

UNIVERSIDAD DEL VALLE DE GUATEMALA  
ALGORITMOS Y ESTRUCTURA DE DATOS

PROYECTO No. 2  
“Grafos”

LUIS ESTURBAN  
GUSTAVO DE LEÓN  
CARLOS CHEW

## RETROESPECTIVA

### 1. Descripción del problema

Los clientes que pertenecen al mercado de películas han demostrado que dejan el servicio porque no encuentran películas que ver, por lo tanto, surgió la necesidad de realizar un programa .java con el objetivo de sugerir películas según el género que prefieran.

Este problema será implementado con grafos, donde cada usuario aporta sus gustos de películas y con esa información predecir que películas recomendar a los demás usuarios y así sucesivamente.

### 2. Necesidades detectadas

- Los usuarios no saben que películas ver, por lo tanto, dejan de pagar
- El usuario no sabe cuales son las películas top trending del momento
- Ampliar el mercado al tener mas usuarios
- El administrador del programa tendrá mas información para controlar a los usuarios

### 3. Oportunidades para cada una de las necesidades detectadas

- Es oportuno implementar una interfaz gráfica amigable con el usuario, donde sea interactiva y fácil de utilizar
- Que la implementación de la base de datos sea rápida en respuesta, donde las sugerencias se presenten de una manera inmediata
- Con esta implementación se puede almacenar la información, y con esa información analizarla para que el cálculo de las sugerencias sea más complejo
- Sería oportuno implementar un programa que saque estadísticas por medio de la información de los clientes

### 4. Fase de empatía

Una vez identificado el problema central, se recurrió a atestiguar directamente el panorama de este a través de las confesiones de los afectados directos y personas relacionadas al conflicto. Esto con el objetivo de identificar las necesidades en primera persona y con ello encontrar soluciones pertinentes.

Para este fin se establecieron tres perfiles de interés, los cuales se desarrollan a continuación.

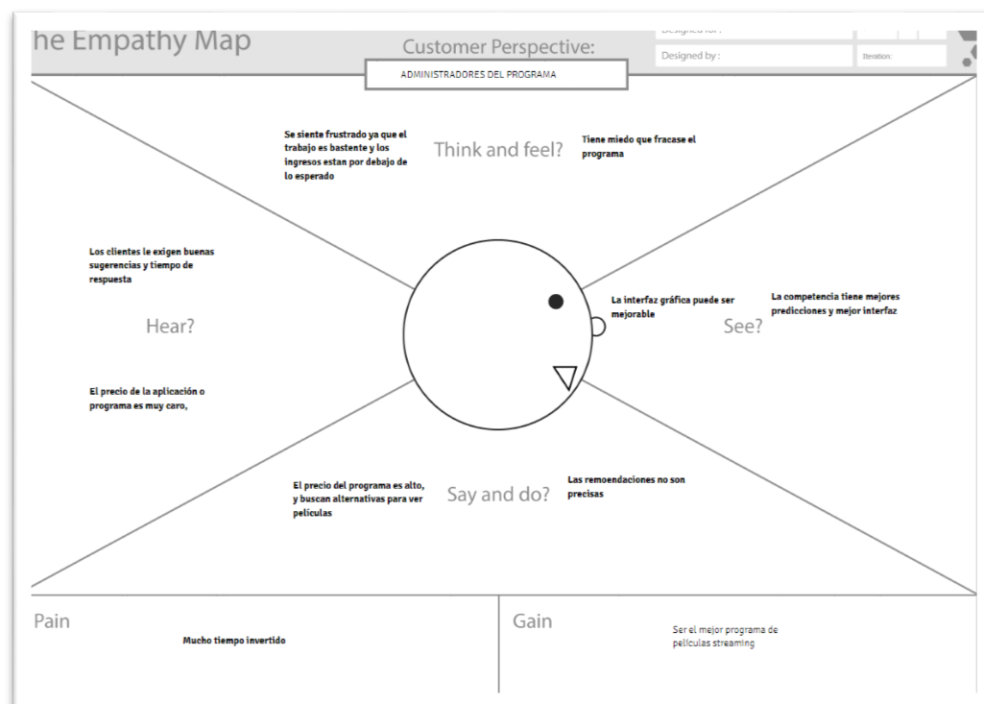
- Administrador del programa: esta persona es el encargo de mantener y desarrollar el programa, donde es el único que tiene acceso al código fuente, su objetivo es optimizar el código de tal forma que sea el mejor de su tipo
- Cliente: esta persona es la persona a quien se le presentan las sugerencias, por medio de los cálculos hechos por el administrador

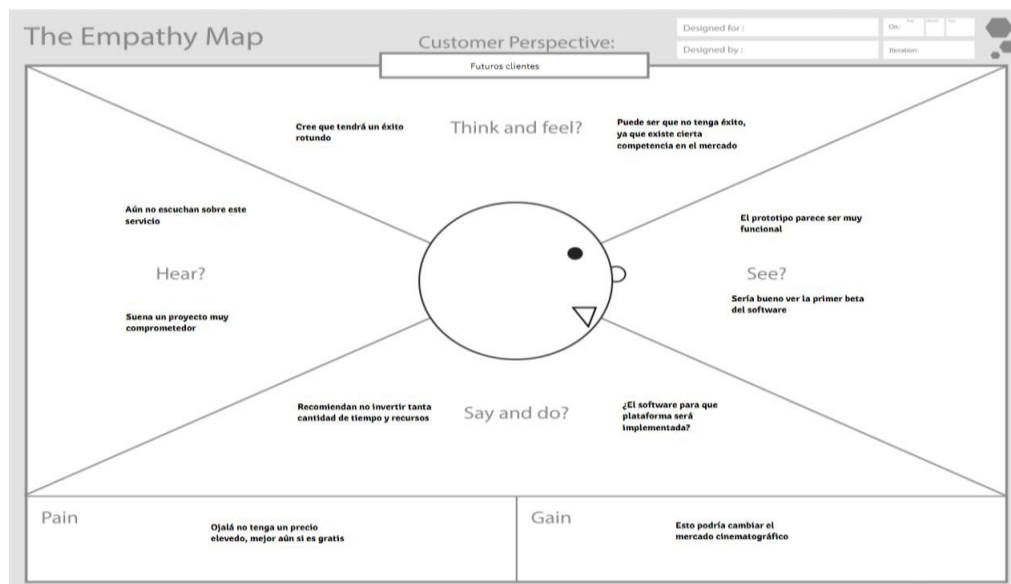
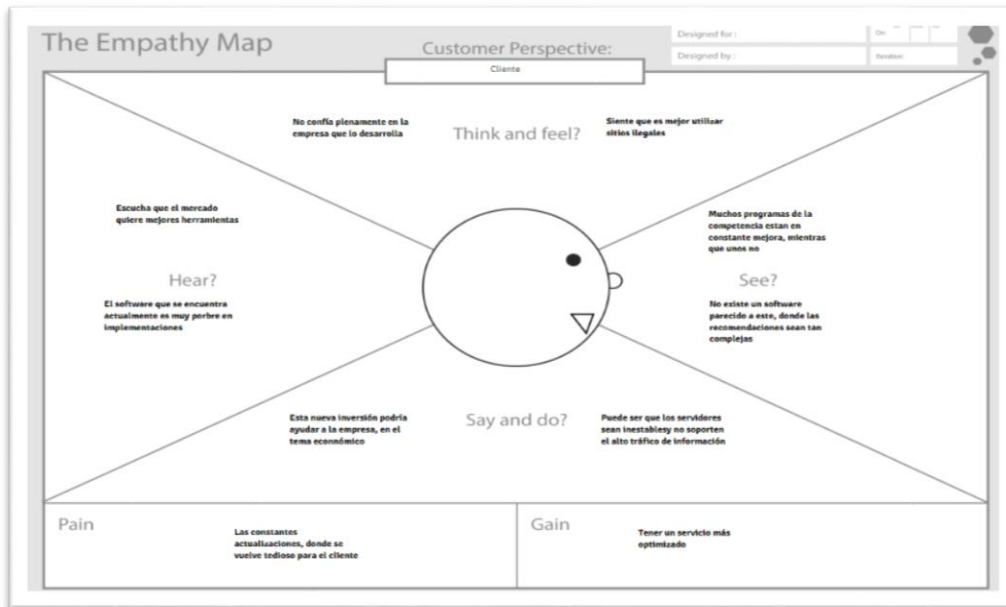
- Futuros clientes: este perfil es aquel que no utilizar el programa, sin embargo, está incluido como potencial consumidor, ya que la idea es tener la máxima cantidad de clientes posible. Esto servirá ya que se analizan sus comportamientos por medio de los mapas de empatía.

Se abordaron a los entrevistados con las siguientes preguntas, las cuales tuvieron ciertas variantes de acuerdo con el tipo de perfil entrevistado.

- ¿Usted cree que el streaming de películas online ha tenido un fuerte impacto en la sociedad?
- ¿La falta de películas u/o análisis de la información cree que afecta en el análisis de predicción de películas sugeridas?
- ¿Considera que las películas son el futuro de la cinematografía?
- ¿Usted le recomendaría este servicio a un amigo o familiar para que lo utilice?
- ¿Considera que los clientes necesitan algún sitio de recomendaciones de películas?

## 5. Mapas de empatía





## 6. Ideación

A partir de la información proporcionada al analizar las grabaciones de las personas entrevistadas se realizó una lluvia de ideas para poder encontrar una estrategia que ayude al desarrollador a mantener el software atractivo, pero además que satisfaga las necesidades de los clientes.

### *Descripción de las ideas más votadas:*

Con base a la lluvia de ideas, se seleccionaron las tres ideas que, desde nuestro criterio, satisfacen gran parte de las necesidades encontradas, y que además son factibles y viables de ejecutar desde nuestra posición. Finalmente se decidió fusionar dos de las ideas seleccionadas ya que esto permitirá cumplir de mejor manera las demandas analizadas. Por lo tanto, las dos ideas seleccionadas son:

1. *Realizar un sistema de recomendaciones sobre un área seleccionada, en este caso recomendaciones sobre restaurantes*

En base a las oportunidades surgidas de las necesidades encontradas se consideró que la solución que mejor se adapta al caso es sobre recomendaciones de restaurantes, siendo los restaurantes ubicados en la ciudad de Guatemala. Además, se tendrán todos los registros de los clientes, siendo estos catalogados por categorías para implementar una búsqueda según los gustos del cliente.

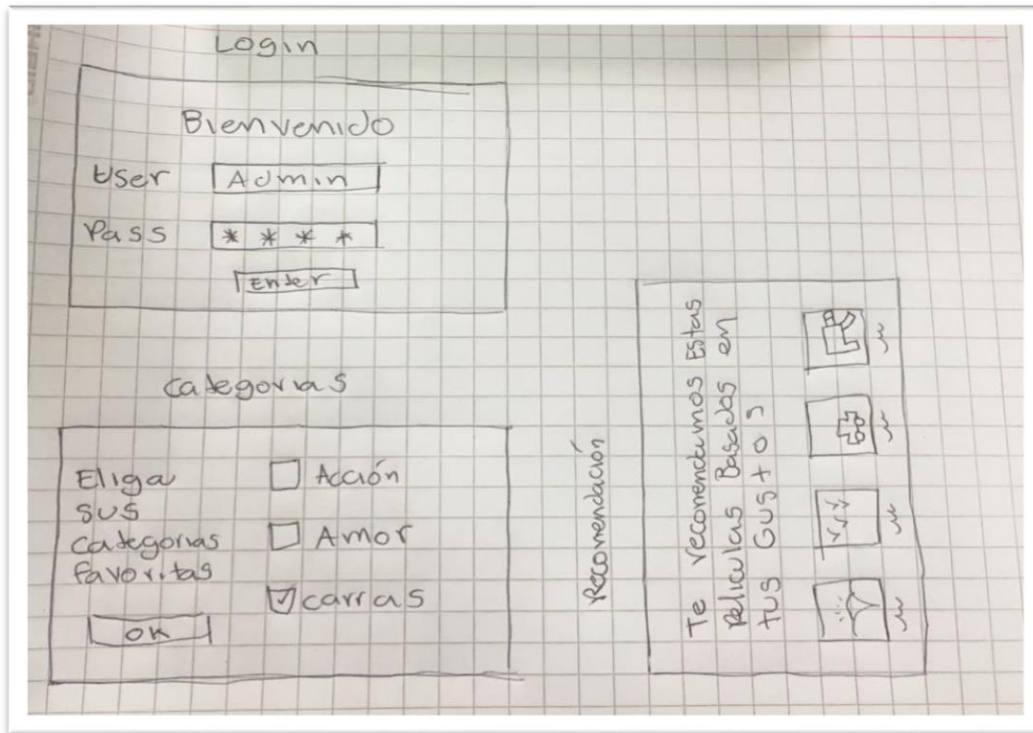
2. *Realizar un sistema de recomendaciones sobre un área seleccionada, en este caso recomendaciones sobre películas*

Se pensó en implementar un programa computacional con el objetivo de tener el registro de la información de los usuarios, esto para generar recomendaciones sobre posibles películas que al cliente le podrían gustar, esto implementado lo siguiente:

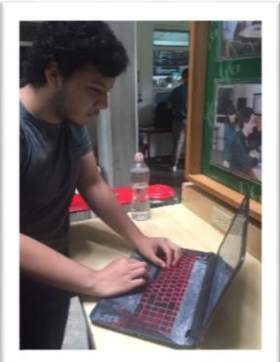
- Algoritmo de agrupamiento
- El método de vecinos K más cercanos
- Algoritmo de eliminación de candidatos

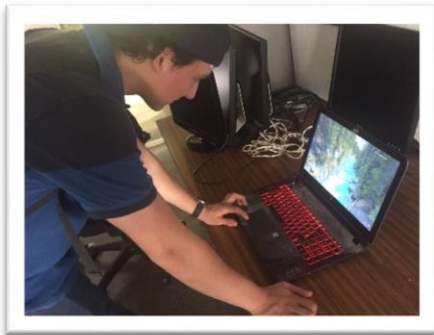
Con estas implementaciones consideramos que el usuario tendrá los resultados mas certeros posibles.

## 7. Prototipo de baja fidelidad



## 8. Script resumido de las entrevistas

IMAGEN	RESPUESTA
	<p>José Ayala Ing. Eléctrica</p> <p>"El programa ha implementar claramente es algo único, donde puede tener un éxito rotundo en el mercado, ya que es algo que la gente viene pidiendo de hace ratos".</p>



Rodolfo Asturias  
Ing. Mecatrónica

“Interesante idea... esto en realidad se ve un poco complicado de implementar, sin embargo, la idea esta muy impresionante. Me gustaría ser de los primeros en probar su beta”



Josué JumpStonick  
Ing. CCTI

“Esta idea es el primero que veo en la carrera, muy interesante que hayan escogido esto ya que es útil para los usuarios, suerte...”



Francisco Lindor  
Ing. Mecatrónica

“Creo que es una implementación muy interesante, ya que el algoritmo a implementar claramente se ve complejo, pero de igual forma que divertido se ve la programación, en especial porque esta en el lenguaje de Java”



Alberto Guzmán  
Ing. Alimentos

“Que cool su idea de proyecto, se ve super interesante y complicado de hacer, pero me gustaría cuando lo saquen al público, porque me encanta ver películas, felicidades y suerte”

## 9. Requisitos funcionales

- Ingresar con usuario y contraseña
- Tener la opción de cambiar la contraseña
- Poder seleccionar categorías
- E imprimir el resultado de las predicciones
- Ingresa películas

## 10. Planificación

FECHA	INICIO	FIN	TIEMPO INTERRUPCIÓN	DELTA TIEMPO	TAREA	COMENTARIOS
8 MAYO	6 p.m.	8 p.m.	20 min	100 min	DISEÑO ALGORITMO	
10 MAYO	6 p.m.	8 p.m.	20 min	100 min	IMPLEMENTACIÓN	
12 MAYO	6 p.m.	8 p.m.	20 min	100 min	IMPLEMENTACIÓN	
14 MAYO	6 p.m.	8 p.m.	20 min	100 min	REVISIÓN	

## 11.Pseudocódigo

- Se ingresa un usuario
- Se le pregunta cuáles son sus 3 categorías de películas favoritas.
- Se van a los nodos los cuales contienen los géneros de películas.
- Se eligen los 3 nodos con la categoría seleccionada.
- Y se muestra cada película relacionada a estos nodos de categorías
- Se realizaría con el Breadth-first search.
- Dado un vértice fuente  $s$ , explorará todos los vértices de  $G$  para “descubrir” todos los vértices alcanzables desde  $s$ .
- Calcula la distancia desde  $s$  a todos los vértices alcanzables.



- Después produce un árbol BF con raíz en  $s$  y que contiene a todos los vértices alcanzables.
- El camino desde  $dt$  a cada vértice en este recorrido contiene el mínimo número de vértices. Es el camino más corto medido en número de vértices.
- Su nombre se debe a que expande uniformemente la frontera entre lo descubierto y lo no descubierto. Llega a los nodos de distancia  $k$ , sólo tras haber llegado a todos los nodos a distancia  $k-1$ .

## 11. Investigación

### Algoritmo de agrupamiento

Un algoritmo de agrupamiento es un procedimiento de agrupación de una serie de vectores de acuerdo con un criterio, la similitud. La cercanía se define en términos de una determinada función de distancia. La medida más utilizada para medir la similitud entre los casos es la matriz de correlación entre los  $n \times n$  casos. Generalmente, los vectores de un mismo grupo comparten propiedades comunes. El conocimiento de los grupos puede permitir una descripción sintética de un conjunto de datos multidimensional complejo. Esta descripción sintética se consigue sustituyendo la descripción de todos los elementos de un grupo por la de un representante característico del mismo.

### $k$ vecinos más cercanos

El método de los  $k$  vecinos más cercanos es un método de clasificación supervisada que sirve para estimar la función de densidad  $F(x/C_j)$  de las predictoras  $x$  por cada clase  $C_j$ . Este es un método de clasificación no paramétrico, que estima el valor de la función de densidad de probabilidad o directamente la probabilidad a posteriori de que un elemento  $x$  pertenece a la clase  $C_j$  a partir de la información proporcionada por el conjunto de prototipos. En el proceso de aprendizaje no se hace ninguna suposición acerca de la distribución de las variables predictoras.

### Algoritmo de eliminación de candidatos

El algoritmo de eliminación de candidatos es utilizado dentro del ámbito de la inteligencia artificial. Su uso se engloba en la búsqueda de hipótesis o conceptos en él dado un conjunto de ejemplos de entrenamiento. Este algoritmo se puede expresar como una extensión del algoritmo find-s. El conjunto de ejemplos deberá estar conformado por una serie de tuplas de valores, cada uno de ellos denominados atributos. Adicionalmente uno de los atributos ha de ser de tipo binario, el cual es el atributo objetivo que clasificar que diferencia el concepto. De esta forma el algoritmo trata de obtener un intervalo de hipótesis flanqueadas por un límite más específico y un límite más general de forma que las nuevas instancias se consideran consistentes con el intervalo aprendido si están dentro de él.

## 12. Funcionamiento del sistema

El sistema de recomendaciones funciona mediante un grafo que consta de  $n$  películas relacionadas entre sí que varían dependiendo de los gustos seleccionados por las personas para ello utilizaremos el algoritmo de  $k$  vecinos más cercanos para poder revisar qué películas parecidas concuerdan con los gustos del usuario para poder encontrar emparejamientos de una manera más rápida.

## 13. Link del repositorio

<https://github.com/CarlosChew/Proyecto-2-ADT>