

Tema 2

Sistemas Recomendadores



SCAR
**Sistema Complejos Adaptativos y
Recomendación**

Sistema Recomendador (SR)

Tipo específico
de filtro de
información
adaptativo

Trata de
presentar al
usuario,
únicamente,
información
sobre la que
cree que va a
estar interesado

Se adelanta a
las necesidades
del usuario

Recomendar
supone hacer
una predicción
sobre los
intereses del
usuario

Definición

Sistema Recomendador (Resnick, 1997)

Tipo específico de filtro de información **adaptativo**

Trata de presentar al usuario, únicamente, información sobre la que cree que va a estar interesado

Década 1990: primeros SR

Tapestry

- Goldberg, 1992
- Sistema de correo electrónico experimental que utiliza filtrado colaborativo

GroupLens

- Resnick, 1994
- Noticias
- Primer SR que utiliza ratios para los datos

MovieLens

- Konstan, 1997
- Recomendador de películas
- El dataset se sigue utilizando

Firefly

- Maes, 1997
- SR de música, permitía a los usuarios calificar y etiquetar música
- Estas calificaciones y etiquetas se utilizaban para construir perfiles de usuario y generar recomendaciones de música similar que podría gustarles a esos usuarios.

Definición formal

Sea

U Conjunto de usuarios

I Conjunto de posibles ítems a recomendar

R Conjunto totalmente ordenado de enteros o reales (ratios)

u Usuario sobre el que se realiza la recomendación

i Ítem para el cual se quiere generar la predicción de la preferencia de u

fu, max Función de utilidad que mide la utilidad de i para u

$$fu : U \times I \longrightarrow R$$

$$\forall u \in U, \exists i' \in I \mid I'_u = \max_{i \in I}(u, i)$$

Definición formal

Obtiene un ratio para **todos** los ítems

Para obtener el ratio se utiliza una función de utilidad **personalizada** al usuario

Selecciona el subconjunto de ítems con mayor ratio (los ítems que **maximizan** la utilidad)

$$f_u : U \times I \longrightarrow R$$

$$\forall u \in U, \exists i' \in I \mid I'_u = \max_{i \in I} (u, i)$$

Personalización

Eje central de un SR

Personalización

Basado en las preferencias e intereses del usuario

Clasifica los ítems a recomendar según la utilidad o interés para el usuario

Intentan

Simular la interacción con un humano

Influir en el comportamiento del consumidor

Facilitar la interacción del usuario con aplicaciones y servicios

Aplicaciones de los SR

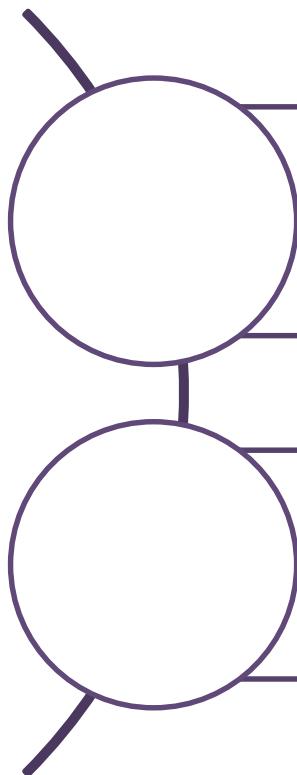
Razones para incluir SR en una aplicación o servicio

Publicidad
personalizada:
más efectiva

Fidelizar a los
clientes

Compartir
información
entre
aplicaciones

Información obtenida



Muestra una **cantidad** de información que el usuario pueda manejar, que se adapte a sus necesidades y gustos, que sea de su interés

Algunos SR permiten definir el **número de ítems** a recomendar (1, una lista de n elementos, todos)



gatos divertidos



Imágenes

Vídeos

Memes

Nombres

Para niños

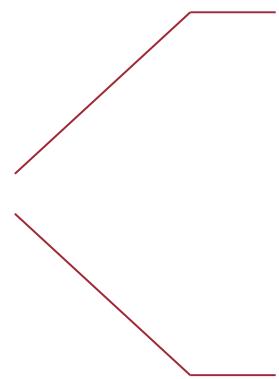
Juegos

Fotos

Juguetes

Usuarios

Pueden ir
dirigidos
a...



Un individuo

- Tienen en cuenta las preferencias e intereses de un único usuario

Un grupo

- Tienen en cuenta las preferencias e intereses del grupo en conjunto



Recomendadores para grupos

Gustos y preferencias del grupo en su conjunto

Restricciones (niños, discapacitados,...)

Mezcla de preferencias de los usuarios

Deben satisfacer por igual a todos los miembros del grupo



Ítems

Los elementos a recomendar se denominan **ítems**

La recomendación permite seleccionar, entre todos los ítems disponibles, aquellos que son de mayor interés para el usuario

Un ítem puede ser cualquier cosa

- Un producto a comprar, una película que ver, un destino turístico, una dirección web,...

Ítems

Un SR calcula, para todos los ítems, su **ratio de utilidad o interés** para el usuario

Selecciona los ítems de mayor ratio

Cuando hay gran cantidad de ítems, se puede hacer un filtrado previo por tipo



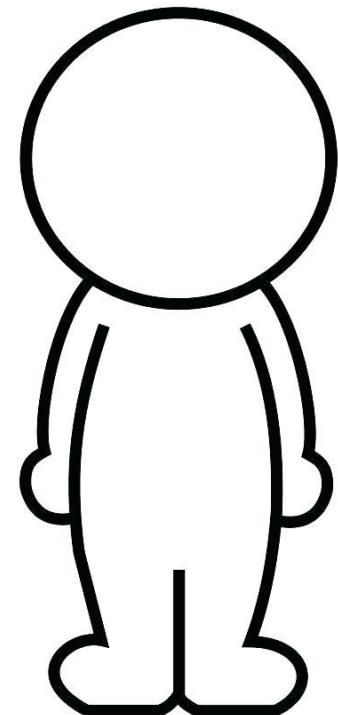
Ratio de interés



Ratio de interés



Ratio de interés

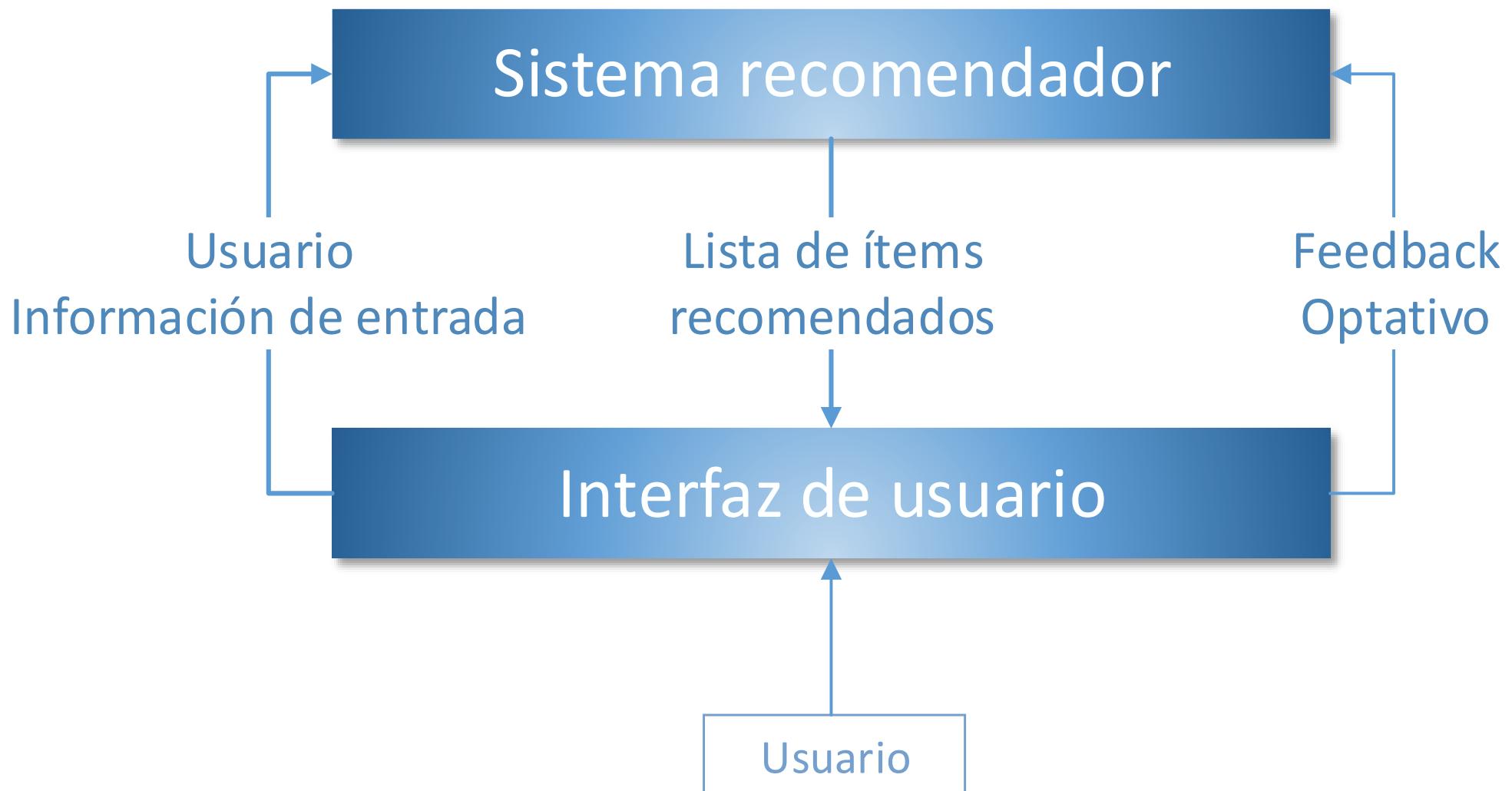


Proceso de recomendación

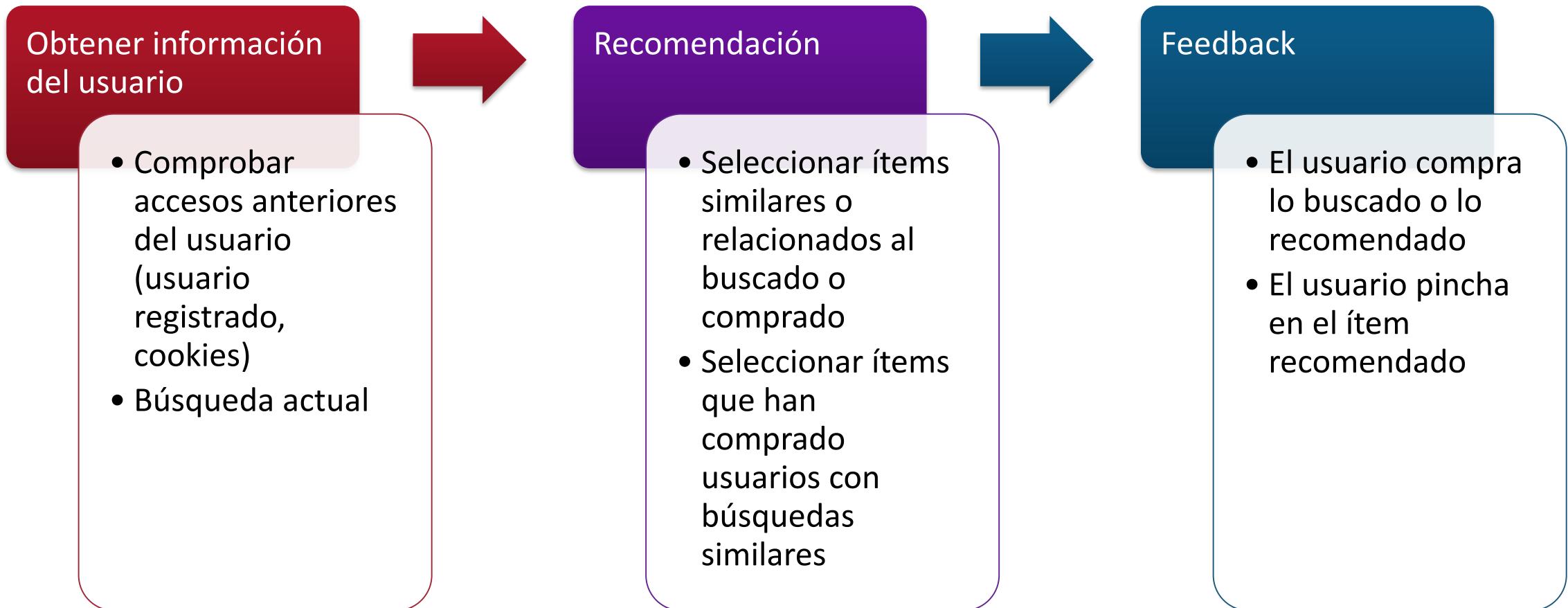
Proceso de recomendación



Proceso de recomendación



Ejemplo: Amazon



Feedback



Se puede obtener información sobre el proceso de recomendación

- Si el usuario compra/visita/pincha/le gustan elementos que se le han recomendado
- Se le puede pedir al usuario explícitamente que valore la recomendación ofrecida

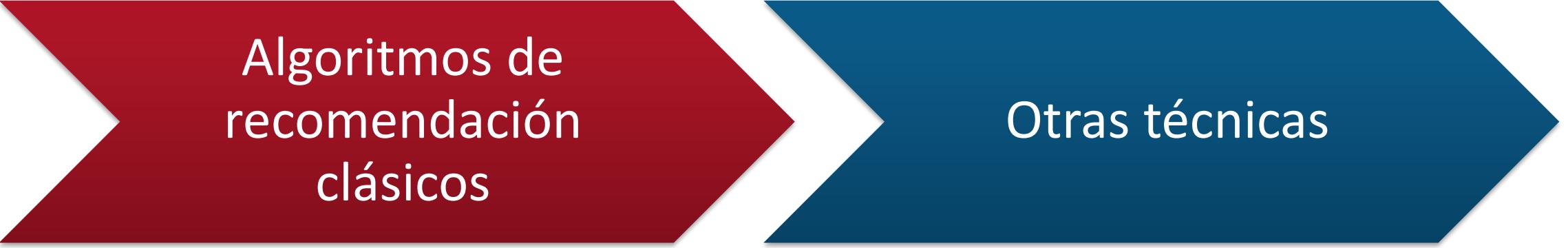
Con este feedback, se actualiza la información sobre los gustos y preferencias del usuario

Este proceso se suele denominar “**puntuación de los ítems**” por parte del usuario

Técnicas de recomendación básicas

BRTs

Técnicas de recomendación básicas (BRTs)

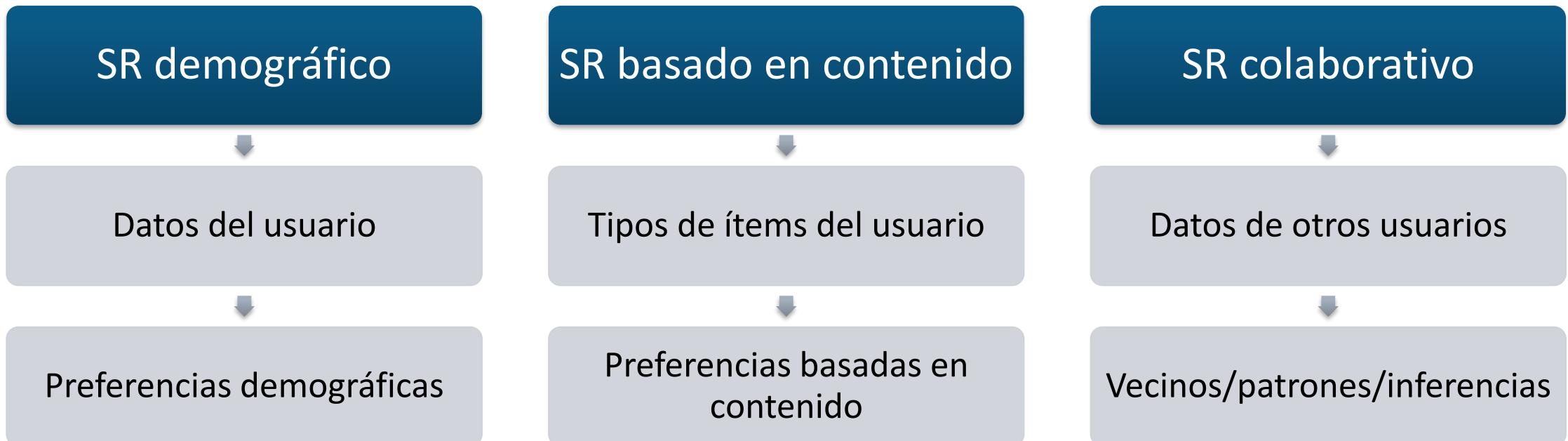


Algoritmos de
recomendación
clásicos

Otras técnicas

- Demográfica
(demographic)
- Basada en contenido
(content-based)
- Colaborativa
(collaborative)

BRTs



SR Demográfico

Clasifica al usuario en una categoría demográfica (tipo) basándose en sus características personales (parte del perfil de usuario)

A cada grupo demográfico se le asigna una lista de preferencias o de ítems

Se recomienda según el grupo demográfico al que pertenece

Se debe definir una clasificación de usuarios común a todo el sistema

SR Basado en Contenido

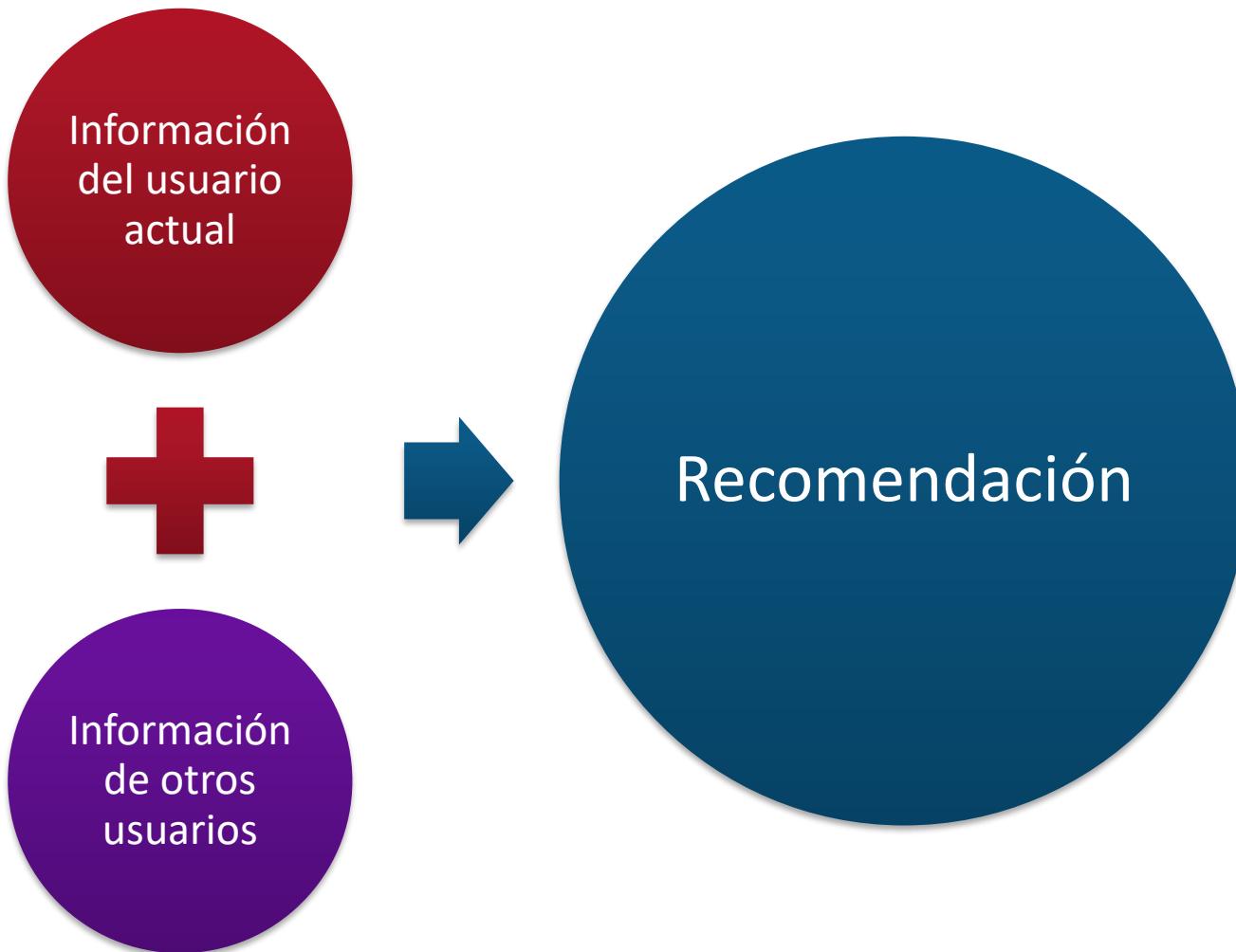
Se definen los ítems por sus características o preferencias asociadas

Se recomienda al usuario ítems con **características similares** a los que él ha puntuado favorablemente en el pasado, o bien, los que están clasificados según sus preferencias

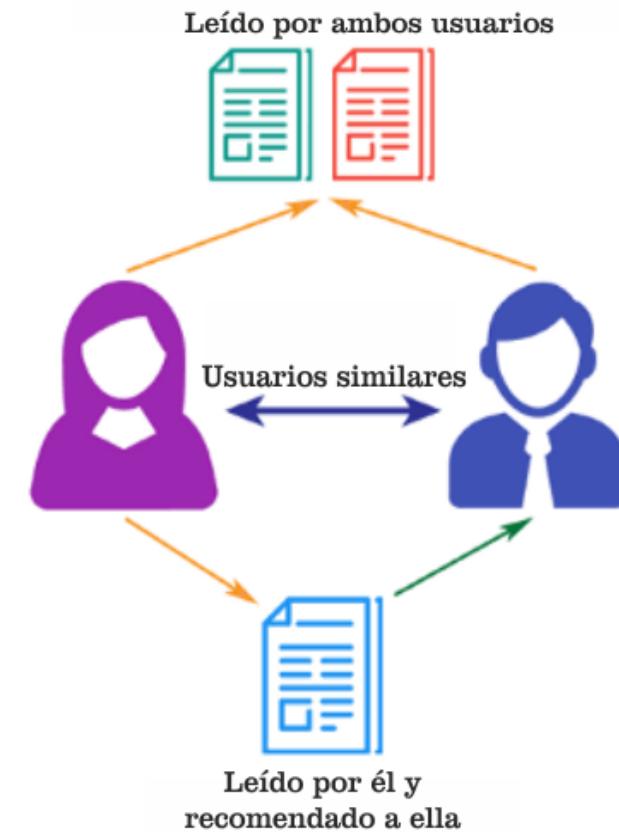
Se tienen en cuenta los ítems previamente puntuados por el usuario (histórico) para conocer el tipo de ítems en los que está interesado

Apropiado cuando los ítems se generan dinámicamente y se debe decidir si son de interés para el usuario o no: email, noticias,...

SR Colaborativo



Filtro colaborativo



Otras BRT

SR basados en conocimiento

SR basados en utilidad

SR basados en casos

SR basado en restricciones

SR basados en crítica

SR sociales

SR multi-criterio

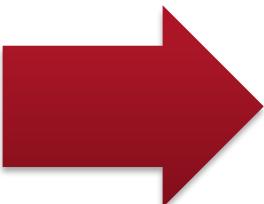
SR context-aware

SR conversacional

Técnicas de recomendación

Técnicas básicas

- Demográfica
- Basada en contenido
- Colaborativa
- ...



Técnicas híbridas

- Mezcla de las técnicas básicas

Información para el proceso de recomendación

Información utilizada en la recomendación

Información disponible sobre el **usuario**

Información sobre **otros usuarios** (ítems que interesan a usuarios similares al usuario)

Información sobre el **entorno y características** del usuario (datos demográficos)

Información sobre los ítems a recomendar:
características de los ítems

Toda la
información
que
podamos
recopilar

¿Qué se necesita para que la recomendación funcione bien?

Ítems

Muchos ítems

Bien clasificados

- Clasificación (taxonomía u ontología rica)

Ítems variados, capaces de satisfacer a cualquier usuario

Usuarios

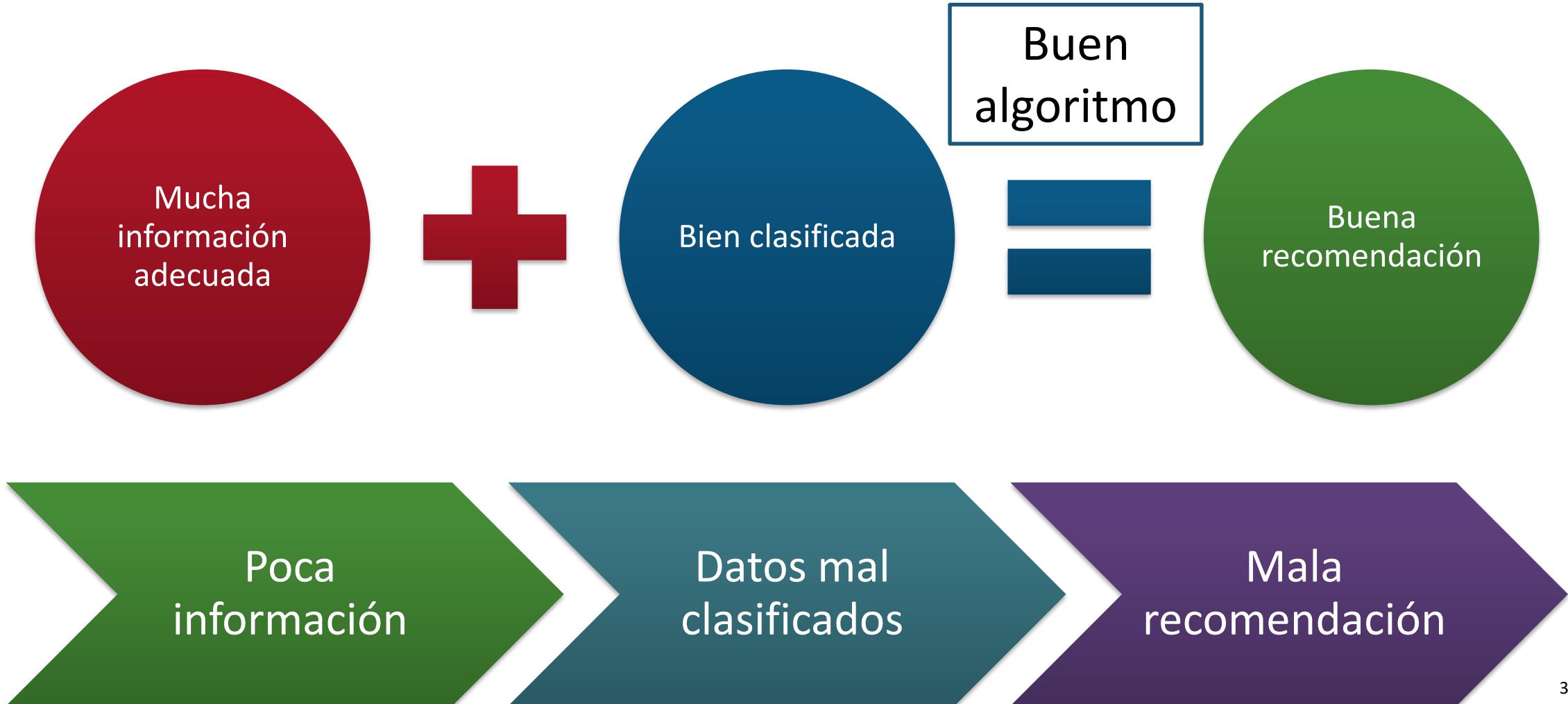
Muchos usuarios

Mucha información de cada usuario

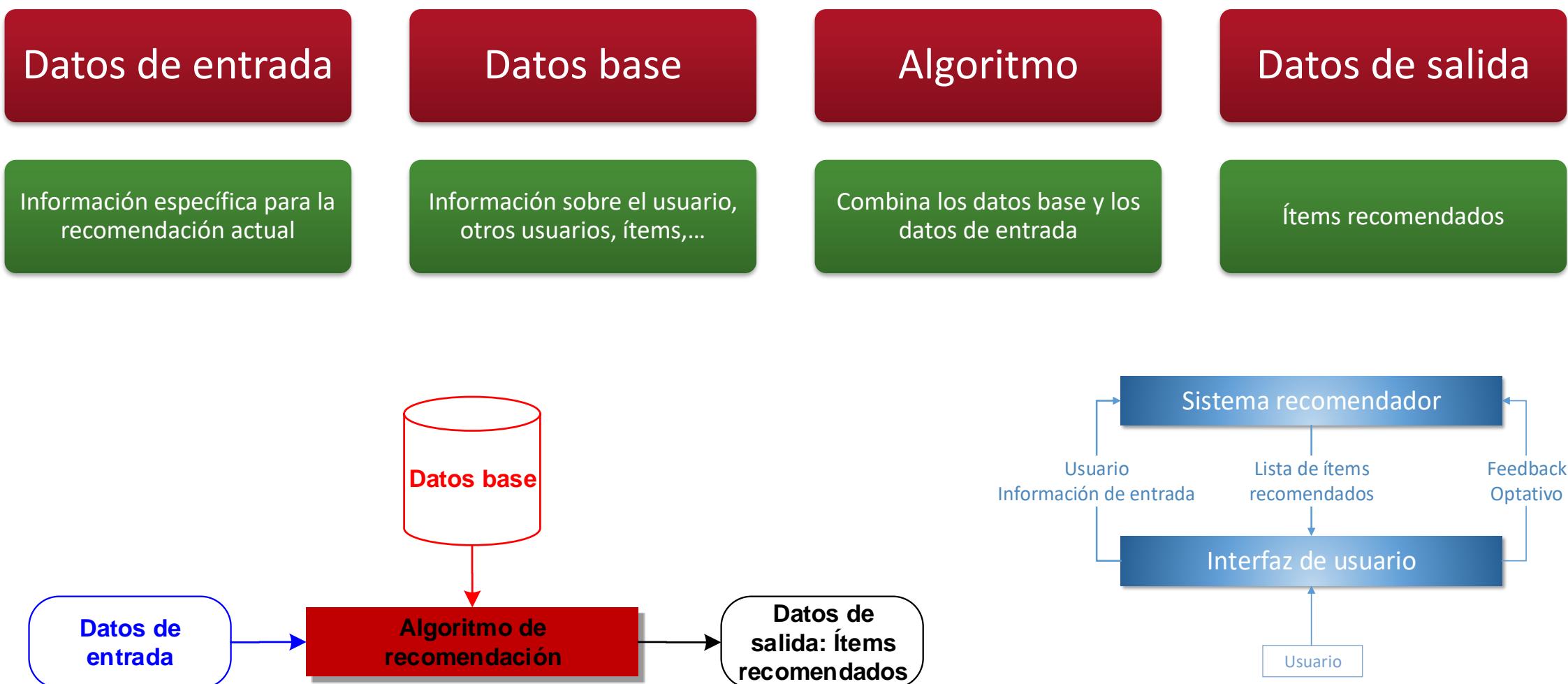
Mucha interacción de los usuarios con el servicio

Usuarios capaces de ser satisfechos

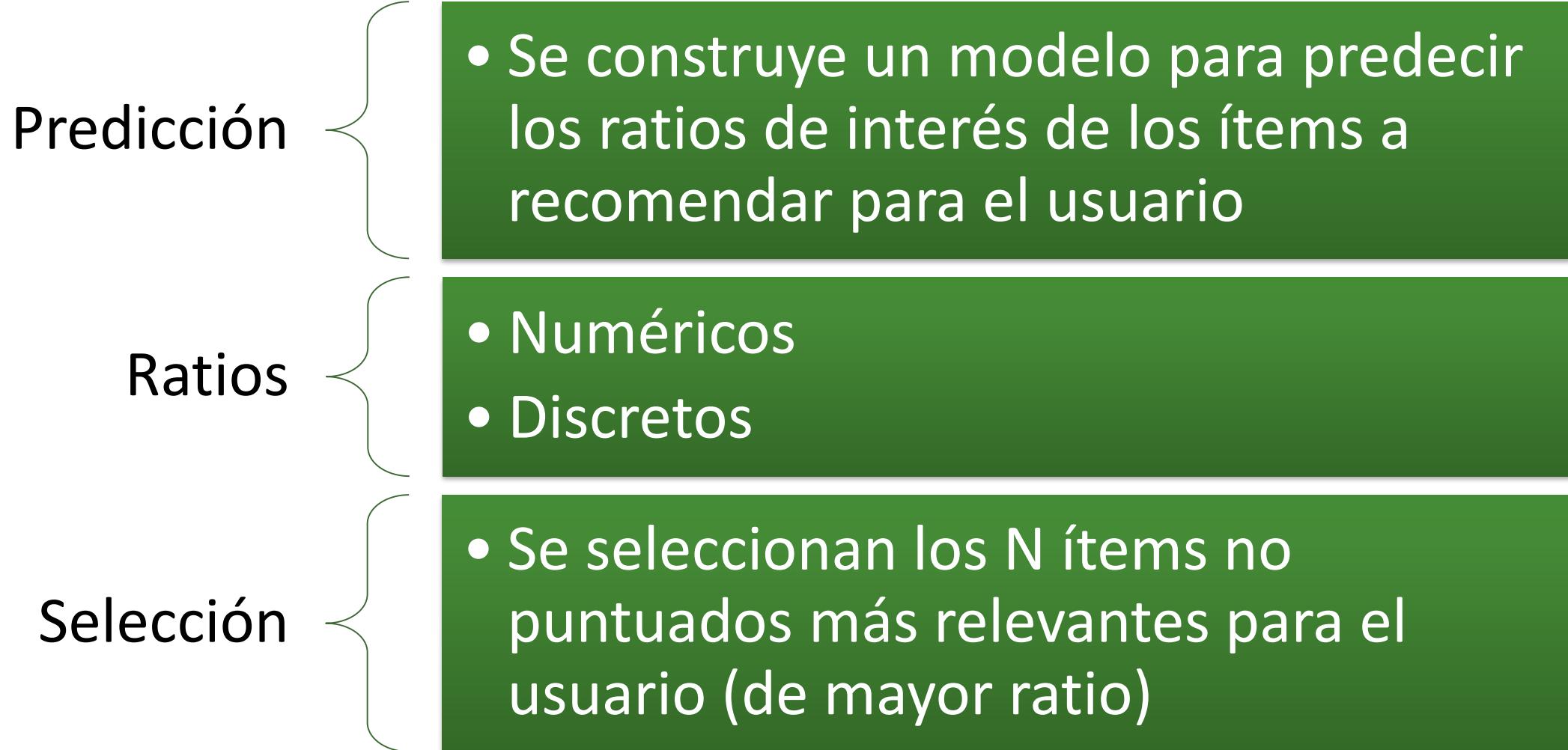
¿Qué se necesita para que la recomendación funcione bien?



Obtención de la recomendación



Algoritmo de recomendación



Algoritmo de recomendación

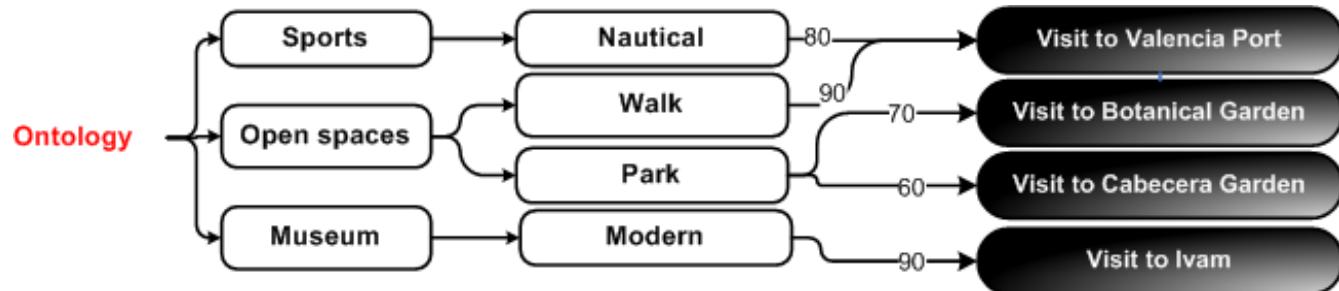
Los SR pueden trabajar de dos formas

Obtener un conjunto de preferencias (clasificación de los datos en la ontología, etiquetas,...) y de éstas, un conjunto de ítems

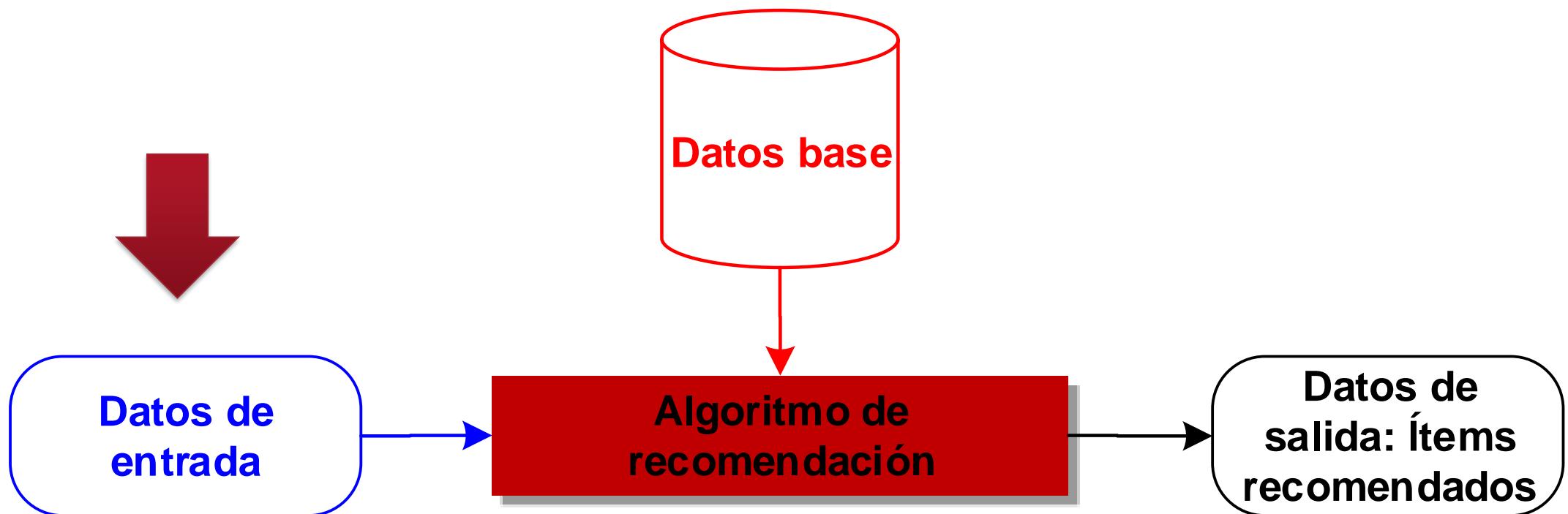
- Paso 1: preferencias
- Paso 2: ítems

Obtener directamente un conjunto de ítems

- Paso 1: ítems



Datos de entrada



Datos de entrada

El usuario proporciona datos sobre lo que busca en ese momento

- Por ejemplo, si busca un libro en Amazon, los datos de entrada puede ser autor, título

La información de entrada guía y restringe la recomendación (aunque se le pueden hacer sugerencias de elementos diferentes)

Algunos SR aprovechan este momento para solicitar al usuario alguna clase de información adicional

- Si está buscando un viaje se le solicitan fechas, por ejemplo

Visita actual

cruz, por favor, digita toda la información que te sea posible, para obtener una mejor recomendación.

Te gustaría hacer una visita en particular hoy ?

- Styles and periods
- Shopping
- Museums
- Open Spaces
- Religious architecture
- Defensive architecture
- Civil Architecture
- Gourmet
- Sports
- Monuments
- Leisure
- Health & SPA
- Events
- Children
- None

Fecha de la visita

Dia *

Hora de comienzo *

Hora de finalización *

Datos de entrada

Se puede aprovechar para que el usuario rellene otra información sobre él

Perfil de usuario

Update Profile - Part 1 of 2

cguz, Remember that to get a good recommendation is very important to enter all or most of the data.

Data User (Required)

Username *

CGUZ

Password *

Have Children ?

Age of eldest son

Age of youngest son

Personal data

Name

Cesar Augusto Guzman Alvarez

Your age

25

E-mail

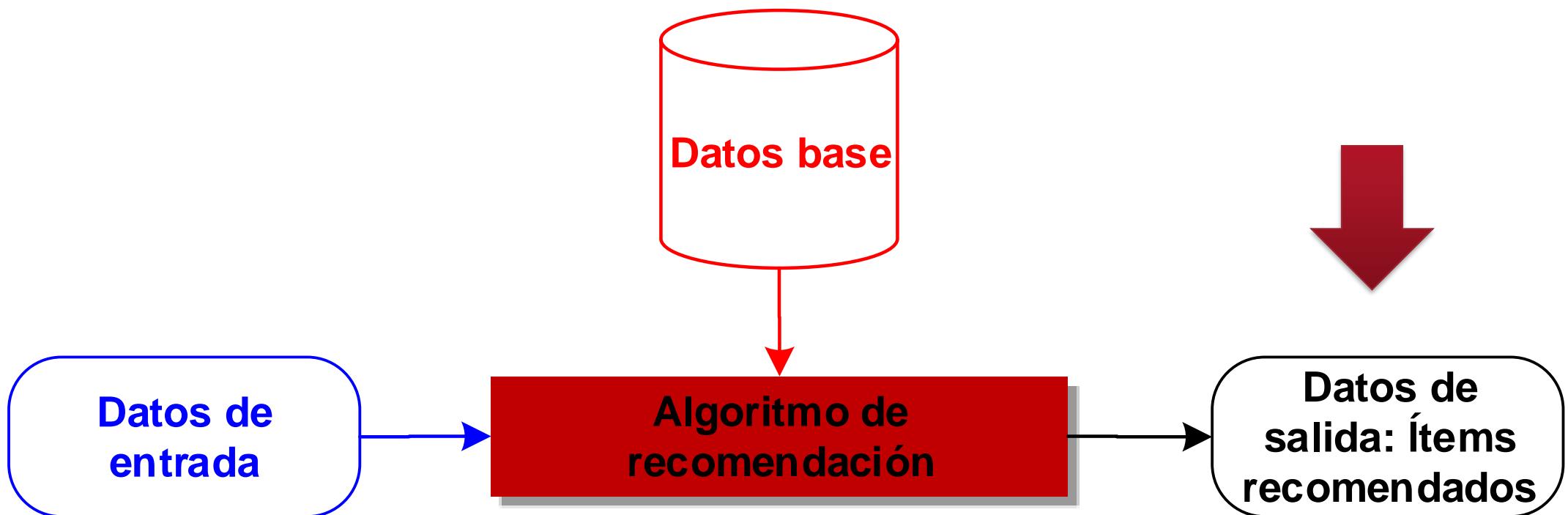
Have partner ?

Profession

Save

Next

Datos de salida



Información resultante de la recomendación

Resultado de la recomendación

- Un único ítem (el de mayor interés o ratio para el usuario)
- Lista

El conjunto de ítems se muestra al usuario de la manera en la que se haya definido en la interfaz

Una vez el usuario ha seleccionado un ítem, el SR puede ofrecer ítems similares al seleccionado:
“también te podría gustar...”

Datos de salida

Número de ítems

Un único ítem

Lista ordenada por ratios de todos los ítems

Lista de los N elementos de mayor ratio de la lista

Lista de los ítems con un ratio mayor que un valor determinado

Datos de salida

La recomendación calcula para cada ítem,
un ratio

- Estimación de la adecuación del ítem al usuario
- En la lista de recomendaciones, puede mostrarse solamente el ítem, o el ítem con su ratio

Cuando se muestra sólo el ítem y no se muestra el ratio, no se sabe si el ítem es adecuado para el usuario o sólo el menos malo

- Lo único que importa es el orden de los ítems en la lista, pero no el ratio con el que se ha calculado el orden
- Por ejemplo, se recomienda “**El Quijote**” en primer lugar y “**Sin noticias de Gurb**” en segundo, pero podría ser que, en una escala del 1 al 100, el ratio de “El Quijote” fuese **100** y el de “Sin noticias de Gurb” **1**

Datos de salida

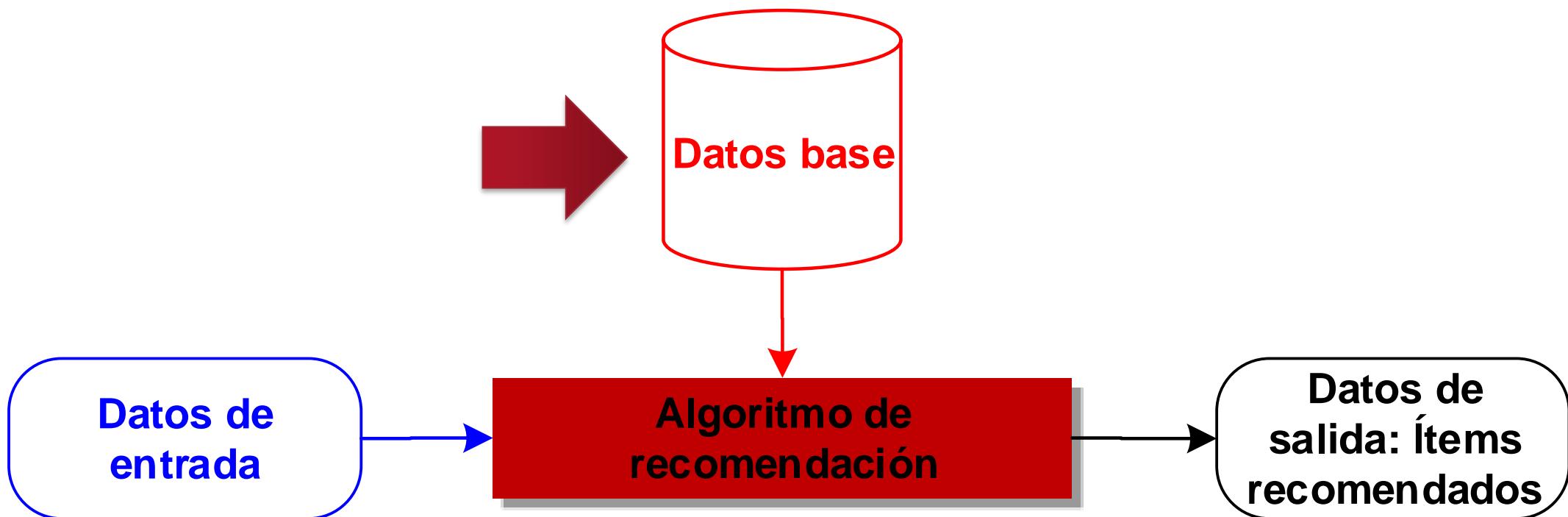
Puede interesar ofrecer ítems variados



- No ofrecer los N ítems de mayor ratio

- Ofrecer n1 ítems de mayor ratio de la categoría c1, n2 ítems de mayor ratios de la categoría c2,... hasta completar N ítems

Datos base



Datos base

Conjunto de datos que se usan para obtener la recomendación

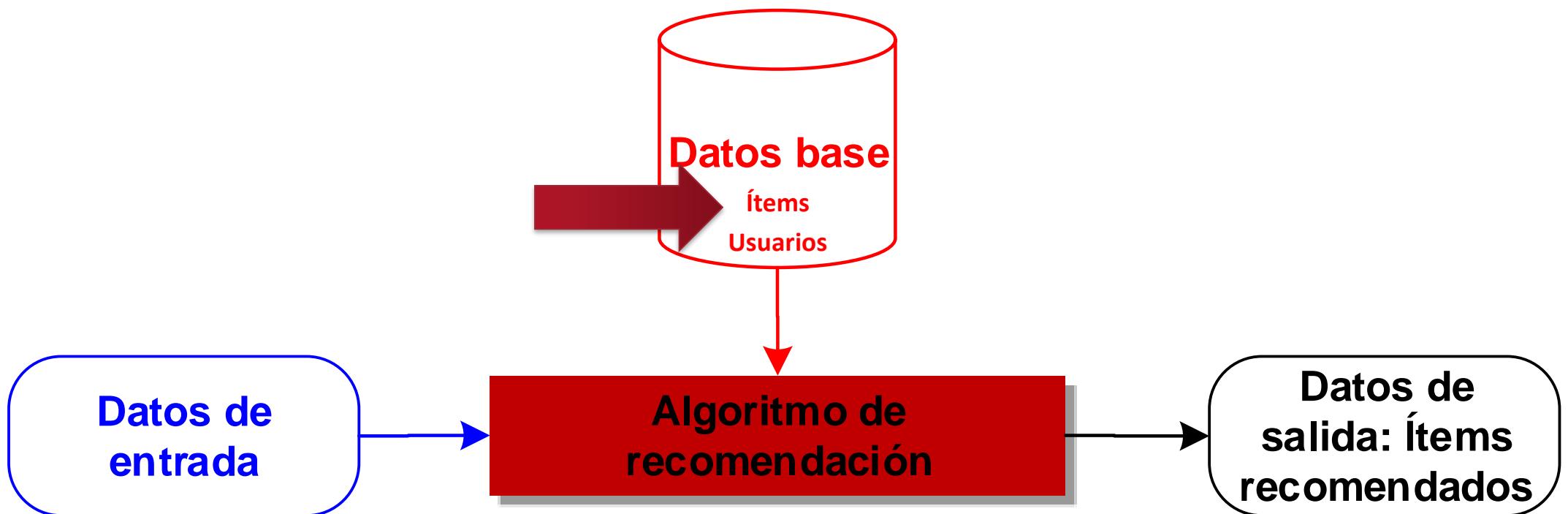
Contiene dos tipos de información



Información sobre los ítems

Información sobre los usuarios

Datos base: ítems



Información sobre los ítems

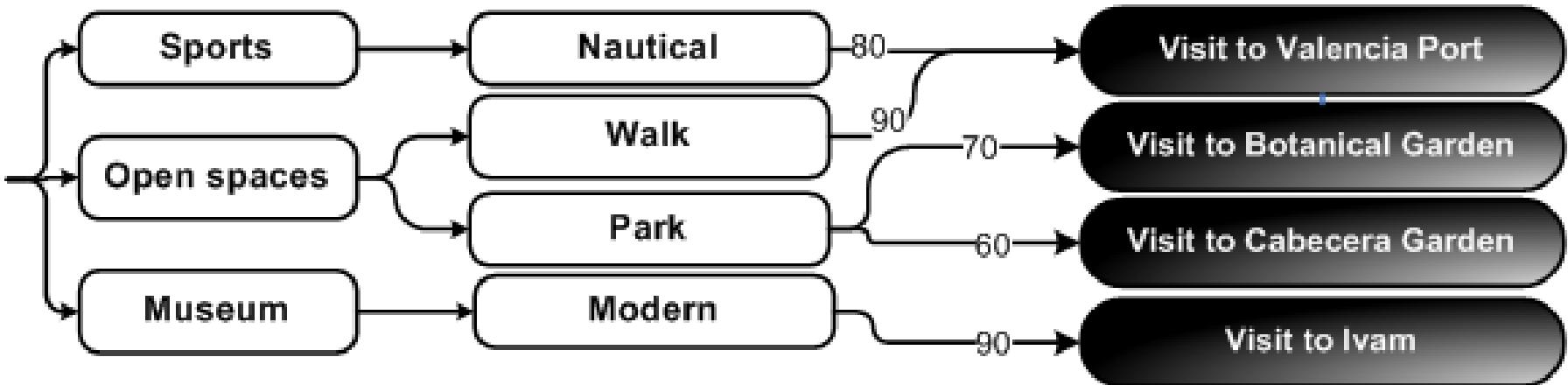
Ítems a recomendar

Los ítems deben estar clasificados según una ontología o taxonomía

Cada ítem tiene asociado una serie de características o preferencias

A mejor clasificación, y mayor variedad de ítems, mejor recomendación

Taxonomía



Ejercicio

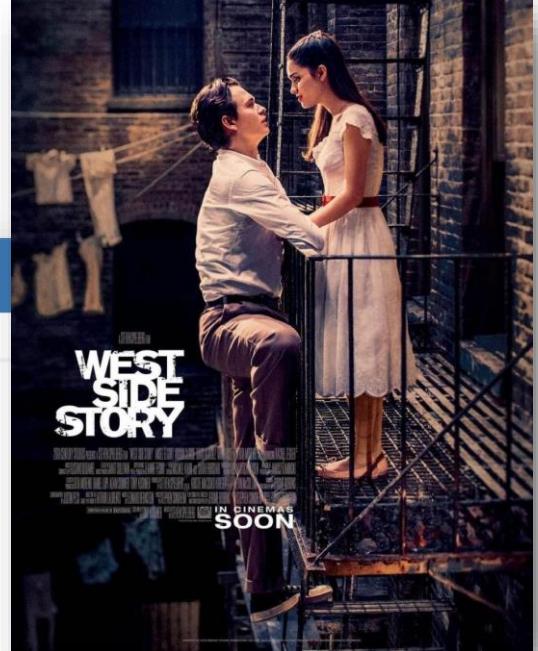
Clasificar estas películas

MovieLens clasifica las películas en 19 categorías

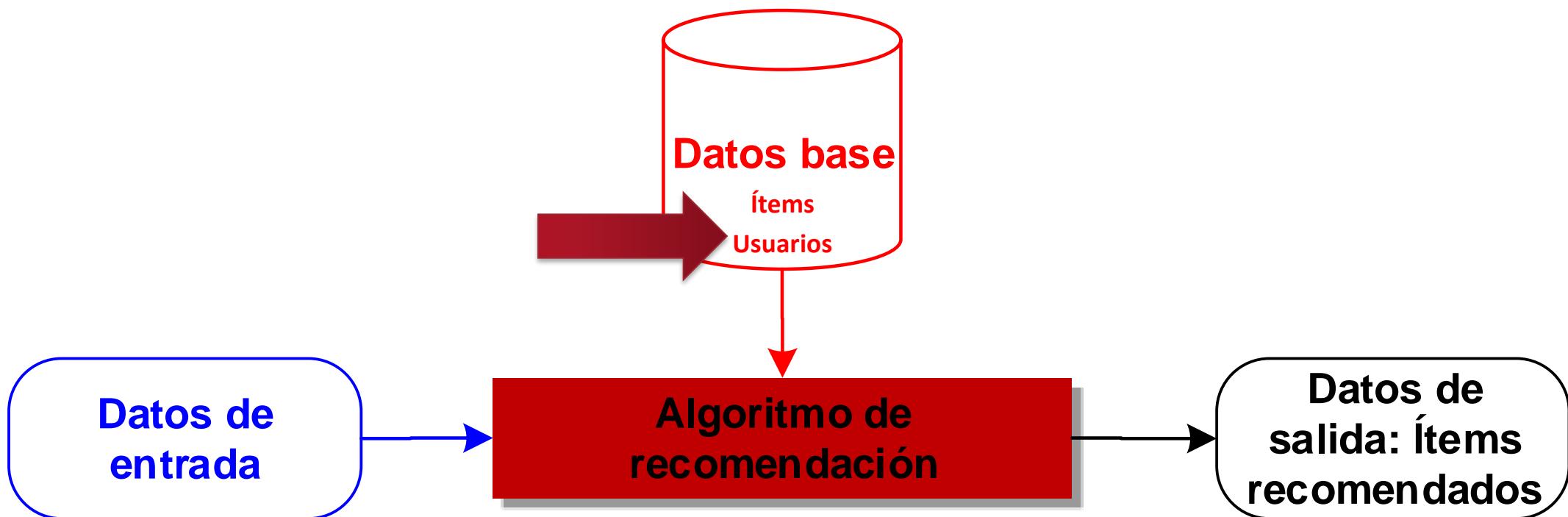
Puede clasificarse en varias categorías

Se puede añadir ratios a la clasificación (grado en que la película pertenece a esa categoría)

	genre: Bloc ...
0	unknown
1	Action
2	Adventure
3	Animation
4	Children's
5	Comedy
6	Crime
7	Documentary
8	Drama
9	Fantasy
10	Film-Noir
11	Horror
12	Musical
13	Mystery
14	Romance
15	Sci-Fi
16	Thriller
17	War
18	Western



Datos base: usuarios



Perfil de usuario

La información que se recopila sobre el usuario se almacena en el **perfil de usuario**

La información se refina y completa a medida que el usuario interacciona con el sistema

Según la técnica de recomendación utilizada, se puede necesitar información de otros usuarios

- Esta información, si se tiene, se suele almacenar en perfiles

Se tiene una base de datos de usuarios donde se almacena la información

Suele contener información sobre un usuario, aunque la recomendación sea para un grupo

- Se podría mantener, además, un perfil de grupo

Perfil de usuario

La información del perfil se obtiene de dos formas

El usuario la introduce
manualmente

Se infiere mediante la forma
en que utiliza el sistema

Introducir la
información no debe
ser demasiado tedioso
para el usuario

Puede completarla o
modificarla en
cualquier momento

Permite obtener
nueva información del
usuario

Refinar la información
existente

Perfil de usuario

Si el usuario
se registra en
el sistema

- Introduce los datos o accede a su perfil
- Mejora la recomendación
- Mantiene su perfil actualizado

Si el usuario
no se registra

- Se usa la información disponible
 - Cookies
 - Historial de accesos o búsquedas

Perfil de usuario

Los SR pueden inferir información del usuario de varias formas

Cuando el usuario indica si le gusta o no un ítem

El sistema observa a que ítems accede y con que frecuencia

El sistema observa que le gusta, basándose, por ejemplo en redes sociales, accesos a internet,...

Solicitándole directamente información demográfica o sobre sus gustos

Información contenida en el perfil de usuario

Información demográfica

- Información personal: edad, género, familia, país,...

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,...

Más información
Mejor información



Mejor recomendación

Información demográfica

Cualquier dato que tengamos del usuario

Edad, sexo, profesión, familia, nivel cultural,...

Talla, color favorito, estilo,...

Lugar de residencia, país, religión,...

Cualquier cosa puede ser de interés

Preferencias del usuario

Descripción de los tipos de ítems en los que está interesado el usuario

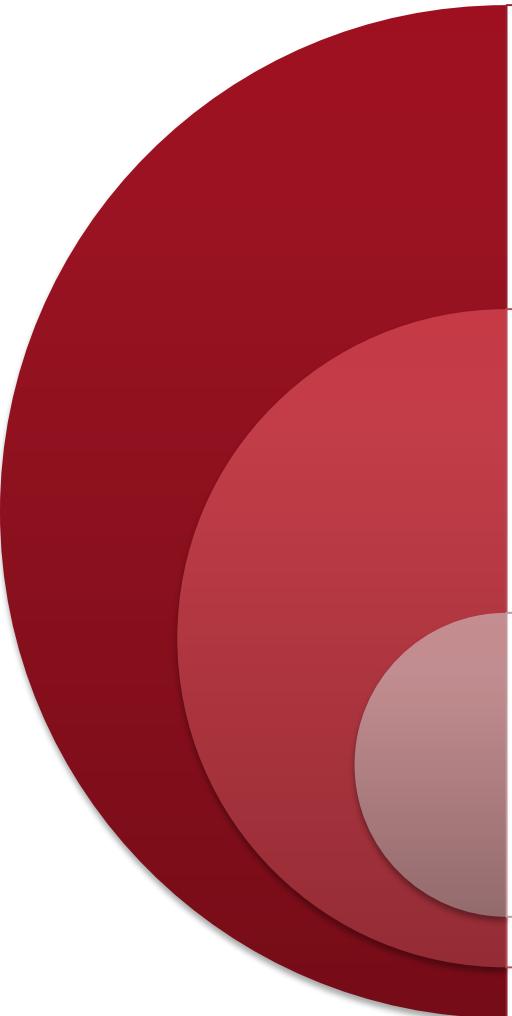
El modelo de preferencias se basa en la clasificación de los ítems a recomendar

Conjunto de características

Para cada característica se almacena el **ratio de interés o utilidad** de la característica para el usuario

El cálculo de la utilidad se basa en heurísticas predefinidas

Preferencias del usuario



El modelo de preferencias del usuario puede no ser único

Ejemplo: en el trabajo vamos a tener tres modelos de preferencias, demográficas, colaborativas y basadas en contenido

Nota: realmente no es así, los gustos son las preferencias basadas en contenido y las demográficas y colaborativas son parte de la información necesaria para el proceso de recomendación. Lo usamos así por simplicidad

Ejemplo

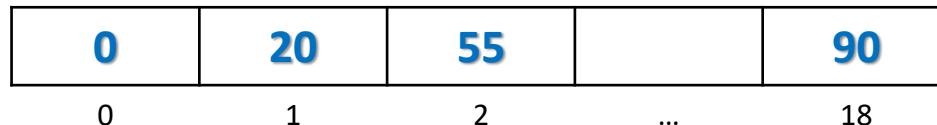
En la BD de MovieLens los ítems (películas) están clasificados en 19 categorías

Una película puede estar clasificada en más de una categoría

El modelo de preferencias de usuario sería un vector de rango 19

Cada posición del vector podría ser un booleano indicando si está interesado o no en esa categoría

O bien un valor numérico indicando en qué grado le gusta la categoría

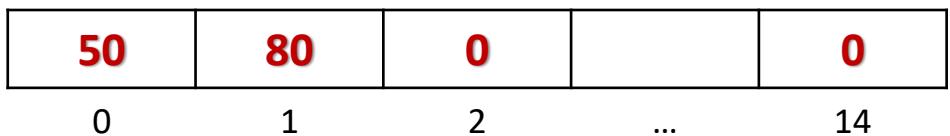


	genre:	Bloc ...	-	□	X	
		Archivo	Edición	Formato	Ver	Ayuda
0	unknown					
1	Action					
2	Adventure					
3	Animation					
4	Children's					
5	Comedy					
6	Crime					
7	Documentary					
8	Drama					
9	Fantasy					
10	Film-Noir					
11	Horror					
12	Musical					
13	Mystery					
14	Romance					
15	Sci-Fi					
16	Thriller					
17	War					
18	Western					

Ejemplo

En un dominio de turismo, por ejemplo, se podrían tener las siguientes 15 categorías

En este caso el vector tendría rango 15



- Styles and periods
- Shopping
- Museums
- Open Spaces
- Religious architecture
- Defensive architecture
- Civil Architecture
- Gourmet
- Sports
- Monuments
- Leisure
- Health & SPA
- Events
- Children
- None

Ratio de las preferencias del usuario

Ratio asociado a la preferencia

Indica el grado de interés del usuario en la característica

true	false	false		true
0	1	2	...	N

90	0	100		40
0	1	2	...	N

-1	1	0.3		0
0	1	2	...	N

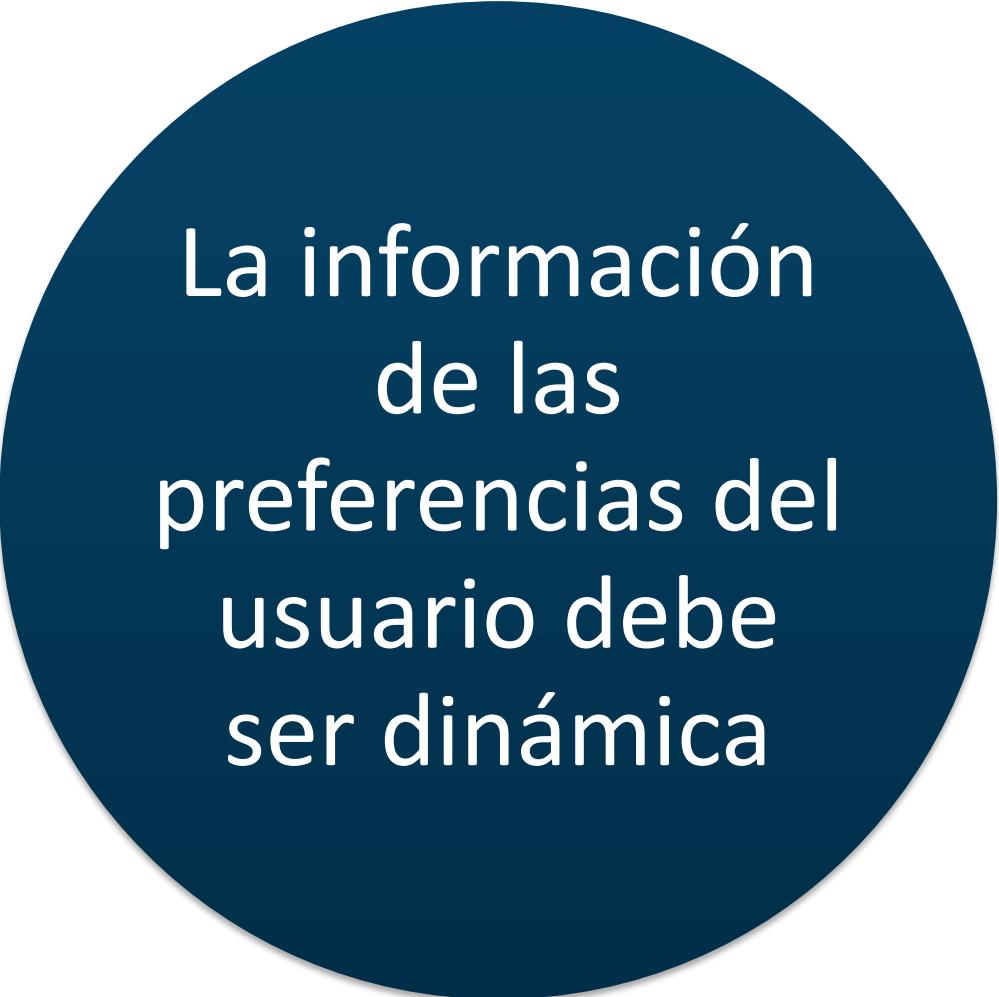
Puede ser

Un valor booleano (tiene interés/no tiene interés)

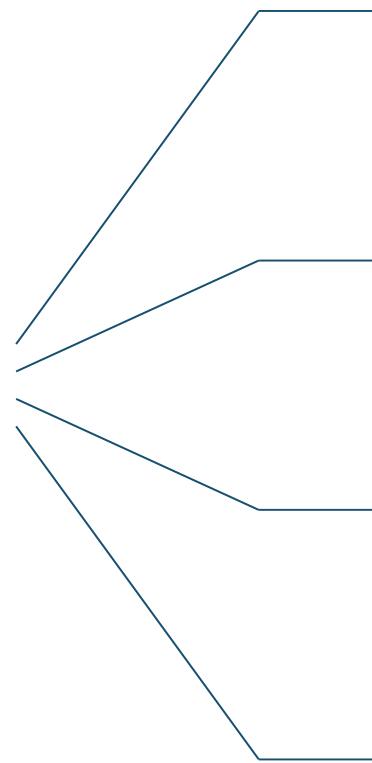
Un valor que indique el grado de interés, por ejemplo valores entre 0 y 100, entre 0 y 1,...

Podrían usarse valores tanto positivos como negativos

Preferencias del usuario



La información
de las
preferencias del
usuario debe
ser dinámica



Se modifica durante las sucesivas recomendaciones

Se añaden nuevas características

Se modifica el interés del usuario sobre la característica (ratio)

Se eliminan características obsoletas

Preferencias del usuario

Ejemplo



Si tienes un bebé,
compras cosas para bebés

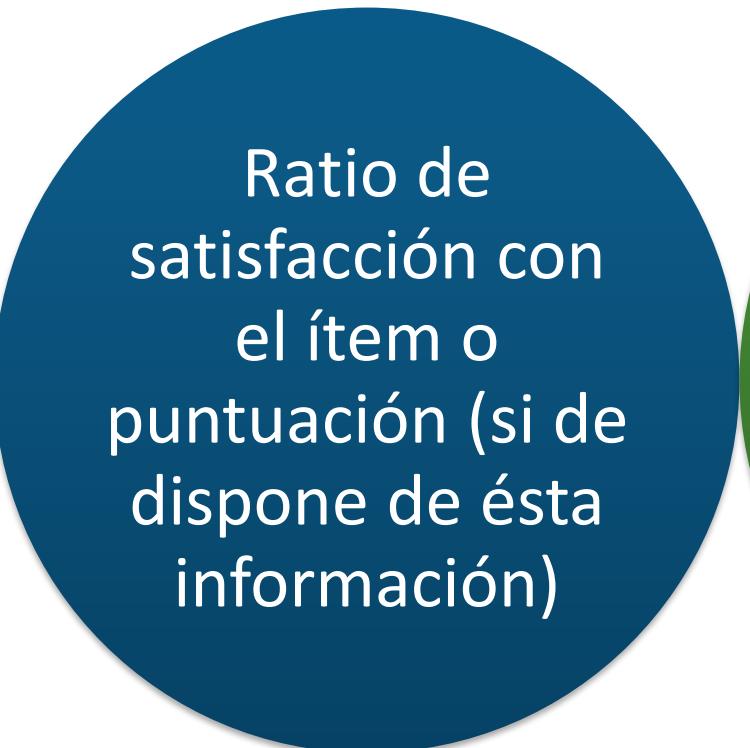


El niño crece y ya no
compras las mismas cosas

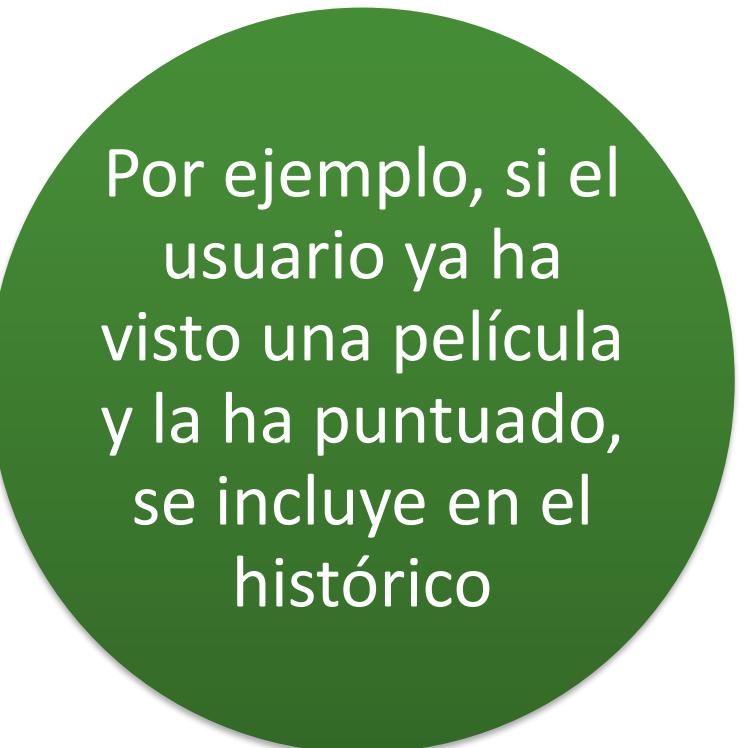
Histórico de interacción del usuario con el sistema



Ítems que el usuario ha seleccionado previamente



Ratio de satisfacción con el ítem o puntuación (si de dispone de ésta información)



Por ejemplo, si el usuario ya ha visto una película y la ha puntuado, se incluye en el histórico

Información interna

Cualquier otra información del proceso de recomendación que se obtiene para utilizar las técnicas y que no está dentro de las categorías anteriormente vistas

Perfil de usuario compartido

Ha habido iniciativas para que las aplicaciones compartan perfil de usuario

Han sido más o menos bien acogidas...

OpenId

APML

Pero la tendencia actual es usar los perfiles de las redes sociales, que aportan mucha información

Facebook

Linkedin

...



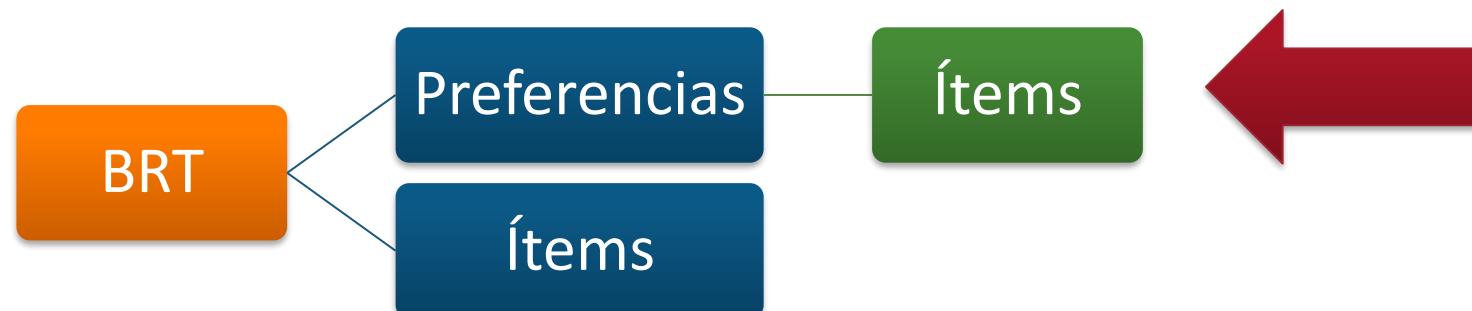
Uso de preferencias o ítems

Preferencias o ítems

Un SR calcula una lista de ítems recomendados

Puede trabajar de dos formas

- Calcular la lista de ítems directamente
- Calcular una lista de preferencias (categorías de la ontología) y después calcular los ítems



Preferencias o ítems

Basada en preferencias

- Primero se calculan las preferencias recomendadas
- Luego los ítems que satisfacen dichas preferencias

Basada en ítems

- Algunas implementaciones de las técnicas de recomendación obtienen directamente un conjunto de ítems recomendados

Problemas del proceso de recomendación

Características deseables en un SR

Debe ser capaz de dar una recomendación independientemente de los datos de los que se disponga

Incluso a un usuario que acaba de entrar en el sistema y del que no se dispone de información

Debe introducir novedad en la recomendación

Aunque siempre debe recomendar ítems que considere que puedan ser de interés para el usuario

Debe ser capaz de dar una recomendación a cualquier usuario

Por muy “poco comunes” que sean sus gustos (oveja negra)

Debe recomendar ítems parecidos, pero no demasiado parecidos

Por ejemplo, no se pueden recomendar dos versiones de “El Quijote”

Pero si se puede si es una edición especial

Problemas

Nuevo usuario

Nuevo ítem

Sparsity, dispersión
de ratios u oveja
negra

Efecto portfolio

- ¿Ítems parecidos o no?

Serendipity o
novedad en la
recomendación

Problemas

Nuevo usuario o start-up problem

- Cuando un usuario es nuevo en un sistema
 - No se conoce nada de él
 - No ha interaccionado con el sistema
 - No ha puntuado ítems
- No se dispone de información sobre sus gustos
- Si el usuario rellena manualmente su perfil de usuario o si se obtienen datos sobre él de otra forma, ya no sería un problema

Problemas

Nuevo ítem

- Hasta que un elemento no ha sido puntuado por un número suficiente de usuarios, no es recomendado (depende de la técnica)
- Un ítem que tenga pocas puntuaciones de los usuarios, es difícil que sea recomendado
- Sólo es un problema si la técnica de recomendación se basa en los ítems que otros usuarios puntúan
 - Por ejemplo, en Amazon, “*otros usuarios a los que les gustan los libros de cocina, también les gusta los robots de cocina*”

Problemas

Sparsity

Dispersión de ratios

Oveja negra

- Un usuario con **gustos poco comunes** será difícil de satisfacer
- Los ratios del usuario actual pueden no coincidir con los ratios de otros usuarios
 - Pocos usuarios con los que comparar
 - Pocos ítems con los que buscar similitudes
- Un usuario con **gustos comunes** tiene mayor probabilidad de que se acierte en la recomendación ofrecida

Problemas

Efecto portfolio

En algunos SR, no se deben recomendar ítems demasiado parecidos a otros que ya ha visto/comprado antes

- No se puede sugerir que compre un libro que ya tiene, una nueva edición, una encuadernación diferente,... o si...

En otros SR, la similitud si es un criterio para recomendar ítems

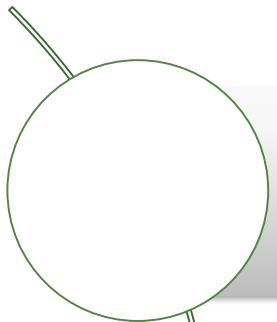
- Una noticia parecida a otra que ya ha visto anteriormente,...
- En este caso, ¿durante cuento tiempo mostrar ítems parecidos?

Problemas

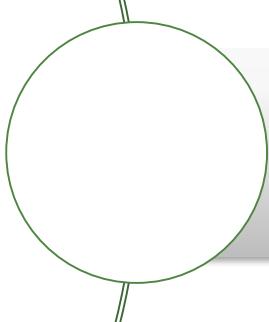
Serendipity (introducir novedades)

- Algunos SR ofrecen al usuario únicamente ítems similares a los ya vistos anteriormente
 - Incrementa la satisfacción del usuario
 - Sería deseable ofrecerle cosas nuevas que le puedan gustar: **novedad**
 - ¿Cómo ofrecer novedades que al mismo tiempo se adecúen a sus gustos?

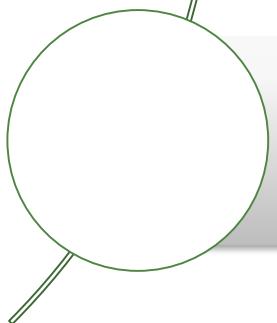
Otros problemas



Cómo recopilar información sobre el usuario



Falta o exceso de resultados de la recomendación



Alta o baja sensibilidad al cambio de preferencias

Problemas

Recopilar información del usuario

- A veces, el usuario debe introducir **información personal**
 - Miedo a la falta de privacidad o poca confianza en el servicio
 - Puede que el usuario no esté dispuesto a proporcionar al sistema información que podría resultar valiosa para la recomendación
- Otras veces es necesario que de información sobre sus **gustos** e intereses
 - Proceso tedioso
 - A mayor información proporcionada, de mayor calidad será la recomendación
- **Solución**
 - Determinar la cantidad óptima de información que se debe solicitar (problema de optimización)
 - Utilizar técnicas no intrusivas (feedback, otras aplicaciones, observar el comportamiento)

Problemas

Falta o exceso de resultados de la recomendación

- Los SR asumen que
 - Los requerimientos del usuario van a poder satisfacerse utilizando las técnicas de recomendación adecuadas
 - Siempre se va a poder dar una recomendación
- **Pocos resultados**
 - El proceso de recomendación puede fallar, no produciendo ninguna recomendación que satisfaga los requerimientos
- **Demasiados resultados**
 - La recomendación no tiene sentido
 - Solución: SR conversacionales

Problemas

Sensibilidad al cambio de preferencias

Si el usuario cambia de preferencias, ¿el SR es capaz de darse cuenta?

Solución: feedback

**Gracias por vuestra
atención...**