

Redes neuronales

ESCUELA DE VERANO DE APRENDIZAJE PROFUNDO 2019

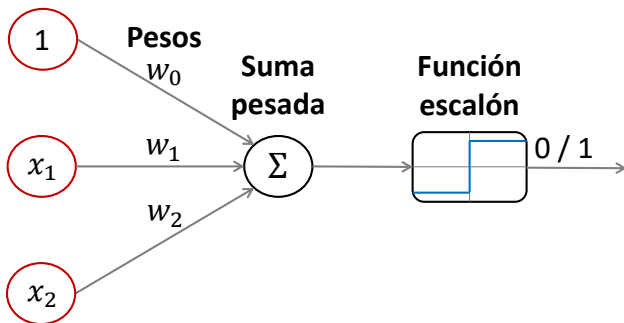
Gibran Fuentes-Pineda

IIMAS, UNAM

Junio 24, 2019

Neurona artificial

Entradas



- Modelo

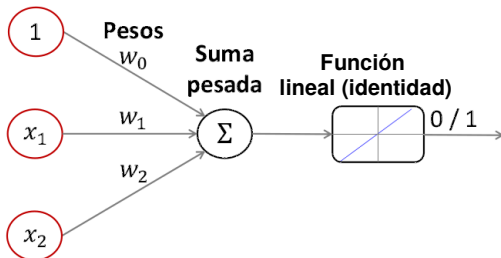
1. Pesos sinápticos
2. Estímulo cumulativo
3. Todo o nada (activación)

- Salida

$$s = \phi\left(\sum_{i=1}^m w_i \cdot x_i\right)$$

Función de activación lineal o identidad

Entradas



$$\begin{aligned} \text{lineal}(z) &= z \\ \frac{d\text{lineal}(z)}{dz} &= 1 \end{aligned}$$

Prediciendo el precio de casas

- Podemos emplear neurona artificial con función de activación lineal o identidad para problemas de regresión

Tamaño (m^2)	Precio (USD)
433.71	61590.60
446.84	69741.66
582.16	56292.42
833.68	113220.61
920.39	145932.16
931.33	126830.14
1326.25	171278.58
⋮	⋮

Característica o
atributo

Respuesta

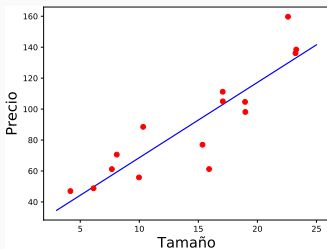


Tomada de <https://en.wikipedia.org/wiki/House>

Prediciendo el precio de casas

- La neurona con activación lineal calcula

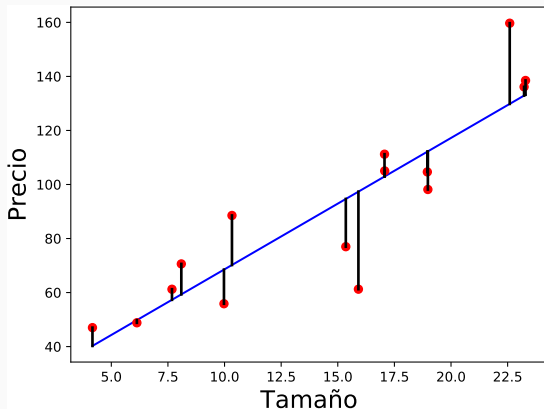
$$\hat{y} = b + w_1 \cdot x$$



Función de pérdida

- Error cuadráticos medio

$$ECM(y, \hat{y}) = \frac{1}{N} \sum_{i=1}^N (y - \hat{y})^2$$



Entrenando el modelo: algoritmo del gradiente descendente

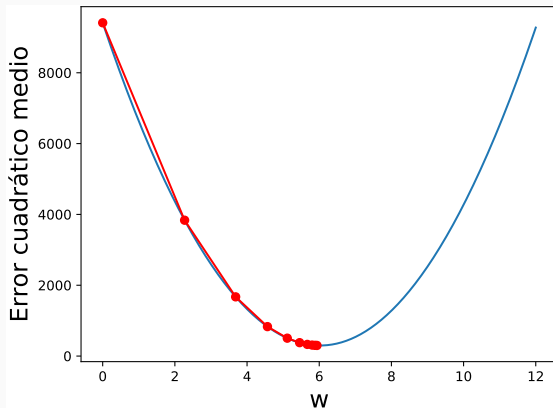
- Mueve sucesivamente pesos y sesgo ($\theta = \{\mathbf{w}, \mathbf{b}\}$) hacia donde más descienda la pérdida

$$\theta^{(k+1)} = \theta^{(k)} - \alpha \mathbf{g}(\theta^{(k)})$$

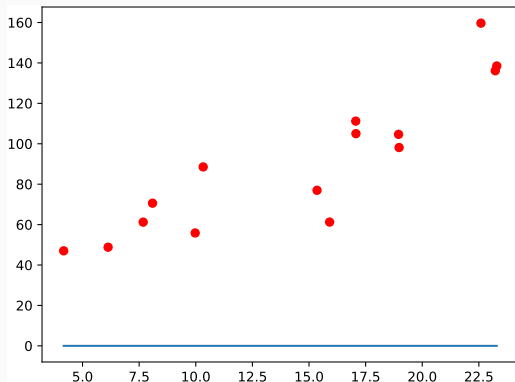
$$\mathbf{g}(\theta) = \frac{d}{d\theta} E = \mathbf{X}^T (\hat{\mathbf{y}} - \mathbf{y})$$

- Gradiente descendente estocástico: calcula gradiente con lote pequeño de datos

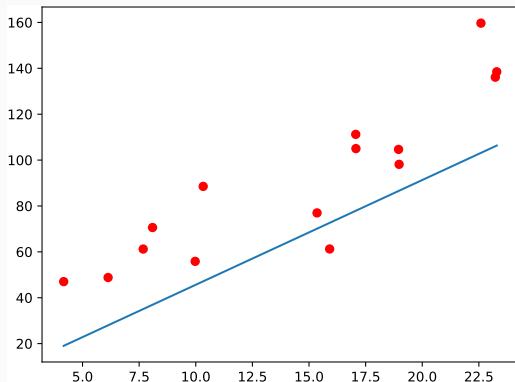
Entrenando el modelo: algoritmo del gradiente descendente



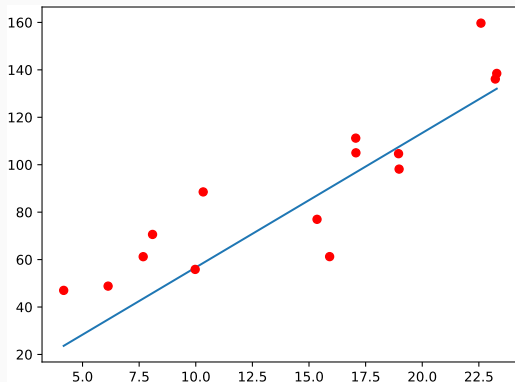
Evolución de modelo: época 0



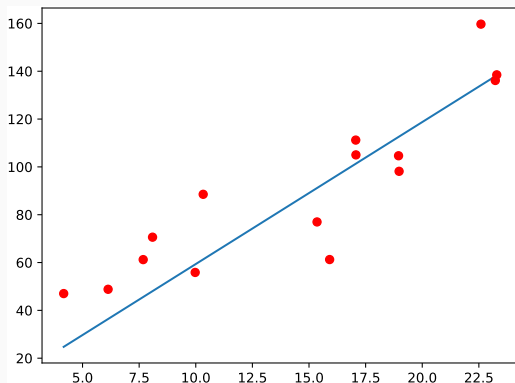
Evolución de modelo: época 3



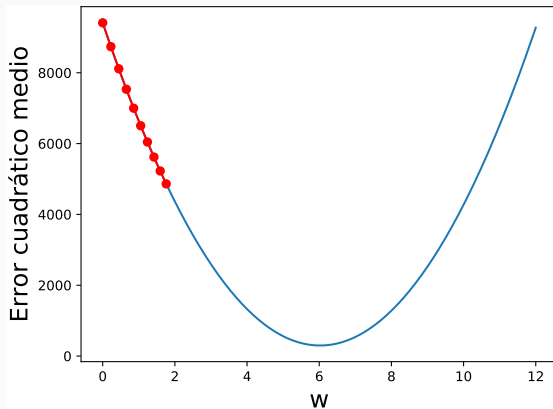
Evolución de modelo: época 6



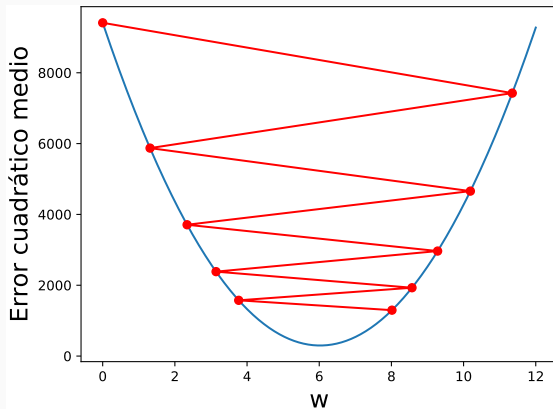
Evolución de modelo: época 10



Sensibilidad a coeficiente de aprendizaje α

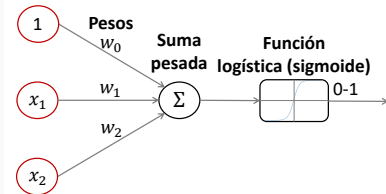


Sensibilidad a coeficiente de aprendizaje α



Función de activación sigmoide o logística

Entradas



$$\text{sigm}(z) = \frac{1}{1 + \exp(-z)}$$
$$\frac{d\text{sigm}(z)}{dz} = \text{sigm}(z)(1 - \text{sigm}(z))$$

- Función de pérdida se conoce como *entropía cruzada binaria*

$$ECB = - \sum_{i=1}^N y^{(i)} \log \hat{y}^{(i)} + (1 - y^{(i)}) \log(1 - \hat{y}^{(i)})$$

Ejemplo de clasificación

- Clasificar sub-especies de la flor Iris basado en el ancho y largo de su pétalo

Ancho	Largo	Especie
1.4	0.2	Setosa
1.7	0.4	Setosa
1.5	0.1	Setosa
⋮	⋮	⋮
4.7	1.4	Versicolor
4.5	1.5	Versicolor
3.3	1.0	Versicolor
⋮	⋮	⋮

Características o
atributo

Respuesta

Setosa

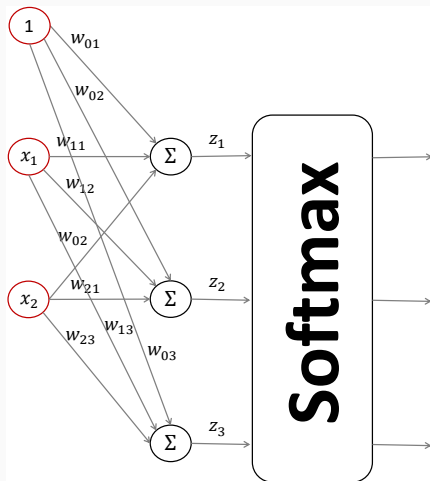


Versicolor



Tomada de https://en.wikipedia.org/wiki/Iris_flower_data_set

Clasificación Softmax (1)



Clasificación Softmax (2)

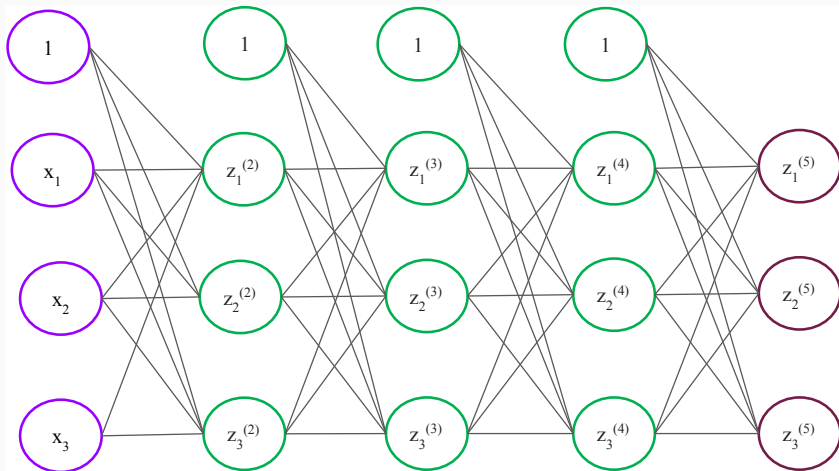
- Salidas pasan por una función Softmax dada por

$$\text{Softmax}(\mathbf{z})_i = \frac{e^{z_i}}{\sum_j e^{z_j}}$$

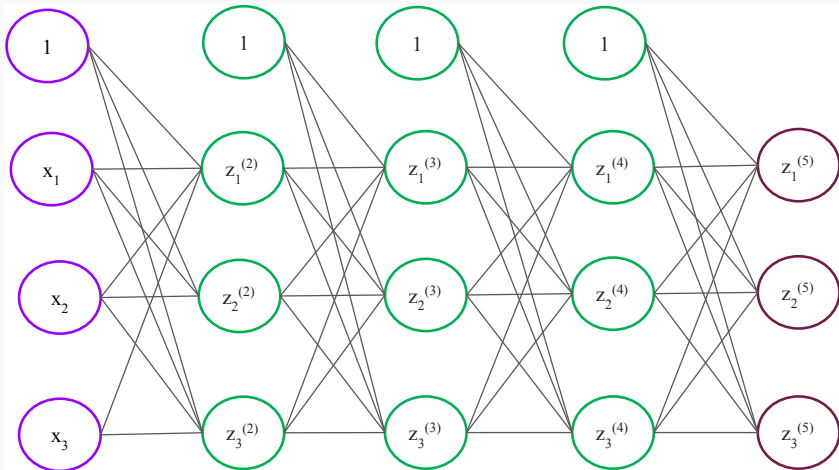
- Función de pérdida de clasificación multiclase: entropía cruzada categórica

$$E(y, \hat{y}) = - \left[\sum_{i=1}^N \sum_{k=1}^K 1\{y^{(i)} = k\} \log \frac{e^{z_k}}{\sum_j e^{z_j}} \right]$$

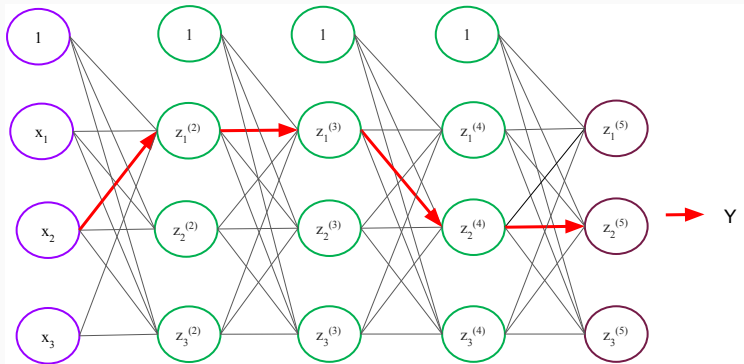
Red neuronal densa



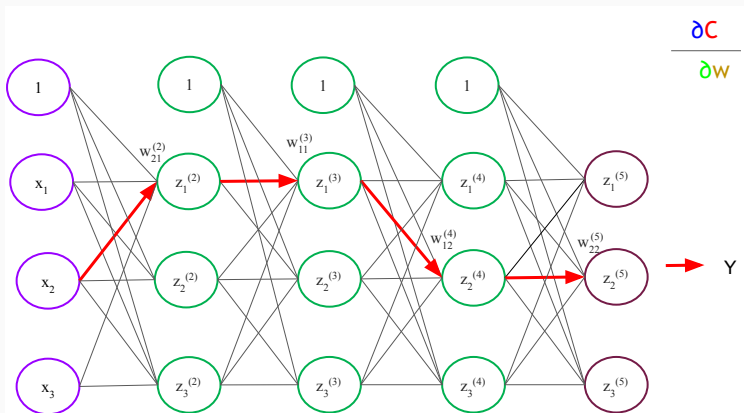
Propagación hacia adelante



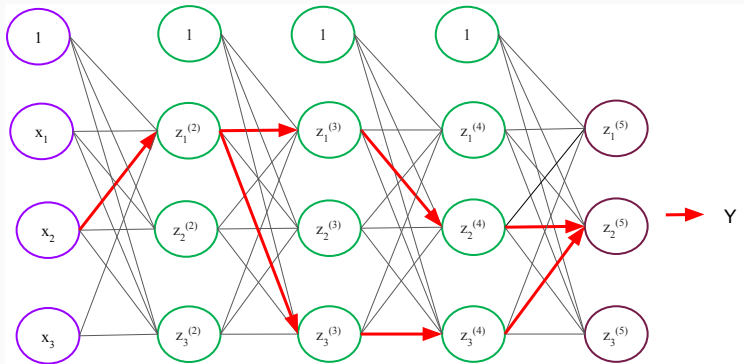
Entrenando por gradiente descendente



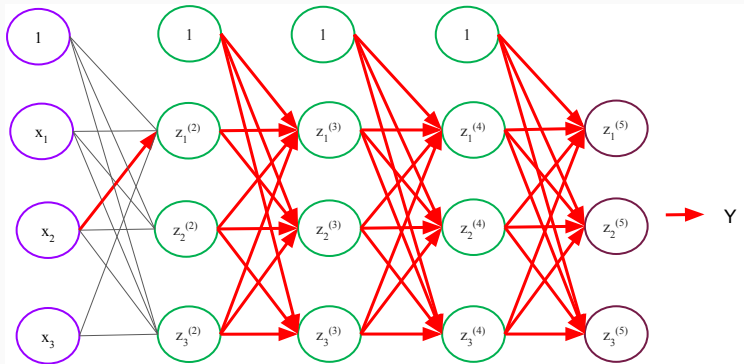
Entrenando por gradiente descendente



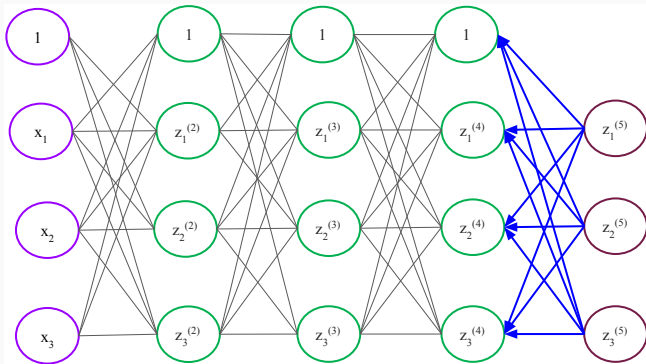
Entrenando por gradiente descendente



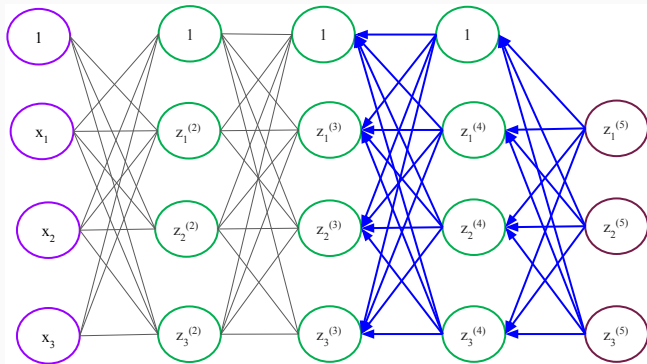
Entrenando por gradiente descendente



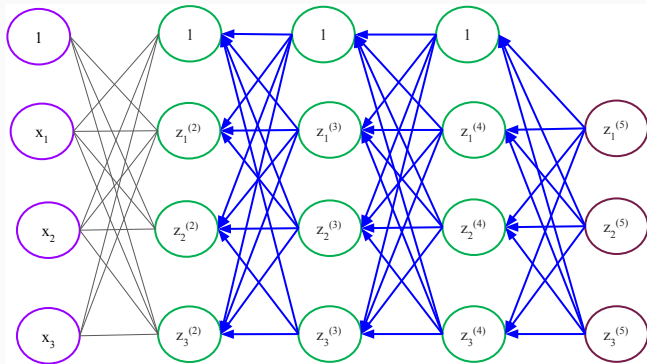
Algoritmo de retropropagación



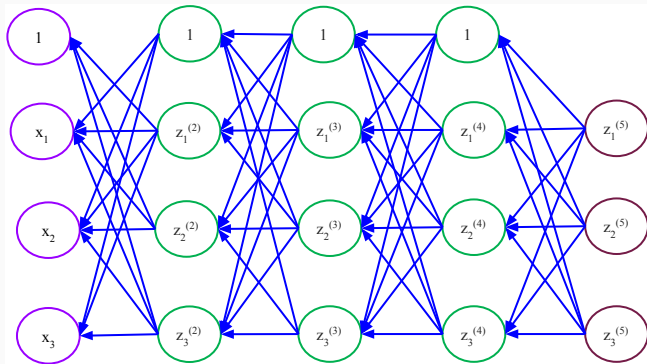
Algoritmo de retropropagación



Algoritmo de retropropagación

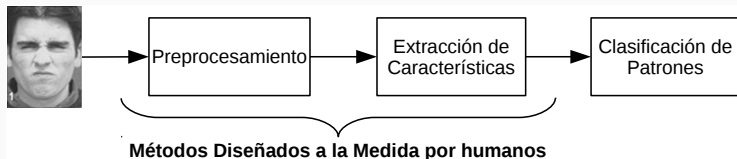


Algoritmo de retropropagación

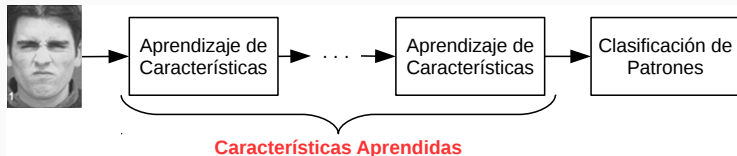


Aprendizaje de características

Sistema estándar



Con aprendizaje de características



Jerarquías de características

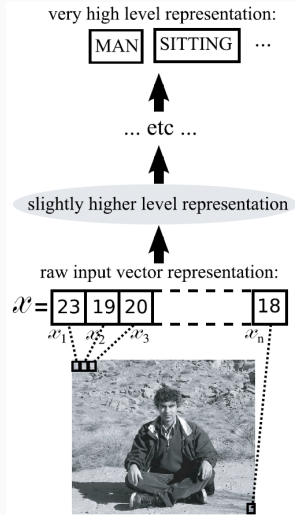


Figura tomada de Bengio, 2009

La importancia de los GPUs

IMAGENET

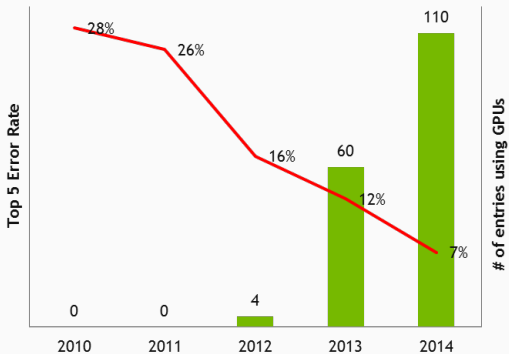


Figura tomada de <https://devblogs.nvidia.com/parallelforall>

Aprendizaje profundo y datos masivos

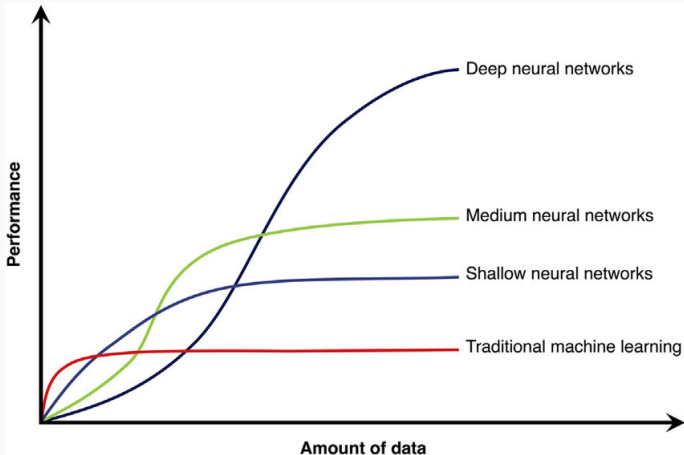


Figura tomada de Tang et al., Canadian Association of Radiologists Journal 69(2), 2018

Marcos de trabajo

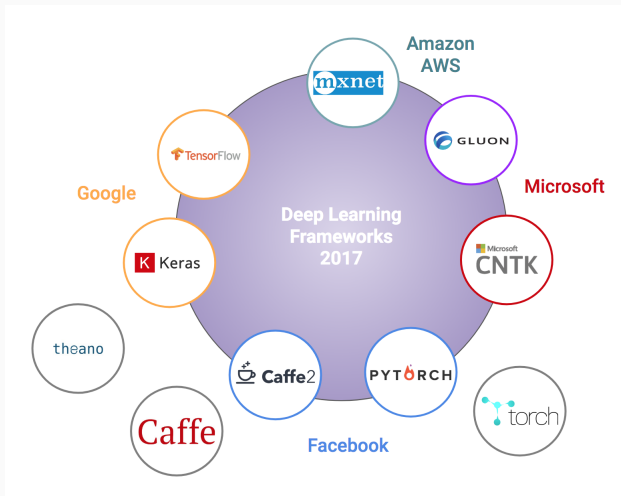


Figura tomada de <https://towardsdatascience.com/battle-of-the-deep-learning-frameworks-part-i-cff0e3841750>