

Sistemas de recomendación

ANEXOS DE LOS RECURSOS DE SECUENCIACIÓN

Anexo 1: Contenido y aprendizaje para la contextualización del funcionamiento de un sistema de recomendación

Anexo 2: Creación de actividad para un sistema de recomendación basado en Filtrado Colaborativo desde las matemáticas.

Anexo 3: Creación de actividad para un sistema de recomendación basado en Contenidos desde las matemáticas

Anexo 4: Creación de actividad para un sistema de recomendación basado en Contenidos desde la programación con Python

Anexo 5: Plantilla en notebook Jupyter para proyecto Netflix

Anexo 6: Reflexión y estudio de los aspectos éticos de un sistema de recomendación

ANEXO 1

Contenido y aprendizaje para la contextualización del funcionamiento de un sistema de recomendación

Objetivo de situación de aprendizaje

El objetivo de esta situación de aprendizaje es el desarrollo por parte del alumnado de un sistema de recomendación contextualizado en un entorno real. Para el desarrollo de este prototipo se plantean 3 actividades, de las cuales 2 son en formato proyecto de una forma más analógica y la tercera actividad, dará la oportunidad al alumnado de crear un prototipo basado en un lenguaje de programación como Python. Esta tercera actividad no solo llevará el desarrollo y entendimiento del algoritmo sino también la base de programación necesaria para desarrollar un sistema más generalista y aplicable a diferentes entornos o datos. El objetivo es que el alumno acabe la situación de aprendizaje con el desarrollo real de un prototipo totalmente utilizable y aplicable de manera real en el contexto elegido.



Fig 1 Pexel.com

Introducción y contextualización

Los sistemas de recomendación son servicios que actualmente están muy cerca de la sociedad debido a su uso y beneficio, no solo para el usuario si no también principalmente para la empresa que lo implementa.

Es objetivo principal de esta actividad entender la funcionalidad de estos servicios desde una vertiente matemática y lógica, además el conocimiento de esta funcionalidad permitirá al alumnado tener unos mayores fundamentos para la comprensión del impacto e influencia que pueden llegar a tener estas plataformas sobre el usuario final.



Fig 2 pexel.com combinación propia.

¿Qué es un sistema de recomendación?

Un sistema de recomendación es una herramienta que utiliza datos sobre las preferencias y comportamientos del usuario para sugerir elementos que pueden interesarle. Estos sistemas se usan comúnmente en una variedad de entornos, como plataformas de compras en línea, servicios de transmisión de video y sitios de redes sociales.

El objetivo de un sistema de recomendación es mejorar la experiencia del usuario proporcionando sugerencias personalizadas que sean relevantes y útiles para el usuario individual. En muchos casos, los sistemas de recomendación utilizan algoritmos de aprendizaje automático para analizar los datos del usuario y hacer predicciones sobre lo que le puede interesar al usuario. Estos algoritmos pueden tener en cuenta factores como las compras anteriores del usuario, los artículos que le han gustado o en los que ha hecho clic, y los artículos que son populares entre otros usuarios con intereses similares.

También la ética es una cuestión muy importante en el diseño e implementación de sistemas de recomendación de las grandes empresas. Estos sistemas pueden tener un gran impacto en la forma en que las personas descubren y consumen productos,

servicios y contenidos, por lo que es crucial asegurarse de que se utilizan de manera justa y equitativa.

En particular, es importante evitar la discriminación y la manipulación en los sistemas de recomendación. Esto significa diseñar los sistemas de manera que no favorezcan a ciertos grupos o individuos en detrimento de otros, y que no se utilicen para influir en la conducta de las personas de manera subliminal o engañosa.

Además, es crucial tener en cuenta el impacto de los sistemas de recomendación en la sociedad en su conjunto. Por ejemplo, es importante evitar que estos sistemas refuercen los prejuicios o estereotipos existentes, o que contribuyan a la polarización o al aislamiento social

Cold start (Arranque en frío)

El arranque en frío, también conocido como **«cold start»**, es un problema común en los sistemas de recomendación. Se produce cuando un usuario nuevo se registra en el sistema y no tiene suficientes datos o información para ser recomendado. Esto puede dificultar la generación de recomendaciones precisas y relevantes para el usuario en cuestión, lo que puede afectar negativamente su experiencia.

Para abordar este problema, existen diferentes estrategias que pueden utilizarse. Por ejemplo, se pueden solicitar al usuario que proporcione información adicional sobre sus intereses o preferencias, como sus gustos musicales o literarios. De esta manera, el sistema puede tener una base de datos más completa y precisa para generar recomendaciones.

Otra estrategia es utilizar el aprendizaje automático para analizar el comportamiento del usuario y detectar patrones que puedan ayudar a predecir sus gustos y preferencias. De esta manera, el sistema puede mejorar sus recomendaciones a medida que el usuario interactúa con él.

En resumen, el arranque en frío es un problema común en los sistemas de recomendación, pero existen diferentes estrategias que pueden ayudar a abordarlo y mejorar la precisión y relevancia de las recomendaciones.

¿Cómo funciona?



Fig 3 3 fotos libres de [Klipartz - Descargar imágenes png de alta calidad](#)

¿Cómo funcionan?

Los sistemas de recomendación suelen funcionar mediante el análisis de datos sobre los usuarios y los artículos que les interesan. Estos datos se pueden recopilar de varias maneras, como rastrear las interacciones de los usuarios con el sistema (por ejemplo, en qué artículos hacen clic, qué compran, etc.), o recopilando comentarios explícitos de los usuarios (por ejemplo, calificaciones o reseñas que proporcionan para los artículos).

Una vez que se han recopilado los datos, se utilizan para entrenar un algoritmo de aprendizaje automático. Este algoritmo está diseñado para encontrar patrones en los datos que se pueden usar para hacer predicciones sobre qué artículos pueden interesarle a un usuario. Por ejemplo, el algoritmo puede buscar puntos en común entre los usuarios que han mostrado interés en un artículo en particular, o puede analizar los tipos de elementos con los que un usuario ha interactuado en el pasado para hacer recomendaciones para elementos similares.

Una vez que se ha entrenado el algoritmo, se puede utilizar para hacer recomendaciones para usuarios individuales. Por ejemplo, si un usuario está navegando en un sitio web de compras y mira un artículo en particular, el sistema de recomendación podría usar el algoritmo para sugerir otros artículos que podrían interesarle al usuario en función de su comportamiento anterior y el comportamiento de otros usuarios con intereses similares.

En algunos casos, los sistemas de recomendación también pueden tener en cuenta factores adicionales, como la ubicación del usuario o la hora del día, para hacer recomendaciones más personalizadas y relevantes. En general, el objetivo de un sistema de recomendación es mejorar la experiencia del usuario proporcionando

sugerencias personalizadas que se adaptan a los intereses y preferencias individuales del usuario.



Fig 4 PExel.com

Los sistemas de recomendación se pueden estudiar y analizar desde dos tipos, por un lado, el filtro basado en contenidos, el cual pone el foco en el producto más que el usuario y por otro lado el filtrado colaborativo donde su focalización es más en el usuario. A continuación, se hará una descripción de ambos y se explicará la actividad desde el punto de vista funcional y matemático.

Filtrado Colaborativo

Un filtrado colaborativo es el proceso de predecir los intereses de un usuario mediante la identificación de preferencias e información de muchos usuarios. Esto se hace filtrando datos en busca de información o patrones utilizando técnicas que involucran la colaboración entre múltiples agentes, fuentes de datos, etc. La intuición detrás del filtrado colaborativo es que, si los usuarios A y B tienen gustos similares en un producto, entonces es probable que A y B tienen un gusto similar en otros productos también.

La principal ventaja de usar modelos de filtrado colaborativo es su simplicidad de implementación y el alto nivel de cobertura que brindan. También es beneficioso porque captura características sutiles y no requiere la comprensión del contenido del ítem.

Desventajas

La principal desventaja de este modelo es que no es amigable para recomendar nuevos artículos, esto se debe a que no ha habido interacción usuario/artículo con él. Esto se conoce como el problema de arranque en frío (anteriormente explicado).

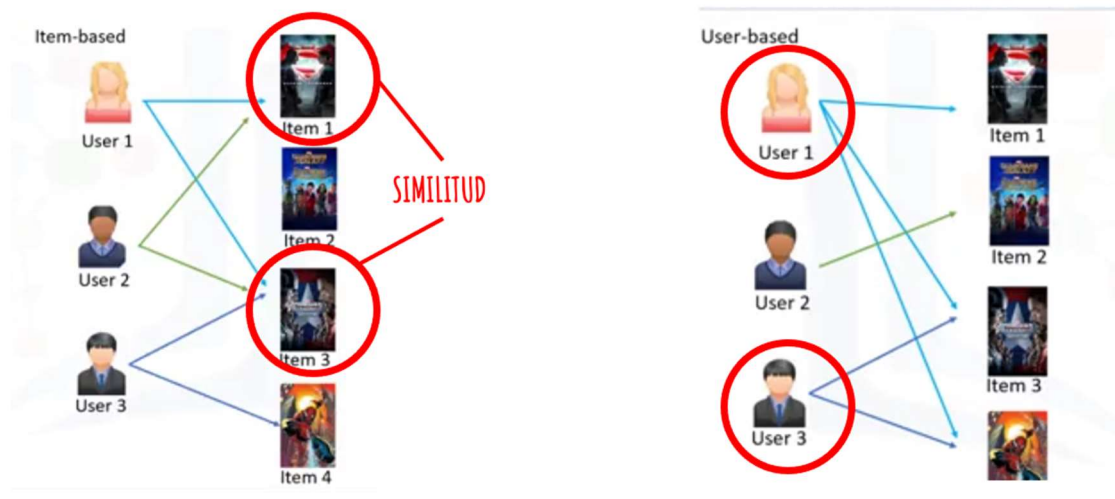
Ejemplos

Algunos ejemplos de algoritmos de filtrado colaborativo:

- **Recomendación de contenido de YouTube a los usuarios:** recomendarle videos en función de otros usuarios que se han suscrito o visto videos similares a los suyos.
- **Recomendación de curso de Coursera:** recomendarle cursos en función de otras personas que han terminado los cursos existentes que usted ha terminado.

Anexo 2: Creación de actividad para un sistema de recomendación basado en Filtrado Colaborativo desde las matemáticas.

Como funciona un Filtrado Colaborativo



Filtrado Colaborativo

Fig 5 Elaboración propia en local

Una vez que se ha explicado en el aula el concepto y contexto de un filtro colaborativo en un sistema de recomendación, el siguiente paso es entender de manera práctica como se puede observar esa similitud enfocada en el usuario.

Se realiza un ejercicio como se muestra en la Fig 6, donde se permite una intuición de cuál puede ser la similitud de gustos por películas de cada usuario.

Observaciones:

Columna Izquierda: Se observa una similitud basada en la elección del producto por parte del usuario 1 y el usuario 2.

Columna derecha: Se observa la similitud de los usuarios 1 y 2 como consecuencia de elecciones muy similares en sus productos.

Caso Real

Se crea una tabla como la que se muestra en la Fig. 7, en las filas tenemos los usuarios y en las columnas las películas que ha visto o no. Las películas están marcadas por una puntuación de 0 a 10, que marca la «similitud o cercanía».

Esas puntuaciones se pueden explicar que provienen de la unión de varias interacciones, por ejemplo, la veces que ha visto esa película, si le ha dado a «me gusta» dentro de la plataforma, si al recomendaría, etc. Para el caso de esta actividad simplemente daremos un valor por concretar las operaciones matemáticas.

Objetivo de la actividad.

- Conocer el valor de similitud de nuestro usuario 4 en esas dos películas rodeadas con un círculo en rojo y signo de interrogación.
- Son valores que no se conocen debido a que no ha existido ninguna interacción entre ese usuario y esa película, por lo tanto, el objetivo es conocer ese valor para recomendar esa película en mayor o menor grado.

Debemos recordar que la importancia de un sistema de recomendación no está solo en saber que recomendar, si no en qué grado hacerlo.

Y las matemáticas

Buscamos la similitud...por medio de la cercanía.

Distancia Euclidiana

Similitud del Coseno


| | |  |  |  |  |  |
|---|---|---|---|--|---|---|
| 1 |  | 9 | 6 | 8 | 4 | × |
| 2 |  | 2 | 10 | 6 | × | 8 |
| 3 |  | 5 | 9 | × | 10 | 7 |
| 4 |  | ? | 10 | 7 | 8 | ? |

Fig. 6 Elaboración propia en local

Para poder calcular ese valor de recomendación del usuario 4 con esos dos productos debemos entender un concepto muy interesante dentro de las matemáticas, «la distancia».

En un conjunto de datos basados en el mismo contexto podemos calcular la cercanía que existe entre ellos gracias a técnicas de distancia como la **distancia Manhattan, la distancia Euclídea o la similitud del coseno**.

En esta actividad usaremos la distancia euclídea para resolver las distancias, pero podría usarse cualquier otra.

Distancia Euclídea

La distancia euclídea es una medida de la distancia entre dos puntos en un espacio bidimensional o tridimensional. Se calcula utilizando la fórmula de la distancia euclídea, que es la raíz cuadrada de la suma de las diferencias entre las coordenadas de cada punto, elevadas al cuadrado.

Por ejemplo, si tenemos dos puntos A y B en un espacio bidimensional con coordenadas (x_1, y_1) y (x_2, y_2) respectivamente, podemos calcular la distancia euclídea entre ellos con la siguiente fórmula:

$$\text{distancia euclídea} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Por lo tanto, si queremos calcular la distancia entre dos puntos en un espacio tridimensional, tendríamos que añadir otro término a la fórmula para tener en cuenta la coordenada z. La fórmula quedaría así:

$$\text{distancia euclídea} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$$

En resumen, la distancia euclídea es una forma de medir la distancia entre dos puntos en un espacio bidimensional o tridimensional utilizando la fórmula matemática de la distancia euclídea.

La similitud del coseno

La similitud del coseno es una medida de similitud entre dos vectores en un espacio vectorial. Se calcula utilizando el coseno del ángulo entre los dos vectores y su magnitud.

Para calcular la similitud del coseno entre dos vectores A y B, podemos utilizar la siguiente fórmula:

$$\text{similitud del coseno} = \frac{A \cdot B}{|A| \cdot |B|}$$

En esta fórmula, "A * B" se refiere al producto escalar de los vectores A y B, mientras que "|A|" y "|B|" son las magnitudes de los vectores A y B, respectivamente.

El resultado de esta fórmula es un número entre -1 y 1, donde un valor cercano a 1 indica una alta similitud entre los vectores, mientras que un valor cercano a -1 indica una baja similitud. Si el resultado es 0, significa que los vectores son ortogonales (es decir, que forman un ángulo de 90 grados).

En resumen, la similitud del coseno es una forma de medir la similitud entre dos vectores utilizando el coseno del ángulo entre ellos y su magnitud. Es una medida muy útil en áreas como el procesamiento del lenguaje natural y la recomendación de contenido.

Ahora debemos conocer la distancia que hay entre el usuario 4, con el resto de los usuarios usando como valores los tres productos que tienen en común

Entenderemos cada producto (película) como la coordenada de un vector, en este caso tenemos un vector de 3 coordenadas, es decir en R^3 . Los valores que no existan (cruz roja) colocaremos un cero.

| | | |
|----|---|----|
| 6 | 8 | 4 |
| 10 | 6 | X |
| 9 | X | 10 |
| 10 | 7 | 8 |

Distancia entre Usuario 4 y Usuario 1

Elaboración propia en local

$$V_1(\text{usuario1}) = (6,8,4)$$

$$V_4(\text{usuario4}) = (10,7,8)$$

$$\sqrt{(6-10)^2 + (8-7)^2 + (4-8)^2} = 5,74$$

Distancia entre Usuario 4 y Usuario 2

$$V_2(\text{usuario2}) = (10,6,0)$$

$$V_4(\text{usuario4}) = (10,7,8)$$

$$\sqrt{(10-10)^2 + (6-7)^2 + (0-8)^2} = 8,06$$

Distancia entre Usuario 4 y Usuario 3

$$V_3(\text{usuario3}) = (9,0,10)$$

$$V_4(\text{usuario4}) = (10,7,8)$$

$$\sqrt{(9-10)^2 + (0-7)^2 + (10-8)^2} = 7,34$$

Posteriormente para que estos valores se mantengan entre 0 y 10 en las recomendaciones, deberemos normalizar nuestras distancias de similitud entre 0 y 1.

¿Por qué debemos normalizar?

Normalizar los datos es una técnica que se utiliza para ajustar las características de los datos a un rango común. Esto puede ser útil por varias razones:

1. **Evitar sesgos de escala:** Algunos algoritmos de *machine learning* son sensibles a la escala de los datos de entrada. Por ejemplo, si tenemos una columna con el número de «clicks» que hace el usuario dentro de Netflix y otra con el número de accesos que tiene ese usuario al día, es evidente que la primera columna tendrá valores mucho más elevados que la segunda y no por esto debe ser más importante o relevante. La normalización ayuda a evitar este problema al ajustar todas las características a un rango común.
2. **Facilitar la comparación:** La normalización también puede facilitar la comparación de características con diferentes escalas. Si dos características tienen diferentes unidades o escalas, es difícil compararlas directamente. La normalización las ajusta a una escala común, lo que facilita la comparación.
3. **Mejorar la eficiencia:** Algunos algoritmos, como los de clasificación, son más eficientes cuando las características están normalizadas. Si una característica tiene una escala muy alta, puede dominar a las demás y hacer que el algoritmo sea menos eficiente. La normalización ayuda a evitar este problema.

¿Cómo normalizamos?

<https://www.youtube.com/watch?v=fczgaWdXr-E>

Los valores normalizados son:

Distancia entre Usuario 4 y Usuario 1

$$V_1(\text{usuario1}) = (6,8,4)$$

$$V_4(\text{usuario4}) = (10,7,8)$$

$$Dst_n = \frac{5,74-0}{0-10} = 0,57$$

Distancia entre Usuario 4 y Usuario 2

$$V_2(\text{usuario2}) = (10,6,0)$$

$$V_4(\text{usuario4}) = (10,7,8)$$

$$Dst_n = \frac{8,06-0}{0-10} = 0,80$$

Distancia entre Usuario 4 y Usuario 3

$$V_3(\text{usuario3}) = (9,0,10)$$

$$V_4(\text{usuario4}) = (10,7,8)$$

$$Dst_n = \frac{7,34-0}{0-10} = 0,73$$

Utilizamos las distancias






En la siguiente fase se usarán las distancias obtenidas en cada usuario para multiplicar con las películas que no se tenían en común con el usuario al que queremos recomendar.

Como se observa en la imagen inferior, tenemos estas dos películas

| |  |  |  |  |  |
|---|---|---|---|---|--|
|  | 9 | 6 | 8 | 4 | × |
|  | 2 | 10 | 6 | × | 8 |
|  | 5 | 9 | × | 10 | 7 |
|  | ? | 10 | 7 | 8 | ? |

Elaboración propia en local

Se realizan las multiplicaciones y obtenemos los resultados para cada producto. Como se puede observar no importa si no se tienen todos los valores completos.

| |  |  | | |
|---|---|---|---|------|
|  | 9 | | X | 0,57 |
|  | 2 | 8 | X | 0,80 |
|  | 5 | 7 | X | 0,73 |

| | Superman Vs Batman | Batman Begins |
|--|--------------------|---------------|
| | 5,13 | |
| | 1,6 | 6,4 |
| | 3,65 | 5,11 |

Elaboración propia en local

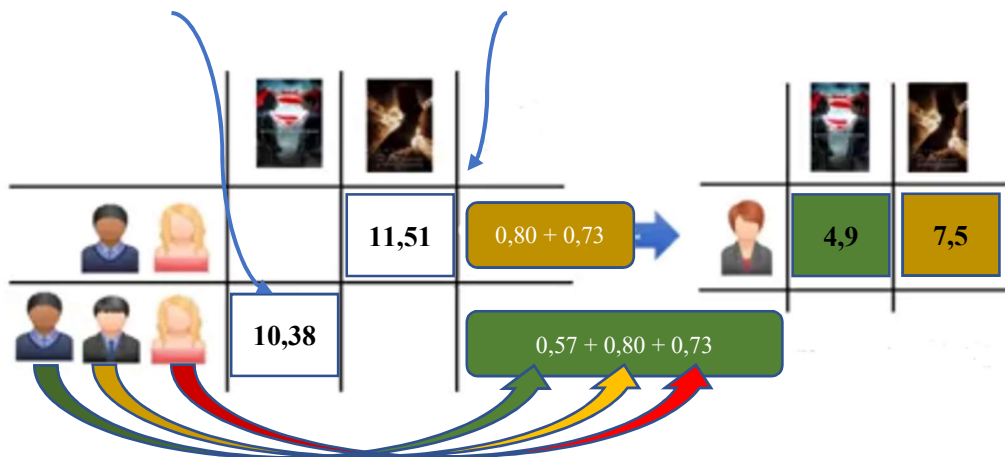
Obteniendo los valores de recomendación

Para terminar la actividad se debe obtener los valores finales de recomendación para el usuario objetivo. Para ello se realizarán los siguientes pasos.

- Sumar las cantidades obtenidas en el punto anterior por producto.
- Multiplicar nuevamente las distancias, pero en este caso en función a la suma de los productos y los usuarios relacionados.

Elaboración propia en local

| Superman Vs Batman | Batman Begins |
|--------------------|---------------|
| 5,13 | |
| 1,6 | 6,4 |
| 3,65 | 5,11 |



Elaboración propia en local

Explicación numérica de las operaciones

$$5,13 + 1,6 + 3,65 = 10,38 \text{ (valor grupal del producto)}$$

$$6,4 + 5,11 = 11,51 \text{ (valor grupal del producto)}$$

$$\frac{10,38}{0,80+0,73} = 7,5 \text{ Valor final de recomendación}$$

$$\frac{11,51}{0,75+ 0,80+0,73} = 4,9 \text{ Valor final de recomendación}$$

Anexo 3: Creación de actividad para un sistema de recomendación basado en Contenidos desde las matemáticas.

En esta segunda actividad se trabajará el otro sistema de recomendación, conocido como “Filtro basado en contenidos”. Como se ha explicado anteriormente es un algoritmo centrado más en el producto que en el usuario, como hemos desarrollado en la anterior actividad.

En este caso se verá desde un punto de vista matemático, entendiendo su funcionamiento y significado funcional.

Escenario de trabajo

En este caso se tiene un único usuario conociendo ya unas recomendaciones hacia unos productos. Por el contrario, tenemos otros tres productos los cuales se desconoce su grado de recomendación, el cual es el objetivo de la actividad.



Elaboración propia en local

Estructurar los datos

Para empezar con el análisis y las operaciones lo primero que se necesita es que los productos estén categorizados de alguna manera para que permita cierta distinción entre ellos en base a la recomendación del usuario.

En este ejemplo que se está usando se puede observar como las películas están categorizadas por diferentes géneros.

- Aventuras
- Comedia
- Ciencia Ficción
-

Esto permite realizar la tabla que se muestra en la imagen inferior.
Colocando un 1 donde esa película esta categorizada y un cero donde no lo está.

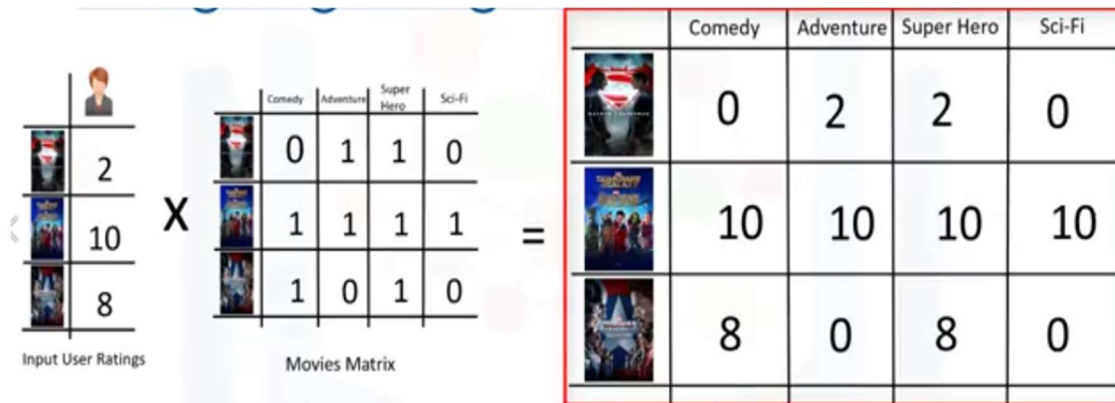
| |  | | Comedy | Adventure | Super Hero | Sci-Fi |
|---|---|---|--------|-----------|------------|--------|
|  | 2 |  | 0 | 1 | 1 | 0 |
|  | 10 |  | 1 | 1 | 1 | 1 |
|  | 8 |  | 1 | 0 | 1 | 0 |
| Input User Ratings | | | | | | |

Elaboración propia en local

En este primer paso únicamente nos centramos en las películas que tenemos evaluadas/recomendadas por el usuario.

Operaciones matemáticas

- Se realiza la multiplicación de los géneros por los valores de recomendación del usuario, para ponderarlos.
- Posteriormente se hace una suma por columnas de los géneros.
- Finalmente se hace una normalización de esos datos obtenidos en la suma para que los valores vuelvan a estar entre 0 y 1 como originalmente se había partido.



NORMALIZAR los datos
 $X_n = (X - X_{\min}) / (X_{\max} - X_{\min})$

18 12 20 10

$X_{\max} = 60$ (todas las películas del dataset, 10x 6 películas)
 $X_{\min} = 0$

Elaboración propia en local

¿Por qué debemos normalizar?

Normalizar los datos es una técnica que se utiliza para ajustar las características de los datos a un rango común. Esto puede ser útil por varias razones:

1. **Evitar sesgos de escala:** Algunos algoritmos de machine learning son sensibles a la escala de los datos de entrada. Por ejemplo, si tenemos una columna con el número de «clicks» que hace el usuario dentro de Netflix y otra con el número de accesos que tiene ese usuario al día, es evidente que la primera columna tendrá valores mucho más elevados que la segunda y no por esto debe ser más importante o relevante. La normalización ayuda a evitar este problema al ajustar todas las características a un rango común.
2. **Facilitar la comparación:** La normalización también puede facilitar la comparación de características con diferentes escalas. Si dos características tienen diferentes unidades o escalas, es difícil compararlas directamente. La normalización las ajusta a una escala común, lo que facilita la comparación.
3. **Mejorar la eficiencia:** Algunos algoritmos, como los de clasificación, son más eficientes cuando las características están normalizadas. Si una característica tiene una escala muy alta, puede dominar a las demás y hacer que el algoritmo sea menos eficiente. La normalización ayuda a evitar este problema.

¿Qué es normalizar?

<https://youtu.be/fczgaWdXr-E>

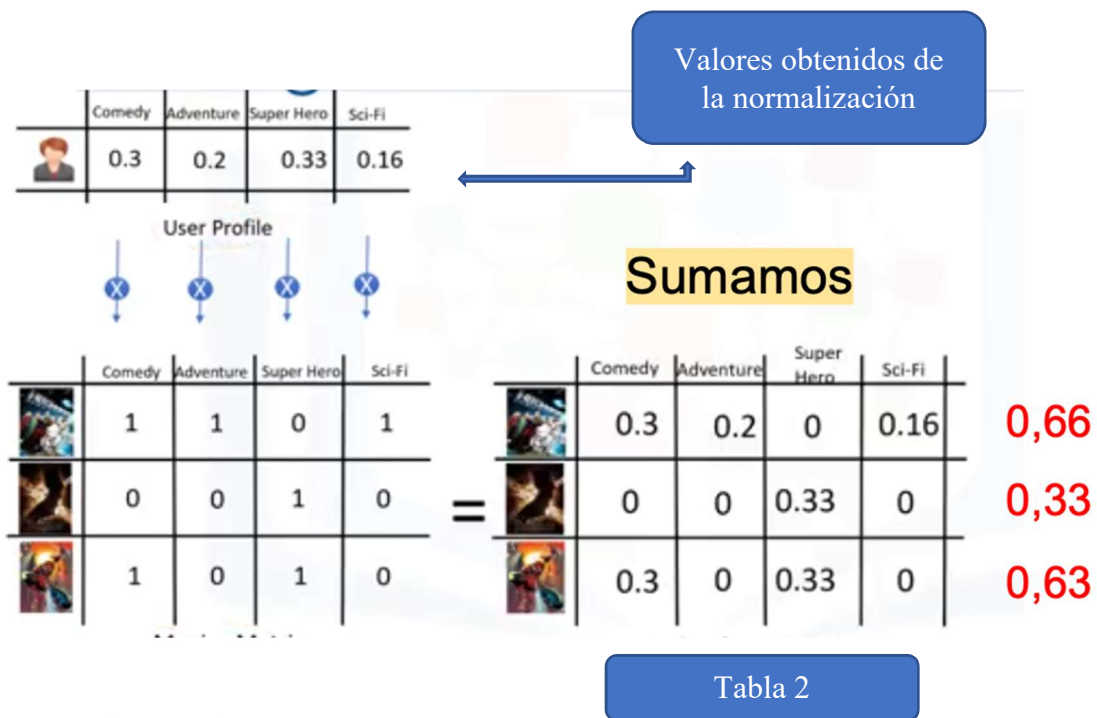
$$\frac{X - X_{min}}{X_{max} - X_{min}} = \frac{18 - 0}{60 - 0} = 0,3$$

$$\frac{X - X_{min}}{X_{max} - X_{min}} = \frac{12 - 0}{60 - 0} = 0,2$$

$$\frac{X - X_{min}}{X_{max} - X_{min}} = \frac{20 - 0}{60 - 0} = 0,33$$

$$\frac{X - X_{min}}{X_{max} - X_{min}} = \frac{10 - 0}{60 - 0} = 0,16$$

Obtener valores de recomendación



Elaboración propia en local

Una vez obtenidos los valores normalizados se usarán con los productos los cuales se quiere recomendar. En este caso multiplicamos esos valores por lo géneros y obtenemos la tabla 2, como se puede observar en la imagen superior.

Finalmente, para acabar la actividad **se suman las columnas para obtener el valor total del producto** y así poder recomendarlo como se muestra en la imagen inferior.

Final de la actividad



Elaboración propia en local

Anexo 4: Creación de actividad para un sistema de recomendación basado en Contenidos desde la programación con Python.

Para la realización de esta actividad se necesita un mínimo de competencia en el uso de Jupyter Notebook. Este proyecto lo realizaremos con la herramienta Google Colaboratory.

Google Colaboratory es un servicio gratuito de Google que permite a los usuarios ejecutar código en vivo en línea en un entorno de Jupyter Notebook. Es útil para la exploración y el aprendizaje de datos, el desarrollo y prueba de código, y muchas otras aplicaciones.

Un Jupyter Notebook es una aplicación web que permite a los usuarios crear y compartir documentos que contienen código, ecuaciones, visualizaciones y texto explicativo. Los notebooks son muy útiles para la presentación de ideas y el análisis de datos porque permiten combinar código, visualizaciones y texto en un mismo documento.

Google Colaboratory proporciona un entorno de Jupyter Notebook en la nube, lo que significa que no es necesario instalar ningún software en su computadora. Además, Colaboratory proporciona recursos de procesamiento y almacenamiento gratuitos, lo que lo convierte en una herramienta muy útil para el aprendizaje y la exploración de datos.

Para utilizar Google Colaboratory, es necesario tener una cuenta de Google y acceder a Colaboratory a través de Google Drive. Una vez que se ha accedido, se pueden crear nuevos notebooks o abrir notebooks existentes. Los notebooks se pueden compartir fácilmente con otros usuarios de Colaboratory y se pueden guardar en Google Drive para su uso posterior.

En resumen, Google Colaboratory es una herramienta útil para la exploración y el análisis de datos, el aprendizaje de código y el desarrollo de proyectos en línea. Es fácil de usar y proporciona recursos de procesamiento y almacenamiento gratuitos, lo que lo convierte en una excelente opción para muchas aplicaciones.

Video tutorial.

https://www.youtube.com/watch?v=8VFYs3Ot_aA&list=RDCMUCTGfq71XODsBLtU3Ci--g-g&start_radio=1

PROYECTO TV-TÚ

Sistemas de recomendación basados en contenidos

Programación Python

Objetivo

Crear un sistema de recomendación que pueda ofrecer a nuestros clientes de TV-Tú un conjunto de películas en función a las puntuaciones que ya tenemos en nuestra basa de datos.

El algoritmo se basará en un mínimo de 3 películas y un máximo de 10 que haya visto el usuario. Deberá recomendar un mínimo de 50 películas y un máximo de 250.

Dataset de trabajo

Para crear este sistema de Recomendación, usaremos un *Dataset* oficial de IDBM con 1000 películas. (Adjunto también en la documentación)

<http://educa.europeanvalley.es/tools/datasets/Movies.csv>

Columns

- *Rank*: Movie rank order.
- *Title*: The title of the film.
- *Description*: Brief one-sentence movie summary.
- *Director*: The name of the film's director.
- *Actors*: A comma-separated list of the main stars of the film.
- *Year*: The year that the film released as an integer.
- *Runtime (Minutes)*: The duration of the film in minutes.
- *Rating*: User rating for the movie 0-10.
- *Votes*: Number of votes.
- *Revenue (Millions)*: Movie revenue in millions.
- *Metascore*: An aggregated average of critic scores. Values are between 0 and 100. Higher scores represent positive reviews.
- *Genre*: 1-2-3 list of genres used to classify the film.

Videos explicativos del Algoritmo de Recomendación Basado en Contenidos

- Parte1 <https://youtu.be/AB6i8RSFaoM>
- ¿Qué es normalizar? <https://youtu.be/fczgaWdXr-E>
- Parte 2 <https://youtu.be/-XyjWGa82eA>

El notebook de la plantilla se adjunta en el Anexo 5.

[Anexo 6: Reflexión y estudio de los aspectos éticos de un sistema de recomendación.](#)

Introducción general

La ética de la inteligencia artificial es un tema de gran importancia en la actualidad, ya que la IA está cada vez más presente en nuestras vidas y puede tener un impacto significativo en la sociedad y en el individuo. Algunos de los principales temas que se deben abordar en una situación de aprendizaje sobre la ética de la IA son:

1. **Qué es la IA y cómo funciona:** Es importante que los estudiantes tengan una comprensión básica de qué es la IA y cómo funciona, para poder entender los desafíos éticos que plantea.
2. **Desafíos éticos en el desarrollo de la IA:** Algunos de los principales desafíos éticos en el desarrollo de la IA incluyen el riesgo de discriminación, el uso indebido de datos personales, el impacto en el empleo y el potencial de aumentar la brecha entre ricos y pobres.
3. **Responsabilidad y regulación de la IA:** ¿Quién es responsable de asegurar que la IA se utiliza de manera ética y responsable? ¿Cómo deben regularse los sistemas de IA?
4. **Usos beneficiosos de la IA:** A pesar de los desafíos éticos, la IA también puede tener usos beneficiosos, como mejorar la eficiencia, ayudar en la toma de decisiones y facilitar el acceso a la información.
5. **Consideraciones éticas en el diseño de sistemas de IA:** Al diseñar sistemas de IA, es importante considerar cómo pueden afectar a los individuos y a la sociedad en su conjunto y asegurar que se utilizan de manera ética y responsable.

Preguntas para elaborar y reflexionar en el aula

1. ¿Cómo podemos asegurar que los sistemas de recomendación sean éticamente responsables?
2. ¿Cómo los desarrolladores de IA pueden prevenir la discriminación y los sesgos en la recomendación de productos?
3. ¿Qué medidas de seguridad y protección de datos deben implementarse para garantizar la privacidad y la seguridad de los usuarios?
4. ¿Cómo podemos garantizar que los sistemas de recomendación no sean utilizados para fines inmorales o ilegales?
5. ¿Cómo podemos garantizar que los sistemas de recomendación no sean utilizados para fines de manipulación?