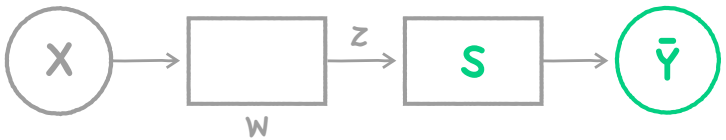
LECTURE 9-1

NEURAL NETS(NN) FOR XOR

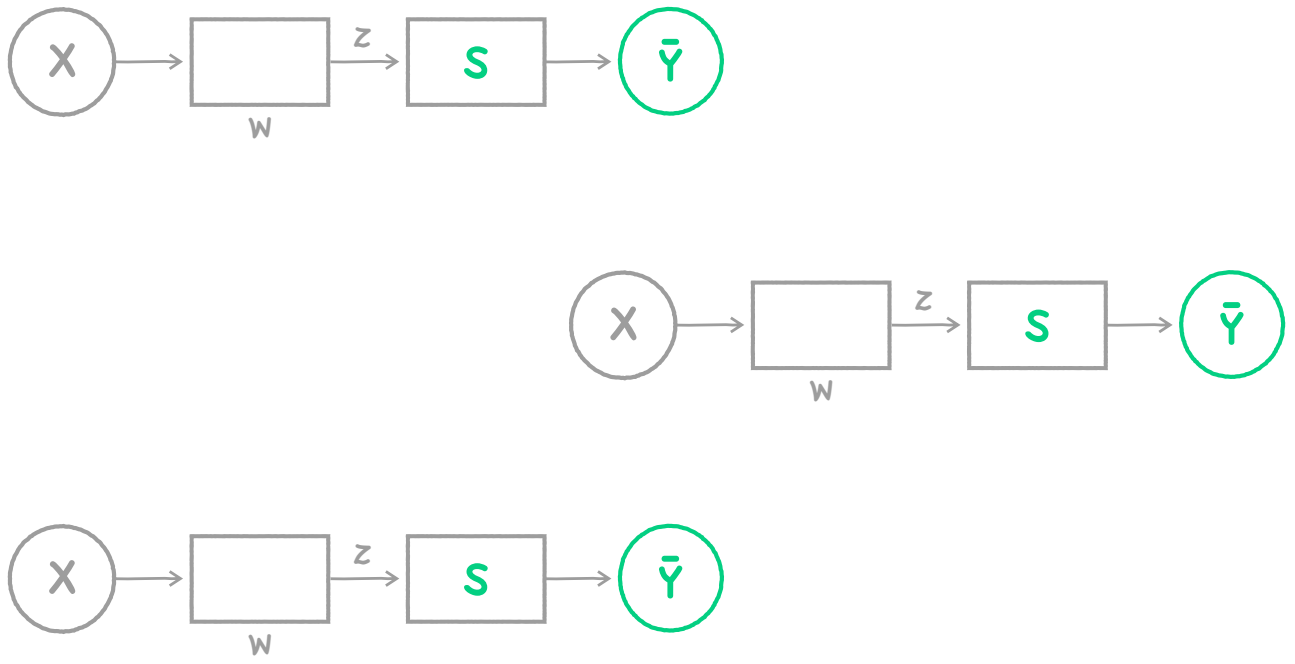
Sung Kim <hunkim+ml@gmail.com>
http://hunkim.github.io/ml

NAVER | Clova

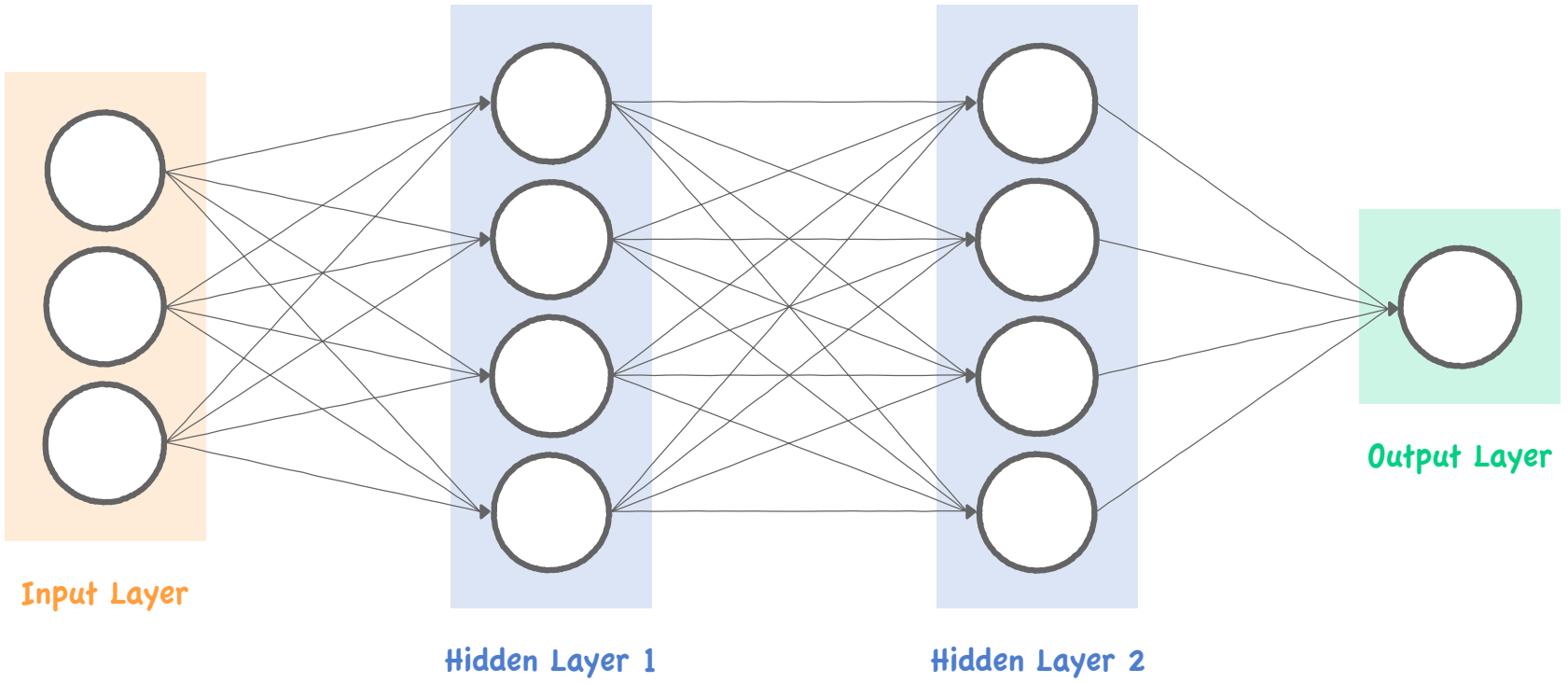
One Logistic Regression Unit Cannot Separate XOR



Multiple Logistic Regression Units



Neural Network(NN)



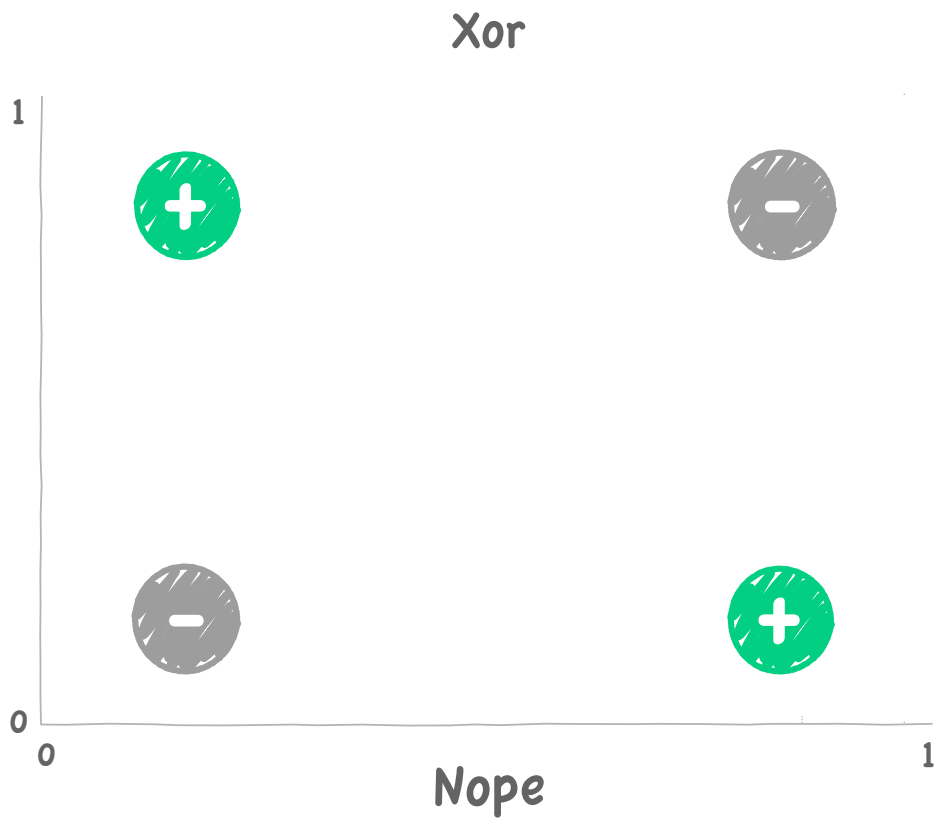
"No one on earth had found a viable way to train"

<Marvin Minsky>

http://cs231n.github.io/convolutional-networks)

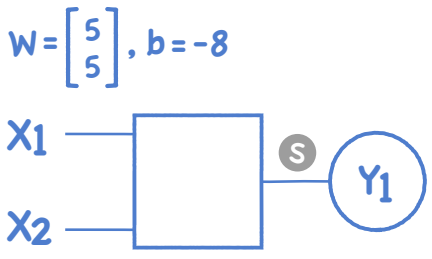
XOR Using NN

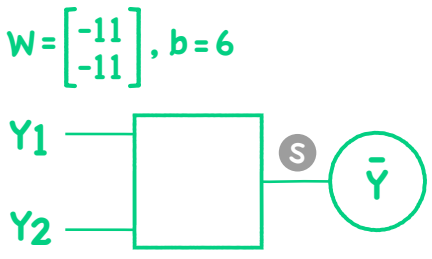
X_1	X_2	XOR
0	0	0(-)
0	1	1(+)
1	0	1(+)
1	1	0(-)

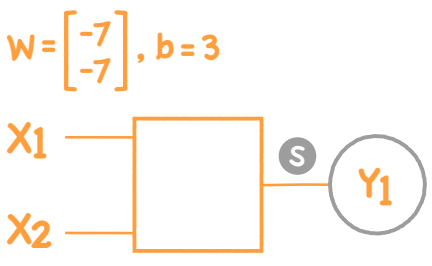


Neural Net

Neural Net

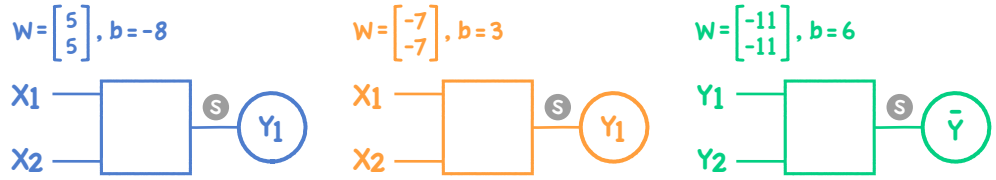






X_1	X_2	Y_1	Y_2	$\bar{Y}$	XOR
0	0				0
0	1				1
1	0				1
1	1				0

Neural Net

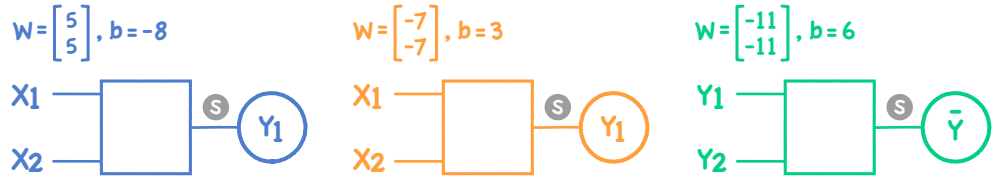


$$[0\ 0]\begin{bmatrix}5\\5\end{bmatrix}-8=$$

$$[0\ 0]\begin{bmatrix}-7\\-7\end{bmatrix}+3=$$

X_1	X_2	Y_1	Y_2	$\bar{Y}$	XOR
0	0				0
0	1				1
1	0				1
1	1				0

Neural Net



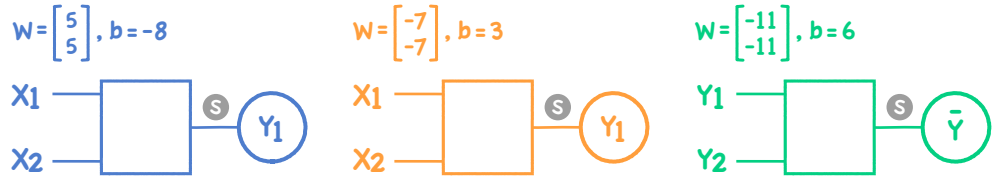
$$[0\ 0]\begin{bmatrix}5\\5\end{bmatrix} -8 = 0+0-8 = -8, \text{Sigmoid}(-8) = 0$$

$$[0\ 0]\begin{bmatrix}-7\\-7\end{bmatrix} +3 = 0+0+3 = 3, \text{Sigmoid}(3) = 1$$

$$[0\ 1]\begin{bmatrix}-11\\-11\end{bmatrix} +6 = 0+(-11)+6 = -5$$
$$\text{Sigmoid}(-5) = 0$$

X_1	X_2	Y_1	Y_2	$\bar{Y}$	XOR
0	0	0	1	0	0
0	1				1
1	0				1
1	1				0

Neural Net



$$[0\ 1]\begin{bmatrix}5\\5\end{bmatrix} -8 = 0 + 5 - 8 = -3, \text{Sigmoid}(-3) = 0$$

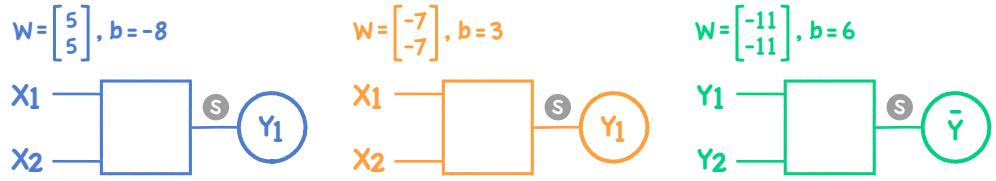
$$[0\ 1]\begin{bmatrix}-7\\-7\end{bmatrix} +3 = 0 + (-7) + 3 = -4, \text{Sigmoid}(-4) = 0$$

$$[0\ 0]\begin{bmatrix}-11\\-11\end{bmatrix} +6 = 0 + 0 + 6 = 6$$

Sigmoid(6) = 1

X_1	X_2	Y_1	Y_2	$\bar{Y}$	XOR
0	0	0	1	0	0
0	1	0	0	1	1
1	0				1
1	1				0

Neural Net



$$[1\ 0]\begin{bmatrix}5\\5\end{bmatrix} - 8 = 5 + 0 - 8 = -3, \text{Sigmoid}(-3) = 0$$

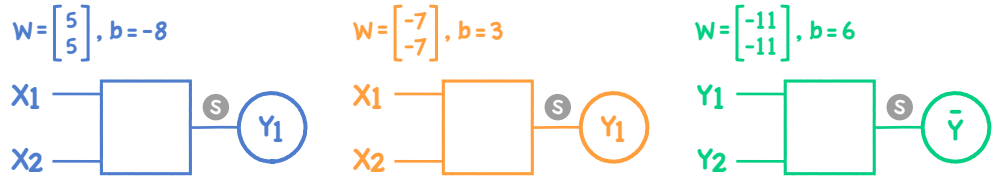
$$[1\ 0]\begin{bmatrix}-7\\-7\end{bmatrix} + 3 = (-7) + 0 + 3 = -4, \text{Sigmoid}(-4) = 0$$

$$[0\ 0]\begin{bmatrix}-11\\-11\end{bmatrix} + 6 = 0 + 0 + 6 = 6$$

$$\text{Sigmoid}(6) = 1$$

X_1	X_2	Y_1	Y_2	$\bar{Y}$	XOR
0	0	0	1	0	0
0	1	0	0	1	1
1	0	0	0	1	1
1	1				0

Neural Net



$$\begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}5\\5\end{bmatrix} - 8 = 5 + 5 - 8 = 2, \text{Sigmoid}(2) = 1$$

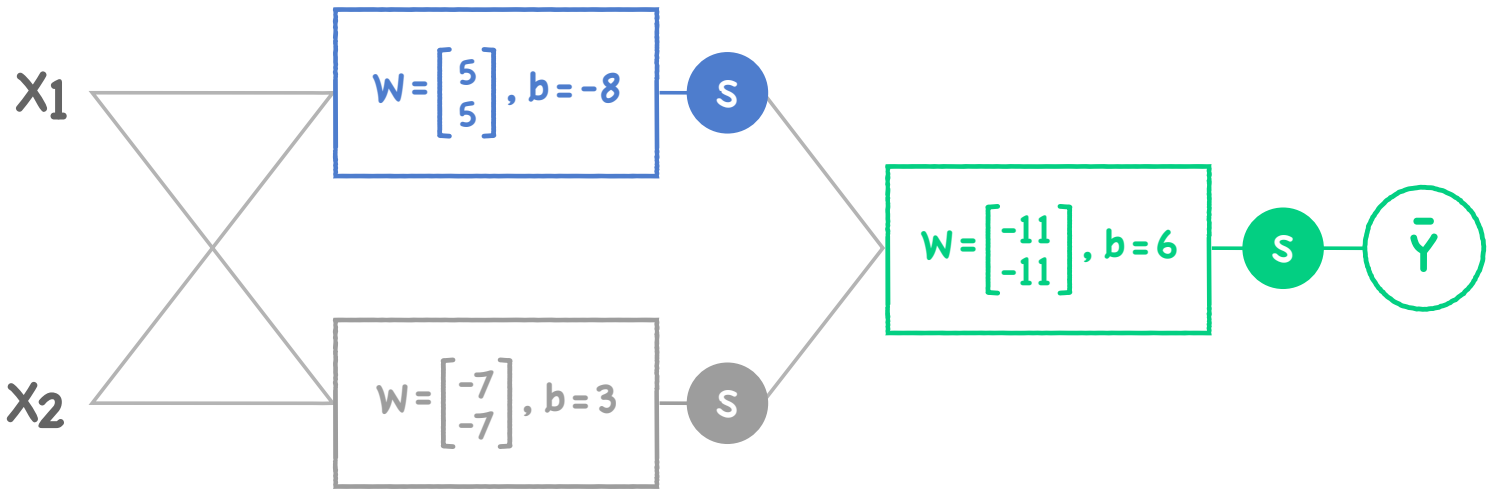
$$\begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}-7\\-7\end{bmatrix} + 3 = (-7) + (-7) + 3 = -11, \text{Sigmoid}(-11) = 0$$

$$\begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}-11\\-11\end{bmatrix} + 6 = (-11) + 0 + 6 = -5$$

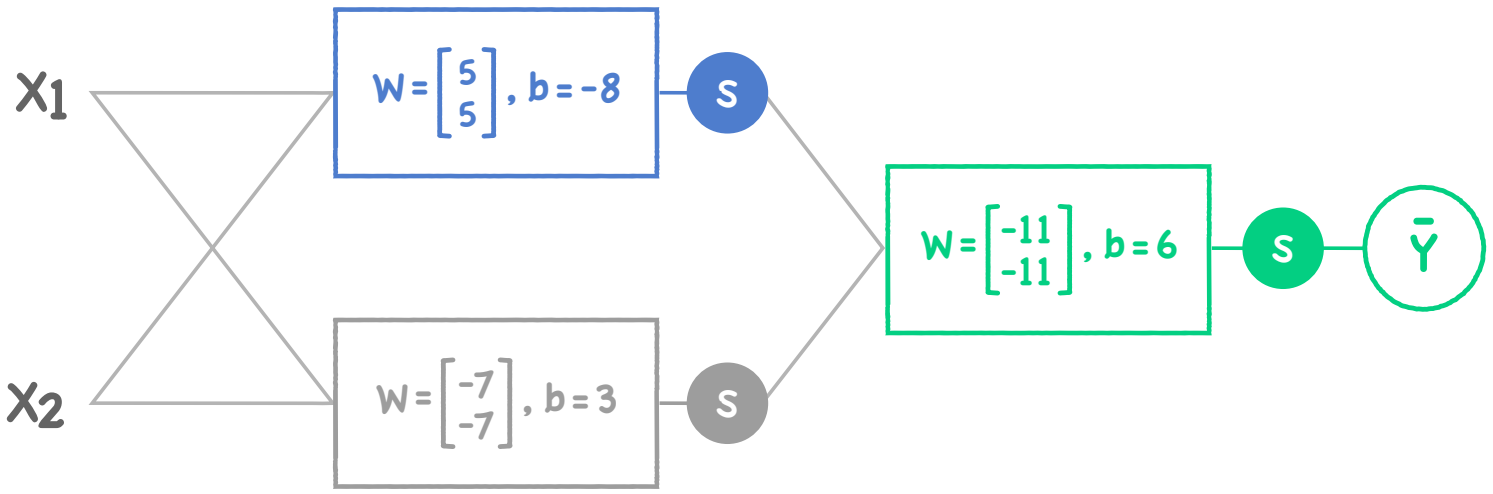
$$\text{Sigmoid}(-5) = 0$$

X_1	X_2	Y_1	Y_2	$\bar{Y}$	XOR
0	0	0	1	0	0
0	1	0	0	1	1
1	0	0	0	1	1
1	1	1	0	0	0

Forward Propagation

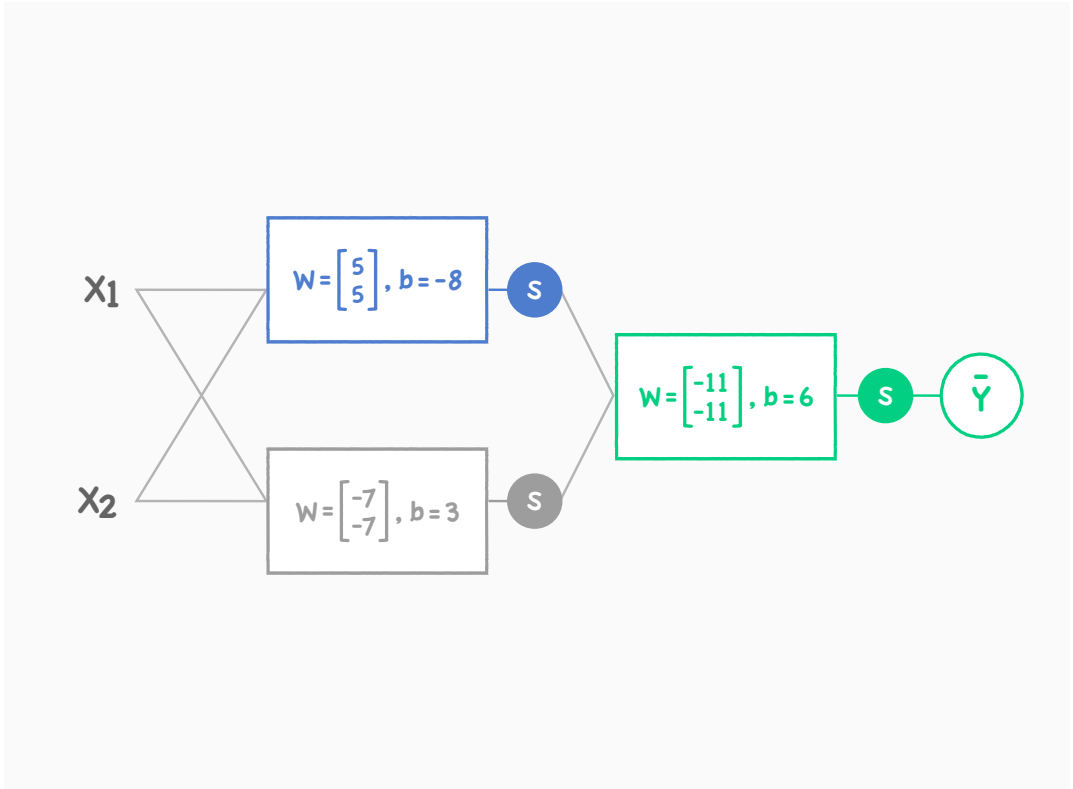


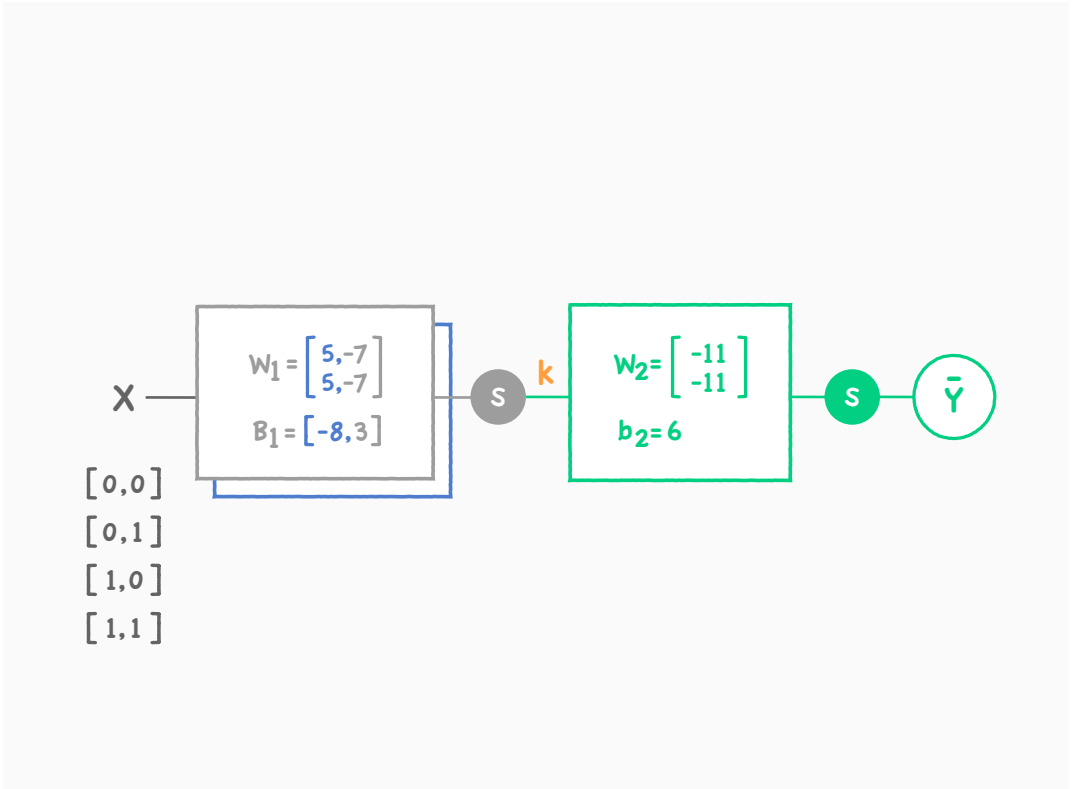
Forward Propagation



Can you find another W and b for the XOR?

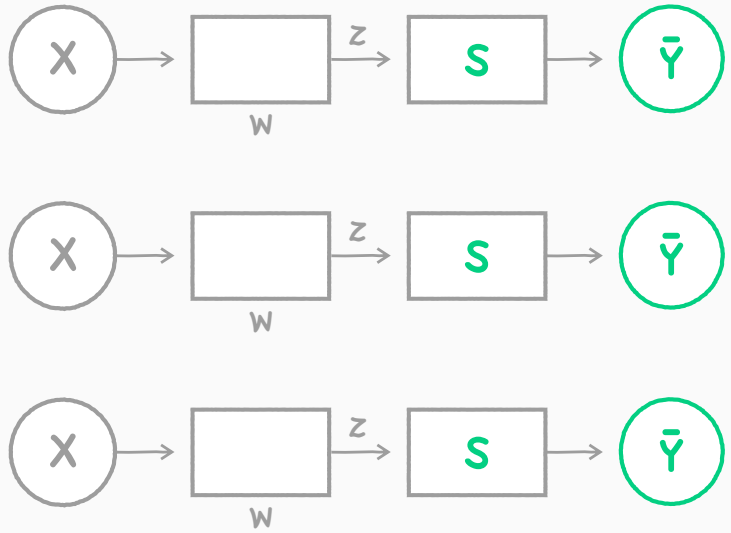
NN


X1
X2
W = [5, 5], b = -8
W = [-7, -7], b = 3
S
S
W = [-11, -11], b = 6
S
Ȳ

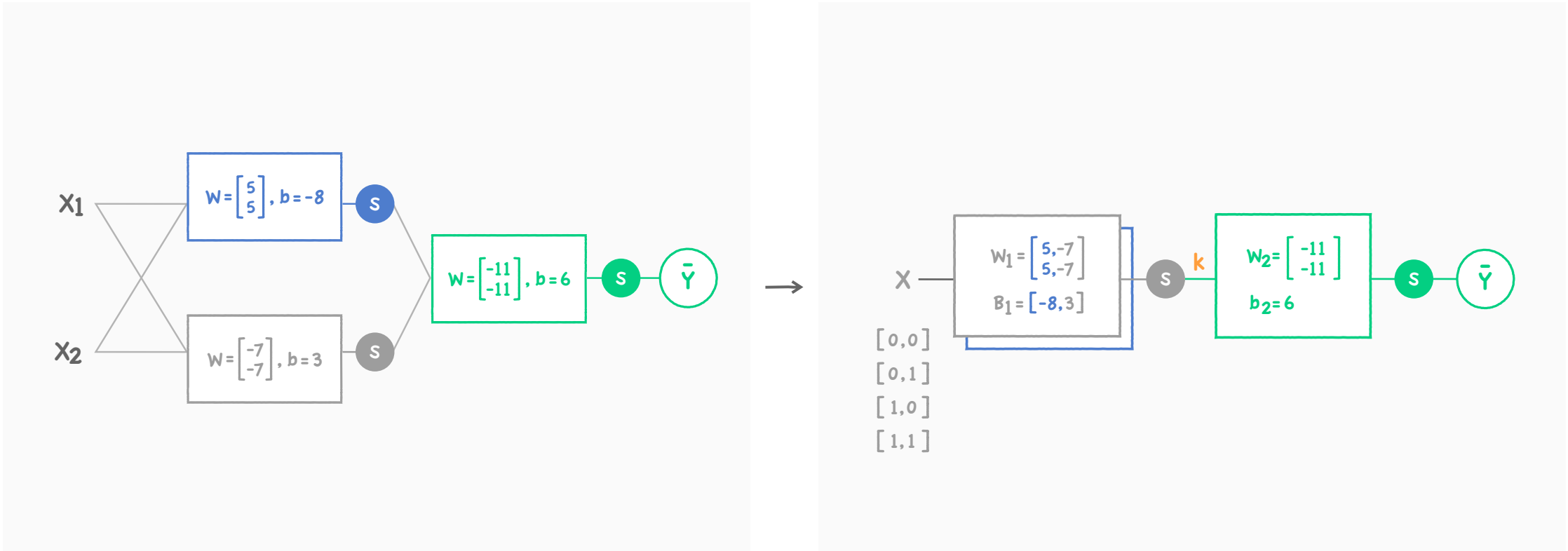

X
[0,0]
[0,1]
[1,0]
[1,1]
W1 = [5,-7; 5,-7]
B1 = [-8,3]
S
k
W2 = [-11; -11]
b2 = 6
S
Ȳ

Recap: Lec 6-1 Multinomial Classification

$$\begin{bmatrix} w_{A1} & w_{A2} & w_{A3} \\ w_{B1} & w_{B2} & w_{B3} \\ w_{C1} & w_{C2} & w_{C3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} w_{A1}x_1 + w_{A2}x_2 + w_{A3}x_3 \\ w_{B1}x_1 + w_{B2}x_2 + w_{B3}x_3 \\ w_{C1}x_1 + w_{C2}x_2 + w_{C3}x_3 \end{bmatrix} = \begin{bmatrix} \bar{y}_A \\ \bar{y}_B \\ \bar{y}_C \end{bmatrix}$$

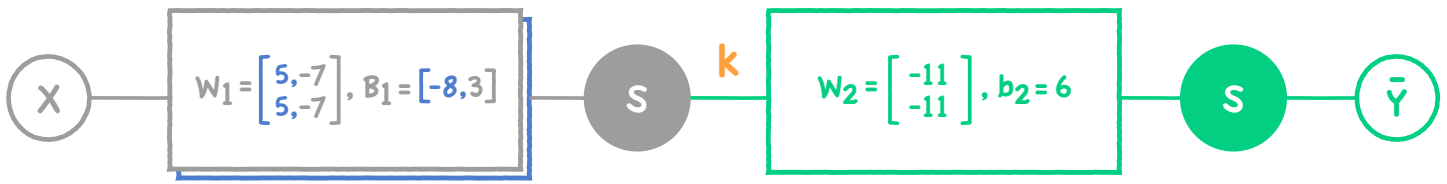


NN



How can we learn W, and b from training data?

NN



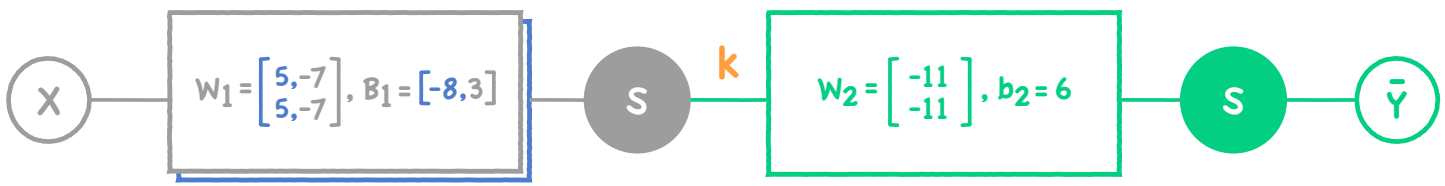
$K(X) = \text{Sigmoid}(XW_1 + B_1)$

$\bar{Y} = H(X) = \text{Sigmoid}(K(X)W_2 + b_2)$

[0,0]
[0,1]
[1,0]
[1,1]

NN

```
# NN
K = tf.sigmoid(tf.matmul(X, W1) + b1)
hypothesis = tf.sigmoid(tf.matmul(K, W2) + b2)
```

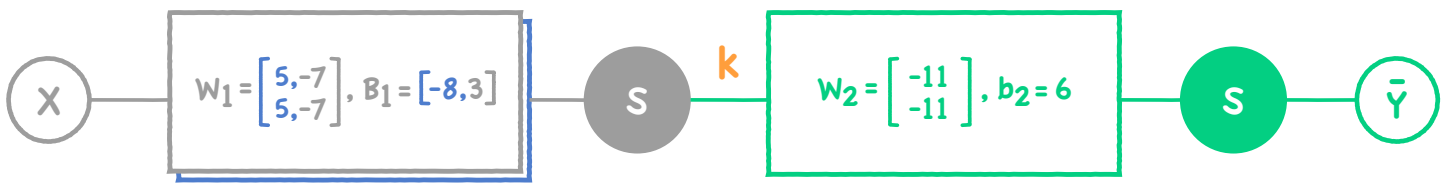


[0,0]
[0,1]
[1,0]
[1,1]

$$K(X) = \text{Sigmoid}(XW_1 + B_1)$$
$$\bar{Y} = H(X) = \text{Sigmoid}(K(X)W_2 + b_2)$$

NN

```
# NN
K = tf.sigmoid(tf.matmul(X, W1) + b1)
hypothesis = tf.sigmoid(tf.matmul(K, W2) + b2)
```



$K(X) = \text{Sigmoid}(XW_1 + B_1)$

$\bar{Y} = H(X) = \text{Sigmoid}(K(X)W_2 + b_2)$

[0,0]
[0,1]
[1,0]
[1,1]

How can we learn W1, W2, B1, b2 from training data?

NEXT LECTURE

BACKPROPAGATION