(*) Neural Networks.

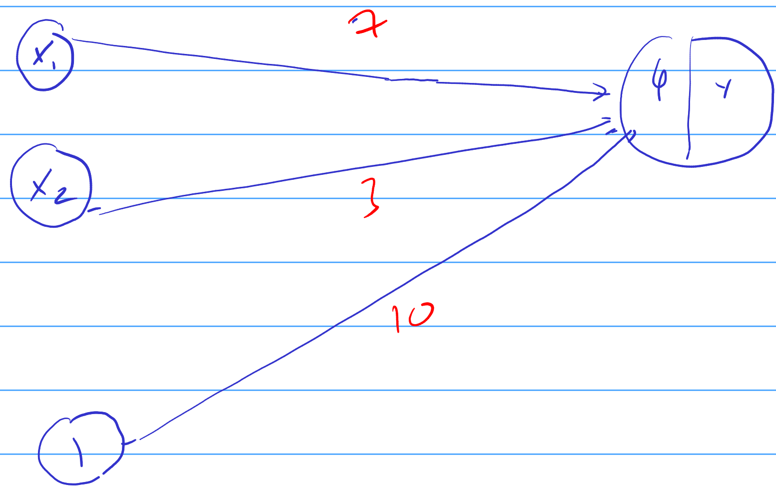
① A Graphical presentation of linear models

$$y = 7 \cdot x_1 + 3x_2 + 10$$

Data

x_1	x_2	y
⋮	⋮	⋮

Input layer output layer



$\varphi(t) = t$

$$\Longleftrightarrow \quad x_1 \cdot 7 + x_2 \cdot 3 + 1 \cdot 10 = y$$

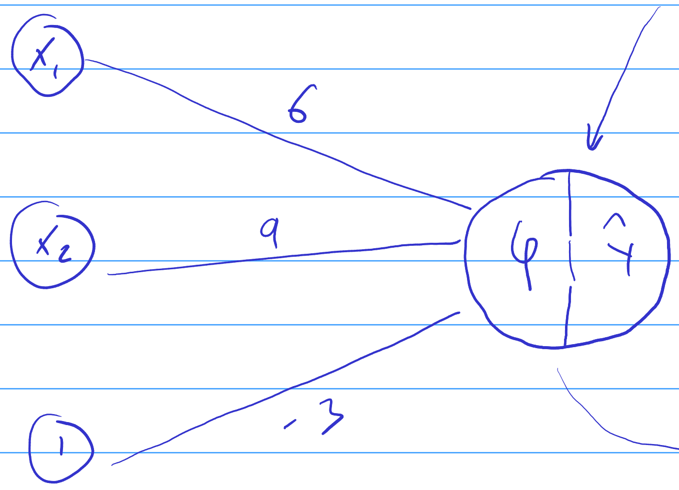
② A Graphical presentation of logistic Regression

x_1	x_2	y

$$\hat{y} = \frac{1}{1 + e^{-\left[\underbrace{6 \cdot x_1 + 9 x_2 - 3}\right]}}$$

Input layer

output layer

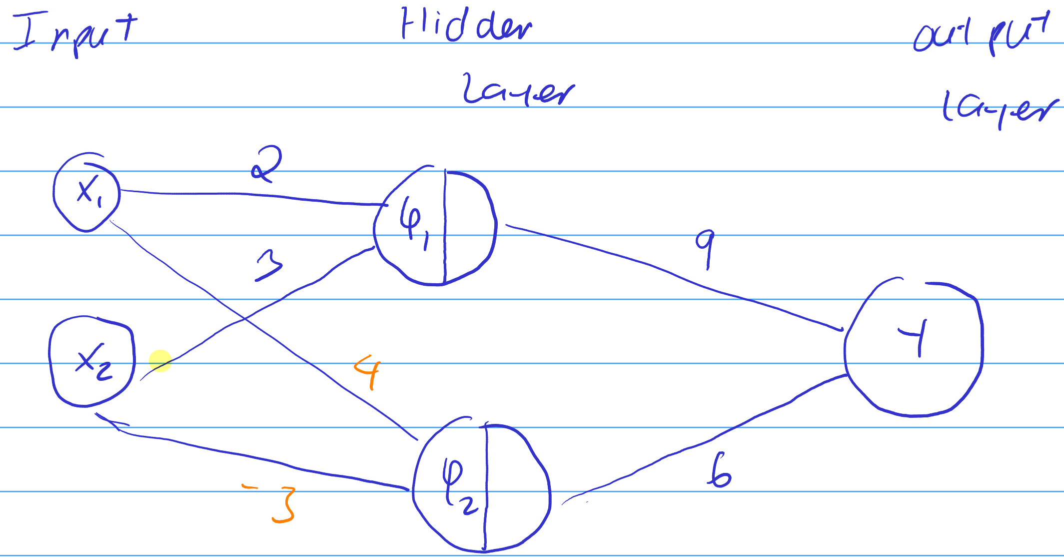


$$\Big(\underbrace{x_1 \cdot 6 + x_2 \cdot 9 + 1 \cdot (-3)}\Big) \cdot \frac{1}{1 + e^{-(\dots)}} = \varphi$$

$$\varphi(t) = \frac{1}{1 + e^{-t}}$$

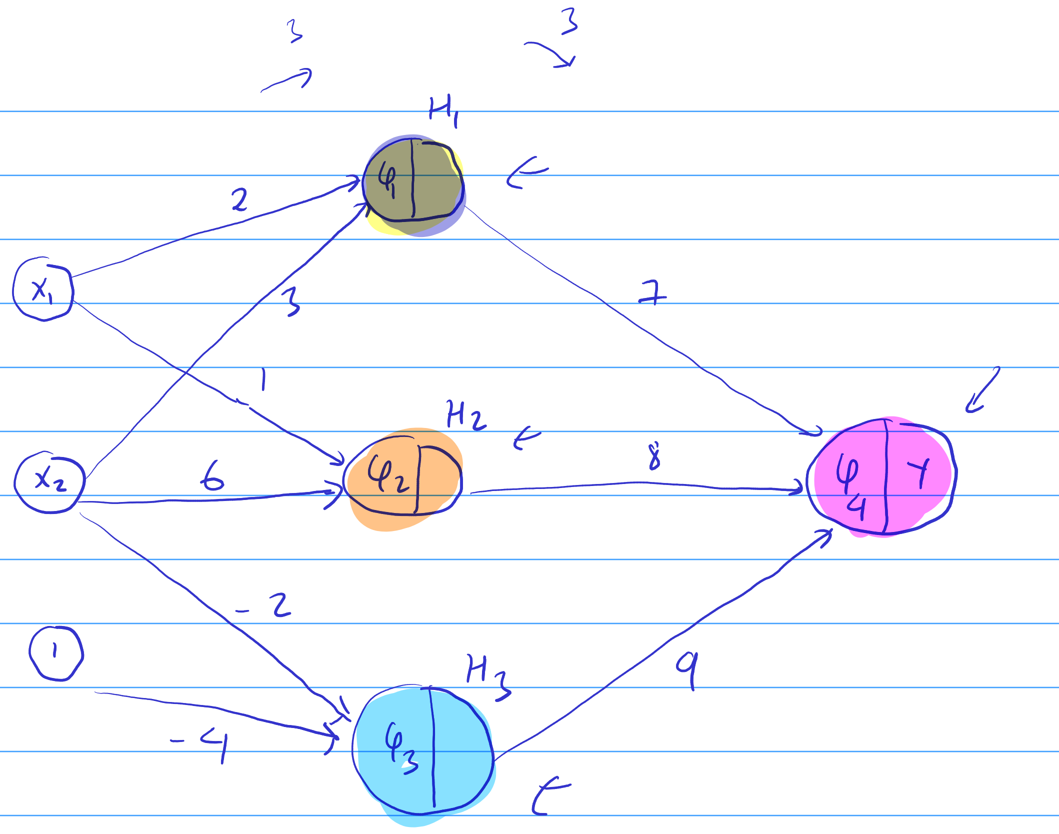
$$y = 9 \cdot (2x_1 + 3x_2)^2 + 6 \cdot \cos(4x_1 - 3x_2)$$

x_1	x_2	y



$$\varphi_1(t) = t^2 \; ; \quad \varphi_2(t) = \cos(t)$$

φ_1 and φ_2 are called activation functions



$\varphi_1(t) = t^3$

$\varphi_2(t) = \log(t)$

$\varphi_3(t) = \sqrt{t}$

$\varphi_4 = \sin(t)$

$$\sin\left[7\cdot\underbrace{(2\cdot x_1 + 3x_2)^3}_{H_1} + 8\cdot\underbrace{\log(x_1 + 6x_2)}_{H_2} + 9\cdot\underbrace{\sqrt{(-2x_2 - 4)}}_{H_3}\right]$$

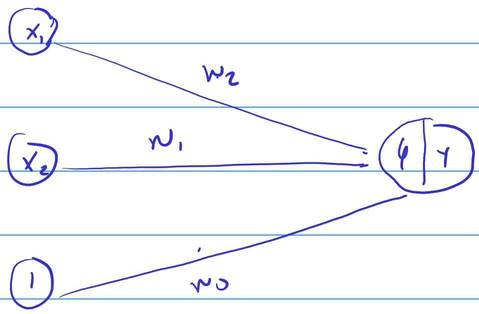
$$= Y$$

$$(Y = \sin(7\cdot H_1 + 8\cdot H_2 + 9\cdot H_3))$$

(*) φ_1, φ_2, φ_3 and φ_4 are called activation functions

(*) Parameters in NN.

input layer

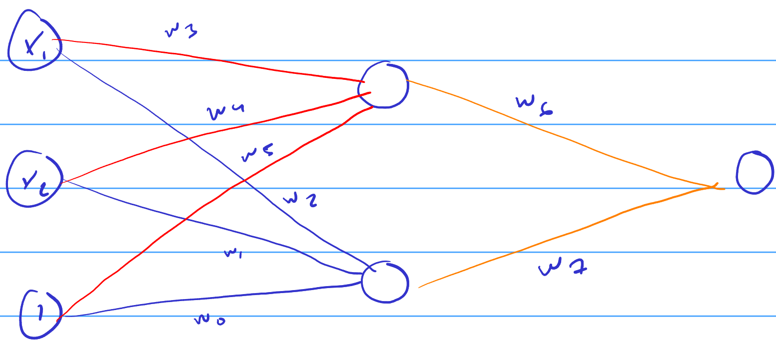


$\varphi(t) = t$

$Y = \underline{w_0} + \underline{w_1}\ \underline{x_1} + \underline{w_2}\ \underline{x_2}$

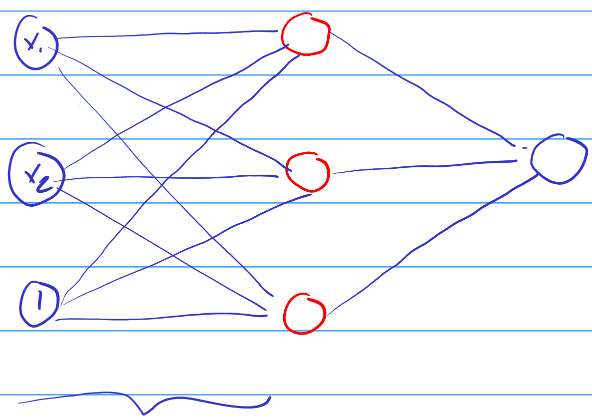
(untrained network) : 3 parameters. Training this network means finding these 3 parameters.

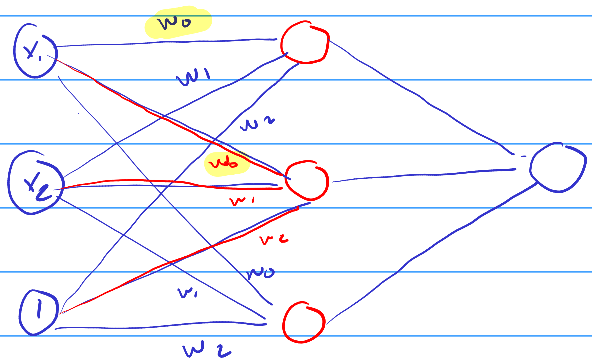
Input layer



8 parameters

(*)





Fully connected layer
(3x3 = 9 parameters)

we can set conditions on parameters to reduce the number of parameters.

Data

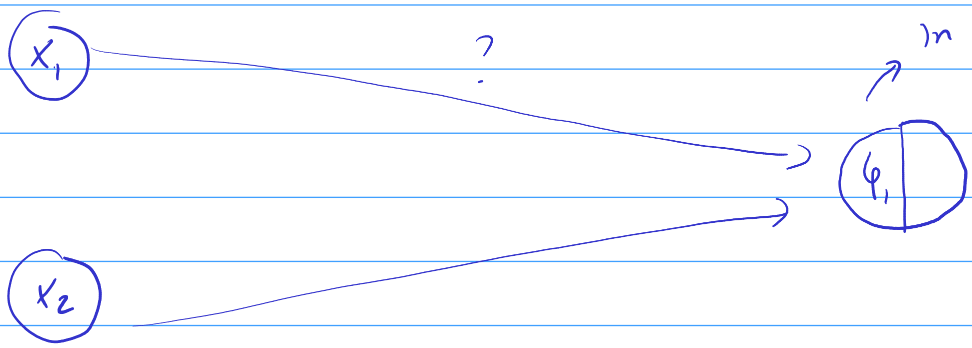
u_1	u_2	u_3
x_1	$\ln x_2$	$x_1 \cdot \ln x_2$

$\Leftarrow$

x_1	x_2	y

$$Y = X_1^{X_2} \Rightarrow \ln Y = X_2 \cdot \ln X_1$$

$\varphi_1(t) = \ln t$



$x_1 x_2$	x_1	x_2	Y

$$Y = X_1 + X_2 + X_1 X_2$$

$$\boxed{Y = (\alpha X_1)^{(\beta X_2)}} \Rightarrow \ln Y = \beta X_2 \cdot \ln(\alpha X_1)$$

$$\ln(\ln Y) = \ln(\beta X_2) + \ln(\ln(\alpha X_1))$$