



Evaluación de interfaces de explicación en los sistemas de recomendación

Evaluation of Explanation Interfaces in Recommender Systems

Evaluação de Interfaces de Explicação em Sistemas de Recomendação

Sergio Cleger-Tamayo¹

Marcela Sávia Picanço-Bastos²

Marcella Sarah Figueiras-de-Farias³

Fecha de recepción: febrero 2017

Fecha de aceptación: abril 2017

Para citar este artículo: Cleger-Tamayo, S., Picanço-Bastos, M.S. y Figueiras de Farias, M.S. (2017). Evaluación de interfaces de explicación en los sistemas de recomendación. *Revista Científica*, 29 (2), 129-139. **Doi:** [10.14483/udistrital.jour.RC.2016.29.a2](https://doi.org/10.14483/udistrital.jour.RC.2016.29.a2)

Resumen

Las interfaces de explicación son una herramienta útil para los sistemas con una alta cantidad de contenido a ser evaluado por los usuarios. Las diferentes interfaces representan una ayuda para los usuarios indecisos o aquellos que consideran los sistemas una caja cerrada inteligente, a su vez, estos sistemas muestran recomendaciones a los usuarios basados en diferentes modelos. En el presente trabajo se manifiestan los diferentes objetivos perseguidos con las interfaces y algunos de los criterios que pudieran ser analizados, así como una propuesta de métricas para registrar resultados. Finalmente, se muestran los principales resultados de un estudio con usuarios reales y su interacción con sistemas de uso cotidiano. Dentro de las principales conclusiones destaca el impacto positivo en relación al tiempo de interacción con los aplicativos y la aceptación de las recomendaciones recibidas.

Palabras Clave: filtrado colaborativo, métricas de evaluación.

Abstract

Explaining interfaces become a useful tool in systems that have a lot of content to evaluate by users. The different interfaces represent a help for the undecided users or those who consider systems as boxed black smart. These systems present recommendations to users based on different learning models. In this paper, we present the different objectives of the explanation interfaces and some of the criteria that you can evaluate, as well as a proposal of metrics to obtain results in the experiments. Finally, we showed the main results of a study with real users and their interaction with e-commerce systems. Among the main results, highlight the positive impact in relation to the time of interaction with the applications and acceptance of the recommendations received.

Keywords: collaborative filtering, evaluation metrics.

¹. Instituto de Ensino Superior. Manaus, Brasil. Contacto: sergio.tamayo@fucapi.br

². Universidad del Estado de la Amazonas (UEA), Manaus, Brasil. Contacto: mspessoa@uea.edu.br

³. Instituto de Ensino Superior. Manaus, Brasil. Contacto: marcella.farias@fucapi.br

Resumo

As interfaces de explicação tornam-se uma ferramenta útil nos sistemas que tem muito conteúdo para ser avaliados pelos usuários. As diferentes interfaces representam uma ajuda para os usuários indecisos ou aqueles que consideram os sistemas como caixa fechada inteligentes. Estes sistemas apresentam recomendações aos usuários com base em diferentes modelos de aprendizagem. Neste trabalho, são apresentados os diferentes objetivos das interfaces de explicação e alguns dos critérios que podem ser avaliados, assim como uma proposta de métricas para obter resultados nas experimentações. Finalmente, são apresentados os principais resultados de um estudo com usuários reais e sua interação com sistemas de comércio eletrônico. Entre os principais resultados destacam o impacto positiva em relação ao tempo de interação com as aplicações e aceitação das recomendações recebidas.

Palavras-chave: filtragem colaborativa, métrica de avaliação.

Introducción

Hoy se cuenta con una diversidad creciente de sitios web de comercio electrónico, donde resaltan las redes sociales y las aplicaciones de ocio/entretenimiento. Sin embargo, en los últimos años un fuerte competidor ha crecido exponencialmente: las aplicaciones móviles. Este aumento está dado ya sea por la facilidad y practicidad de uso o por el incremento del número de dispositivos móviles. La variedad y diversidad de estas aplicaciones, junto con la tendencia a interconectar los objetos que son usados de manera continua y/o cotidiana, término acuñado como internet de las cosas (IoT, por sus siglas en inglés), lleva a la importante utilización de técnicas de personalización y recomendación.

Estas técnicas, pertenecientes al campo de la recuperación de información, resultaban impensables dos décadas atrás, pero hoy se tornan necesarias e imprescindibles por el volumen de información que se maneja a cada segundo. Aunque pudieran mostrarse ejemplos de utilización en

disímiles contextos, para la mayoría de los usuarios, el comportamiento de los modelos, traducido en recomendaciones o predicciones, continúa siendo una caja negra en la que es preferible confiar antes de navegar entre las muchas opciones disponibles. En muchas ocasiones este aspecto resulta el principal motivo de abandono/pérdida de usuarios en las diferentes aplicaciones.

Diferentes sitios de comercio electrónico como Amazon, Facebook, Walmart, Ebay o Jester utilizan sistemas de recomendación (SR) para lograr incrementar las ventas y fidelizar usuarios ya que cada vez es más notable la preferencia creciente de usuarios de comprar vía digital. El empleo en estas aplicaciones de interfaces de explicación los dota de un plus para que los usuarios comprendan mejor el funcionamiento de las sugerencias/recomendaciones recibidas. Numerosos estudios realizados demuestran la utilidad de la explicación en los diferentes sistemas y la utilidad, sobre todo, para aquellos usuarios que se encuentran en el umbral de indefinición.

Los SR suelen clasificarse en basados en contenido (Pazzani y Billsus, 2007) o de filtrado colaborativo, a pesar de que pueden encontrarse otras clasificaciones como demográficas, económicas y otras muchas clasificaciones, sobre todo enfocada en la función de ganancia o restricción que se quiere aplicar. Cada una de las clasificaciones tiene un conjunto de limitantes que suelen ser atenuadas al combinarse con otras técnicas, dando paso a los sistemas de recomendación híbridos.

Todas las aplicaciones que implementan técnicas de recomendación e intentan explicar mediante diferentes recursos las sugerencias ofrecidas para el usuario se basan en la información que es utilizada por los diferentes modelos. Dentro de las interfaces más comunes en la figura 1 se muestra un histograma con los votos de los vecinos/usuarios más similares. Esta interface es muy común en los SR de filtrado colaborativo, dado que dichos modelos evalúan la similitud de todos los usuarios para encontrar el mejor vecindario capaz de hacer una sugerencia. Entonces, resulta muy útil

para los usuarios conocer el juicio de aquellos usuarios más similares, con relación al ítem que es analizado.

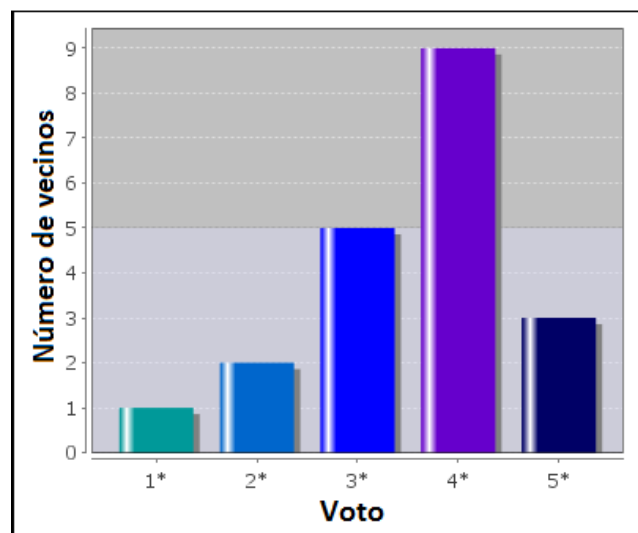


Figura 1. Interfaz de explicación. Histograma con los votos del vecindario.


Destacan otras propuestas de interfaces como la de la figura 2, muy utilizada en sistemas que emplean filtrado basado en contenido y efectúan sugerencias después de realizar el análisis del contenido de los ítems seleccionados previamente, con los que aún no han sido accedidos.

Sin lugar a dudas, la utilidad de una u otra propuesta de explicación depende del SR y del objetivo que se persiga con la interfaz diseñada. En este trabajo se muestran los resultados obtenidos de la utilización de interfaces de explicación en sistemas con usuarios reales.

En la próxima sección se muestra una breve descripción de trabajos relacionados con el campo de la explicación en los SR. Seguidamente, se analizan objetivos y medidas empleadas en la experimentación. Por último, se describen las principales conclusiones del estudio realizado, las recomendaciones para futuros trabajos en el área y el material bibliográfico utilizado.

Recommendations Based on Your New Likes Return to the Betterizer

Page 1 of 15




Twilight
 > Kristen Stewart, Robert Pattinson, ...
 ★★★★★ (1,634) \$7.99

Recommended because you said you liked:

The Twilight Saga
 > Kristen Stewart, Robert Pattinson, ...

The Twilight Saga
 > Kristen Stewart, Robert Pattinson




Harry Potter and the Goblet of Fire
 > Daniel Radcliffe, Emma Watson, ...
 ★★★★★ (782) \$6.99

Recommended because you said you liked:

Harry Potter and the Order of the Pho...
 > Daniel Radcliffe, Emma Watson, ...

Harry Potter and the Half-Blood Prince
 > Daniel Radcliffe, Rupert Grint, Emma...




Twilight
 > Robert Pattinson, Kristen Stewart, ...
 ★★★★★ (16) \$10.99

Recommended because you said you liked:

The Twilight Saga
 > Kristen Stewart, Robert Pattinson, ...

Twilight in Forks
 > Various, Jason Brown



Twilight
 > Kristen Stewart, Robert Pattinson
 ★★★★★ (53)

Recommended because you said you liked:

The Twilight Saga
 > Kristen Stewart, Robert Pattinson, ...

The Twilight Saga
 > Kristen Stewart, Robert Pattinson

[> See all of your recommendations](#)

Figura 2. Interfaz de explicación.

Fuente: Amazon.

Contenido

De forma general, la explicación encuentra su origen en los sistemas expertos (Lacave y Diez, 2004) y siempre parte de la idea de que estos sean transparentes para los usuarios. Estos sistemas deben mostrar las razones de las recomendaciones centradas en la explicación de las conclusiones arribadas por los modelos empleados. Al dar una explicación se intenta ser transparente, exponiendo las razones y los datos considerados en la recomendación. Dados los modelos empleados en muchos sistemas, resulta difícil lograr una adecuada transparencia mediante algún mecanismo de explicación. Evidentemente, pudieran encontrarse campos de aplicación donde resulte crítico o complejo este aspecto, por ejemplo, cuando se considera el entretenimiento, la aceptación de una recomendación o no, e incluso su posterior evaluación negativa, solo involucra generalmente gasto de tiempo como recurso. En otros campos donde el usuario deba comprar un ítem (por ejemplo, un destino turístico, un boleto aéreo o un equipo electrónico) sí se muestran razones adecuadas, pudiendo incluso persuadirse al usuario llevándolo a adquirir el ítem, además de confiar y volver a emplear el sistema.

Las interfaces de explicación en los SR cuentan con múltiples objetivos que van desde lograr inspirar confianza al usuario hasta ayudarlo a tomar una adecuada decisión, todo depende de lo que se persiga con el sistema. En Tintarev y Masthoff (2007 y 2011) se mencionan y describen siete de los dichos objetivos:

- Transparencia: explicar cómo funciona el SR.
- Escrutabilidad: permitir al usuario decir si el SR funciona bien o mal.
- Confianza: incrementar la confianza de los usuarios en el SR.
- Eficacia: ayudar a los usuarios a tomar buenas decisiones.
- Persuasión: convencer a los usuarios para que prueben o compren los ítems.

- Eficiencia: ayudar a los usuarios a tomar decisiones rápidas.
- Satisfacción: incrementar la facilidad de uso o disfrute del SR.

De forma general, las propuestas de explicación persiguen mostrar cuán relevante es el ítem recomendado. En Tintarev y Masthoff (2007 y 2011) destacan la presentación y justificación de: el mejor ítem, los mejores n ítems en un listado y los ítems más similares. Estudios como los de Zanker y Martin (2014) evalúan el efecto persuasivo de varios estilos de explicación enfocados en el uso de palabras independientes u oraciones que argumentan el ítem sugerido por el sistema. Los resultados evidencian mejores resultados con el último de los estilos mencionados. Con un enfoque similar, Ben-Elazar y Koenigstein (2014) proponen un *framework* de explicación híbrida con alta aceptación.

Hetlocker, Konstan y Riedl (2000) desarrollaron un estudio con MovieLens donde describen cómo la explicación de las recomendaciones puede persuadir a los usuarios a tener mayor confianza en el sistema. Los autores revelan cómo pueden diseñarse explicaciones para cada posible etapa del proceso de recomendación. Respecto al criterio de los usuarios, el estudio desarrollado reveló que de las 21 interfaces utilizadas, donde se muestra un histograma con los votos agrupados de los vecinos resultó la más efectiva y aceptada por los usuarios.

Otros estudios como los de Chen y Pu (2005) y Pu y Chen, (2006 y 2007) abordan la construcción de confianza a través del correcto diseño de interfaces para proveer explicaciones. Los autores se enfocan en que dichas interfaces se autoexpliquen. Por su parte, Bridge y Dunleavy (2014) proponen un algoritmo para obtener reglas que son utilizadas en una interface de explicación, obteniendo positivos resultados.

El empleo de algoritmos para obtener reglas explicables (Bridge y Dunleavy, 2014) y la utilización de nube de etiquetas (Zhang, 2014; Zhang *et al.*, 2014), son otros ejemplos de mecanismos

diversos de explicación de gran utilidad. De forma general, en todos los ejemplos anteriores se intenta evaluar uno o un conjunto de objetivos para determinar la efectividad de las interfaces.

Con los nuevos avances tecnológicos, elementos como la geo localización, redes sociales, información en la nube, entre otros, se convierten en variables y objetivos evaluables en la utilidad de los diferentes SR e interfaces de explicación para dichos sistemas.

Luego de realizado un breve comentario sobre las tendencias de las interfaces de explicación se pasa a detallar para cada uno de los objetivos algunas medidas y métricas de utilidad consideradas en este estudio.

Medidas para evaluar objetivos de las interfaces de explicación

Cada uno de los objetivos marcados en el momento de la evaluación es muy subjetivo y depende del grado de satisfacción de cada usuario. Pudiera preguntarse directamente a cada uno de los usuarios y luego considerar métricas globales o ciertas tendencias.

Dado que la *transparencia* intenta explicar al usuario cómo funciona el SR, sin lugar a dudas la opinión de este es muy importante. Aunque bien pudieran obtenerse algunas medidas para evaluar el uso del usuario y con ello asociarlo a la utilidad de la transparencia de la interfaz de explicación, por ejemplo:

- Tiempo de interacción con la aplicación después de observar la explicación.
- Argumentos empleados luego de evaluar la explicación.
- Aceptación de la recomendación actual y futuras.
- Variación en el valor de la recomendación.
- Fidelidad del usuario y utilización de la aplicación.
- Estudio de intención de los usuarios y estudio de usuarios (monitorear).

En el caso de la variación de la recomendación se pueden emplear las métricas clásicas Error absoluto medio (MAE), Raíz del error medio cuadrático (RMSE) y las versiones normalizadas de ambas, precisión, exhaustividad, y F-Measure (Avazpour, Pitakrat, Grunske y Grundy, 2014). Para los restantes casos sería necesario análisis de texto y/o medidas de registro en función de tiempo o en función de otros criterios.

Para evaluar la *escrutabilidad*, y considerando que esta permite al usuario expresar el grado de conformidad con el funcionamiento del sistema (funciona bien o mal), se pueden considerar los siguientes elementos:

- Valoración o puntuación de los usuarios.
- Variación del valor de la recomendación.
- Aceptación de la recomendación actual y futuras.
- Argumentos empleados luego de evaluar la explicación.
- Tiempo de interacción con la aplicación después de observar la explicación.
- Corrección del voto de recomendación.

En los casos en los que interviene el valor de la recomendación pueden considerarse las métricas clásicas señaladas en el caso de la transparencia.

Mientras, en el caso de la *confianza*, como opción para incrementar la confianza de los usuarios en el SR pueden emplearse un conjunto de discreciones para evaluar el criterio de los usuarios. Las interfaces pueden dar información para que el usuario certifique el uso o no del sistema; por tanto, sería importante considerar los siguientes criterios:

- Valoración o puntuación de los usuarios.
- Variación del valor de la recomendación.
- Aceptación de la recomendación actual y recomendaciones futuras.
- Gastos de recursos (tiempo, dinero, etc.) luego de la explicación.

- Selección de ítems dentro de un listado de recomendados.
- Corrección del voto de recomendación.

Uno de los elementos más interesantes radica en evaluar el gasto de los usuarios luego de valorar la explicación. Es importante evaluar el perfil del usuario y las tendencias de forma general. En este objetivo es importante considerar el empleo de métricas de evaluación de ranking, por ejemplo, desempeño basado en la distancia normalizada (NDPM), correlación de Spearman, correlación de Kendall y ganancia acumulada del descuento normalizado (NDCG), además de las métricas evaluadas en los objetivos anteriores.

Con el fin de ayudar a los usuarios a tomar correctas o buenas decisiones, se define el objetivo *eficacia* de las interfaces de explicación. Para ello pueden considerarse los siguientes elementos:

- Aceptación y utilidad posterior de la recomendación.
- Argumentos empleados luego de evaluar la explicación.
- Estudio de intención de los usuarios y estudio de usuarios (monitorear).

La mayoría de estos elementos son de tipo cualitativo y su colección requiere de la captura y monitoreo del comportamiento de los usuarios para identificar su valoración posterior.

La *persuasión* es la capacidad que se da al sistema de convencer a los usuarios para que acepten, prueben o compren los elementos que están siendo recomendados o sugeridos. En este caso las interfaces de explicación juegan un papel importante, pues muchos usuarios se encuentran con dudas de aceptar o no el ítem recomendado y, por tanto, una adecuada interfaz puede ayudarlo y hasta convencerlo. Así, se consideran importantes los aspectos que a continuación se relacionan:

- Aceptación de la recomendación actual y recomendaciones futuras.

- Gastos de recursos (tiempo, dinero, etc.) luego de la explicación.
- Selección de ítems dentro de un listado de recomendados.
- Corrección del voto de recomendación.
- Fidelidad del usuario y utilización de la aplicación.
- Estudio de intención de los usuarios y estudio de usuarios (monitorear).

Al igual que los primeros objetivos, se pueden utilizar las métricas MAE, RMSE y las versiones normalizadas de ambas, además de la precisión, exhaustividad y F-Measure. Para los restantes casos sería necesario analizar y monitorear a los usuarios para evaluar su gasto y comportamiento futuro.

Se considera el objetivo de *eficiencia* de la interfaz de explicación como la característica que el sistema tiene de ayudar a los usuarios a tomar decisiones rápidas. Dentro de los elementos más importantes a considerar se encuentran:

- Selección de ítems dentro de un listado de recomendados.
- Aceptación rápida de la recomendación actual y recomendaciones futuras.
- Análisis de comportamientos actuales con respecto a comportamientos anteriores.
- Gasto efectivo de recursos costosos (dinero, bonos, etc.) inmediatamente posterior de valorar la explicación.

En este objetivo se puede evaluar la NDCG relacionada con el posicionamiento de los ítems seleccionados en el ranking mostrado, además de otras métricas relacionadas con la importancia de los elementos del ranking y otras que monitorean el comportamiento del usuario.

Como último objetivo se tiene la *satisfacción*, con la idea de incrementar la facilidad de uso o disfrute del SR. Una vez los usuarios comprenden el funcionamiento del sistema logran interactuar mucho más y confiar en el sistema, pues les resulta más fácil el uso y comprensión de las recomendaciones

y pueden disfrutar del material recibido. Por tanto, se considera importante valorar:

- Gastos de recursos (tiempo, dinero, etc.) luego de la explicación.
- Número de clic para acceder a ítems recomendados.
- Aceptación de la recomendación actual y recomendaciones futuras.
- Fidelidad del usuario y utilización de la aplicación.
- Valoración o puntuación de los usuarios.
- Variación del valor de la recomendación.
- Estudio de intención de los usuarios y estudio de usuarios (monitorear).

Como en mucho de los objetivos anteriores donde se evalúan criterios cuantitativos, se pueden utilizar las métricas MAE, RMSE y las versiones normalizadas de ambas, además de la precisión, exhaustividad y F-Measure. Para los restantes casos sería necesario analizar y monitorear a los usuarios para evaluar su gasto y comportamiento futuro. Otras métricas de registro pueden ser empleadas también de forma general.

Aplicación de interfaces de explicación en sistemas de recomendación

En la experimentación fueron aplicadas las interfaces de explicación en tres aplicativos móviles que realizan recomendaciones con diferentes enfoques. Un estudio inicial mostró declaraciones de usuarios que manifestaron no entender las recomendaciones recibidas en dichos sistemas.

- Aplicativo de recomendación de películas.
- Aplicativo de recomendación de restaurantes.
- Aplicativo de recomendación de productos.

Las aplicaciones seleccionadas no disponían de ningún mecanismo de argumentación o explicación. Fueron utilizadas las interfaces de la figura 3 y figura 4.

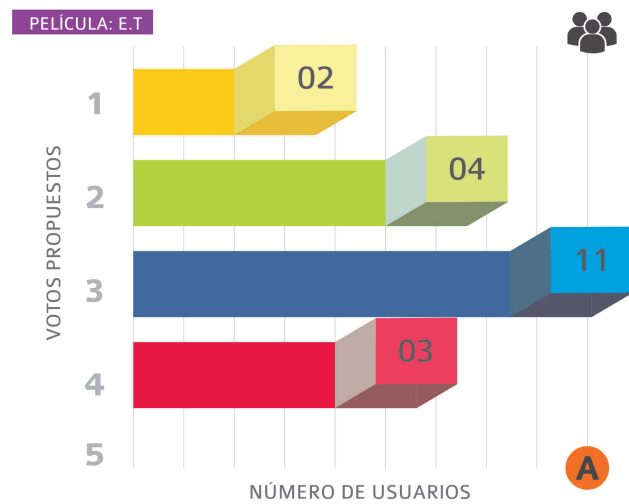


Figura 3. Propuesta de interfaz de explicación. Histograma con los votos de usuarios más similares.

Fuente: los autores.

La primera (figura 3) muestra la clásica interfaz del comportamiento de los usuarios más similares en relación con el ítem que se está recomendando. La segunda interfaz refleja en un mapa bidimensional el comportamiento de los usuarios en función de los ítems valorados en el pasado. Esta última interfaz refleja cómo fue la predicción de la vecindad en relación a la del usuario activo, tanto a nivel colectivo como individual. En el lado izquierdo se refleja el grado de similitud de esos usuarios que hubieran recomendado ese ítem en relación a la valoración real efectuada por el usuario. Esta valoración es realizada tanto a nivel colectivo como individual. Mientras, por el lado derecho se muestra un resumen tabular de acierto, sobre y subestimación realizada por la vecindad.

Para todos los ítems valorados por el usuario se muestra la calidad de predicción, empleando cuadrados o diamantes para representar el acierto o falla del vecindario, respectivamente. Las figuras tienen un color el cual representa el criterio dado por el usuario. Dicho color está en correspondencia con otros utilizados en los histogramas que complementan la propuesta. La propuesta permite capturar como:

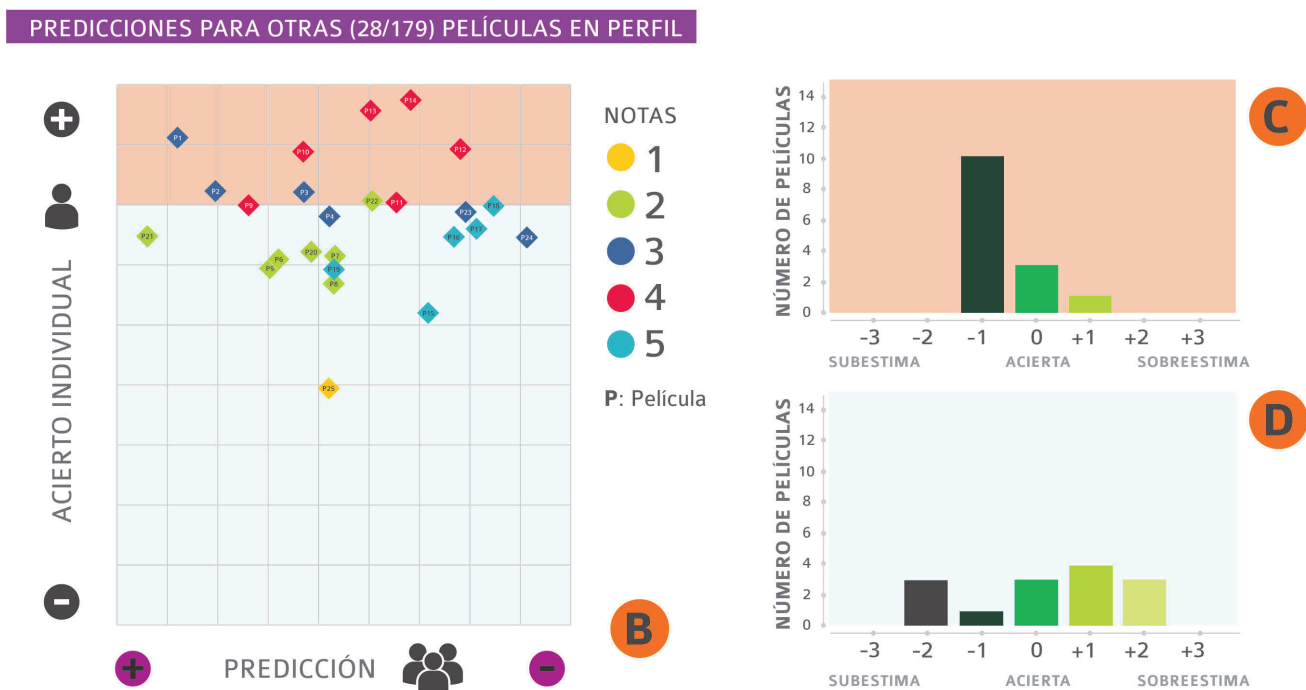


Figura 4. Propuesta de Interfaz de explicación. Mapa de acierto de usuarios más similares.

Fuente: los autores.

Tabla 1. Resultados de la experimentación con usuarios.

Criterios	Sistemas de recomendación			
	Películas	Restaurantes	Productos	
Usuarios monitoreados	Hombres	20	18	5
	Mujeres	18	8	12
	Total	38	26	17
Actividad de uso inicial	Baja	8	8	6
	Media	12	12	6
	Alta	18	6	5
Actividad de uso 25 días	Baja	5	8	4
	Media	12	8	6
	Alta	21	10	7
Actividad de uso 5 días finales	Baja	3	6	4
	Media	11	6	3
	Alta	24	14	10
Aceptación inicial de las recomendaciones	20	10	7	
Aceptación final de las recomendaciones	29	16	11	
Gasto medio de recursos (segundos)	Inicial	46	30	112
	5 días finales	34	57	173

Fuente: los autores.

- Si una película X le gusta a los usuarios similares (vecindario), seguro que la que me recomiendan me gustará.
- ¿Qué opinión tiene este vecindario sobre la película Y? Esa película fue de mis preferidas.
- ¿Cuáles de las películas que he visto son valoradas en conjunto por el vecindario igual que esta que me recomiendan?

En la experimentación fueron considerados los elementos mostrados en la tabla 1, que de forma general registran la continuidad de uso de los usuarios con los diferentes sistemas monitoreados. El estudio fue realizado en un periodo de tiempo de 48 días.

Los resultados muestran la incidencia en todos los sistemas analizados de las interfaces de explicación. Aunque la evaluación del tiempo pudiera resultar una medida subjetiva, fue considerado porque los usuarios eran conscientes del estudio. Resaltar que el tiempo de visualización y análisis de la interfaz de explicación no fue considerado.

En relación al uso de las aplicaciones se puede observar que aumentó el número de usuarios con actividad alta en los tres sistemas (de un 36 % para un 59 %). Reduciendo también el número de usuarios con actividad baja (de un 27 % para un 16 %). Aquí pudiera considerarse que tal aumento de actividad está provocado porque las recomendaciones están facilitando al usuario:

- Una mayor transparencia en relación al conocimiento de funcionamiento del sistema respecto a las recomendaciones recibidas.
- Tener mayor confianza en las recomendaciones recibidas y usar continuamente el sistema.
- Comprar o adquirir los ítems recomendados y usar nuevamente los sistemas.

A su vez, la aceptación de las recomendaciones por parte de los usuarios también aumentó (de un 46 % para un 69 %). Evidenciando un incremento de la confianza por parte del usuario y el

considerar buenas las recomendaciones recibidas. Es importante que el usuario sienta que es reconocido por el sistema y que este lo conoce bien.

En la última de las variables se nota que, en el sistema de recomendación de películas, los usuarios terminaron gastando menos tiempo en la aplicación, lo que pudiera considerarse que invierten menos tiempo en buscar aquella película que gustarían ver. Mientras que, en los otros dos sistemas, los usuarios aumentaron el tiempo medio de uso, lo que representa que, aunque no compraran, los usuarios se convirtieron en usuarios más frecuentes de las aplicaciones.

Conclusiones y trabajos futuros

Sin lugar a dudas, diseñar interfaces de explicación siempre está condicionado a los objetivos que se persiguen en un SR. Un sinnúmero de métricas suele utilizarse para evaluar los criterios que resultan más importantes dentro del sistema, siendo en muchos casos medidas diferentes en una u otra investigación, aunque sean similares. En el trabajo se abordó un conjunto de elementos que permiten partir de un punto inicial en cuanto a la evaluación de los diferentes objetivos de las interfaces de explicación.

Para los diferentes objetivos se comentan los principales elementos que pueden considerarse y se sugieren métricas para que cada uno de ellos sea evaluado. Al final, y con el auxilio de dos interfaces de explicación, se obtiene una evaluación de usuarios reales en cuanto a la utilización de interfaces de explicación en sistemas de amplio uso en la actualidad.

El aumento de la actividad de los usuarios en el sistema, la aceptación de las recomendaciones recibidas y la variación (reducción/aumento) del gasto medio de recursos reflejan la importancia de la utilización de interfaces de recomendación en sistemas de recomendación. Las mismas permiten personalizar el contenido a los usuarios y que cada uno se sienta, en un medio virtual, un usuario único y exclusivo.

La continuidad de las experimentaciones permitirá evaluar otros criterios y conocer el comportamiento de los usuarios en relación a otros objetivos. Ampliar el número de interfaces de explicación, así como experimentar en otros sectores de aplicativos de comercio electrónico, se convierte en los próximos desafíos de los investigadores.

Referencias

- Avazpour, I., Pitakrat, T., Grunske, L. y Grundy, J. (2014). Dimensions and Metrics for Evaluating Recommendation Systems. En M. P. Robillard et al. (eds.), *Recommendation Systems in Software Engineering* (pp. 245-273). Berlin: Springer Berlin Heidelberg. Doi: 10.1007/978-3-642-45135-5.
- Ben-Elazar S. y Koenigstein, N. (2014, octubre). *A Hybrid Explanations Framework for Collaborative Filtering Recommender Systems*. Poster presentado en el 8th ACM Conference on Recommender Systems (RecSys), Foster City, Silicon Valley, EE.UU.
- Bridge, D. y Dunleavy, K. (2014). *If you liked Herlocker et al.'s explanations paper, then you might like this paper too*. En Tintarev et al. (eds.), *Procs. of the Workshop on Interfaces and Human Decision Making for Recommender Systems: 8th ACM Conference on Recommender Systems* (pp. 22-27), Foster City, Silicon Valley, EE.UU.
- Chen L. y Pu, P. (2005, octubre). *Trust building in recommender agents*. En 1st International Workshop on Web Personalization, Recommender Systems and Intelligent User Interfaces (WPRSIUI05) (pp. 135-145), Reading, UK.
- Herlocker, J. L., Konstan, J. A. y Riedl, J. (2000). Explaining collaborative filtering recommendations. En *CSCW '00: Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work* (pp 241-250). Nueva York: Association for Computing Machinery.
- Lacave, C. y Diez, F. J. (2004). A review of explanation methods for heuristic expert systems. *The Knowledge Engineering Review*, 19(02), 133-146.
- Pazzani, M.J. y Billsus, D. (2007). Content-based recommendation systems. En Brusilovski et al. (eds.), *The Adaptive Web: Methods and Strategies of Web Personalization* (pp. 325-341). Berlin: Springer-Verlag Berlin Heidelberg.
- Pu P. y Chen, L. (2006, enero). Trust building with explanation interfaces. En *Proceeding of the 11th international conference on Intelligent User Interfaces* (pp 93-100). Sydney, Australia: Association for Computing Machinery Press.
- Pu P. y Chen, L. (2007). Trust-inspiring explanation interfaces for recommender systems. *Knowledge-Based Systems*, 20(6), 542-556.
- Tintarev, N. y Masthoff, J. (2007, abril). A survey of explanations in recommender systems. En *Proceedings of the 2007 IEEE 23rd International Conference on Data Engineering Workshop* (pp. 801-810). Washington D.C.: IEEE Computer Society.
- Tintarev, N. y Masthoff, J. (2011). Designing and evaluating explanations for recommender systems. En Ricci et al. (eds.), *Recommender Systems Handbook* (pp. 479-510). Nueva York: Springer Science+Business Media.
- Zanker, M. y Schoberegger, M. (2014, octubre). *An Empirical Study on the Persuasiveness of Fact-based Explanations for Recommender Systems*. En Tintarev et al. (eds.), *Procs. of the Workshop on Interfaces and Human Decision Making for Recommender Systems: 8th ACM Conference on Recommender Systems*, (pp. 33-36), Foster City, Silicon Valley, EE.UU.
- Zhang, Y. (2014, octubre). *Browser-Oriented Universal Cross-Site Recommendation and Explanation based on User Browsing Logs*. En *Proceedings of the 8th ACM Conference Series on Recommender Systems (RecSys)*, Foster City, Silicon Valley, EE.UU.
- Zhang, Y., Lai, G., Zhang, M., Zhang, Y., Liu, Y. y Ma, S. (2014, julio). *Explicit factor models for*

explainable recommendation based on phrase-level sentiment analysis. En Proceedings of the 37th international ACM SIGIR conference

on Research & development in information retrieval (pp 83-92), Gold Coast, Queensland, Australia.

