

# Introducción al aprendizaje automático

# Aprendizaje automático

- La idea del aprendizaje automático (*machine learning*) consiste en programar una computadora para que mejore en la realización de una tarea a partir de datos de ejemplo o de la experiencia
- ¿Por qué puede ser necesario?
  - Tareas en las que el algoritmo se debe adaptar a circunstancias particulares (hace falta un “entrenamiento”): detección de Spam, reconocimiento del habla,...
  - Algoritmos difíciles de programar “a mano”: reconocimiento del habla, de escritura manuscrita, reconocimiento de objetos, ...

# Ejemplos de áreas de aplicación

- “Minería de datos” (*data mining*): obtener conclusiones a partir del análisis estadístico de grandes cantidades de datos
- Aplicaciones muy variadas, desde marketing a medicina
- Facilitado por las enormes bases de datos *online* y *offline* existentes en la actualidad

## Customers Who Bought This Item Also Bought



Book Title	Author	Rating	Price
La Conspiración / Deception Point (Spanish Edition)	by Dan Brown	★★★★☆ (4)	\$15.61
Angeles y demonios (Spanish Edition)	by Dan Brown	★★★★★ (25)	\$10.17
La Fortaleza Digital / Digital Fortress (Spanish Edition)	by Dan Brown	★★★★★ (3)	\$16.06
La isla bajo el mar (Vintage Espanol) (Spanish...)	by Isabel Allende	★★★★★ (32)	\$10.74
El Código Da Vinci / Da Vinci Code (Spanish Edition)	by Dan Brown	★★★★★ (1)	\$10.17

parris jilton

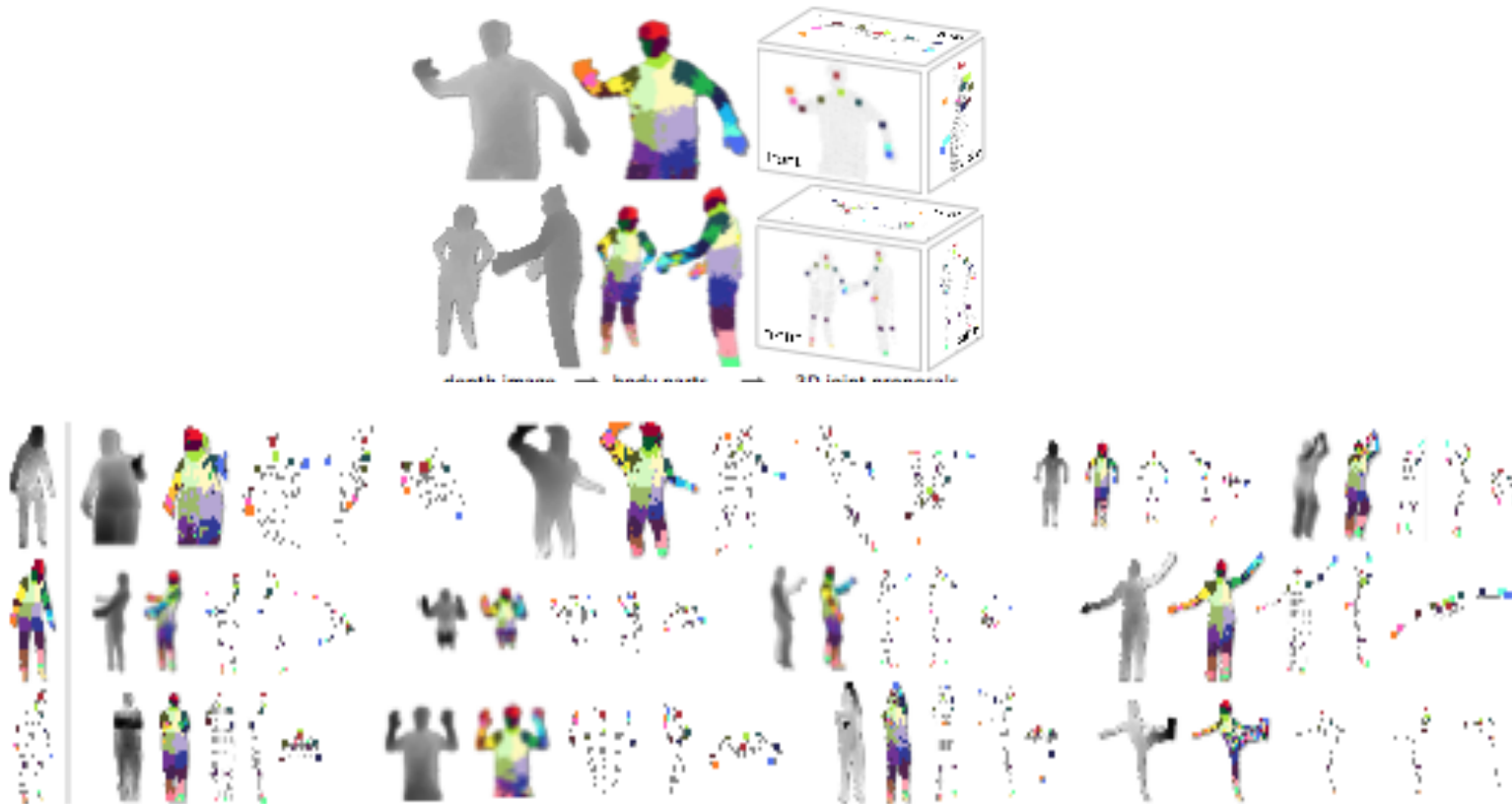
Aproximadamente 49.600 resultados (0,36 segundos)

► Quizás quiso decir: [paris hilton](#)

# Ejemplos de áreas de aplicación (II)

- Algoritmos difíciles de programar “a mano”
  - Reconocimiento de partes del cuerpo con Kinect

<http://research.microsoft.com/pubs/145347/BodyPartRecognition.pdf>





# Tipos de aprendizaje

- **Aprendizaje supervisado:** conocemos la clase a la que pertenece cada ejemplo
  - Clasificación o reconocimiento de patrones: asignamos una clase de un conjunto de clases discreto. Ejemplo: reconocer una cara, un carácter
  - Regresión: estimar una función continua
- **Aprendizaje no supervisado:** tenemos un conjunto de datos que queremos agrupar en *clusters*
- **Aprendizaje por refuerzo:** tenemos una medida de lo bien o mal que está funcionando el algoritmo, pero no sabemos exactamente qué falla.

# Aprendizaje Supervisado

- Aprendemos a partir de ejemplos conocidos (etiquetados según su clase).
- Conjunto de validación:
  - Aleatoriamente se parte el conjunto inicial de ejemplos en dos grupos.
  - Uno se usa como **conjunto de entrenamiento**, para ajustar los parámetros de aprendizaje del clasificador.
  - El otro es el **conjunto de validación** y se usa para estimar el error de generalización.
  - Se entrena hasta alcanzar el **error de validación** mínimo.
  - Entrenar hasta alcanzar un **error de entrenamiento** mínimo produciría sobreentrenamiento (overfitting)
  - Cuando los conjuntos de entrenamiento no son lo suficientemente grandes usamos la técnica de validación cruzada (k-fold cross validation)



# Bibliografía

- Escolano et al. [Inteligencia Artificial](#). Thomson-Paraninfo 2003.
- Duda, Hart & Stork. [Pattern Classification](#). Wiley 2001.
- Hastie, Tibshirani, Friedman, [The Elements of Machine Learning](#). Springer Series in Statistics. 2001.
- Reed, Marks, [Neural Smithing](#). MIT Press, CA Mass 1999
- Mitchell, [Machine Learning](#). McGraw Hill, Computer Science Series. 1997