

SCALAB

Universidad Carlos III de Madrid

Aprendizaje Automático

Demanda laboral

- ▶ <http://blog.indeed.com/2019/03/14/best-jobs-2019/>
- ▶ <https://www.indeed.com>

Cada vez más empresas utilizan técnicas de aprendizaje automático



Definición y objetivos

- ▶ Programas que **mejoran automáticamente** basados en la experiencia de acuerdo a alguna medida

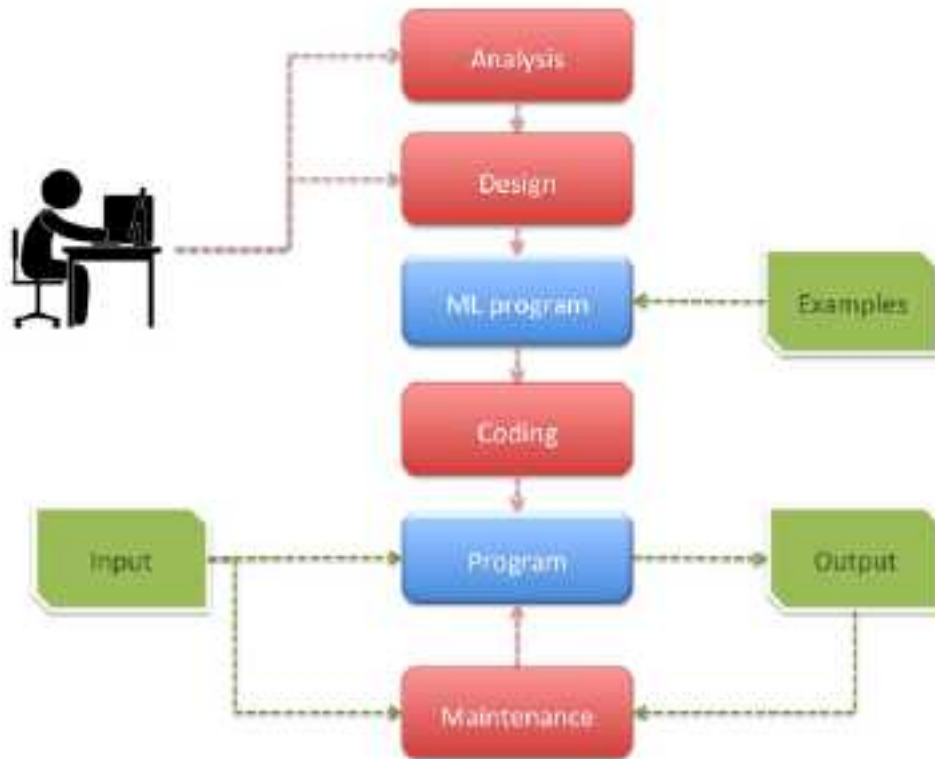
Definición y objetivos

- ▶ Programas que mejoran automáticamente basados en la experiencia de acuerdo a alguna medida
- ▶ Objetivos
 - ▶ Aprender conocimiento nuevo
 - ▶ Mejorar el comportamiento de un sistema

Motivación

- ▶ El **desarrollo de software** es un cuello de botella
- ▶ Introducir **conocimiento a través de ejemplos** es atractivo
- ▶ El aprendizaje automático es útil cuando:
 - ▶ No existe experiencia humana (*navegación en Marte*)
 - ▶ Los humanos no saben explicar su experiencia (*reconocimiento de habla*)
 - ▶ Los modelos deben ser personalizados/adaptados (*medicina personalizada*)
 - ▶ Los modelos se basan en enormes cantidades de datos (*genómica*)

Programación mediante aprendizaje automático



Algunos ejemplos

- ▶ **Sanidad**: cómo diagnosticar mejor una enfermedad

Algunos ejemplos

- ▶ **Sanidad**: cómo diagnosticar mejor una enfermedad
- ▶ **Ciudades inteligentes**: cómo autoregular eficientemente el consumo de energía

Algunos ejemplos

- ▶ **Sanidad:** cómo diagnosticar mejor una enfermedad
- ▶ **Ciudades inteligentes:** cómo autoregular eficientemente el consumo de energía
- ▶ **Domótica:** cuándo apagar las luces

Algunos ejemplos

- ▶ **Sanidad:** cómo diagnosticar mejor una enfermedad
- ▶ **Ciudades inteligentes:** cómo autoregular eficientemente el consumo de energía
- ▶ **Domótica:** cuándo apagar las luces
- ▶ **Banca:** por qué conceder un crédito hipotecario, cuando aceptar el cargo a una tarjeta

Algunos ejemplos

- ▶ **Sanidad:** cómo diagnosticar mejor una enfermedad
- ▶ **Ciudades inteligentes:** cómo autoregular eficientemente el consumo de energía
- ▶ **Domótica:** cuándo apagar las luces
- ▶ **Banca:** por qué conceder un crédito hipotecario, cuando aceptar el cargo a una tarjeta
- ▶ **Marketing:** qué tipo de personas compran cerveza los viernes, qué perfil de compra tiene la familia Martínez

Algunos ejemplos

- ▶ **Sanidad:** cómo diagnosticar mejor una enfermedad
- ▶ **Ciudades inteligentes:** cómo autoregular eficientemente el consumo de energía
- ▶ **Domótica:** cuándo apagar las luces
- ▶ **Banca:** por qué conceder un crédito hipotecario, cuando aceptar el cargo a una tarjeta
- ▶ **Marketing:** qué tipo de personas compran cerveza los viernes, qué perfil de compra tiene la familia Martínez
- ▶ **Personalización:** qué tipo de noticias le gusta leer los viernes por la mañana al usuario

Algunos ejemplos

- ▶ **Sanidad:** cómo diagnosticar mejor una enfermedad
- ▶ **Ciudades inteligentes:** cómo autoregular eficientemente el consumo de energía
- ▶ **Domótica:** cuándo apagar las luces
- ▶ **Banca:** por qué conceder un crédito hipotecario, cuando aceptar el cargo a una tarjeta
- ▶ **Marketing:** qué tipo de personas compran cerveza los viernes, qué perfil de compra tiene la familia Martínez
- ▶ **Personalización:** qué tipo de noticias le gusta leer los viernes por la mañana al usuario
- ▶ **Robots:** cómo programar un robot para coger objetos

Algunos ejemplos

- ▶ **Sanidad:** cómo diagnosticar mejor una enfermedad
- ▶ **Ciudades inteligentes:** cómo autoregular eficientemente el consumo de energía
- ▶ **Domótica:** cuándo apagar las luces
- ▶ **Banca:** por qué conceder un crédito hipotecario, cuando aceptar el cargo a una tarjeta
- ▶ **Marketing:** qué tipo de personas compran cerveza los viernes, qué perfil de compra tiene la familia Martínez
- ▶ **Personalización:** qué tipo de noticias le gusta leer los viernes por la mañana al usuario
- ▶ **Robots:** cómo programar un robot para coger objetos
- ▶ **Seguridad:** qué perfil de uso de Linux tiene un determinado usuario, cuándo se produce una entrada ilegal

Algunos ejemplos

- ▶ **Sanidad:** cómo diagnosticar mejor una enfermedad
- ▶ **Ciudades inteligentes:** cómo autoregular eficientemente el consumo de energía
- ▶ **Domótica:** cuándo apagar las luces
- ▶ **Banca:** por qué conceder un crédito hipotecario, cuando aceptar el cargo a una tarjeta
- ▶ **Marketing:** qué tipo de personas compran cerveza los viernes, qué perfil de compra tiene la familia Martínez
- ▶ **Personalización:** qué tipo de noticias le gusta leer los viernes por la mañana al usuario
- ▶ **Robots:** cómo programar un robot para coger objetos
- ▶ **Seguridad:** qué perfil de uso de Linux tiene un determinado usuario, cuándo se produce una entrada ilegal
- ▶ **Videojuegos:** en qué orden le gusta explorar el terreno al usuario

Algunos ejemplos

- ▶ **Sanidad:** cómo diagnosticar mejor una enfermedad
- ▶ **Ciudades inteligentes:** cómo autoregular eficientemente el consumo de energía
- ▶ **Domótica:** cuándo apagar las luces
- ▶ **Banca:** por qué conceder un crédito hipotecario, cuando aceptar el cargo a una tarjeta
- ▶ **Marketing:** qué tipo de personas compran cerveza los viernes, qué perfil de compra tiene la familia Martínez
- ▶ **Personalización:** qué tipo de noticias le gusta leer los viernes por la mañana al usuario
- ▶ **Robots:** cómo programar un robot para coger objetos
- ▶ **Seguridad:** qué perfil de uso de Linux tiene un determinado usuario, cuándo se produce una entrada ilegal
- ▶ **Videojuegos:** en qué orden le gusta explorar el terreno al usuario
- ▶ **Astromomía:** determinar si una imagen contiene algo interesante

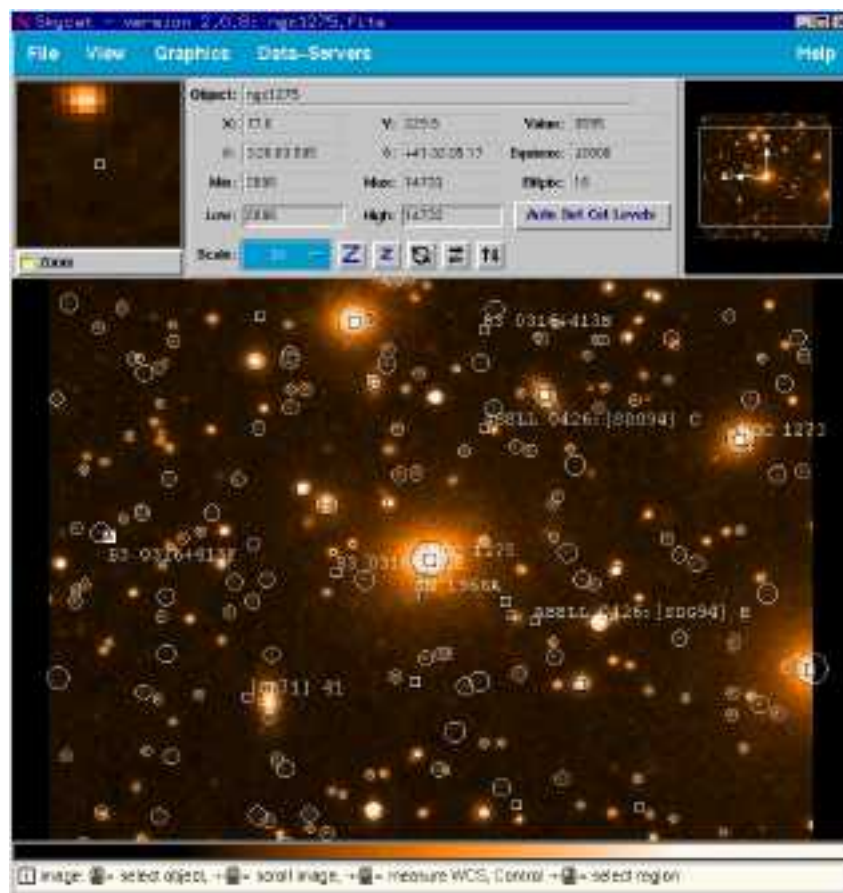
Algunos ejemplos

- ▶ **Sanidad:** cómo diagnosticar mejor una enfermedad
- ▶ **Ciudades inteligentes:** cómo autoregular eficientemente el consumo de energía
- ▶ **Domótica:** cuándo apagar las luces
- ▶ **Banca:** por qué conceder un crédito hipotecario, cuando aceptar el cargo a una tarjeta
- ▶ **Marketing:** qué tipo de personas compran cerveza los viernes, qué perfil de compra tiene la familia Martínez
- ▶ **Personalización:** qué tipo de noticias le gusta leer los viernes por la mañana al usuario
- ▶ **Robots:** cómo programar un robot para coger objetos
- ▶ **Seguridad:** qué perfil de uso de Linux tiene un determinado usuario, cuándo se produce una entrada ilegal
- ▶ **Videojuegos:** en qué orden le gusta explorar el terreno al usuario
- ▶ **Astromomía:** determinar si una imagen contiene algo interesante
- ▶ **Análisis de imágenes:** determinar si una imagen dada contiene una determinada persona

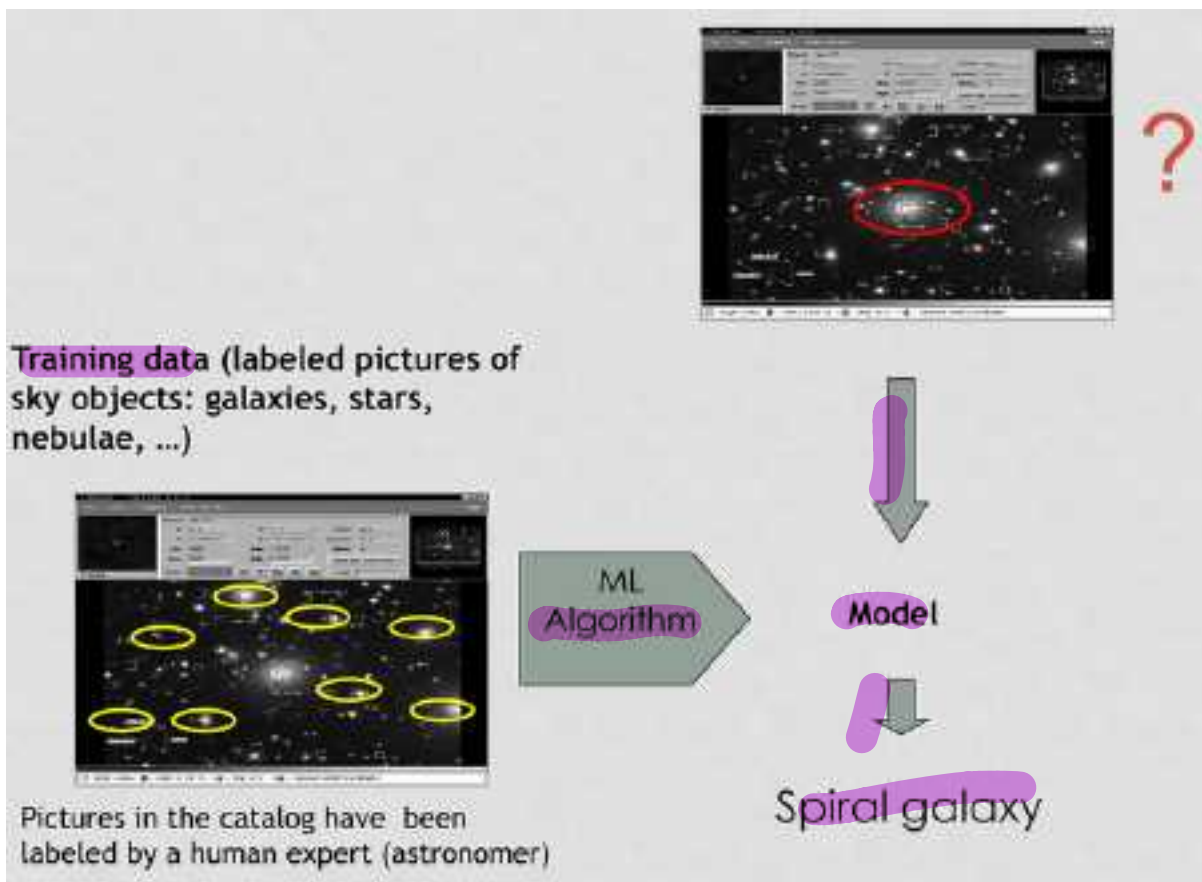
Skycat. Catalogación de imágenes del espacio

- ▶ Los expertos humanos no pueden mantener el ritmo de generación actual de imágenes por día
- ▶ Ejemplo: el segundo estudio celeste realizado en el observatorio del monte Palomar ha generado tres terabytes de datos en imágenes
- ▶ Problema: cómo catalogar automáticamente los dos billones de objetos
- ▶ Solución: aprendizaje automático aplicada a la catalogación
- ▶ Resultado: 94% de acierto en la catalogación
- ▶ El catálogo es tres veces más grande de lo que se podría tener sin ayuda del sistema
- ▶ <http://www.eso.org/sci/observing/tools/skycat.html>

Skycat



Skycat



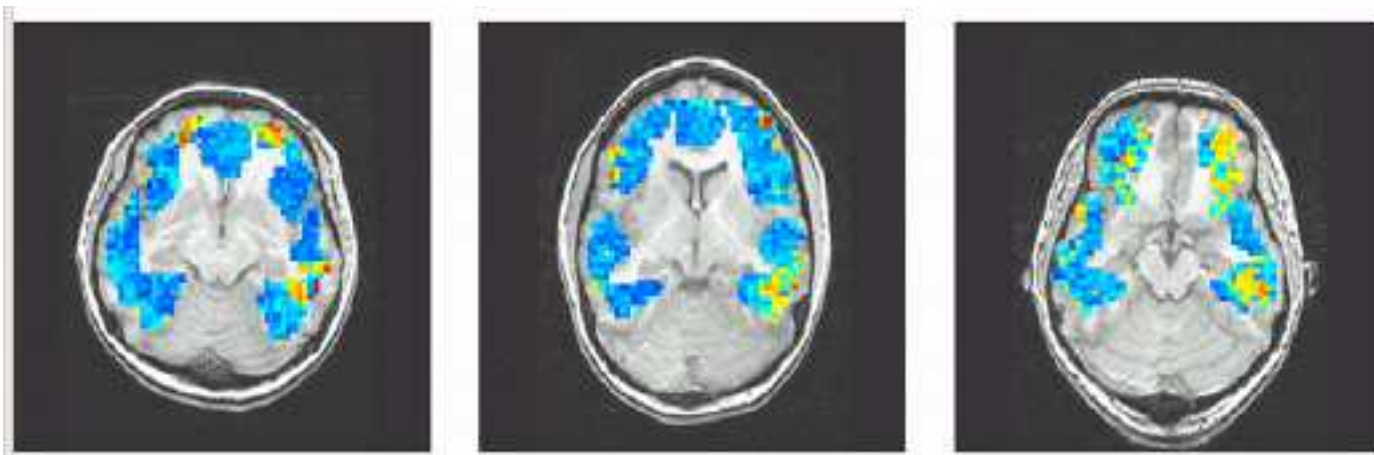
Análisis de imágenes cerebrales

Un fMRI permite obtener imágenes de la actividad de un cerebro humano

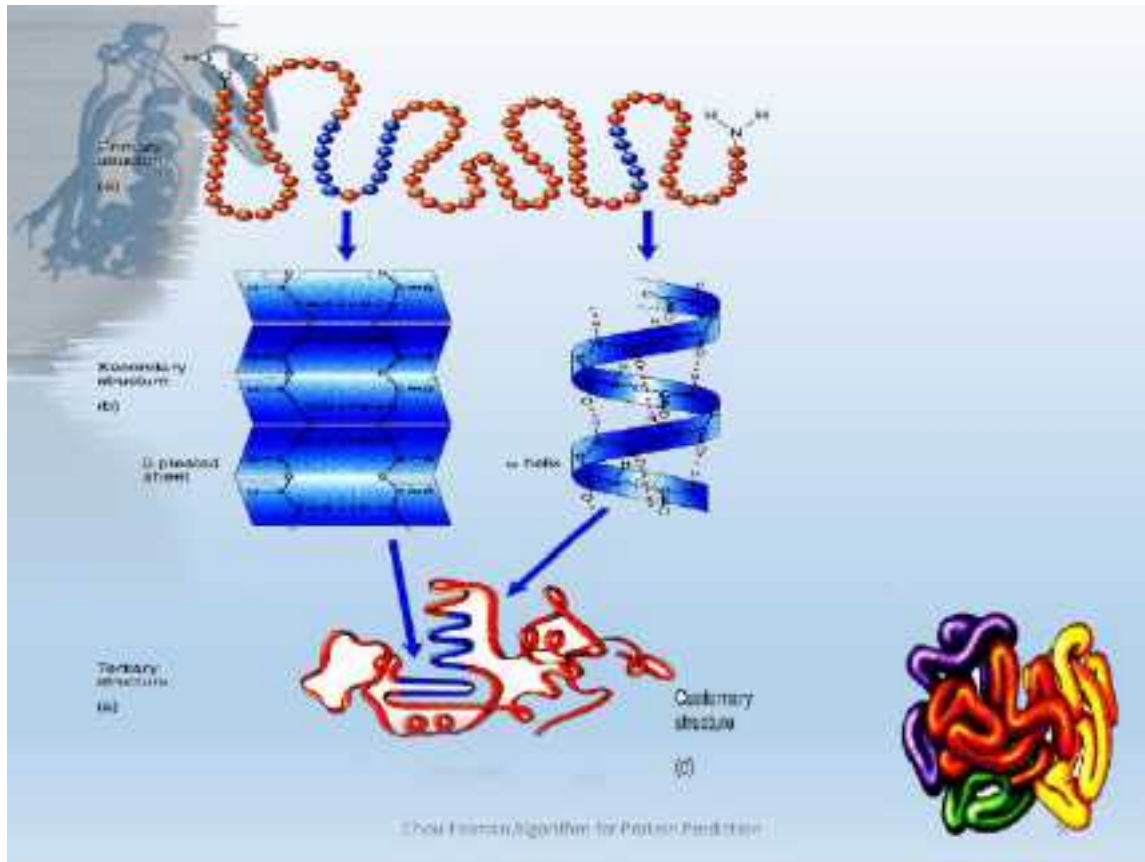
- ▶ Nuevas oportunidades para estudiar los mecanismos cerebrales
- ▶ Ejemplo:
 - ▶ decodificar si las palabras que está leyendo un humano son sobre herramientas, edificios, comida, u otras categorías semánticas
 - ▶ el clasificador obtenido acierta la categoría semántica un 90% de las veces cuando, por ejemplo, el sujeto en estudio lee palabras sobre herramientas o edificios
- ▶ <http://www.cs.cmu.edu/afs/cs.cmu.edu/project/theo-73/www/index.html>

Análisis de imágenes cerebrales

La figura muestra, para tres sujetos distintos, el grado con que diferentes localizaciones del cerebro ayudan a predecir la categoría semántica de la palabra. Las zonas rojas y amarillas son las que más ayudan. Las regiones predictivas en cada sujeto son similares.



Predicción de la estructura secundaria de las proteínas



Minería de datos (data mining) → Big data

► Minería de datos

- ¿Cuál es el perfil de los clientes que se gastan al mes más de 10 Keuros?
- ¿Qué producto de nuestra empresa es el que compran los clientes junto con pañales?
- ¿Debo invertir en la empresa X mañana y cuánto?
- ¿Qué libro recomendar después de haber leído X?
- ¿A qué tipos de números de teléfono llaman nuestros clientes?
- ¿Cuándo concedo un crédito hipotecario? ¿por cuánto?

Minería de datos (data mining) → Big data

► Minería de datos

- ¿Cuál es el perfil de los clientes que se gastan al mes más de 10 Keuros?
- ¿Qué producto de nuestra empresa es el que compran los clientes junto con pañales?
- ¿Debo invertir en la empresa X mañana y cuánto?
- ¿Qué libro recomendar después de haber leído X?
- ¿A qué tipos de números de teléfono llaman nuestros clientes?
- ¿Cuándo concedo un crédito hipotecario? ¿por cuánto?

► Big data

- ¿Cuál es el transporte público más utilizado en un año?
- ¿Qué patrones de tráfico son raros a partir de datos GPS?
- ¿Cómo controlar los semáforos para reducir el consumo de energía?

Creditos bancarios



Marketing y compras



Text mining: regularidades en textos

Aprende de textos en leng. natural, a base de ejemplos y es capaz de tomar decisiones.

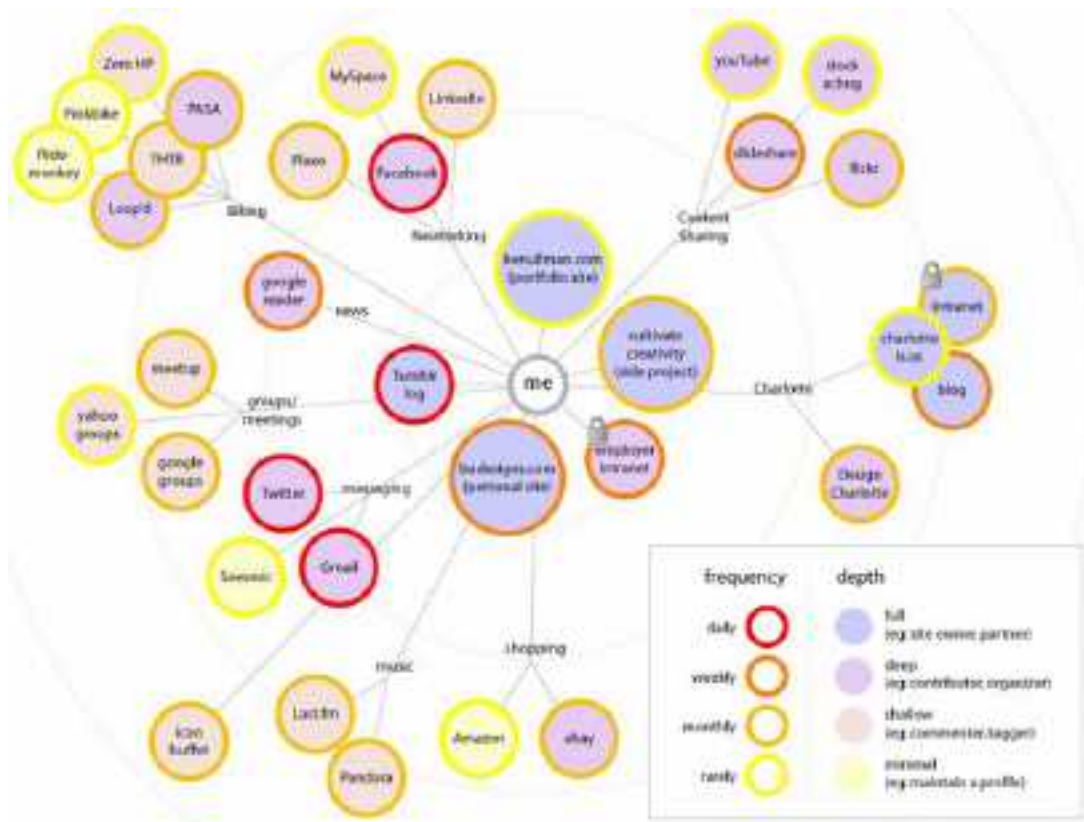
Atributo - valor
 Edad Salario gasto ...
 ○ ○ ○

- ▶ Sistemas automáticos de filtrado (email, noticias)
- ▶ Noticias personalizadas
- ▶ Análisis de opiniones
- ▶ Construcción de buscadores semánticos
- ▶ Buscadores personalizados
- ▶ Venta personalizada

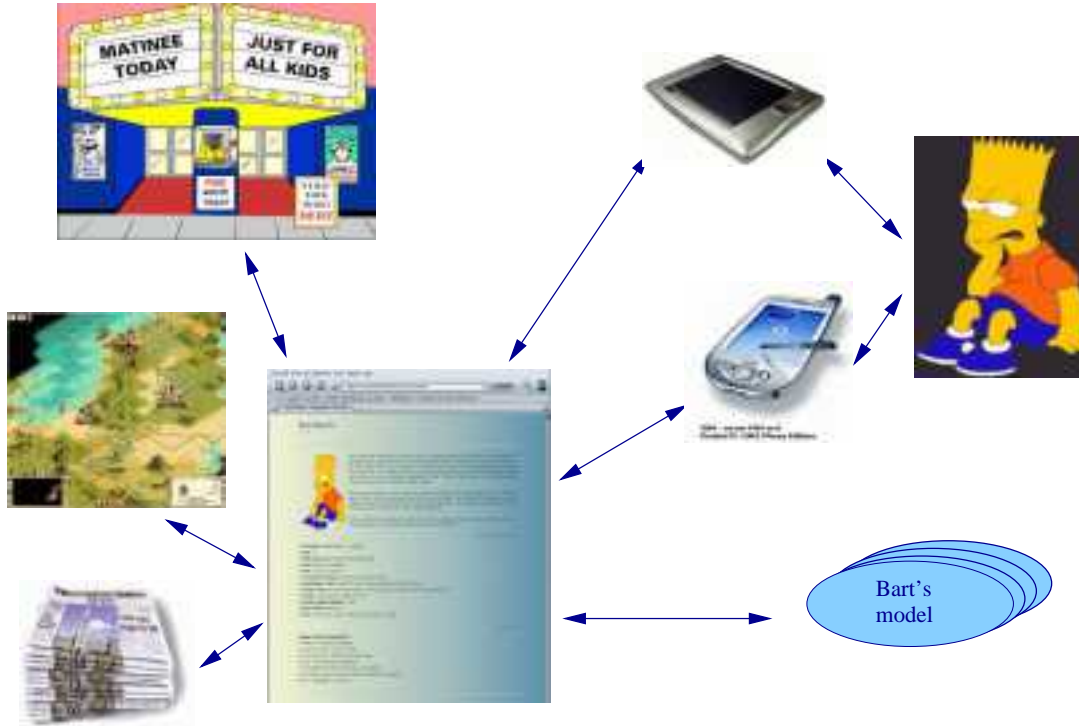
Ejemplos → AA
 ↓
textos, en lenguaje natural

Un humano hace una muestra de ejemplos.

Social Mining (redes sociales)



Personalización



RALPH: aprendiendo a conducir

*Aprendizaje por
refuerzo*



Visión

Deep learning
Redes neuronales
con muchas capas.



a street sign on a pole in front of a building



a plate with a sandwich and a salad

<https://cs.stanford.edu/people/karpathy/neuraltalk2/demo.html>

Visión

*No siempre
funciona
correctamente*



a man riding a bike on a beach with a dog
in the water

<https://cs.stanford.edu/people/karpathy/neuraltalk2/demo.html>

Visión



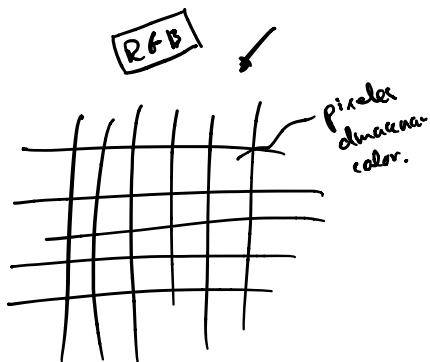
a man riding a skateboard down a street

<https://cs.stanford.edu/people/karpathy/neuraltalk2/demo.html>

Reconocimiento de imágenes



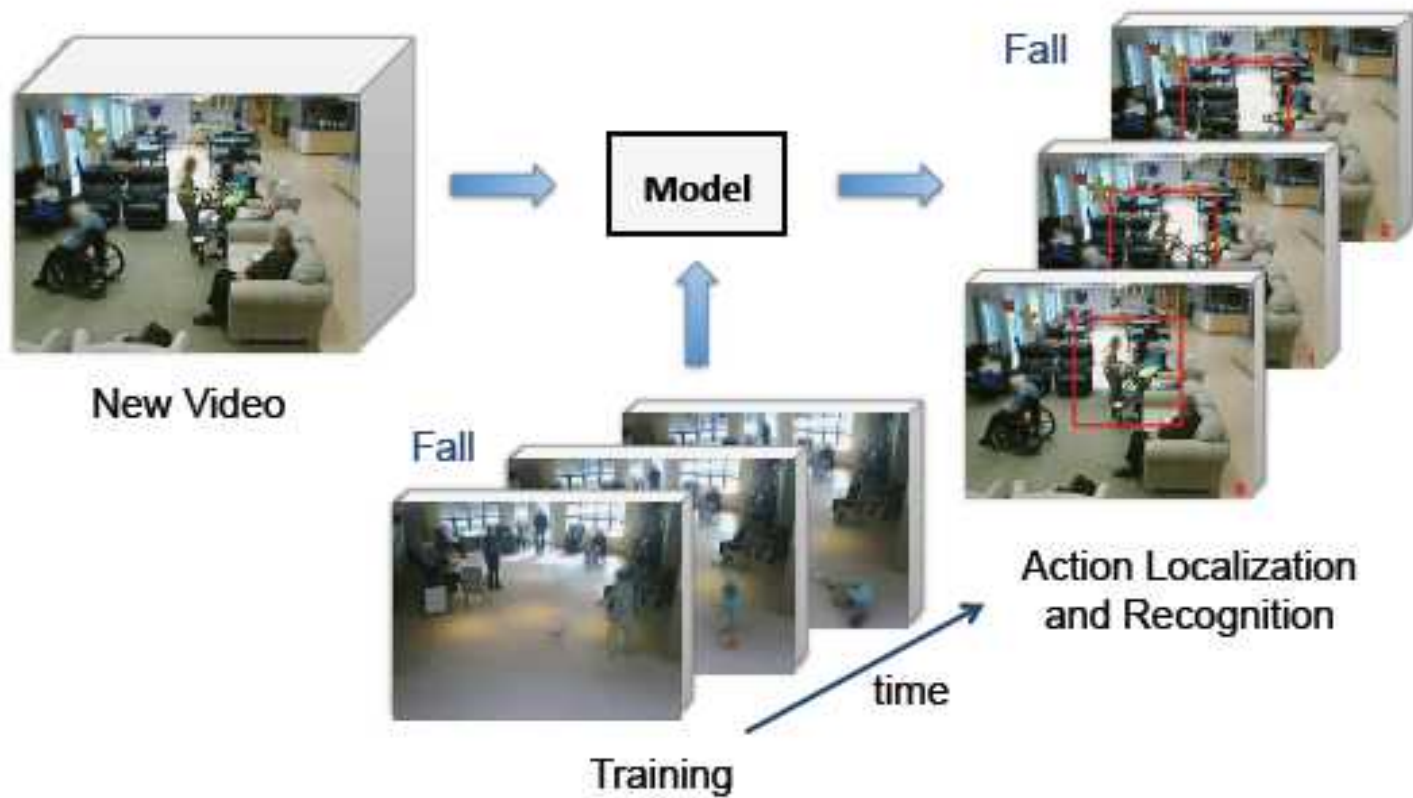
Reconocimiento de imágenes



Atributo - valor abstracta



Reconocimiento de actividades



Tareas de Aprendizaje Automático

- ▶ **Clasificación** *Supervisado*
dados unos datos sobre un cliente, determinar si devolverá el crédito o no
- ▶ **Predicción - Regresión** *~ algo numerico*
dados varios valores de la bolsa de varios días, determinar el valor que tendrá un determinado índice mañana
- ▶ **Agrupación - Clustering**
dados datos sobre las compras realizadas por un conjunto de personas, determinar qué subconjuntos de personas se parecen entre sí
no supervisado

Tareas de aprendizaje automático (II)

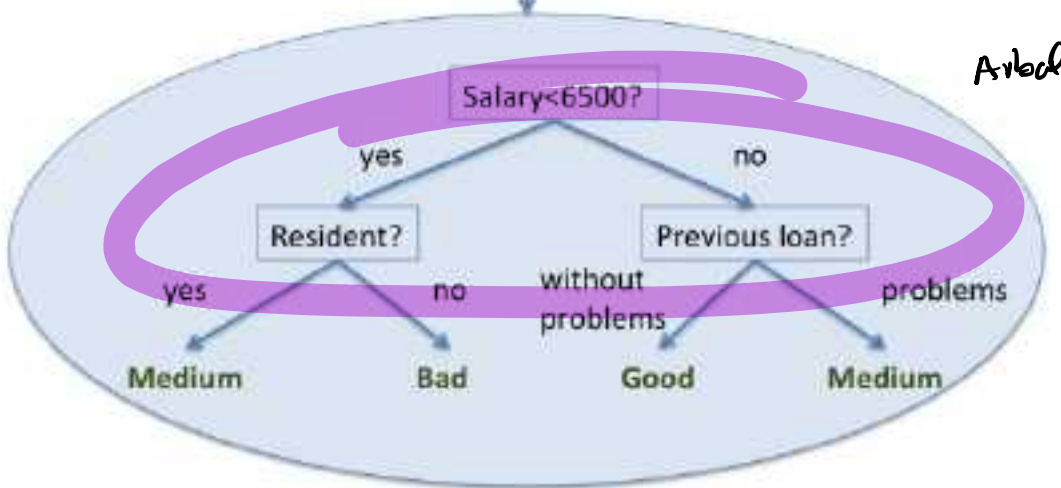
- ▶ **Diagnóstico** *→ similar a agrupación.*
dados unos síntomas de un paciente, determinar la enfermedad que tiene
- ▶ **Caracterización**
dados varios conjuntos de personas que compran de forma parecida, determinar en qué se parecen entre sí
- ▶ **Selección de acciones** *→ Aprendizaje por refuerzo.*
dados varios ejemplos de ejecución de acciones en un entorno determinar cuál es la mejor acción a ejecutar para cada estado
- ▶ **Mejorar la eficiencia**
dado un sistema que planifica las rutas óptimas de un conjunto de transportes para llevar paquetes de un lugar a otro, conseguir que lo haga en menos tiempo y con mejores soluciones

Algunas técnicas de clasificación: árboles de decisión

Tabla Atributo-valor de ejemplos.

Instances	Age	Salary	...	Postal code	Class
Instance 1	30	6000	...	28911	Good
Instance 2	45	8000	...	28001	Bad
...
Instance E	23	1200	...	08002	Medium

ID3, C4.5



Algunas técnicas de clasificación: reglas

Instances	Age	Salary	...	Postal code	Class
Instance 1	30	6000	...	28911	Good
Instance 2	45	8000	...	28001	Bad
...
Instance E	23	1200	...	08002	Medium

C4.5, AQ, CN2

Condiciones

If Salary<6500=yes AND Resident=yes
Then Class=Medium
If Salary<6500=yes AND Resident=no
Then Class=Bad
If Salary<6500=no AND Previous Loan=yes
Then Class=Good
If Salary<6500=no AND Previous Loan=no
Then Class=Medium

Algunas técnicas de clasificación: bayesiano

Instances	Age	Salary	...	Postal code	Class
Instance 1	30	6000	...	28911	Good
Instance 2	45	8000	...	28001	Bad
...
Instance E	23	1200	...	08002	Medium

Naïve Bayes

Class	P(Class)	Resident	Class	P(Resident=V Class)
Good	0.6	yes	Good	0.7
Bad	0.2	yes	Bad	0.1
Medium	0.2	yes	Medium	0.2
		no	Good	0.3
		no	Bad	0.3
		no	Medium	0.4



Algunas técnicas de clasificación: aprendizaje basado en instancias (IBL)

Instances	Age	Salary	...	Postal code	Class
Instance 1	30	6000	...	28911	Good
Instance 2	45	8000	...	28001	Bad
...
Instance E	23	1200	...	08002	Medium

IBL

Instances	Age	Salary	...	Postal code	Class
Instance 1	30	6000	...	28911	Good
Instance 2	45	8000	...	28001	Bad
...
Instance E	23	1200	...	08002	Medium

Mediante comparación
entre ejemplos resueltos
y los casos nuevos

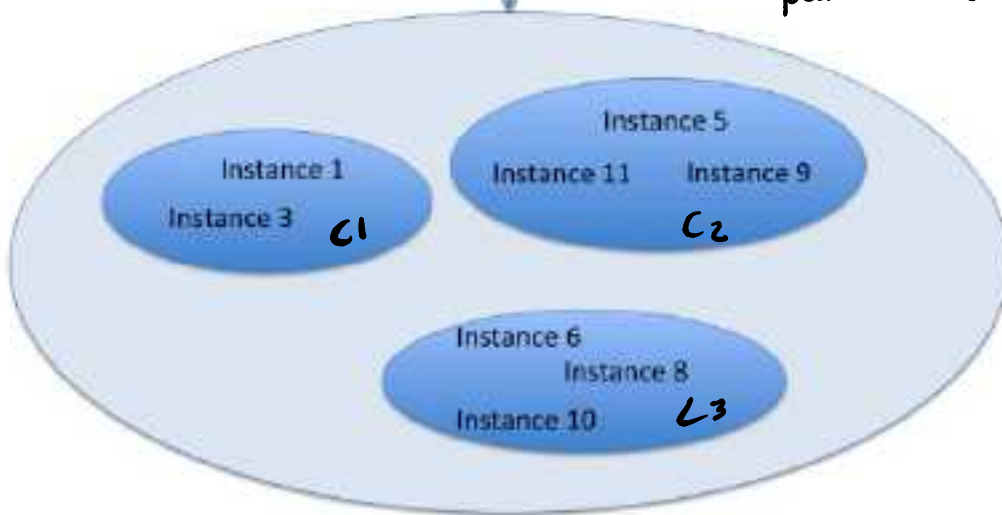
Agrupamiento: k-medias

↳ No supervisado.

Instances	Age	Salary	...	Postal code
Instance 1	30	6000	...	28911
Instance 2	45	8000	...	28001
...
Instance E	23	1200	...	08002

Clustering

Analiza cuáles H
parecen más



Selección de acciones: aprendizaje por refuerzo

Por debajo
Hay un MDP para hallar la política óptima.
Proceso de
Decision de
Markov

- ▶ Mario Bros: https://www.youtube.com/watch?v=oC3IJADC_1s
- ▶ Péndulo: <https://www.youtube.com/watch?v=Lt-KLtkDlh8>
- ▶ Robot: <https://www.youtube.com/watch?v=SJS22jQTdSU>
- ▶ Helicóptero: https://www.youtube.com/watch?v=3mvE_CeH9q0
- ▶ Imitación: <https://www.youtube.com/watch?v=qtqubguikMk>