ETH zürich

Introduction to Scientific Computation
Lecture 4
Fall 2018

Data related problems, Supervised learning
Perceptron

Mikhail Usvyatsov October 17, 2018

ETH zürich

Basic definitions

We have a set of objects: X

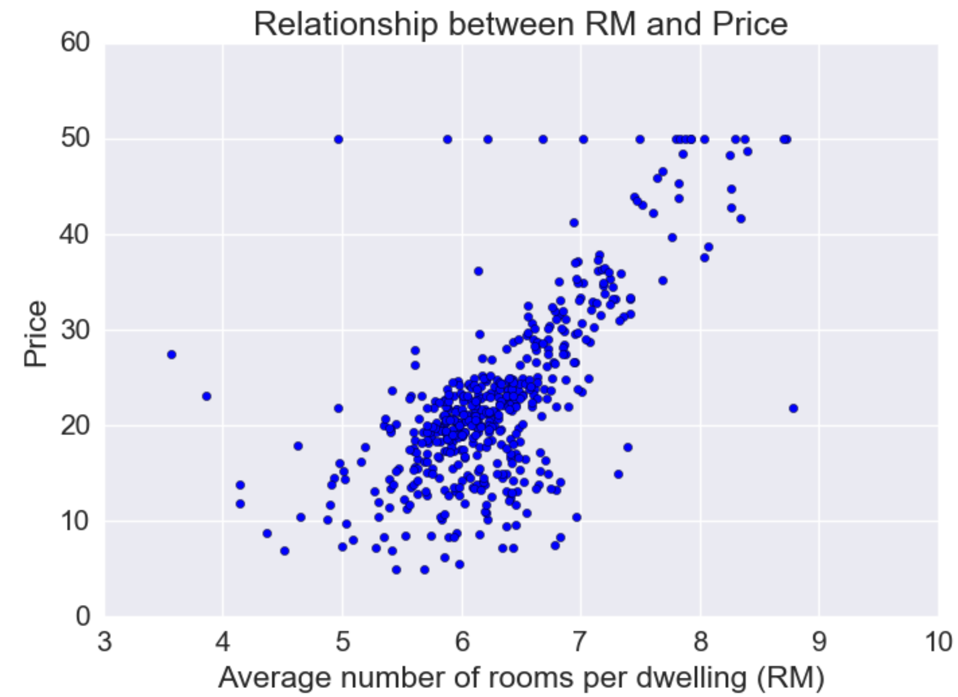
And the set of possible answers: Y

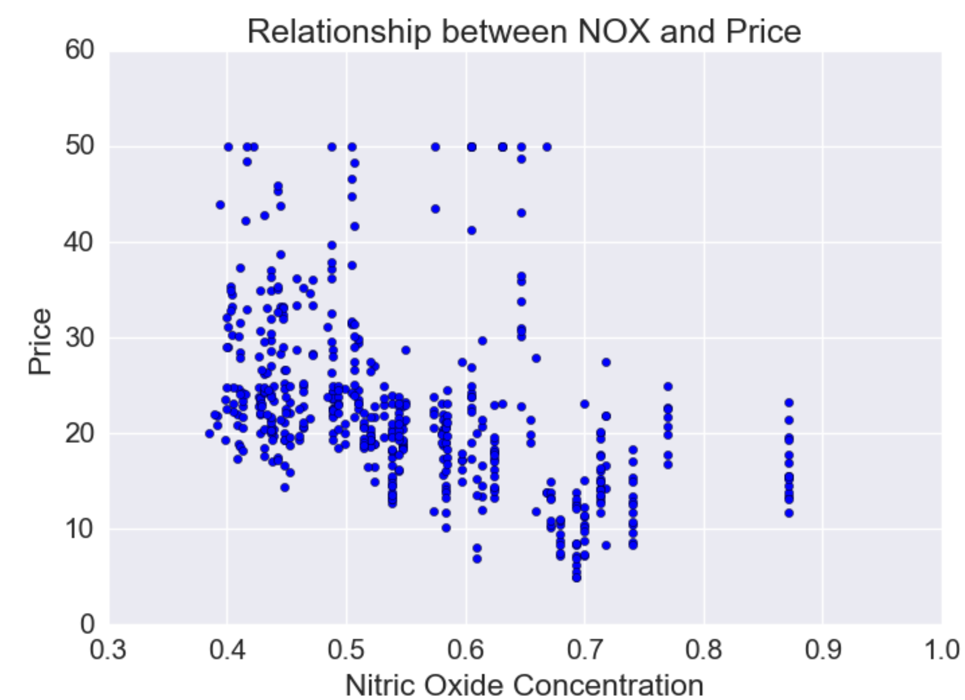
Basic definitions

We have a set of objects: X

And the set of possible answers: Y

We define the target function: $y^* : X \to Y$





Basic definitions

We have a set of objects: X

And the set of possible answers: Y

We define the target function: $y^* : X \to Y$

Supervised

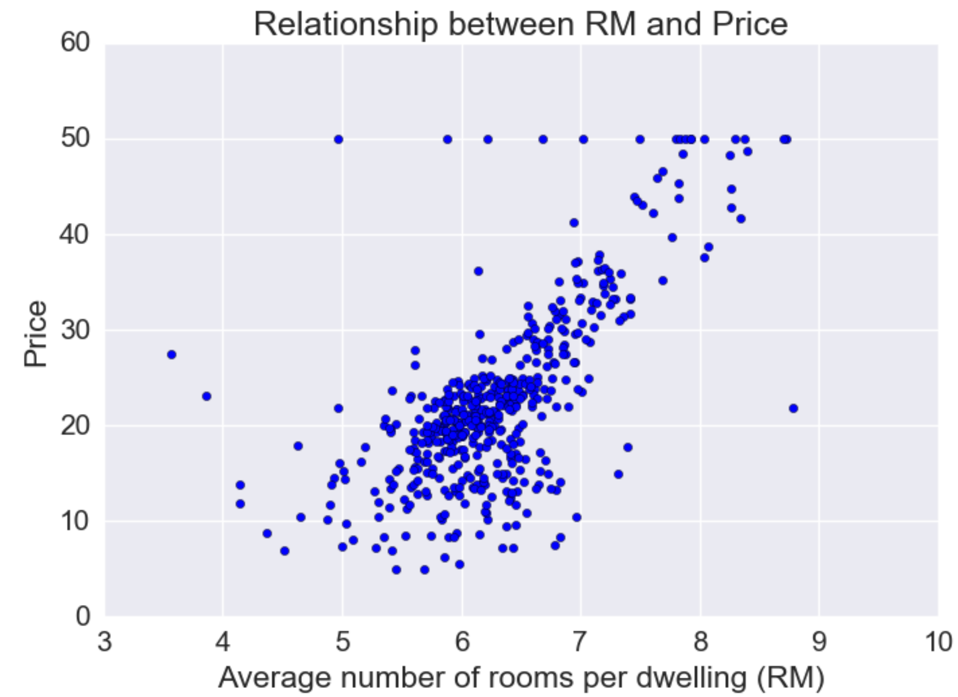
Unsupervised $Y \in \varnothing$

Supervised

We have a set of objects: X

And the set of possible answers: Y

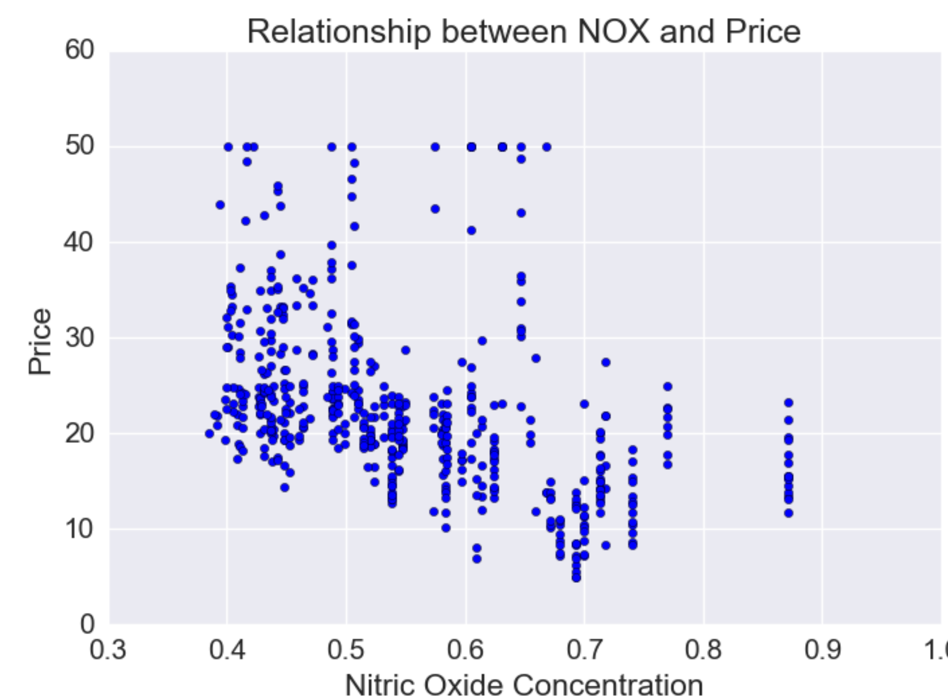
We define the target function: $y^* : X \to Y$



And we know values of y^* only on a finite subset $\{x_1, \ldots, x_l\} \subset X$

Training sample: $X^l = (x_i, y_i)_{i=1}^{l}$

Precedent



ETH zürich

Let’s look into this guy

$$X^l = \left(x_i, y_i\right)_{i=1}^{l}$$

Let's look into this guy

$$X^l = \left(x_i, y_i\right)_{i=1}^{l}$$

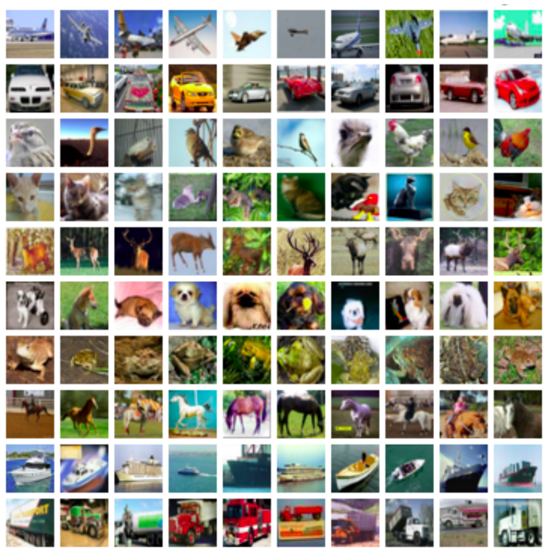
	N_Rooms	Floor	Smokers?
Hous1	1	1	Yes
...	...	...	...
Hous N	3	-	No

We have a set of objects: X

And the set of possible answers: Y

We define the target function: $y^* : X \to Y$

If Y is discrete $\{1, \ldots, N\}$

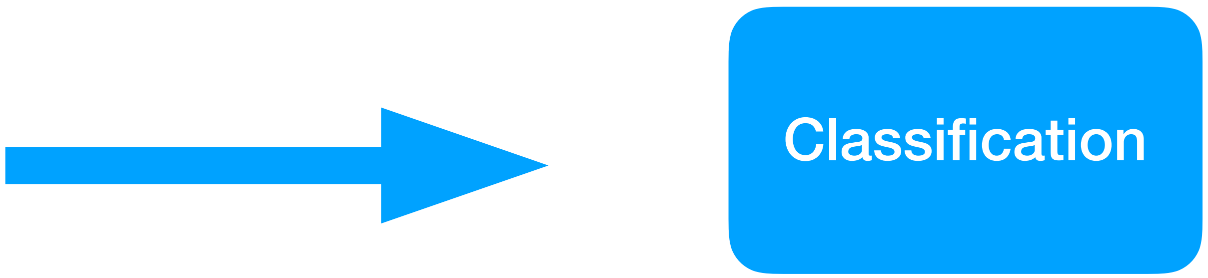


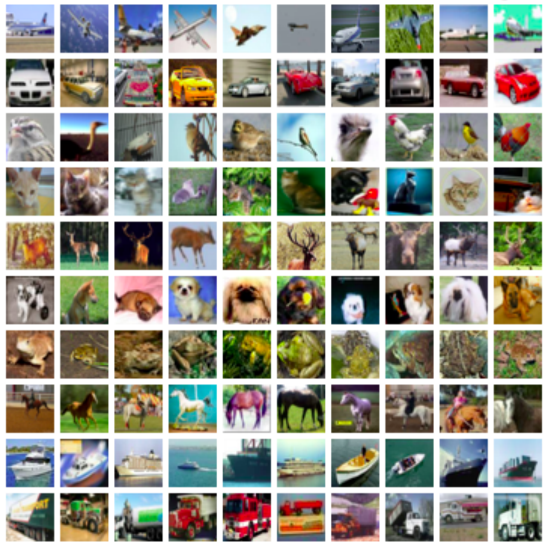
We have a set of objects: X

And the set of possible answers: Y

We define the target function: $y^* : X \to Y$

If Y is discrete $\{1, \ldots, N\}$





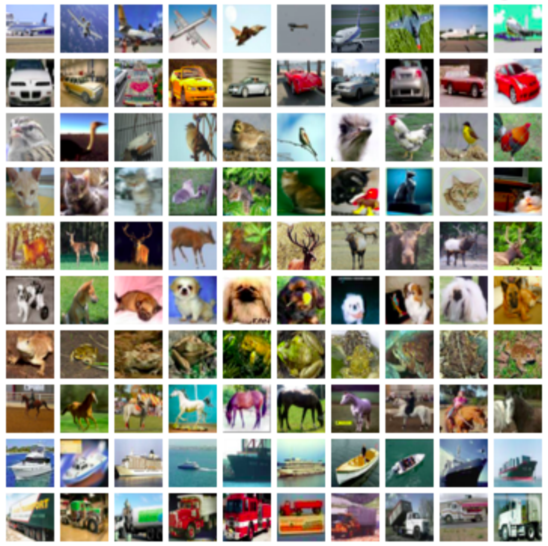
We have a set of objects: X

And the set of possible answers: Y

We define the target function: $y^* : X \to Y$

If Y is discrete $\{1, \ldots, N\}$ → **Classification**

If Y is continuous $[-25, 100]$



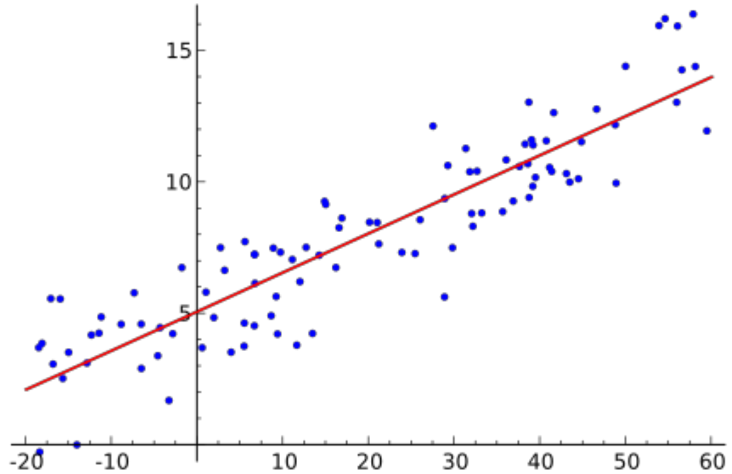
We have a set of objects: X

And the set of possible answers: Y

We define the target function: $y^* : X \to Y$

If Y is discrete $\{1, \ldots, N\}$ → **Classification**

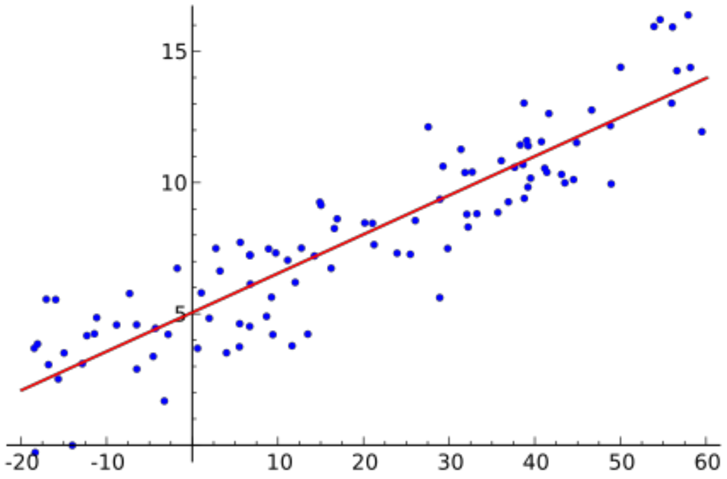
If Y is continuous $[-25, 100]$ → **Regression**



ETH zürich

How to say if one target function is better?

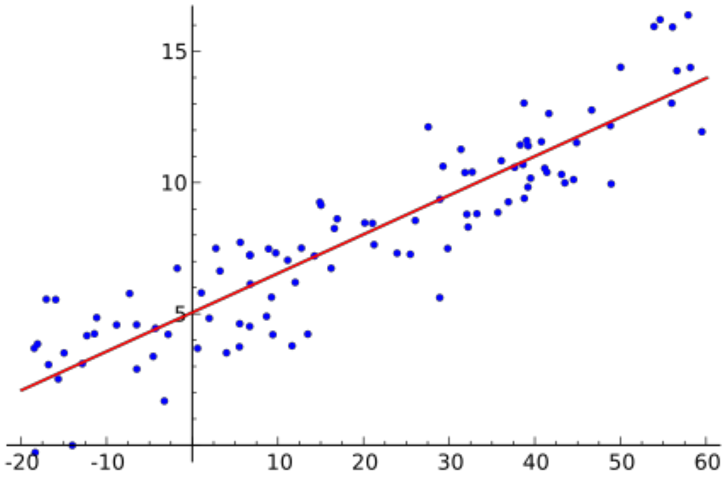
$y_1^* = 5x + 6$ or $y_2^* = 4x^3 - 2x + 1$



How to say if one target function is better?

$y_1^* = 5x + 6$ or $y_2^* = 4x^3 - 2x + 1$

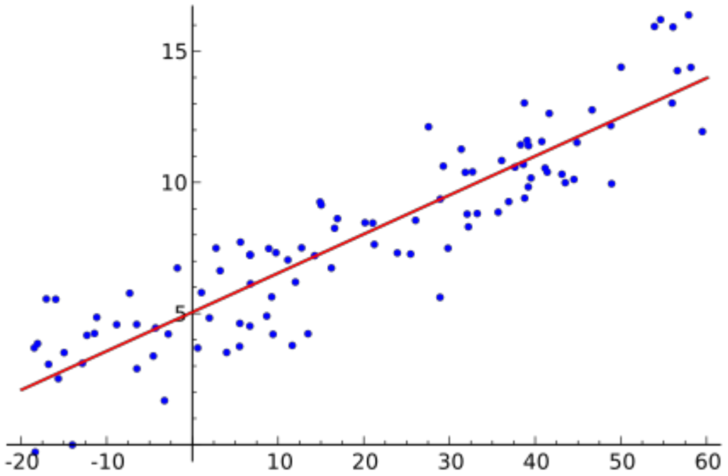
We need to compare them.



How to say if one target function is better?

$y_1^* = 5x + 6$ or $y_2^* = 4x^3 - 2x + 1$

We need to compare them.



$L(\theta, x_i, y_i)$

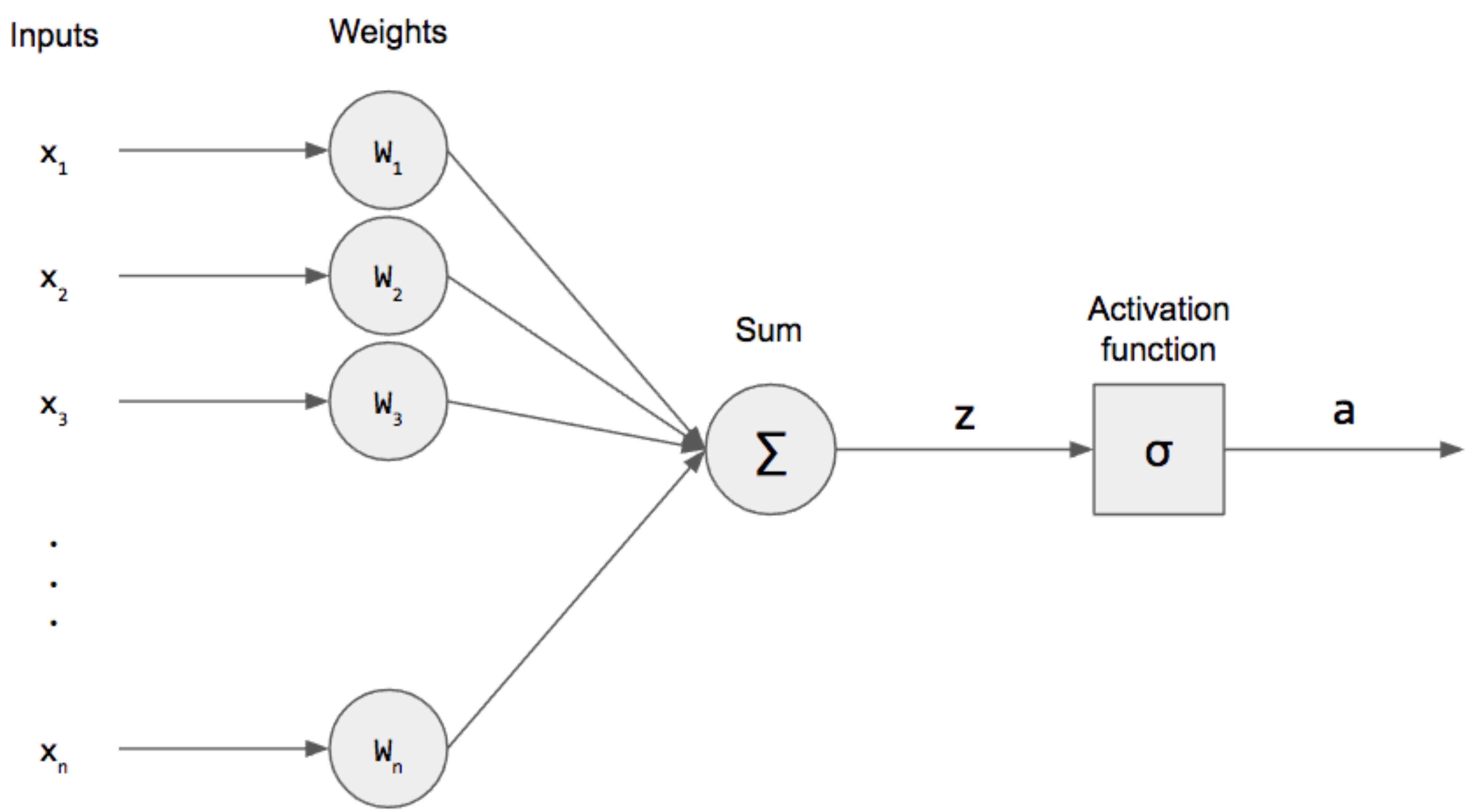
Loss Function

Popular loss functions:

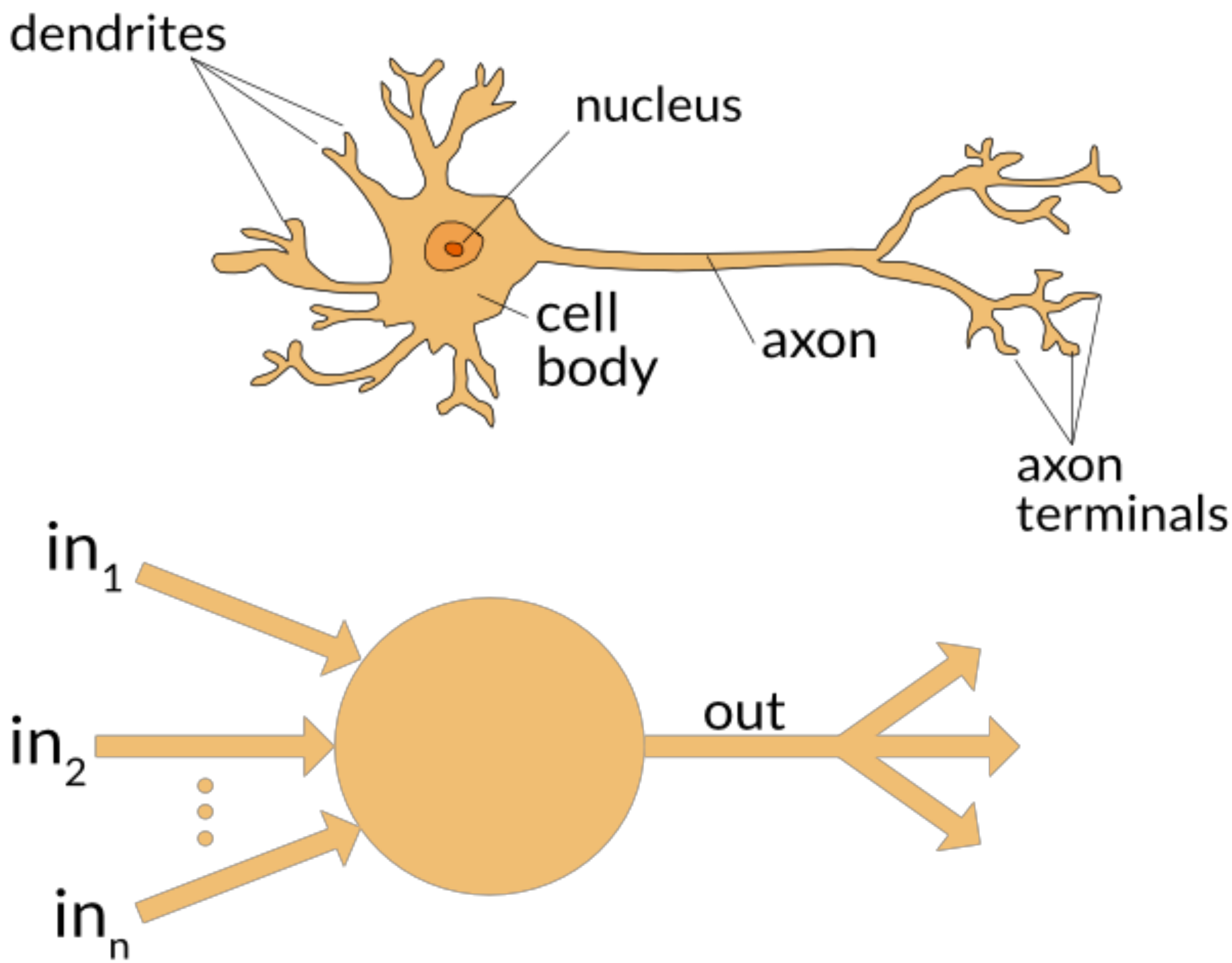
Classification $\rightarrow$ $-(y_i \log(p_i) + (1 - y_i)\log(1 - p_i))$

Regression $\rightarrow$ $(y_i - p_i)^2$

Perceptron:



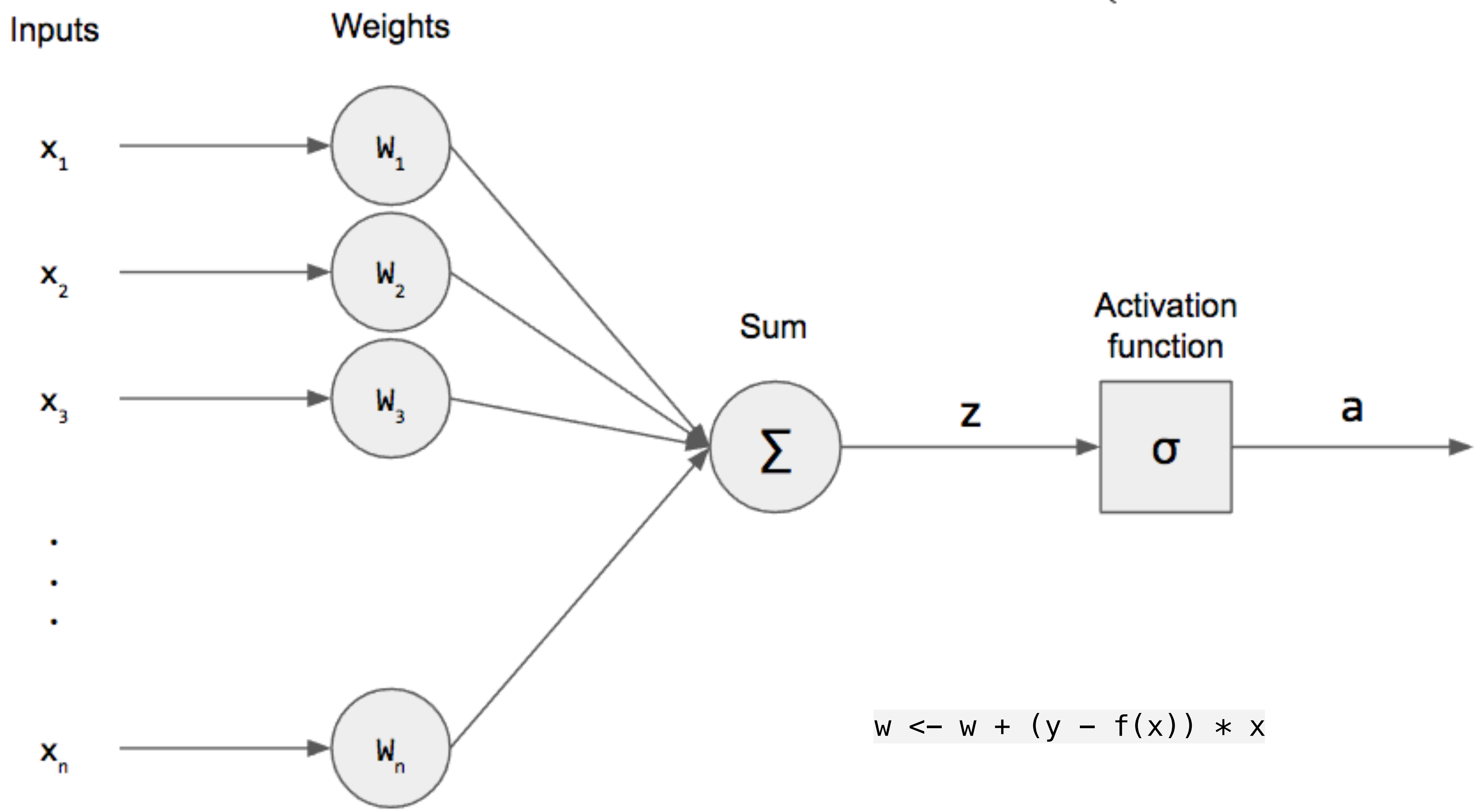
(c) https://towardsdatascience.com/what-the-hell-is-perceptron-626217814f53



(c) https://appliedgo.net/perceptron/

Perceptron:

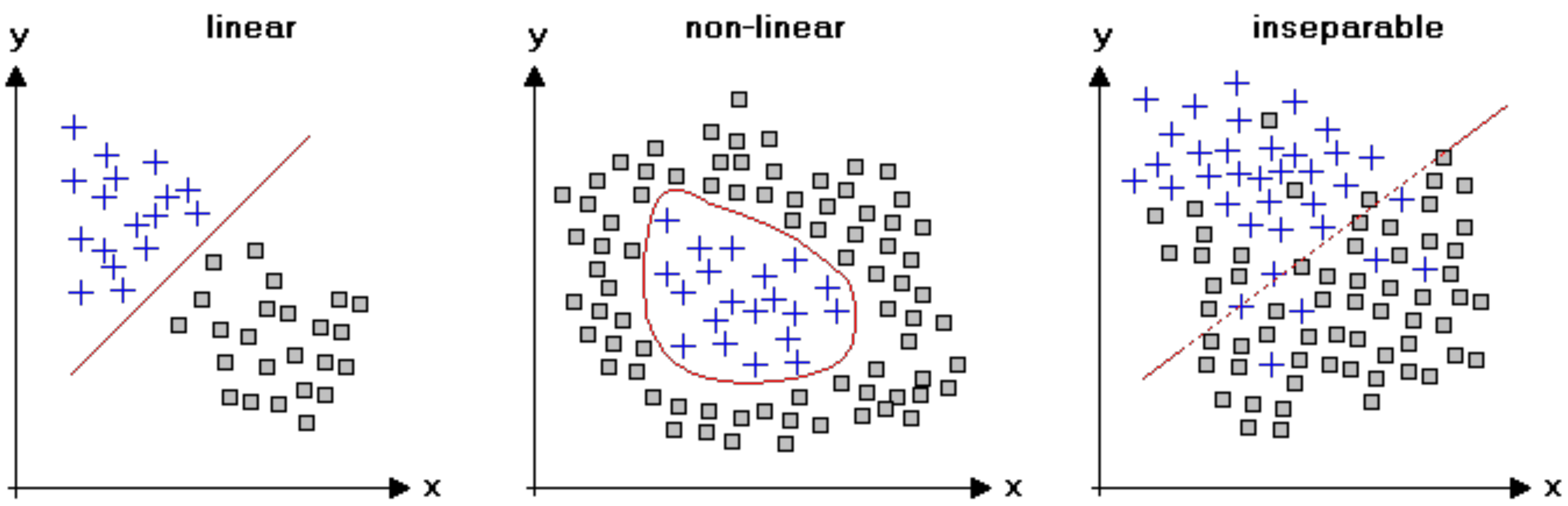
$$f(x) = \begin{cases} 1 & \text{if } w \cdot x + b > 0 \\ 0 & \text{otherwise} \end{cases}$$




```
w <- w + (y - f(x)) * x
```

(c) https://towardsdatascience.com/what-the-hell-is-perceptron-626217814f53

Perceptron:



(c) http://www.vias.org/tmdatanaleng/cc_data_structure.html