1

Práctica 1: Toma de decisiones

1.1 Introducción

En nuestro día a día nos enfrentamos a la toma de decisiones, por ejemplo, al elegir las palomitas en el cine o elegir un destino para las vacaciones de verano. La ciencia de datos se nutre de la ingeniería y el procesamiento de los datos, junto con la minería de datos, para realizar una toma de decisiones basada en los datos que tenemos. Por lo tanto, podemos definir la toma de decisiones como:

"El proceso de pensamiento de seleccionar una opción lógica de las opciones disponibles."

Cuando se intenta tomar una buena decisión, una persona debe sopesar los aspectos positivos y negativos de cada opción, y considerar todas las alternativas. Para una toma de decisiones efectiva, una persona debe ser capaz de predecir el resultado de cada opción también, y en base a todos estos elementos, determinar qué opción es la lo mejor para esa situación particular. Esta definición es demasiado ideal y poco real. Es más matemática. Por tanto podemos dar otra definición:

"La toma de decisiones puede considerarse como una actividad de resolución de problemas determinada por una solución considerada óptima, o al menos satisfactoria. Por lo tanto, es un proceso que puede ser más o menos racional o irracional y puede basarse en conocimientos y creencias explícitas o tácticas."

1.2 El proceso de toma de decisiones

El proceso de toma de decisiones se desarrolla en las siguientes etapas:

1. Definición y análisis del problema.
2. Desarrollar soluciones alternativas.
3. Evaluación de las soluciones alternativas.
4. Seleccionar la mejor solución.
5. Implementar la decisión.
6. Se analizan los resultados.

Aquí, el mayor desafío está en escoger la mejor solución. Para ello trataremos de buscar una forma de valorar cada alternativa y elegir la mejor de ellas.

1.2.1 El problema de la elección.

Esta etapa trata de comparar todas las posibles soluciones que hemos seleccionado en etapas anteriores para realizar un filtrado y quedarnos con la solución óptima para nuestro problema. Para ello, realizamos una evaluación de alternativas, seleccionando la alternativa que prefiramos por encima de las otras, o dicho de otro modo aplicando la regla de elección con preferencias:

$$RE(A, \succsim) = \{a' \in A | a' \succsim a, \forall a \in A\} \tag{1.1}$$

Donde A es el conjunto de alternativas que tenemos y $\succsim$ es la relación de equivalencia para comparar dos elementos. La relación de equivalencia es la que se encarga de establecer si una alternativa es mejor que otra, para ello, tendremos que construir una función para establecer lo útil que es una alternativa.

1.2.2 Utilidad

Podemos definir la utilidad desde dos perspectivas diferentes:

- Perspectiva de Marshall: La utilidad es la funcionalidad de una alternativa. Se estima comparando los atributos de una alterativa con las preferencias del sujeto. Además, utiliza el concepto de utilidad estándar.

- Perspectiva de Bentham: Satisfacción que se espera obtener de una alternativa. Esta perspectiva utiliza el concepto original de utilidad.

Para construir la función de utilidad seguimos los siguientes pasos:

1. Identificar los atributos relevantes de las alternativas.
2. Identificar los valores de los atributos anteriores.
3. Con lo anterior establecer la función.

Teniendo en cuenta lo anterior, podemos dar infinidad de funciones de utilidad. Una de ellas, por simplicidad, es usar una funcion linear:

$$u(c, a) = \sum_{k}^{K} \beta_{c,k} \cdot x_{a,k},$$

donde c es un individuo, a es una alternativa, K es el número de atributos, $x_{a,k}$ es el valor del atributo k, y $\beta_{c,k}$ es la ponderacion del atributo k. Esta función de utilidad variará dependiendo del usuario.

Example 1.2.1. *Queremos comprarnos una camiseta, de entre todas las camisetas de una tienda queremos elegir una. Para nosotros, hay dos factores a tener en cuenta, lo que nos guste su diseño, y su precio. Para simplificar los calculos mediremos el diseño de 0 a 1, y el precio de 0 a 1, siendo 0 muy caro, y 1 muy barato. Nosotros, no estamos en un buen momento economico, por tanto, para nosotros influye más el precio que su diseño, pongamos 60-40. Con esta información podemos construir en este caso la función de utilidad:*

$$u(a) = 0.6 \cdot x_{a,1} + 0.4 \cdot x_{a,2},$$

donde $x_{a,1}$ es el valor del precio, y $x_{a,2}$ es el valor de su diseño. Ahora, tenemos dos camisetas para elegir, la primera, nos gusta su diseño un 0.5, y su precio es 0.6. La segunda, su diseño le damos un 0.7, y un precio de 0.4. Ahora, para elegir con cual nos quedamos calculamos la utilidad de cada una.

$$u(camiseta\,1) = 0.6 \cdot 0.6 + 0.4 \cdot 0.5 = 0.56$$

$$u(camiseta\,2) = 0.6 \cdot 0.4 + 0.4 \cdot 0.7 = 0.52$$

Por tanto, nos quedaremos con la primera.

Con todo esto, podemos definir la relación entre dos alternativas a y b como $a \succsim b \iff u(a) \geq u(b)$, y con ello, podemos redefinir la ecuacion 1.1 como:

$$RE(A, \succsim) = \{a \in A | u(a') \geq u(a), \forall a \in A\} \tag{1.2}$$

1.2.3 Teoría de la utilidad esperada.

La teoría de la utilidad esperada surge para responder preguntas como: ¿Hay completa seguridad de predecir el resultado/utilidad una vez tomada una decisión? ¿Es posible que no siempre se obtenga el mismo resultado/utilidad?

En estos casos se usa lo que se conoce como *utilidad esperada*. La utilidad esperada $\hat{u}$ se calcula en base a los posibles resultados de una decision. Cada resultado r tiene una probabilidad de ocurrencia y una utilidad asociada. Por tanto, la utilidad esperada $\hat{u}$ asociada a la elección de la alternativa a se calcula:

$$\hat{u}(c, a) = \sum_r P(Result = r) \cdot u(c, a) \tag{1.3}$$

Con esto podemos definir el principio MEU (Maximum Expected Utility):

$$RE(A, \succsim) = \{a \in A | \hat{u}(a') \geq \hat{u}(a), \forall a \in A\} \tag{1.4}$$

La **aversión al riesgo** es la tendencia de las personas a preferir los resultados con poca incertidumbre a los resultados con mucha incertidumbre, incluso si el resultado medio de estos últimos es igual o superior en valor monetario al resultado más seguro. La aversión al riesgo explica la inclinación a aceptar una situación con una retribución más predecible, pero posiblemente inferior, antes que otra situación con una retribución muy impredecible, pero posiblemente superior. Existen muchas fórmulas para expresar la aversión al riesgo, nosotros usaremos la siguiente:

$$Risk\,aversion(x) = -\frac{u''(x)}{u'(x)}, \tag{1.5}$$

donde u' y u'' son las derivadas primera y segunda respectivamente.

1.3 Decisión Multiobjetivo

1.3.1 Introducción

El proceso de toma de decisiones multicriterio nace de la necesidad de solucionar un problema, es decir, una situación en conflicto en la se pretende cubrir simultáneamente varios objetivos o criterios.

En términos generales un problema multicriterio se puede plantear de la siguiente manera:

$$\text{Optimizar } F(x) \quad \text{con } x \in X \tag{1.6}$$

donde:

- x es el vector $[x_1, x_2, \ldots, x_n]$ de las *variables* de decisión. El problema de decisión es el de asignar los "mejores" valores a las variables.
- X es la denominada *región factible* del problema (el conjunto de posibles valores que pueden tomar las variables).
- $F(x)$ es el vector $[f_1(x), f_2(x), \ldots, f_p(x)]$ de las p *funciones objetivo* que recogen los criterios u objetivos simultáneos del problema. Dicha función objetivo puede maximizarse o minimizarse según sea la situación que se quiera lograr.

Para que un problema pueda considerarse como un problema de toma de decisiones con múltiples criterios se deben dar las siguientes condiciones:

1. Mínimo dos criterios de evaluación para valorar las alternativas en función de sus atributos. La condición suficiente es que los criterios estén en conflicto.
2. Al menos dos alternativas de decisión.
3. Un sistema de relaciones que permita asignar a cada alternativa ciertas medidas o atributos. Importantes para la posterior aceptación o rechazo de las alternativas disponibles.
4. Un conjunto de requerimientos de información de entrada que se obtendrán del decisor, y esto implica una metodología apropiada.

La clasificación de los problemas multicriterio está basada en la descripción del tipo de alternativas o posibles soluciones con las que cuenta el decisor o grupo decisor. A continuación, se explica la importancia del enfoque de toma de decisiones multicriterio y se cita un ejemplo para mayor claridad.

1.3.2 Concepto de soluciones no inferiores y región de Pareto

En los problemas de un solo objetivo, la meta de la solución es identificar la solución óptima: la solución factible (o soluciones) que proporciona el mejor valor de la función objetivo. Nótese que, incluso cuando existen óptimos alternativos, el valor óptimo de la función objetivo es único. Esta noción de optimalidad debe abandonarse en los problemas con múltiples objetivos porque una solución que maximiza un objetivo no maximizará, en general, ninguno de los otros objetivos. Lo que es óptimo con respecto a uno de los p objetivos suele no ser óptimo para los otros $p-1$ objetivos.

La optimalidad juega un papel importante en la resolución de problemas de un solo objetivo. Permite al analista y a los responsables de la toma de decisiones limitar su atención a una sola solución o a un subconjunto muy reducido de soluciones dentro de un conjunto mucho mayor de soluciones factibles. Para los problemas con múltiples objetivos, se introduce un nuevo concepto llamado *no inferioridad*, que cumple un propósito similar pero menos restrictivo. Así pues, nos centraremos en soluciones no inferiores.

La idea de la no inferioridad es muy similar al concepto de *dominancia*. A la no inferioridad se la denomina "no dominancia" en algunos textos de programación matemática, "eficiencia" en otros (incluidos estadísticos y economistas) y "optimalidad de Pareto" entre los economistas del bienestar. En la imagen A de 1.1 vemos un ejemplo en dos dimensiones. En este caso se quieren maximizar dos funciones objetivo. Es este caso F se le llama área factible, que corresponde a los posibles puntos elegibles. En esta region, vemos que cuanto más avancemos arriba a la derecha más valdrán las dos funciones objetivos. Sin embargo, cuando lleguemos a la *Fontera de Pareto*, que sean los puntos en los que al cambiar de uno a otro no se puede aumentar el valor de una función objetivo sin disminuir otra. Igual con el caso B.

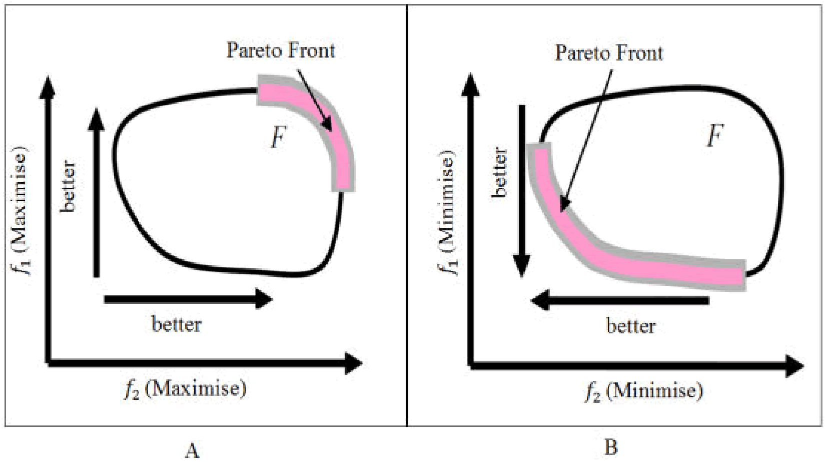

Figure 1.1: Ejemplo frente de pareto

Example 1.3.1. *Una compañía del sector automovilístico posee dos plantas para producir un mismo modelo de coches, la planta en Valencia* (x) *y la planta en Zaragoza* (y)*. Se desea saber en qué planta se fabricará el nuevo modelo de coche que ofrezca un mayor beneficio y cuya producción tenga un menor impacto ambiental.*

El beneficio que se obtiene es:

$$B = 2x + 5y$$

El impacto ambiental total será:

$$I = x + y$$

Las restricciones técnicas están representadas por:

$$\begin{aligned} 2x + y &\leq 95 \\ 40 \leq x + 2y &\leq 55 \\ 5 \leq x &\leq 40 \\ 0 \leq y &\leq 20 \end{aligned}$$

1.4 Teoria de juegos

La Teoría de Juegos es una teoría matemática que estudia las características generales de las situaciones competitivas de manera formal y abstracta. Es útil para tomar decisiones en casos donde dos o mas personas que deciden se enfrentan en un conflicto de intereses. Esta teoria se ha convertido en una valiosa herramienta para analizar situaciones económicas, políticas, y sociales, mediante modelos tomados de los juegos de estrategia. De Hecho, se ha mejorado la descripción mecánica inspirada en las ciencias físicas, incorporando a los modelos la conducta estratégica de los agentes, que cooperan o compiten para lograr sus objetivos. Robert J. Aumann define la teoría de juegos como:*"La teoría de juegos es la toma de decisiones óptimas en presencia de otras personas con objetivos diferentes"*

Tradicionalmente la Teoría de Juegos clásica se ha dividido en dos ramas: Teoría Cooperativa y No Cooperativa. Los juegos cooperativos tienen un papel normativo, buscan los resultados “equitativos”, “justos” que conseguirían agentes “racionales” y “bien informados”. Los diversos conceptos de solución que hay en la teoría que los estudia se establecen con conjuntos de axiomas que responden a una forma de entender esas propiedades de racionalidad, justicia y equidad. Los juegos no cooperativos, en cambio, son un marco teórico adecuado para estudiar si hay una “ley” interna en el conflicto que se estudia, y pueden resultar un importante instrumento de análisis razón por la que nos centraremos en los juegos no cooperativos. La Teoría de Juegos No Cooperativa asume que no hay lugar para comunicación, correlación o acuerdos entre los jugadores, de no ser explícitamente estipulados por las reglas del juego.

Un juego es un modelo estático que describe situaciones interactivas entre varios jugadores. Según este modelo, todos los jugadores tomarán decisiones a la vez e independientemente. Los juegos estan caracterizados por una interdependencia estratégica, gobernada por reglas y con un resultado definido. Los jugadores que participen en el juego van a ser racionales, es decir, van a buscar siempre su bien común siguiendo las reglas. La solución de un juego debería indicar a cada jugador qué resultado esperar y cómo alcanzarlo. Por tanto, un juego va a consistir en:

1. Un conjunto J de jugadores.

2. Un conjunto S_j de estrategias para cada jugador.

3. una función de pagos f que asocia cada elección de estrategias a un

valor.

Nosotros nos centraremos en los juegos con 2 jugadores. Denotaremos como I_i a las estrategias del jugador 1, y como II_j a las estrategias del jugador 2. El juego permite determinar la valoracion que ocasiona el que cada jugador utilice una de sus estrategias, que denotaremos por $v(i,j) = (v_I(i,j), v_{II}(i,j))$, donde $v_I(i,j)$ y $v_{II}(i,j)$ son los pagos para los jugadores 1 y 2 respectivamente. Al variar i y j podremos agrupar los valores una matriz. En el caso de que $v_I(i,j) = -v_{II}(i,j)$ se denomina que el juego es de suma nula.

1.4.1 Equilibrio de Nash

Una solución de estos juegos especifica las estrategias óptimas que jugadores racionales usarán y el pago que se obtiene con ellas. La solución o soluciones de un juego bipersonal de suma nula pueden caracterizarse de dos formas: mediante entrategias de seguridad y el concepto de punto de equilibrio.

En los juegos de suma nula un jugador intenta maximizar su pago, a la vez está intentando minimizar el pago de su oponente. Cada jugador considera el peor resultado que puede conseguir con cada una de sus estrategias y después escoge la estrategia que le proporciona el mejor de los peores resultados.

Definiremos el valor maximin del jugador I como

$$v_I = \max_i v_I(I_i) = \max_i \min_j a_{ij}$$

y el valor minimax del jugador II como

$$v_{II} = \min_j v_{II}(I_i) = \min_j \max_i a_{ij}$$

En general vamos a tener que estos valores v_I y v_{II} son unicos, va a existir al menos una estrategia para cada jugador que sea igual a cada valor, y que $v_I \leq v_{II}$. Cuando $v_I = v_{II} = v$ se dirá que el juego tiene solución o punto de silla, y tendrá valor v.

Para los casos en los que $v_I < v_{II}$ se dice que el juego no tiene solución en estrategias puras. En esta clase de juegos, si un jugador descubre la estrategia del otro, este último puede salir muy perjudicado. Por tanto, lo ideal en este caso es mantener fuera del alcance del otro oponente la manera de elegir las estrategias. Una forma de conseguir esto es introduciendo una distribución de probabilidad en nuestro método de elección. A esto se llamará una estrategia mixta, es decir, a una distribución de probabilidad en el conjunto de estrategias puras. Y gracias al teorema del minimax tenemos la seguridad que el juego bipersonal finito tendrá solucion en soluciones mixtas.

1.5 Problemas

Problem 1.5.1. Dibuja una gráfica donde el eje X represente cantidad de dinero de 0 a 10 millones de euros y el eje Y la utilidad del dinero para ti.

- ¿Cómo representarías la función de utilidad para ti?
- Suponte que ahora tienes 0 euros en la cuenta. ¿Cambiarías la función de utilidad?
- ¿Y si tuvieses 1 millón de euros?¿Cómo sería la función?

Problem 1.5.2. Conisideremos los siguientes juegos:

Juego 1	
Opción A	**Opción B**
2.500 euros con probabilidad 0.33 2.400 euros con probabilidad 0.66 0 con probabilidad 0.01	2.400 euros seguros.

Juego 2	
Opción A	**Opción B**
2.500 euros con probabilidad 0.33 0 con probabilidad 0.67	2.400 euros con probabilidad 0.34 0 con probabilidad 0.66

Juego 3	
Opción A	**Opción B**
4.000 euros con probabilidad 0.8 0 con probabilidad 0.2	3.000 euros seguros.

Juego 4	
Opción A	**Opción B**
Tour 3 semanas por Europa con probabilidad del 0.5.	Tour 1 semana por Europa seguro.

Juego 5	
Opción A	**Opción B**
6.000 euros con probabilidad 0.45.	3.000 euros con probabilidad 0.9.

- Calcula la utilidad utilizando la teoría de la utilidad esperada. ¿Coinciden tus resultados con los resultados teóricos?¿Por qué?

- Lee el paper de Kahneman y Tversky donde explican su teoría de por qué la teoría de la utilidad esperada no siempre funciona correctamente. Los juegos planteados aquí se corresponden con los juegos que plantean en su paper. Fíjate si tus resultados coinciden con la mayoría de encuestados por estos autores.

Problem 1.5.3. Una tienda de abarrotes recibe su suministro semanal de huevos cada jueves por la mañana. Este envío debe durar hasta el jueves siguiente, cuando llega un nuevo pedido. Cualquier huevo no vendido antes del jueves se descarta. Los huevos se venden a \$10 por cada cien y cuestan \$8 por cada cien. La demanda semanal de huevos en esta tienda varía de una semana a otra. Basándose en la experiencia pasada, se asigna la siguiente distribución de probabilidad a la demanda semanal:

Demanda (cientos de huevos)	10	11	12	13	14
Probabilidad	0.1	0.2	0.4	0.2	0.1

Este patrón de demanda se mantiene estable durante todo el año; la demanda de huevos no es estacional y la tendencia es plana. El problema es: ¿Cuántos huevos deben ordenarse para la entrega cada jueves?

Problem 1.5.4. La funcion de utilidad de Ana es $U = \sqrt{w}$, donde w es el dinero. Ella es dueña de una panadería, la cual el año que viene puede valer o bien 0 o bien 100 euros con las mismas posibilidades.

1. Supongamos que su empresa es el único activo que posee. ¿Cuál es el precio P más bajo al que aceptará vender su panadería?

2. Supongamos que tiene tiene 100 euros guardados ¿Como afectaría en este caso al caso anterior?

3. Compara tus resultados de las partes anteriores. ¿Cuál es la relación entre los ingresos de Ana y su aversión al riesgo?

Problem 1.5.5. Resuelve el siguiente problema de decisión multiobjetivo:

$$\begin{aligned}
\min \quad & f_1(x,y) = 3x^2 - 5y, \\
\max \quad & f_2(x,y) = 10x - y^2, \\
\min \quad & f_3(x,y) = 74 - x^2 - 4x - 5y,
\end{aligned}$$

sujeto a:

$$\begin{aligned}
& x + y \geq 15, \\
& x - y \leq 5, \\
& 2x + y \geq 20, \\
& 0 \leq x \leq 20, \\
& 0 \leq y \leq 20.
\end{aligned}$$

Problem 1.5.6. Eres el Científico de Datos de una importante empresa mayorista del sector turístico. La empresa tiene como negocio la creación de paquetes turísticos que luego vende a minoristas, agencias de viaje tanto online como offline. La empresa te pide lo siguiente:

Suponte que se ha hecho un estudio previo sobre toma de decisiones con el siguiente conjunto de atributos y valores:

- Calidad del paquete (Num. Estrellas del hotel): 2, 3, 4, y 5.
- Precio: Valores con rango desde 120 a 870 euros.
- Días del paquete: 5 o 7.
- Aspecto del producto: Alta o Baja.

Se elaboró un cuestionario con 8 situaciones de elección, con 3 paquetes en cada elección. Se preguntó a los encuestados que escogieran la elección más atractiva. Tanto los atributos de cada situación de elección como los resultados de elección para un conjunto de individuos se encuentran en un fichero Excel que se adjunta con el boletín. Se pide crear un modelo que dado 3 paquetes, que prediga cual va a ser el más atractivo.

Problem 1.5.7. Obtener el equilibrio de Nash en estrategias mixtas del juego de piedra, papel, o tijera.

2

Práctica 2: Estrategias de recomendación.

Uno de los objetivos de empresas como Amazon o Netflix es dar recomendaciones al usuario de productos o contenidos que le puedan interesar. Para ello se basan en sugerencias personalizadas, que proceden del análisis de los datos que realizan de sus usuarios. Estas sugerencias son complicadas de conseguir, pues requieren de tener en cuenta multitud de factores para que la propuesta sea interesante al usuario. Pero este problema no hay que resolverlo una única vez, sino que se debe realizar para cada usuario individual, lo que eleva la complejidad del mismo. Esto nos puede llevar a algunas estrategias de recomendación, como las de content-based, collaborative, o una solución híbrida entre ambos.

- Content-based: Se le recomiendan al usuario contenidos similares a los que le han gustado.

- Collaborative: Se le recomiendan al usuario contenidos similares a los que les gustan a sus semejantes.

El funcionamiento de estos sistemas se suelen basar en combinar información del usuario junto a información del contenido que queremos recomendar, aprendiendo las preferencias del usuario y recomendándole dichos contenidos.

2.1 Recomendación content-based

Nosotros partimos de un elemento $A = (a_i)_1^n$ con n atributos y de otro elemento $X = (x_i)_1^n$ del que querremos medir la similitud para poder recomendarlo.

2.1.1 Cálculo de utilidad en base a la distancia

En este caso para medir la similitud de dos elementos la haremos en base a la distancia de los atributos. Definimos la distancia como:

$$d(A, X) = \|A - X\| = \sqrt{\sum_{1}^{n}(a_i - x_i)} \tag{2.1}$$

Por tanto, la similitud entre A y X va a ser:

$$Similidud(A, X) = \frac{1}{1 + d(A, X)} \tag{2.2}$$

2.1.2 Cálculo de utilidad en base a la similaridad

En este caso mediremos la similitud en base del coseno que formen sus coordenadas:

$$Similitud = \cos\theta = \frac{A \cdot X}{\|A\|\|X\|} = \frac{\sum_1^n a_i \cdot x_i}{\sqrt{\sum_1^n a_i^2}\sqrt{\sum_1^n x_i^2}} \tag{2.3}$$

Resumen

Las estrategias que se llevaran a cabo para estas recomendaciones son:

1. Aprender el perfil del usuario (conjunto de términos clave preferentes).
2. Calcular la utilidad del item para el usuario.
3. Clasificar el item en base a una función umbral.

Pero la recomendación basada en contenido tiene ciertas limitaciones:

- **Usuario:** Las recomendaciones suelen ser productos muy similares a los ya consumidos por el usuario

- **Ingeniero:** Necesita conocer en detalle el dominio de la aplicación (productos, atributos, valores...)
- **Ambos:** Los valores de los atributos no aportan información acerca de la calidad del producto. Estimar la utilidad de un producto a través de sus atributos no garantiza la satisfacción o utilidad de la recomendación.

2.2 Recomendación colaborativa

La recomendación colaborativa es la más utilizada por empresas para sus recomendaciones, ya que se puede aplicar en multitud de productos y se entiende fácilmente. Se basa en utilizar el conocimiento común para recomendar items. Asume que los usuarios que tuvieron gustos similares en el pasado tendrán gustos similares en el futuro. Existen dos tipos de recomendación colaborativa: Basada en memoria y basada en modelo

2.2.1 Basada en memoria

En la recomendación colaborativa basada en la memoria, sólo se utiliza la matriz de interacción usuario-elemento para hacer nuevas recomendaciones a los usuarios. Todo el proceso se basa en las valoraciones e interacciones previas de los usuarios.

La recomendación colaborativa basada en la memoria de 3 métodos: la recomendación por *TOP-N*, la recomendación colaborativa basado en el usuario y la recomendación colaborativa basado en items.

Recomendación por *TOP-N*

Este es uno de los recomendadores más basicos que nos podemos encontrar. Se trata de recomendar a un usuario que no ha visto un item i los N items mejor valorados que no haya calificado.

Recomendación colaborativa basada en usuario

Dado un usuario u y un item p que no ha visto:

- Encontrar un conjunto de usuarios con los mismos gustos que u que hayan votado sobre p.

- Usar (por ejemplo) la media para predecir la puntuacion que le dará u al item p.

- Repetir para todo elemento que u no haya visto

Para ello, primero haremos uso de se la correlación de Pearson para medir la similitud entre dos usuarios a y b:

$$sim(a,b) = \frac{\sum_{i \in I}(r_{a,i} - \bar{r}_a)(r_{b,i} - \bar{r}_b)}{\sqrt{\sum_{i \in I}(r_{a,i} - \bar{r}_a)^2}\sqrt{\sum_{i \in I}(r_{b,i} - \bar{r}_b)^2}}, \tag{2.4}$$

donde I es el conjunto de items evaluados tanto por a como por b, $r_{k,i}$ es la valoriación de un usuario k al item i, y $\bar{r}_k$ la media de un usuario k. En este caso, la fórmula de predicción será:

$$pred(a,p) = \bar{r}_a + \frac{\sum_{k \in K} sim(a,k) \cdot (r_{k,p} - \bar{r}_k)}{\sum_{k \in K} sim(a,k)} \tag{2.5}$$

Existen una serie de recomendaciones a tener en cuenta al aplicar este tipo de cálculos:

- No todos los ratings de los vecinos deben tener el mismo valor.

  - Estar de acuerdo en items que les gustan a mucha gente no significa que estén de acuerdo en items más especializados.

  - Una posible solución a esto es darle más peso a los artículos con más varianza.

- Valorar el número de items co-valorados. Utilizar un "peso significativo", por ejemplo, reduciendo linealmente el peso cuando el número de covalorados es bajo.

- Amplificación de caso. Dar más peso a los vecinos "muy similares", pero ejemplo, donde la puntuación de similaridad es cercana a 1.

- Selección de vecinos. Usar threshold de similaridad o un número fijo de vecinos.

Recomendación colaborativa basada en items

La idea básica es usar la similaridad entre items para hacer predicciones. Produce mejores resultados filtrando item-to-item y las valoraciones son vistas como vectores en el espacio n-dimensional. La similitud es calculada basandose en el ángulo entre los vectores, mediante la fórmula del coseno previamente descrita en 2.3, dejando la ecuación para predecir un elemento como:

$$pred(a,p) = \frac{\sum_{i \in I(a)} sim(i,p) \cdot r_{a,i}}{\sum_{i \in I(a)} sim(i,p)}, \tag{2.6}$$

siendo en este caso i y p los vectores de las puntuaciones de los demas usuarios de los items i y p.

El tamaño de vecinos es típicamente acotado a un tamaño específico. No todos los vecinos se tienen en cuenta para la predicción. Un análisis del dataset MovieLense nos indica que una tamaño de vecinos entre 20 y 50 es adecuado para situaciones reales.

2.2.2 Basada en modelo

En el enfoque basado en modelos, se utilizan modelos de aprendizaje automático para predecir y clasificar las interacciones entre los usuarios y los elementos con los que aún no han interactuado. Estos modelos se entrenan utilizando la información de interacción ya disponible en la matriz de interacción mediante el despliegue de diferentes algoritmos como la factorización matricial, el deep learning, clustering, etc. Nosotros nos centraremos en la factorización matricial.

2.2.3 Factorización matricial

La factorización matricial es una forma de generar características latentes al multiplicar dos tipos diferentes de entidades. Los sistemas de recomendación es la aplicación de la factorización matricial para identificar la relación entre las valoraciones de los items y las de los usuarios. Con la información de las valoraciones de los usuarios sobre los items de la tienda, queremos predecir cómo valorarán los usuarios los items para que puedan obtener una recomendación basada en la predicción. Este enfoque ha demostrado ser especialmente efectivo en entornos donde la matriz de interacciones es dispersa, es decir, cuando la mayoría de los usuarios solo han interactuado con un pequeño subconjunto de items disponibles. La factorización matricial aplicada a sistemas de recomendación ha sido empleada con éxito en plataformas de streaming de contenido, comercio electrónico, redes sociales y otros servicios en línea, mejorando la precisión y la relevancia de las recomendaciones personalizadas para los usuarios. Este enfoque continúa siendo objeto de investigación y desarrollo, buscando constantemente maneras de mejorar la escalabilidad, la interpretabilidad y la adaptabilidad a diversas situaciones del mundo real.

La idea que reside detrás de la factorizacion matricial es obtener factorizar la matriz de las valoraciones como multiplicacion de dos matrices de menor dimesión. En el proceso de factorización matricial aplicada a sistemas de recomendación, se parte de una matriz de valoraciones, también conocida como matriz de utilidad, que registra las valoraciones o interacciones de los usuarios con los items. Esta matriz se descompone en dos matrices más pequeñas: una matriz de usuarios y otra de items. La idea central es representar cada usuario y cada item en un espacio de características latentes de dimensionalidad inferior. Cada fila en la matriz de usuarios y cada columna en la matriz de items corresponde a un vector en este espacio latente. La multiplicación de estas dos matrices proporciona una estimación de la matriz original, permitiendo así prever las interacciones faltantes. La optimización del proceso de factorización busca minimizar la diferencia entre las valoraciones reales y las predichas, ajustando los parámetros de las matrices de usuarios y items. Este enfoque captura patrones latentes no explícitos en los datos originales, permitiendo realizar recomendaciones personalizadas al identificar similitudes entre usuarios y items en el espacio latente, incluso cuando las interacciones observadas son limitadas. Este método ofrece una aproximación eficiente y efectiva para enfrentar el desafío de la recomendación personalizada en entornos con grandes conjuntos de datos.

2.3 Problemas

Problem 2.3.1. Tutorial de “recommenderlab” con el dataset Jester5k.

- Descarga el tutorial sobre la librería “recommenderlab”. Enlace: `https://cran.rproject.org/web/packages/recommenderlab/vignettes/recommenderlab.pdf`
- Sección 5.4. Inspecciona las propiedades del datasetJester5k.
- Crea un imagen de las 100 primeras filas y columnas de la matriz con el comando: image(Jester5k[1:100,1:100]). ¿Qué usuarios son los qué aportan más ratings?¿Qué chistes han sido más valorados?
- Sección 5.5. Crea un recomendador siguiendo las indicaciones del tutorial. Crea otra versión del recomendador sólo con los 100 primeros usuarios y vuelve a predecir para los usuarios 1001 y 1002. ¿Detectas diferencias en las recomendaciones?
- Secciones 5.6 y 5.7. Evalúa las recomendaciones siguiendo el tutorial. ¿Qué diferencia hay entre la evaluación de la sección 5.6 y la 5.7? ¿Cuál te parece más apropiada para un sistema de recomendación?
- Sección 5.8. Compara los diferentes métodos de recomendación en esta sección. ¿Cómo quedaría el algoritmo de recomendación ítem-based en comparación con los utilizados en esta sección?¿Cambian los resultados si modificamos el particionamiento utilizado en el procedimiento de evaluación?

Problem 2.3.2. Recomendación de películas con el dataset Movielense:

- Lee la descripción del dataset Movielens en la descripción de la librería “recommenderlab”(Pg. 14). Enlace: `https://cran.r-project.org/web/packages/recommenderlab/recommenderlab.pdf`
- Carga ahora el dataset Movielense:data(MovieLense).
- Inspecciona las propiedades de este dataset. ¿Qué diferencias encuentras respecto al dataset Jester5k?
- Realiza un estudio y comparación similar al del ejercicio anterior