process all values in the truth table through the perceptron, the results are as follows.

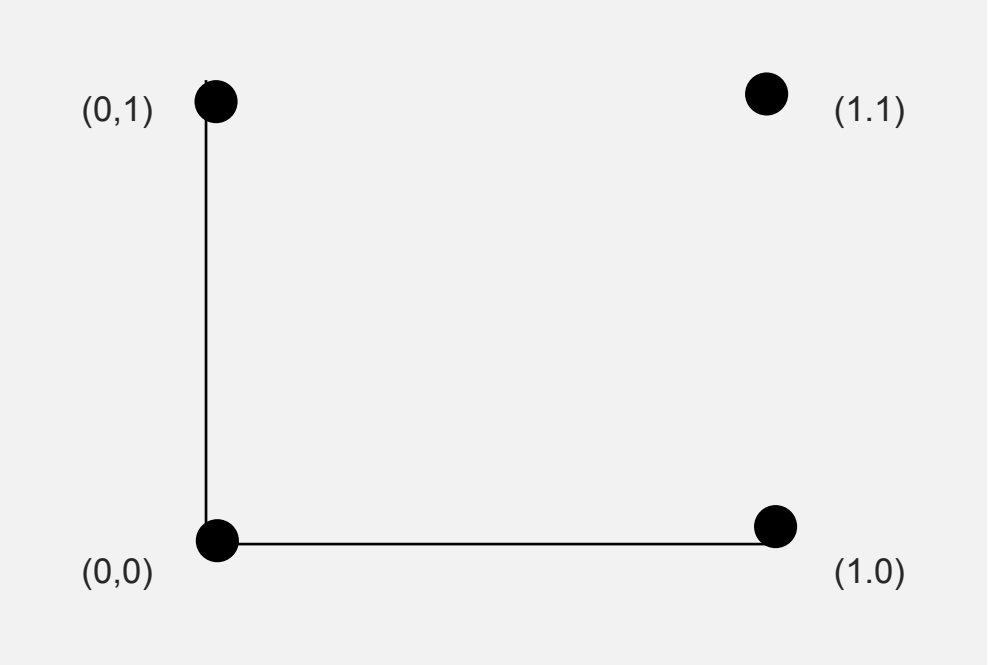


- 4 results can be displayed in a table.

x1	x2		
0	0	-0.5	-1
1	0	0.5	1
0	1	0.5	1
1	1	1.5	1

Perceptron

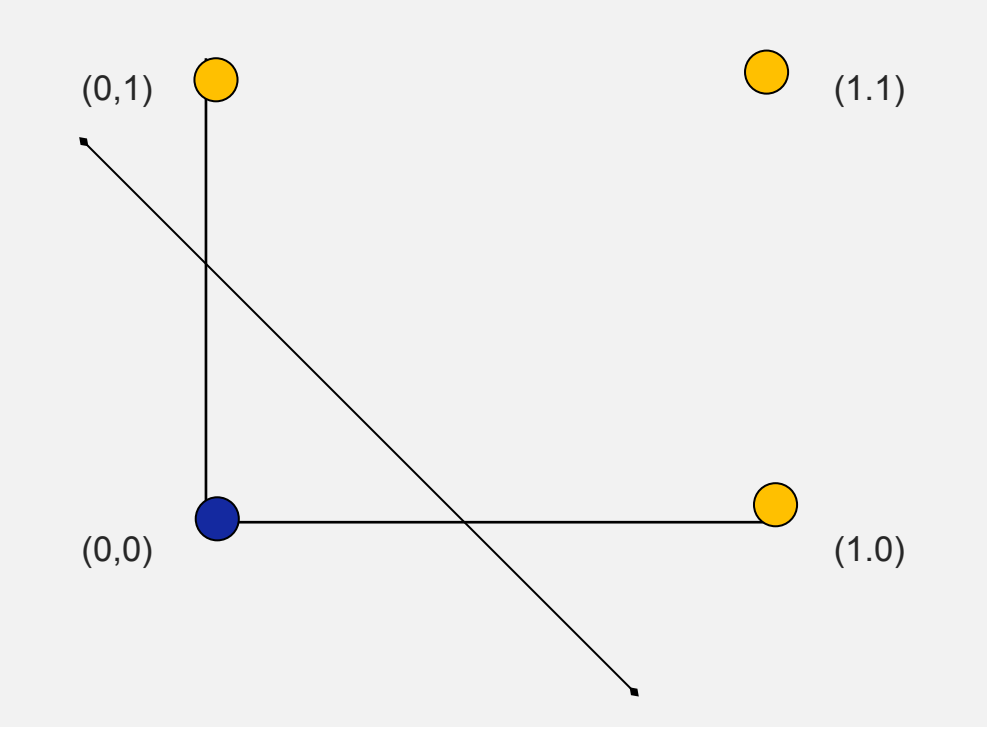
- These results can be shown in the coordinate plane as follows.
- The results can be grouped into that of point (0,0) or the rest.



x1	x2		
0	0	-0.5	-1
1	0	0.5	1
0	1	0.5	1
1	1	1.5	1

Perceptron

- Such classification can be presented as a geometrical figure as follows.
- If you recall the basic notion of machine learning through linear regression, machine learning is a process that builds regression equation with the predicted values of the slope and y-intercept. The figure below shows how linear regression forms a line that separate the two groups.



x1	x2		
0	0	-0.5	-1
1	0	0.5	1
0	1	0.5	1
1	1	1.5	1

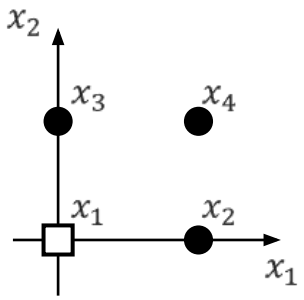
- This table displays the OR operation of a truth table.

OR operation

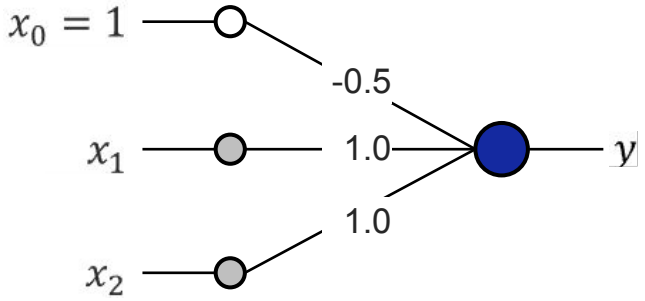
x1	x1	x2	Value before running threshold function	Value after running threshold function
0	0	0	=$A$2*$C$7+B2*$C$8+C2*$C$9	=IF(D2>=0,1,-1)
	0	1	=$A$2*$C$7+B3*$C$8+C3*$C$9	=IF(D3>=0,1,-1)
	1	0	=$A$2*$C$7+B4*$C$8+C4*$C$9	=IF(D4>=0,1,-1)
	1	1	=$A$2*$C$7+B8*$C$8+C5*$C$9	=IF(D5>=0,1,-1)

w1	-0.5
w2	1
w3	1

$$x_1 = \begin{pmatrix}0\\0\end{pmatrix}, y_1 = -1, x_2 = \begin{pmatrix}1\\0\end{pmatrix}, y_2 = 1, x_3 = \begin{pmatrix}0\\1\end{pmatrix}, y_3 = 1, x_4 = \begin{pmatrix}1\\1\end{pmatrix}, y_4 = 1$$



(a) Train set



(b) Perceptron

Example of a perceptron's operation using the OR logic gate

OR operation

Let’s input four samples to the perceptron and check the results.

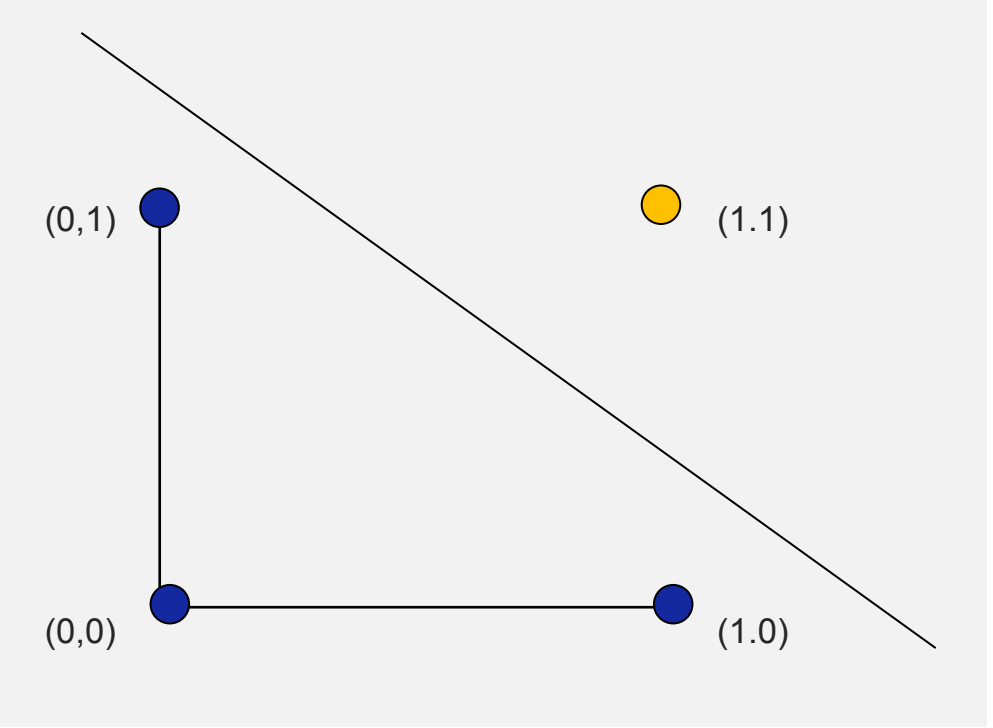
$$x_1: s = -0.5 + 0 * 1.0 + 0 * 1.0 = -0.5, \quad \tau(-0.5) = -1$$
$$x_2: s = -0.5 + 1 * 1.0 + 0 * 1.0 = 0.5, \quad \tau(0.5) = 1$$
$$x_3: s = -0.5 + 0 * 1.0 + 1 * 1.0 = 0.5, \quad \tau(0.5) = 1$$
$$x_4: s = -0.5 + 1 * 1.0 + 1 * 1.0 = 1.5, \quad \tau(1.5) = 1$$

- As you’ve seen in the previous slide, the perceptron delivered correct results for all four samples.
- It can be said that this perceptron classifies the train set with 100 % performance.

AND operation

- To recap, you can solve the OR problem by perceptron with the appropriate w values.
- Now, the AND operation. The truth table is on the left, and the geometrical solution is the line that separates the units into two groups.

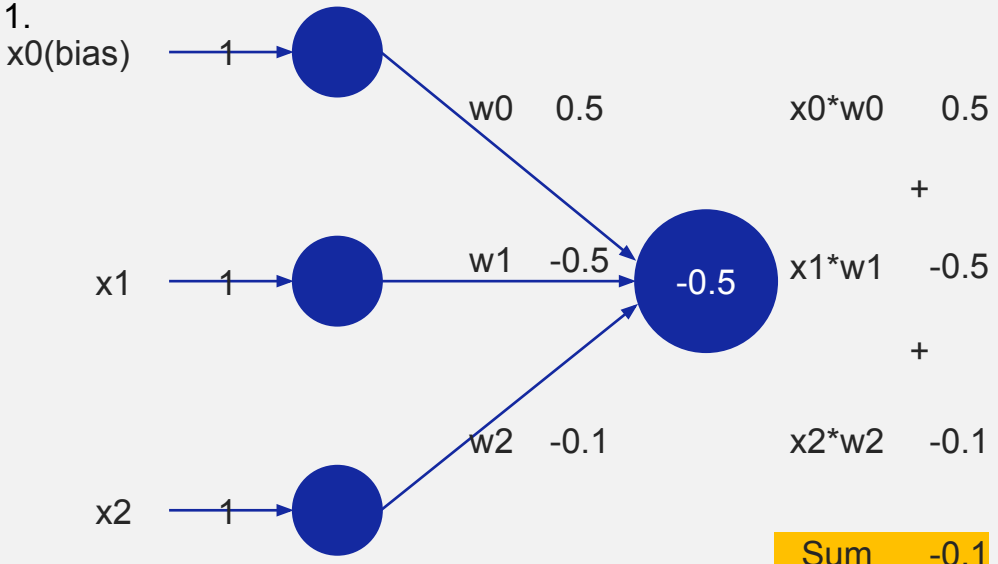
x1	x2	AND operation
0	0	F
1	0	F
0	1	F
1	1	T



AND operation

- Find the values of w0, w1, w2 that return the result from the previous slide.
- First, let's apply a random value, and then gradually change the values of w0, w1, and w2 to find the solution.

The bias is generally fixed as 1.

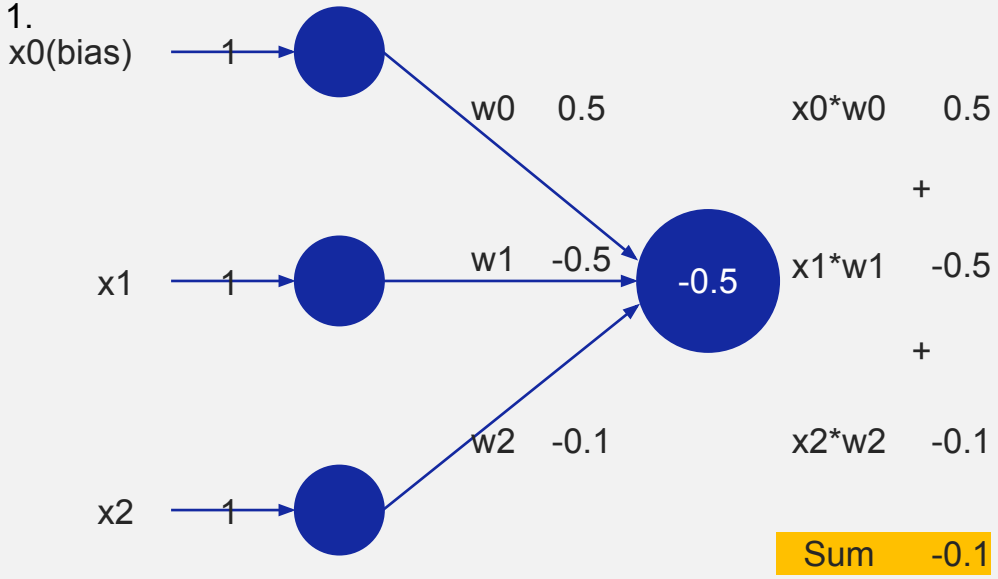


- 4 results can be displayed in a table.

x1	x2		
0	0	0.5	1
1	0	0	1
0	1	0.4	1
1	1	-0.1	-1

- AND operation of a truth table.

AND operation



- 4 results can be displayed in a table.

x1	x2		
0	0	0.5	1
1	0	0	1
0	1	0.4	1
1	1	-0.1	-1

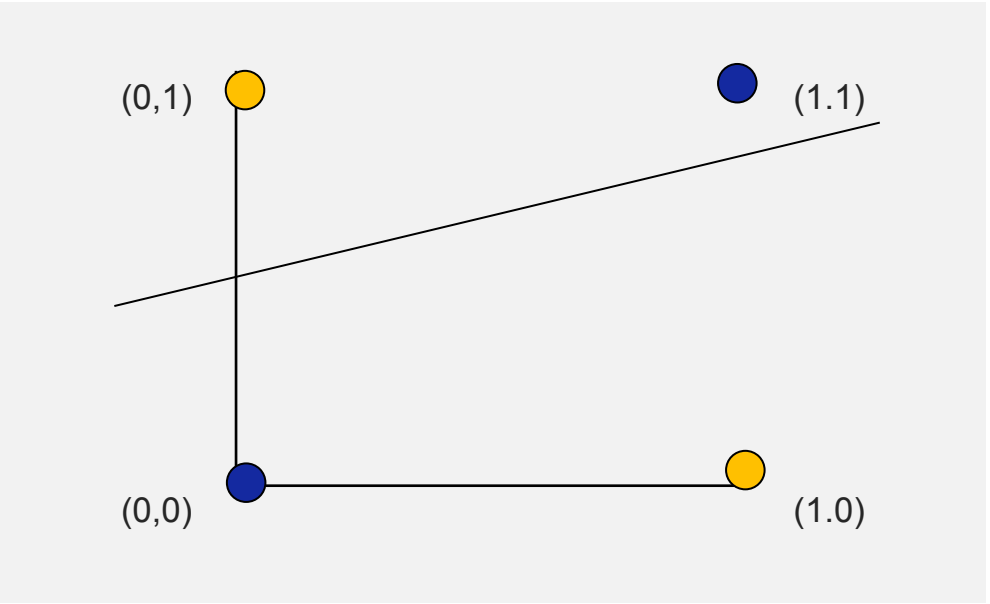
- AND operation of a truth table.

- By substituting w0 with 0.5, w1 with -0.5, w2 with -0.1, the equation solves the AND operation.
- You found a structure that with appropriate values for the w vector, solves a certain problem.
- Although the perceptron solved the OR and AND problem, it could not solve the XOR. Let's find out why.

XOR Operation

- XOR operation in a truth table.
- In XOR, the result is true if two propositions are the opposite.

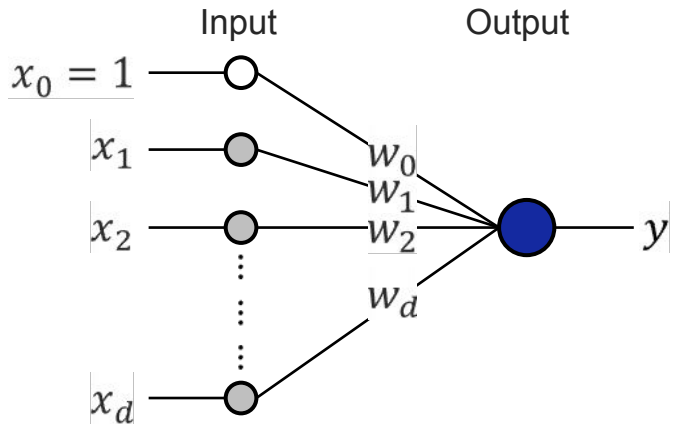
x1	x2	XOR Operation
0	0	F
1	0	T
0	1	T
1	1	F



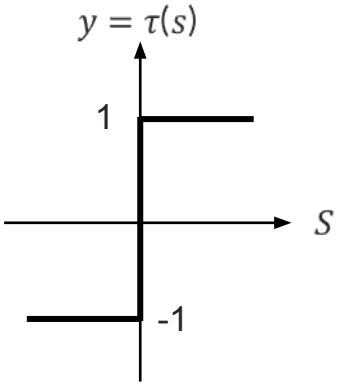
- Since the perceptron is a linear classifier, you cannot find a line that separate the blue dots and the red dots.
- A multilayer perceptron solved the problem.

Summary of Perceptron's Structure

- It has an input layer and an output layer.
- The input layer does not operate, so the perceptron is considered a single layer structure.
- The i th node of the input layer takes x_i from the feature vector $x = (x_1, x_2, \cdots, x_d)^{\mathrm{T}}$.
- The bias node always takes 1 as input.
- The output layer has a single node.
- The connection of the i th node of the input later and the output layer has weight w_i.



(a) Structure of a perceptron



(b) Use threshold function as an activation function τ(s)

Mechanism of a perceptron