

Proyecto 2: Fase 1

Design thinking

Empatizar:

Primera encuesta realizada a los sujetos de investigación:

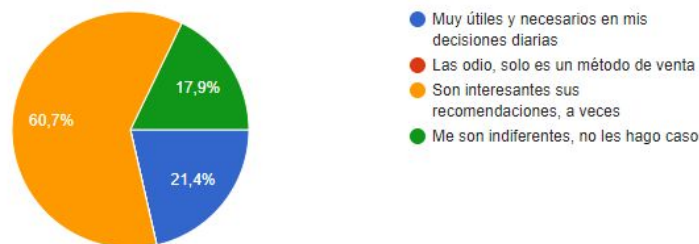
1. ¿Cuál es tu opinión sobre los sistemas de recomendaciones de las páginas web?
2. ¿Sobre qué tipo de productos te interesa más recibir recomendaciones?
3. ¿Consideras que sin las recomendaciones tardarías más en elegir una página, vídeo, canción, película, etc?
4. ¿Qué número le darías a tu indecisión? Donde 1 es lo más bajo y 5 lo más alto.
5. ¿Cuál entre estas dos tipos de recomendaciones preferías? Películas o Libros.

Se decidió realizar esta encuesta para conocer más sobre las distintas problemáticas que tienen los sujetos de investigación relacionadas al tema del proyecto. Para realizar un sistema que resuelva de manera efectiva la mayor parte de las problemáticas y conocer los mayores intereses que tiene el público respecto al tema.

Resultados de la primera encuesta:

¿Cuál es tu opinión sobre los sistemas de recomendaciones de las páginas web?

28 respuestas

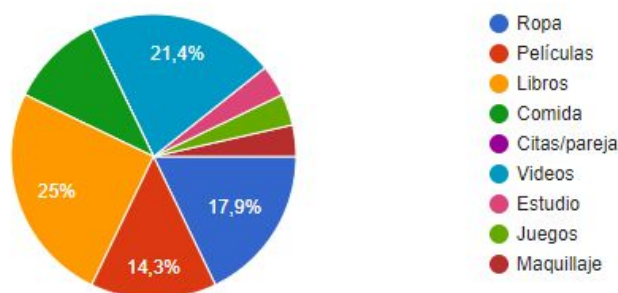


Gráfica 1. Opinión de los sistema de recomendaciones online

De acuerdo con la gráfica uno se pudo determinar que para el 60.7% de la población las recomendaciones realizadas por páginas web a veces son interesantes, mientras que para el 17% no les presta atención

¿Sobre qué tipo de productos te interesa más recibir recomendaciones?

28 respuestas



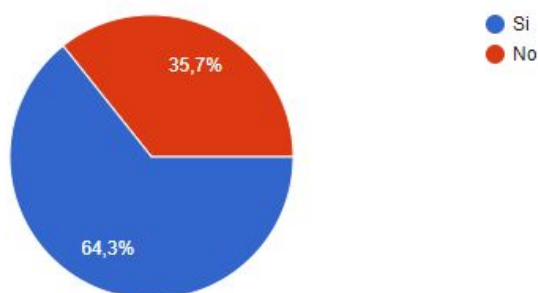
Gráfica 2. Recomendaciones de interés

De acuerdo con la gráfica dos se evidenció que para las personas entrevistadas los libros son los productos en el mercado del que más le interesa recibir recomendaciones. Seguidos por videos

¿Consideras que sin las recomendaciones tardarías más en elegir una página, vídeo, canción, película, etc?



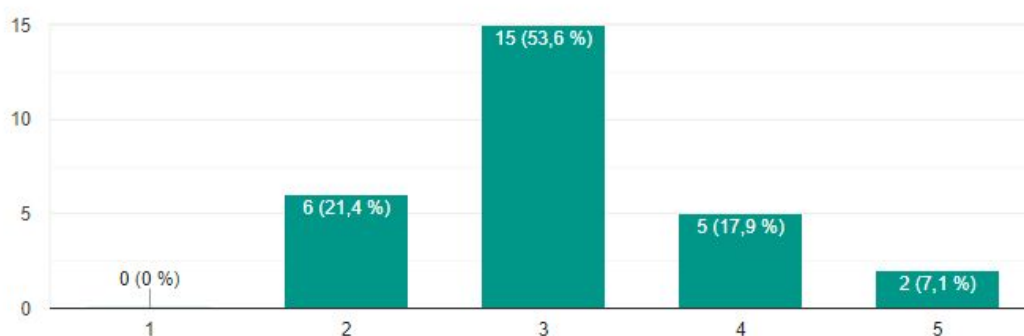
28 respuestas



Gráfica 3. Reducción del tiempo gracias a los sistema de recomendaciones online
De acuerdo con la gráfica tres se muestra que más del 60% de la población está de acuerdo en que sin las recomendaciones de sitios web tardarían más eligiendo el contenido a consumir.

¿Qué número le darías a tu indecisión? Donde 1 es lo más bajo y 5 lo más alto

28 respuestas

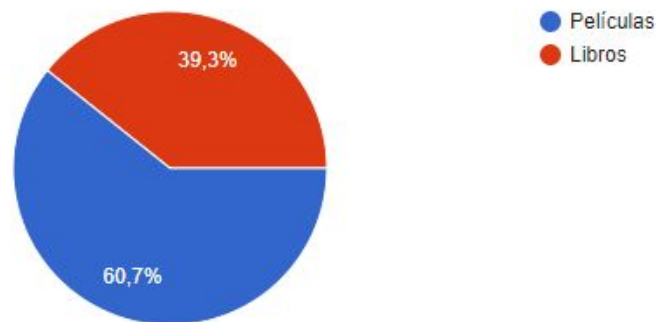


Gráfica 4. Nivel de indecisión

Con esta gráfica se puede ver que la mayoría de personas entrevistadas se encuentran entre un punto medio de indecisión.

¿Cuál entre estas dos tipos de recomendaciones preferías?

28 respuestas



Gráfica 5. Preferencia de recomendaciones

Esta pregunta permitió definir el problema a resolver porque el 60% de la población indicó que prefiere recibir recomendaciones de películas que de libros.

Segunda encuesta realizada a los sujetos de investigación:

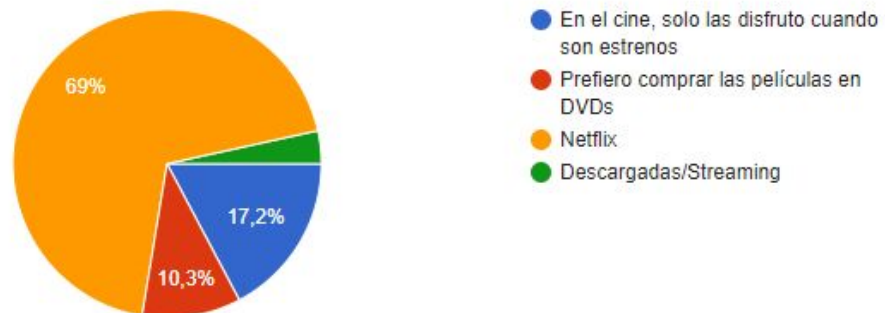
1. ¿Cuál es tu forma favorita de ver películas?
2. ¿En qué te basas para escoger qué película ver?
3. ¿Te gustaría tener un programa que te recomiende películas de forma personalizada?

Se realizó esta segunda encuesta en base a los resultados obtenidos de la primera. El objetivo de esta encuesta es poder conocer si el público está interesado en nuestra propuesta de proyecto y cómo podemos realizar una recomendación personalizada para cada persona interesada. Poder entender cómo se basan al momento de escoger una película y cuánto les influencia las recomendaciones de otras personas o sistemas.

Resultados de la segunda encuesta:

¿Cuál es tu forma favorita de ver películas?

29 respuestas

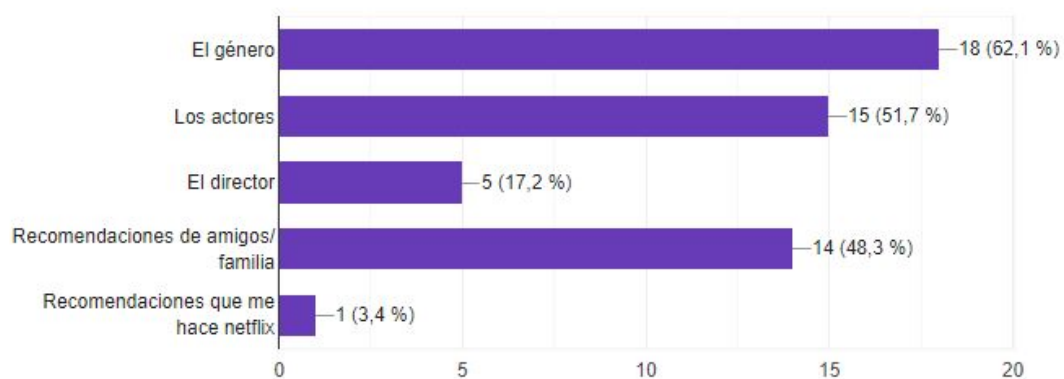


Gráfica 6. Maneras de ver películas

Esta gráfica es muestra del avance en el entretenimiento, ya que casi el 70% de las personas encuestadas prefieren ver películas a través de Netflix en lugar de DVD, cine o descargas.

¿En qué te basas para escoger qué película ver?

29 respuestas

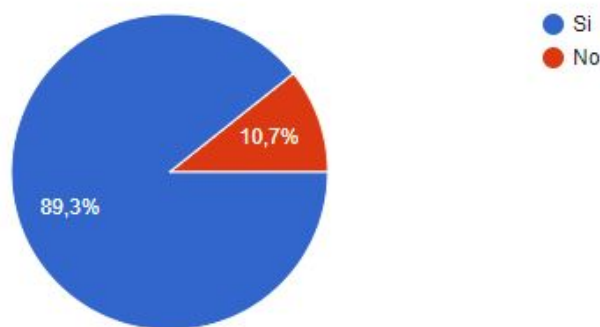


Gráfica 7. Bases seleccionar película

La gráfica siete evidencia que las personas encuestadas se basan principalmente en el género y los actores para ver o no una película

¿Te gustaría tener un programa que te recomiende películas de forma personalizada?

28 respuestas



Gráfica 8. Preferencia de programa para recomendación de películas

Con este gráfica se puede observar que al 89.3% de la población le gustaría acceder a un programa que le recomiende películas de forma personalizada y el 10.7 no mostró interés en esta oferta

Definir:

La compra y consumo de productos es algo habitual en nuestro diario vivir, desde las compras del supermercado hasta visualizar películas o escuchar música. La competitividad ha ido escalando entre las empresas dado porque constantemente salen nuevos productos al mercado. Debido a ello las empresas se vieron en la necesidad de recurrir a la creatividad para hacer sus productos notables al público deseado. El humano está acostumbrado a no salir de su área de confort, por lo cual tendemos a utilizar muchas cosas similares y a solo confiar en recomendaciones confiables de familia y amigos. De esta naturaleza nació el concepto de algoritmos de recomendación. Nuestras búsquedas en internet, amigos, seguidores y likes en redes sociales brindan mucha información acerca de nuestra personalidad, gustos y preferencias. Esta información es valiosa para los vendedores de productos, ya que en base a ella pueden crear una interacción muy personalizada al dar recomendaciones de productos que convienen y encajan en nuestra vida.

Actualmente las personas cuentan con una gran variedad de películas a elegir en distintas plataformas web y pierden un tiempo significativo buscando y eligiendo una película para ver. Algunas veces los usuarios reciben recomendaciones de amigos o conocidos pero no siempre terminan siendo efectivas porque cada persona percibe de manera distinta la misma película de acuerdo a su personalidad. Hay plataformas

como netflix que si realizan una recomendación personalizada a sus usuarios según el contenido previamente visualizado, pero lastimosamente solo lo hace con el contenido disponible en su página.

Investigar:

Existen 3 tipos de recomendación:

1. Filtración colaborativa

Este método se basa principalmente en la recolección y análisis de información del usuario de acuerdo a su comportamiento, actividades o preferencias para predecir sus futuras elecciones basándose en su similitud con otros usuarios.

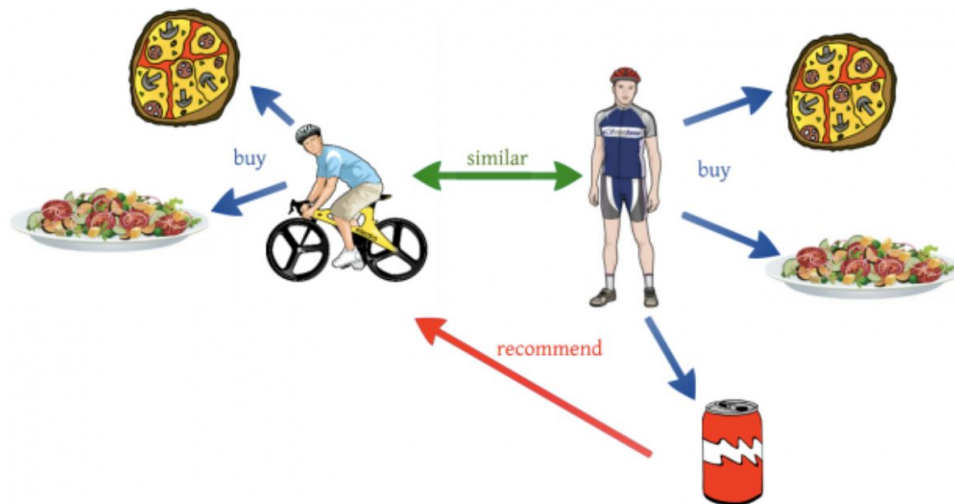


Imagen 1. Descripción gráfica de filtración colaborativa

2. Filtrado basado en contenido

Este método se basa es recomendar productos similares a los cuales el usuario ha mostrado interés en el pasado. Esto es en base a la idea que si te gusta un ítem probablemente también te guste otro muy similar a él. Para este método se requieren palabras claves que se buscan en los usuarios deseados para el producto que se busca consumir.

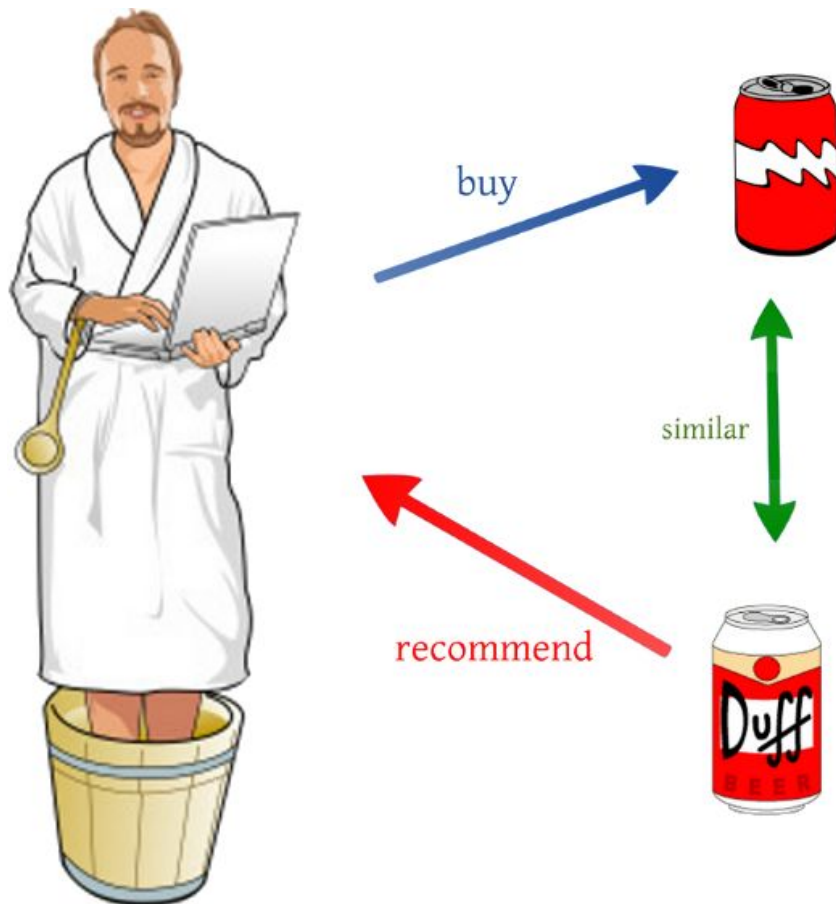


Imagen2. Descripción gráfica de de filtración basada en contenido

3. Sistemas de recomendación híbridos

Es la unión de ambos métodos previamente mencionado, ya que estudios recientes demostraron que la combinación es mucho más efectiva.

Hybrid Recommendations

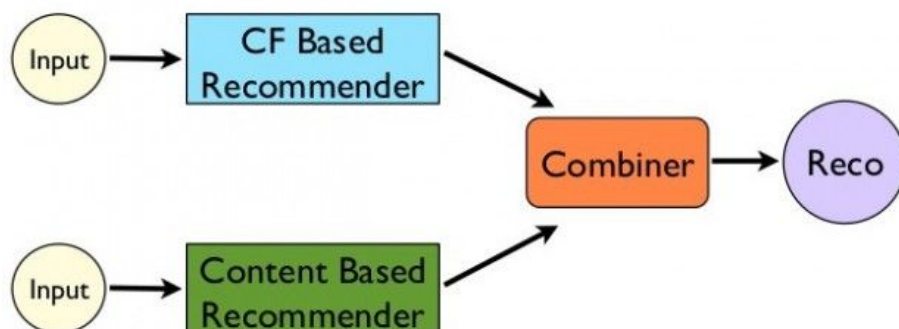


Imagen 3. Descripción gráfica de sistema de recomendación híbrida

¿Cómo funciona?

Fases:

- Recolección de datos:
La data puede ser implícita, por medio de los “clicks”, historial, carros de compra y páginas visitadas del usuario, o explícita, introducida por lo usuarios directamente como “ratings” y comentarios. Data del comportamiento es la más fácil de obtener dado que se obtiene mientras el usuario esté activo y no requiere de una acción extra de su parte.
- Almacenamiento de datos:
Entre más data almacenada, mejor será las recomendaciones. El tipo de data puede formar por si sola el tipo de almacenamiento necesario para su manejo y análisis.
- Análisis de datos:
Existen diferentes tipos de análisis:
 - a. *Real-time systems*: procesa la data mientras es creada. Útil para recomendaciones del momento.
 - b. *Batch analysis*: analiza la data periódicamente. Se requiere cierta cantidad de data para hacer su análisis relevante.
 - c. *Near-real-time analysis*: permite obtener data rápidamente para darle “refresh” al análisis luego de cierto tiempo (minutos o segundos).
- Filtración de datos:
Esta etapa final se basa en filtrar la data relevante para brindar las recomendaciones al usuario. Existen los siguientes algoritmos:
 - *Basado en contenido*: recomienda productos populares con características similares a los de interés del usuario.
 - *Colaborativo*: lo que otro usuarios con gustos similares recomiendan.
 - *“Cluster”*: recomienda un conjunto de productos que “van bien juntos”.

Algoritmos existente:

1. Amazon Prime: al algoritmo de recomendaciones de Amazon Prime se basa en un conocimiento profundo del cliente y cuenta con una base de datos muy extensa. Debido a su gran conocimiento del cliente este algoritmo puede realizar una personalización de la pantalla de inicio de cada cliente, haciendo una serie de recomendaciones según unos filtros o categorías específicas

para cada recomendación. Existen dos enfoques principales en las metodologías adoptadas por los sistemas de recomendación de Amazon Prime: Filtrado Colaborativo y los de Filtrado basado de contenido. Amazon guarda los gustos de los distintos clientes y su historial de reproducción. En base al historial de reproducción y sus distintos gustos por las películas de distintos géneros, actores, directores, etc, el algoritmo puede hacer una predicción en los posibles ítems que el cliente puede consumir. También comparando esta información con otros usuarios que poseen similitudes entre sí. (Linden G., Smith B., York J. 2006)

Amazon cuenta con un algoritmo personalizado de recomendación que ellos llaman filtrado colaborativo ítem a ítem. Se trata de un desarrollo de la casa, patentado en los Estados Unidos, y que nació porque ninguna de las propuestas existentes por aquel entonces (principios de década) servían para grandes conjuntos de datos.(Amazon, 2019)

2. Netflix: Este servicio de películas en streaming emplea algoritmos de machine learning para poder ofrecer al usuario contenido que sea de su interés. En la primera etapa netflix analiza el contenido visualizado de un usuario, el orden, el dispositivo utilizado y las fechas de reproducción. Así como los programas que han dejado a medias o los que ha calificado como favoritos y actividad de otros miembros con preferencias similares. (Netflix, 2018)

En la siguiente etapa la información recopilada del usuario se combina con todas las etiquetas determinadas al reproducir el contenido multimedia. Esto permite que los algoritmos de machine learning averiguen lo más importante. En este punto netflix clasifica a cada usuario con un tipo de gusto entre los 2,000 que tienen disponibles y de acuerdo con la categoría, el sistema muestra los contenidos recomendados. Cabe mencionar que las etiquetas son iguales a nivel mundial y no involucra información demográfica como la edad o género para decidir qué contenido ofrecer al usuario. (Arteaga, 2017)

Netflix además de seleccionar los títulos a incluir en la página de inicio, ejecuta un algoritmo para ordenar cada título dentro de cada fila y finalmente organizar las filas. En cada fila hay tres capas de personalización, la elección de filas (continuar viendo, tendencias, entre otras), títulos mostrados en la fila y orden de los títulos. Las filas más recomendadas se encuentran hasta arriba y en cada una los títulos de izquierda a derecha, en orden de preferencia. Si el sistema es árabe o hebreo, el orden de preferencia de los títulos es de derecha a izquierda. (Netflix, 2018)

Propuesta de ideas de solución del problema:

Sistema de recomendación de películas que no estén disponibles en una sola plataforma para que los usuarios puedan visualizar películas de distintas épocas, directores y países basado en sus elecciones previas. Nuestro mayor enfoque para filtración van a ser los actores y géneros de las películas que ya visualizó el usuario, ya que estos dos son las principales características por la cual un usuario escoge qué película desea ver de acuerdo a los resultados de nuestras encuestas.

Prototipar:

Diseño de base de datos:

Nuestra base de datos fue descargada del link: <https://neo4j.com/developer/movie-database/> la cual ya está diseñada para ser representada en grafos para Neo4j. Es muy similar a la utilizada por IMDB y contiene información de películas, actores, directores, género, entre otros. Fue creada y proporcionada por “The Movie DB” la cual es una comunidad web que recopila información diversa de varias películas desde el 2008, siendo una fuente de información confiable para brindar recomendaciones.

El modelo de datos es de la siguiente forma:

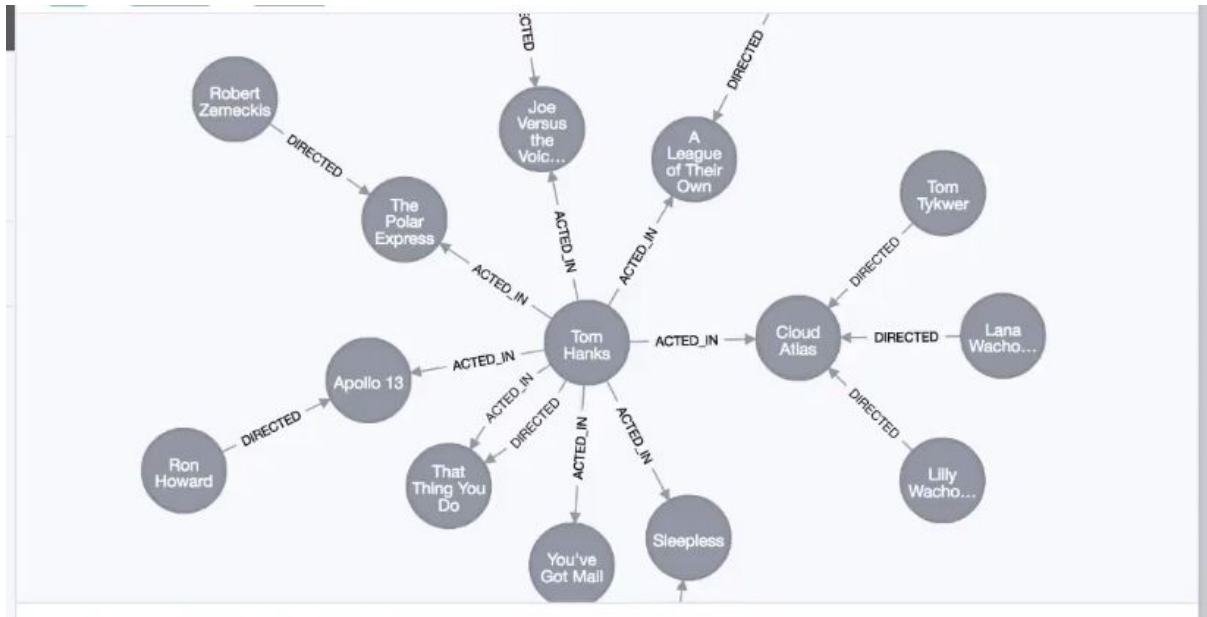
- (:Movie {title, released, ... })
- (:Person {name, born, ...})
- (:Person)-[:ACTED_IN|:DIRECTED|:PRODUCED]→(:Movie)

Ejemplo de su visualización:

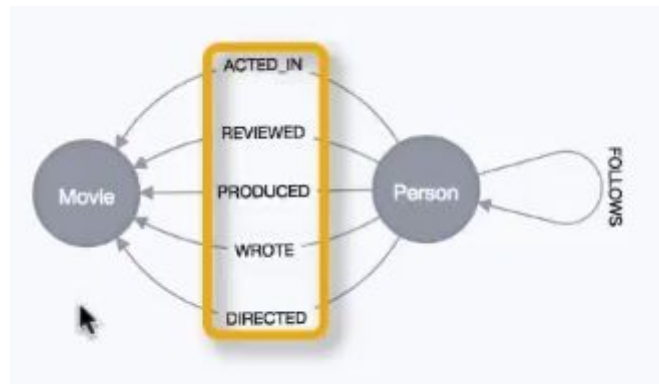
Para el grafo desarrollado a continuación se hizo una búsqueda del actor Tom Hanks como el nodo principal, luego como aristas o conexiones al actor se tiene las películas donde ha actuado y como nodo la respectiva película conectada con el director de la película. La búsqueda se realiza de la siguiente forma:

```
WITH TomH as a
MATCH (a)-[:ACTED_IN]->(m)<-[:DIRECTED]-(d) RETURN a,m,d LIMIT 10
;
```

Resultado:



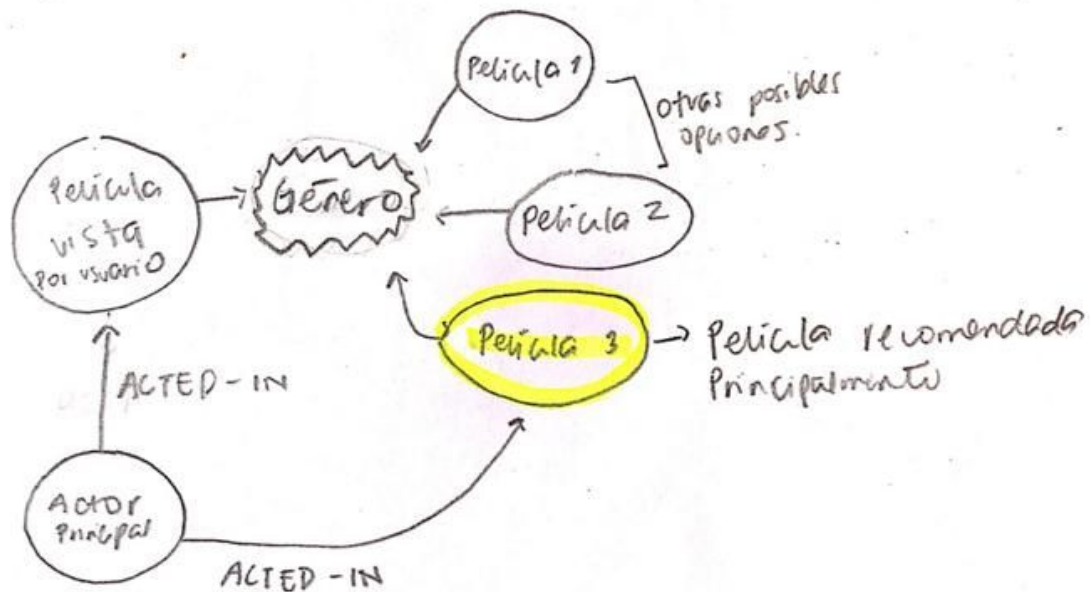
En esta base de datos existen dos tipos de nodos: movie y person. La relación entre ellos se basa en: ACTED_IN, REVIEWED, PRODUCED, FOLLOWS, WROTE y DIRECTED.



La relación siempre va dirigida del nodo Person al nodo Movie o a otra Person.

Algoritmo:

Nuestro programa se basará en filtros colaborativos (collaborative filtering) dado que nuestra lógica se basará en características e historial del usuario. Las características a evaluar son el historial de películas observadas por el usuario que él va a introducir. De las películas su primer filtro será el género, y luego del género se filtrarán los actores más recurrentes en su historial. Se realizará de la siguiente forma:



1. Usuario introduce una película vista y que haya sido de su gusto.
2. El programa filtra su búsqueda de recomendaciones por la elección del usuario con el género de la película
3. El listado de películas con el género es filtrado por lo actores de la película seleccionada.
4. Las películas que coincidan en la búsqueda en el género y el actor serán desplegadas principalmente.

Evaluar:

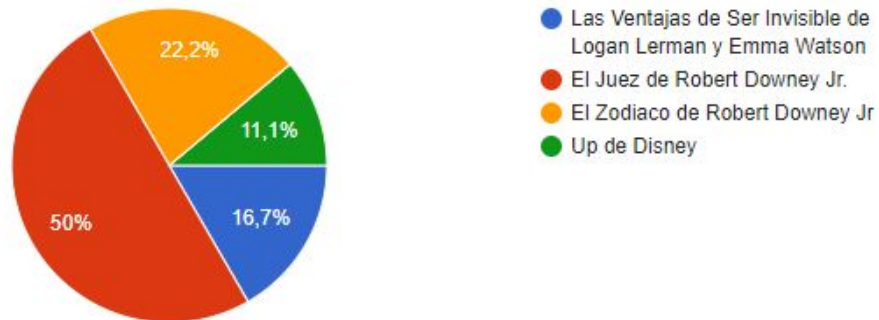
Tercera encuesta a realizar a los sujetos de investigación:

1. Seleccione la película que vería después de haber visto Avengers.
2. Seleccione la película que vería después de haber visto Rapidos y Furiosos.
3. Seleccione la película que vería después de haber visto Crepusculo.
4. Seleccione la película que vería después de haber visto Bajo la misma estrella.

Se realizó esta encuesta con el objetivo principal de comprobar nuestra hipótesis de que los usuarios seleccionan películas principalmente basados en el género y en un segundo plano por los actores principales.

Avengers

18 respuestas

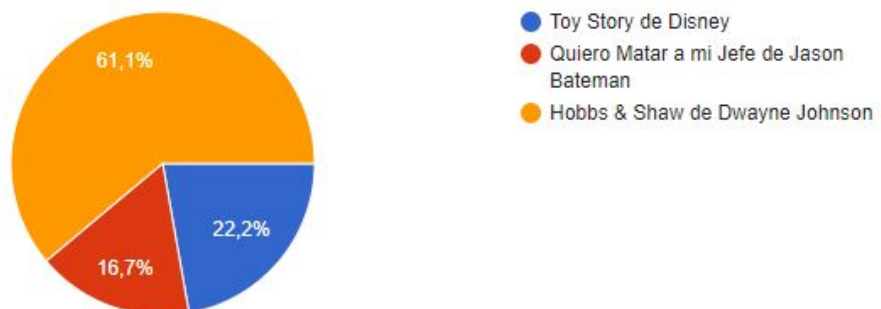


Gráfica 9. Elección de película luego de visualizar Avengers.

El 50% de la población seleccionó El juez de Robert Downey Jr como la película que vería después de ver Avengers. Ya que Avengers es una película de acción-ficción y la otra de acción-drama, se demostró que los encuestados se basaron tanto en el género como en el actor principal para ver una siguiente película.

Rápido y Furioso

18 respuestas

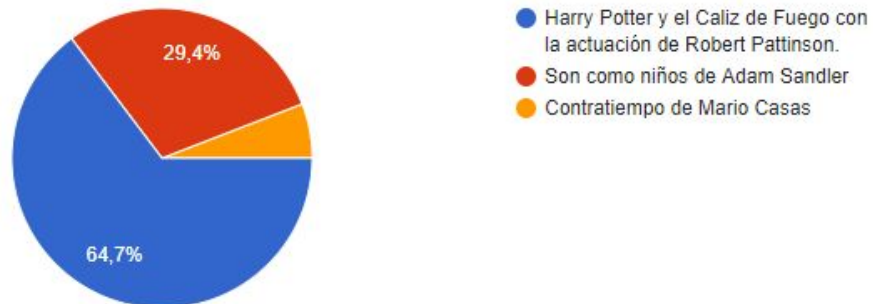


Gráfica 10. Elección de película luego de visualizar Rápido y Furioso.

Esta gráfica demuestra que el 61.1% de las personas encuestadas luego de visualizar Rápido y furioso prefería ver Hobbs & Shaw de Dwayne Johnson. Ambas películas coinciden en género y actor principal, por lo que se confirma una vez más que el usuario se basa en estas dos características para seleccionar una película

Crepúsculo

17 respuestas

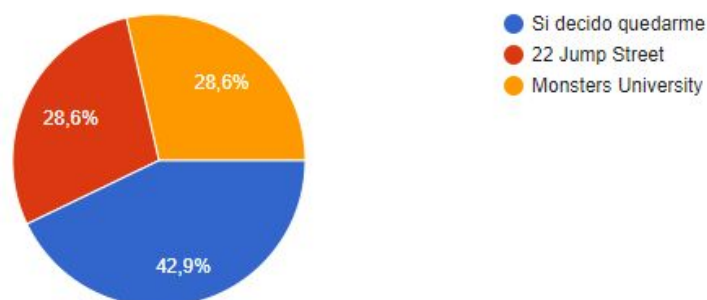


Gráfica 11. Elección de película luego de visualizar Crepúsculo.

Se puede observar que el 64.7% de la población vería Harry Potter y el Cáliz de Fuego con la actuación de Robert Pattison después de visualizar Crepúsculo. Ya que tanto Crepúsculo como Harry Potter son películas de ficción en las que Pattison forma parte del elenco, se muestra que el usuario seleccionó la segunda película de acuerdo al género y actor.

Bajo la misma estrella

7 respuestas



Gráfica 12. Elección de película luego de visualizar Bajo la misma estrella.

Con esta última gráfica se observa que el 42.9% de las personas preferirían ver Si decido quedarme después de ver Bajo la misma estrella, donde ambas películas coinciden en el género de Drama-Romance.

Conclusión:

En base a los resultados de las encuestas realizadas y la investigación de sistemas de recomendaciones, podemos concluir que con este proyecto si solucionaremos el problema que las personas tienen para elegir una película a visualizar. Esto porque para realizar las recomendaciones efectivas nos basaremos en las películas de géneros y actores en común con las películas ingresadas como favoritas por parte del usuario. Para lograr dicha solución contaremos con una base de datos extraída de The movie DB y vincularemos los nodos necesarios en la plataforma de neo4j.

Bibliografía:

- Arteaga, S. (2017, Agosto 22). Así funciona el sistema de recomendaciones de Netflix. Extraído de <https://computerhoy.com/noticias/apps/asi-funciona-sistema-recomendaciones-netflix-66807>
- Maruti TechLab. (2019, Marzo 27). How do Recommendation Engines work? And What are the Benefits? Extraído de <https://www.marutitech.com/recommendation-engine-benefits/>
- Dataconomy. (2016, Mayo 30). An Introduction to Recommendation Engines. Extraído de <https://dataconomy.com/2015/03/an-introduction-to-recommendation-engines/>
- Neo4j. (2019). Dataset: Movie Database. Extraído de <https://neo4j.com/developer/movie-database/>
- The movie DB. (2019, Marzo 04). About TMDb. Extraído de <https://www.themoviedb.org/about>
- Netflix. (2018, Abril 24). Cómo funciona el sistema de recomendaciones de Netflix. Extraído de <https://help.netflix.com/es/node/100639>
- Linden G., Smith B., York J. (2006). Amazon.com Recommendations Item-to-Item. Collaborative Filtering. Extraído de: <http://www.cs.umd.edu/~samir/498/Amazon-Recommendations.pdf>
- Amazon.com (2019). Políticas de uso. Extraído de: https://www.primevideo.com/help?_encoding=UTF8&nodeId=202095490