

Aprendizaje de máquina

Verónica E. Arriola-Rios
(Basado en el curso de Andrew NG)

Aprendizaje de máquina

14 de abril de 2020

Temas

1 Introducción

- Tipos de aprendizaje

2 Sesgo inductivo

Aplicaciones

- Minería en bases de datos
 - Internet
 - Búsquedas en la red
 - Clasificación de correos
 - Registros médicos
 - Biología
 - Ingeniería
- Programas que no se logran hacer a mano.
 - Reconocimiento de escritura.
 - Visión por computadora.
 - Etiquetado de fotografías
 - Procesamiento de lenguaje natural.
- Programas que aprenden a ajustarse al usuario.
 - Sistemas de recomendación (Amazon, Netflix, etc.)
- Estudiar cómo aprende el cerebro humano.

Aprendizaje de máquina

El *aprendizaje de máquina* es el campo de estudio que dota a las computadoras de la habilidad de aprender sin haber sido programadas explícitamente.

Definición

Se dice que un programa de computadora aprende de la experiencia E , con respecto a una tarea T , y una medida de desempeño D si su desempeño en T , como lo mide D , mejora con la experiencia E .^a

^aMitchell 1997

Juego de damas:

T = la tarea de jugar damas.

D = la probabilidad de que el programa gane el próximo juego.

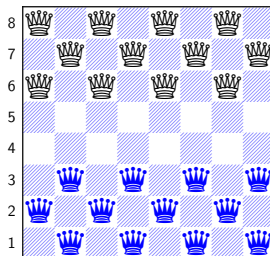


Figura: Juego de damas.

Hipótesis

El aprendizaje se realizará con respecto a una familia de *hipótesis*.


Clasificación de los conjuntos de datos

Para entrenar un algoritmo de aprendizaje se requieren datos de entrenamiento, separados en tres conjuntos:

Entrenamiento Datos con los cuales se ajustan los parámetros de la hipótesis.

Validación Datos utilizados para ajustar los parámetros del algoritmo de entrenamiento.

Prueba Datos utilizados para evaluar la posibilidad de que la hipótesis aprendida generalice^[1] a datos no vistos anteriormente.

^[1]Es decir, que sea válida también para datos nuevos. 

Temas

1 Introducción

- Tipos de aprendizaje

2 Sesgo inductivo

Aprendizaje supervisado

Decimos que el aprendizaje es *supervisado* si para cada ejemplo x entre los datos de entrenamiento X se conoce la respuesta correcta y .

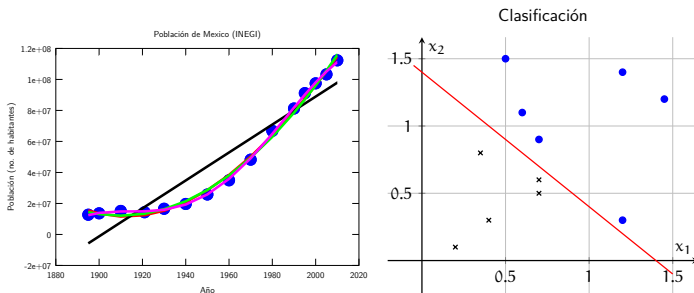


Figura: Ejemplos de aprendizaje supervisado. Izquierda: *Regresión*. Derecha *Clasificación*

Regresión

Un modelo de *regresión* busca predecir valores de salida **continuos**.

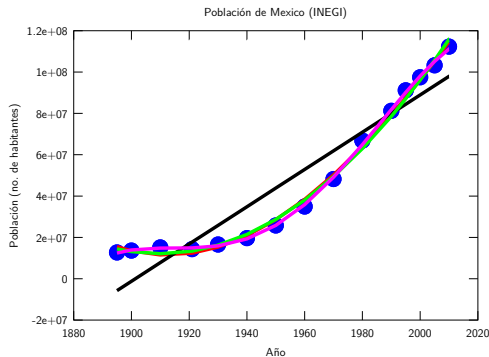


Figura: Función de *Regresión*: Dada la entrada $x = (\text{tamaño})$, ¿en qué precio se puede vender la casa?.

Clasificación

En un problema de *clasificación* se desea predecir una salida discreta.

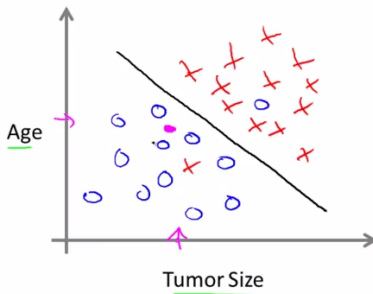


Figura: Problema de *Clasificación*: Dadas las entradas $x = (\text{edad}, \text{tamaño_del_tumor})$, ¿el cáncer es maligno o benigno?

Aprendizaje no supervisado

- No se tienen valores *correctos* o *incorrectos*
- El objetivo del aprendizaje no supervisado es descubrir estructura en los datos.

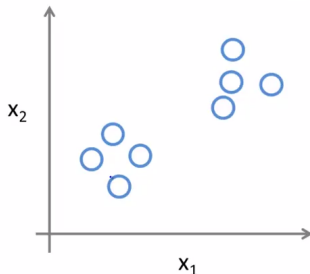
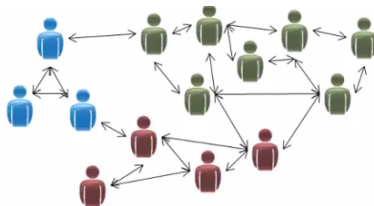


Figura: Aprendizaje no supervisado ¿qué estructura tienen los datos?



Organizar clusters de computadoras



Análisis de redes sociales



Segmentación del mercado



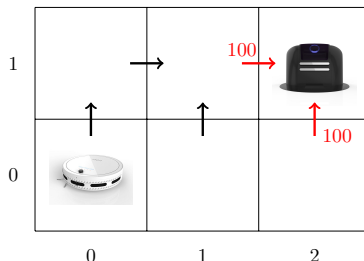
Análisis de datos astronómicos

Figura: Aplicaciones

Aprendizaje por refuerzo

Para cada estado se desea aprender una *política* que indique la acción que maximizará la *recompensa* recibida en el menor tiempo posible.

Política π	
(0,0)	↑
(1,0)	↑
(2,0)	↑
(0,1)	→
(1,1)	→



Temas

- 1 Introducción
 - Tipos de aprendizaje
- 2 Sesgo inductivo

Sesgo inductivo

Definición

Considere un algoritmo de aprendizaje de conceptos L para el conjunto de ejemplares X .

- Sea c un concepto arbitrario definido sobre X y
- sea $D_c = \langle x, c(x) \rangle$ un conjunto arbitrario de ejemplares de entrenamiento de c .
- Sea $L(x_i, D_c)$ la clasificación asignada al ejemplar x_i por L , después de ser entrenado con los datos D_c .

El *sesgo inductivo* de L es cualquier conjunto mínimo de aseveraciones B tales que:

- para cada concepto objetivo c
- y su respectivo conjunto de entrenamiento D_c

$$(\forall x_i \in X)[(B \wedge D_c \wedge x_i) \vdash L(x_i, D_c)] \quad (1)$$

Referencias I



Mitchell, Tom M. (1997). *Machine Learning*. McGrawHill.