

CASO DE ESTUDIO DE MARKETING

Piedrahita Allison; Ramírez Anyi; Vergara María

Departamento de Ingeniería Industrial, Universidad de Antioquia.

Medellín, Colombia.

RESUMEN ~ Los sistemas de recomendación han transformado por completo la forma en que interactuamos con plataformas en línea. Estas herramientas tecnológicas no solo facilitan nuestras decisiones de compra, ya que *“las recomendaciones de terceros son una opción a la que acuden los individuos para tratar de solucionar los problemas derivados de su limitada información y conocimiento respecto a categorías o alternativas de bienes, servicios e información”* (Resnick & Varian, 1997, como se citó en Cruz et al., en 2010) [2]. En este caso de estudio, se plantea una solución analítica para mejorar la experiencia de los usuarios en una plataforma de películas, usando recomendaciones basada en sus gustos o preferencias ofreciendo de forma periódica contenido de interés. Para ello, se propone el uso de varios sistemas de recomendación según ciertos criterios predefinidos.

1. INTRODUCCIÓN

Una plataforma online quiere tener una solución que le permita hacer recomendaciones de películas a sus usuarios, con el objetivo de que estos tengan una mejor experiencia, y esto permita mejorar su fidelización y recomendación a nuevos clientes. Los sistemas inteligentes de recomendación han sido la respuesta a estas necesidades dado su capacidad para analizar grandes volúmenes de datos se ha convertido en una herramienta con mucho poder para predecir. A continuación, se presentan algunas ventajas de estos sistemas

- Experiencia personalizada, donde se reciben recomendaciones basadas en gustos y preferencias de los usuarios.
- Los usuarios descubren contenido de su interés de forma rápida, además incluye opciones que posiblemente los usuarios no habían considerado
- Mayor satisfacción en los usuarios la encontrar algo que estaban buscando o sugerencias nuevas.

“La gran cantidad de sitios especializados en Internet, ofertando millones de productos y/o servicios para su consumo, se convierte en un caos de información sin solución cuando se necesita realizar una adquisición eligiendo entre todas las opciones existentes”. (García & Gil, s.f) [1]. En la Era de la digitalización las opciones en plataformas han transformado la experiencia de la adquisición de bienes y servicios, tanto así que los consumidores se les dificulta tomar decisiones pues encuentran un sin número de posibilidades, es ahí cuando los sistemas de recomendación emergen, brindando una guía personalizada de acuerdo con el consumo y gusto de los usuarios

Teniendo en cuenta lo anterior, y para este caso en particular se analizan bases de datos que contienen información acerca de las películas para proporcionar una solución analítica, que facilite el establecimiento de planes de mejora que le permita a los usuarios diversas recomendaciones basado en la personalización de la cinta de contenido que tiene la plataforma. De esta forma la empresa puede fidelizar a sus clientes en un largo plazo.

2. METODOLOGÍA

En las siguientes secciones se describe cada uno de los pasos desarrollados.

2.1.DISEÑO DE LA SOLUCIÓN

Se plantean los siguientes objetivos para desarrollar el caso de estudio:

- Generar recomendaciones oportunas a los usuarios de la plataforma de películas

Con base a lo anterior, se plantea la solución analítica de *Ilustración 1*, en la cual se evidencian el objetivo anterior.

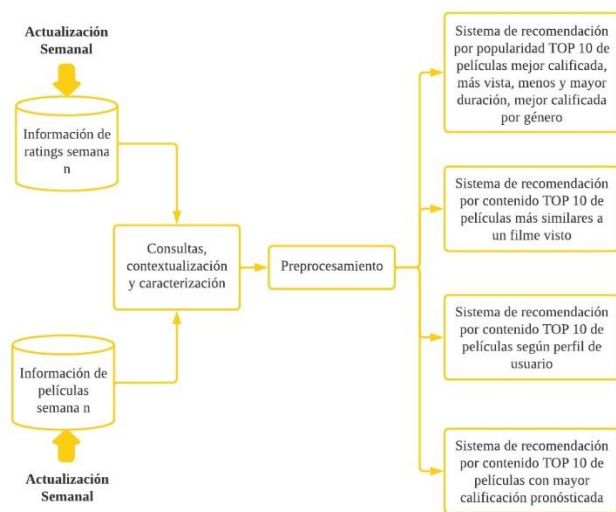


Ilustración 1. Diseño de la solución analítica.

2.2.LIMPIEZA Y TRANSFORMACIÓN

Se cuenta con una base de datos, la cual cuenta con dos tablas:

- “*movies*”: Contiene la información del catálogo de películas disponibles en la plataforma.
- “*movies_ranting*”: Contiene una lista de los usuarios y las películas que vieron, el tiempo de duración y la calificación que le dieron a la película

Las tablas de datos no presentan valores faltantes ni datos nulos, sin embargo, al visualizar la tabla

de “*movies*” se puede ver que hay películas con varios géneros en una misma variable, lo que podría ocasionar poca interpretabilidad de esta y a partir de ella se genera una tabla nueva con todos los géneros volviendo cada uno de ellos una variable tipo “*dummie*”

2.3.ANÁLISIS EXPLORATORIO

Después de realizar un análisis para comprender el contexto de la plataforma y sus películas, se observa que:

- La plataforma tiene un total de 9.742 diferentes películas.
- La cantidad de evaluaciones hechas por usuarios en películas es 100.854.
- El total de usuarios que han calificado películas son 611.
- Existe un total de 18 películas sin una sola evaluación.
- Hay un total de 3.459 películas que solo han sido calificadas una vez.
- La plataforma tiene un total de 20 géneros distintos para las películas.
- El género con menor número de películas es de tipo films noir, la cual tiene un total de 87.
- El género con mayor número de películas es drama, la cual acumula un total de 4361.
- La mayoría de las películas tienen calificaciones entre los tres y los cuatros puntos.
- En promedio los usuarios han calificado 165 veces películas, además que el 75% de los usuarios han calificado 168 veces o menos.
- En promedio una película puede ser calificada hasta 10 veces.

2.4. SELECCIÓN DE FILTROS PARA QUITAR VALORES EXTREMOS

En primer lugar, se observa el número de calificaciones hecha por los usuarios, presentado en el siguiente grafico

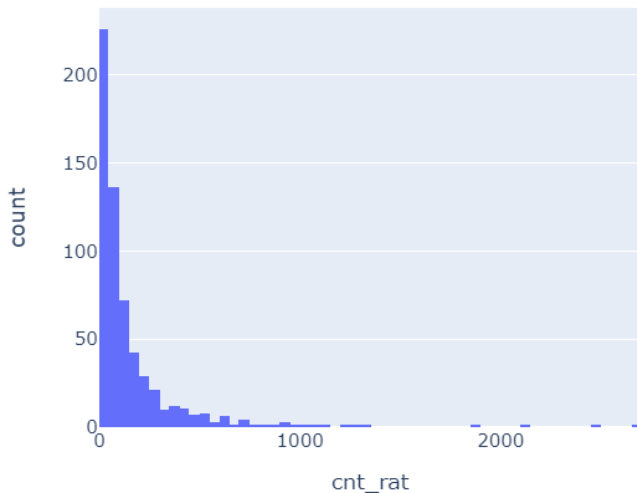


Ilustración 2. Número de calificaciones por usuarios

Se puede observar que la gran mayoría de los usuarios no califican las películas vistas y por otro lado hay quienes han calificado más de 2.000 películas, lo que nos deja muchos datos extremos y con esto se evidencia que el 75% de los usuarios han calificado 168 veces o menos.

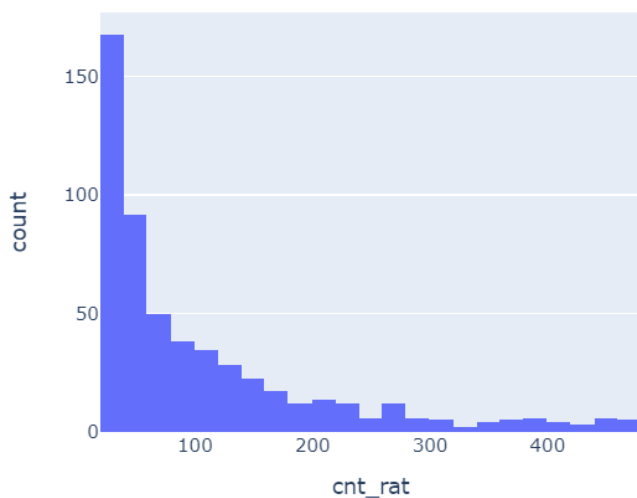


Ilustración 3. Número de calificaciones por usuarios (FILTRADA)

Para quitar los valores extremos que podrían dar recomendaciones errores o sesgadas, se

recomienda tomar los usuarios que han calificado más de 20 películas y menos de 480 películas.

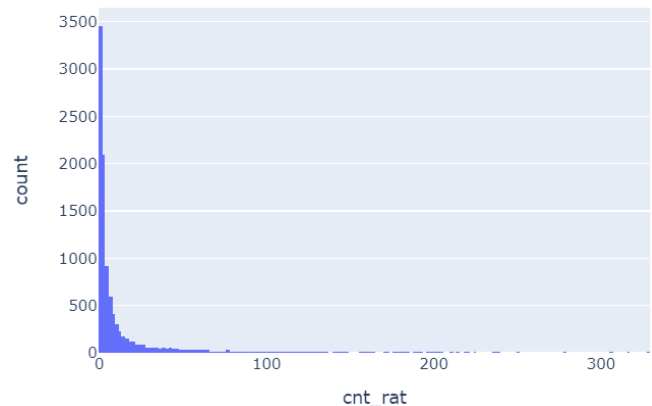


Ilustración 4. Número de calificaciones por película

Cuando la distribución de los datos se encuentra muy al extremo izquierdo de la distribución, y unos pocos a la derecha, es gráficamente se observa que la distribución no es uniforme, la idea es organizar los filtros para que la base muestre información más real.

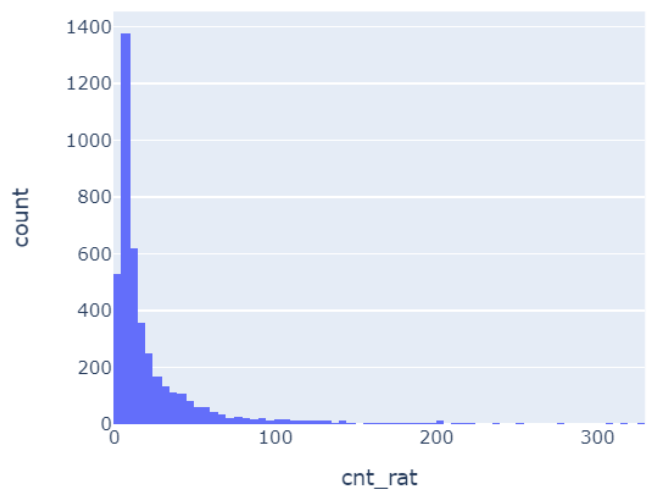


Ilustración 5. Número de calificaciones por película (FILTRADA)

Acá se filtraron las películas que tuvieran más de tres calificaciones debido a que en la *Ilustración 6* se podía identificar que el 50% de las películas tenían al menos 3 calificaciones.

Entonces, los filtros que se tuvieron en cuenta están ligados a la cantidad de usuarios y los puntajes de las películas generando una base de

datos más sólida, de esta forma los usuarios también pueden diferir menos entre ellos.

2.5.SISTEMAS DE RECOMENDACIÓN (MODELOS)

Con las bases de datos filtradas, se llevan a cabo los sistemas de recomendación propuestos en el diseño de la solución, como se muestra a continuación:

2.5.1. SISTEMA DE RECOMENDACIÓN BASADO EN POPULARIDAD

se construye un conjunto de tablas en donde se recomiendan películas de manera generalizada según diferentes parámetros. En primer lugar, se recomienda al usuario las 10 películas mejor calificadas, posteriormente se recomiendan aquellas películas más vistas independientemente de su calificación, de la misma manera se construye una lista con las películas más cortas y largas con el fin de que el usuario pueda seleccionar según sus preferencias de duración y finalmente, se recomiendan las películas mejores calificadas por género, para que el usuario pueda seleccionar según sus gustos. Las recomendaciones realizadas se realizan por medio de consultas SQL.

2.5.2. SISTEMA DE RECOMENDACIÓN BASADO EN UN SOLO PRODUCTO

teniendo en cuenta una película vista por el usuario, se recomienda de manera generalizada las 10 películas más parecidas de acuerdo con dos criterios de similitud. El primero, realiza la recomendación por medio de la correlación existente entre las características de géneros de las películas mientras que el segundo realiza la recomendación haciendo uso del algoritmo de los K vecinos más cercanos, los cuales son determinados por medio de la métrica coseno.

2.5.3. SISTEMA DE RECOMENDACIÓN BASADO EN EL PERFIL DEL USUARIO

al igual que en el sistema anterior, se hace uso de los K vecinos más cercanos para evaluar la similitud de las películas que ha visto el usuario con respecto a aquellas que aún no ha visualizado. Este sistema es personalizado, debido a que se recomiendan 10 películas de acuerdo con los filmes vistos por cada usuario.

2.5.4. SISTEMA DE RECOMENDACIÓN BASADO EN FILTROS COLABORATIVOS

al igual que en el sistema anterior, se establece una recomendación de las 10 películas que podrían ser mejor calificadas por el usuario de acuerdo con las calificaciones otorgadas por usuarios con perfiles similares. Este sistema, tiene en cuenta cuatro modelos de K vecinos más cercanos (KNNBasic, KNNWithMeans, KNNWithZScore, KNNBaseline), los cuales son evaluados a partir del MAE y el RMSE, donde este último es tomado para seleccionar el modelo con mejor desempeño. Finalmente, se establece que el modelo KNNWithMeans presenta métricas óptimas para realizar la recomendación, por lo tanto, se hace un afinamiento de hiperparámetros para mejorar el modelo de recomendación y se implementa para recomendar al usuario aquellos filmes que no ha visto pero que podrían obtener una buena calificación

2.6.DESPLIEGUE DEL MODELO

Luego de tener unas bases filtradas y para la implementación de los sistemas de recomendación la recopilación de la información de dará de la siguiente manera:

- Para los modelos basados en popularidad cada 7 días y a las 6:00AM se tomará una nueva

información de los usuarios y se actualizará en la plataforma desde la pantalla de inicio.

- Cada 15 días, se evaluará cuáles son las nuevas películas agregadas y se muestran cuáles son las películas más parecidas. De esta forma cada vez que la persona termine una película se le recomienden 10 películas nuevas no vistas, información quedará en un menú desplegable de la plataforma.
- La aplicación recopila semanalmente las películas vistas por los usuarios, actualiza sus perfiles y crea una tabla de recomendaciones basada en estas preferencias. La tabla se guarda en un menú desplegable de la plataforma para personalizar las recomendaciones para cada usuario.

3. CONCLUSIONES

De los resultados obtenidos, se evidencia que:

- Los sistemas de recomendaciones les permiten a las empresas adaptar sus productos a las necesidades y preferencias del cliente aumentando el consumo y la satisfacción de este.
- Construir cualquier tipo de sistema de recomendación requiere de una limpieza adecuada de las bases de datos, además de criterios técnicos para establecer tablas confiables que representen el comportamiento de los usuarios, con el fin de obtener mejores resultados.
- Entender las preferencias y acciones de los usuarios puede enriquecer la experiencia del usuario, haciendo que esta sea más pertinente y atractiva gracias a las recomendaciones.
- *Link del proyecto en GitHub*
<https://github.com/Mariavergaraz/analitica3proyecto2>

4. REFERENCIAS

[1]. García, F. J., & Gil, A. B. *Personalización de Sistemas de Recomendación*. Salamanca: Universidad de Salamanca.

https://www.researchgate.net/profile/Francisco-Garcia-Penalvo/publication/268344600_Personalizacion_de_Sistemas_de_Recomendacion/links/54d0ab3e0cf20323c21888ad/Personalizacion-de-Sistemas-de-Recomendacion.pdf

[2]. Cruz, C. C. C., López, F. J. M., & Moreno, V. M. (2010). *Sistemas de recomendación en el Comercio Electrónico y la E-educación*. *Criterio Libre*, 8(12), 161-182.

<https://revistas.unilibre.edu.co/index.php/criteriolibre/article/view/1264/984>