

INFORME CASO DE ESTUDIO MARKETING

ELABORADO POR:

ANDRÉS FELIPE PENNA HERNÁNDEZ

JHONATAN VALENCIA OCAMPO

STEVEN ZAPATA ZULETA

PROFESOR:

JUAN CAMILO ESPAÑA LOPERA

ANALÍTICA



UNIVERSIDAD DE ANTIOQUIA

FACULTAD DE INGENIERÍA

DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

## **a. Diseño de solución propuesto.**

### Problema de negocio

La plataforma online requiere recomendaciones de su contenido (películas), estas le permiten crear una experiencia atractiva y satisfactoria al usuario, puesto que puede hallar películas que se adapten a su preferencia de consumo o similares, generando contenido de alta calidad. Para ello se requiere efectuar modelos de recomendación filtrados por popularidad, contenido basado en un producto (año de publicación, género), colaborativo y por preferencia del usuario.

Si bien, no se lograra crear un sistema de recomendación que impacten al usuario, debido a que las películas enunciadas no son relevantes de acuerdo a su perfil, es posible que este usuario pierda el interés por consumir dicho servicio en la plataforma, en consecuencia, la plataforma online perdería posición en el mercado a causa de la disminución en la retención de usuarios. Por tal razón, la aplicación de un sistema de recomendaciones efectivo originaría elementos de gran interés al usuario y con ello una ventaja competitiva respecto a la competencia, aumentando el goodwill de la plataforma.

### Problema analítico

Se establecen los siguientes apartados de analítica, estos son:

1. Definir las variables relevantes que permitan diseñar un sistema de recomendación que proporcione elementos característicos de las películas que se adapten al usuario.
2. Establecer categorizaciones en los sistemas de recomendaciones. (popularidad, contenido basado en un producto (año de publicación, género), colaborativo y por preferencia del usuario)
3. Realizar un modelo de recomendaciones que evalúe las características y preferencias de cada usuario según su perfil y que se adapte de acuerdo al entorno cambiante.

### Marco

De acuerdo con las necesidades expuestas por la plataforma online de películas de requerir un sistema de recomendaciones que fidelice a sus usuarios a partir de mejores experiencias, se establecen diferentes maneras de recomendar al usuario dichas películas, como se mencionó anteriormente.

Según el artículo *“Hybrid recommender systems: Survey and experiments. User Modeling and User-Adapted Interaction”* esta propuesta busca que la plataforma online utilice algoritmos de recomendación que puedan procesar grandes cantidades de datos de manera eficiente y proporcionar recomendaciones en tiempo real. Con ello se impulsa a que los usuarios sigan consumiendo contenido de la plataforma al proporcionar recomendaciones que se adapten a sus gustos o preferencias particulares.

En el apartado de anexos se muestra el diagrama de procesos donde se plantea la interacción de la plataforma y el sistema de recomendaciones con el usuario.

Ahora bien, la interacción con los datos permite establecer criterios de elección y limpieza de datos de la siguiente manera:

Primero se cargan las bases de datos a Google colab, posteriormente se revisa dicho contenido, obteniendo las bases de datos *movies* y *ratings*. que enlistan las películas con sus nombres, fechas de publicación, género e interacción con usuarios por medio de calificaciones.

## **b. Análisis exploratorio**

En la exploración se realiza el cargue de las bases de datos Movies y Ratings por medio de SQL, se hace una revisión del contenido de cada una de ellas para establecer la forma en que se explora los datos, esto con el objetivo de crear relaciones entre los datos que sean representativas y desarrollar criterios de alto valor para la posterior etapa de creación de tablas que son determinantes en la factibilidad de los sistemas de recomendación.

Después de verificar que las bases de datos no tienen valores duplicados, se hace una consulta para visualizar la distribución de calificaciones que dieron los usuarios sobre las películas. Se puede observar que la tendencia de los usuarios es a dar valoraciones altas siendo un primer indicio positivo para la plataforma con respecto a contar con contenido de alta calidad que pueda mantener la fidelización de los clientes y generar el interés de nuevos clientes.

También se realiza una consulta que calcula la cantidad de valoraciones a películas que ha realizado cada usuario, es importante contar con esta información para saber si es factible tener sistemas de recomendación ya que estos depende de la actividad de los mismos usuarios para brindar calificaciones sobre el contenido consumido. Puntualmente, se halla que el usuario con menor cantidad de películas valoradas tiene un total de 20 y la persona con mayor cantidad de valoraciones tiene 2698. Se hace uso de gráficos que faciliten el análisis por medio de la visualización, dejando a la vista una discontinuidad en los datos a partir de 1055 valoraciones, adicionalmente, por medio de un describe se halla que el 75% de los usuarios tienen valoraciones por debajo de 168, corroborando la importancia de limitar los datos a usuarios que tengan valoraciones por debajo o igual a 1055 valoraciones al momento de crear las tablas que se usarán. Así, se evitará tener sesgos en los sistemas de recomendación por datos atípicos que pueden ser valoraciones falsas.

Finalmente, se hace una última consulta para verificar cuantas valoraciones tiene cada película, obteniendo un rango que tiene como mínimo una valoración y máximo 329 valoraciones. Según el describe el 50% de los datos está por debajo de 3 valoraciones. Para garantizar que las recomendaciones sean apropiadas y el usuario pueda acceder a contenido que realmente cumpla con sus expectativas se limita a que las películas a considerar deben contar por lo menos con 5 valoraciones. De esta manera se pasa de 9719 registros a 3651 registros, la pérdida de datos es considerable sin embargo esta decisión garantiza una mayor fiabilidad en los sistemas además estos serán actualizados con una periodicidad definida que permite agregar las películas que vayan cumpliendo el umbral mínimo, aumentando el volumen de contenido sin perder calidad en el servicio.

## **c. Selección de algoritmos y técnicas de modelado - Sistemas de recomendación**

Los sistemas de recomendación elegidos para ser aplicados para la plataforma online de acuerdo con los datos suministrados son:

- Basado en popularidad: este modelo de recomendación filtra las películas y propone que las películas más atractivas (con mejor calificación, cantidad de calificaciones o vistas) al usuario son las de mayor expectativa general o popularidad, para el caso de estudio se utilizan elementos como mejor rating por año y género. Este tipo de recomendaciones se limita en la carencia de adaptabilidad a las particularidades de determinado usuario. Debido a esta limitación, este sistema es ideal como primer elemento de acercamiento para un usuario que no ha interactuado con la plataforma (usuario nuevo), de esta manera es posible reconocer características de preferencia de acuerdo con su primera elección.
- Contenido basado en un producto: una vez se establece la preferencia respecto a esta primera interacción realizada por el usuario ante la plataforma que se basa en contenido popular, se generan recomendaciones que refieren a películas con contenido similar al elegido previamente por el usuario, la manera en la que se seleccionan las películas a recomendar se basan en género a causa de que existen varias categorías que se ajustan a una película, por lo tanto, el sistema de recomendación basado en un producto selecciona por medio de una medida de relación (correlación), qué elementos en este caso géneros, son más recíprocos entre las películas disponibles por la plataforma.
- Preferencia del usuario: A diferencia del sistema de recomendación basado en popularidad, este sistema funciona a partir de interacciones preliminares entre usuario y la plataforma, entregando contenido similar respecto a las preferencias y patrones de consumo de otros usuarios. La limitación que puede presentar este recomendador puede darse en sesgos generados por búsquedas atípicas del usuario a causa de la gran variedad de elementos con los que se dispone.
- Colaborativo: Este sistema de recomendaciones posee similitudes respecto a las preferencias de usuario, sin embargo, su principal diferencia es la forma de cómo se dispone de los datos, ya que tiene en cuenta las calificaciones e interacciones de los usuarios para recomendar este tipo de contenido. Como resultado, los atributos de este sistema es la flexibilidad, por su capacidad de adaptarse al usuario dependiendo de sus interacciones, además en el artículo mencionado anteriormente, *“Hybrid recommender systems: Survey and experiments. User Modeling and User-Adapted Interaction”*, se hace apreciación de que este sistema de recomendación compara perfiles de usuarios o elementos, por este motivo, es capaz de manejar grandes cantidades de información y con esto generar recomendaciones más precisas.

#### **d. Selección de variables**

En la selección de variables se identificó que la toda composición de las bases de datos agregan valor exceptuando la variable “timestamp”. Durante la creación de la tabla final que se usa para la elaboración de los sistemas de recomendación se originan nuevas variables, estas derivan de la composición de algunas variables originales que se tratan para poder extraer datos puntuales que faciliten su manipulación y sean útiles.

Las variables originales que se deben manipular son “title” y “genres”, ya que para el primer caso la variable consta del título y el año de estreno, como resultado se obtiene una nueva variable llamada “year” y la variable “title” se reestructura con solo el título de la película. Por otra parte, la variable “genres” debe ser tratada ya que algunas películas se encuentran categorizadas en más de un género, dando como resultado un total de 19 variables que hacen alusión a cada género.

Comprendiendo la nueva composición de la base de datos, es importante aclarar que para este ejercicio puntual se aplican diversos sistemas de recomendación por lo que, de acuerdo con sistema que el usuario quiera usar para filtrar las películas que pretende recibir como recomendación, las variables ganan o pierden valor, sin embargo la captación de información de cada una de ellas para alimentar la base de datos y generar los sistemas es de vital importancia para el éxito del servicio.

#### **e. Evaluación y análisis de los sistemas de recomendación**

En este apartado se pone a consideración las recomendaciones realizadas por cada sistema, analizando si las películas recomendadas tienen relación a los criterios de búsqueda y filtro que el usuario establece:

- Basado en popularidad: Los sistemas de recomendación basados en popularidad que se elaboran usan diversos criterios, cada uno de ellos tiene una alta precisión en las recomendaciones arrojadas, ya que los puntajes asociados a los títulos son altos, adicionalmente al comparar entre las recomendaciones se pueden asociar varios títulos siendo consistentes en la información arrojada. Aunque existen títulos en común también se identifica diversidad en cada modelo garantizando que la experiencia basada en el consumo de todos los usuarios de la plataforma cumpla con satisfacer al usuario en cuestión.
- Contenido basado en un producto: Se identifica que las recomendaciones arrojadas por este sistema son precisas, por medio de características de cada película se logra asociar otros títulos que presentan patrones similares. De esta manera se obtiene un modelo novedoso que puede brindar una experiencia diferente al usuario al tener recomendaciones basadas en datos que distan de la subjetividad.
- Colaborativo: La recomendación arrojada por este sistema puede tener una de las mejores precisiones al usar una construcción colaborativa y un valor agregado al filtrar las recomendaciones a títulos que el usuario no haya visto.
- Preferencia del usuario: Al filtrar la información por cada usuario se brinda una experiencia personalizada estableciendo como valor de referencia su historial de consumo, esto puede ayudar a mejorar la precisión del sistema en comparación al contenido basado en un producto.

#### **f. Despliegue del modelo**

El proceso para generar un sistema de recomendaciones para la plataforma online, varía dependiendo de la clase de modelo que se implemente, en caso de aplicarse el sistema de recomendación basado en contenido, se recomendarán películas que contengan semejanzas con contenido antes visualizado, si fuese un modelo colaborativo, las películas que se recomiendan son las de mejor rating de acuerdo a usuarios con gustos semejantes. Se recomienda actualizar o generar nuevas recomendaciones cada mes, con el fin de mantener el interés por parte de los usuarios de la plataforma e identificar nuevas películas de interés. Además se hace la apreciación de realizar un filtro respecto a la calidad de las calificaciones con la finalidad de pulirlas y hacer recomendaciones más precisas y útiles.

En cuanto a la visualización de la plataforma se realiza por medio de una interfaz intuitiva, que presenta fácil acceso a cualquier tipo de persona (niños), al loguearse el usuario se visualizan listas con títulos de acuerdo a sus características, tales como: películas populares, contenido basado en búsquedas históricas de acuerdo a calificaciones, además se podrá buscar películas añadiendo ayudas que le faciliten ser hallada, por medio del año de publicación, género y rating.

Además, podría agregarse una lista que se actualice de manera mensual, donde se muestran las novedades en la plataforma. Al poseer interés en una película puede obtener una previsualización con información a detalle de la película, para lograr que cada usuario obtenga datos que de manera previa familiarice la película a sus gustos y a la hora de terminar de visualizar determinado contenido, puntuar las películas con el fin de mejorar el algoritmo de recomendación.

#### **g. Conclusiones**

- Los sistemas de recomendaciones utilizan de diferentes maneras la información con la que se cuenta, por lo que poseen diferencias en sus características y limitaciones, por tal razón el uso de sistemas de recomendación híbridos, abre la posibilidad de acomodarse a las diferentes preferencias que pueda tener un usuario en determinado momento, la personalización y mostrar contenido diferente, popular y relacional respecto a otros usuarios, impactaría positivamente en el interés de los usuarios.
- La implementación de un sistema de recomendación adecuado, puede mejorar la posición de la empresa en el mercado a causa de una retención y buena calificación de los usuarios hacia la plataforma.
- Si bien los comportamientos del usuario son diversos y cambiantes respecto a algunas características que puedan tener las películas, es preciso adaptarse y pulir elementos que faciliten y acierten ante dichas preferencias.

#### **Bibliografía:**

Burke, R. (2002). *Hybrid recommender systems: Survey and experiments*. *User Modeling and User-Adapted Interaction*. **Recuperado de:**  
<https://doi.org/10.1023/A:1021240730564>

## Anexo:

