
Recomendación Basada en Contenido

Alondra Berzunza



Recomendación Basada en Contenido

Sugieren ítems a los usuarios basándose en sus preferencias previas y las características de los ítems, en lugar de en el comportamiento de otros usuarios.

Estos sistemas analizan los metadatos de los ítems que le han gustado al usuario (como el género de una película o el autor de un libro) para generar un perfil de usuario y recomendarle ítems similares.

Recomendación Basada en Contenido

- Sistemas basados en los atributos de los productos/elementos.
 - Combina calificaciones con el contenido.
 - Muy útil cuando no hay suficientes calificaciones de algún producto/elemento.
 - Desventajas:
 - Las recomendaciones tienden a no ser tan diversas.
 - No es efectiva para nuevos usuarios.
-

Perfiles de Productos/Elementos

Representación de cada producto usando atributos distintivos

Por ejemplo una película:

- Actores
- Director
- Año de realización
- Género

Otros dominios: videojuegos, música, libros, noticias, etc.

Funcionamiento

1. **Análisis de Ítems:** Se extraen características de los ítems disponibles.
 2. **Creación de un Perfil de Usuario:** Se construye un perfil del usuario a partir de los ítems con los que ha interactuado previamente, ya sea explícitamente.
 3. **Generación de Recomendaciones:** El sistema compara las características de los nuevos ítems con el perfil del usuario para recomendar aquellos que tengan similitudes.
-

Ejemplo



Si a un usuario le han gustado y visto varias películas de ciencia ficción y del director Christopher Nolan, un sistema basado en contenido podría recomendarle otra película de ciencia ficción de un director con un estilo similar al de Nolan, o incluso otra película del mismo Nolan.

Ventajas

- **Independencia de otros usuarios:** Las recomendaciones se basan únicamente en la información y preferencias del usuario, lo que es útil para usuarios nuevos o con gustos muy específicos.
 - **Recomendación de contenido nuevo:** Puede recomendar ítems que nunca han sido consumidos por otros usuarios, ya que se enfoca en las características del ítem en sí.
-

Recomendación de Documentos de Texto

No es fácil asignar directamente un conjunto de atributos general.

Ejemplos:

- Artículos científicos
- Noticias
- Blogs

Cálculo de similitud de documentos.

- Una estrategia para representar los documentos es usar las n palabras con mayor pesos *tf-idf*.
 - Métricas comunes: Jaccard y coseno.
-

Perfiles de Usuarios

- Tienen la misma representación que perfiles de productos/elementos.
 - Agregado de los perfiles de productos/elementos que le gustan al usuario.
 - Una estrategia es usar el promedio.
 - Por ej. si el 30 % de las películas que le gustan a un usuario son de acción, ese atributo tendría el valor 0.3 en el perfil del usuario.
 - Es posible calcular la similitud o distancia de perfiles de usuarios y productos/elementos.
-

Clasificadores

Otra estrategia es usar los perfiles de productos/elementos y las matrices de utilidad para entrenar clasificadores.

- Para cada usuario se entrena un clasificador que predice las calificaciones que le dará a cualquier producto/elemento.

Ejemplos de entrenamiento:

- Entrada: perfil de producto/elemento calificado por el usuario.
 - Salida: calificación otorgada.
-

Práctica

Construir un sistema que recomiende películas similares a una película concreta.

Calcular las puntuaciones de similitud por pares para todas las películas basándose en las descripciones de sus argumentos y recomendar en función de un umbral de puntuación de similitud.

<https://www.kaggle.com/datasets/grouplens/movielens-20m-dataset>
