



Python Avanzado

Día 3

Curso Proteco 2020-2

Instructores:

- Samuel Garrido Sánchez
- Alicia Carbalindo García
- Armando Rivera



Inteligencia Artificial: Help *Kill* humans

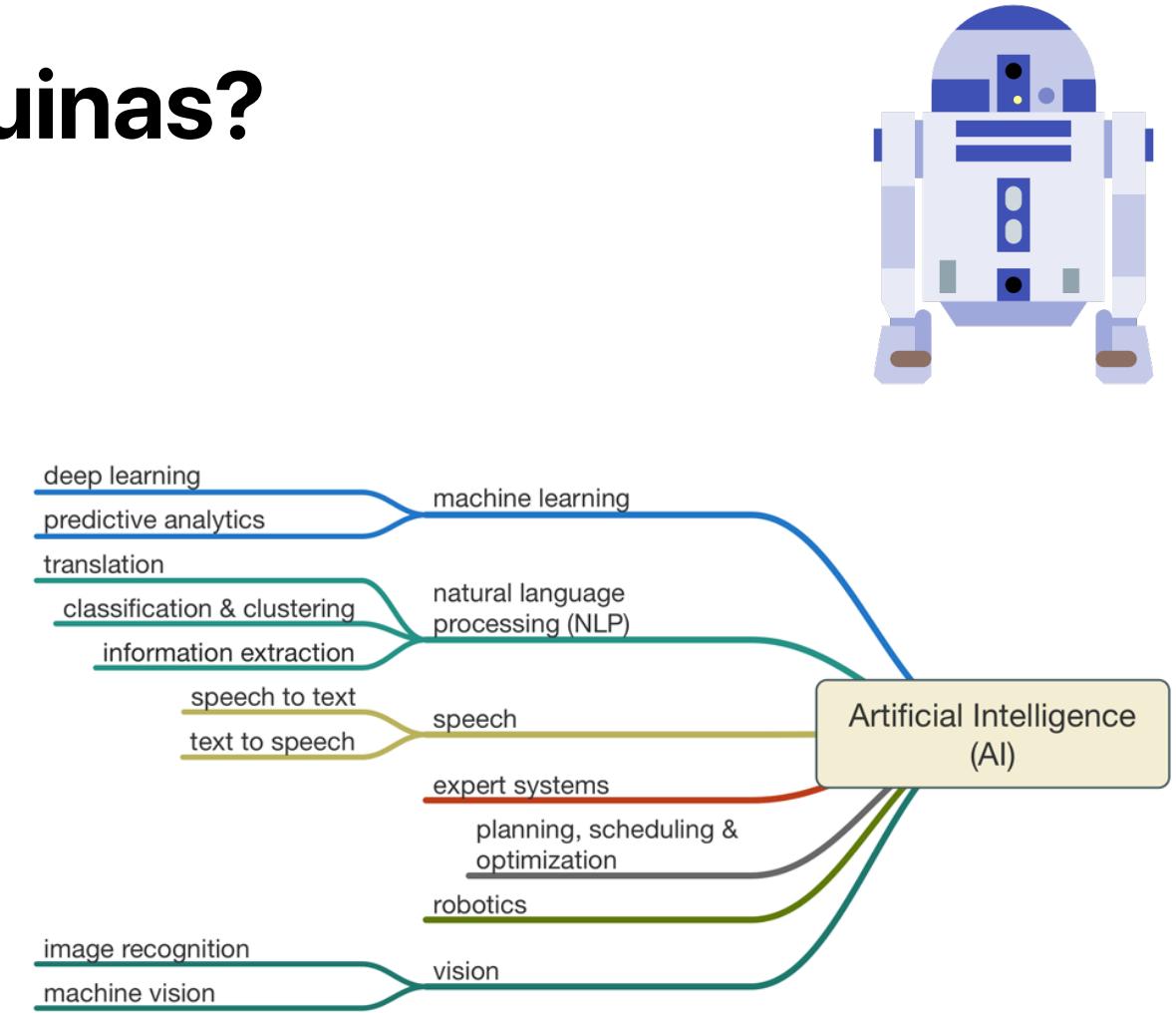


- La inteligencia artificial nos lo han articulado en películas como algo que nos va a destruir en un futuro o nos va a ayudar como nunca antes algo lo ha hecho.
- Alan Turing fue el creador de este término aunque no se le había tomado la importancia tanto como ahora.
- ¿Qué es exactamente IA hablando en términos computacionales?
 - R: Matemáticas:
 - Cálculo (Diferencial, Integral, Vectorial, Tensorial)
 - Probabilidad y estadística inferencial
 - Procesos estocásticos
 - Procesos de decisión de Markov
 - Análisis numérico
 - Ecuaciones diferenciales
 - Sistemas Dinámicos



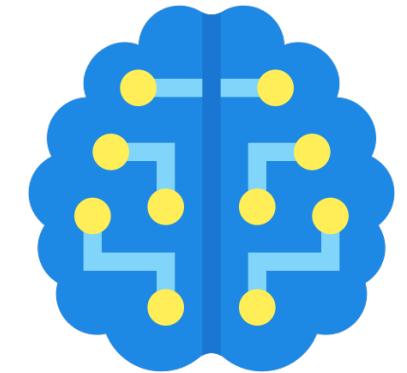
¿Cómo aprenden las máquinas?

- Clasificaciones de IA encontrarán muchas, sin embargo cada una de las ramas no existiría sin Machine Learning o Aprendizaje Máquina que denotaremos como ML
- La clasificación que les ofrecemos para ML es:
 - Aprendizaje Supervisado
 - Aprendizaje No Supervisado
 - Aprendizaje Por Refuerzo



Definiciones

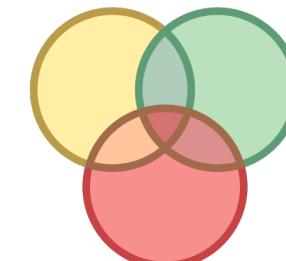
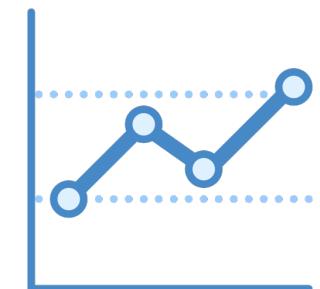
- El Aprendizaje Máquina (Machine Learning) permite a los ordenadores aprender sin necesidad de haber sido explícitamente programados.
- Son los avances en este campo, especialmente en el **deep learning**, los que han producido la reciente proliferación de la IA.
- El ML funciona entrenando a los sistemas informáticos a utilizar algoritmos para detectar patrones en los datos y posteriormente actuar de forma predictiva.



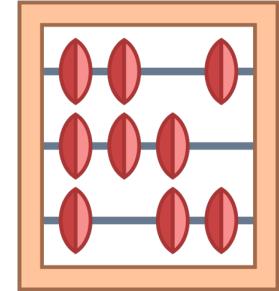
Aprendizaje Supervisado: Llevar de la manito



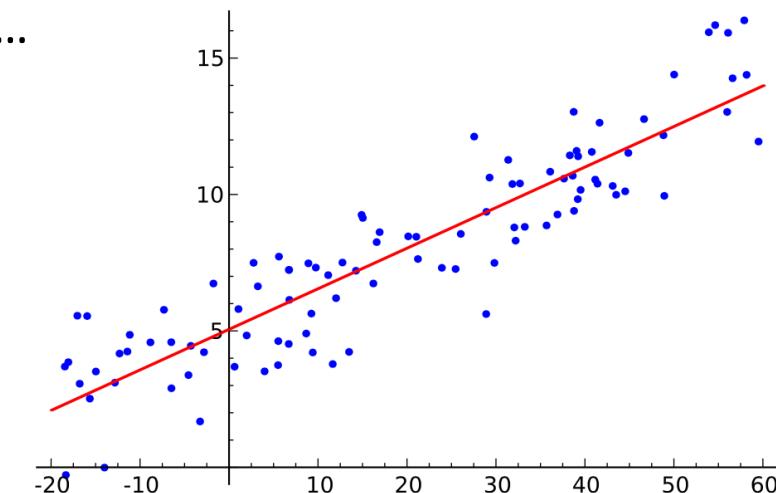
- Se le llama aprendizaje supervisado ya que nosotros tenemos un conjunto de elementos con etiquetas, por así decirlo, y así como mostramos distintos objetos punzantes a los niños, en un futuro si ven algo punzante, aunque no sean de los que les mostramos inicialmente, intuirán que corta o lastima.
- Dentro del aprendizaje supervisado podemos encontrar 2 grandes categorías:
 - **Algoritmos de Regresión o Predictivos.**
 - Nos permiten determinar el resultado de un valor en condiciones desconocidas.
 - **Algoritmos de Clasificación o Logísticos.**
 - Nos permite determinar a qué conjunto al que pertenece un elemento.



Supervisado: Algoritmos de regresión



- Los algoritmos de regresión consisten en predecir un valor, futuro o pasado, dados algunos que tengamos.
 - Existe:
 - Regresión Lineal Simple $\rightarrow f(x) = mx + b$
 - Regresión Lineal Multiple $\rightarrow f(x,y) = mx + ny + b$
 - Regresión Polinomial $\rightarrow f(x) = x^n$ o $f(x_1,x_2,\dots,x_n) = a_1x_1^n+a_2x_2^n+\dots$
 - Un ejemplo clásico es Regresión Lineal Simple
 - Tenemos una serie de puntos y queremos calcular la pendiente de la recta m que mejor se ajusta a todos los puntos.

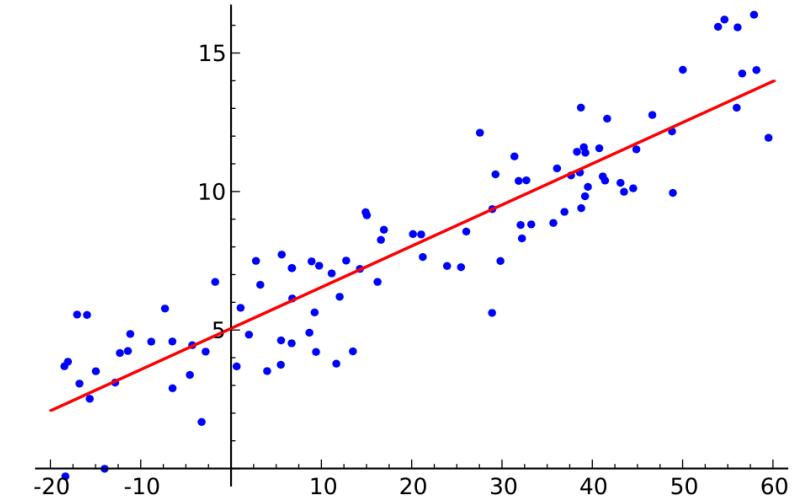


Supervisado: Algoritmos de regresión

- Ok, ya tengo la recta y la ecuación de la recta supongamos es:

$$f(x) = 4x + 4$$

- Es la ecuación que mejor se ajusta a todos los puntos, entonces con esa ecuación podría predecir, si recibo un valor más allá de 60 digamos 300, ¿Qué valor tendría el resultado?
- Ej. Imaginemos que x en la recta es la cantidad de productos que llevas de una tienda, y el total a pagar. Si ahora quiero llevar 200 productos, ¿cuánto debería de pagar?

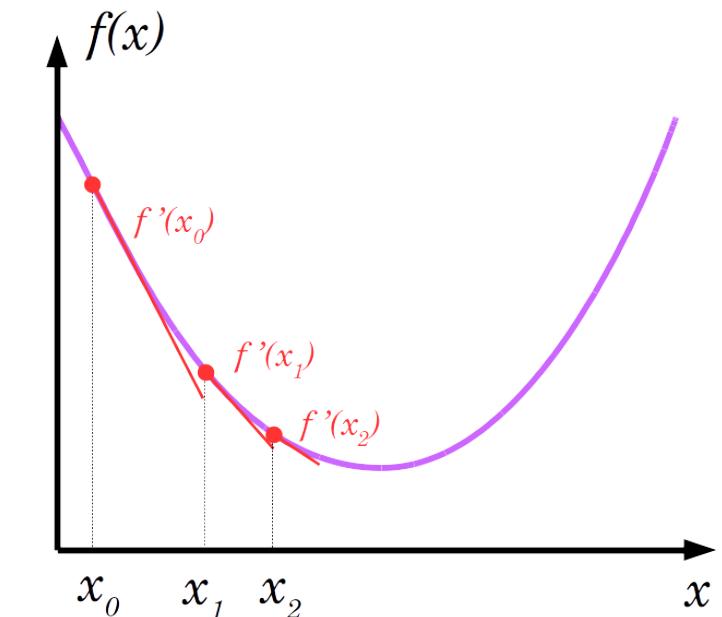
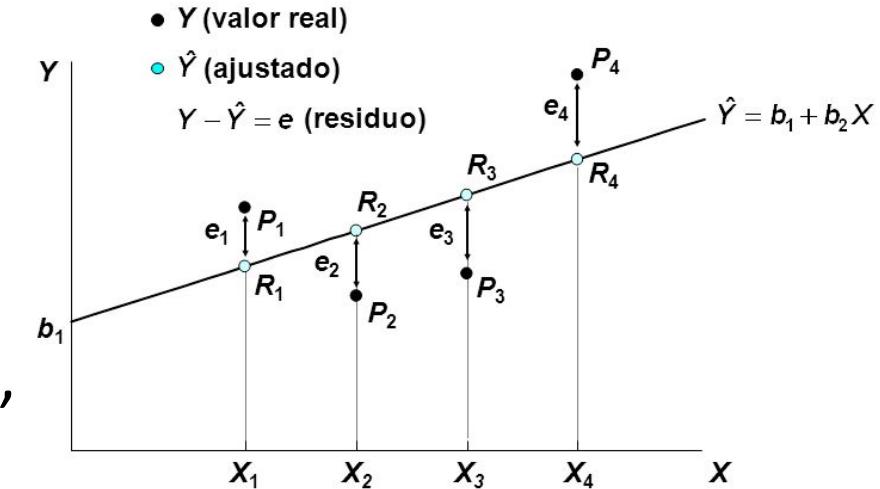


ECM: Error Cuadrático Medio

- La resta de P_n menos R_n nos da como resultado lo alejado que se encuentra el punto de la predicción, a esto lo llamaremos error e_n
- Existe algo llamado el error cuadrático medio, que es la suma de cada uno de estos errores cuadrado.

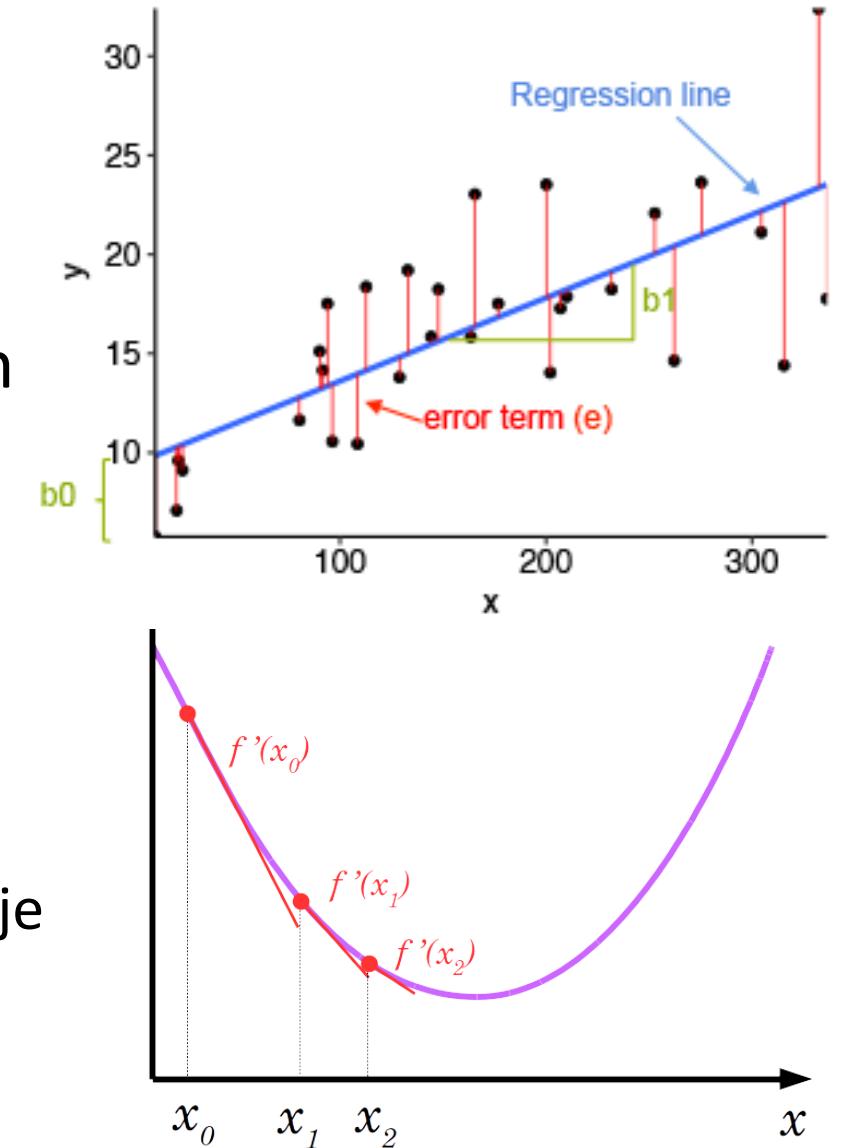
$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

- Por la palabra cuadrática se nos viene a la mente una parábola y justamente, tratamos de hacer mínimo este valor para tener la recta que mejor se ajuste a todos los puntos y pueda predecir mejor.



Función Gradiente

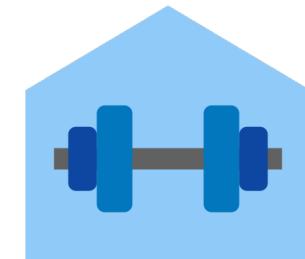
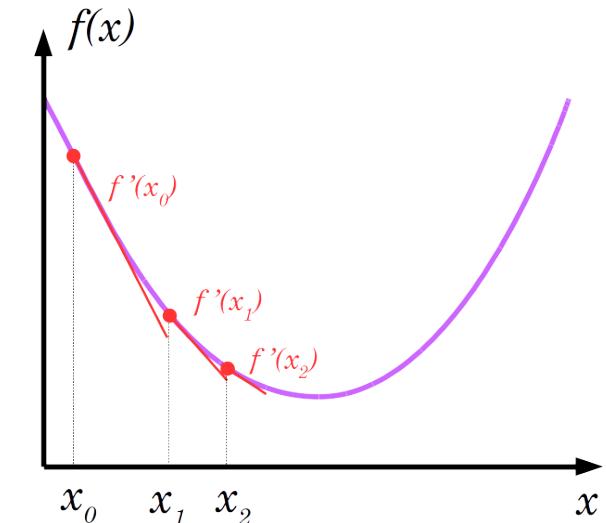
- La función gradiente está dada por la siguiente expresión, digamos tiene el vector de dirección al punto mínimo de la función ECM.
- Esta sale de pensar en la pendiente de la función original. Para moverla o cambiarla usaremos la expresión: (Digamos le quitamos un pedacito a m)
 - $m' = m - am$
- La variable a será nuestro coeficiente de aprendizaje y quiere decir qué tan fuerte vamos iterando para llegar hasta el punto mínimo.



Entrenamiento y Prueba (Train y Test)

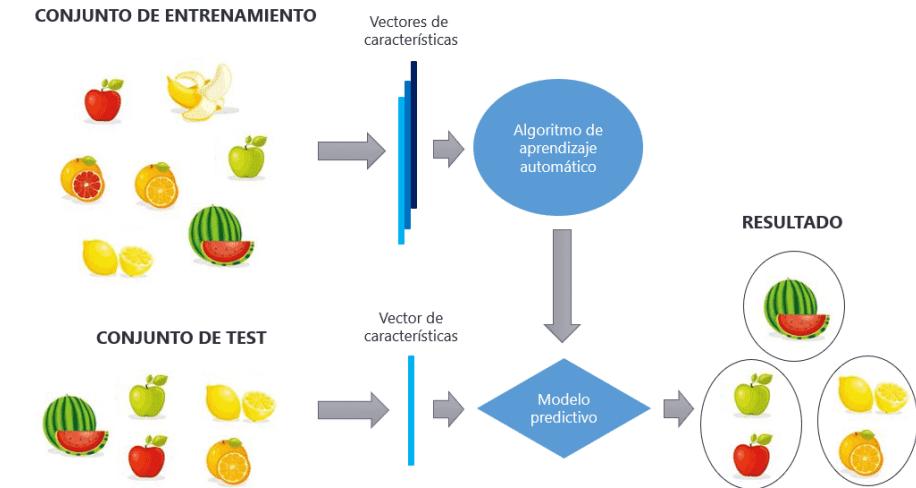
- Digamos que tenemos un conjunto de datos. Con variables x y y , “ x ” será la cantidad de útiles escolares que necesitas y “ y ” la cantidad de dinero que se tiene que pagar.
- Digamos que si tenemos un DataSet, separaremos 30% de los datos para que nos funcionen para probar la efectividad de nuestro algoritmo.
- 70% para entrenar, 30% para probar

*Esto puede variar dependiendo del entrenamiento pero eso es lo clásico.



Supervisado: Algoritmos de clasificación

- Un sistema de clasificación predice una categoría, mientras que una regresión predice un número.
- Les colocamos etiquetas y así poder determinar si es o no es un término.
 - Técnicas de ML para clasificación
 - Regresión Logística
 - Máquina de soporte Vectorial
 - Árboles de decisión
 - Bosques Aleatorios

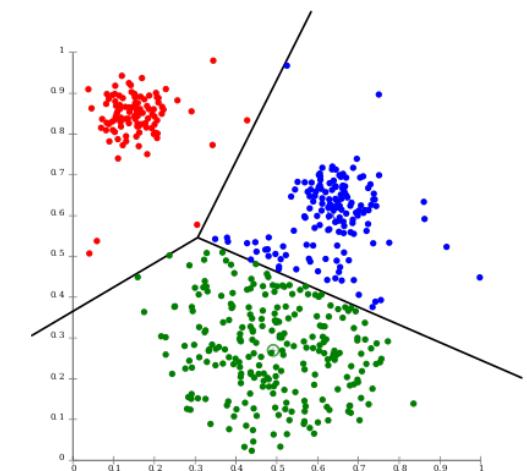


Supervisado: Algoritmos de clasificación

- Para determinar pertenecemos a un conjunto u otro es importante darle un significado matemático, por ejemplo si tenemos varios conjuntos de datos de un tumor:
 - Rigidez, forma, volumen, ancho, alto, textura, color, etc.
 - Si nos entregan un dataset con estos podremos decir desde qué punto exactamente un tumor tiene muchas probabilidades de ser cáncer.
 - Al lado derecho vemos que usando 7 algoritmos diferentes de clasificación, el porcentaje de precisión que tuvieron. Aquí ganó Regresión Logística.
- | Algoritmo | Porcentaje de precisión |
|---------------------------------|-------------------------|
| 1. Logistic Regression | 95.8% |
| 2. Nearest Neighbor | 95.1% |
| 3. Support Vector Machines | 97.2% |
| 4. Kernel SVM | 96.5% |
| 5. Naive Bayes | 91.6% |
| 6. Decision Tree Algorithm | 95.8% |
| 7. Random Forest Classification | 98.6% |

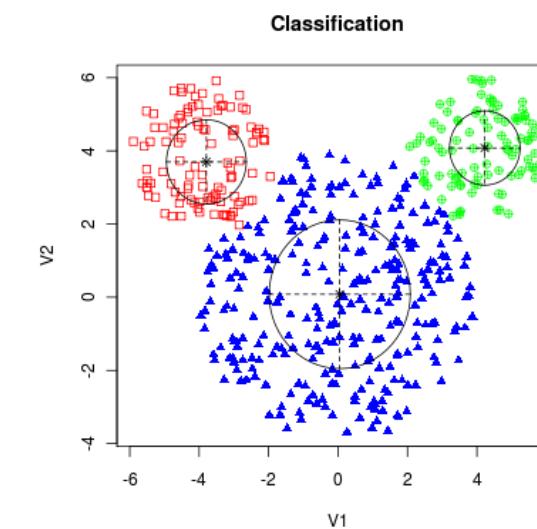
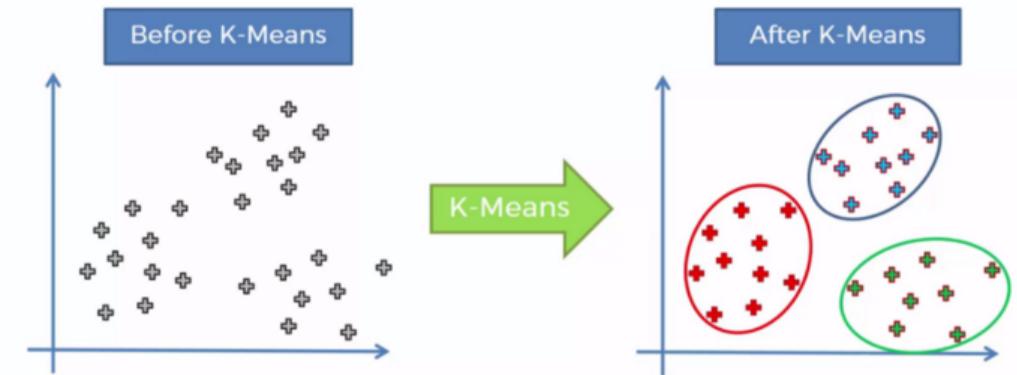
Aprendizaje No Supervisado: Que lo descubra el mismo !

- Un sistema de aprendizaje no supervisado consiste en que al algoritmo no le daremos etiquetas iniciales, quiere decir que no tendrá un conjunto de entrenamiento donde le mostremos que esto es un perro, esto un gato, esto un 2 o un 3. Simplemente por medio de características que logre reconocer comunes dirá en qué conjunto pertenecen.
- El más famoso dentro de ésta categoría es **K-Means**.



Aprendizaje No Supervisado: Clustering K-Means

- El método K-M (K nearest neighbors Fix y Hodges, 1951) es un método de clasificación no supervisada
- En el proceso de aprendizaje no se hace ninguna suposición acerca de la distribución de las variables predictoras.
- K-Means necesita llamado curva de Elbow, la cuál digamos que cuando se aplana ya tendremos nuestros conjuntos clasificados



Aprendizaje por refuerzo: PACMAN

- Hay algunas ocasiones, como en los videojuegos, que el aprendizaje de un algoritmo o mejor dicho, la adecuación del mismo viene hasta que interactuamos con él.
- Por ejemplo, los fantasmitas de PacMan después de tanto jugar hasta predicen qué vamos a hacer



Aprendizaje por refuerzo: Pégale de manotazos



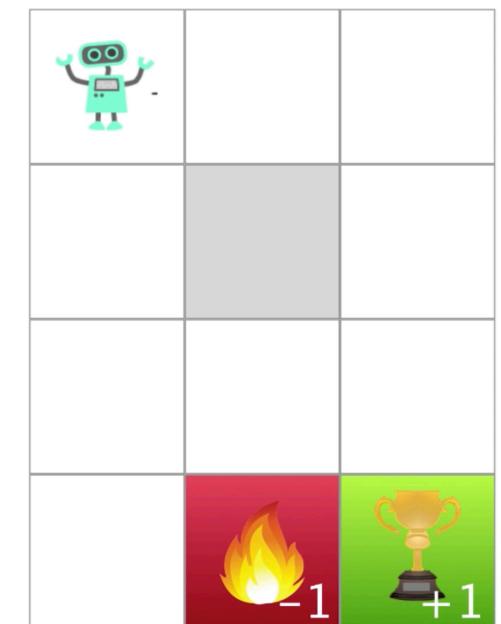
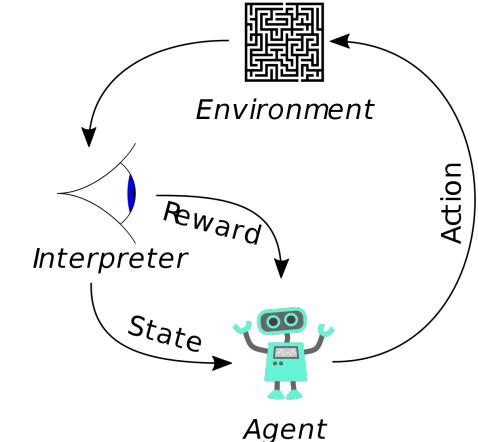
- Para ello se maneja un sistema de recompensas y castigos.
- Una ecuación matemática me va a decir dependiendo el paso que tome la probabilidad de fallar o ser recompensado.
- Así si un enemigo en un videojuego nos logra ganar, la computadora le dará una recompensa a esa serie de pasos que siguió para ganarnos y tratará de hacer esos pasos otra y otra vez. Si le logramos ganar la castigaremos y no seguirá esos pasos.



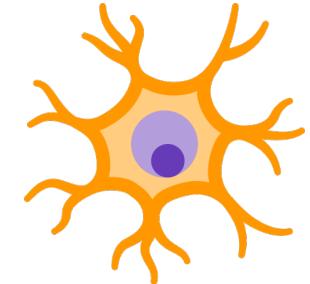
Agente



- Un agente inteligente, es una entidad capaz de percibir su entorno, procesar tales percepciones y responder o actuar en su entorno de manera racional, es decir, de manera correcta y tendiendo a maximizar un resultado esperado
- Supongamos un laberinto como el siguiente:
- Si nuestro robot trata de recorrer el laberinto puede quemarse o ganar un premio. A nivel computadora veremos que sumaremos +1 o restaremos -1



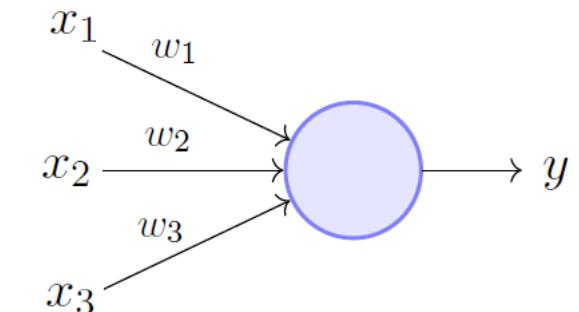
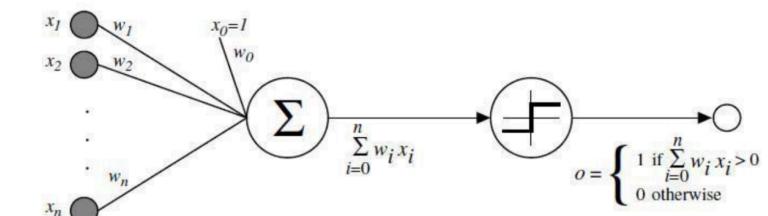
Introducción a Redes Neuronales.



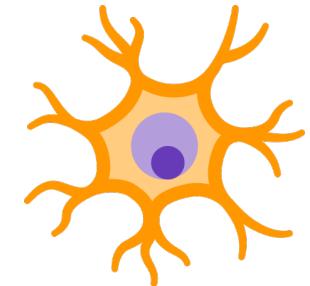
- Suena medio rimbombante el nombre y la mayoría del tiempo el nombre atrae más que lo que es realmente.
- Las redes neuronales artificiales son un modelo computacional inspirado en el comportamiento observado en su homólogo biológico, las neuronas.

Su unidad es la neurona, llamada **perceptrón**:

- Donde el famoso perceptron solo es una suma.
- Una suma de variables ponderadas y que luego decidimos si pasamos esos valores a la siguiente capa o no.



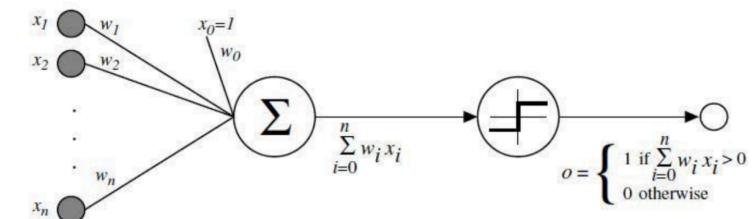
¿Qué busca hacer una Red Neuronal?



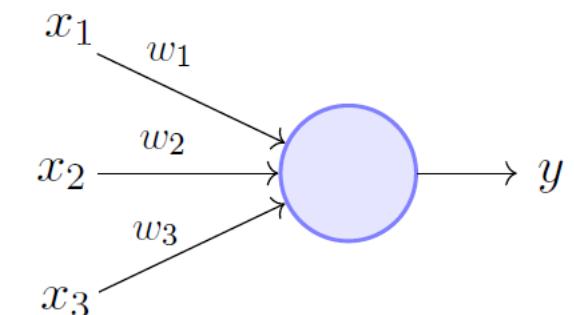
- Resolver una ecuación de este tipo:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \epsilon$$

$$y = \beta_0 + \sum_{i=1}^n \beta_i x_i + \epsilon$$

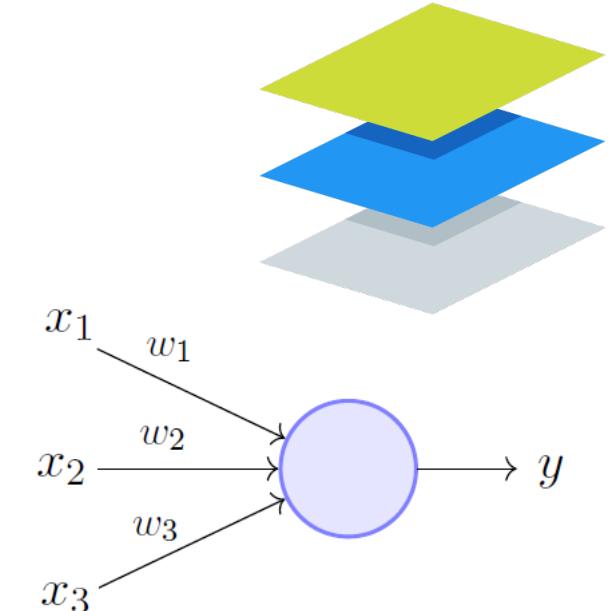
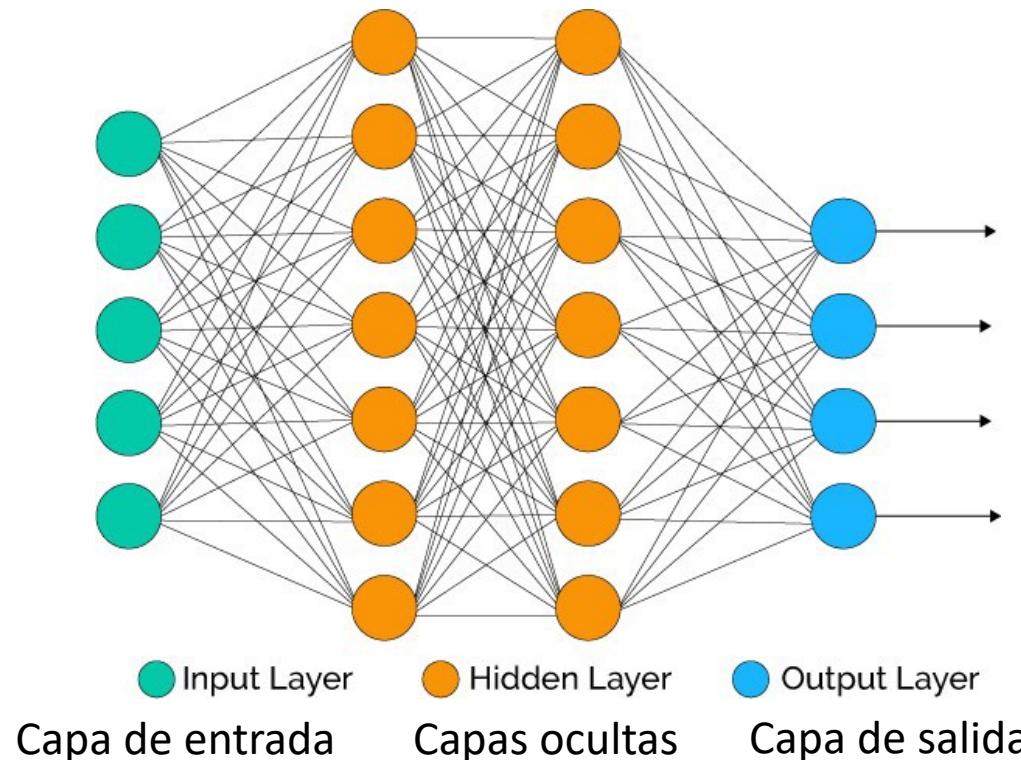


- Donde los betas son llamados pesos o también se colocan como w_1, w_2, \dots, w_n por **weight**
- $X_1, X_2, X_3, \dots, X_5$ son los valores que nos arrojaron las neuronas anteriores.



Composición de una Red Neuronal: Capas

- Una red neuronal se compone de capas

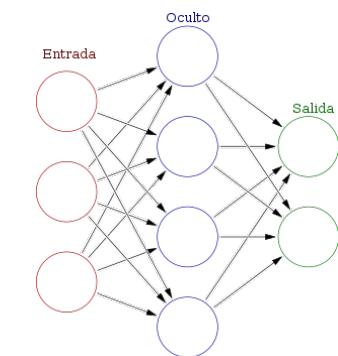


Los ogros son como las cebollas. [...] Las cebollas tienen capas. Los ogros tienen capas. ¿Entiendes? Los dos tenemos capas
-Shrek

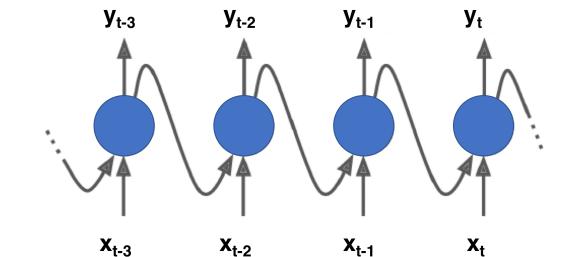
- Estas capas están compuestas por un conjunto de neuronas. (Layer = Capas)

Tipos de RN : Redes Neuronales

- Existen 3 principales tipos de redes neuronales:
 - **Redes Neuronales Artificiales**
 - Capa inicial de un tamaño definido.
 - Usos: Predicción de números, con entradas definidas.
 - **Redes Neuronales Convolucionales**
 - Capa inicial de un tamaño indefinido, se realiza la operación de convolución.
 - Usos: Reconocimiento de imágenes, reconocimiento de escritos y patrones.
 - **Redes Neuronales Recurrentes**
 - Forman un gráfico dirigido a lo largo de una secuencia temporal.
 - Comportamiento Dinámico Temporal.
 - Usos: Procesamiento del lenguaje natural.



$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau$$

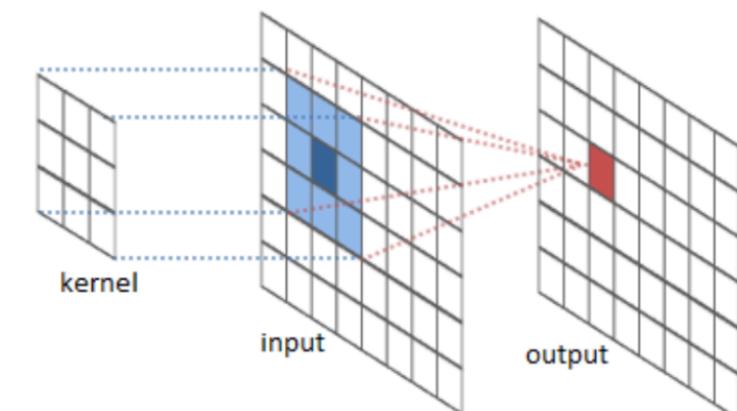


Ejemplo de Redes Neuronales Convolucionales

- ¿Una imagen qué es?
- Todo dentro de nuestra computadora son números entonces ¿qué sería una imagen?
- Un ARREGLO de tuplas (r,g,b,a). Cada pixel tiene 3 colores para formar todos los colores disponibles.
 - Rojo
 - Azul
 - Verde
 - Y por ultimo una transparencia alpha.

Entonces si va vemos números, podremos ponerlos en una función matemática

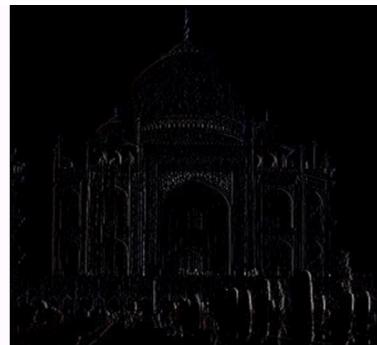
$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau$$



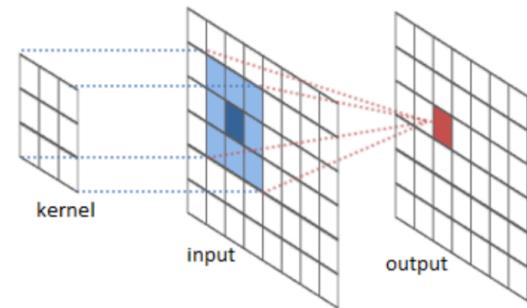
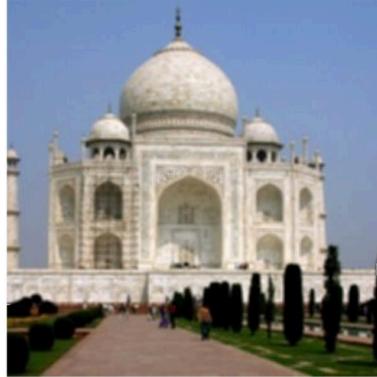
Así también funcionan sus filtros de SnapChat o PhotoShop

- Como las imágenes son una matriz de números, podríamos aplicar álgebra lineal
- Si "tallamos" una matriz como la de la derecha de las imágenes sobre la original, numéricamente las estaremos poniendo borrosas u oscureciéndolas para marcar contornos.

0	0	0	0	0	0
0	0	0	0	0	0
0	-1	1	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0



0	0	0	0	0	0
0	1/9	1/9	1/9	0	0
0	1/9	1/9	1/9	0	0
0	1/9	1/9	1/9	0	0
0	0	0	0	0	0



$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau$$



SnapChat o PhotoShop o apps de fotos Pura Álgebra Lineal y cálculo.

- Cómo podemos recordar fácilmente la convolución
- La manera más gráfica y descriptiva que he encontrado

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau$$

