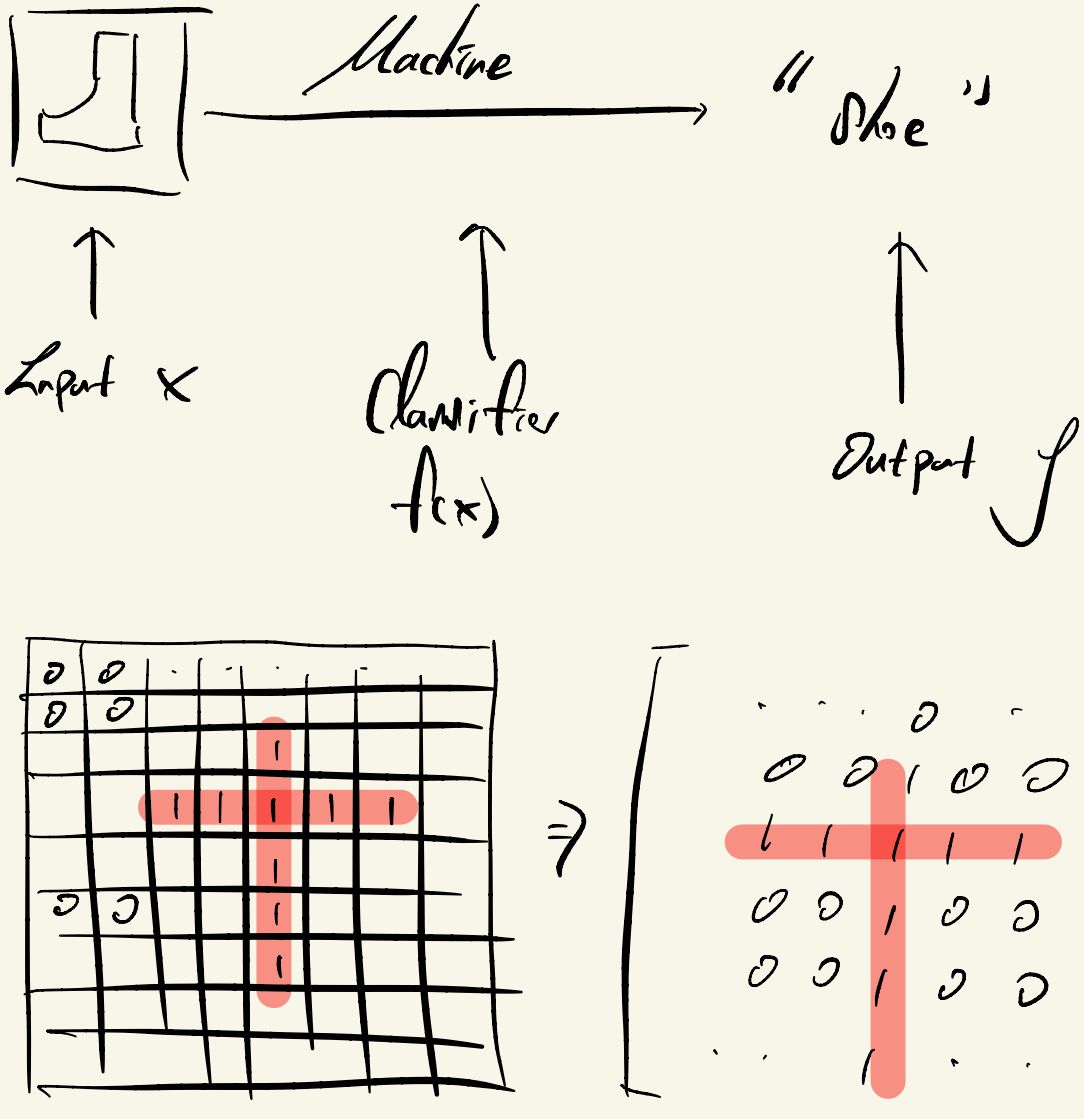
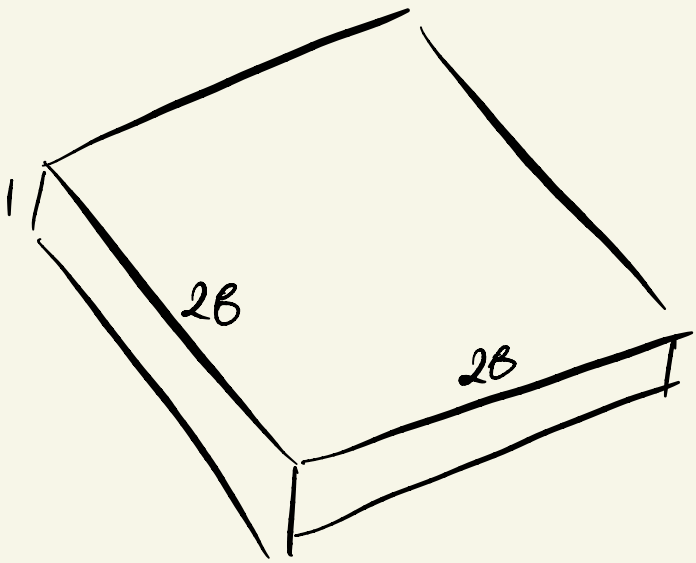
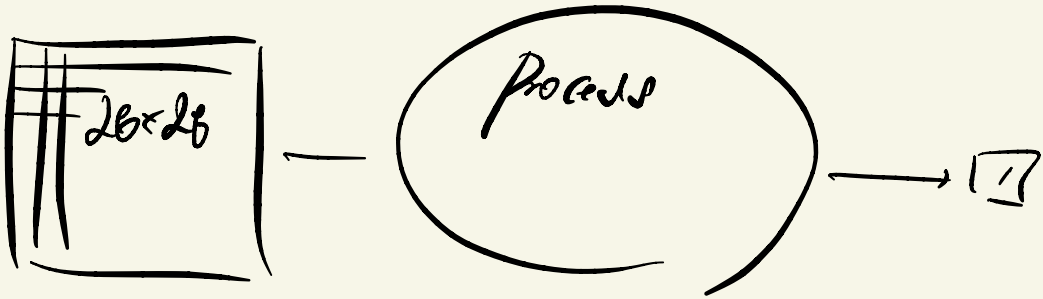

Machine
"Shoe"
Input x
Classifier
f(x)
Output y

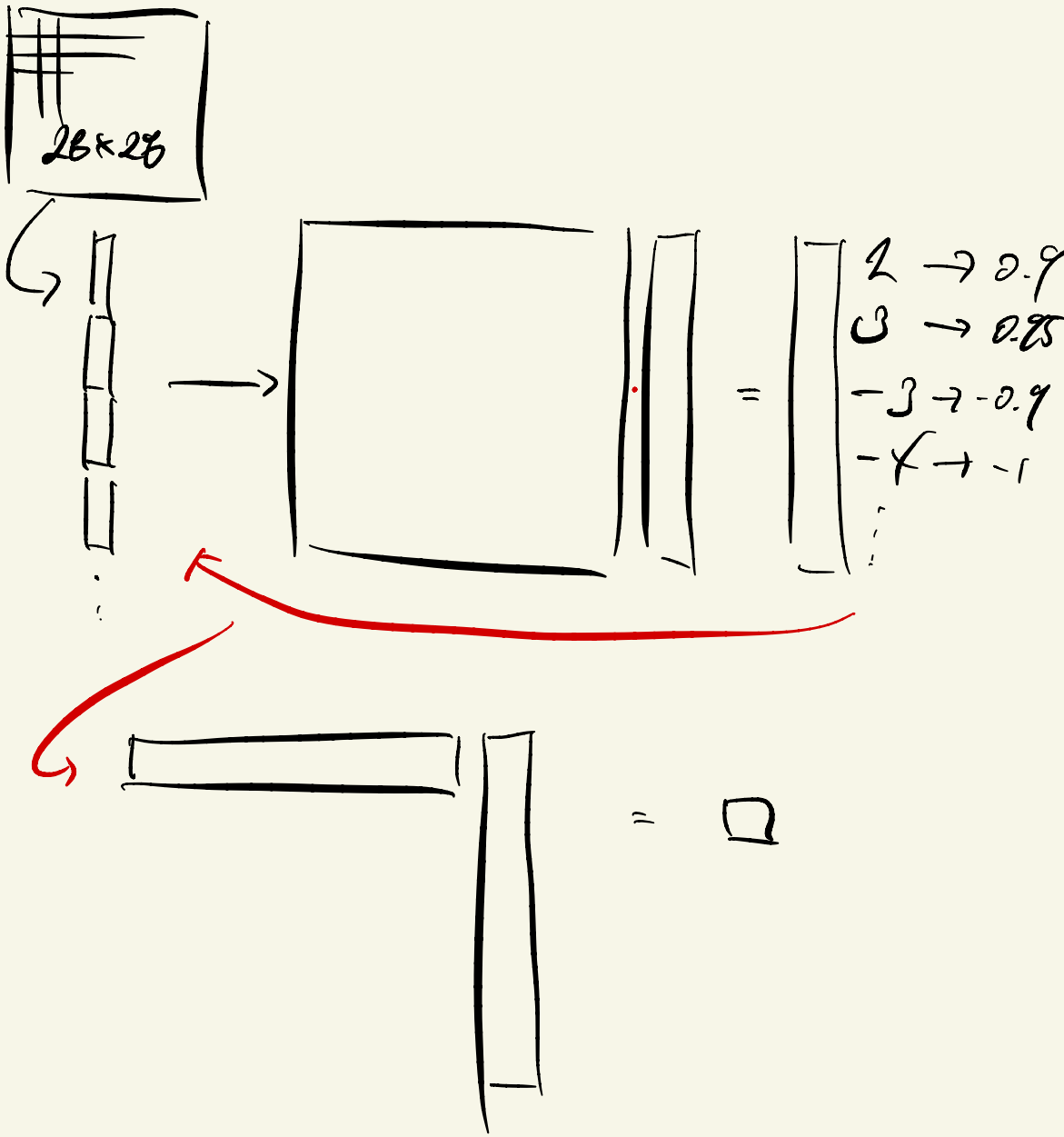


$1 \times 28 \times 28 \longrightarrow 1 \times 1$

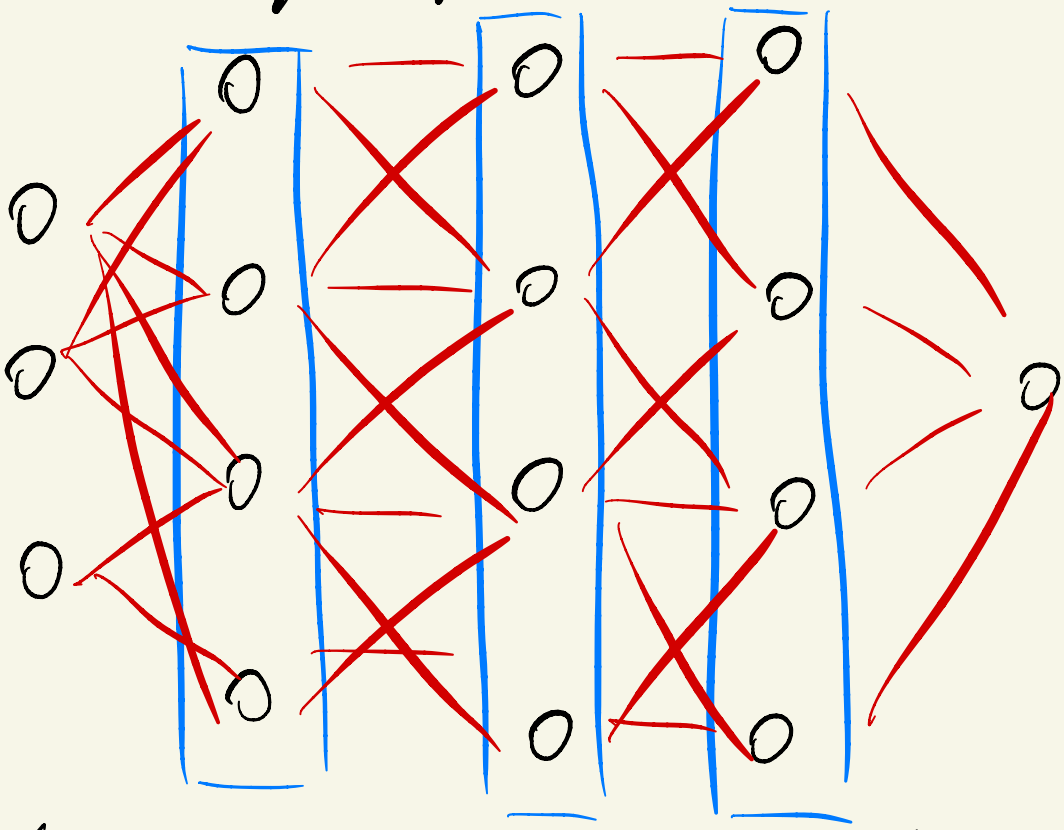
$f : \mathbb{R}^{28 \times 28} \longrightarrow \mathbb{R}$



Neural Network is a kind of f(x)



Multi - Layer Perceptron



$$W_n\ \sigma(W_{n-1}\ \sigma(W_2\ \sigma(W_1 x)))$$

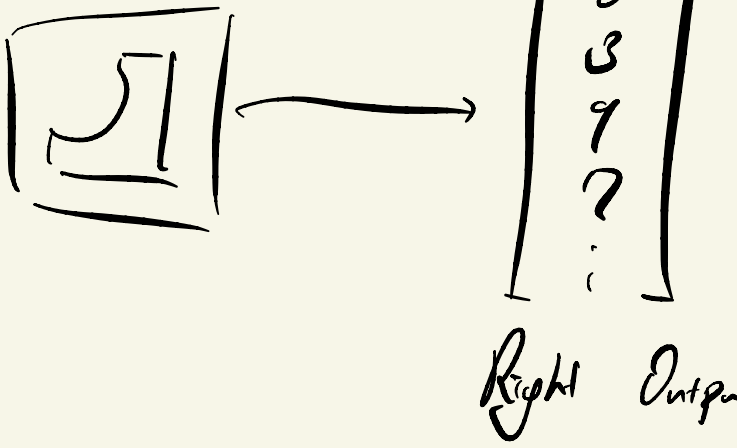
Input data (array/s)

Output data (array/s)

Neural network (function: Input → Output)

$$W_n \sigma(W_3 \sigma(W_2 \sigma(W_1 x + b_1) + b_2) + b_3)$$

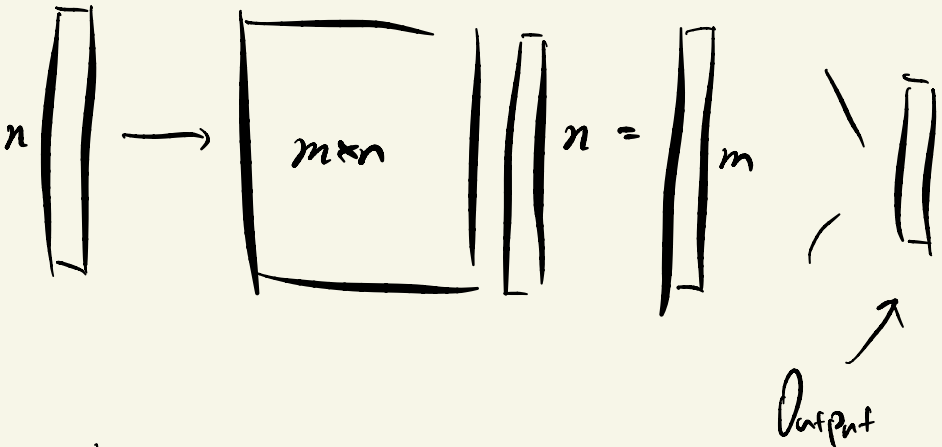
Training



Right Output ?

y_{train} = training output

$y_{predict}$ = NN prediction



Loss

$$\min_{w,b} \sum_{\text{training data}} \| y_{train} - y_{pred} \|^2$$

Minimize Loss → Optimization Algorithm