

Trabajo 6

SR colaborativo

SCAR

**Sistemas Complejos Adaptativos y
Recomendación**



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

MIARFID

Objetivos

Añadir al SR la técnica de recomendación colaborativa mediante vecinos

Para la búsqueda de vecinos se utilizará el coeficiente de correlación de Pearson

Objetivos

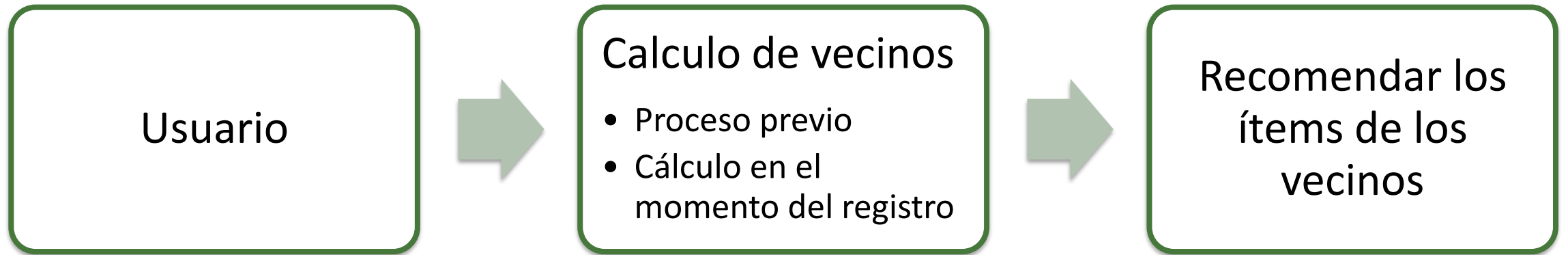
El funcionamiento de esta técnica de recomendación será algo diferente a las dos anteriores

En las anteriores rellenábamos un vector de preferencias (demográficas o basadas en contenido) y obteníamos los ítems que satisfacían las preferencias

En este caso buscaremos vecinos y usaremos directamente los ítems de los vecinos

Tendremos un vector de preferencias colaborativas que se usan para calcular los vecinos

Objetivos



Obtener vecinos

El proceso será diferente si el usuario está registrado en el sistema o acaba de registrarse

Un usuario registrado ya debe tener los vecinos calculados

- Pre-proceso que calcula todos los vecinos de movielens
- Se han calculado sus vecinos cuando el usuario se ha registrado

Un usuario que acaba de registrarse no tiene vecinos calculados

- Cuando el usuario se registra, se deben calcular los vecinos

Obtener vecinos

Usuarios de movielens

- Proceso previo
 - Obtener las preferencias colaborativas de todos los usuarios
 - Obtener los vecinos de todos los usuarios
- Proceso normal
 - Un usuario solicita la recomendación (ya tiene calculados los vecinos)
 - Se le muestra la recomendación obtenida

Usuario registrado

- Cuando el usuario se registra
 - Se obtienen las preferencias colaborativas
 - Se obtienen sus vecinos
- Proceso normal
 - Un usuario solicita la recomendación (ya tiene calculados los vecinos)
 - Se le muestra la recomendación obtenida

Usuario nuevo

- Registrar al usuario
 - Se obtienen las preferencias colaborativas
 - Se obtienen sus vecinos
- Proceso normal
 - Un usuario solicita la recomendación (ya tiene calculados los vecinos)
 - Se le muestra la recomendación obtenida

Obtener las preferencias colaborativas

Usuario del dataset de movielens

En el recomendador basado en contenido ya obtuvimos las preferencias colaborativas

Obtener las preferencias basadas en contenido usando las películas que ha puntuado favorablemente el usuario

- Se obtienen aprendiendo de los gustos de los usuarios
- Se utilizan las películas que ha puntuado un usuario y como están clasificadas
- Cuantas más posiciones del vector de preferencias tengan datos, mejor (aunque sean valores bajos)

Realizar un proceso previo a la recomendación que calcule las preferencias basadas en contenido de todos los usuarios del dataset

- Mayor rapidez en la recomendación
- Es necesario hacerlo para la recomendación colaborativa (como se verá en la siguiente parte del trabajo)

Preferencias

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

Obtener las preferencias colaborativas

Usuario del dataset de movielens

Para cada película del histórico
puntuada favorablemente

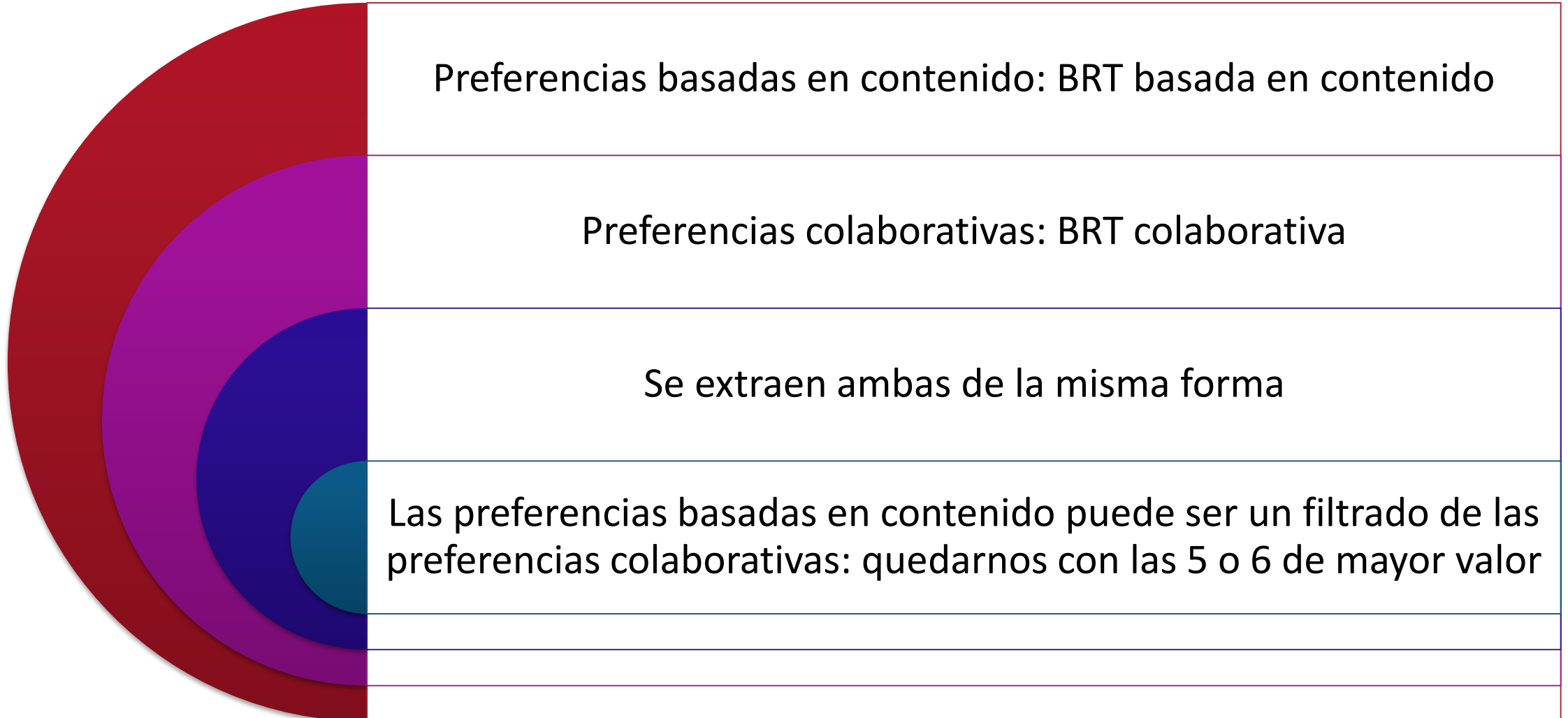
Optativo

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 favorablemente

Se puede realizar el proceso
contrario con las películas
puntuadas desfavorablemente,
disminuyendo el valor de los
ratios

Preferencias basadas en contenido vs preferencias colaborativas



Preferencias basadas en contenido vs preferencias colaborativas

Proceso de obtener los géneros en los que está clasificada la película

Obtendremos todos los géneros que son de interés para el usuario: preferencias colaborativas

De las preferencias colaborativas nos quedaremos con las de mayor ratio

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

Preferencias basadas en contenido vs preferencias colaborativas

La razón de tener las preferencias colaborativas como las preferencias basadas en contenido pero con todos los valores posibles está en la BRT colaborativa

Es mejor que hayan cuantas menos posiciones vacías,
mejor

Preferencias colaborativas

Si usásemos Pearson con las preferencias con muchas posiciones a 0

Si se usan las 19 preferencias

- Habría muchos ratios similares (los 0)
- Alta correlación falsa

Si se usasen sólo las preferencias distintas de 0 de ambos usuarios

- Habría muy pocas preferencias coincidentes
- Pearson no funciona bien con muestras pequeñas

Obtener vecinos

Los vecinos se obtienen utilizando la correlación de Pearson entre las preferencias colaborativas de los usuarios del sistema

Los vecinos serán aquellos cuyo coeficiente de correlación entre sus preferencias y la del usuario, sea mayor

Obtener vecinos

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 de obtención de vecinos



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

2. Cada fila es un usuario
Cada columna es una preferencia

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

4. Seleccionar los vecinos con mayor afinidad (valores más cercanos a 1)

Proceso de obtención de vecinos

Si el proceso se hace
para un usuario

Si el proceso se hace
para todos los
usuarios

Calcular Pearson
entre sus
preferencias y
las de todos los
demás usuarios

Rellenar sus
vecinos

Repetir el
proceso con
todos los
usuarios del
sistema

Dejar rellenos
los vecinos de
todos los
usuarios

Proceso de obtención de vecinos

Si el usuario para el que se solicita la recomendación es u

Para cada uno de los usuarios $u' \in U$

Calcular el ratio de afinidad entre u y u' usando el coeficiente de correlación de Pearson entre las preferencias de u y las de u'

Seleccionar como vecinos los usuarios de mayor coeficiente de correlación

Para elegir los vecinos:

- Se pueden elegir los N de mayor coeficiente
- O bien, los vecinos cuyo coeficiente supere un umbral
- Ambas cosas

Pasos

1

- Obtener los ítems recomendados, usando los ítems puntuados favorablemente por los vecinos del usuario

2

- Mostrar los ítems recomendados en la interfaz

Paso 1. Obtener los ítems recomendados

Una vez se han calculado los vecinos, obtener los ítems de los vecinos para recomendarlos al usuario

Para cada vecino

- Obtener los ítems del histórico del vecino puntuados favorablemente
 - Eliminar ítems repetidos. Combinar los ratios de los ítems repetidos (por ejemplo, usar la media de los ratios)
 - Calcular el ratio de interés del ítem para el usuario
- Eliminar de la lista los ítems ya vistos por el usuario (histórico)

Paso 1. Obtener los ítems recomendados

El ratio de interés del ítem para u puede depender de

Si el ítem está
repetido en más
de un vecino

Ratio de
afinidad con el
vecino

Puntuación del
vecino al ítem

Paso 1. Obtener los ítems recomendados

Una vez se obtiene la lista de ítems de los vecinos

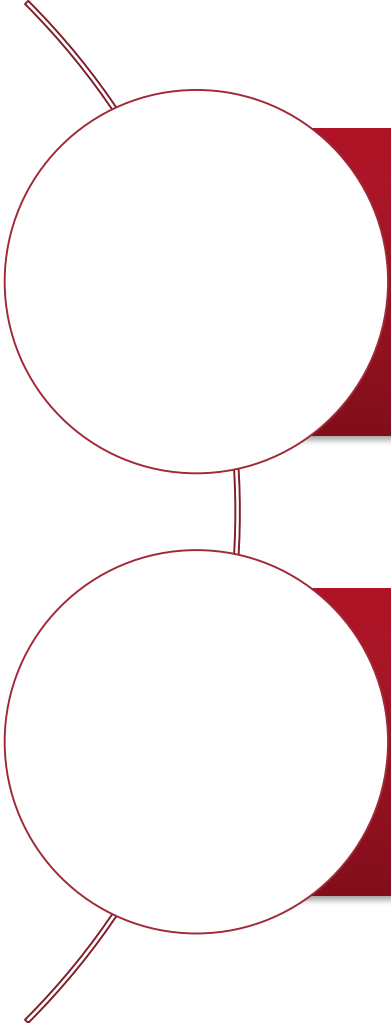
Se ordena por ratio

Se muestran al usuario los N ítems de mayor ratio

Paso 2. Mostrar los ítems recomendados

Proceso similar al del
resto de técnicas de
recomendación

Optativo



Obtener la similitud de los usuarios, usando los ítems puntuados favorablemente (puntuación mayor que X)

Usar sólo los ítems comunes entre los dos usuarios que se está comparando si pueden ser vecinos

Resumen

Búsqueda de vecinos
con preferencias

Usar todas las
preferencias,
rellenando a 0

Búsqueda de vecinos
con ítems

Usar sólo los
ítems que
ambos vecinos
han puntuado