

30231 - Aprendizaje Automático

Curso 2014-2015

■ Profesores:

Juan D. Tardós D1.16: M 9-11 y 12-13h, J 10-13h
tardos@unizar.es

Luis Montesano D0.15: L 12-14h, M 10-13h, X 16-17
montesano@unizar.es

■ Material (teoría y prácticas): <http://moodle2.unizar.es/>

Índice

- ¿Qué es el aprendizaje automático?
- Aprendizaje supervisado
- Aprendizaje no supervisado
- Organización de la asignatura
 - Programa
 - Prácticas
 - Evaluación
 - Bibliografía

Créditos de transparencias y figuras:

- Andrew Ng, *Machine Learning*, Stanford AI Lab. <https://www.coursera.org/course/ml>
- Kevin P. Murphy, *Machine Learning: A Probabilistic Perspective*, MIT Press, 2012

¿Qué es el Aprendizaje Automático?

- Arthur Samuel (1959):
 - Campo de estudio que proporciona a los computadores la capacidad de aprender sin ser programados explícitamente
- Tom Mitchell (1998):
 - Un programa aprende de la experiencia E ,
 - con respecto a una tarea T
 - y una medida de prestaciones P
 - si sus prestaciones en T , medidas por P , mejoran con la experiencia E
- Kevin P. Murphy (2012):
 - Conjunto de métodos capaces de descubrir automáticamente patrones en datos y usarlos para predecir datos futuros o para tomar decisiones bajo incertidumbre.

Estamos en la Era de los Datos

- El mundo es inmensamente rico en datos
 - La web: 10^{12} páginas web, +10 años de video al día en Youtube
 - Se ha secuenciado el genoma humano ($1000s \times 3.8 \times 10^9$ pares)
 - Bases de datos químicas, farmacéuticas, financieras, ...
- Objetivo del aprendizaje: sacar sentido a los datos, extraer información útil de ellos.
- Innumerables aplicaciones usan modelos aprendidos a partir de datos:
 - Google: para saber cómo responder mejor a una búsqueda, para minería de la web, para traducción automática, etc.
 - Netflix: para saber qué DVD recomendarte
 - Amazon: para saber cómo colocar los productos en su página web de forma que los clientes respondan mejor.

Taxonomía del Aprendizaje

¿Qué?	Parámetros	Estructura	Conceptos ocultos	
¿De dónde?	Supervisado	No supervisado	Por Refuerzo	Semi-Supervisado
¿Para qué?	Predicción	Diagnóstico	Resumen	Descubrimiento
¿Cómo?	Pasivo	Activo	Online	Offline
¿Salida?	Clasificación	Regresión	Clustering	
¿Detalles?	Generativo	Discriminativo	Suavizado	

Aprendizaje Supervisado

- Datos de entrenamiento: parejas entrada-salida conocida

$$\mathcal{D} = \{(\mathbf{x}^{(i)}, y^{(i)})\}_{i=1}^N$$

- La salida es generada por una función desconocida $y = f(\mathbf{x})$

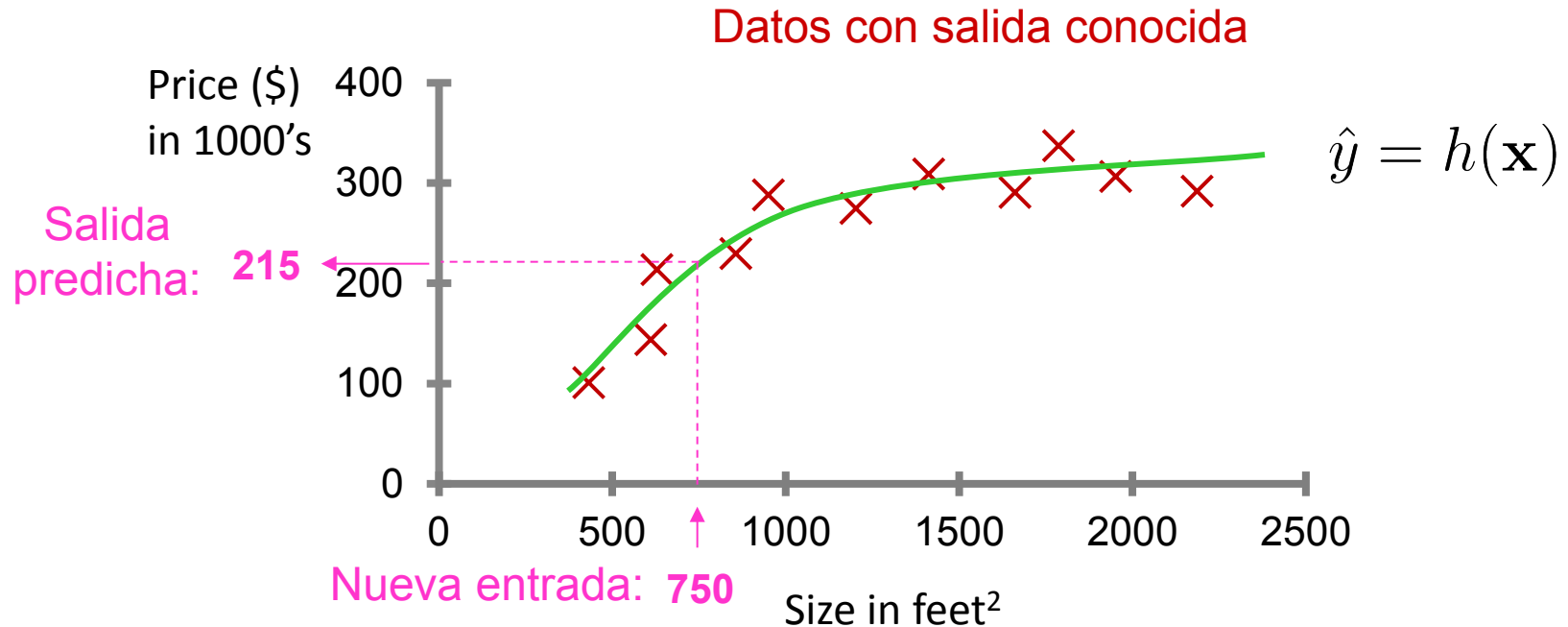
Atributos (<i>features</i>)				Salida	} Datos de Entrenamiento
$x_1^{(1)}$	$x_2^{(1)}$	\dots	$x_D^{(1)}$	$\rightarrow y^{(1)}$	
$x_1^{(2)}$	$x_2^{(2)}$	\dots	$x_D^{(2)}$	$\rightarrow y^{(2)}$	
\vdots	\vdots	\ddots	\vdots	\vdots	
$x_1^{(N)}$	$x_2^{(N)}$	\dots	$x_D^{(N)}$	$\rightarrow y^{(N)}$	

- Descubrir una función h que aproxime la función real f y que permita predecir la salida para datos futuros

$$\hat{y} = h(\mathbf{x})$$

Aprendizaje Supervisado: Regresión

- Regresión: salida continua










- Atributos: superficie, nº habitaciones, nº baños, antigüedad,....

Ejemplos de Regresión

- Predecir el valor de un coche de segunda mano
 - Para el vendedor
 - Para detectar chollos

16 anuncios de coches CITROEN C3 de segunda mano en Zaragoza

Fecha	Foto	Marca y Modelo	Provincia	Combustible	Km	Año	Precio
Ayer 15:52		★ CITROEN C3 1.4 HDi Exclusive 5p.	Zaragoza	Diesel	120.000	2004	4.000 €
Ayer 12:06		★ CITROEN C3 1.4 HDi Furio 5p.	Zaragoza	Diesel	26.700	2009	9.500 €
10/02 12:39		★ CITROEN C3 HDi 90 SX 5p.	Zaragoza	Diesel	37.000	2010	9.500 €
08/02 13:00		★ CITROEN C3 1.4 HDi Furio 5p.	Zaragoza	Diesel	70.000	2009	7.500 € <small>financiado</small>
08/02 12:26		★ CITROEN C3 1.4i Collection 5p.	Zaragoza	Gasolina	95.500	2006	5.800 €
05/02 20:13		★ CITROEN C3 1.4 HDi Furio 5p.	Zaragoza	Diesel	70.000	2009	7.500 € <small>financiado</small>
05/02 11:31		★ CITROEN C3 1.4 HDi Audace 5p.	Zaragoza	Diesel	67.120	2008	6.500 €

Ejemplos de Regresión

- Predecir la temperatura en cualquier punto de un edificio en función de la temperatura externa, hora, sensores en las puertas,...
- Predecir el número de unidades que se van a vender de un producto
- Predecir el número de días que va a pasar cada paciente en el hospital el próximo año
 - ♦ Premio: \$3.000.000 - <http://www.heritagehealthprize.com/c/hhp>
- Predecir el valor de la bolsa mañana en función de las condiciones actuales y otros factores
 - ♦ “Difficult to see. Always in motion is the future” (Yoda)

Aprendizaje Supervisado: Clasificación

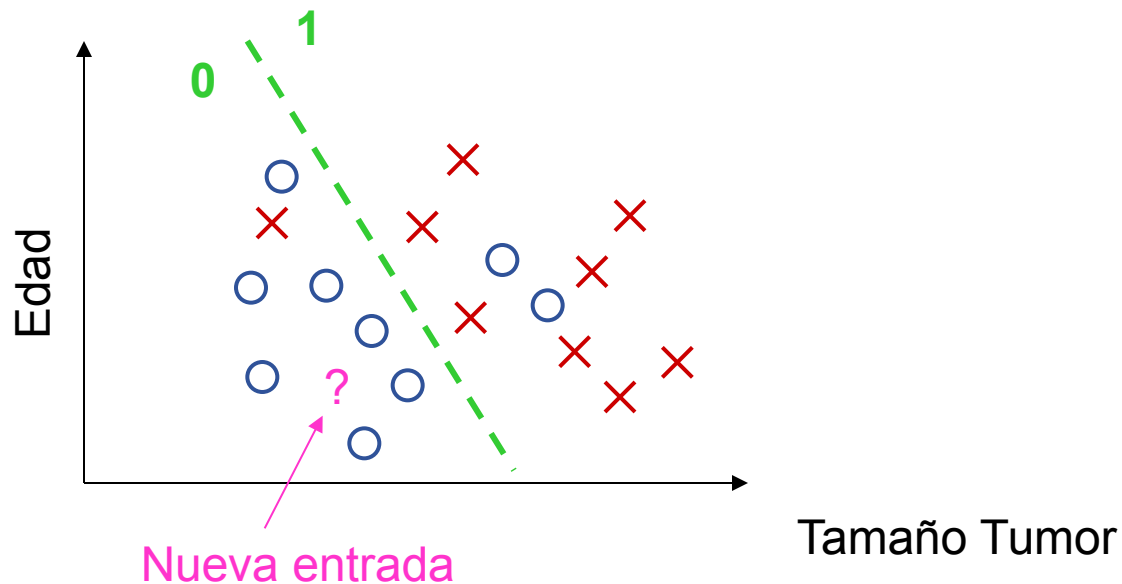
■ Clasificación: salida discreta

- Binaria: $y \in \{0, 1\}$
- Multi-clase: $y \in \{1, \dots, C\}$



Aprendizaje Supervisado: Clasificación

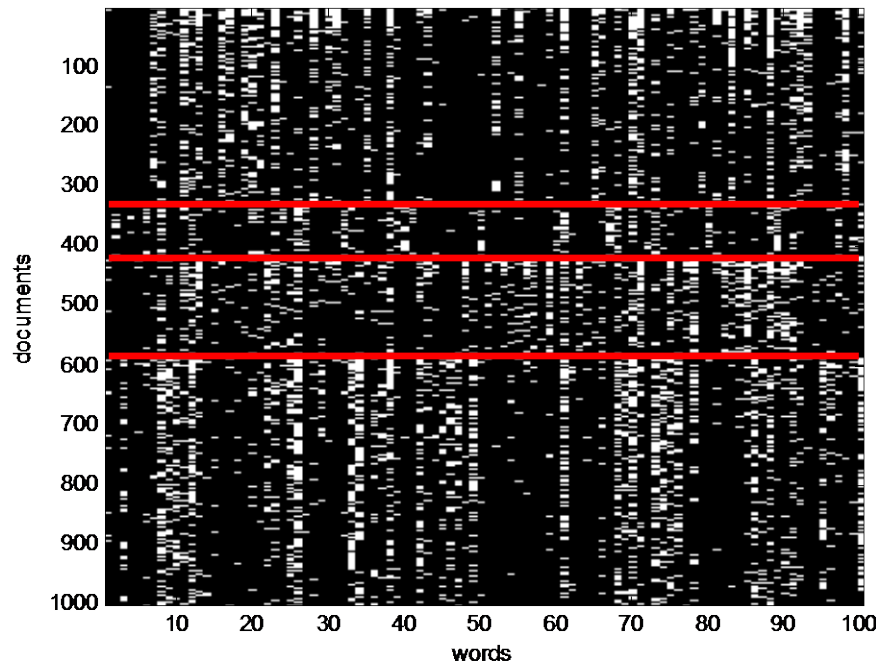
■ Clasificación con varios atributos



Ejemplos de Clasificación

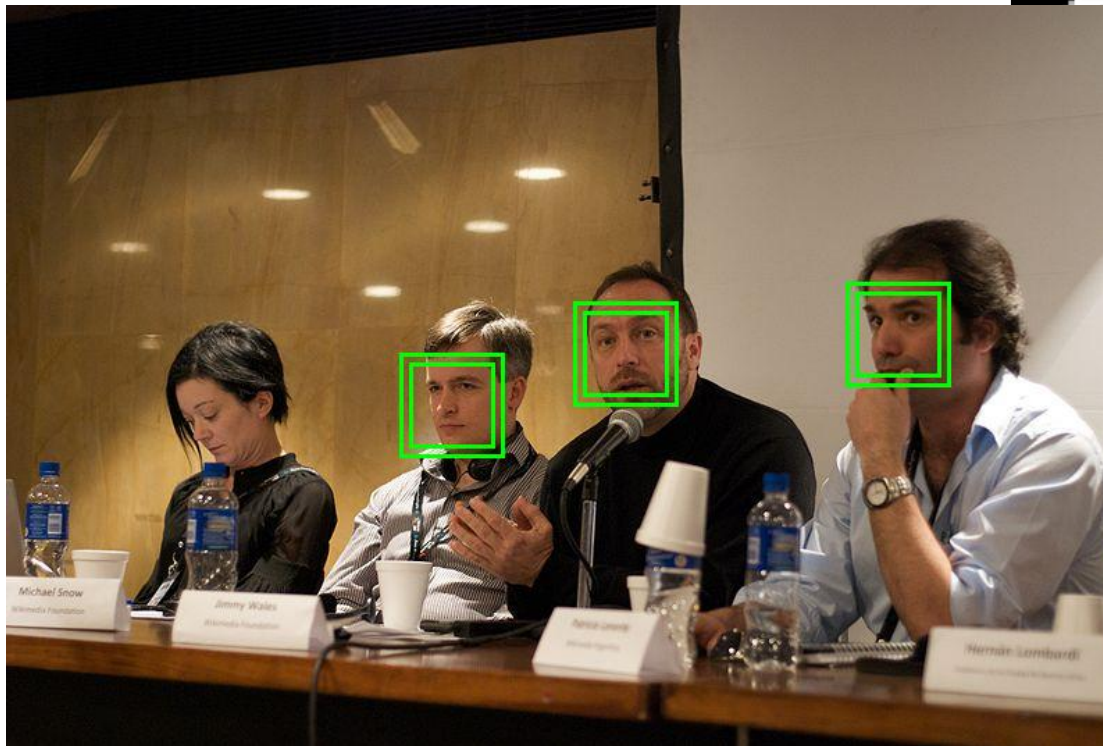
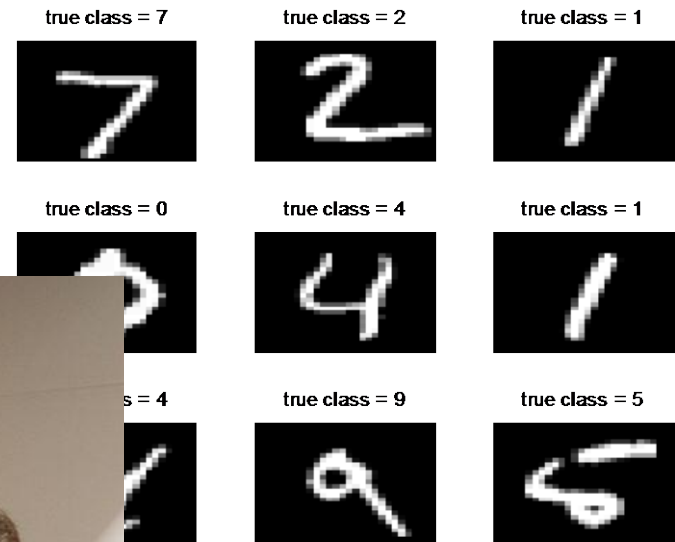
- Clasificación de documentos
 - Representación con bolsa de palabras
 - Caso particular: filtro de SPAM

USENET
newsgroups



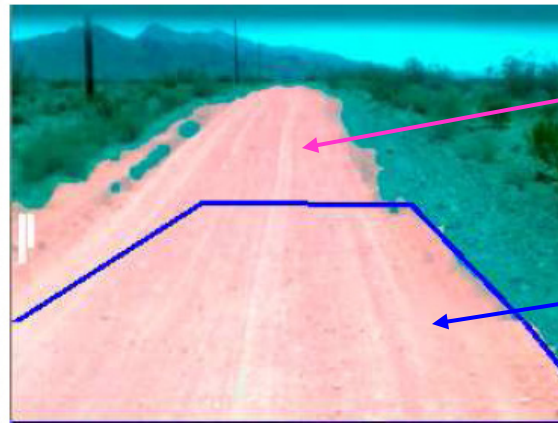
Ejemplos de Clasificación

- Reconocimiento de caracteres
- Detección de caras



Ejemplos de Clasificación

- Conducción Automática
 - Navegable / No navegable



Zona navegable predicha

Zona vista por el láser → datos de entrenamiento

Self-supervised Monocular Road Detection in Desert Terrain
H. Dahlkamp, A. Kaehler, D. Stavens, S. Thrun, G. Bradski

Aprendizaje Supervisado

- Requiere datos etiquetados
- Obtener las etiquetas es costoso
 - Muchas compañías optan por “*crowd sourcing*”
- Las etiquetas contienen poca información

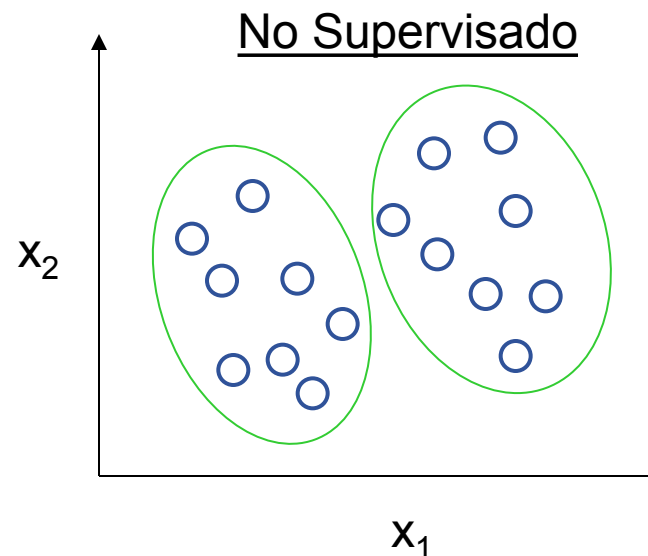
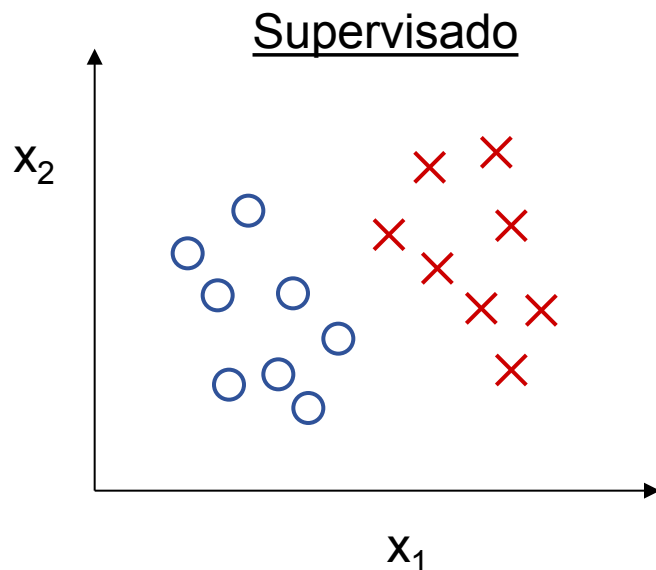
Cuando aprendemos a ver, nadie nos dice las respuestas, sólo miramos. A veces tu madre dice “eso es un perro” pero esa es muy poca información. Tendrás suerte si consigues unos pocos bits, incluso un bit por segundo, de esa manera. El sistema visual del cerebro tiene 10^{14} conexiones neuronales. Y solo vivimos 10^9 segundos. Así que un bit por segundo no sirve de nada. Necesitas más, unos 10^5 bits por segundo. Y solo hay un lugar de donde puede venir tanta información: de la entrada. *G. Hinton (1996)*

Aprendizaje No Supervisado

- También llamado “*knowledge discovery*”
- Datos de entrenamiento: sólo las entradas

$$\mathcal{D} = \{\mathbf{x}^{(i)}\}_{i=1}^N$$

- Objetivo: descubrir “estructura interesante” en los datos



Aprendizaje No Supervisado: Clustering

- Clustering: descubrir agrupaciones en los datos
 - ¿Cuántos clusters hay?
 - ¿A qué cluster pertenece cada dato?
- Ejemplos
 - Astronomía: el sistema *autoclass* (1988) descubrió un nuevo tipo de estrella agrupando medidas astrofísicas
 - Comercio electrónico: agrupar a los usuarios basado en su navegación y sus compras, para enviarles propaganda

Ejemplo de Clustering

Noticias destacadas



La Vanguardia

Cobertura
tiempo real

Ensayo nuclear Corea del Norte

La Vanguardia - hace 18 minutos

Corea del Norte ha anunciado oficialmente haber realizado una tercera prueba nuclear, con un dispositivo "miniatura", desafiando los llamamientos de China, del resto de potencias internacionales y de la ONU para que no llevara a cabo este ensayo y ...

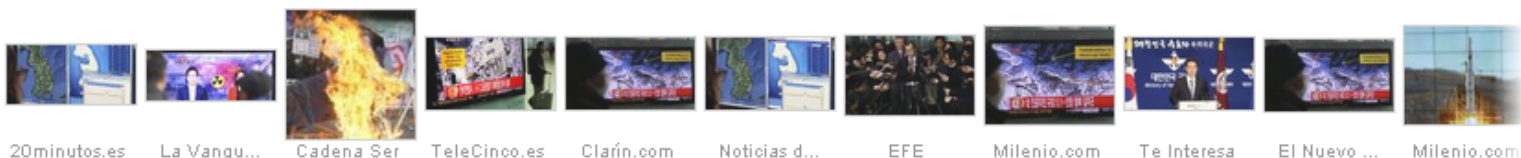
Corea del Norte realiza una tercera prueba nuclear que provoca un ... RTVE

Corea del Norte desafía al mundo con una prueba nuclear El Mundo.es

De Estados Unidos: Prueba norcoreana es una provocación, dice EEUU El Nuevo Herald

De fondo: Corea del Norte realizó una prueba nuclear y desafió la prohibición ... Lanacion.com (Argentina)

Relacionados
[Corea del Norte »](#)



20minutos.es

La Vangu...

Cadena Ser

TeleCinco.es

Clarín.com

Noticias d...

EFE

Milenio.com

Te Interesa

El Nuevo ...

Milenio.com



Periodista Digital

Cobertura
tiempo real

Benedicto XVI, un pontífice decepcionado de su rol

El Universal (Venezuela) - hace 43 minutos

Fue sorpresivo el anuncio pero no la idea. En 2010, entrevistado por Peter Seewald para el libro Luz del mundo, Benedicto XVI habló de la posibilidad de renuncia en caso de que fuese necesario y desde hace tiempo se habla en Roma del deterioro de la ...

Benedicto XVI renuncia: "No tengo ya fuerzas para seguir como Papa" ecodiario

Dimisión del Papa Benedicto XVI Latercera

Opinión: El legado que deja Benedicto XVI SIGLO21.com.gt

Relacionados
[Benedicto XVI »](#)
[Papa »](#)



YouTube

Reporte Indigo

CanalINTN24

La Vangu...

La Vangu...

levante.e...

La Razón

La Razón

La Razón

La Razón



Europa Press

El Congreso debate la iniciativa popular de la dación en pago ...

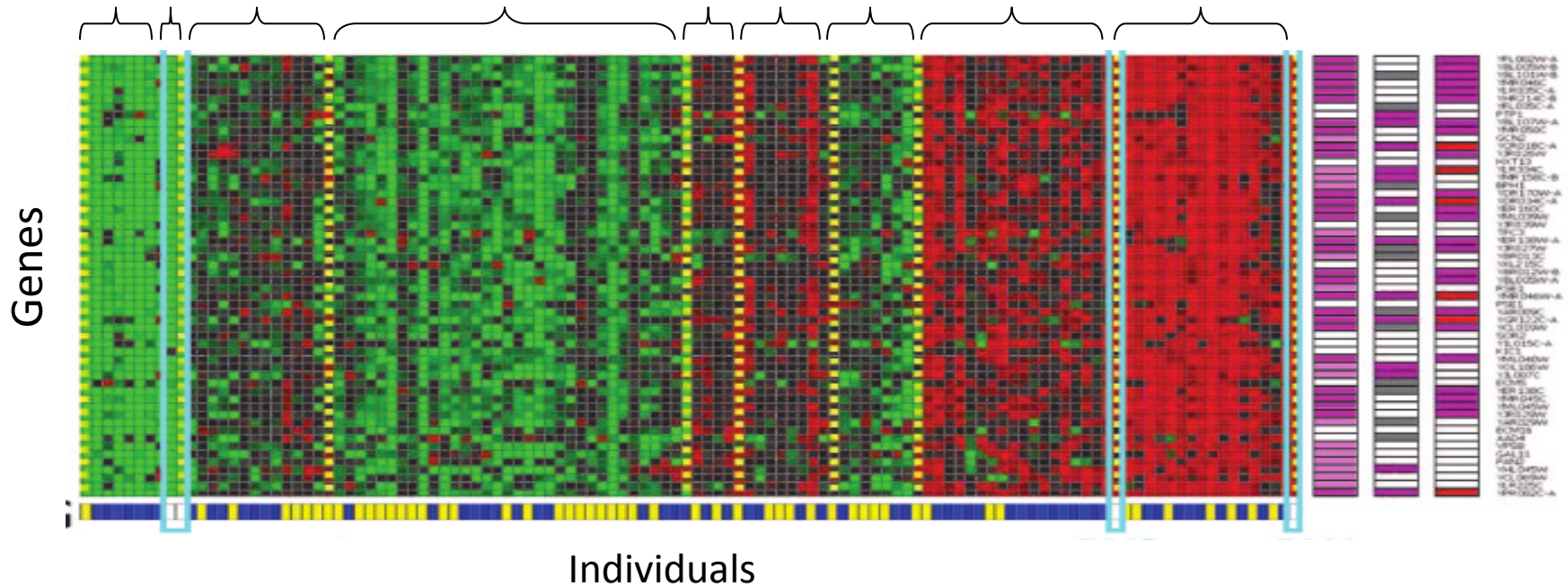
Europa Press - hace 46 minutos

Los promotores de la iniciativa legislativa popular (ILP) a favor de la dación en pago retroactiva y de la paralización de las ejecuciones hipotecarias se reunirán este martes con el PP y con CiU pocas horas antes de la votación en el Pleno del Congreso de la ...



Universidad
Zaragoza

Ejemplos de Clustering

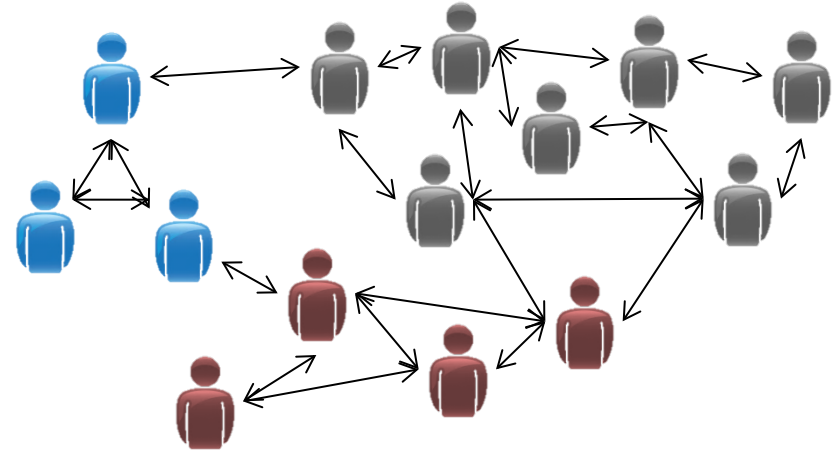


[Source: Su-In Lee, Dana Pe'er, Aimee Dudley, George Church, Daphne Koller]

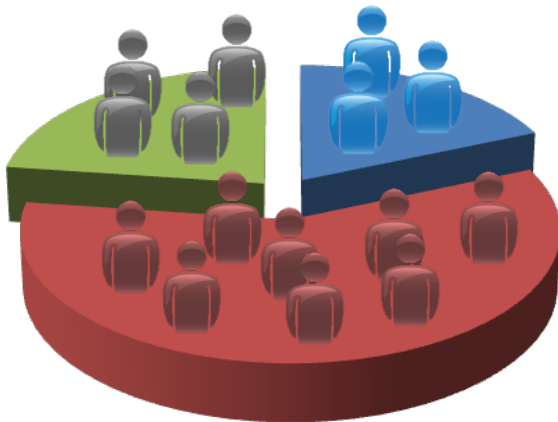
Ejemplos de Clustering



Organize computing clusters



Social network analysis



Market segmentation



Astronomical data analysis

Reducción de la Dimensionalidad

- Proyectar los datos en un espacio de menor dimensión que capture su “esencia”

- PCA: Análisis de componentes principales



mean



principal basis 1



principal basis 2



principal basis 3



reconstructed with 2 bases



reconstructed with 10 bases



reconstructed with 100 bases

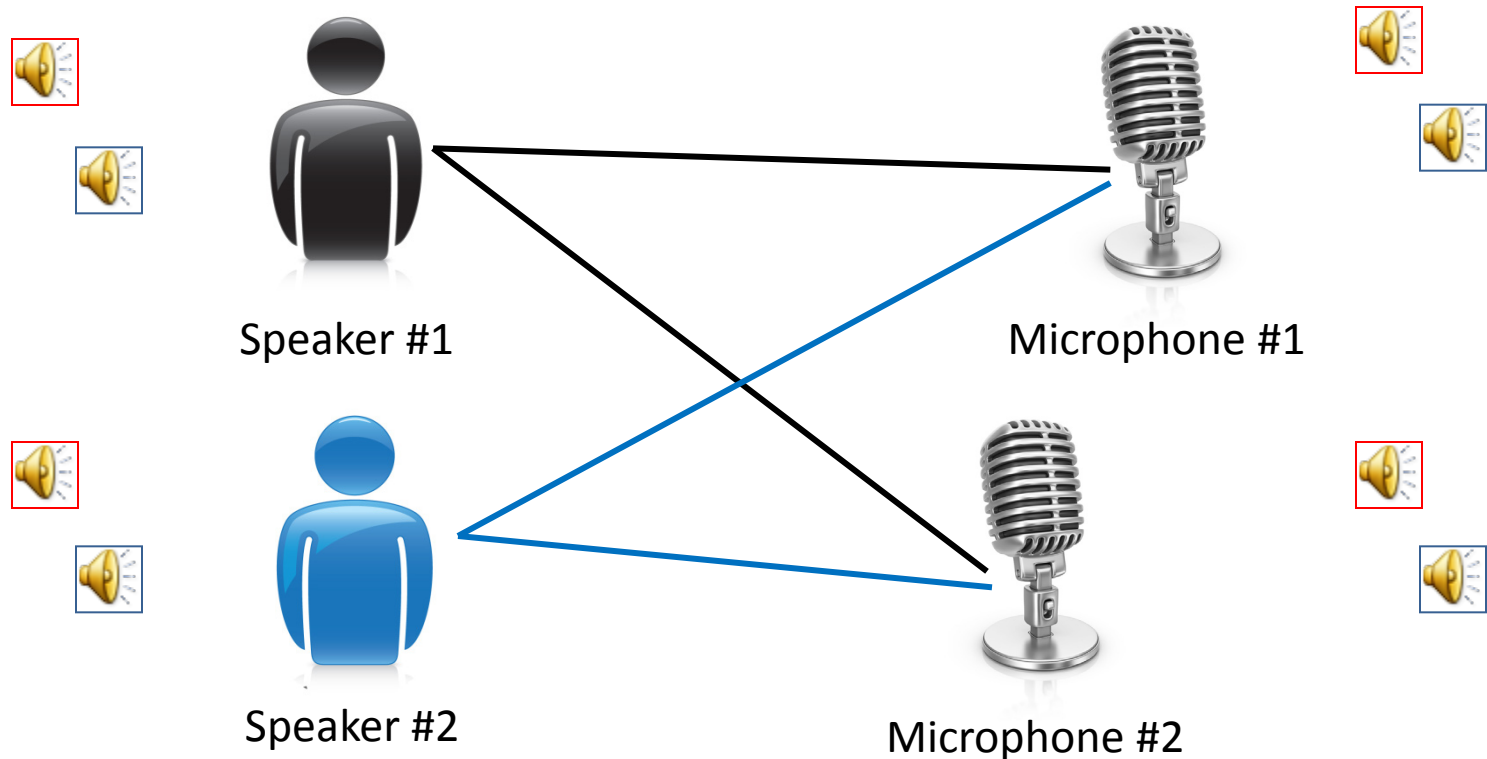


reconstructed with 400 bases



Ejemplo: Separación de Fuentes

■ Cocktail Party Problem:



Examples by Te-Won Lee: http://cnl.salk.edu/~tewon/Blind/blind_audio.html

Ejemplo: Separación de Fuentes

- Kurtosis Maximization ICA (variante de PCA)

- Código en Matlab:

```
yy = sqrtm(inv(cov(xx')))*(xx-repmat(mean(xx,2),1,size(xx,2)));  
[W,ss,vv] = svd((repmat(sum(yy.*yy,1),size(yy,1),1).*yy)*yy');
```

- xx son las señales mezcladas de los micrófonos (una por columna)
- W es la matriz de des-mezclado

[Fuente: Sam Roweis, Yair Weiss & Eero Simoncelli, 2001]

Programa de la Asignatura

Aprendizaje Supervisado

- Regresión y Clasificación

Aprendizaje No Supervisado

- Clustering, EM, PCA,...

Aplicaciones

- Reconocimiento de patrones
- Filtrado de SPAM
- Sistemas de recomendación
- Detección de anomalías
- ...

Prácticas de Laboratorio

■ Objetivos

- Comprender las ventajas y dificultades de las técnicas de aprendizaje estudiadas
- Aplicarlas a casos sencillos

■ 10 prácticas x 3 horas

- Prácticas individuales
- Programación en Matlab sobre Windows
- L0.01, Lunes 9:00-12:00
- L0.04, Viernes 8:00-11:00 y 11:00-14:00
- **Ver calendario en Moodle**

Evaluación de la Asignatura

- Nota final:
 - Prácticas 40%
 - Examen 60%
 - Notas mínimas para promediar: examen **4.0**; prácticas: **4.0**
 - Se guardan las notas de cada parte si son \geq **5.0**
- Las prácticas son **individuales**
 - Las irregularidades en la realización de las prácticas (por ejemplo, código copiado) supondrán el **suspenso** en la correspondiente convocatoria

Evaluación de las Prácticas

■ Evaluación Continua (recomendada)

- Acudir a la sesión con estudio previo por escrito
- Corrección al final de cada sesión
- ¿Qué pasa si no la termino a tiempo?
 - ✦ Disponéis de un total de 12 días de retraso, que podéis utilizar como queráis a lo largo de todo el cuatrimestre.
 - ✦ Cuenta el día de entrega en moodle.unizar.es
 - ✦ Para casos **excepcionales**, hablar con el profesor.

■ Evaluación Global

- Memoria de todas las prácticas por escrito
- Defensa en el laboratorio ante el profesor
- Durante el periodo de evaluación de cada convocatoria

Bibliografía

- Transparencias disponibles en moodle.unizar.es
- Bibliografía complementaria
 - R.O. Duda, P.E. Hart, D.G.Stork, Pattern Classification, 2nd ed., John Willey & Sons, 2001
 - C.M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006
 - Kevin P. Murphy, Machine Learning: A Probabilistic Perspective, MIT Press, 2012
 - T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd ed., Springer, 2009
<http://www-stat.stanford.edu/~tibs/ElemStatLearn/download.html>
- Curso Online
 - Andrew Ng, Machine Learning, Stanford AI Lab.
<https://www.coursera.org/course/ml>