

Sistemas de recomendación y cómo hacerlos

Un sistema de recomendación es una herramienta o algoritmo que sugiere elementos relevantes a los usuarios. Podemos verlo aplicado en la recomendación de productos en casi todas las tiendas en línea, en la recomendación de contenido de redes sociales, en la recomendación de anuncios, etc. Estas recomendaciones están basadas en las preferencias del usuario para maximizar su satisfacción.

Si lo pensamos de cierta manera, es como tener un asistente personalizado que nos ayuda a optimizar el tiempo. Muchas aplicaciones famosas como Netflix lo usan a manera de mantener la atención del usuario, es decir, si recomiendan buenas películas, pueden maximizar la interacción con la página antes de que el usuario pierda interés. El mismo principio es aplicado a YouTube y Spotify.

Existen varios tipos de sistemas de recomendación, según los algoritmos:

Los basados en filtrado colaborativo analizan el comportamiento entre varios usuarios para encontrar similares. El sistema recomienda productos que uno de ellos consume y el otro no.

Los basados en contenido recomiendan elementos similares con los que interactuamos previamente. Como cuando has visto películas de terror antes y se te recomiendan más películas de terror.

Los basados en conocimiento utilizan información que tienen sobre los usuarios (como la edad) para hacer recomendaciones.

Los de aprendizaje profundo utilizan redes neuronales para analizar grandes volúmenes de información y hacer recomendaciones según el entrenamiento.

Para medir la eficacia de un sistema de recomendación se toman en cuenta tres factores:

- Precisión: hacer recomendaciones que estén lo suficientemente personalizadas según los intereses y gustos del usuario final.

- Capacidad de respuesta: la capacidad de recomendar nuevas opciones rápidamente y mantener el interés del usuario.
- Costo: las inversiones iniciales en comparación con los gastos operativos.

¿Cómo funcionan?

Son un canal de varios modelos de IA y flujos de trabajo de análisis de datos diferentes. En cualquier flujo de trabajo generalmente hay tres etapas:

- Clasificación: usan la visión artificial y el procesamiento de lenguaje natural para clasificar los elementos de un contenido.
- Recordatorio y búsqueda de similitudes: reúnen diferentes categorías de características similares entre elementos u objetos.
- Clasificación: el recomendador ordena los artículos por relevancia, a menudo con modelos de aprendizaje profundo de Wide & Deep o DLRM.

¿Cómo construir uno?

La creación de un sistema de recomendación es un proceso estratégico que implica comprender los datos de los usuarios, aprovechar los algoritmos de aprendizaje automático y perfeccionar continuamente el sistema en función de las interacciones de los usuarios.

1. Recolección de datos: datos del usuario, los ítems como las películas o canciones y las interacciones que ha tenido como calificaciones, clicks, tiempo de visualización, etc.
2. Preprocesamiento de datos: antes de entrenar el sistema los datos deben estar limpios, eliminar datos nulos, normalizar los valores.
3. Elección del algoritmo: elegir cómo se harán las recomendaciones, ya sea basado en contenido, colaborativo o híbrido.
4. Entrenar el modelo: ajustar parámetros para mejorar la precisión
Para filtrado basado en contenido:
 1. Crea vectores de características para los ítems.

2. Usa una métrica de similitud (por ejemplo, similitud del coseno).
3. Calcula los productos similares y recomienda.

Para filtrado colaborativo:

1. Construye una matriz de calificaciones (usuarios vs. productos).
 2. Usa SVD (Singular Value Decomposition) para factorizar la matriz.
 3. Predice las calificaciones faltantes y recomienda los productos mejor puntuados.
5. Evaluación y ajustes: usar métricas como RMSE o precisión
 6. Despliegue: crear una API o integrarlo en una plataforma web

Herramientas de Amazon Web Services

Amazon Personalize: Es un servicio que utiliza machine learning y es específico para recomendaciones. Está creado para simplificar el proceso de creación de modelos aunque no tengas experiencia en machine learning. Se encarga desde la preparación de datos hasta implementar el modelo. Viene con algoritmos pre entrenados y algunos casos de uso son productos, contenidos, ofertas, etc.

Amazon SageMaker: Sirve para construir, entrenar e implementar modelos en general. Se usa para sistemas de recomendación complejos y modelos personalizados.

Existen otros servicios como Lambda que implementan el código de recomendación según eventos, DynamoDB o Redshift que se encarga de grandes volúmenes de datos, etc. Pero principalmente son Personalize y SageMaker para creación de modelos de recomendación.

Herramientas de Google Cloud Platform

Recommendations AI: Servicio específico para construir sistemas de recomendación personalizados. Se usa sobre todo para productos y contenido, ya que es fácil de integrar con catálogos e interacción.

Vertex AI: Plataforma para machine learning que permite construir, entrenar e implementar todo tipo de modelos de IA. Soporta frameworks populares.

BigQuery ML (Machine Learning): Para crear y entrenar modelos de machine learning directamente desde el almacén de datos de google (bigquery).

Frameworks

TensorFlow es uno de los frameworks más populares de machine learning y deep learning. Es una de las formas de trabajar proporcionando una estructura para construir y entrenar los modelos.

Algunos otros son PyTorch (de Facebook), Keras (ahora parte de tensorflow), Scikit-learn (código abierto), LightFM (de Lyst), etc.

Algoritmos

Arriba fueron mencionados algunos enfoques de los algoritmos, pero dentro de estos enfoques existen otros algoritmos.

Por ejemplo, los basados en filtrado colaborativo son los que agrupan comportamientos similares. Dentro de esta categoría está el algoritmo de usuario a usuario (recomendar lo que compran usuarios parecidos), los de item a item (recomendar más productos similares a los que ya compraste), los de factorización de matrices (descompone las interacciones en un valor para predecir preferencias).

Los algoritmos basados en contenido contienen el de similitud de características y el de clasificación (clasifica los usuarios en grupos de interés).

Algoritmos Híbridos que combinan distintos enfoques, como el de ponderación que suma las puntuaciones, o el de conmutación que los intercambia según el contexto.

Además, están los algoritmos de deep learning que meten los datos a una red neuronal y la caja negra se encarga de tomar decisiones, como las redes neuronales de factorización, modelos de secuencia y autoencoders.