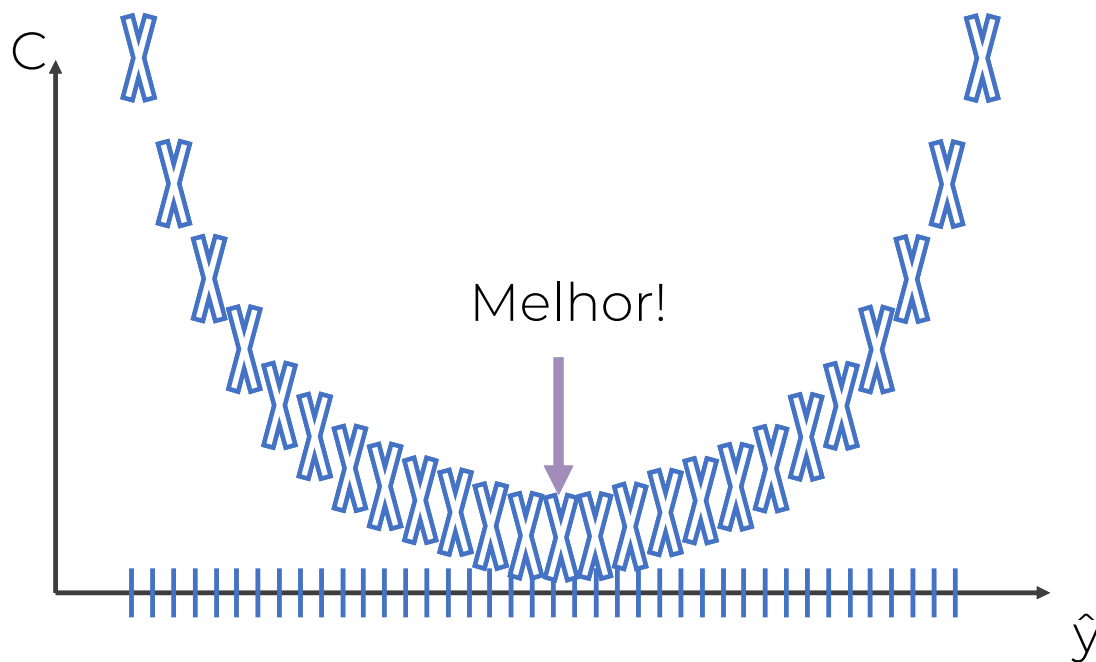


Ajustando os Pesos

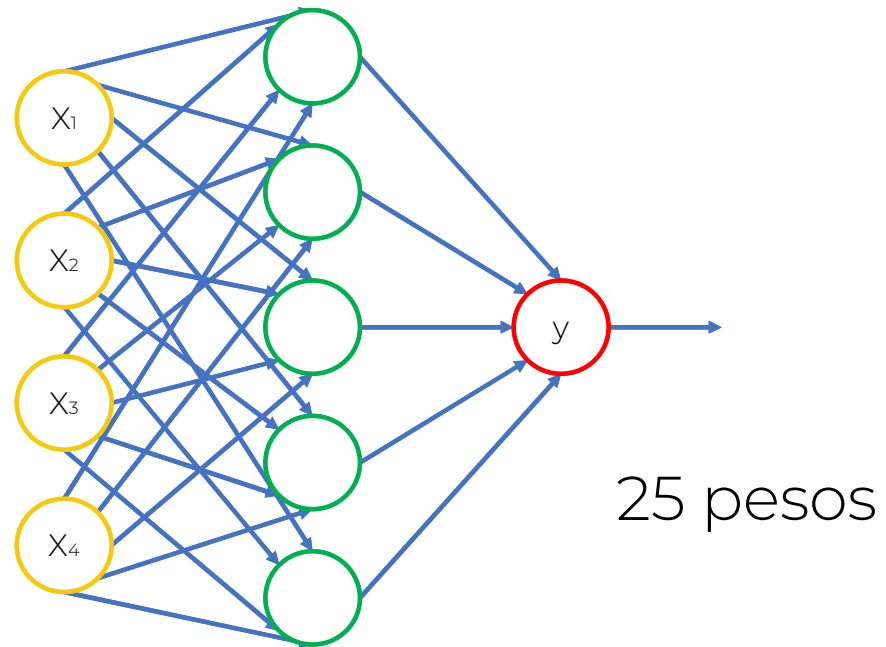
- Learning Rate: taxa de aprendizado
- Como ajustar os pesos? Aumentar, diminuir? Em quanto?
- Lembrando que podemos ter muitos pesos para ajustar!



Força Bruta...



Gradient Descent



Gradient Descent

$1.000 \times 1.000 \times \dots \times 1.000 = 1.000^{25} = 10^{75}$ combinações

Sunway TaihuLight: o super computador mais rápido do mundo

93 PFLOPS

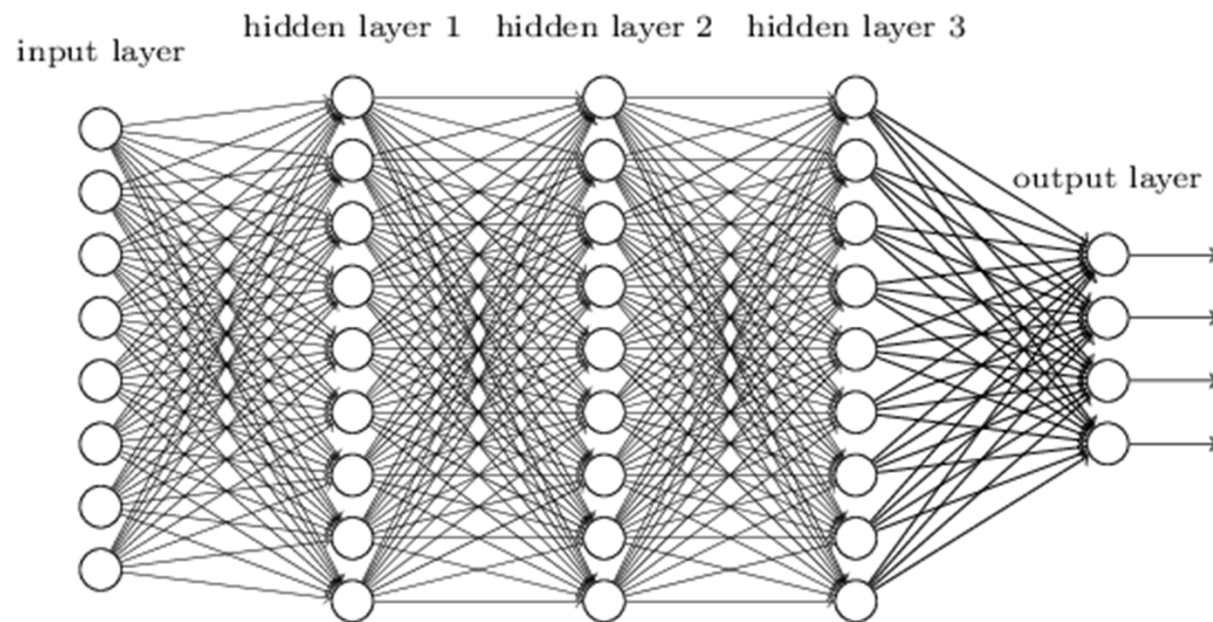
93×10^{15}

$10^{75} / (93 \times 10^{15})$

$= 1,08 \times 10^{58}$ segundos

$= 3,42 \times 10^{50}$ anos

Gradient Descent

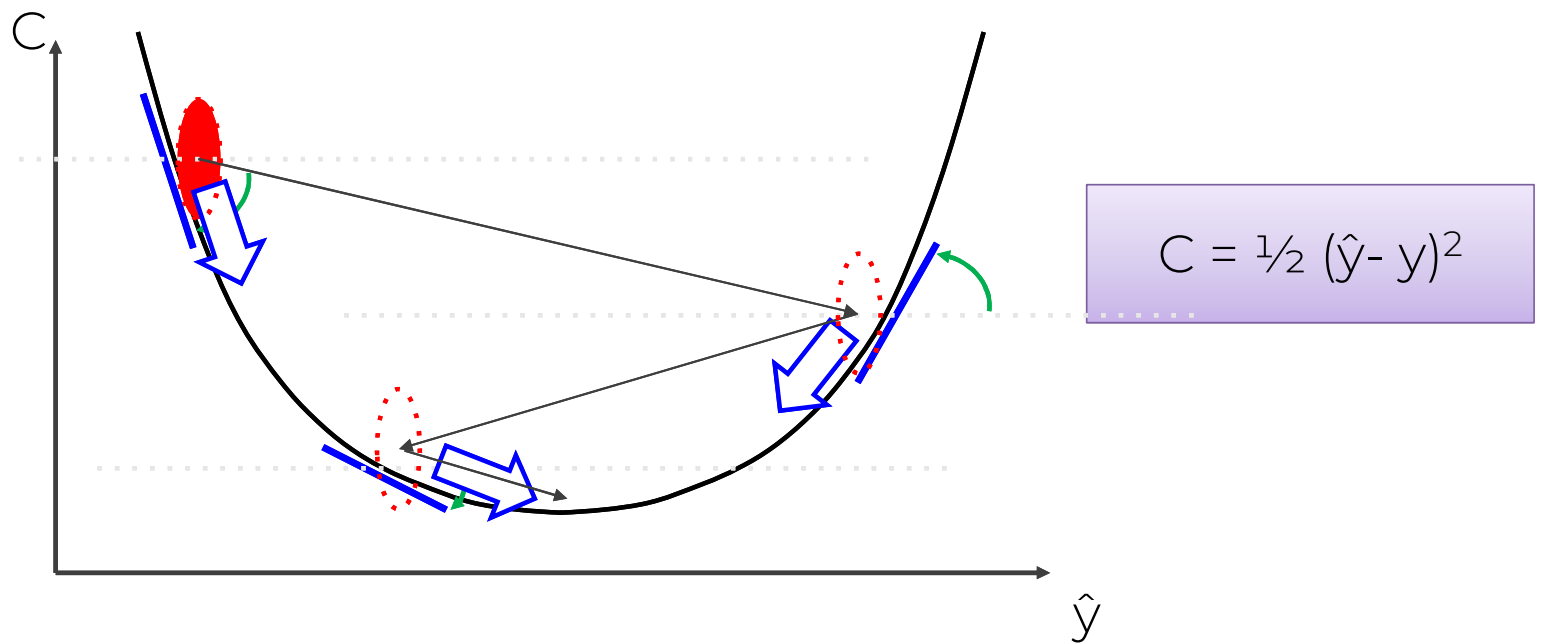


Fonte da imagem: neuralnetworksanddeeplearning.com

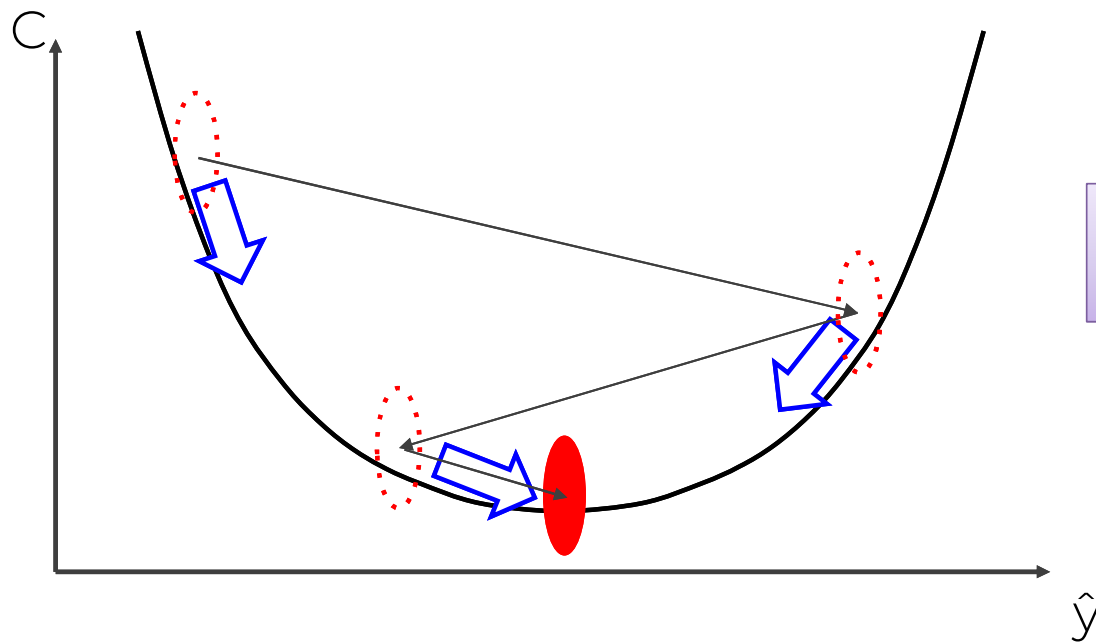
Gradient Descent

Gradient Descent

Gradient Descent



Gradient Descent



$$C = \frac{1}{2} (\hat{y} - y)^2$$

Stochastic Gradient Descent

