

Resumen técnicas de recomendación

SCAR

**Sistemas Complejos Adaptativos y
Recomendación**



Official Master's Degree in Artificial Intelligence,
Pattern Recognition and Digital Imaging

MIARFID

SR Demográfico

0	80	30	0	0	60	100	0		50
0	1	2	3	4	5	6	7	...	18

Proceso previo

Este proceso se realiza al crear el recomendador y es general para todos los usuarios del sistema

Definir la clasificación de usuarios con sólo sus datos demográficos

Para cada tipo de usuario, se define un **vector** de 19 enteros, que indica, mediante un valor entre 0 y 100, el interés del usuario por el género (preferencias demográficas), 5 o 6 posiciones no 0

Los datos de los que se dispone para un usuario son: Edad, Sexo y Ocupación

Ejemplo

- Los hombres mayores de 70 años les gusta con un 90 los documentales, las películas del oeste con 100, las de acción con un 50
- Los niños y niñas menores de 10 años, les gustan las películas infantiles con un 100, la animación con un 90, las comedias con un 50 y la ciencia ficción con un 60

Proceso previo

Ejemplo

Tipo 1: niños <ul style="list-style-type: none">• Edad ≤ 10 años	Tipo 2: adolescentes <ul style="list-style-type: none">• Edad entre 11 y 18	Tipo 3: interesados en la acción <ul style="list-style-type: none">• Hombres• Edad entre 18 y 40• Profesión...	Tipo 4: interesados en ciencia ficción <ul style="list-style-type: none">• Edad mayor que 18• Profesión: informático, ...	Tipo 5: <ul style="list-style-type: none">• Edad: mayor de 30• Profesión abogados	Tipo 6: <ul style="list-style-type: none">• Edad: mayor de 35• Profesión amo de casa	...
---	--	---	---	---	--	-----

Proceso previo

Usuario tipo 1

0	80	30	0	0	60	100	0		50
0	1	2	3	4	5	6	7	...	18

Usuario tipo 2

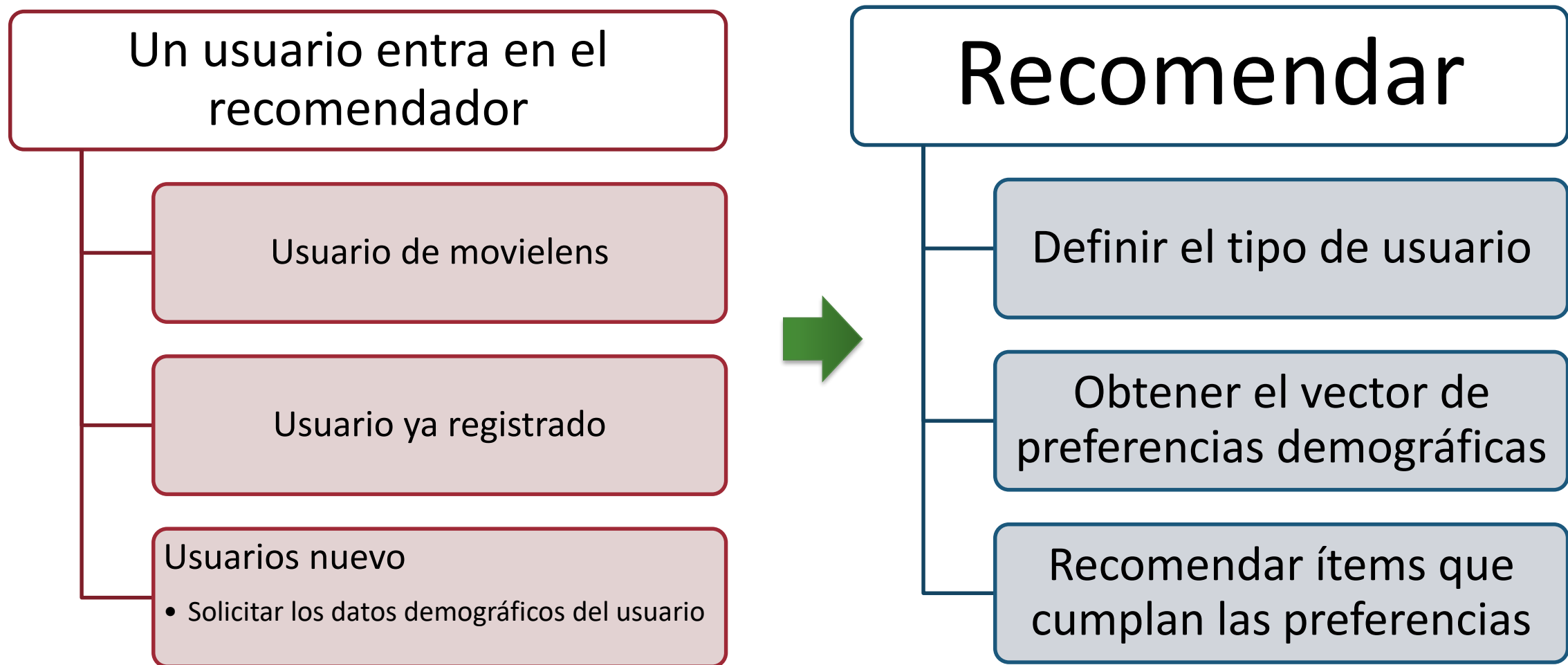
0	0	10	90	0	4	25	0		0
0	1	2	3	4	5	6	7	...	18

...

Usuario tipo N

0	100	0	0	10	90	0	50		0
0	1	2	3	4	5	6	7	...	18

Recomendación



Recomendación



Información demográfica

- Información personal: edad, género, familia, país, religión,...
- Cualquier característica que se pueda obtener del usuario



Tipo X



```
int preferencias_demograficas[NUM_GENEROS];
```

Modelo de preferencias del usuario

- Gustos o preferencias
- Tipo de ítems en los que el usuario está interesado

Histórico de interacción del usuario con el sistema

- Conjunto de ítems recomendados, visitados, comprados, ... anteriormente por el usuario
- Grado de satisfacción con la recomendación, visita, compra, ...

Información interna del SR

- Información que el SR calcula y que facilita el proceso de recomendación
- Clasificación del usuario, usuarios similares, ...

Recomendación

Dadas las preferencias demográficas

Obtener los ítems que satisfacen las preferencias (todos los ítems que están clasificados en alguna/todas las preferencias)

Eliminar los ítems ya vistos por el usuario (los que están en su histórico, en el perfil de usuario)

Calcular para cada ítem, el ratio de interés del usuario en el ítem (puede que el ítem esté clasificado en varias preferencias, con varios ratios)

Recomendación

El ratio de interés del usuario en el ítem se calcula en función de

Ratio de las preferencias del usuario que coinciden con las que está clasificado el ítem

Número de preferencias por las que está clasificado el ítem

Esta función la definís vosotros, no hay una fórmula

Ajustar la fórmula según lo que vosotros consideréis que se debería obtener

Recomendación

Clasificación de la película p1

0	1	1		0
0	1	2	...	18

Preferencias del u1

0	20	55		90
0	1	2	...	18

Obtener el ratio de interés del u1 en p1

$$r_{u1_{p1}} = f(f2(p1[0], u1[0]), \dots, f2(p1[18], u1[18]))$$

Recomendación

Ordenar los ítems
por ratio calculado

- De mayor a menor

Seleccionar los ítems
que se mostrarán en
la interfaz y que
información se
mostrará

SR Basado en Contenido

Proceso previo

Sólo para los usuarios de Movielens

Para todos los usuarios de Movielens

Obtener las preferencias basadas en contenido

Almacenarlas en el perfil

```
int preferencias_basadas_contenido[NUM_GENEROS];
```

Proceso previo

Preferencias basadas en contenido

Se utilizan las películas que ha puntuado un usuario favorablemente y como están clasificadas

Para cada película del histórico puntuada favorablemente

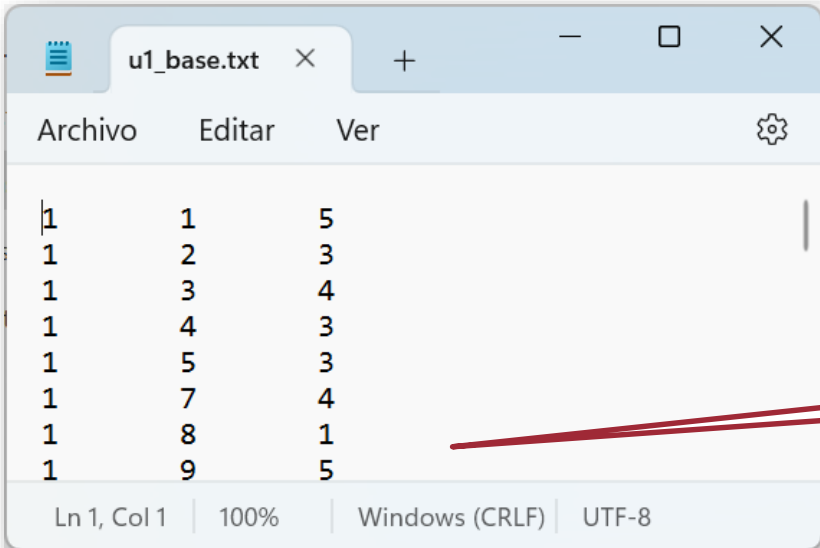
- Obtener los géneros en los que está clasificada la película
- Actualizar el valor de la preferencia del usuario dependiendo de si el género se ha repetido en más de una película puntuada favorablemente

Proceso previo

Seleccionar las películas del usuario puntuadas con 3 o más

Seleccionar los géneros en los que está clasificada cada película

Con cada película, modificar el vector de preferencias basadas en contenido



1	1	5
1	2	3
1	3	4
1	4	3
1	5	3
1	7	4
1	8	1
1	9	5

items: Bloc de notas

ArchivoEdiciónFormatoVerAyuda

1	0	0	0	1	1	1	0	0	0	0	0	0
2	0	1	1	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0							
4	0	1	0	0	0							
5	0	0	0	0	0							
6	0	0	0	0	0							
7	0	0	0	0	0	0	0			0	0	0

Con cada película, modificar el vector de preferencias en contenido



0	80	30	0	0	60	100	0		50
0	1	2	3	4	5	6	7	...	18

```
int preferencias_basadas_contenido[NUM_GENEROS];
```

Proceso previo

Intentar que casi todas las posiciones del vector de preferencias basadas en contenido tengan valor, aunque sea pequeño

Copiar este vector en el vector de preferencias colaborativas (lo usaremos más adelante)

Dejar en el vector de preferencias basadas en contenido sólo las 5-6 posiciones de mayor valor

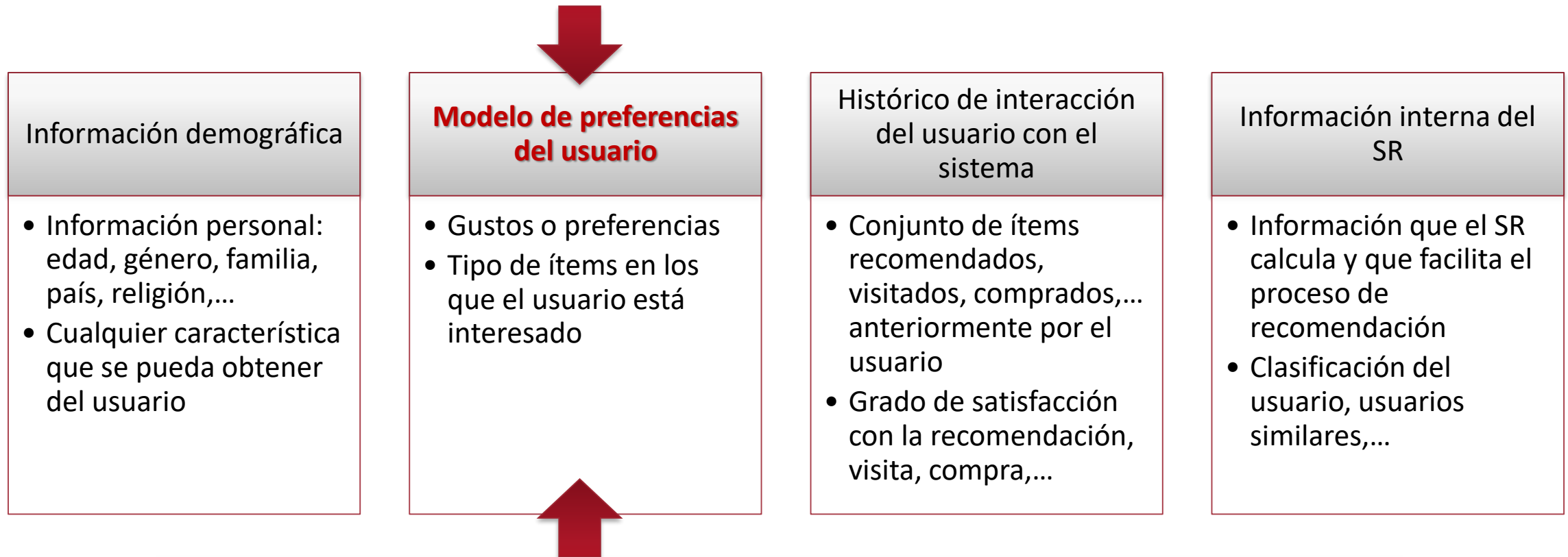
Preferencias colaborativas

	90	10	90	7	43		14		7
0	1	2	3	4	5	6	7	...	18

Preferencias basadas en contenido

	90		90		43	25	14		
0	1	2	3	4	5	6	7	...	18

Proceso previo



```
int preferencias_basadas_contenido[NUM_GENEROS];
```

Recomendación

Un usuario entra en el recomendador

- Usuario de movielens
- Usuario ya registrado
- Usuario nuevo
 - Pedirle las preferencias (en cuantos más géneros, mejor)
 - Guardarlos en las preferencias colaborativas
 - Seleccionar las 5-6 más altas y guardarlas en las preferencias basadas en contenido



Recomendar

Obtener el vector de preferencias BC

Recomendar ítems que cumplan las preferencias

SR Colaborativo

Proceso previo

Sólo para los usuarios de Movielens

- Para los usuarios de Movielens, la búsqueda de vecinos será un proceso previo
- Para el resto de usuarios, se realizará en el momento del registro (cuando haya rellenado sus preferencias)

Obtener los vecinos de cada uno de los usuarios de Movielens

Almacenar la lista de usuarios vecinos en el perfil de cada usuario

Para este proceso, es necesario tener las preferencias colaborativas del usuario

- Proceso realizado en la recomendación basada en contenido

Proceso previo

Crear una matriz donde se incluyan las preferencias colaborativas de los usuarios del sistema

Cada fila es un usuario
Cada columna es una preferencia

Para buscar los vecinos afines se usa el Coeficiente de Correlación de Pearson entre las filas de la matriz

Seleccionar los vecinos con mayor afinidad (valores más cercanos a 1 o más altos)

Repetir el proceso con todos los usuarios del sistema

Dejar rellenos los vecinos de todos los usuarios de Movielens

Proceso previo

Crear una matriz

- Filas: usuarios
- Columnas: preferencias colaborativas

Los valores que no tienen datos, rellenarlos a 0

	0	1	2	...	18
Usuario 1	0	10	40		70
Usuario 2	0	0	60		0
...					
Usuario n	0	20	30		90

Proceso previo

	0	1	2	...	18
Usuario 1	0	10	40		70
Usuario 2	0	0	60		0
...					
Usuario n	0	20	30		90

$$Var(X) = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

$$Var(Y) = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n}$$

$$Cov(X, Y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n}$$

$$\rho_{XY} = \frac{Cov(X, Y)}{\sqrt{Var(X) \cdot Var(Y)}}$$

Proceso previo

Rellenar los vecinos en el perfil de usuario

Número de vecinos del usuario

Para cada vecino

Variable, depende del usuario

Identificador

Grado de afinidad con el vecino (valor entre 0 y 100)

```
// Vecino
typedef struct
{
    int id_vecino, afinidad;
} vecino;
```



```
// Recomendación colaborativa
int num_vecinos;
vecino vecinos[MAX_VECINOS];
```


Proceso previo

Lista de vecinos

$$L_u = \{(u', ra_{u'}) / u' \in U, ra_{u'} \in [0..100]\}$$

- $ra_{u'}$ es el ratio de afinidad entre u y u' . Puede ser un valor entre 0..100 o en el rango que se desee

Recomendación

Un usuario entra en el recomendador

- Usuario de movielens
- Usuario ya registrado
- Usuario nuevo
 - Pedirle las preferencias (en cuantos más géneros, mejor)
 - Guardarlos en las preferencias colaborativas
 - Calcular los vecinos del usuarios



Recomendar

Recomendar ítems de los vecinos

Recomendación

Recopilar todos los ítems puntuados favorablemente por los vecinos (eliminar repetidos)

- Eliminar los ítems puntuados por u



Calcular el ratio de interés r_i de cada ítem para u

- Considerar de que usuario proviene el ítem (ratio de afinidad con el vecino)
- Considerar el ratio que el vecino dio al ítem
- Considerar si el ítem está puntuado por más de un vecino



Obtener una lista ordenada de ítems a recomendar

- Ordenada por ratio de interés del ítem para el usuario u : r_i

Recomendación

Calcular el ratio de interés r_i de cada ítem para u

Es la clave del proceso de recomendación

Depende de varios factores y es importante ajustarla para cada sistema en concreto

No hay una fórmula a aplicar, supone entender como se puntúan los datos en cada sistema en concreto y como afectan elementos como: afinidades de usuarios, repeticiones de ítems,...