



Introduction to Scientific Computation

Lecture 4

Fall 2020

Data related problems, Supervised learning
Perceptron

Basic definitions

We have a set of objects: X

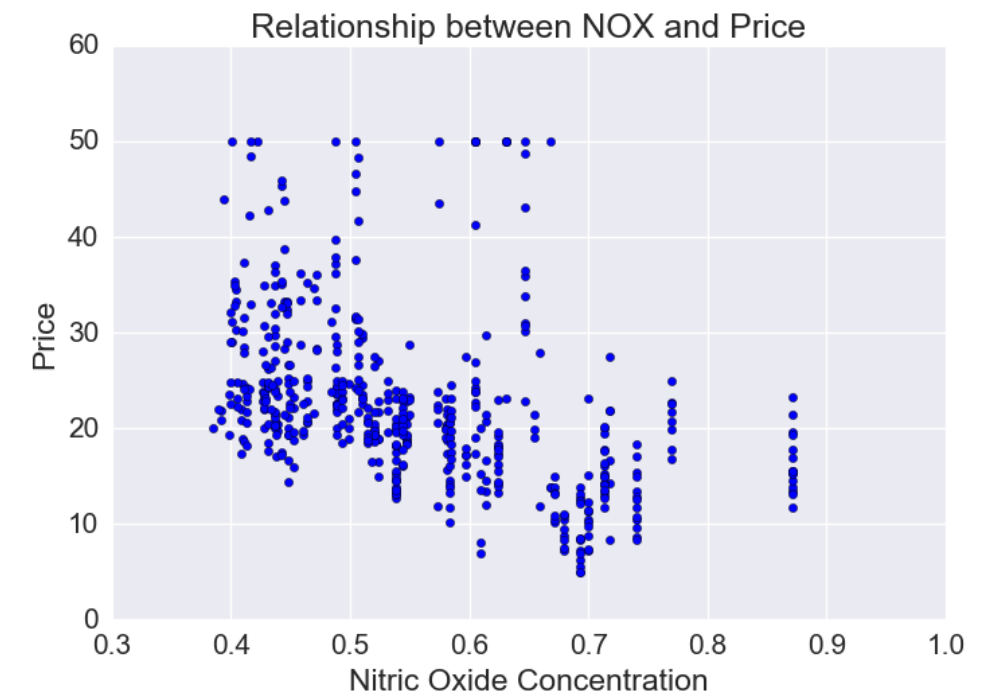
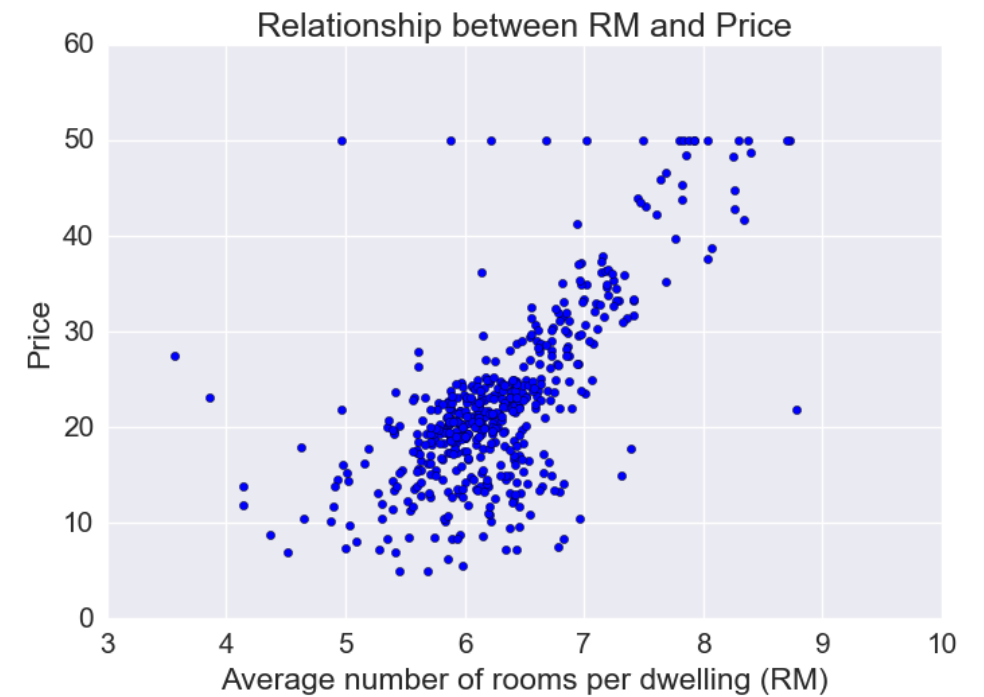
And the set of possible answers: Y

Basic definitions

We have a set of objects: X

And the set of possible answers: Y

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

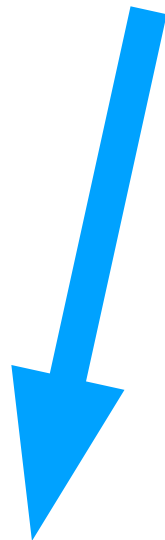


Basic definitions

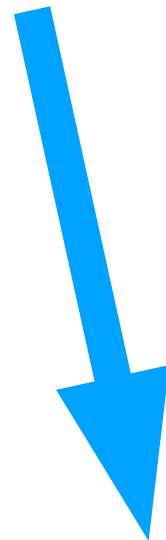
We have a set of objects: X

And the set of possible answers: Y

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



Supervised



Unsupervised $Y \in \emptyset$

Supervised

We have a set of objects: X

And the set of possible answers: Y

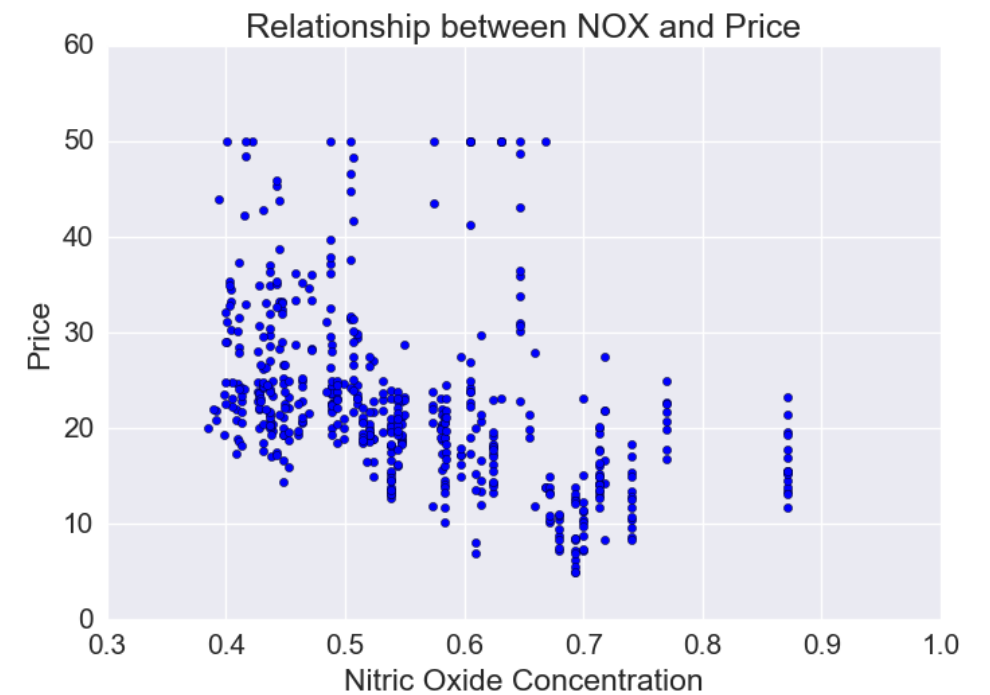
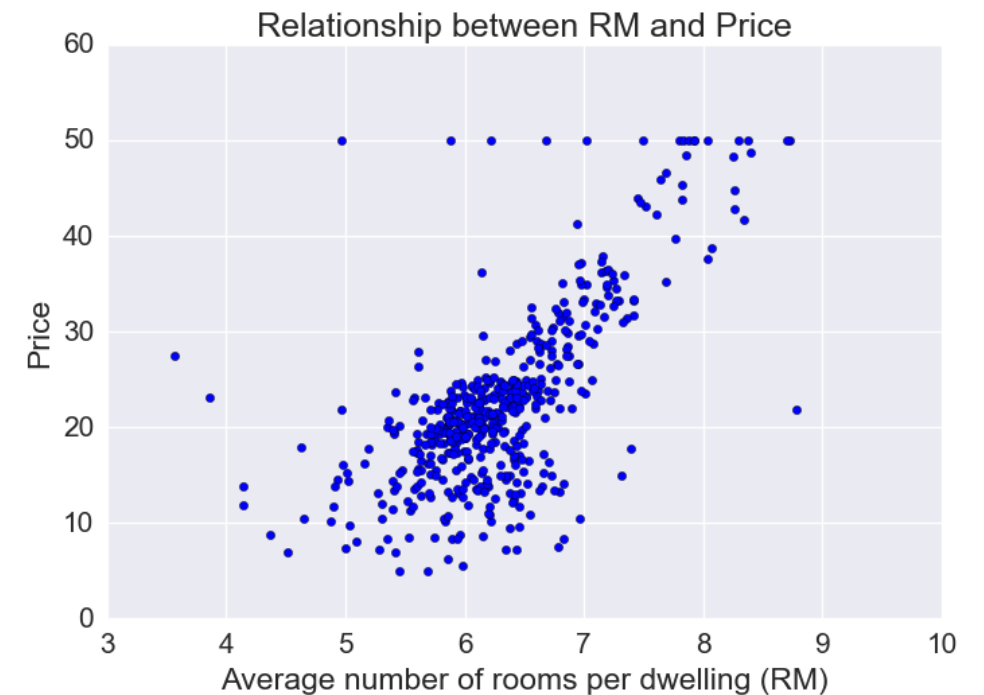
We define the target function: $y^* : X \rightarrow Y$

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

Training sample:

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

Precedent



Let's look into this guy



↓

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

Let's look into this guy



↓

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

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

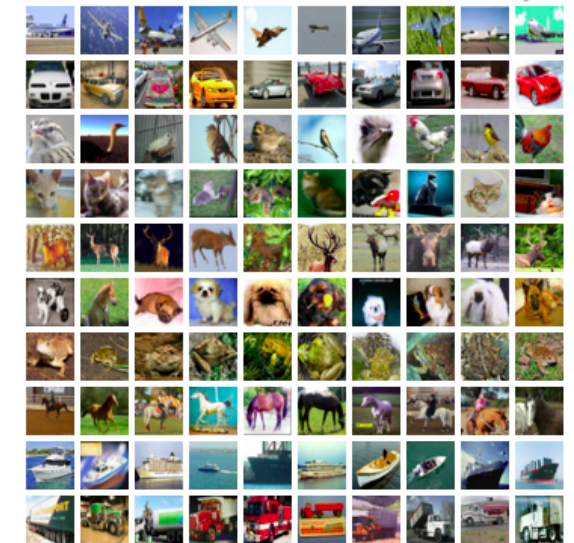
We have a set of objects: X

And the set of possible answers: Y

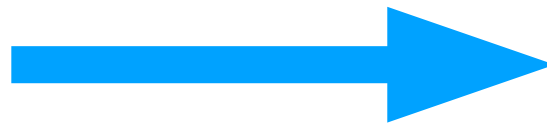
We define the target function: $y^* : X \rightarrow Y$

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

We have a set of objects: X
And the set of possible answers: Y
We define the target function: $y^* : X \rightarrow Y$

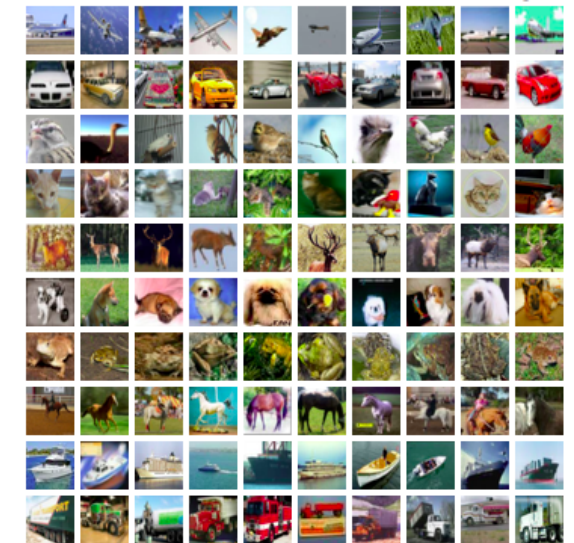


If Y is discrete $\{1, \dots, N\}$

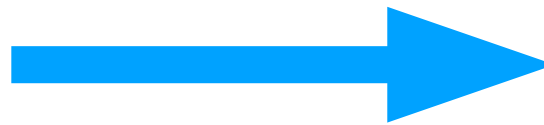


Classification

We have a set of objects: X
And the set of possible answers: Y
We define the target function: $y^* : X \rightarrow Y$

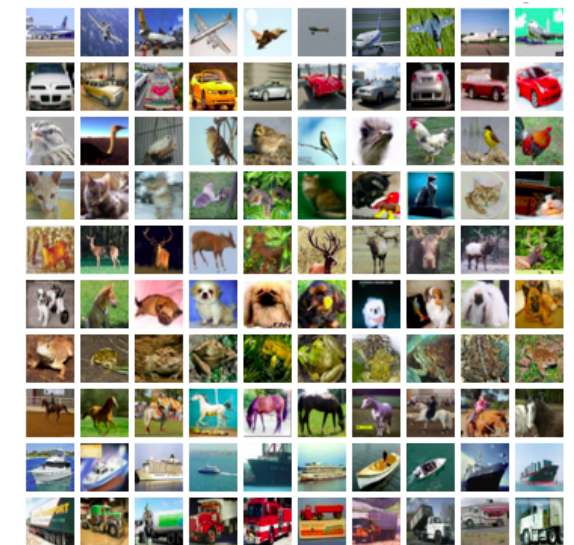


If Y is discrete $\{1, \dots, N\}$



Classification

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

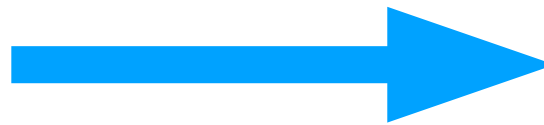


We have a set of objects: X

And the set of possible answers: Y

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

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

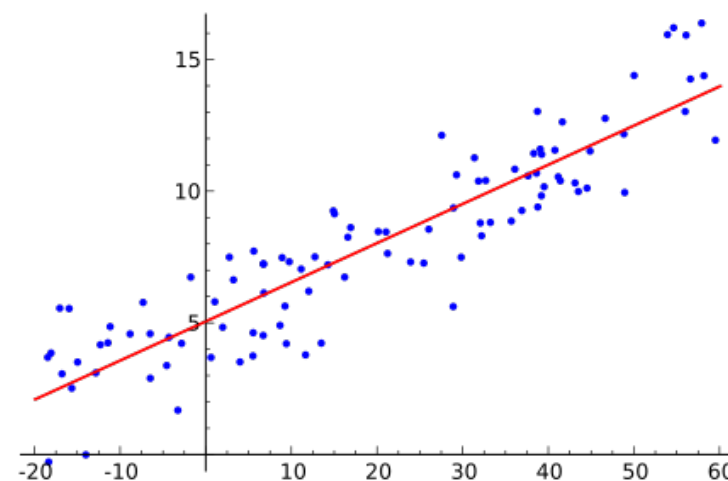


Classification

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

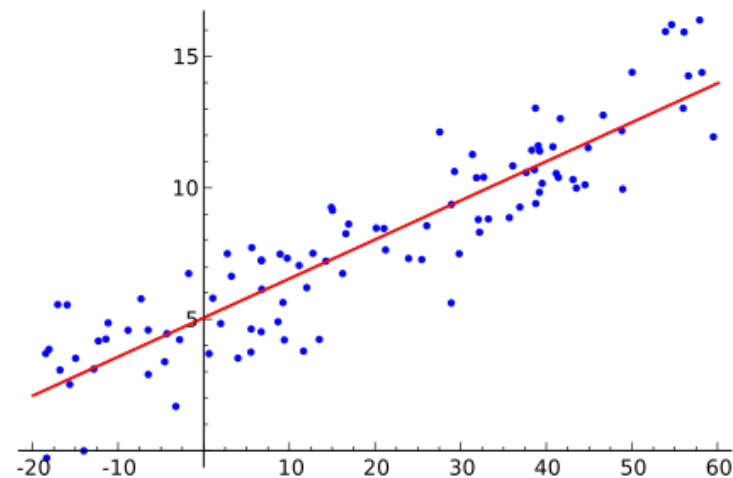


Regression



How to say if one target function is better?

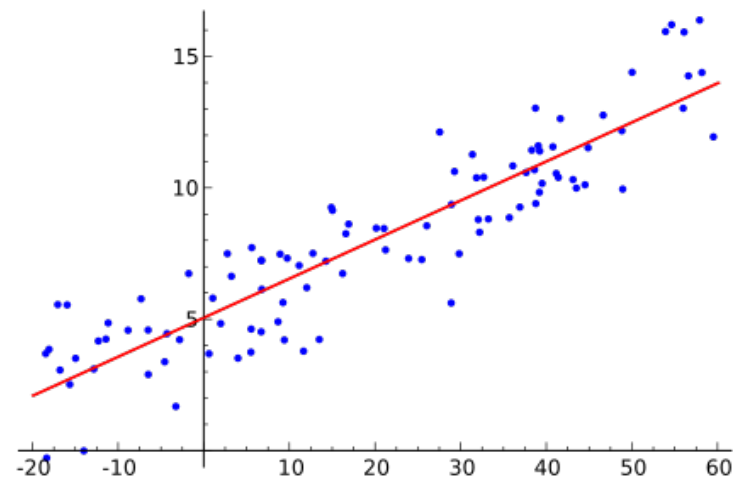
$$y_1^* = 5x + 6 \quad \text{or} \quad y_2^* = 4x^3 - 2x + 1$$



How to say if one target function is better?

$$y_1^* = 5x + 6 \quad \text{or} \quad y_2^* = 4x^3 - 2x + 1$$

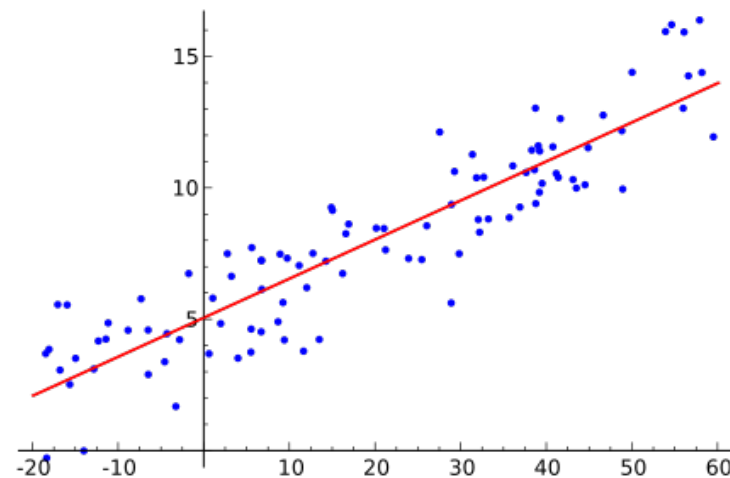
We need to compare them.



How to say if one target function is better?

$$y_1^* = 5x + 6 \quad \text{or} \quad y_2^* = 4x^3 - 2x + 1$$

We need to compare them.



$$L(\theta, x_i, y_i)$$

Loss Function

Popular loss functions:

Classification



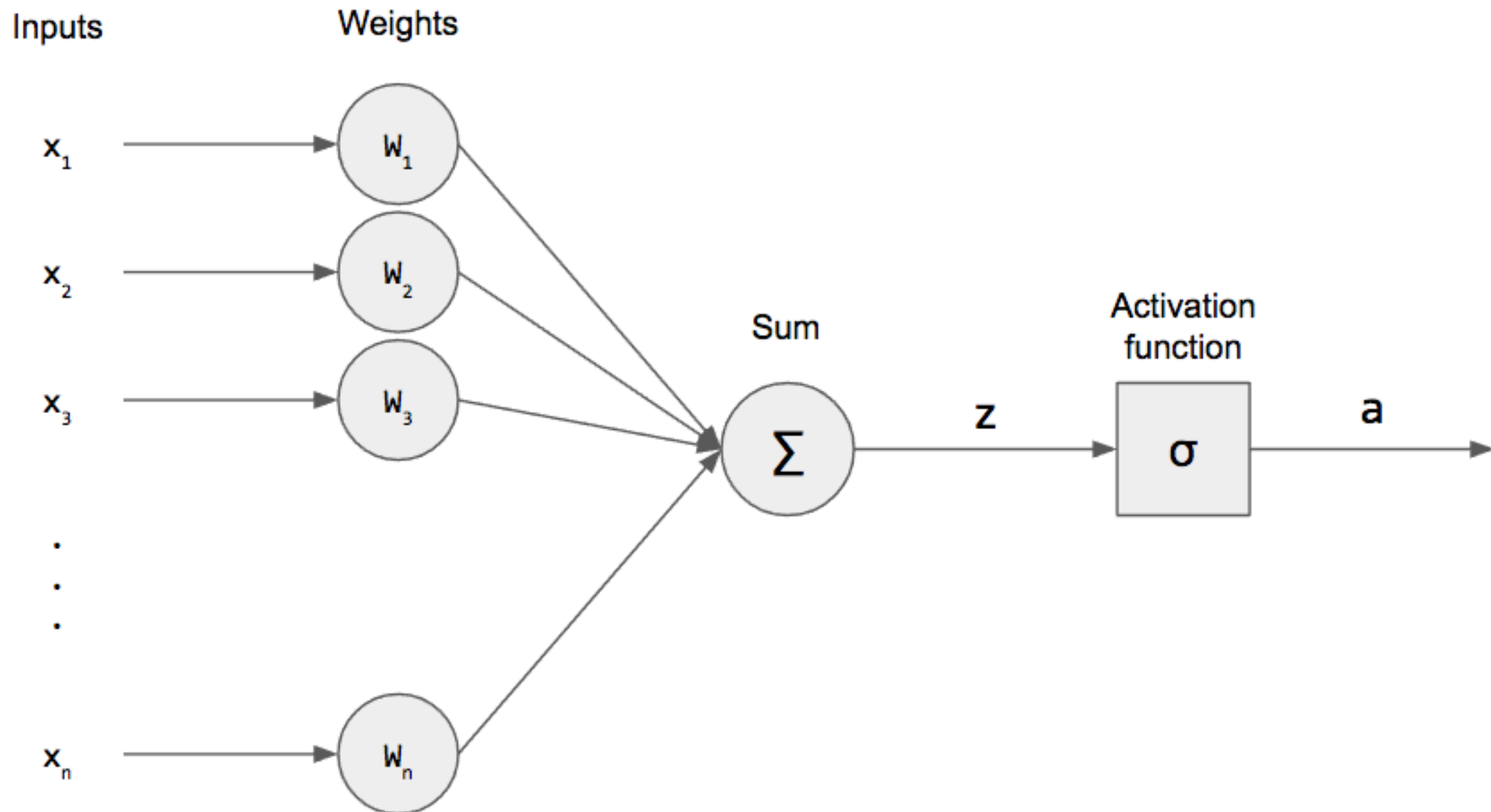
$$-(y_i \log(p_i) + (1 - y_i) \log(1 - p_i))$$

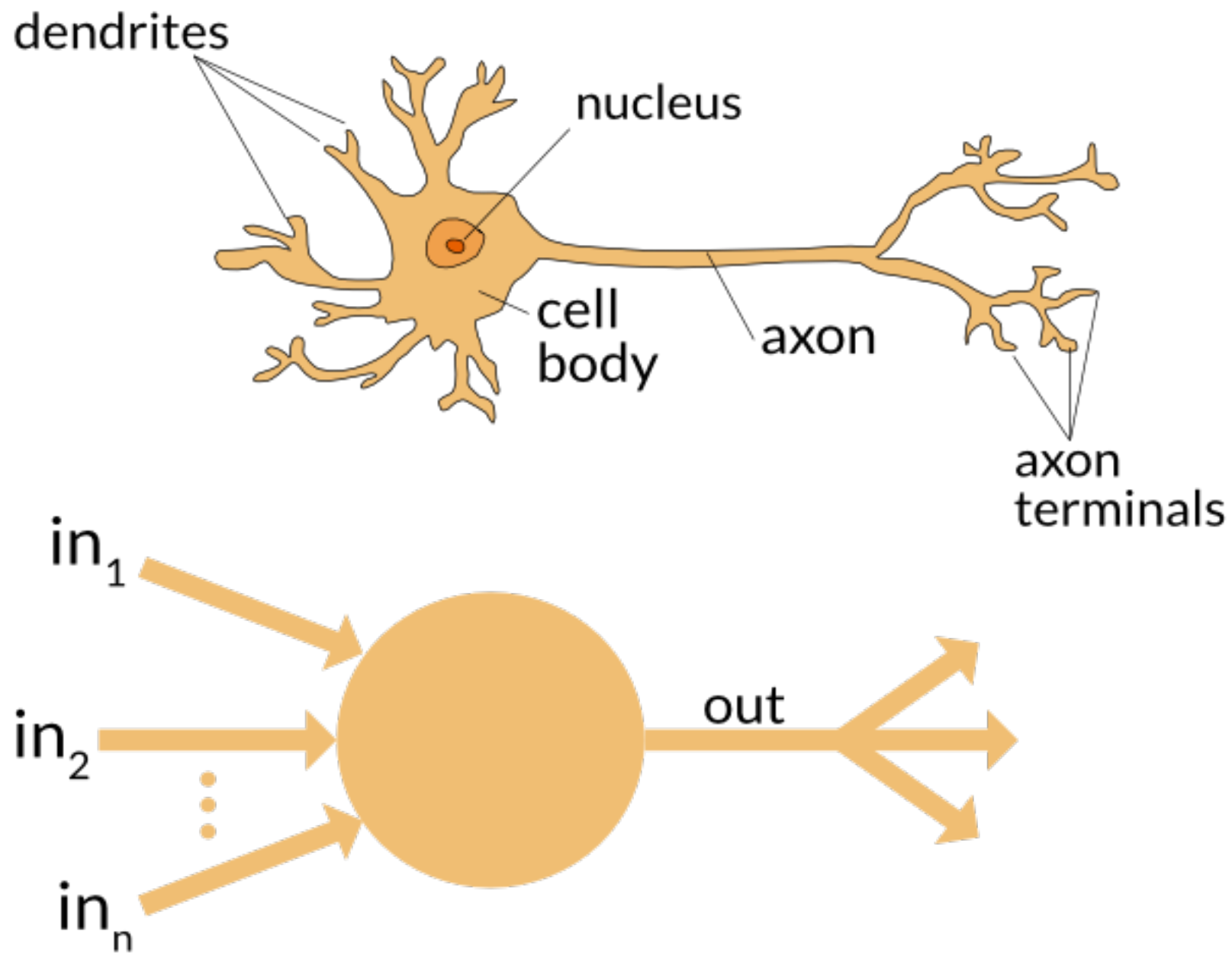
Regression



$$(y_i - p_i)^2$$

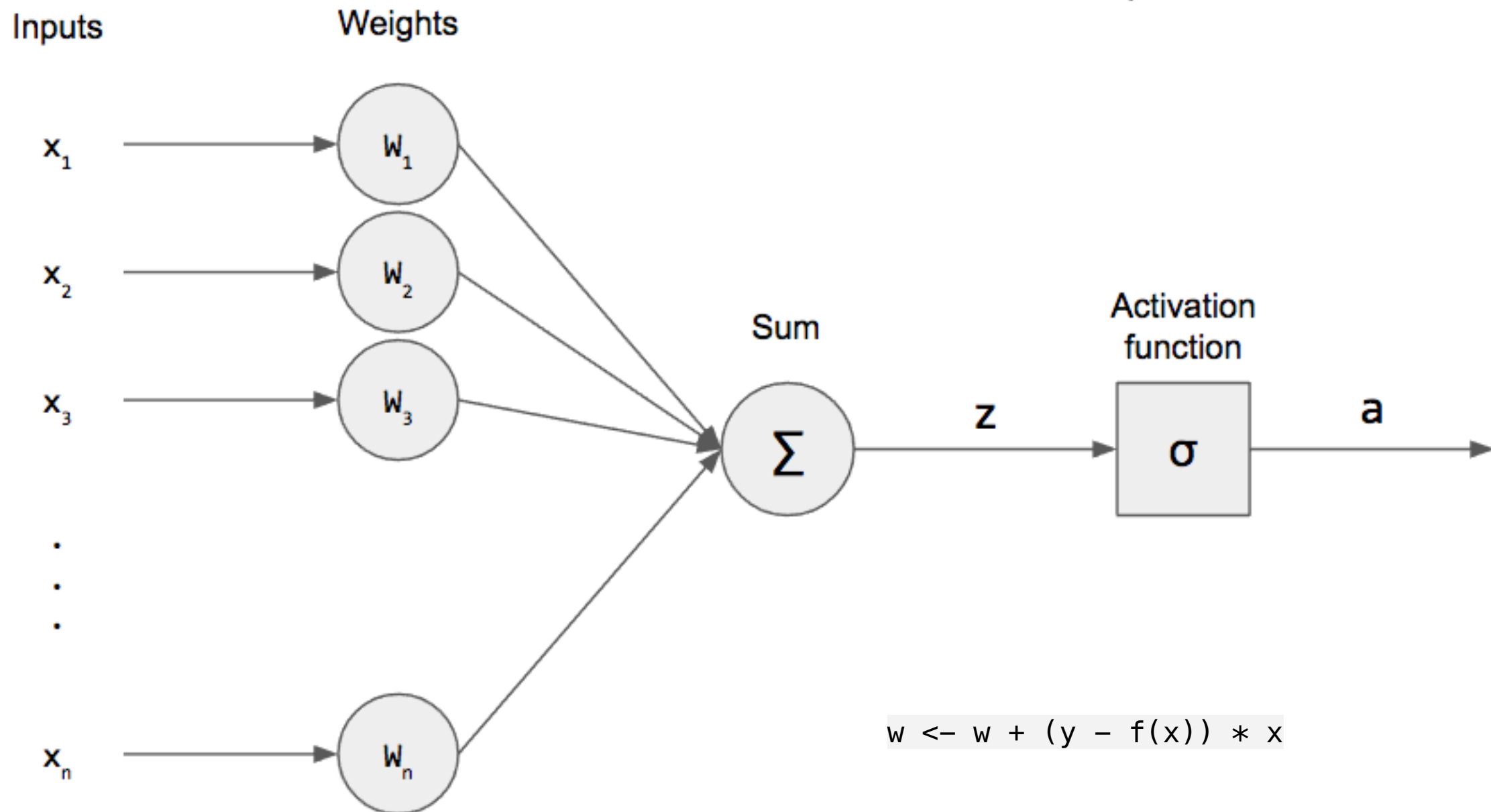
Perceptron:





Perceptron:

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



$$w \leftarrow w + (y - f(x)) * x$$

Perceptron:

