

# Introducción a Machine Learning

## Módulo 1

Ronald Cárdenas Acosta

Setiembre, 2016

# Outline

## 1 Introducción

- El curso a grandes rasgos
- ¿Qué es Machine Learning?

## 2 Tipos de aprendizaje

- Aprendizaje Supervisado
- Aprendizaje No Supervisado
- Otros tipos de aprendizaje

# Outline

## 1 Introducción

- El curso a grandes rasgos
- ¿Qué es Machine Learning?

## 2 Tipos de aprendizaje

- Aprendizaje Supervisado
- Aprendizaje No Supervisado
- Otros tipos de aprendizaje

# El curso a grandes rasgos...

- Duración: 32 horas, 8 sesiones, 2 módulos
- Cada sesión con parte teórica y práctica (laboratorio)
- Evaluación:
  - Proyecto Parcial (50%)
  - Proyecto Final (50%)
- Requisitos para certificación:
  - Nota mínima: 14
  - Asistencia mínima: 75
- Conocimientos previos necesarios:
  - Programación en Python
  - Estadística y Probabilidades
  - Álgebra Lineal

# Outline

## 1 Introducción

- El curso a grandes rasgos
- ¿Qué es Machine Learning?

## 2 Tipos de aprendizaje

- Aprendizaje Supervisado
- Aprendizaje No Supervisado
- Otros tipos de aprendizaje

# ¿Qué es Machine Learning?

## Definition

"Conjunto de métodos para análisis de datos...que pueden detectar automáticamente patrones y usarlos para predecir futura data o realizar otro tipo de decisión en un entorno de incertidumbre".

*Kevin P. Murphy*

# ¿Qué es Machine Learning?

## Definition

"Conjunto de métodos para análisis de datos...que pueden detectar automáticamente patrones y usarlos para predecir futura data o realizar otro tipo de decisión en un entorno de incertidumbre".

*Kevin P. Murphy*

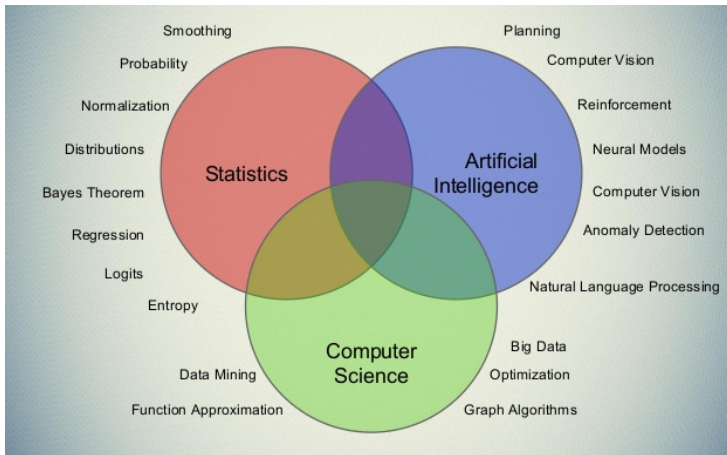
## Definition

"Se dice que un programa de computadora aprende de la experiencia  $E$  con respecto a una tarea  $T$  y métrica de rendimiento  $P$ , si su rendimiento en la tarea  $T$ , medido por  $P$ , mejora con la experiencia  $E$ ".

*Tom Mitchell*

# ¿Qué áreas cubre Machine Learning?

Figure: Machine Learning toma conocimientos de muchas otras áreas.





# Machine Learning y Estadística

- Ambos campos buscan responder a la misma pregunta:  
¿Cómo aprendemos de la data?
- La diferencia radica en los enfoques.
- La estadística se enfoca en la inferencia formal en bajas dimensiones (intervalos de confianza, pruebas de hipótesis, estimadores).
- Machine Learning se enfoca en problemas de inferencia en altas dimensiones.
- Las diferencias entre ambos campos cada vez se hace menor.

# Outline

- 1 Introducción
  - El curso a grandes rasgos
  - ¿Qué es Machine Learning?
- 2 Tipos de aprendizaje
  - **Aprendizaje Supervisado**
  - Aprendizaje No Supervisado
  - Otros tipos de aprendizaje

# Aprendizaje Supervisado

- Objetivo: estimar una función  $f(x, y)$  que mapee un entrada  $x$  a una salida  $y$ , dado un conjunto de pares  $D = (x, y)_{i=1}^N$ .
- $D$  es llamado data de entrenamiento
- $x$  puede ser: una imagen, un documento, datos meteorológicos, etc.
- $y$  puede ser: un objeto reconocido, clase de documento, estimación de parámetro ambiental.

# Clasificación y Regresión

Dos modalidades:

- Clasificación:  $y$  es una clase, una variable cualitativa.  
Ejemplo:  $y \in A, B, C$
- Regresión:  $y$  es un número real, una variable cuantitativa.  
Ejemplo:  $y$  es el precio de una casa, riesgo crediticio, etc.

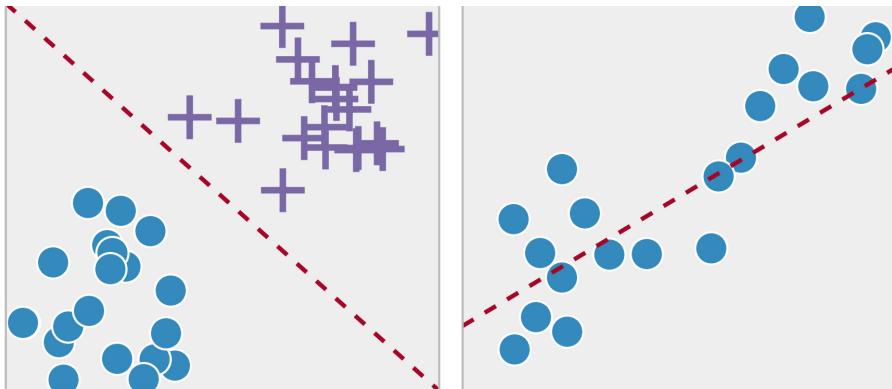


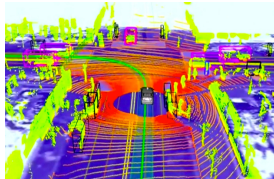
Figure: Clasificación (izquierda) vs Regresión (derecha)

# Ejemplos

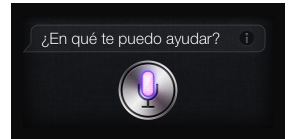


A stop sign is on a road with a mountain in the background.

(a) Descripción automática de imágenes con mecanismos de atención.



(b) Navegación autónoma de vehículos.



(c) Reconocimiento de voz, entendimiento de lenguaje.

# Hipótesis y modelado

## Hipótesis del modelo (ejemplo lineal)

$$f(x, y) \approx g(x, y) * w$$

$$f(x, y) \approx w_0 * x_0 + w_1 * x_1 + ..w_m * x_m$$

Donde:

- $w$ : parámetros del modelo estimar
- $g(x, y)$ : función que representa cada par de entrenamiento como un vector de características

Por ahora:  $g(x, y) = x$

## Optimización

$$\hat{w} = \operatorname{argmax}_w p(y|x, w)$$

$p(y|x, w)$  se puede modelar como una función objetivo o de costo.

# Outline

## 1 Introducción

- El curso a grandes rasgos
- ¿Qué es Machine Learning?

## 2 Tipos de aprendizaje

- Aprendizaje Supervisado
- **Aprendizaje No Supervisado**
- Otros tipos de aprendizaje



# Aprendizaje No Supervisado

- Objetivo: descubrir patrones o estructuras inherentes a la data (*knowledge discovery*)
- Generalmente modelado como estimación de densidad:  $p(x|w)$
- Descubre agrupaciones de la data de acuerdo las características definidas

## Example

Segmentación de imágenes, filtrado de voz, extracción de tópicos semánticos de un documento, clusterización, etc.

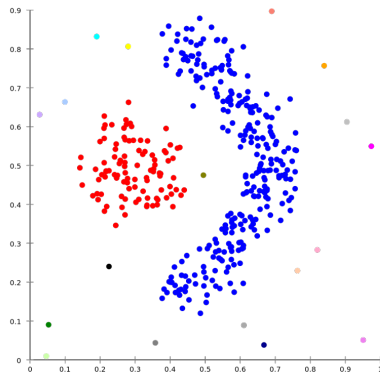
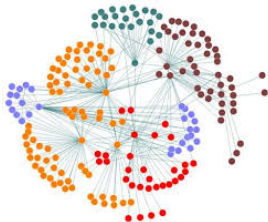


Figure: Aprendizaje no supervisado

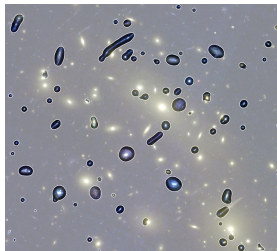
# Ejemplos



(a) Segmentación de mercado



(b) Analisis de redes sociales



(c) Astronomía.  
Agrupación de estrellas y galaxias.

# Outline

## 1 Introducción

- El curso a grandes rasgos
- ¿Qué es Machine Learning?

## 2 Tipos de aprendizaje

- Aprendizaje Supervisado
- Aprendizaje No Supervisado
- Otros tipos de aprendizaje

# Otros tipos de aprendizaje

- Aprendizaje por reforzamiento: agente aprende a tomar acciones que maximicen una función acumulativa de premiado.
- Sistemas de Recomendación: modelo infiere productos de interés para un usuario comparándolo con otros usuarios.