

Aprendizaje, hallazgo y satisfacción en un sistema de recomendación en turismo.

Jasur Ismoilov¹, Fernando Sánchez-Vilas¹, Fabián P.Lousame¹, Eduardo Sánchez² y Manuel Lama²

Dept. de Electrónica y Computación. Campus Vida

Universidad de Santiago de Compostela

Resumen:

En este artículo se presenta un sistema de recomendación de tapas y restaurantes validado en la III edición de Santiago(é)Tapas celebrada en el año 2010. Santiago(é)Tapas es un concurso gastronómico que se celebra en Santiago de Compostela y en el que 56 restaurantes compitieron por elaborar la tapa que más gustase al público.

Nuestro sistema de recomendación utiliza dos algoritmos diferentes para recomendar tapas y tres para la recomendación de restaurantes. Para la validación del recomendador se realiza una encuesta entre un conjunto de usuarios de prueba para determinar si las recomendaciones proporcionadas han sido originales, han supuesto un hallazgo y si les han satisfecho. Al contrario que la mayoría de las propuestas de sistemas de recomendación no evaluamos nuestro recomendador en términos de precisión de predicción de las valoraciones sino que lo hacemos atendiendo a la utilidad de las recomendaciones. Para ello proponemos una ordenación en 3 niveles de utilidad de los distintos casos atendiendo a su originalidad, el hallazgo que suponen y la satisfacción para el usuario. Los resultados obtenidos nos muestran que un 72% de las recomendaciones producen experiencias con una utilidad alta, un 22% de las recomendaciones producen experiencias con una utilidad media y tan solo un 6% de las recomendaciones la utilidad es baja.

1. Estudiante de doctorado
2. Profesor Doctor

Aprendizaje, hallazgo y satisfacción en sistema de recomendación en turismo.

1. Introducción

Los sistemas de recomendación son aplicaciones inteligentes que intentan facilitar a un usuario la selección de ítems relevantes de entre una gran cantidad de alternativas. Existen dos aproximaciones principales [ADOMAVICIUS 2005]. La primera, que recibe el nombre de basada en contenido describe tanto a los usuarios como a los ítems por medio de perfiles y recomendará a cada usuario aquellos ítems que coincidan con su perfil. La segunda recibe el nombre de filtrado colaborativo, CF por sus siglas en inglés, y compara los patrones de valoración de un conjunto de usuarios para ofrecer recomendaciones a un usuario dado. Para ello predicen la valoración que el usuario k asignará a su interacción con un determinado ítem i . Estas predicciones permitirán ordenar los ítems no consumidos todavía por los usuarios según el grado de satisfacción que se espera que experimente el usuario al consumir cada ítem. Finalmente se le recomendarán al usuario o bien el primer ítem de la lista o bien un número reducido de ítems que ocupen las primeras posiciones de la lista.

Los recomendadores tradicionales basados en CF utilizan un única valoración para capturar la satisfacción global del usuario en su interacción, estos recomendadores son conocidos como monocriterio, por contra los recomendadores multicriterio [ADOMAVICIUS 2007] trabajan sobre la idea que un única valoración es insuficiente y por lo tanto permiten al usuario valorar distintos aspectos del ítem con el fin de describir mejor la interacción y obtener una mayor precisión. En la línea de los recomendadores multicriterio en [LOUSAME 2010] se define el concepto de vista de ítem, siendo una vista el conjunto de componentes de un ítem que son relevantes para una determinada interacción de un usuario.

Aunque existe una gran diversidad de métricas para medir la utilidad de los sistemas de recomendación [HERLOCKER 2004], la mayoría de las aproximaciones actuales basadas en

filtrado colaborativo utilizan tan solo métricas de precisión de las predicciones para validar la calidad de las recomendaciones. Así, una de las más utilizadas es el Error Absoluto Medio, MAE por sus siglas en inglés, que se calcula como la media de las diferencias entre los valores predichos para las experiencias de un usuario con un ítem y las valoraciones que el usuario realmente asigna a esas experiencias.

Sin embargo el uso de estas métricas presenta una serie de problemas como ponen de manifiesto los autores de [MCNEE 2006] y [HERLOCKER 2004]. Así un sistema de recomendación que proponga recomendaciones muy precisas pero triviales y que el usuario ya consumiría por iniciativa propia o que proponga recomendaciones siempre muy parecidas que terminen por “aburrir” al usuario será de poca utilidad. Para intentar evitar estos problemas han surgido propuestas como [CELMA 2008], [MCNEE 2002], [WENG 2007] o [ZIEGLER 2005] donde los autores tienen en cuenta la novedad y el hallazgo que proporcionan las recomendaciones bien en la elaboración del algoritmo de recomendación, bien para proponer nuevas métricas de evaluación de algoritmos o bien para evaluar los algoritmos propuestos.

2. Objetivo

El objetivo de este artículo es proponer un sistema de recomendación de restaurantes y tapas que será utilizado en el concurso Santiago(é)Tapas para generar recomendaciones a un conjunto de usuarios de prueba. Santiago(é)Tapas es un concurso gastronómico real celebrado todos los años en Santiago de Compostela en el que los restaurantes participantes compiten por elaborar la tapa mejor valorada por los consumidores.

Finalmente para comprobar la calidad del recomendador se analizará su utilidad en función de la novedad, el hallazgo y la satisfacción que proporcionan las experiencias que se generan a partir de las recomendaciones sugeridas a los usuarios. Para poder hacer esta evaluación

asumimos la hipótesis de que aquellas recomendaciones originales son las más útiles siempre y cuando la experiencia sea buena, ya que suponen un aprendizaje por parte del usuario de nuevas preferencias, además también consideramos de utilidad alta aquellas experiencias originales y que el usuario no tuviese interés a priori en consumir aunque la valoración de la experiencia sea media ya que suponen un aprendizaje que el usuario en otro caso no adquiriría. Por el contrario consideraremos de utilidad baja aquellas recomendaciones con una valoración baja, porque resultan en una experiencia negativa para el usuario. En el resto de los casos consideramos que tienen una utilidad media. En la Tabla 1 se muestra la clasificación de los distintos casos

Tabla 1. Utilidad de las recomendaciones

Nivel de Utilidad	Casos
Alta	1.1 Original, no existía interés a priori, valoración alta 1.2 Original, no existía interés a priori, valoración media 1.3 Original, sí existía interés a priori, valoración alta
Media	2.1 Original, sí existía interés a priori, valoración media. 2.2 No Original, no existía interés a priori, valoración alta. 2.3 No original, no existía interés a priori, valoración media. 2.4 No original, sí existía interés a priori, valoración alta. 2.5 No original, sí existía interés a priori, valoración media.
Baja	3.1 Original, no existía interés a priori, valoración baja. 3.2 Original, sí existía interés a priori, valoración baja. 3.3 No original, sí existía interés a priori, valoración baja. 3.4 No original, no existía interés a priori, valoración baja.

3. Métodos

a. Algoritmos

En el recomendador de Santiago(é)Tapas se combinaron 5 algoritmos distintos. Dos para recomendar tapas y otros 3 para recomendar una determinada vista de un restaurante.

Recomendación de Tapas:

Los recomendadores utilizados para la recomendación de tapas son:

- **Bipolar Slope-1 Scheme:** Se trata de una aproximación muy sencilla, trabaja sobre la idea de que dada una valoración de un usuario para un ítem k , la valoración de ese usuario para otro ítem m no valorado puede predecirse sumando a la valoración de k una desviación que puede ser positiva o negativa. Esta desviación se obtiene como la diferencia entre las valoraciones medias de ambos ítems. Para una descripción más detallada de su funcionamiento puede consultar [LEMIRE 2005].
- **Similitud de usuario:** Se trata de un algoritmo tradicional de filtrado colaborativo basado en similitud de usuario. Estos algoritmos siguen una serie de pasos para generar una recomendación a un usuario: (1) Crean un vecindario para el usuario a recomendar, este vecindario estará compuesto por un conjunto de usuarios que han consumido un número mínimo de ítems en común con el usuario a recomendar, valorándolos de forma similar, (2) Predicen la valoración que el usuario a recomendar le dará a un ítem asumiendo que si los usuarios de su vecindario lo han valorado bien, su satisfacción también será alta y si por el contrario los usuarios del vecindario lo han valorado mal, la satisfacción del usuario a recomendar con ese ítem será

baja, (3) Devuelven una lista de recomendaciones con los K ítems con mejor predicción de valoración. Una descripción más detallada de su funcionamiento puede ser encontrada en [ADOMAVICIUS 2005].

Recomendación de Vistas

Para la recomendación de restaurantes nos basamos en el concepto de vista de ítem, desarrollado dentro de la línea de trabajo de nuestro grupo y presentado en [LOUSAME 2010]. El concepto de vista de ítem intenta abordar una serie de problemas detectados en muchos de los recomendadores tradicionales. Así la mayoría de algoritmos de recomendación basados en filtrado colaborativo utilizan una única valoración para caracterizar la interacción entre un usuario y un ítem y generan recomendaciones asumiendo que si dos usuarios valoran de forma similar un conjunto de ítems, sus gustos serán similares. Sin embargo esta aproximación presenta dos inconvenientes:

- Valoraciones iguales no tienen porque implicar que los gustos de los usuarios sean iguales ya que los usuarios pueden valorar un ítem de forma similar pero por razones distintas.

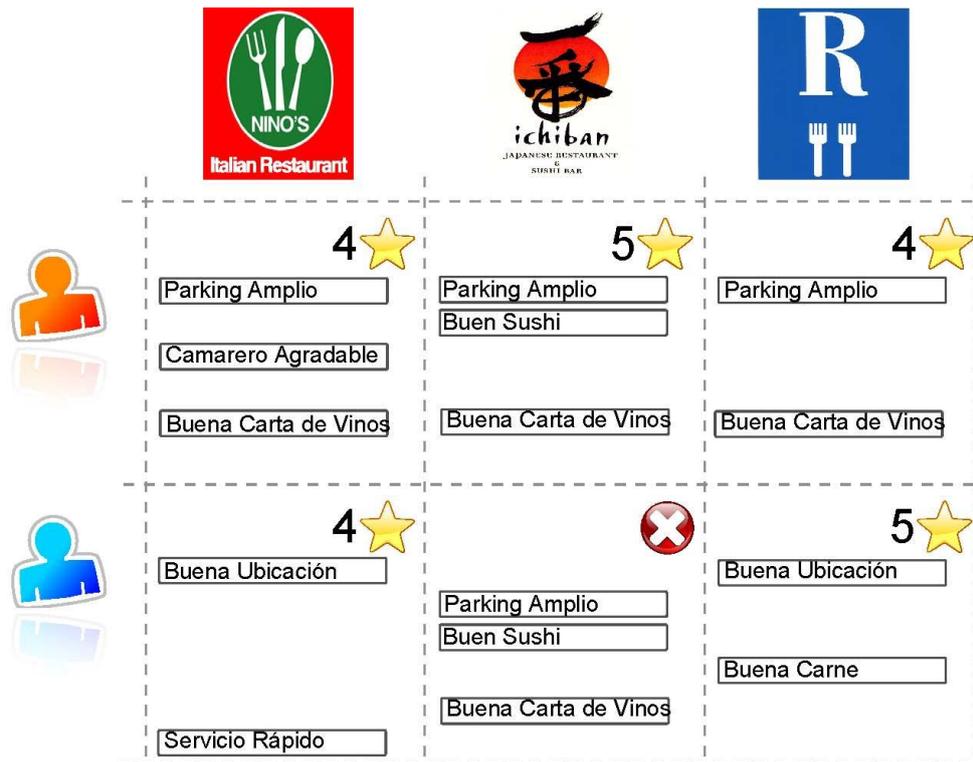


Figura 1. Ejemplo de usuarios que valoran de forma similar aunque tienen preferencias distintas. En este caso el Restaurante Japonés no debería ser recomendado al usuario 2

Imaginemos el caso que se muestra en la Figura 1, en el que el usuario 1 le da mucha importancia a que el restaurante tenga una buena selección de vinos y no le importa tener que coger el coche para desplazarse si el restaurante dispone de un buen parking. Sin embargo, al usuario 2 no le gusta conducir y tan solo quiere restaurantes que estén bien ubicados, además le gusta el restaurante italiano porque el servicio es rápido y el restaurante de lujo porque le encanta como preparan la carne. Si asumimos que valoraciones similares implican gustos similares sería lógico recomendar al usuario 2 el restaurante 2, sin embargo esta recomendación no sería acertada si a éste no le gusta ni sushi ni el vino, que es en lo que se basó el usuario 1 para valorar bien ese restaurante.

- La existencia de ítems configurables, es decir formados por un conjunto de componentes que pueden ser configurados y consumidos de forma diferente hace que la experiencia del usuario dependa de la configuración que esté

experimentado. Así en el dominio de la restauración hay distintos componentes configurables como por ejemplo, el plato elegido de entre todos los de la carta o la bebida seleccionada. En este escenario, parece más efectivo no solo recomendar un restaurante sino una determinada vista o configuración de ese restaurante. Sin embargo para poder hacer esto es necesario permitir al usuario evaluar además de al restaurante globalmente, también aquellas propiedades del restaurante que han influido en su experiencia, permitiendo así capturar diferencias entre experiencias que de otra manera permanecerían ocultas.

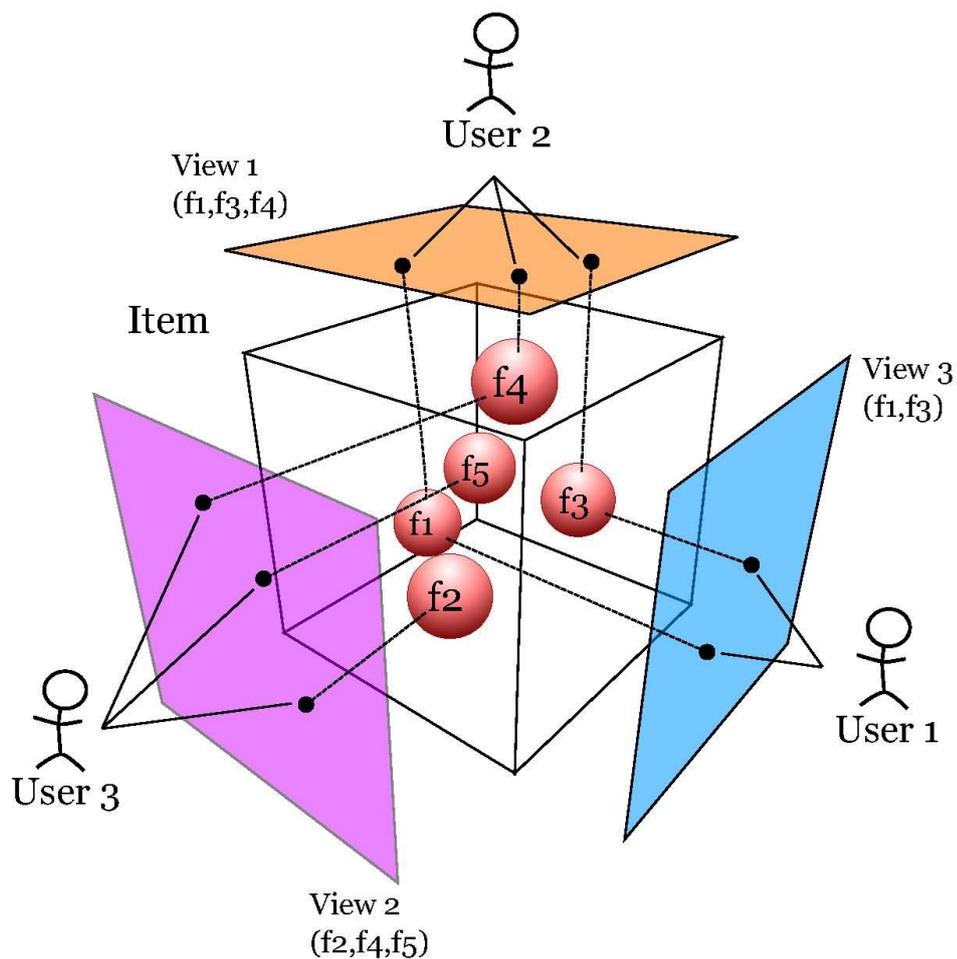


Figura 2. Vista de Ítem. El ítem se representa como un cubo y las distintas propiedades que lo conforman como esferas. Las vistas pueden ser interpretadas como cada una de las caras del cubo que experimenta cada usuario.

Como ya hemos comentado, los algoritmos que proponemos para la recomendación de restaurantes están basados en el concepto de vista de ítem. La vista de un ítem es el

conjunto de componentes de un ítem que serán relevantes para una determinada experiencia de un usuario. Dos usuarios distintos o incluso el mismo usuario en distintos momentos pueden experimentar el ítem de forma completamente distinta, bien porque las propiedades consumidas sean distintas, bien por las características del usuario que hacen que dé mayor relevancia a unas propiedades que a otras, o bien por el contexto del usuario que hace que éste dé más importancia a unas propiedades que a otras. Cada vista estará definida por un conjunto de propiedades y una colección de valoraciones, las propiedades definen los componentes del ítem que han sido relevantes para el usuario mientras que las valoraciones miden la satisfacción del usuario con ese componente. La Figura 2 muestra una representación conceptual de vista de un ítem. La figura representa a tres usuarios que están “mirando” a un ítem desde distintos lados del cubo. Así mientras que la interacción del usuario 3 es descrita por medio de la $Vista_2 = \{f_2, f_4, f_5\}$, la del usuario 2 está descrita por la $Vista_1 = \{f_1, f_3, f_4\}$.

De esta forma, la vista de un ítem vendrá determinada no sólo por el propio ítem sino también por la interacción usuario-ítem. Dos usuarios distintos o incluso el mismo usuario en distintos momentos y contextos, pueden consumir componentes distintas de un mismo ítem o dar distinta importancia a cada componente del ítem. Para poder recomendar al usuario la mejor vista, es necesario trabajar con valoraciones multicriterio, donde el usuario valorará el ítem de forma global y además un conjunto de propiedades del ítem. En el experimento realizado en Santiago(é)Tapas, se le pide al usuario que valore la tapa, el servicio y el ambiente del local. En el concurso cada restaurante podía ofrecer hasta un máximo de 3 tapas, que en este caso serán los elementos configurables del restaurante. Así y suponiendo un restaurante que tenga 3 tapas, al usuario se le podrán recomendar 4 vistas de ese restaurante, una por cada tapa

y una última sin ninguna tapa donde se le deja al usuario la opción de escoger la tapa que desee. Para generar las recomendaciones de las vistas en nuestro recomendador utilizamos tres algoritmos diferentes para predecir las funciones de agregación de las componentes del restaurante. Estas funciones de agregación permitirán predecir el peso que un usuario otorgará a cada uno de los elementos de un ítem todavía no consumido. Se pretende por tanto determinar cómo influyen los distintos elementos del ítem en la vista que experimentará el usuario:

- Aproximación de agregación con esquema de pesado basado de distancia inversa: En esta aproximación para obtener las funciones de agregación se trabaja con la idea de que aquellas componentes del ítem que estén más próximos a la valoración global tendrán una mayor importancia sobre ésta que las componentes que estén lejos. Puede consultarse [LOUSAME 2010] para más detalles sobre este algoritmo.
- Aproximación de agregación con esquema de pesado basado de distancia de soporte: Se trata de una variación de la anterior pero se definen un conjunto de funciones de agregación por defecto y se observa lo bien que se ajustan a las interacciones pasadas del usuario, aquellas que mejor se ajusten serán utilizadas para predecir las funciones de agregación futuras. Puede consultarse [SANCHEZ-VILAS 2011] para más detalles sobre este algoritmo.
- Similitud de usuario basado con filtrado de perfil de prioridad: Es una variación del algoritmo de filtrado colaborativo basado en usuario para incluir información de las valoraciones multicriterio. Para calcular la similitud entre usuarios en lugar de trabajar con sus valoraciones globales se trabaja con sus perfiles de prioridad. El perfil de prioridad de una interacción usuario-ítem determina la importancia que cada una de las componentes del restaurante tienen en la experiencia del usuario y

viene dado por los pesos de cada componente en la función de agregación. Así para ver cuán similares son dos usuarios se observa la similitud de sus funciones de agregación. Puede consultarse [SANCHEZ-VILAS 2011] para más detalles sobre este algoritmo.

b. Experimento

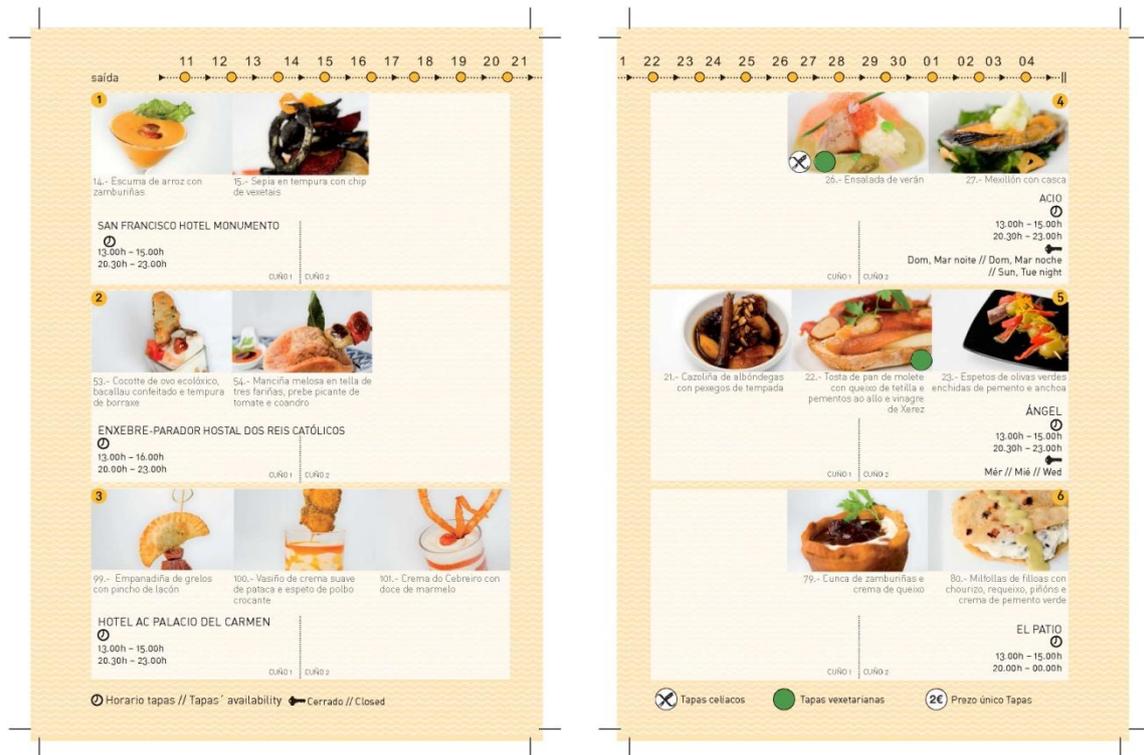
Para evaluar nuestro recomendador se realizó un experimento durante Santiago(é)Tapas, un concurso gastronómico celebrado en Santiago de Compostela todos los veranos. En el año 2010 tuvo lugar la tercera edición con un total de 56 restaurantes participantes que proponían y elaboraban hasta un máximo de tres tapas que eran vendidas a un precio común para todos los restaurantes de dos euros. El experimento fue diseñado para mantener el espíritu del concurso y se realizó de la siguiente manera:

The image displays two panels of the Santiago(é)Tapas menu, showing a grid of tapas items. Each item is numbered and includes a photograph, a title, a brief description, the restaurant name, and operating hours. The menu is organized into three rows on each panel, with two columns of items per row. A legend at the bottom explains the symbols used: a circle with an 'X' for gluten-free (Tapas celiacos), a green circle for vegetarian (Tapas vexetarianas), and a circle with '2€' for a unique price (Prezo único Tapas). The left panel shows items 1-3, and the right panel shows items 4-6. The top of each panel features a timeline of numbers from 1 to 21 (left) and 1 to 04 (right), representing the sequence of items.

Item	Restaurant	Operating Hours	Special Notes
1. Escuma de arroz con zamburñias	SAN FRANCISCO HOTEL MONUMENTO	13.00h - 15.00h 20.30h - 23.00h	CUÑO1 CUÑO2
2. Cocotte de ovo ecolóxico, bacallau confitado e tempura de borraaxe	ENXEBRE-PARADOR HOSTAL DOS REIS CATÓLICOS	13.00h - 16.00h 20.00h - 23.00h	CUÑO1 CUÑO2
3. Empanadiña de grelos con pincho de lácón	HOTEL AC PALACIO DEL CARMEN	13.00h - 15.00h 20.30h - 23.00h	CUÑO1 CUÑO2
4. Ensalada de verán	ACIO	13.00h - 15.00h 20.30h - 23.00h	CUÑO1 CUÑO2 Dom, Mar noite // Dom, Mar noite // Sun, Tue night
5. Espetos de olivas verdes enchidas de pimento e anchoa	ÁNGEL	13.00h - 15.00h 20.30h - 23.00h	Már // Miá // Wed
6. Milloilas de filloas con chourizo, queixo, piñón e crema de pemento verde	EL PATIO	13.00h - 15.00h 20.00h - 00.00h	CUÑO1 CUÑO2

Figura 3. El Tapasporte proporcionaba a los participantes las reglas del concurso y la descripción de las tapas

1. Nociones básicas sobre el concurso: A los participantes se les animaba a consumir y evaluar tapas a cambio de diversos regalos que se conseguían tras consumir un número mínimo de tapas. En la



2. Figura 3 se muestra el Tapasporte, una pequeña guía con varias funciones: (1) explicar las reglas del concurso a los participantes, (2) proporcionar la localización de los restaurantes y mostrar las tapas que se ofrecen en ellos, (3) permitir demostrar que un participante había visitado el número mínimo de restaurantes necesarios para conseguir alguno de los premios. Los empleados de los restaurantes debían cuñarlos en el lugar correspondiente para certificar que los dueños del Tapasporte habían consumido una tapa en su local.

ORGANIZA
TURISMO DE SANTIAGO DE COMPOSTELA
SANTIAGO(É)TAPAS
III CONCURSO DE TAPAS DE SANTIAGO DE COMPOSTELA

PUNTÚA A TAPA E DEPOSITA O TEU VOTO NA URNA
PUNTÚA LA TAPA Y DEPOSITA TU VOTO EN LA URNA
EVALUATE YOUR TAPA AND PLACE YOUR VOTE IN BALLET BOX

Tapa Nº//Nr. ↓ 0 1 2 3 4 5 ↑

Tapa Nome//Nombre//Name ↓ 0 1 2 3 4 5 ↑
Servizo/atención//Servicio/atención//
Service/treatment

↓ 0 1 2 3 4 5 ↑
Instalación/local//Instalaciones/local//
facilities/premises

↓ 0 1 2 3 4 5 ↑
experiencia global//overall experience

Nome&apellidos*//Nombre&apellidos*//Name&surname*
DNI/Pasaporte*//ID/Passport*
Lugar de procedencia*//Place of origin*
Tel./e-mail**

*A efectos de garantir a veracidade das votacións, só serán válidos os pousavos que teñan todos os datos que se solicitan do votante//A efectos de garantizar la veracidad de las votaciones, sólo serán válidos los pousavos que tengan todos los datos que se solicitan del votante//To guarantee accuracy of voting, only beer-mats with all requested data will be considered valid.
Se desexas participar no sorteo das Cristalarías Kwarx de cristal irrompible deberás ademais facilitar tel. e-mail, co fin de que podamos contactar contigo de ser o caso//Si desexas participar en el sorteo de las Cristalarías Kwarx de cristal irrompible deberás ademais facilitar tel. y e-mail, con el fin de que podamos contactar contigo de ser el caso//**To participate in prize draw to win Kwarx set of unbreakable glasses, you should also provide tel. and e-mail so that we can contact you.
Os datos son confidenciais e só serán usados para santiago@etapas
//Los datos son confidenciales y sólo serán usados para santiago@etapas//Your data is confidential and will only be used for santiago@etapas.
www.santiagoetapas.com

Figura 4. Voto de Tapa. Los participantes tenían que cubrir el voto para evaluar su experiencia en un restaurante.

- Obtención de las valoraciones de los usuarios: Tras consumir la tapa, los participantes debían evaluar su experiencia cubriendo el voto que se muestra en la Figura 4. Tenían que proporcionar 4 valoraciones entre 0 y 5: (i) una valoración global de la experiencia, (ii) una valoración de la tapa, (iii) una valoración del servicio proporcionado por el personal del restaurante y (iv) una valoración del ambiente y decoración del restaurante.

Del conjunto de participantes en el concurso se seleccionaron 100 para que evaluaran las recomendaciones que les proporcionamos. Las condiciones para convertirse en uno de estos usuarios de prueba fueron anunciadas en la página web de Santiago(é)Tapas. Los usuarios seleccionados tenían una semana y media para probar y evaluar un mínimo de 7 tapas. Después de ese tiempo se

recogieron y procesaron las valoraciones de todos los usuarios que participaron el concurso.

4. Generación de recomendaciones: La información recogida nos permitió generar 5 recomendaciones para cada usuario utilizando los 5 algoritmos descritos en la sección 3.a. En el caso concreto de que alguno de estos algoritmos produjese una recomendación ya proporcionada por otro algoritmo se seleccionaba la siguiente recomendación de su lista. Cuando alguno de estos recomendadores no era capaz de proporcionar una recomendación para un usuario dado, bien porque no había información disponible en la base de datos o bien porque este algoritmo no era capaz de generar una recomendación diferente a las ya proporcionadas por los otros, se recomendaban aquellos ítems con una mayor valoración media.

ID of participant: 9
Name&surname:

INSTRUCCIONES:

Special note for allergic people:

- Although we have filtered the tapas that could be problematic according to the information filled-in in your profile, be sure to check at the restaurant that you are eating tapas that do not have any ingredient that may be problematic for you.

Guidelines for consuming the recommendations:

- Some recommendations have a restaurant and a tapa, and in others, only the restaurant itself. In the former, you should taste the tapa that appears in the recommendation. In the latter, you can choose the tapa you want among the contest's tapas offered by the restaurant, and specify afterwards the id of the chosen tapa in the appropriate place provided in this questionnaire.
- There is no need to taste the tapas in the order that is listed in this document. The recommendations can be tasted in the order that fits you.

Guidelines for reimbursement:

- Address: Municipal Tourist Office located in the Plaza de Galicia.
- Documentation to be submitted:
 - Tickets de los restaurantes donde se han consumido las tapas.
 - This form, correctly filled in, where you are asked to evaluate some aspects of the recommendations and answer some general questions.

MAPA:



1. Indicate the order in which you visited the restaurants
Indicate the sequence of the letters A, C, D, ... according to your case.

2. What criteria did you follow in consuming the recommendations?

Route planning Better priority aspects Random choice Affinity with the local rate _____

3. Estás satisfecho con las recomendaciones?

No Little Quite Very much

In case if you are little or not at all satisfied with the recommendations, please indicate the reason:

Any comments:

RECOMENDACIONES:

A) CASA OURENSE:

RÚA DO HORREO, 172

OVERALL EXPERIENCE: 0 1 2 3 4 5

Tapa: Sorrentinos con prebe de chorizo: 0 1 2 3 4 5

Quality of Service: 0 1 2 3 4 5

Local: 0 1 2 3 4 5

Did you find this recommendation original? SI Non

Would you consume this recommendation spontaneously? SI Non

B) CALDERÓN:

RÚA DA CARREIRA DO CONDE, 8

OVERALL EXPERIENCE: 0 1 2 3 4 5

Tapa: Ostra con gin-tonic: 0 1 2 3 4 5

Quality of Service: 0 1 2 3 4 5

Local: 0 1 2 3 4 5

Did you find this recommendation original? SI Non

Would you consume this recommendation spontaneously? SI Non

C) MESÓN DE LÁZARO:

RÚA DE SAN LÁZARO, 73

OVERALL EXPERIENCE: 0 1 2 3 4 5

Tapa: Zamburifias con verduras confeitadas: 0 1 2 3 4 5

Quality of Service: 0 1 2 3 4 5

Local: 0 1 2 3 4 5

Did you find this recommendation original? SI Non

Would you consume this recommendation spontaneously? SI Non

D) CASTELAO:

RÚA NOVA DE ABAMO, 7

OVERALL EXPERIENCE: 0 1 2 3 4 5

Tapa: Hamburguesa de xoubriña: 0 1 2 3 4 5

Quality of Service: 0 1 2 3 4 5

Local: 0 1 2 3 4 5

Did you find this recommendation original? SI Non

Would you consume this recommendation spontaneously? SI Non

E) ACIO:

RÚA DAS GALERAS, 28

OVERALL EXPERIENCE: 0 1 2 3 4 5

Tapa: Mexillón con casca: 0 1 2 3 4 5

Quality of Service: 0 1 2 3 4 5

Local: 0 1 2 3 4 5

Did you find this recommendation original? SI Non

Would you consume this recommendation spontaneously? SI Non

Figura 5. Recomendaciones de restaurantes. Cada participante en el experimento debía descargar un archivo PDF del sitio web del concurso con las recomendaciones personalizadas

5. Evaluación de recomendaciones: El precio de las tapas recomendadas era devuelto a los usuarios pero solo una vez que estos hubiesen rellenado y entregado el formulario que se muestra en la **¡Error! No se encuentra el origen de la referencia.** donde se les pedía que proporcionasen cierta información sobre sus recomendaciones. En concreto se les pedía que evaluaran los restaurantes de la misma manera que lo hacían en el concurso y que proporcionasen información referente a: (i) originalidad de las recomendaciones, (ii) hallazgo que proporcionan las recomendaciones, (iii) el criterio que siguieron para consumir las recomendaciones, (iv) su satisfacción con las tapas recomendadas, (v) en caso de que no quedasen satisfechos con las recomendaciones, cuál era la razón y (vi) comentarios adicionales. En concreto, la valoración global de la experiencia y la información referente a las preguntas (i) y (ii) nos permitieron hacer la clasificación de las distintas recomendaciones generadas a los usuarios en función de los niveles de utilidad definidos en la sección 2. Además en la Tabla 2 se muestra la asociación entre las preguntas planteadas en el formulario y los tres parámetros (aprendizaje, hallazgo y utilidad) en los que nos basamos para definir la función de utilidad.

Tabla 2. Asociación entre preguntas y parámetros de utilidad

Pregunta	Parámetro
¿Esta recomendación te ha parecido original?	Aprendizaje
¿Consumirías esta recomendación por ti mismo?	Hallazgo
Valoración global de la experiencia	Satisfacción

Por último en la Tabla 3 se muestran algunos datos relevantes del experimento. El número total de recomendaciones evaluadas fue menor del esperado ya que un número importante de usuarios de prueba no pudo completar el experimento. El escaso tiempo disponible para completar todas las fases del experimento fue la principal razón de esto.

Tabla 3. Información del Experimento

Restaurantes participantes	56
Número de tapas diferentes ofrecidas	109
Tapas consumidas	35000
Participantes en el experimento	100
Participantes que completaron el experimento	18
Recomendaciones generadas	500
Recomendaciones validadas	81

4. Resultados

Gracias a las preguntas realizadas en la encuesta a los 18 participantes que han finalizado el proceso y trabajando con las 81 recomendaciones generadas, se ha estudiado cómo se comporta el recomendador en términos de utilidad. Para poder realizar este estudio consideraremos que una recomendación ha sido:

- Bien valorada si el usuario le asigna un 5 o 4 sobre 5
- Regular valorada si el usuario le asigna un 3 sobre 5
- Mal valorada si el usuario le asigna un 2,1 o 0 sobre 5

Realizamos inicialmente un estudio en términos de aprendizaje y hallazgo:

Caso 1: Tapa que el usuario no consumiría de forma espontánea y que le resulta original.

Este sería el mejor caso ya que en el caso de que la tapa sea del agrado del consumidor suele implicar satisfacción con el nuevo hallazgo y además un aprendizaje por parte del usuario y un aumento de sus preferencias.

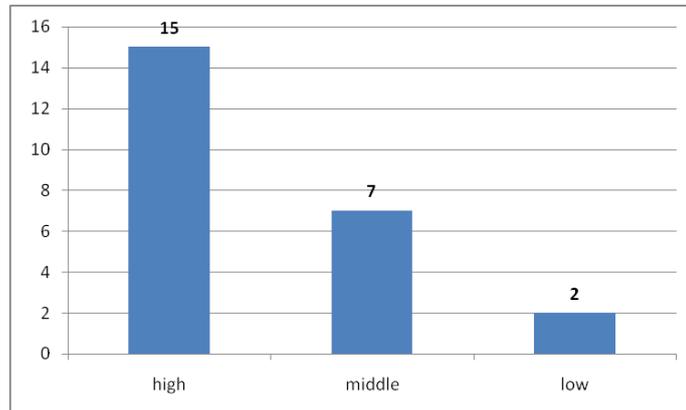


Figura 6. Valoraciones de las recomendaciones del caso 1

En la Figura 6 se puede observar que nuestro recomendador funciona muy bien para estos casos ya que la mayoría de las recomendaciones proporcionan una satisfacción elevada.

Caso 2: Tapa que el usuario no consumiría espontáneamente pero que no es original. Se trata de tapas con las que el usuario ya está familiarizado, pero que no consumiría por sí mismo, tienen menos valor que el caso anterior ya que aunque implica hallazgo no un aprendizaje y ampliación de las preferencias del usuario.

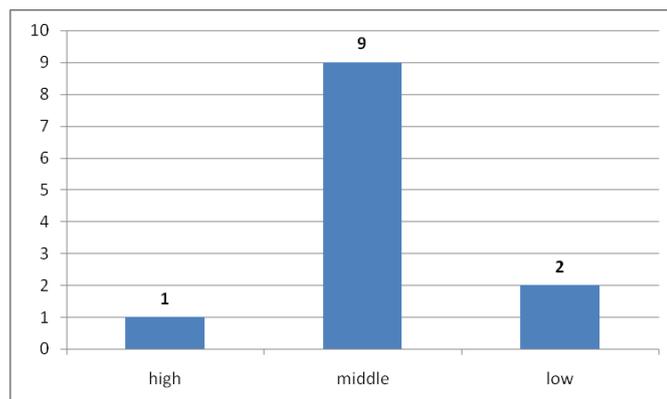


Figura 7. Valoraciones de las recomendaciones para el caso 2

Como vemos en la Figura 8 estos son los casos en los que peor se comporta el recomendador y que concuerda con el hecho de que es una tapa que el consumidor ya conocía pero que no consumiría, posiblemente porque o bien no le gusta o bien porque no le resulta atractiva.

Caso 3: Tapa que el usuario consumiría de forma espontánea y que es original. En este caso el valor de la recomendación es ligeramente superior al caso anterior ya que puede suponer aprendizaje y por tanto un aumento de las preferencias del usuario, pero inferior al primer caso ya que el usuario ya la consumiría por sí mismo.

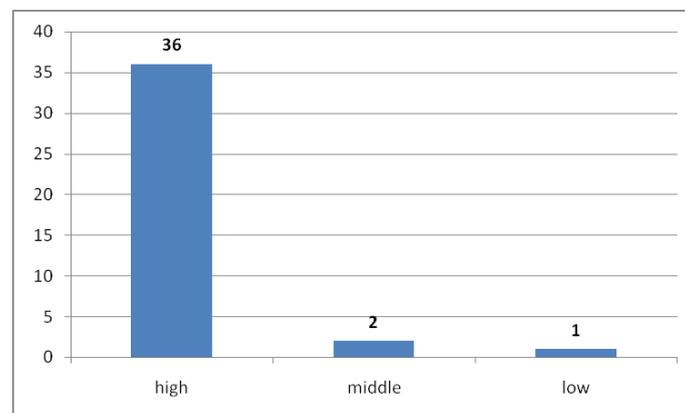


Figura 8. Valoraciones de las recomendaciones para el caso 3

Como puede observarse en la Figura 8 nuestro recomendador se comporta muy bien en estos casos.

Caso 4: Tapa que el usuario consumiría de forma espontánea y que no es original. Son las recomendaciones que menos valor aportan al no proporcionar ni hallazgo ni aprendizaje, pero pueden ser necesarias en las etapas iniciales para establecer la confianza del usuario en el recomendador al ver que el sistema recomienda tapas que el usuario sabe que le gustan.

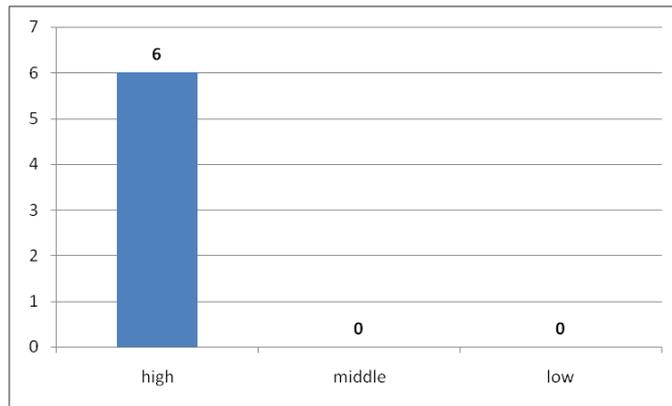


Figura 9. Valoraciones de las recomendaciones para el caso 4

Como se observa en la Figura 9 en estos casos nuestro recomendador funciona muy bien, ya que de las 6 recomendaciones generadas, todas fueron valoradas de forma positiva por el usuario, algo que concuerda con el hecho de que son tapas conocidas por el usuario y que consumiría por sí mismo.

Finalmente mostramos en la Figura 10 la utilidad del recomendador atendiendo a la clasificación presentada en la Tabla 1.

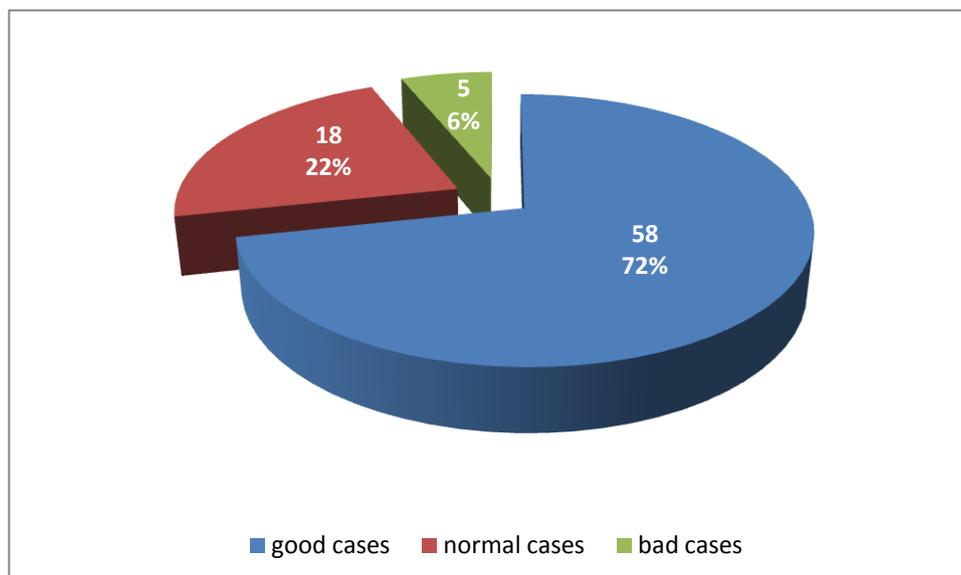


Figura 10. Utilidad de las recomendaciones

Como se puede observar, los resultados son muy buenos, con un 58% de casos que presentan una utilidad alta para el usuario y tan solo un 5% de utilidad baja.

Desglosando por casos:

Utilidad Alta:

Caso	Número
1.1 Original, no existía interés a priori, valoración alta	15
1.2 Original, no existía interés a priori, valoración media	7
1.3 Original, sí existía interés a priori, valoración alta	36

Utilidad Media:

Caso	Número
2.1 Original, sí existía interés a priori, valoración media.	2
2.2 No original, no existía interés a priori, valoración alta.	7
2.3 No original, no existía interés a priori, valoración media.	9
2.4 No original, sí existía interés a priori, valoración alta.	0
2.5 No original, sí existía interés a priori, valoración media.	0

Utilidad Baja:

Caso	Número
3.1 Original, no existía interés a priori, valoración baja.	2
3.2 Original, sí existía interés a priori, valoración baja.	1
3.3 No original, sí existía interés a priori, valoración baja.	2
3.4 No original, no existía interés a priori, valoración baja.	0

5. Discusión

Métricas como MAE, MSE permiten medir la exactitud con la que los sistemas de recomendación son capaces de predecir la valoración que va a asignar un determinado usuario a su interacción con un conjunto de ítems no consumidos. Es innegable el valor que aporta este tipo de métricas para medir la calidad de los sistemas de recomendación y permitir compararlos con otros. Sin embargo, si los datos son obtenidos en un espacio limitado de tiempo, en nuestro caso las tres semanas que duraba el concurso, se corre el riesgo de no tener en cuenta la evolución del usuario. Generalmente durante su interacción con el dominio y con el propio sistema de recomendación un usuario va experimentado cambios en sus preferencias. Así, en su toma de contacto con el dominio, el usuario buscará minimizar el riesgo y maximizar la satisfacción de las experiencias, sin embargo y a medida que va conociendo el dominio, el usuario suele estar dispuesto a arriesgarse un poco más con el fin de aprender y descubrir nuevas preferencias para evitar consumir siempre ítems del mismo tipo. De esta manera puede darse el caso de que un recomendador que funcionaba muy bien las primeras semanas pierda su utilidad a largo plazo ya que siempre proporciona recomendaciones muy similares y obvias, que aunque proporcionan una satisfacción para el usuario no implican ni aprendizaje ni hallazgo, provocando que el usuario abandone el sistema de recomendación, bien porque le aburren las recomendaciones o bien porque son recomendaciones que ya consumiría por sí mismo. Para minimizar este riesgo es necesario utilizar métricas de utilidad de las recomendaciones que incorporen el aprendizaje y el hallazgo que supone para el usuario las recomendaciones proporcionadas. En este artículo hemos presentado un sistema de recomendación validado durante el concurso Santiago(é)Tapas, y hemos analizado su comportamiento en base a la novedad, hallazgo y satisfacción que proporcionaban las recomendaciones a los usuarios, obteniendo unos resultados muy satisfactorios.

6. Agradecimientos

Este trabajo fue financiado por la Xunta de Galicia a través del proyecto 09TUR001E.

También queremos agradecer a SantiagoTurismo, y en concreto a Tono Mugico y Yolanda Ferro, por la gestión de la organización del concurso Santiago(é)Tapas y por sus valiosas sugerencias como expertos en turismo.

7. Referencias

[ADOMAVICIUS 2005] ADOMAVICIUS, G. AND TUZHILIN, A. 2005 Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Engineering.

[ADOMAVICIUS 2007] ADOMAVICIUS, G. AND KWON, Y. 2007 New recommendation techniques for multicriteria rating systems. IEEE Intelligent Systems.

[CELMA 2008] CELMA, O. AND HERRERA, P. 2008. A new approach to evaluating novel recommendations. In Proceedings of the 2008 ACM conference on Recommender systems (RecSys '08). New York.

[LEMIRE 2005] LEMIRE, F. AND MACLACHLAN A., 2005. Slope One Predictors for Online Rating-Based Collaborative Filtering. Proceedings of SIAM Data Mining.

[LOUSAME 2010] LOUSAME, F. AND SANCHEZ E. 2010 Multicriteria Predictors using Aggregation Functions based on Item Views. Proceedings of 10th International Conference on Intelligent Systems Design and Applications

[HERLOCKER 2004] HERLOCKER, J. L., KONSTAN, J. A., TERVEEN, L. G. AND RIEDL, J. T. 2004. Evaluating collaborative filtering recommender systems. In ACM Transactions on Information Systems. 22, 1

[MCNEE 2002] MCNEE, S. M., ALBERT, I., COSLEY, D., GOPALKRISHNAN, P., RASHID, A. M., KONSTAN, J. A., AND RIEDL, J. 2002. On the recommending of citations for research papers. In Proceedings of ACM CSCW 2002. New York.

[MCNEE 2006] MCNEE, S. M., RIEDL, J. AND KONSTAN, J. A. 2006. Being accurate is not enough: how accuracy metrics have hurt recommender systems. In Computer Human Interaction. Human factors in computing systems, New York.

[SANCHEZ-VILAS 2011] SANCHEZ-VILAS, F., ISMOILOV, J., LOUSAME, F., SANCHEZ, E., LAMA, M. 2011 Applying Multicriteria Algorithms to Restaurant Recommendation. In The 2011 IEEE/WIC/ACM International Conference on Web Intelligence [To be published]

[WENG 2007] WENG, L.T., XU, Y., LI, Y. AND NAYAK, R. 2007. Improving recommendation novelty based on topic taxonomy. In Proceedings of the IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, Washington.

[ZIEGLER 2005] ZIEGLER, C.-N., MCNEE, S. M. , KONSTAN, J. A. AND LAUSEN, G. 2005 Improving recommendation lists through topic diversification. In Proceedings of the 14th International Conference on World Wide Web, New York.