



Sistemas de Recomendación

Sarabia Guzmán Jesús Aldahir

Zamudio Lizárraga Bryan Martín

¿Qué es un sistema de recomendación?

- Es una herramienta que recomienda productos, contenido o conceptos en base a las preferencias del usuario.
- Hay varios tipos, como los basados en contenido, los colaborativos o los basados en interacciones.

FILTRADO COLABORATIVO




FILTRADO BASADO CONTENIDO



Construir un sistema de recomendación

Se recopilan los datos sobre los usuarios (edad, género, historial) y los productos (descripción, categoría).



Se elige el tipo de sistema de recomendación que se usará.



Se escoge el algoritmo a utilizar, por ejemplo en un sistema colaborativo se podría usar K-Nearest Neighbors o factorización matricial.

Ejemplo Sistema de recomendación en Python (Películas)

- Utilizando la librería pandas cargamos en un dataframe un csv con información de miles de películas (su nombre, descripción, géneros, etc.)

```
# Import Pandas
import pandas as pd

# Load Movies Metadata
metadata = pd.read_csv('movies_metadata.csv', low_memory=False)
```

tiempo de ejecución	spoken_languages	estado	eslogan	título	video	vote_average
81.0	[{'iso_639_1': 'es', 'name': 'Español'}]	Publicado en	NaN	Toy Story	Falso	7.7
104.0	[{'iso_639_1': 'es', 'name': 'Inglés'}, {'iso...	Publicado en	¡Tira los dados y desata la emoción!	Jumanji	Falso	6.9
101.0	[{'iso_639_1': 'es', 'name': 'Español'}]	Publicado en	Sigue gritando. Sigue luchando. Aún listo para...	Viejos gruñones	Falso	6.5

- Hacemos un sistema de rec. basado en la similitud de la descripción de las películas.
- Transformamos el texto de las descripciones (overview) en una matriz numérica usando TF-IDF. Esto representa qué tan importantes son las palabras en un texto dado.
- En esta matriz, cada fila es una descripción, y cada columna corresponde a las palabras. Cada valor que tenga la matriz será un valor numérico basado en la relevancia de cada palabra.

```
#Import TfIdfVectorizer from scikit-learn
from sklearn.feature_extraction.text import TfidfVectorizer

#Define a TF-IDF Vectorizer Object. Remove all english stop words such as '
tfidf = TfidfVectorizer(stop_words='english')

#Replace NaN with an empty string
metadata['overview'] = metadata['overview'].fillna('')

#Construct the required TF-IDF matrix by fitting and transforming the data
tfidf_matrix = tfidf.fit_transform(metadata['overview'])
```

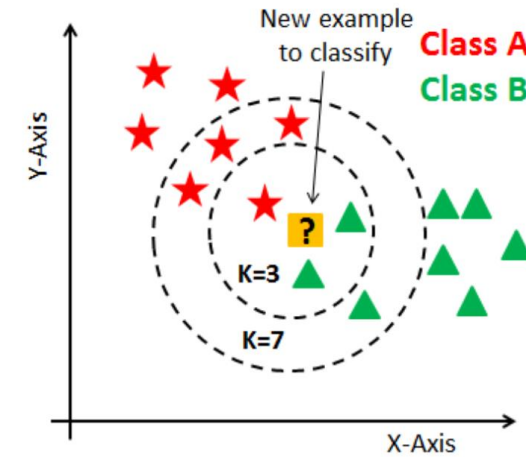
```
# Import linear_kernel
from sklearn.metrics.pairwise import linear_kernel

# Compute the cosine similarity matrix
cosine_sim = linear_kernel(tfidf_matrix, tfidf_matrix)
```

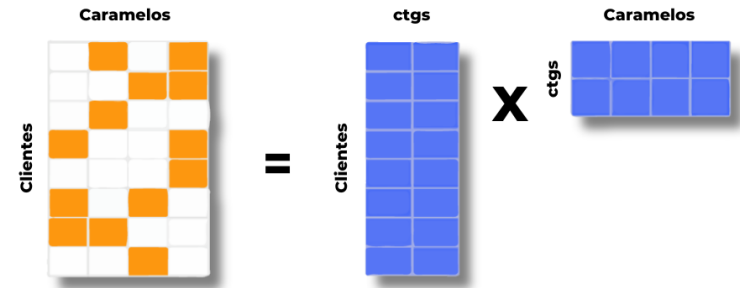
- Teniendo la matriz anterior, calculamos el valor de similitud de cada película, lo cual nos devuelve una matriz cuadrada. En este ejemplo, cada película será un vector de 1x45466, donde cada columna será el valor de similitud con las demás películas.
- Después, es cuestión de ordenar de mayor a menor para saber cuáles películas son más similares, y dar una recomendación basándose en esta información.

Algoritmos utilizados

- Los algoritmos utilizados dependen del tipo de sistema de recomendación.
- En el filtrado colaborativo se pueden usar los anteriormente mencionados KNN y Factorización Matricial



Factorización de matrices



Frameworks utilizados

- **Scikit-learn**, es una librería de Python de código abierto, ampliamente utilizada para el machine learning. Aunque no está diseñada específicamente para sistemas de recomendación, puede aplicarse en métodos básicos como el filtrado colaborativo o el clustering.
- **TensorFlow**, desarrollado por Google, se usa para desarrollar aplicaciones avanzadas de machine learning, como redes neuronales profundas
- **LightFM**, es una librería de Python especializada en sistemas de recomendación. Es ideal para recomendar productos o servicios en los que se dispone de información adicional sobre los usuarios o los artículos (como género, categoría, etc.).

Herramientas disponibles en Amazon Web Services (AWS)



- **Amazon Personalize** es un servicio gestionado de machine learning ofrecido por AWS que permite crear sistemas de recomendación personalizados sin necesidad de tener experiencia en machine learning. Utiliza los mismos algoritmos de aprendizaje automático que Amazon usa internamente para sus propias recomendaciones.
- **Amazon SageMaker** es un servicio gestionado de AWS para construir, entrenar y desplegar modelos de machine learning a escala. Permite crear sistemas de recomendación personalizados mediante algoritmos avanzados y herramientas de ML.

Herramientas disponibles por Google Cloud Platform (GCP)

- BigQuery es una plataforma de almacenamiento y análisis de datos en la nube de Google, optimizada para manejar grandes volúmenes de datos. Se suele usar principalmente para analizar grandes cantidades de datos de interacciones entre los usuarios y los productos.
- AI Platform es un conjunto de herramientas y servicios de Google Cloud hecho para crear y entrenar los modelos de machine learning a gran escala. Es ideal para crear modelos de recomendación personalizados utilizando datos de interacción de usuarios y artículos.



BigQuery ML

Referencias

- <https://www.smartpanel.com/en/como-funciona-la-inteligencia-artificial-en-los-sistemas-de-recomendacion/>
- <https://aprendeia.com/sistema-de-recomendaciones-inteligencia-artificial/>
- <https://www.datacamp.com/es/tutorial/recommender-systems-python>
- <https://www.incentro.com/es-ES/blog/que-es-tensorflow>
- <https://medium.com/@williamwarley/mastering-ai-development-with-google-cloud-platform-a-comprehensive-guide-to-building-ai-solutions-7490c61a21f2>