

PROYECTO FINAL

Sistema de recomendación de películas

Proyecto final Data Science

Alumno:
Albert Vila



Sistema de recomendación de películas

Un proceso más eficiente y personalizado para fidelizar los usuarios de las plataformas de streaming

< Índice >

Introducción

1. Personalización
2. Navegación eficiente
3. Aumento de retención de usuarios
4. Incremento de ventas y utilización
5. Innovación y competitividad

Importación de librerías

1. Introducción
2. Metodología
 - 2.1. Preprocesamiento de datos
 - 2.2. Exploración de datos
 - 2.3. Filtrado colaborativo
3. Resultados
 - 3.1. Modelo de recomendación
 - 3.2. Filtrado colaborativo
4. Discusión
5. Conclusiones
6. Referencias



Sistema de recomendación de películas

Un proceso más eficiente y personalizado para fidelizar los usuarios de las plataformas de streaming

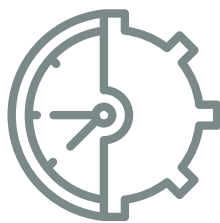
< Necesidades clave >

En la era digital actual, la cantidad de información disponible en línea es abrumadora y continúa creciendo a un ritmo exponencial. Ya sea en plataformas de streaming, tiendas online o redes sociales. Los usuarios se encuentran ante una vasta cantidad de opciones y, por lo tanto, tomar decisiones informadas sobre qué película o serie ver, se vuelve cada vez más desafiante. Aquí es donde los sistemas de recomendación se vuelven esenciales, respondiendo a varias necesidades clave:



1. Personalización:

- Contexto del Usuario: Los sistemas de recomendación permiten que las plataformas ofrezcan experiencias personalizadas a cada usuario, adaptando su contenido según las preferencias y comportamientos individuales.
- Relevancia: Aseguran que las opciones presentadas al usuario sean relevantes y de su interés, mejorando así su experiencia de usuario y satisfacción.



2. Navegación Eficiente:

- Facilitar Decisiones: Ayudan a los usuarios a navegar a través de la amplia gama de opciones disponibles, proporcionándoles selecciones que probablemente encuentren atractivas.
- Ahorro de Tiempo: Reducen el tiempo que los usuarios necesitan para buscar y decidir sobre un producto o servicio, filtrando y destacando las opciones más pertinentes.

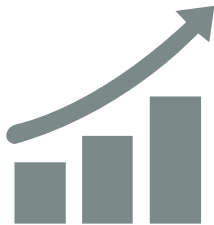


3. Aumento de la Retención de Usuarios:

- Engagement: Al ofrecer recomendaciones personalizadas, los usuarios son más propensos a interactuar y pasar más tiempo en la plataforma.
- Lealtad: Los sistemas de recomendación mejoran la lealtad del usuario al hacer que se sientan valorados y comprendidos en sus preferencias y necesidades.

Sistema de recomendación de películas

Un proceso más eficiente y personalizado para fidelizar a los usuarios de las plataformas de streaming



4. Incremento en Ventas y Utilización:

- Descubrimiento: Facilitan el descubrimiento de nuevos contenidos que los usuarios podrían no haber encontrado por sí mismos.
- Conversión: Mejoran las tasas de conversión al dirigir a los usuarios hacia contenidos o productos que son más propensos a comprar o consumir.



5. Innovación y Competitividad:

- Innovación: Los sistemas de recomendación son una herramienta de innovación, permitiendo que las plataformas ofrezcan nuevas y mejoradas experiencias a sus usuarios.
- Ventaja Competitiva: En mercados saturados, tener un sistema de recomendación eficiente puede diferenciar a una plataforma de sus competidores, proporcionando una ventaja en términos de retención de usuarios y satisfacción del cliente.

Los sistemas de recomendación, por lo tanto, no solo enriquecen la experiencia del usuario, sino que también es una herramienta estratégica para las empresas y plataformas online, permitiéndoles maximizar su engagement, optimizar sus inventarios y mejorar estrategias de marketing. Lo que finalmente se traduce en mayores ingresos y una mejor posición en el mercado.



Sistema de recomendación de películas

Un proceso más eficiente y personalizado para fidelizar los usuarios de las plataformas de streaming

<Importación de Librerías>

Este proyecto desarrolla un sistema de recomendación de películas utilizando técnicas de aprendizaje automático y filtrado colaborativo. A través del análisis y manipulación de un conjunto de datos y de la implementación de modelos de factorización matricial (como SVD). Con esto buscaba predecir las calificaciones de los usuarios hacia las películas, para hacer buenas recomendaciones que satisfagan al usuario.

1. Introducción:

El objetivo de un sistema de recomendación es sugerir productos, servicios o información relevantes a los usuarios basándose en sus preferencias y patrones. En este proyecto, se exploran películas, buscando recomendar títulos que puedan ser de interés para los usuarios **basándose en sus calificaciones previas**. Tal y como funciona el mundo del streaming, cada vez más al alza, y con nuevas plataformas que se suman a la vasta oferta ya existente. Un buen sistema de recomendación es esencial, para agilizar y potenciar la experiencia del usuario, y destacar ante la competencia.

2. Metodología:

2.1. Preprocesamiento de Datos:

Cargué cuatro conjuntos de datos: **movies_metada, credits, ratings y links** descargados desde Kaggle y que proviene de **The Movies Dataset**, una base de datos de 45000 películas y 26 millones de puntuaciones de aproximadamente 270.00 usuarios que contiene información sobre películas y calificaciones dadas por los usuarios. Llevé a cabo la limpieza y transformación de los datos, eliminando duplicados, gestionando valores nulos y convirtiendo tipos de datos para que pudieran ser utilizados.

2.2. Exploración de Datos:

Exploré las características de las películas y las calificaciones, verificando la distribución de las calificaciones, la cantidad de películas por género, y la cantidad de calificaciones por usuario, entre otros.

2.3. Filtrado Colaborativo:

Este método emplea las calificaciones de los usuarios para predecir cómo un usuario podría calificar una película que aún no ha calificado. Usé el algoritmo SVD para la factorización de la matriz de usuario-ítem y realicé una validación cruzada para evaluar el rendimiento del modelo en términos de RMSE y MAE.

Sistema de recomendación de películas

Un proceso más eficiente y personalizado para fidelizar los usuarios de las plataformas de streaming

3. Resultados:

3.1. Modelo de Recomendación basado en Contenido:

El modelo de recomendación basado en contenido se centra en la similitud de los ítems, haciendo recomendaciones de películas que son similares en términos de género, director y actores. Hice uso del método `TfidfVectorizer` para convertir la información textual a vectores y `linear_kernel` para calcular las similitudes del coseno entre las películas. Los resultados fueron bastante buenos en términos de recomendación.

3.2. Filtrado Colaborativo:

Logré entrenar y validar el modelo SVD. El modelo fue capaz de predecir calificaciones con un RMSE de aproximadamente 0.906 y un MAE de 0.7019, lo que indica diferencia moderada entre las calificaciones predichas y las reales. Aquí los resultados fueron correctos, aunque el modelo tiene margen de mejora.

4. Conclusión

El proyecto ha demostrado cómo implementar un sistema de recomendación utilizando técnicas de aprendizaje automático. El modelo basado en contenido proporciona recomendaciones razonables basadas en la similitud del contenido, mientras que el modelo de filtrado colaborativo es capaz de predecir calificaciones de usuario, pero la similitud entre los resultados plantea dudas.

El RMSE y MAE obtenidos en el modelo de filtrado colaborativo indican que, aunque el modelo tiene un desempeño decente, aún hay margen amplio de mejora. He ajustado los hiperparámetros para mejorar el modelo, pero no he obtenido una mejora sustancial. Quizás usar un conjunto de datos más grande, o probar con diferentes algoritmos, podrían ayudar a mejorar las predicciones.

Sistema de recomendación de películas

Un proceso más eficiente y personalizado para fidelizar los usuarios de las plataformas de streaming

6. Referencias:

- <https://www.kaggle.com/datasets/rounakbanik/the-movies-dataset>
- <https://www.kaggle.com/code/ibtesama/getting-started-with-a-movie-recommendation-system>
- <https://www.aprendemachinellearning.com/sistemas-de-recomendacion/>
- <https://hescaso.github.io/recomendador/>