

Fraud Detection Using Deep Learning

Rubén Martínez Sánchez



CCCCCCCCCCCCCCCCCCCC



CyberCamp.es

whoami



Rubén Martínez Sánchez (@eldarsilver)

- Ingeniero Informático Universidad Politécnica Madrid
- Máster Data Science Datahack
- Certified Etical Hacker (CEH)
- Cloudera Developer Training for Apache Spark
- Cloudera Developer Training for Apache Hadoop
- Título Desarrollo UML y Java por la UPM
- Miembro de MundoHacker
- Data Scientist desarrollando chatbots mediante Deep Learning



Agenda# Is()



- Te interesa el Deep Learning
- Ciclo de vida del Machine Learning
- Introducción a las Redes Neuronales
- Deep Learning Supervisado
- Demo: Detección de fraude con Deep Learning Supervisado
- Bonus. Deep Learning No Supervisado: Autoencoders
- Bonus. Demo: Detección de fraude con Autoencoders

The Best Jobs of 2016: 1. Data Scientist

2016 Jobs Rated Score: 91

Annual Median Salary: \$128,240

Growth Outlook: 16%

Opportunities across a variety of fields make data scientist not just a high-growth job, but also one of the most lucrative tracked by the Jobs Rated report.

Scientists say deep learning is the future of Silicon Valley tech

CHRISTINA FARR NOVEMBER 25, 2012 10:42 AM

TAGS: ARTIFICIAL INTELLIGENCE, DEEP LEARNING, FACIAL RECOGNITION SOFTWARE, FEATURED, FUTURISTIC TECH, IMAGE RECOGNITION, MACHINE LEARNING, ROBOTICS, SILICON VALLEY, SILICON VALLEY TECH, SIRI, VOICE RECOGNITION



it may sound like something out of a dystopian novel, but scientists are confident about a machine learning technology that can recognize and replicate human activities like seeing and thinking.



Press Releases

Steve Altman Joins Enevate's Advisory Board

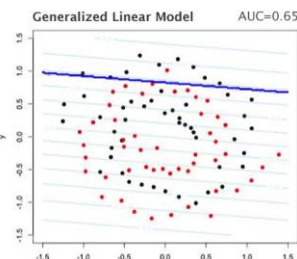
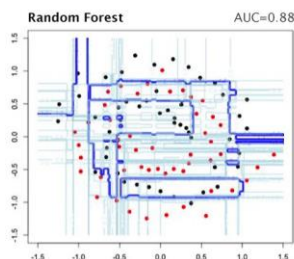
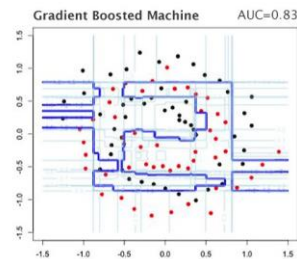
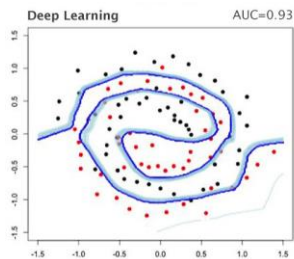
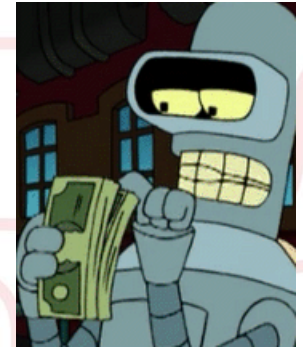
Virtual Reality and Augmented Reality Leader Marzent Secures \$5 Million



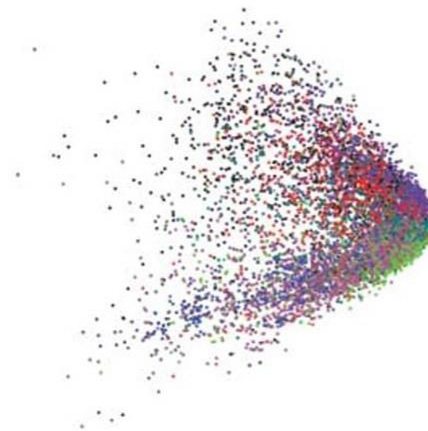
Te interesa el Deep Learning

Algunas de sus capacidades:

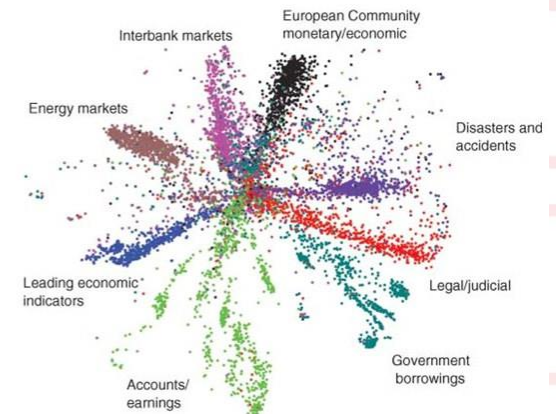
- Ayuda a tomar mejores decisiones ----->
- Es capaz de detectar patrones complejos no lineales
- Aprende conceptos abstractos
- Aprende la importancia de las variables automáticamente



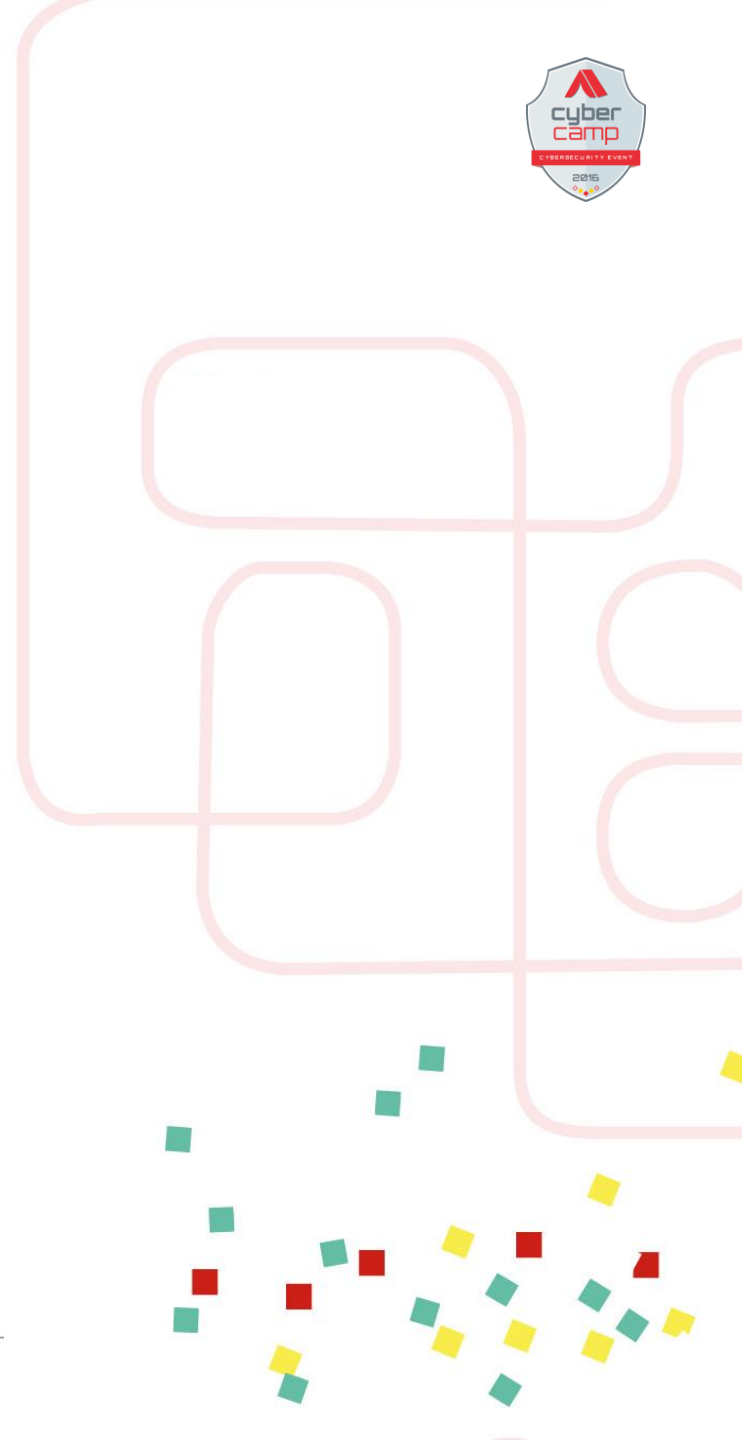
Principal Component Analysis



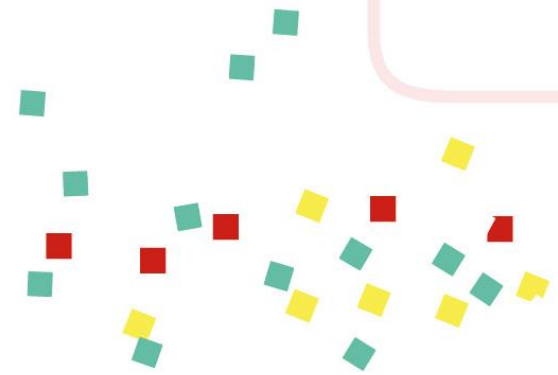
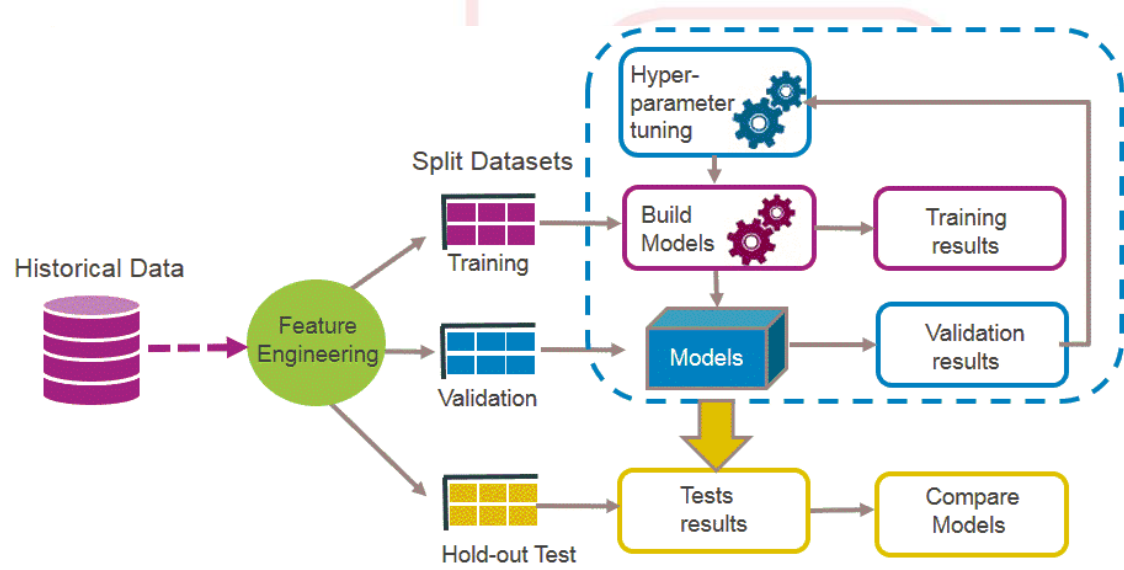
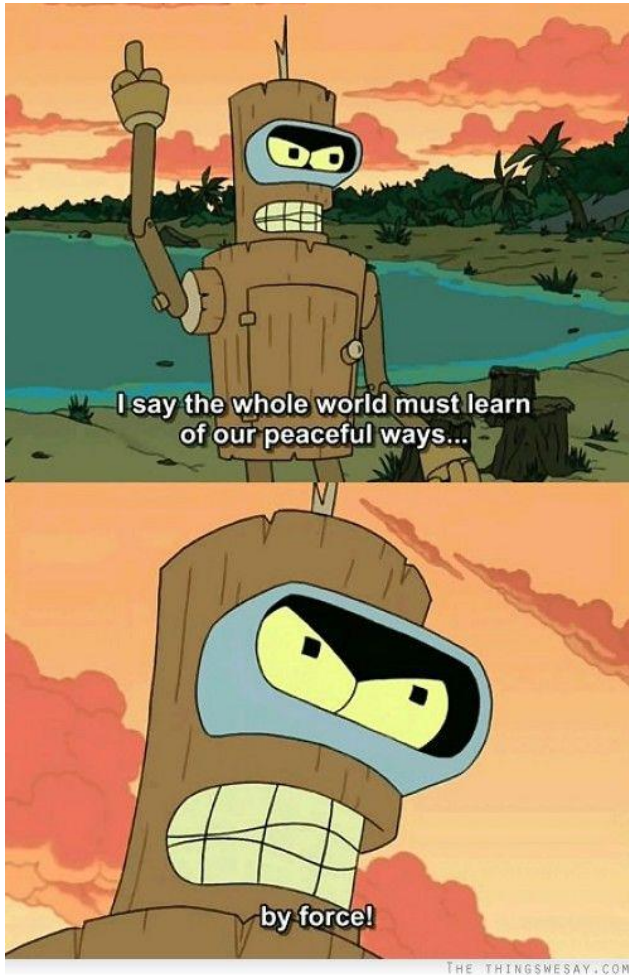
Deep Autoencoder



Te interesa el Deep Learning

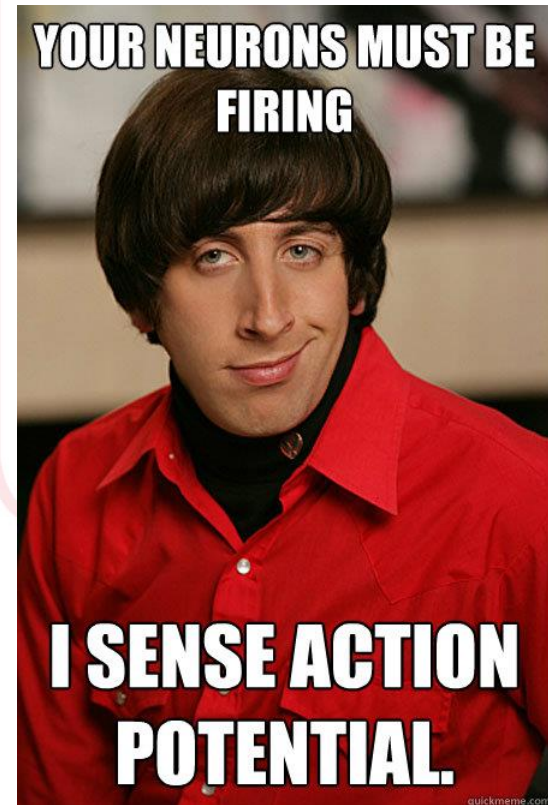
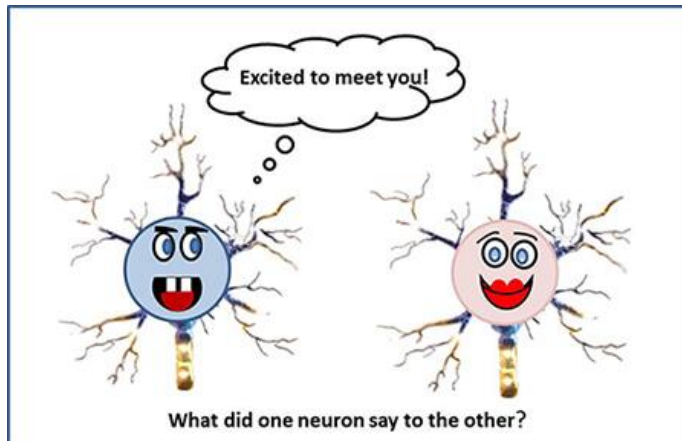
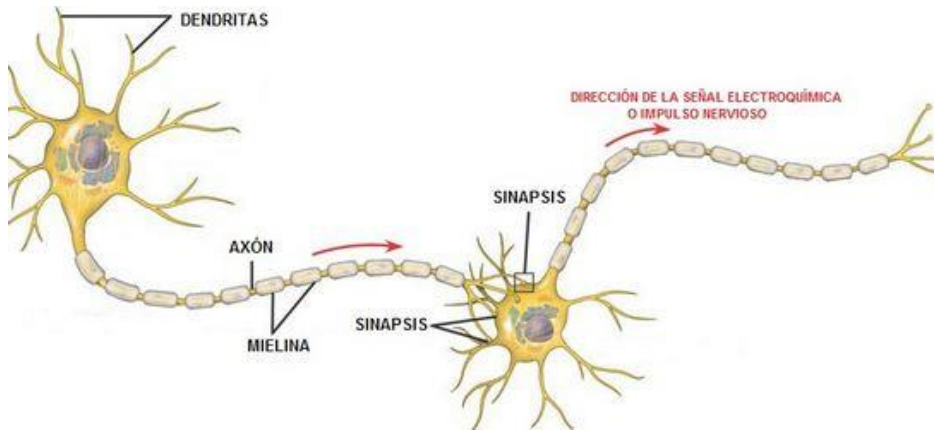


Ciclo de vida del Machine Learning



Introducción a las redes neuronales

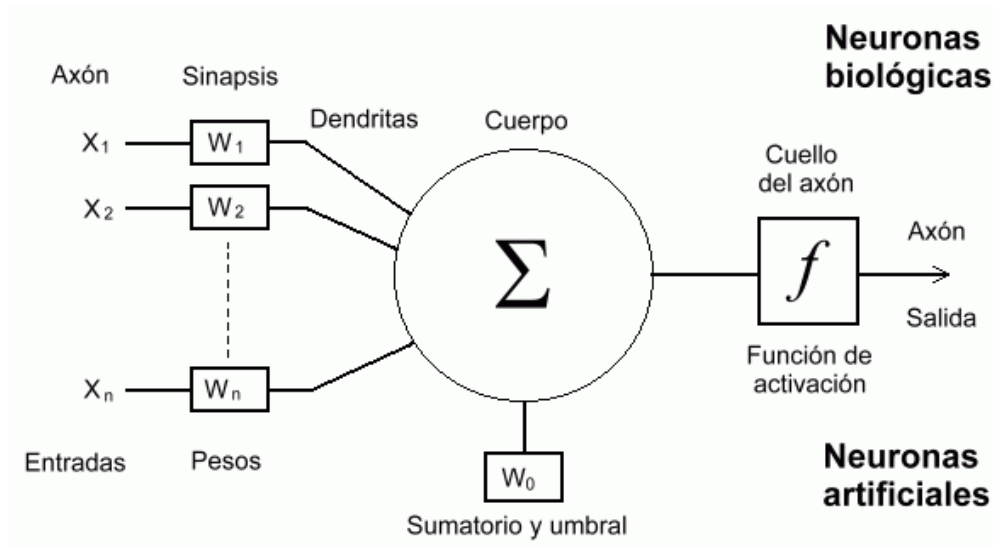
Neurona biológica



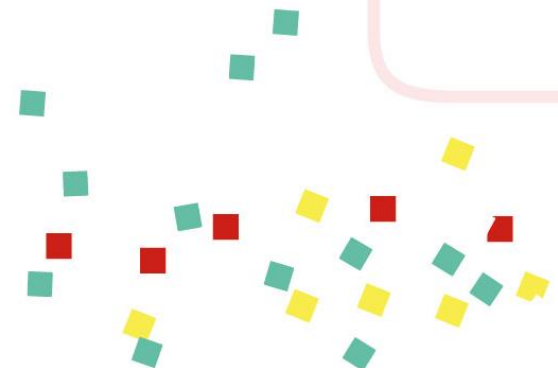
Introducción a las redes neuronales



Neurona Artificial (Perceptrón simple)



$$\sum_i w_i x_i + b$$

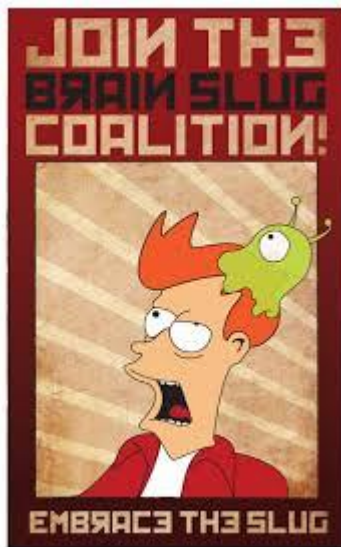


Introducción a las redes neuronales

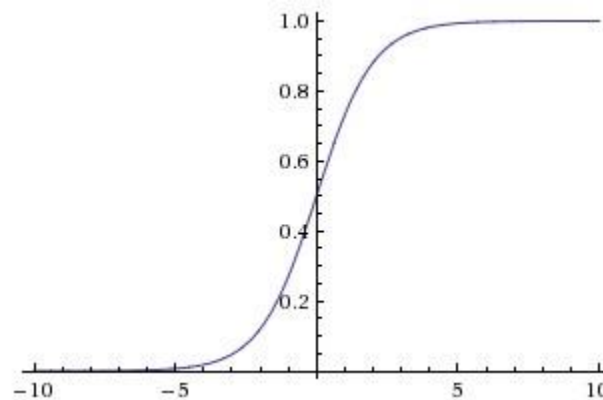


Función de activación Sigmoidal

- Acota un número real entre 0 y 1



$$\sigma(x) = 1 / (1 + e^{-x})$$



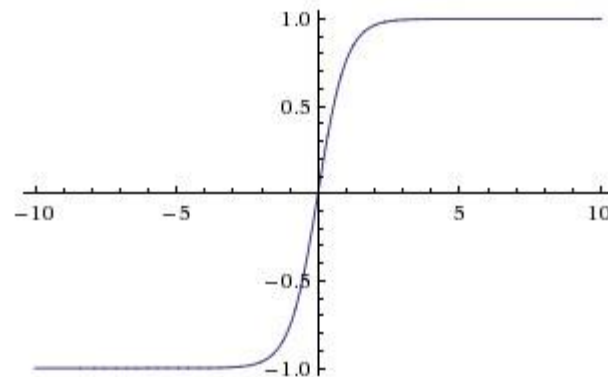
Introducción a las redes neuronales



Función de activación Tanh

- Acota un número real entre -1 y 1

$$\tanh(x) = 2\sigma(2x) - 1$$

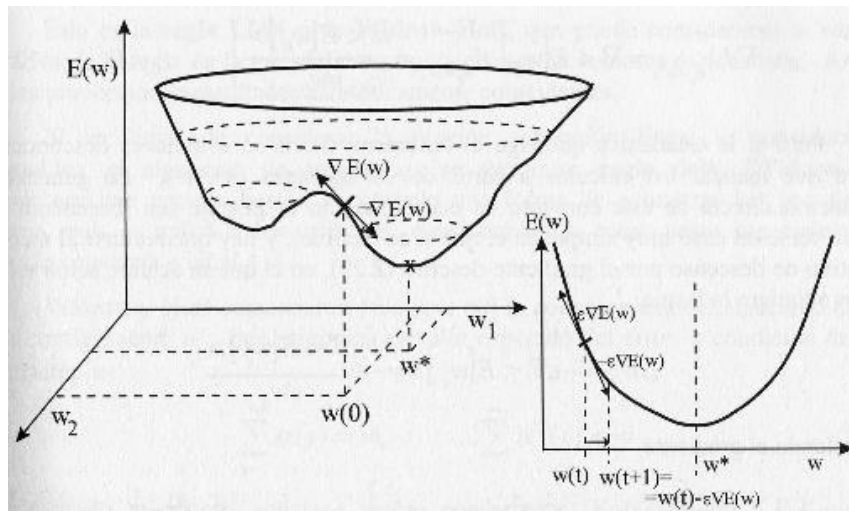


Introducción a las redes neuronales



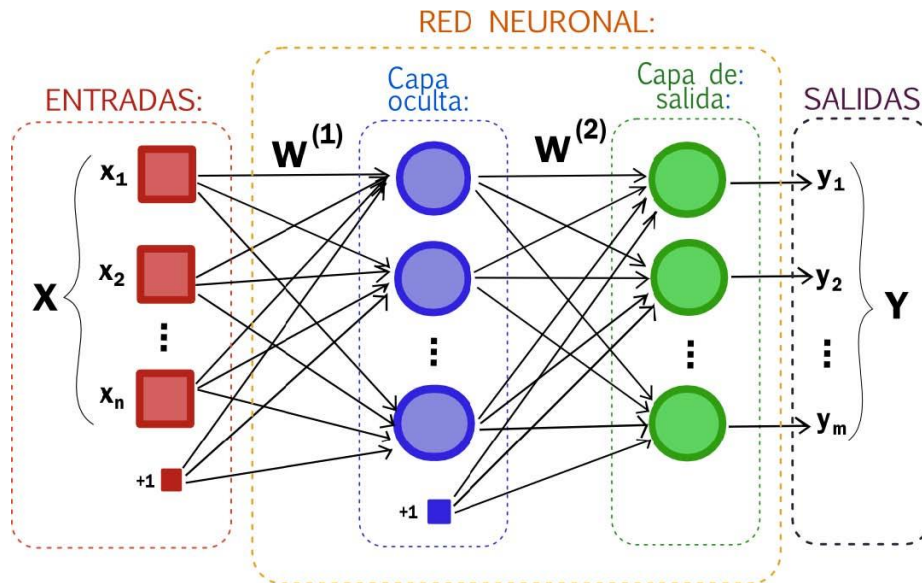
Aprendizaje por Descenso de Gradiente

- Aprendizaje de los pesos (w) mediante la optimización de una función de coste $E(w)$
 - Método de optimización: Descenso de Gradiente del error cometido actualmente
1. Inicio en $t=0$ con una cierta configuración de pesos
 2. Determinar la dirección de la máxima variación de $E(w)$ en w_0 , que vendrá dada por su gradiente para w_0
 3. Actualizar los pesos w siguiendo el sentido opuesto al indicado por el gradiente de $E(w)$
 4. Volver a realizar los pasos 2 y 3 hasta encontrar un mínimo local



Deep Learning Supervisado

Red Neuronal Artificial (Simple vs Deep)



Deep Learning Supervisado

Entrenamiento (Backpropagation I)

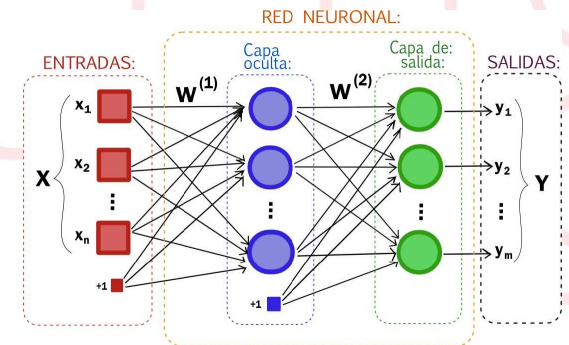
- Inicializar los pesos con valores pequeños aleatorios
- Dar patrón de entrada y proporcionar la salida deseada
- Calcular la salida de la red para esa entrada

- Valor de salida de neurona oculta

• Valor de salida de neurona oculta $sum_{pj}^h = \sum_{i=1}^N w_{ij}^h x_{pi} + \theta_j^h$ $y_{pj} = f_j^h (sum_{pj}^h)$

$$sum_{pk}^0 = \sum_{j=0}^L w_{jk}^0 y_{pj} + \theta_k^0$$

$$y_{pk} = f_k^0 (sum_{pk}^0)$$



Deep Learning Supervisado



Entrenamiento (Backpropagation II)

■ Calcular el error en cada neurona

| Error en neurona de salida

$$\delta_{pk}^0 = (d_{pk} - y_{pk}) f_k^{0'}(net_{pk}^0)$$

| Error en neurona oculta

$$\delta_{pj}^h = f_j^{h'}(net_{pj}^h) \sum_k \delta_{pk}^0 w_{jk}^0$$

■ Actualizar los pesos

| Pesos en neurona de salida

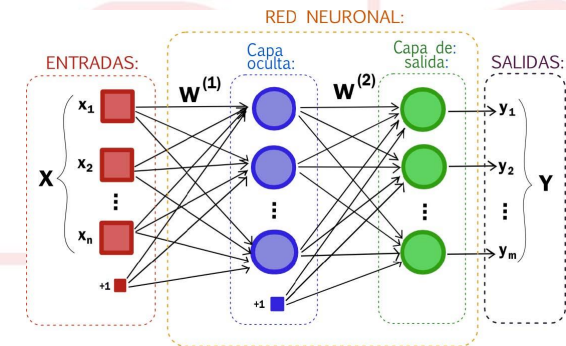
$$w_{jk}^0(t+1) = w_{jk}^0(t) + \Delta w_{jk}^0(t+1);$$

$$\Delta w_{jk}^0(t+1) = \alpha \delta_{pk}^0 y_{pj}$$

| Pesos en neurona oculta

$$w_{ij}^h(t+1) = w_{ij}^h(t) + \Delta w_{ij}^h(t+1);$$

$$\Delta w_{ij}^h(t+1) = \alpha \delta_{pj}^h x_{pi}$$

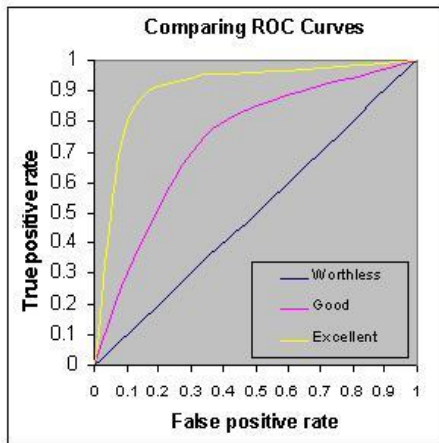


Deep Learning Supervisado



Area Under Curve ROC

- ROC (Receiver Operating Characteristic)



| predicted → real ↓ | Class_pos | Class_neg |
|-----------------------|-----------|-----------|
| Class_pos | TP | FN |
| Class_neg | FP | TN |

$$\text{TPR (sensitivity)} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{FPR (1-specificity)} = \frac{\text{FP}}{\text{TN} + \text{FP}}$$

Deep Learning Supervisado



LogLoss

- Objetivo Logloss próximo a 0.
- Mide la falta de certeza de las probabilidades de cada clase comparándolas con las etiquetas correctas.

$$L(\mathbf{w}) = \frac{1}{N} \sum_{n=1}^N H(p_n, q_n) = -\frac{1}{N} \sum_{n=1}^N \left[y_n \log \hat{y}_n + (1 - y_n) \log(1 - \hat{y}_n) \right],$$

Deep Learning Supervisado



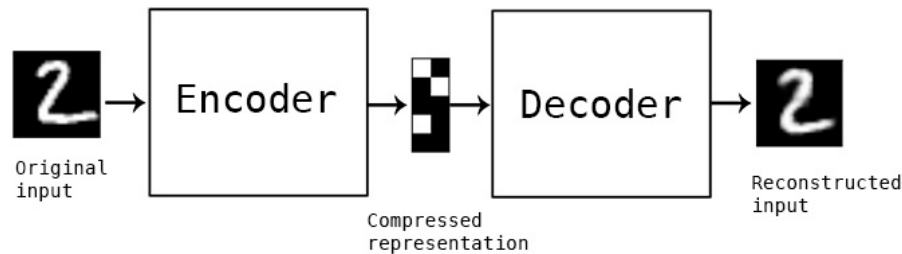
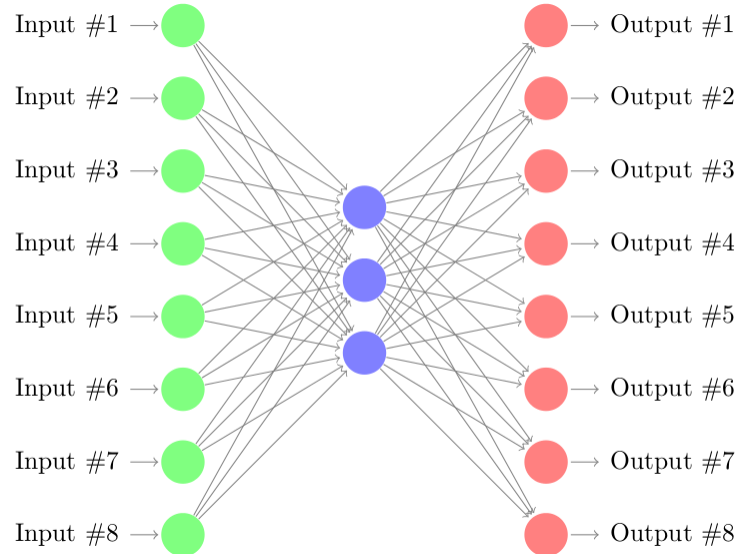
Demo: Detección de Fraude usando Deep Learning Supervisado



WeKnowMemes

Deep Learning No Supervisado

Autoencoders



Brace yourself..



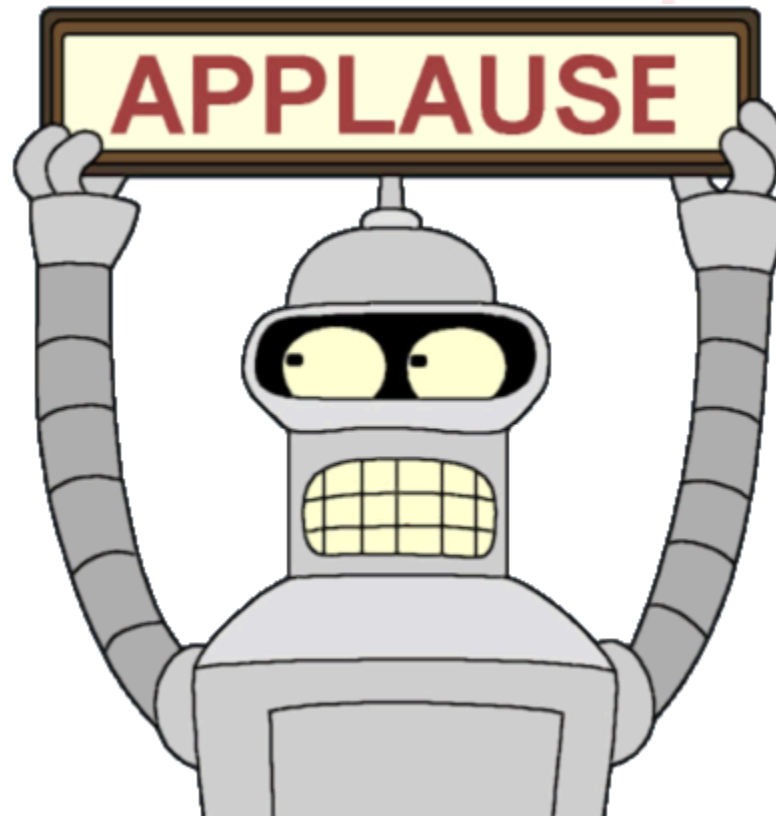
Deep Learning No Supervisado



Demo: Uso de Autoencoders



Tiempo de preguntas





Gracias por
su atención



GOBIERNO
DE ESPAÑA

MINISTERIO
DE ENERGÍA, TURISMO
Y AGENDA DIGITAL

