

Tema 1

Introducción al Aprendizaje Automático

Gonzalo A. Aranda-Corral

Ciencias de la Computación e Inteligencia Artificial
Universidad de Huelva

19 de octubre de 2020

Por favor, registrar vuestra asistencia en la siguiente dirección:

<http://opendatalab.uhu.es/aplicaciones/asistencia>



¿Las máquinas pueden aprender?

Ada Augusta, filósofa de la computación (1961) dijo que la máquina analítica no pretende crear nada, puede hacer lo que sea si se le ha indicado la forma de hacerlo.

“Aprender es hacer cambios útiles en nuestra mente” [Marvin Minsky]

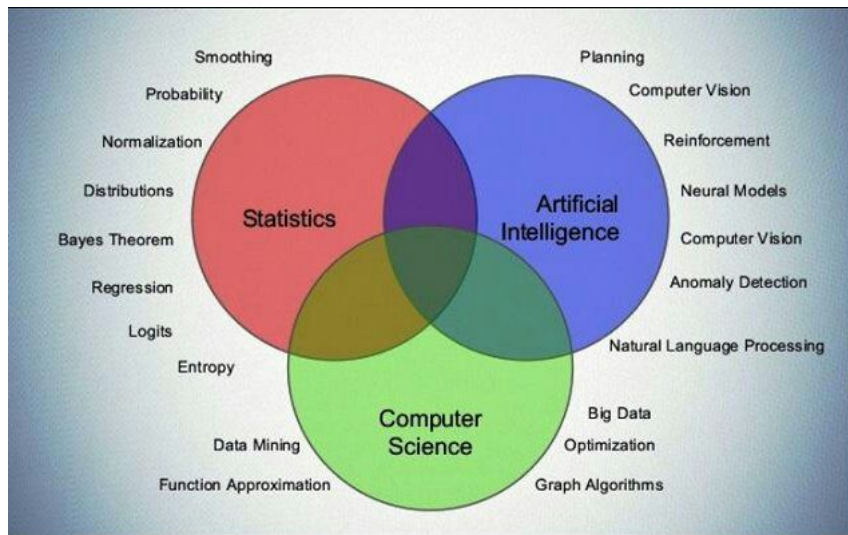
- Simon(1983),define el aprendizaje como **cambios** en el sistema que se **adaptan** de manera que permiten llevar a cabo la **misma tarea**, o una tarea parecida, de un modo **más eficiente y eficaz**.
- “(Del lat. apprehendere) tr. **Adquirir** el conocimiento de alguna cosa por medio del estudio o la **experiencia** ” [Real Academia Española]

- “**Decimos** que una máquina puede **aprender**, con respecto a una **tarea T**, una **medida** de control **P** y un tipo de **experiencia E**, si el sistema **mejora** su rendimiento **P** en la tarea **T** en las **siguientes** experiencias de **E**” [Mitchell]
- Dependiendo de cómo especifiquemos la tarea, el aprendizaje podrá ser “data mining”, “autonomous discovery”, “programming by example”, etc.

Resumen:

- Aprender significa: Modificación del comportamiento del sistema o de su representación interna, a partir de una experiencia previa, con la mejora del sistema, de acuerdo a algún criterio de evaluación.

- Las Redes sociales, IoT (Internet of Things), Ordenadores. . . producen GRANDES cantidades de datos
- Es la era del **Big Data**
- Vuelven a tomar MUCHA fuerza las redes neuronales: **Deep Learning**, . . .
- Nuevos métodos para el tratamiento paralelo y masivo de datos: **Random Forests**. . .
- Adaptación de los algoritmos antiguos al **paralelismo masivo**.
- Aparece la “**Ciencia del Dato**” (Data Science)



El proceso que sigue un Data Scientist para responder a las cuestiones que se le plantean se pueden resumir en estos 5 pasos:

- **Extraer** los datos, independientemente de su fuente (webs, csv, logs, apios, etc.) y de su volumen (Big Data o Small Data).
- **Limpiar** los datos, para eliminar lo que distorsiona las mismas.
- **Procesar** los datos usando diferentes métodos estadísticos (inferencia estadística, modelos de regresión, pruebas de hipótesis, etc.).
- **Diseñar** nuevos tests o experimentos en caso necesario.
- **Visualizar** presentar gráficamente los datos.

El proceso que sigue un Data Scientist para responder a las cuestiones que se le plantean se pueden resumir en estos 5 pasos:

- **Extraer** los datos, independientemente de su fuente (webs, csv, logs, apios, etc.) y de su volumen (Big Data o Small Data).
- **Limpiar** los datos, para eliminar lo que distorsiona las mismas.
- **Procesar** los datos usando diferentes métodos estadísticos (inferencia estadística, modelos de regresión, pruebas de hipótesis, etc.).
- **Diseñar** nuevos tests o experimentos en caso necesario.
- **Visualizar** presentar gráficamente los datos.

El proceso que sigue un Data Scientist para responder a las cuestiones que se le plantean se pueden resumir en estos 5 pasos:

- **Extraer** los datos, independientemente de su fuente (webs, csv, logs, apios, etc.) y de su volumen (Big Data o Small Data).
- **Limpiar** los datos, para eliminar lo que distorsiona las mismas.
- **Procesar** los datos usando diferentes métodos estadísticos (inferencia estadística, modelos de regresión, pruebas de hipótesis, etc.).
- **Diseñar** nuevos tests o experimentos en caso necesario.
- **Visualizar** presentar gráficamente los datos.

El proceso que sigue un Data Scientist para responder a las cuestiones que se le plantean se pueden resumir en estos 5 pasos:

- **Extraer** los datos, independientemente de su fuente (webs, csv, logs, apios, etc.) y de su volumen (Big Data o Small Data).
- **Limpiar** los datos, para eliminar lo que distorsiona las mismas.
- **Procesar** los datos usando diferentes métodos estadísticos (inferencia estadística, modelos de regresión, pruebas de hipótesis, etc.).
- **Diseñar** nuevos tests o experimentos en caso necesario.
- **Visualizar** presentar gráficamente los datos.

El proceso que sigue un Data Scientist para responder a las cuestiones que se le plantean se pueden resumir en estos 5 pasos:

- **Extraer** los datos, independientemente de su fuente (webs, csv, logs, apios, etc.) y de su volumen (Big Data o Small Data).
- **Limpiar** los datos, para eliminar lo que distorsiona las mismas.
- **Procesar** los datos usando diferentes métodos estadísticos (inferencia estadística, modelos de regresión, pruebas de hipótesis, etc.).
- **Diseñar** nuevos tests o experimentos en caso necesario.
- **Visualizar** y presentar gráficamente los datos.

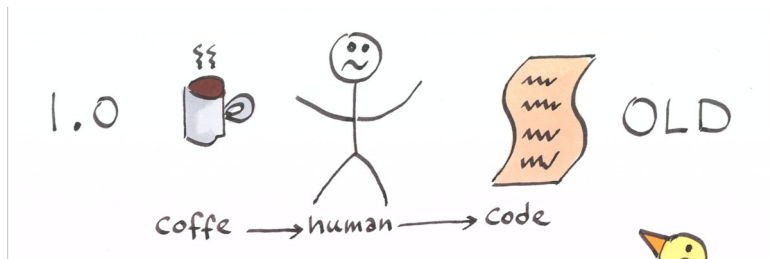
Cada día se integra en más aspectos cotidianos

Algunas aplicaciones:

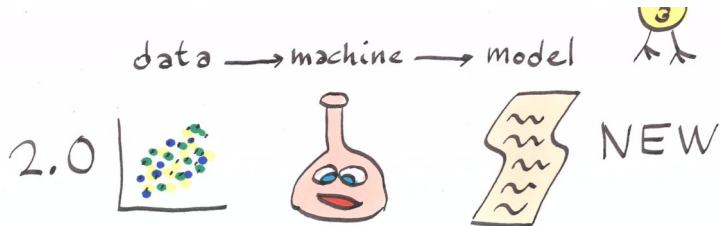
- **Urbanismo**: Estudio de datos para la mejor planificación de la ciudad: Smart Cities
- **Arte**: Digital Humanities, o tratamiento digital de información artística para la extracción de patrones o características
- **Periodismo**: Detrás de los datos hay historias y los periodistas quieren saber interpretarlas
- ...

- El aprendizaje es una **tarea para mejorar a otras tareas**.
- No tiene un objetivo específico, **depende** de la tarea a mejorar.
- Es una **alternativa** a la tarea de la especificación y **programación**.
- Es de **aproximación** sujeta a valoración.

Nueva Tarea



Nueva Tarea



Ejemplo

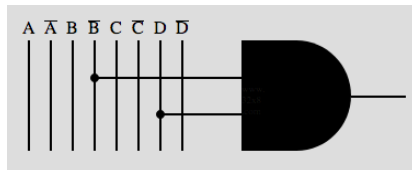
- Tenemos la siguiente máquina digital con la descripción del comportamiento, pero no del interior
- ... y tenemos que replicar su comportamiento:



Example	x_1	x_2	x_3	x_4	y
0	0	1	1	0	0
1	0	0	0	0	0
2	0	0	1	1	1
3	1	0	0	1	1
4	0	1	1	0	0
5	1	1	0	0	0
6	0	1	0	1	0

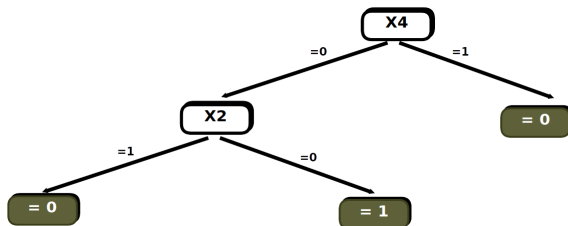
Ejemplo

Probablemente, con electrónica



Ejemplo

con Aprendizaje Automático (algoritmo ID3):



Dentro del aprendizaje hay que distinguir 2 etapas:

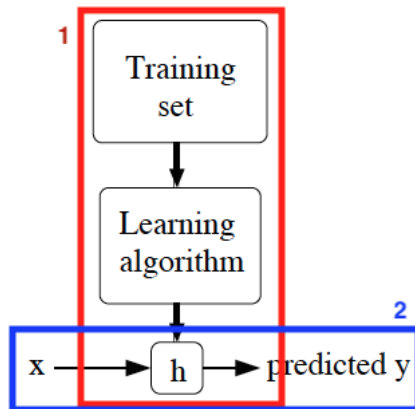
- **Aprendizaje o Adquisición:** las entradas del sistema son el conocimiento disponible los resultados disponibles y un objetivo de mejora.
- **Validación y Utilización:** Se usa el sistema aprendido para aplicar a otros objetos. Debe de incluir algún método de valoración o crítica sobre los resultados.

Dentro del aprendizaje hay que distinguir 2 etapas:

- **Aprendizaje o Adquisición:** las entradas del sistema son el conocimiento disponible los resultados disponibles y un objetivo de mejora.
- **Validación y Utilización:** Se usa el sistema aprendido para aplicar a otros objetos. Debe de incluir algún método de valoración o crítica sobre los resultados.

Elementos del aprendizaje

Gráficamente:



- **Dataset:** Conjunto de instancias completo
- **Instancia:** (o Ejemplo) concreto, definido por un conjunto de valores de atributos
- **Atributo:** Característica que define a un elemento de un conjunto
- **Clase:** Cada uno de los subconjuntos en los que se quiere dividir el conjunto de instancias.
- **Ejemplo (positivo):** Instancia que pertenece a la clase que queremos definir.
Ejemplo negativo: Instancia que no pertenece a la clase. instancias de otras clases.

- **Dataset:** Conjunto de instancias completo
- **Instancia:** (o Ejemplo) concreto, definido por un conjunto de valores de atributos
- **Atributo:** Característica que define a un elemento de un conjunto
- **Clase:** Cada uno de los subconjuntos en los que se quiere dividir el conjunto de instancias.
- **Ejemplo (positivo):** Instancia que pertenece a la clase que queremos definir.
Ejemplo negativo: Instancia que no pertenece a la clase. instancias de otras clases.

- **Dataset:** Conjunto de instancias completo
- **Instancia:** (o Ejemplo) concreto, definido por un conjunto de valores de atributos
- **Atributo:** Característica que define a un elemento de un conjunto
- **Clase:** Cada uno de los subconjuntos en los que se quiere dividir el conjunto de instancias.
- **Ejemplo (positivo):** Instancia que pertenece a la clase que queremos definir.
Ejemplo negativo: Instancia que no pertenece a la clase. instancias de otras clases.

- **Dataset:** Conjunto de instancias completo
- **Instancia:** (o Ejemplo) concreto, definido por un conjunto de valores de atributos
- **Atributo:** Característica que define a un elemento de un conjunto
- **Clase:** Cada uno de los subconjuntos en los que se quiere dividir el conjunto de instancias.
- **Ejemplo (positivo):** Instancia que pertenece a la clase que queremos definir.
Ejemplo negativo: Instancia que no pertenece a la clase. instancias de otras clases.

- **Dataset:** Conjunto de instancias completo
- **Instancia:** (o Ejemplo) concreto, definido por un conjunto de valores de atributos
- **Atributo:** Característica que define a un elemento de un conjunto
- **Clase:** Cada uno de los subconjuntos en los que se quiere dividir el conjunto de instancias.
- **Ejemplo (positivo):** Instancia que pertenece a la clase que queremos definir.
Ejemplo negativo: Instancia que no pertenece a la clase. instancias de otras clases.

Definiciones

Feature or Attribute

Class

Instance

	X0	X1	X2	class
	0.5351795492	0.9443102776	0.1582435145	1
	0.2372136153	0.6406416746	0.2375491596	1
	0.9115356348	0.3311024322	0.5615073269	0
	0.5634070287	0.4183148035	0.151904445	0
	0.3728975195	0.3816657621	0.616341473	1
	0.6783527289	0.938524515	0.5269012505	1
	0.09568660734	0.04465749689	0.0133451798	0
Positive Example	0.2173318229	0.6170559076	0.3122273853	1
	0.818890594	0.7459451367	0.9026713492	0
Negative Example	0.6064854042	0.5945985792	0.2188024961	0
	0.1546966824	0.1579937453	0.1333579164	0

Dataset

Definiciones

- **Training Set:** Conjunto de instancias a partir de los cuales se aprende
- **Validation Set:** Conjunto de instancias para validar el conocimiento aprendido
- **Test Set:** Conjunto de instancias para validar el conocimiento aprendido



Definiciones

- **Training Set:** Conjunto de instancias a partir de los cuales se aprende
- **Validation Set:** Conjunto de instancias para validar el conocimiento aprendido
- **Test Set:** Conjunto de instancias para validar el conocimiento aprendido



- **Training Set:** Conjunto de instancias a partir de los cuales se aprende
- **Validation Set:** Conjunto de instancias para validar el conocimiento aprendido
- **Test Set:** Conjunto de instancias para validar el conocimiento aprendido



- **Hipótesis:** Generalización o descripción de un conjunto de ejemplos de una clase. Describe a la clase y no al resto.
- **Ruido:** Instancias mal clasificadas o desviadas.
- **Bias:** Permisividad en la predicción. Aprendizaje demasiado abierto. Mala predicción.
- **Sobreajuste:** Aprendizaje de los datos demasiado ajustado. Mala predicción.

- **Hipótesis:** Generalización o descripción de un conjunto de ejemplos de una clase. Describe a la clase y no al resto.
- **Ruido:** Instancias mal clasificadas o desviadas.
- **Bias:** Permisividad en la predicción. Aprendizaje demasiado abierto. Mala predicción.
- **Sobreajuste:** Aprendizaje de los datos demasiado ajustado. Mala predicción.

- **Hipótesis:** Generalización o descripción de un conjunto de ejemplos de una clase. Describe a la clase y no al resto.
- **Ruido:** Instancias mal clasificadas o desviadas.
- **Bias:** Permisividad en la predicción. Aprendizaje demasiado abierto. Mala predicción.
- **Sobreajuste:** Aprendizaje de los datos demasiado ajustado. Mala predicción.

- **Hipótesis:** Generalización o descripción de un conjunto de ejemplos de una clase. Describe a la clase y no al resto.
- **Ruido:** Instancias mal clasificadas o desviadas.
- **Bias:** Permisividad en la predicción. Aprendizaje demasiado abierto. Mala predicción.
- **Sobreajuste:** Aprendizaje de los datos demasiado ajustado. Mala predicción.

Los problemas de Aprendizaje Automático se pueden dividir atendiendo a diversos criterios:

- Objetivo
- Datos de entrada
- Datos de salida
- ...

Podemos encontrar 2 tipo de objetivos:

- **Predicción:**

Tratar de, dado una serie de ejemplos, adivinar que valor tendrá una función objetivo.

- **Clasificación:**

Agrupar objetos en función a un objetivo, en función de su valor.

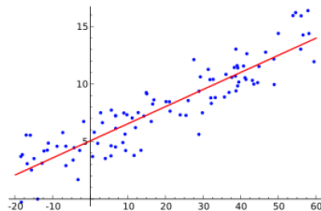
Ambas tareas son equivalentes.

Predicción

El problema de predicción consiste en encontrar una ecuación $y = g(x)$ que nos modele de una forma más simple el sistema.

Generalmente se especifica mediante una regresión lineal.

Ejemplo: Precio de los coches en función de su kilometraje.



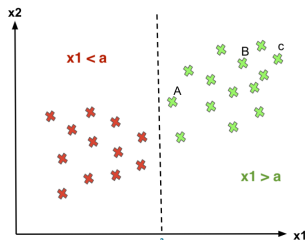
$$y = w_1 * x + w_0$$

Clasificación

Consiste en encontrar fórmulas o reglas que nos digan a que grupo pertenece un “individuo”

Ejemplo: Clasificación del riesgo

Tarea: Diferenciar los clientes de “alto riesgo” o “bajo riesgo” en función de los ingresos y los ahorros



```
IF ( $x_1 > a$ ) THEN  
    “bajo riesgo”  
ELSE  
    “alto riesgo”
```


Dependiendo del tipo de dato de entrada podremos aplicar unos algoritmos u otros y nos va a condicionar mucho el proceso.

Las variables pueden ser de tipos:

- **Reales:** Variable numérica que puede almacenar cualquier número real.
- **Enteros o Discretos:** Variable numérica con valores únicamente enteros.)
- **Categorica o Nominal:** Variable que no es representada por números ni tiene algún tipo de orden.

Según el problema, podemos conocer el resultado que debería darnos el sistema en algunos casos, o no, o, al menos, tener alguna función que nos oriente al aprendizaje.

- **Aprendizaje supervisado:** Se aprende a partir de un conjunto de ejemplos de los cuales sabemos el comportamiento ideal. Tenemos un atributo de clase.
- **Aprendizaje no supervisado:** No tenemos medida de cuál es el comportamiento ideal del sistema, pero el criterio se basa en la similitud de grupos de datos. No tenemos clase.
- **Aprendizaje por refuerzo:** El comportamiento deseado no se representa mediante ejemplos sino mediante una cierta evaluación sobre los resultados que genera el sistema en su entorno. Función de Recompensa/castigo

Según el problema, podemos conocer el resultado que debería darnos el sistema en algunos casos, o no, o, al menos, tener alguna función que nos oriente al aprendizaje.

- **Aprendizaje supervisado:** Se aprende a partir de un conjunto de ejemplos de los cuales sabemos el comportamiento ideal. Tenemos un atributo de clase.
- **Aprendizaje no supervisado:** No tenemos medida de cuál es el comportamiento ideal del sistema, pero el criterio se basa en la similitud de grupos de datos. No tenemos clase.
- **Aprendizaje por refuerzo:** El comportamiento deseado no se representa mediante ejemplos sino mediante una cierta evaluación sobre los resultados que genera el sistema en su entorno. Función de Recompensa/castigo

Según el problema, podemos conocer el resultado que debería darnos el sistema en algunos casos, o no, o, al menos, tener alguna función que nos oriente al aprendizaje.

- **Aprendizaje supervisado:** Se aprende a partir de un conjunto de ejemplos de los cuales sabemos el comportamiento ideal. Tenemos un atributo de clase.
- **Aprendizaje no supervisado:** No tenemos medida de cuál es el comportamiento ideal del sistema, pero el criterio se basa en la similitud de grupos de datos. No tenemos clase.
- **Aprendizaje por refuerzo:** El comportamiento deseado no se representa mediante ejemplos sino mediante una cierta evaluación sobre los resultados que genera el sistema en su entorno. Función de Recompensa/castigo

- Un factor muy importante es cómo queremos obtener nuestro aprendizaje.
- Los requerimientos del problema nos puede obligar a tener unas estructuras de salida que, además, nos puede condicionar, incluso, el algoritmo de aprendizaje a aplicar

Aproximación Lineal

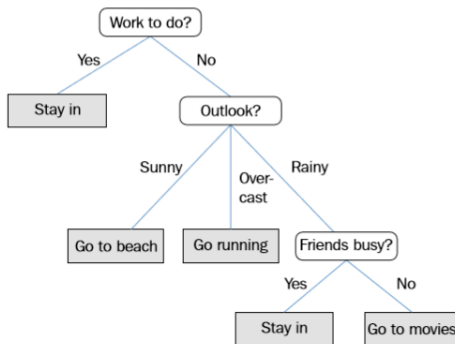
- En la mayoría de los casos numéricos, lo más fácil es suponer que la salida va a ser una función lineal.
- Puede ser una regresión, o una máquina Vector-Soporte o cualquier otro, pero la hipótesis siempre se escribe de la misma forma.

$$H(x) = \sum_{i=0}^{n-1} \omega_i x_i \quad \text{Caso } n - \text{dimensional}$$

- Aprender significa... encontrar esos ω_i que nos hace falta para completar la ecuación

Árboles de decisión

- Permiten clasificar ejemplos basados en valores de atributos
 - Los nodos hoja representan una clasificación
 - Los nodos internos son atributos a evaluar y las ramificaciones corresponden a los diferentes valores
- El aprendizaje consiste en la creación y modificación de estos árboles



- Reglas IF-THEN en las que el consecuente es la elección de una clase
- El antecedente de las reglas se refiere a valores de atributos y suelen ser una conjunción de selectores
 - Un selector es una condición de la forma “atributo-relación-valor” (Ej: $x > 2$)
- El aprendizaje consiste en la creación y modificación de estas reglas

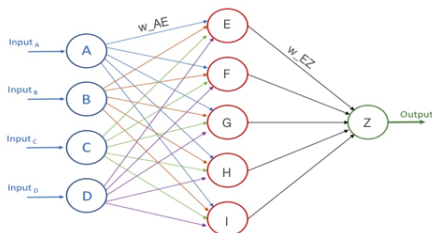
Queremos obtener como resultado un código ejecutable y, en este caso, en algún lenguaje lógico.

- El conocimiento está representado como una base de reglas y hechos, escritos generalmente en lógica de primer orden.
- El aprendizaje consiste en la construcción de hechos y reglas de forma automática a partir de ejemplos.

```
plus(0,X,X).  
plus(X,0,X).  
plus(A,B,C) :- dec(A,D), inc(B,E), plus(D,E,C).  
  
mult(0,X,0).  
mult(A,B,C) :- dec(A,D), mult(D,B,E), plus(B,E,C).
```

Redes neuronales

- Redes formadas por elementos de procesamiento simple (neuronas) con conexiones que poseen un peso asociado



- El aprendizaje consiste en la modificación de los pesos para adaptar el comportamiento de la red