

Guillermo Alfonso Furlan - 20713
Sebastián Solorzano - 21826
Alejandro Pallais Garcia - 20093
Wilfredo Gallegos - 20399

Proyecto 2 Algoritmos y Estructuras de Datos

Fase 1

Investigación Algoritmos

Un algoritmo de recomendación es un algoritmo que filtra toda la información disponible sobre un usuario en particular para predecir qué es lo que tendrá valor para él dentro del catálogo disponible en la página con el fin de que la sugerencia también se convierta en una transacción o un motivo para evitar la pérdida de ese cliente.

Dependiendo del análisis de información realizado por el algoritmo, los filtros para los sistemas de recomendaciones se clasifican de la siguiente manera:

- Contenido: Las sugerencias dadas por el algoritmo se basan en las características mismas de los productos, buscando similitudes entre lo que se ha consumido o gustado en el pasado. Es un modelo sencillo porque solamente se requiere la información de los productos para armarlo, pero no toma en consideración la información proveniente de las transacciones lo que dificulta su mejora.
- Híbrido: El sistema procesa tanto las características del producto o servicio como las interacciones realizadas para dar mejores recomendaciones al usuario. El sistema de Netflix es un excelente ejemplo puesto que da sugerencias basadas en lo que las personas ven, sus hábitos de búsqueda y características similares entre las películas.
- Implementar adecuadamente un sistema de recomendaciones puede dar un impulso a las ventas de un producto o servicio desconocido para el usuario, además de ser una forma efectiva de marketing personalizado al crear una experiencia única para cada individuo. ¿Sientes que las sugerencias que dan son cosas que tú comprarías?
- Popularidad: Aconseja por la “popularidad” de los productos. Por ejemplo, “los más vendidos” globalmente, se ofrecerán a todos los usuarios por igual sin aprovechar la personalización. Es fácil de implementar y en algunos casos es efectiva.

- Colaborativo: El sistema presenta recomendaciones procesando las calificaciones otorgadas por los usuarios. Los resultados se interpretan por métodos basados en la memoria, encontrando usuarios con perfiles parecidos para dar sugerencias de acuerdo con lo que ellos han comprado o les ha gustado, o en modelos que realizan predicciones como lo haría una máquina cualquiera.

Este último en particular se puede realizar de dos formas, basado en los objetos, que se dedica a calcular la similitud en los objetos y encontrar los objetos no evaluados similares a los que tengan su mejor evaluación o, basado en los usuarios, que calcula similitud entre los usuarios y te recomienda los objetos que tú no has evaluado y el usuario semejante evaluó bien.

Para el método basado en los usuarios, hay diversas maneras de medir o calcular la similitud entre usuarios y de ello dependerá que se den buenas recomendaciones. Una de las maneras de medir esa similitud se llama “distancia por coseno de los vectores” y por simplificar el concepto, digamos que crea un espacio vectorial con n dimensiones correspondientes a los n objetos y sitúa los vectores siendo su medida el “valor rating” de cada usuario a ese objeto. Luego calcula el ángulo entre los vectores partiendo de la “coordenada cero”. A “poca distancia” entre ángulos, se corresponde con usuarios con mayor similitud. Una vez que tenemos la matriz de similitud, nos valdremos de otra operación matemática para calcular las recomendaciones.

Fórmula para calcular los ratings faltantes: sería algo así como “Matriz de similitud PROD. VECTORIAL ratings / (sumatoria de cada fila de ratings) Transpuesta

Cómo lo haremos: cada rating se multiplica por el factor de similitud de usuario que dio el rating. La predicción final por usuario será igual a la suma del peso de los ratings dividido por la “suma ponderada”.

Esta es solo una forma de hacerlo para un tipo específico de algoritmo, puede haber otras, y lo que cambia es el enfoque y las fórmulas, pero la idea es la misma para todos.

Design Thinking

Resultados Encuesta y resumen de entrevistas

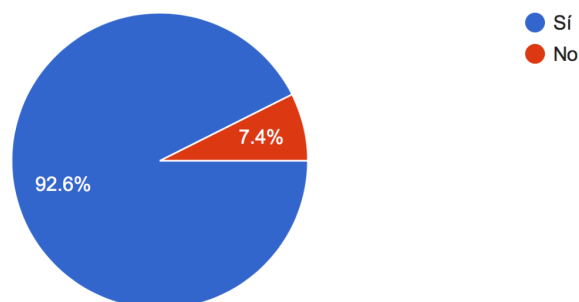
Entrevistas: Los puntos más importantes de las entrevistas fueron:

- Las recomendaciones de los servicios de streaming de series relacionadas no suelen ser acertadas.
- Cuando se tienen más de 3 servicios de streaming se complica mucho el buscar una serie ya que hay muchas series entre los servicios de streaming y las series con los parámetros que se buscan se vuelve una tarea tediosa.
- Depende que se quiere la cantidad de temporadas es importante, ya que si se busca una serie para pasar el rato se suele buscar una con muchas temporadas , pero si se desea una serie para un maratón un poco más pesada suele buscarse series más cortas
- Una serie se considera larga si tiene más de 5 temporadas normal entre 4-5 y corta si tiene menos de 3 temporadas.
- La temática de la serie es importante al momento de decidir si verla.
- Los actores que participan en la serie no es un factor definitivo para ver o no una serie.
- El género de la serie es importante.
- Los servicios de streaming solo permiten buscar series a partir de un único parámetro(nombre genero, actor etc)

Resultados de encuesta

Se te dificulta escoger una serie nueva para ver?

27 responses



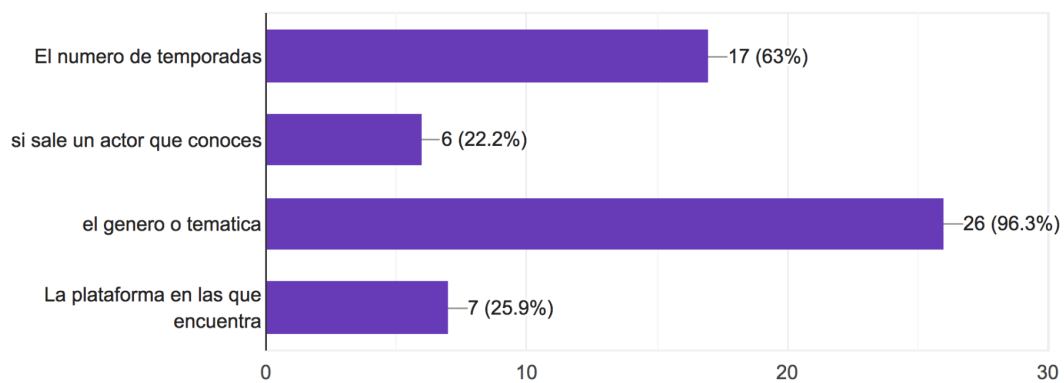
Cuando acabas una serie prefiero que la siguiente serie que veas ...

27 responses



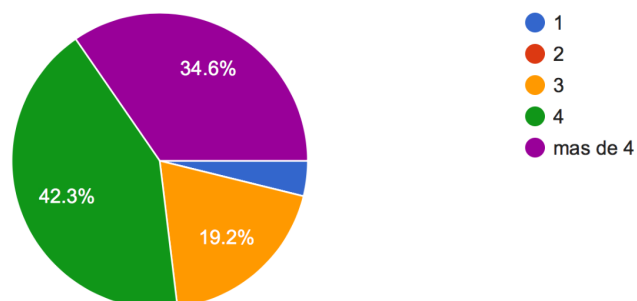
Que factores influyen en tu decisión de empezar una serie?

27 responses



Cuantos servicios de Streaming tienes contratados

26 responses



Necesidades identificadas

- El tiempo de las personas es limitado y con el boom de los servicios de streaming se pierde cada vez más tiempo en buscar una serie para ver por lo que se ***necesita reducir el tiempo que se invierte en buscar una serie*** que satisfaga las que se tengan necesidades.
- El hecho que existan muchos servicio de streaming implica que hay muchas series que solo se encuentran en uno o dos de esto por lo que cuando le recomiendan una serie a alguien es importante ***saber en qué servicio de streaming pueden encontrar la serie***
- En ocasiones no se tiene claro que se quiere ver por lo que solo ***se busca algo parecido a algo que ya se vio.***
- **Realizar una búsqueda más puntual** ya que las aplicaciones tiene sistemas de recomendaciones pero son más generales y muchas veces por intentar predecir basado en todo lo que se vio las recomendaciones son muy variadas

Problema Identificado

Después de haber encuestado y entrevistado a varias personas como grupo pudimos identificar una serie de problemas. El principal es que la mayoría de personas se les dificulta el escoger una serie nueva para ver y que la mayoría de personas no solo se les dificulta encontrar si no que por lo general buscan una serie similar a la que acaban de terminar. Esta tarea se dificulta aún más tomando en cuenta que más del 40% de los encuestados tiene más de 4 de servicios de streaming. Además de esta información gracias a las encuestas y entrevistas también se pudo identificar que los dos factores que definen si ver una serie o no es la temática y la cantidad de temporadas. El problema en resumen es que a las personas les cuesta encontrar una serie de temática similar a la última que se vio y con un número de temporadas entre un rango específico.

Posibles soluciones

- Un programa al cual se le ingrese la última serie que se vio (o una serie en particular) y en base la información de esta buscar series similares y recomendarlas. (A)
- Al programa se le pueden ingresar qué plataformas se tienen y recomendar solo serie en las plataformas que se tienen. (B)
- Un programa que se le ingresan parámetros (cantidad de temporadas, temática , etc) y recomendar en base a esto. (C)

Necesidades cubiertas por la solución

Un programa que tenga características de A,B y C podría cubrir las necesidades identificada; A permite cubrir la búsqueda por default por así decirlo , en la cual no se tiene tan claro que se quiere ver pero se tiene una serie de referencia por lo que se busca algo similar, B permite eliminar el factor plataforma ya que ahora se obtendrán solo recomendaciones que están dentro de las posibilidades que realmente tenemos pero sin limitarnos a recomendar a una sola plataforma, C nos permite el escoger parámetros un poco más específicos si no se tiene una serie de referencia. Y todos logran reducir el tiempo invertido en búsqueda al enfocar las recomendaciones de manera más simple.

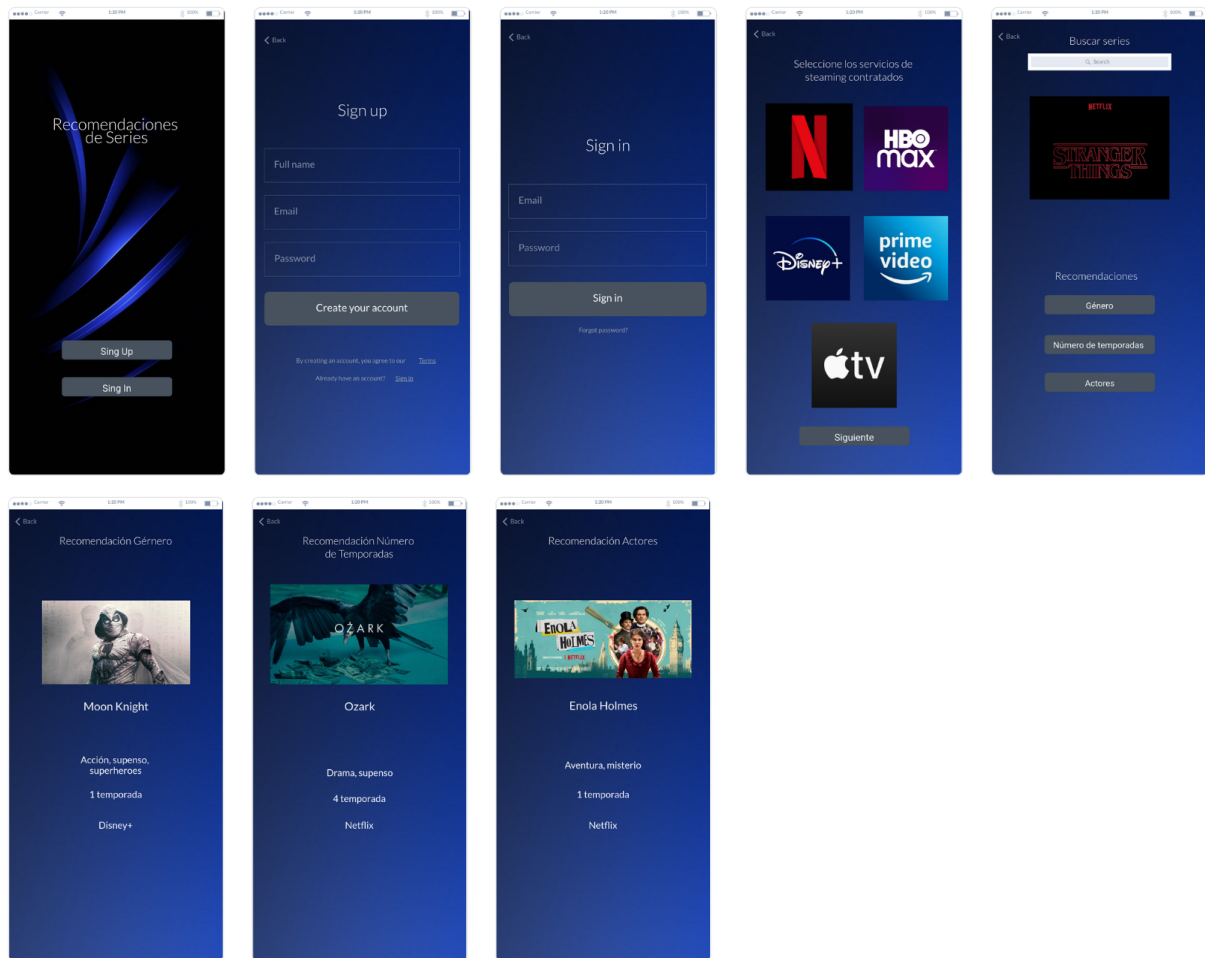
Pseudocódigo

Idea general: El programa que se desea desarrollar para resolver los problemas identificados es el siguiente. Un programa al cual se le pueda ingresar la última serie que se vio y el programa buscará esta en la base de datos y recolecta programas con características similares usando los parámetros identificados en el diseño thinking. También se tendrá la opción de ingresar directamente los parámetros que se desean y de misma forma recomendando series que cumplan los parámetros.

funciones a desarrollar:

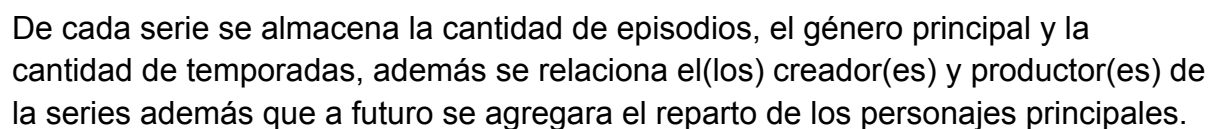
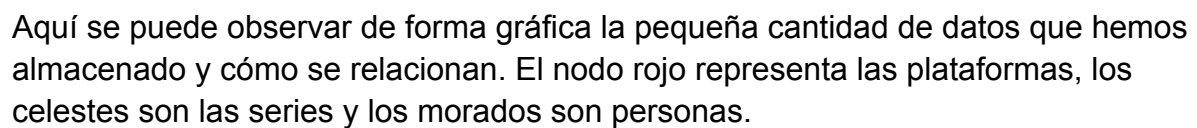
- Función que recoja el nombre de la serie ingresada y la busque en la base de datos.
- Función que traiga de la base de datos los parámetros deseados de una serie determinada.
- Función que compare los parámetros de todas las series de la base de datos con unos parámetros indicados y que almacene las que lo cumplen.
- función que mande la lista de series que cumplen los parámetros a la vista gráfica.
- (opcional) Método para crear cuentas
- opcional) método que permita almacenar series que se desean ver más tarde

Prototipo de baja resolución



Base de datos inicial

Se utilizó Neo4j Desktop para llevar a cabo una base de datos pequeña, para mostrar de qué forma van a estar distribuido y organizado los datos a utilizar en nuestro proyecto, a continuación se muestran capturas de pantalla que muestran básicamente como funciona:



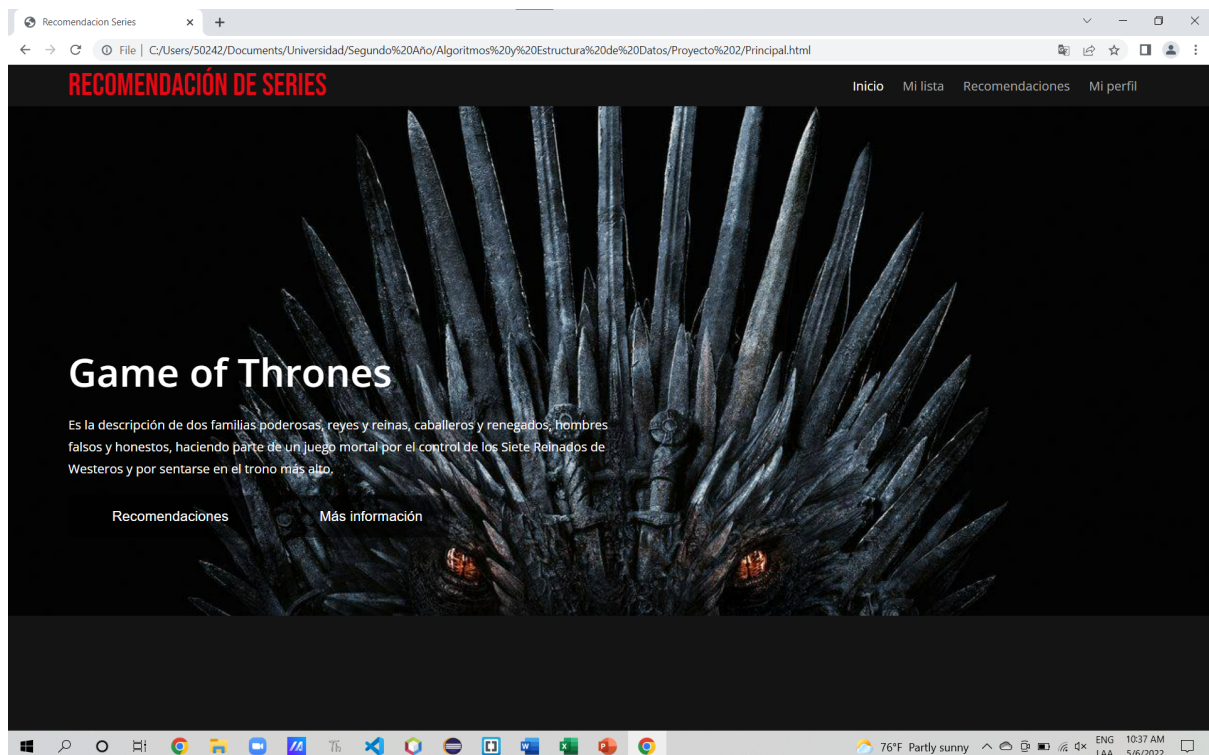

```
selfert, function: Creator })
10 Create(MiklosWeigert:Person{name:'Miklos
Weigert',function:'Creator'})
11 Create(TerriMinsky:Person{name:'Terri
Minsky',function:'Creator'})
12 Create(MichelleManning:Person{name:'Michelle
Manning',function:'Producer'})
13 Create
14 (DalmatianStreet)-[:Show_de]→(DisneyPlus),
15 (AndiMack)-[:Show_de]→(DisneyPlus),
16 (AustinAlly)-[:Show_de]→(DisneyPlus),
17 (TobiasConanTrost)-[:Productor_de]→(GravityFalls),
18 (AlexHirsch)-[:Productor_de]→(GravityFalls)
```

Para ingresar los datos primero se crearon los objetos o nodos(personas, series y plataformas) y luego se realizaron las relaciones entre ellos.

```
1 MATCH
2 (a:Serie),
3 (b:Plataform)
4 WHERE a.title = 'Hannah Montana' AND b.title = 'Disney Plus'
5 CREATE (a)-[:Show_de]→(b)
```

Y para agregar nuevos datos, se agregan los nodos y luego se utiliza la función match para relacionarlos con los datos preexistentes.

Página principal



Enlace Prototipo: <https://marvelapp.com/prototype/177h137j/screen/86442241>

Enlace Git: <https://github.com/Furlan11/Proyecto-2-Estructuras>