A Single Neuron

Outline

- A quick tour of machine learning
- A description of linear regression
- A description of binary classification
- Description of a single neuron

A Quick Tour of Machine Learning

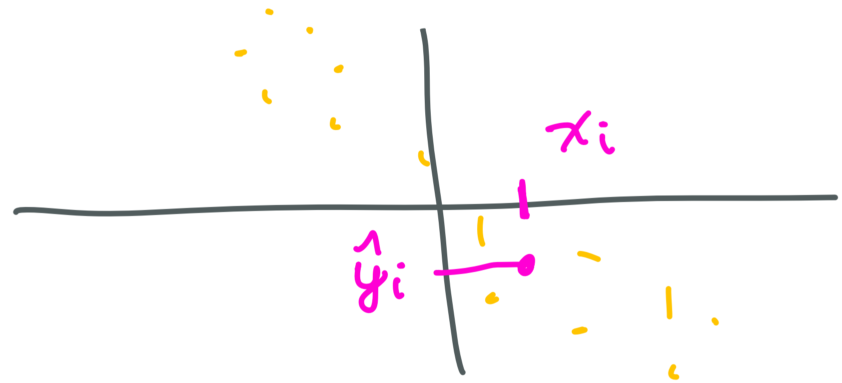
Machine learning – Wikipedia
(https://en.wikipedia.org/wiki/Machine_learning)

(https://blinpete.github.io/wiki-graph/?lang=en&wordle&query=Machine%20learning)

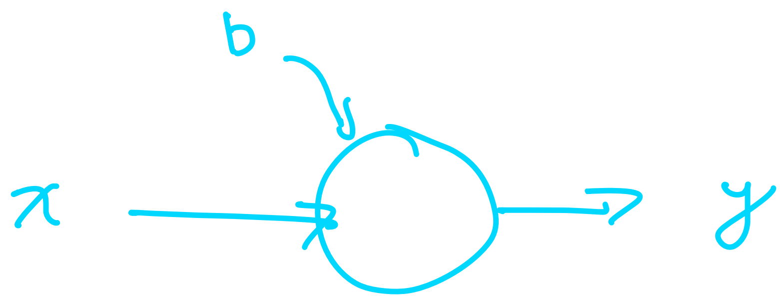
Linear Regression (Scalar Output)

$x_i = 0.2$

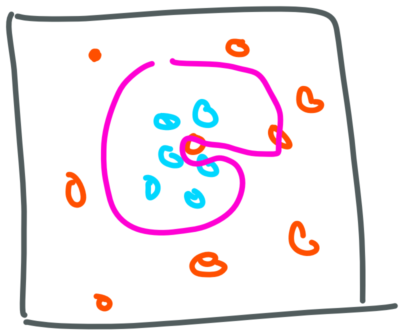
$\hat{y}_i$?



[Linear Regression - MIT Mathlets (https://mathlets.org/mathlets/linear-regression/)](https://mathlets.org/mathlets/linear-regression/)



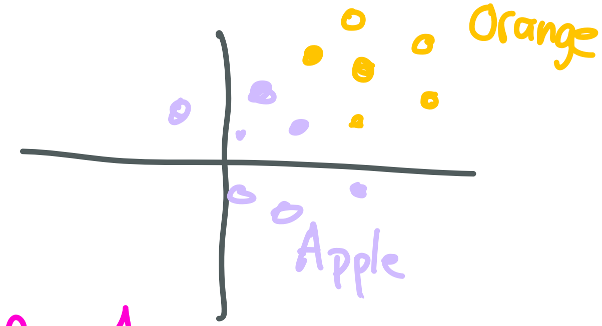
Classification

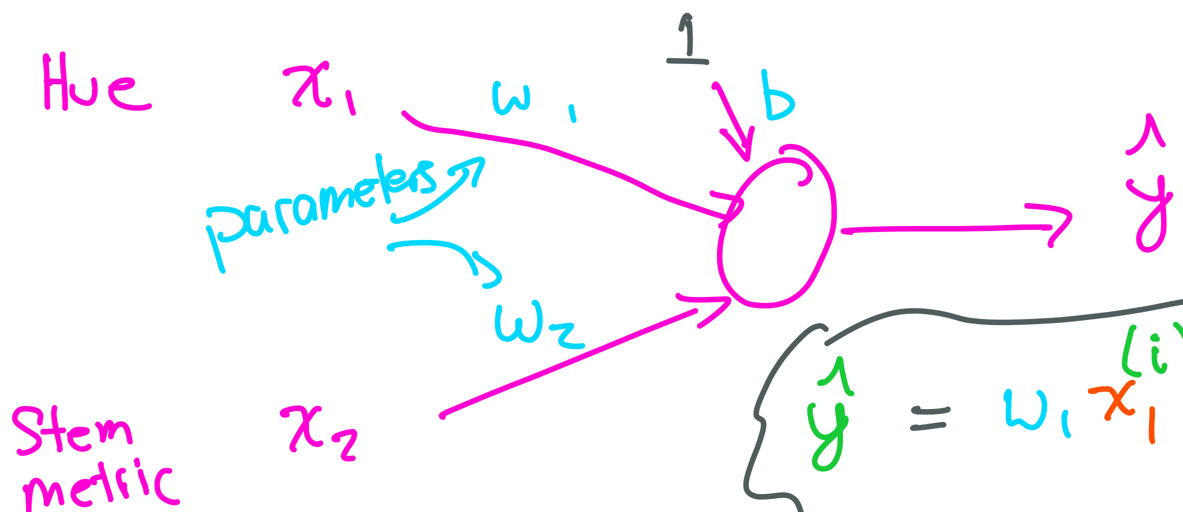


[A Neural Network Playground (https://playground.tensorflow.org)](https://playground.tensorflow.org)

A Single Neuron

- AKA perceptron or logistic-regression (more or less)
- What would you all like to use as an example?





if $\hat{y} > 0 \rightarrow$ apple

$\hat{y} \leq 0 \rightarrow$ orange

$$\hat{y} = w_1 x_1^{(i)} + w_2 x_2^{(i)} + b > 0$$

Model
↳ params
↳ arch

$$X = \begin{bmatrix} x_1^{(1)} & x_1^{(2)} & \cdots & x_1^{(N)} \\ x_2^{(1)} & x_2^{(2)} & \cdots & x_2^{(N)} \end{bmatrix}$$

$$\hat{Y} = \begin{bmatrix} \hat{y}^{(1)} & \hat{y}^{(2)} & \cdots & \hat{y}^{(N)} \end{bmatrix}$$

What do we optimize?

Supervised Learning

↓

$Y = \langle \text{correct values} \rangle$

$\hat{Y} = \langle \text{predictions} \rangle$

$$\hat{y} = w_1 x_1^{(i)} + w_2 x_2^{(i)} + b > 0$$

Optimize the parameter

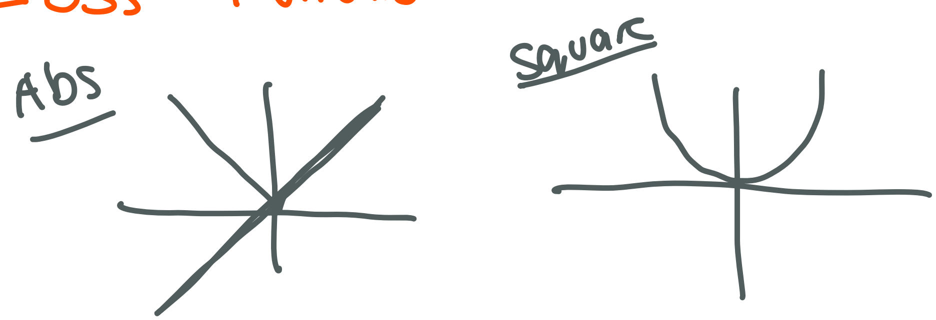
~~Objective Function~~

~~Cost Function~~

$\mathcal{L}$oss Function

$$\mathcal{L}(\hat{y}, y) = (\hat{y} - y)^2$$

How do I know if I should increase or decrease w_1 based on the given loss function?



Gradient Descent

$\mathcal{L}(\hat{y}, y) = (\hat{y} - y)^2$ ‖ $\hat{y} = w_1 x_1^{(i)} + w_2 x_2^{(i)} + b > 0$

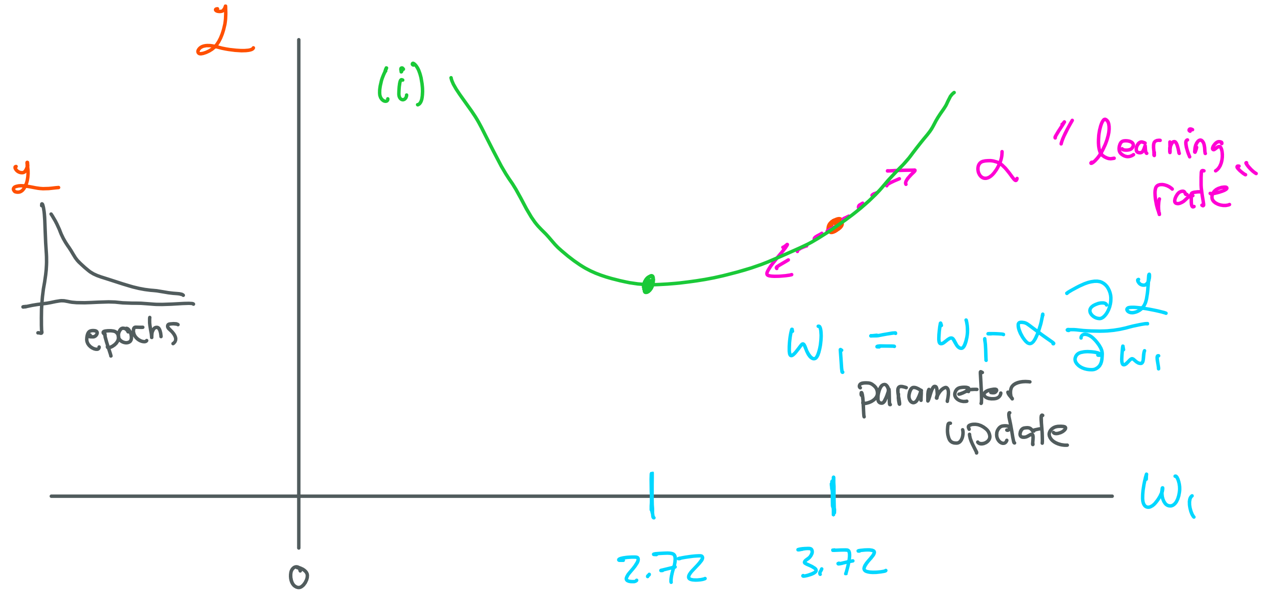
$$\frac{\partial \mathcal{L}}{\partial w_1} = \frac{\partial}{\partial w_1} (\hat{y} - y)^2$$ <power rule>

$$= 2(\hat{y} - y) \frac{\partial}{\partial w_1} (w_1 x_1^{(i)} + w_2 x_2^{(i)} + b - y)$$

$$= 2(\hat{y} - y) x_1^{(i)}$$

$\frac{\partial y}{\partial w_1} =$

$\frac{\partial \mathcal{L}}{\partial y} \cdot \frac{\partial y}{\partial w_1}$



$$\hat{y} = w_1 x_1^{(i)} + w_2 x_2^{(i)} + b$$