

# Contextualizando el Conocimiento de Usuarios desde sus Comentarios en Foros

Silvana Aciar y Gabriela Aciar<sup>1</sup>

<sup>1</sup>Universidad Nacional de San Juan, Argentina.  
saciar@iinfo.unsj.edu.ar - gaby\_aciar@yahoo.com.ar

**Abstract.** Este artículo propone una solución al problema de buscar usuarios que respondan preguntas sobre un tema. La solución propuesta es un sistema recomendador que sugiere usuarios en base al conocimiento sobre el tema. Este conocimiento es obtenido desde las publicaciones realizadas en los foros. El valor del conocimiento es diferente dependiendo del contexto en el que se utilice. En este trabajo se presenta una forma de contextualizar el conocimiento de una persona para luego utilizarlo en un sistema recomendador. Un caso de estudio en el dominio de educación, demuestra que se obtienen recomendaciones más precisas si tiene en cuenta información del contexto.

**Keywords:** Text Mining, Sistemas recomendadores, Foros, Interacciones de usuarios.

## 1. Introducción

Las tecnologías de la información y comunicación (TICs) han cambiado el modo de interacción entre las personas. La penetración masiva de Internet en los hogares hace posible que los usuarios tengan acceso a la gran cantidad de información disponible en la web. Esto provoca que en muchas ocasiones se produzca la paradoja de no saber qué escoger a pesar de que a nuestro alrededor se acumula información que podría ser de nuestro interés. Esta situación también se observa en el ámbito educativo. Hoy en día, la fuente de información más consultada es Internet. Actualmente las plataformas virtuales de aprendizaje presentan la misma información a los usuarios que entran en ella, información estándar con links a otras páginas [1]. Un estudiante o docente puede estar horas mirando las diferentes alternativas y al final desistir en su búsqueda por la cantidad de tiempo que le lleva visitar cada página, o quedarse con una opción que quizás no sea la mejor para su perfil. Es por eso que el problema ya no es de no tener información, si no, qué escoger entre un número inmanejable de alternativas disponibles.

Herramientas, con posibilidades sincrónicas y asincrónicas incorporadas a la formación y la educación, propician nuevas opciones de interacción y retroalimentación, para ayudar a los pedagogos y estudiantes a adquirir/enseñar conocimiento a más sectores de la sociedad [2] [3][4].

Una forma de facilitar la interacción de docentes/estudiantes con las plataformas virtuales de aprendizajes es mediante el uso de los sistemas recomendadores [5][6][7]. El propósito de estos sistemas es el de simplificar el proceso de búsqueda de una persona solucionando el problema de sobrecarga de información[8].

Los sistemas recomendadores en e-learning brindan sugerencias a los profesores/estudiantes durante su aprendizaje para maximizar su desempeño, proveyendo la recuperación y presentación personalizada de los recursos de aprendizaje en base a las necesidades, intereses, preferencias y gustos.

Este artículo se enfoca en el problema de recomendaciones de personas. Por ejemplo un estudiante que tiene una inquietud sobre un tema en particular y necesita que otro usuario que conoce del tema lo ayude. Situación que es frecuente en un ambiente no virtual. En una plataforma virtual saber quien conoce del tema es un poco más difícil debido a que puede haber muchos estudiantes o docentes registrados y no se cuenta con el contacto presencial que favorece tener conocimiento acerca de la experiencia y conocimiento que poseen las personas. En [9] se presentó la forma de obtener ese conocimiento que poseen los usuarios de las publicaciones en foros en una plataforma virtual de aprendizaje. En base a ese conocimiento son recomendadas a otras personas para responder sus inquietudes sobre un tema particular.

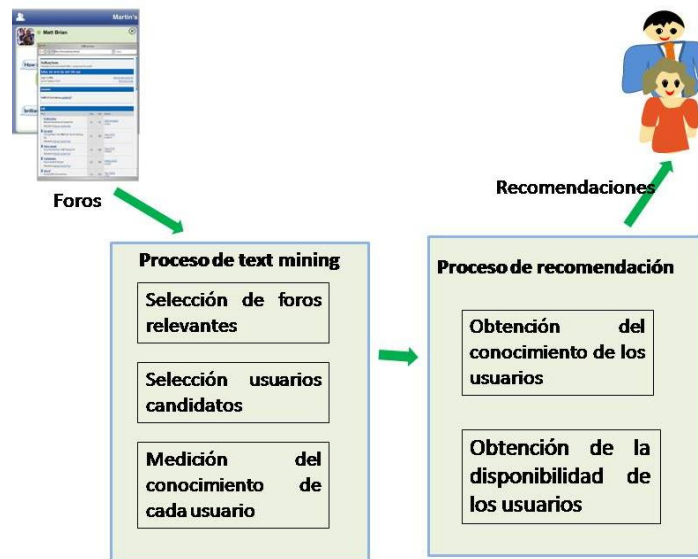
El presente trabajo es una extensión de [9], se integra en las recomendaciones información del contexto del conocimiento de la persona recomendada. Información del contexto es obtenida del análisis de las publicaciones realizadas en los foros. El conocimiento es sensible al contexto, la información que posee una persona tiene valores diferentes dependiendo del contexto en la cual se vaya a utilizar.

La estructura de este artículo es la siguiente: En la Sección 2 se presenta el trabajo previo realizado para obtener el grado de conocimiento de los usuarios desde los textos escritos en un foro. El proceso realizado para la obtención del contexto desde publicaciones en los foros es descrito en la Sección 3. Sección 4 presenta el proceso de realizar las recomendaciones de usuarios integrando información del contexto. En la sección 5 se presenta un caso de estudio que muestra la efectividad de la propuesta. Finalmente las conclusiones y trabajo futuro son presentados en la Sección 6.

## 2. Trabajo previo

En [9] se propone el uso de sistemas de recomendación para sugerir usuarios adecuados para responder preguntas de otras personas. En la literatura hay varios estudios que recomiendan a los usuarios interactuar en una red social [5] [6] [7] muy pocos analizan la información escrita y publicada por los usuarios en las redes sociales. Parte del problema radica en la complejidad de la extracción de información de texto y convertir esa información en recomendaciones. En [9] se analiza el contenido del comentario del usuario para obtener su conocimiento sobre un tema en particular, y luego se recomienda los usuarios en base a ese valor.

La figura 1 muestra el proceso utilizado para recomendar usuarios en base al conocimiento y disponibilidad.



**Fig. 1.** Proceso para recomendar usuarios en base a las interacciones que realizan por medio de Foros

Se ha desarrollado un método de selección y recuperación de información desde los comentarios de los usuarios, métricas para obtener un valor que representa el conocimiento de los usuarios sobre un tema y se realizó la recomendación de un usuario en base a ese valor.

## 2.1 Selección de comentarios e identificación de los usuarios candidatos

Dada una petición de usuario  $Q$  donde  $Q$  está compuesta de palabras clave de modo que  $Q = (q_1, q_2, \dots, q_n)$ , el primer paso es encontrar todas las publicaciones que contienen  $Q$ . La búsqueda de publicaciones relevante una búsqueda simple, se seleccionan todas las publicaciones que contienen las palabras de  $Q$  ( $q_i$ ). El comentario seleccionado constituye una colección de publicaciones relevantes  $Pr = \{p_1, p_2, p_3, \dots, p_n\}$  donde cada publicación puede tener asociado un conjunto de comentarios y usuarios que las escribieron.

La Figura 2 muestra la estructura de los datos obtenidos a partir de las publicaciones pertinentes.

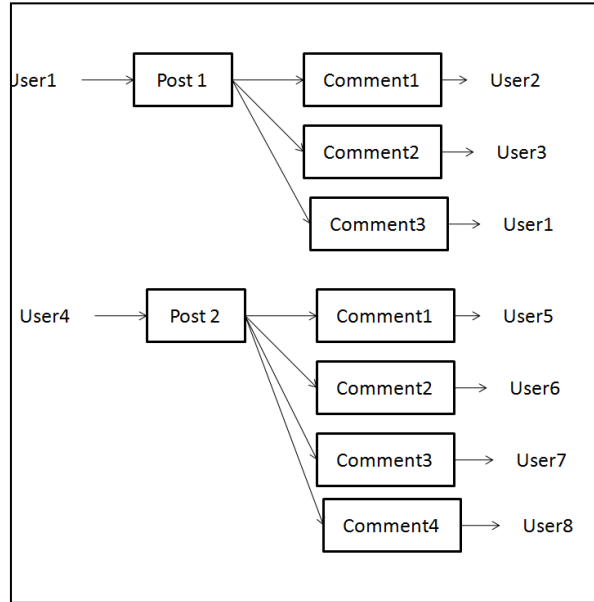


Fig. 2. Información obtenida de comentarios relevantes

Los usuarios que hicieron comentarios y los que hicieron la publicación constituyen un conjunto de usuarios candidatos  $UC = \{u_1, u_2, u_3, \dots, u_n\}$ . Para cada usuario candidato, se obtienen los comentarios realizados y las veces que interactuaron en la comunidad virtual.

## 2.2 Análisis de los comentarios del usuario

Un usuario puede responder a una pregunta  $Q$  si tiene conocimiento sobre  $Q$  y está disponible para responder preguntas. Se han definido un conjunto de métricas para obtener dicha información. La medida Información del Usuario sobre un Tema (**UIT**) se calcula con la información obtenida desde lo que han escrito en los comentarios, mientras que las medidas Interacción de Usuario (**UI**) y la Voluntad de Responder (**WR**) se calculan utilizando la información de las interacciones en la comunidad virtual. Más detalle del cálculo de estas medidas se encuentran en [9]

El Conocimiento del Usuario sobre el Tema (**UKT**) se obtiene en base a la información que tiene el usuario sobre el tema (**UIT**) multiplicado por el número de interacciones del usuario en la red virtual (**UI**). Ecuación 1 se utiliza para obtener el conocimiento de cada usuario candidato.

$$UKT(u) = UIT * IU \quad (1)$$

El usuario más adecuado para recomendar es seleccionado en base al grado de conocimiento (**UKT**) y la voluntad que tiene para responder una pregunta (**WR**). **WR** es definida como la probabilidad que tiene un

usuario para responder una pregunta en base a su comportamiento en el pasado. Una persona que ha respondido frecuentemente preguntas de otros usuarios es más probable que esa persona responda otra pregunta respecto a otro usuario que no ha respondido preguntas en el pasado aunque se lo hayan preguntado. Más detalles del proceso obtención del conocimiento y recomendación de usuarios relevantes se encuentra en [9].

### 3. Obteniendo el contexto de los comentarios

El conocimiento que posee un usuario es sensible al contexto, por ejemplo: en el caso que un estudiante tenga dudas de un tema pide ayuda a un colega suyo que le pueda ayudar a resolver la duda, si la duda es respecto a un tema de ingeniería le pedirá ayuda a una persona que trabaja o se desempeña en esa área. En la solución planteada en la sección anterior no se tiene en cuenta el contexto en el cual un usuario tiene conocimiento sobre un tema.

En este trabajo el contexto se define como cualquier condición que permita identificar el área, la disciplina, las palabras claves relacionadas con el conocimiento que posee un usuario.

El contexto condicionará la aceptación o no de las recomendaciones realizadas a los usuarios. Figura 3 muestra el proceso de recomendación donde se añade información del contexto.

El contexto se modelará como un conjunto de conceptos involucrados en las publicaciones de cada usuario candidato. Estos conceptos serán identificados como aquellas palabras que fueron más mencionadas en las publicaciones. Se espera que el conjunto de palabras más mencionadas, presentadas a un usuario le permita seleccionar a la persona recomendada más adecuada para su contexto. En la Figura 4 se muestran las recomendaciones que se presentan al usuario con información del contexto en forma de palabras relacionadas.

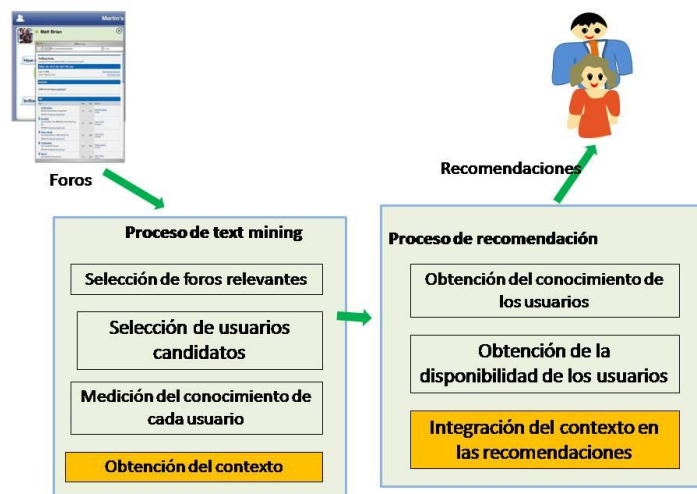


Fig. 3. Proceso para recomendar usuarios integrando información del contexto



Fig. 4. Recomendaciones resultantes con información del contexto en forma de palabras relacionadas

A continuación se definen los pasos a seguir en el proceso de obtención de información del contexto desde los comentarios escritos por los usuarios en un foro.

### 3.1 Preprocesamiento

Para cada usuario candidato se obtienen todas las publicaciones realizadas. Se eliminan todos aquellos términos que son irrelevantes, como lo son los artículos, las preposiciones, los adjetivos, etc. Para ello se utiliza una lista de palabras “stop words” elaborada para tal fin.

### 3.2 Selección de los términos relevantes

Las palabras claves que permitirán conocer el contexto en el cual un usuario tiene conocimiento sobre un tema se obtiene por medio del cálculo de las frecuencias de las palabras, las palabras más frecuentes son las que se presentarán junto con el nombre y contacto del usuario recomendado.

El contexto  $CX$  del conocimiento de un usuario  $i$  es un conjunto de duplas compuesta por las palabras más frecuentes en los textos escritos por  $i$   $CX_i = ((p1, w1), (p2, w2), \dots (pn, wn))$ . Donde  $w_j$  es el peso que tiene la palabra  $j$  en todas las publicaciones del usuario  $i$ .

$w_j$  se obtiene por medio de la medida  $tf-idf$  de la Ecuación 2.  $F_{jk}$  es la frecuencia normalizada del término  $j$  en los comentarios realizados por el usuario  $i$ , los cuales los denominamos  $k$ .  $idf_j$  representa la frecuencia invertida de la palabra  $j$ . Estas medidas se obtienen mediante las Ecuaciones 3 y 4.

$$w_j = f_{jk} * idf_j \quad (2)$$

$$f_{jk} = \frac{frec_{jk}}{MaxFrec_k} \quad (3)$$

$$idf_j = \log \frac{N}{n_j} \quad (4)$$

Donde  $frec_{jk}$  es la frecuencia de la palabra  $j$  en los comentarios  $k$ .  $MaxFrec_k$  representa la máxima frecuencia sobre todas las palabras de los comentarios  $k$ .  $N$  es el número de fóruns que existen y  $n_j$  es la cantidad de fóruns que contienen el término  $j$ .

#### 4. Proceso de recomendación

Los usuarios recomendados son aquellos que tienen el valor de conocimiento más alto y que de otro respecto al contexto. Para saber la diferencia entre los usuarios según el contexto, se modela el contexto del usuario como un vector. El vector está formado por las 5 palabras más frecuentes. Luego se realiza una comparación entre vectores formados por cadenas. Los tres usuarios menos similares respecto al contexto son recomendados, ordenados por grado de conocimiento.

#### 5. Caso de estudio

Un caso de estudio en el dominio de e-learning fue realizado. Se crearon diferentes foros en la plataforma Moodle. En ellos participaron 12 estudiantes y 4 profesores. Se motivó a que participaran discutiendo varios temas en las áreas de Lenguaje, Matemática, Computación y Ciencias Sociales. Se obtuvieron 256 publicaciones de los usuarios.

##### 5.1 Experimentos

La consigna del experimento fue que durante 15 días los estudiantes utilizaran el recomendador de usuarios aplicado en la plataforma Moodle para ingresar preguntas al recomendador y el mismo le sugerirá personas a las cuales contactar.

Cada vez que el usuario recibía una recomendación, luego de contactar a la persona recomendada se le pedía que evalué la recomendación. Recibía un email con el texto “La recomendación recibida ha satisfecho sus requerimientos?”. Esta información se utilizó para evaluar las recomendaciones.

Se dividió a los estudiantes en dos grupos y se realizaron dos experimentos:

**Experimento 1:** 6 estudiantes utilizó el recomendador solamente teniendo en cuenta el conocimiento del usuario [9].

**Experimento 2:** Los restantes 6 estudiantes utilizaron el recomendador que realizó las sugerencias en base al conocimiento e incorporando información contextual de la persona recomendada.

##### 5.2 Evaluación

La evaluación se realizó teniendo en cuenta las recomendaciones exitosas sobre todas las recomendaciones realizadas. Recomendaciones exitosas es definida como: dado un requerimiento del usuario, la recomendación que satisfaga sus expectativas es considerada exitosa, para ello se utiliza la información obtenida del email enviado para evaluar las recomendaciones. Si el usuario ha respondido afirmativamente que se ha satisfecho sus expectativas, la recomendación es considerada exitosa. Esta información es utilizada en la ecuación 5 para obtener un valor de precisión.

$$Precisión = \frac{NR}{N} \quad (5)$$

Donde NR es el número de recomendaciones exitosas y N el número total de recomendaciones realizadas. Esta medida se ha obtenido por cada día que se realizó el experimento (15 días) y para ambos experimentos. Los resultados se pueden observar en la figura 5. Como se puede observar las recomendaciones realizadas agregando información del contexto tienen mayor precisión. Los usuarios se encuentran más satisfechos si el recomendador les brinda más información referente a la persona a quien contactar.

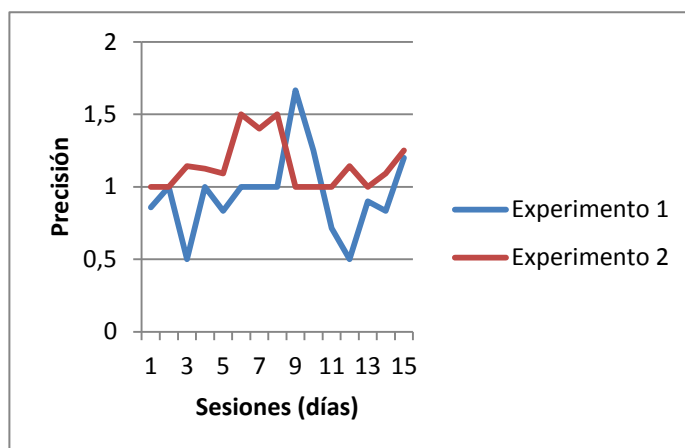


Fig. 5. Precisión de las recomendaciones obtenida en la ejecución de los dos experimentos

## 6. Conclusiones

En este trabajo se presentó una solución al problema de cómo obtener el contexto en el cual una persona posee conocimiento sobre un tema. El valor del conocimiento es diferente en diferentes contextos. Si la finalidad de un sistema recomendador es sugerir usuarios en base al conocimiento que tengan de un tema, es muy relevante tener en cuenta el contexto donde se generó o aplicó el conocimiento que posee la persona a recomendar. Una de las causas de la baja precisión de las recomendaciones puede ser debido a que no se tiene en cuenta el contexto en el cual se va a realizar la recomendación. La solución propuesta en este trabajo fue brindar información del contexto en las diferentes alternativas recomendadas, para ayudar al usuario a seleccionar una. El caso de estudio demuestra que al integrar dicha información se obtienen recomendaciones más precisas. La forma de obtener información del contexto, presentada en este trabajo, fue mediante el análisis de los textos escritos en las publicaciones de los usuarios en los foros.

Como trabajo futuro se prevé realizar dos mejoras: mejorar el método de obtención del contexto mediante la elaboración de mapas conceptuales con información de los foros; integrar la información contextual en el momento del cómputo de las recomendaciones y no dejar al usuario que realice la decisión final de qué alternativa recomendada seleccionar.

## Referencias

1. Sekhavatian, M. Mahdavi (2011) APPLICATION OF RECOMMENDER SYSTEMS ON E-LEARNING ENVIRONMENTS, EDULEARN11 Proceedings, pp. 2679-2687.
2. Gaudioso, E.; Hernandez-del-Olmo, F.; Montero, M. "Enhancing E-Learning Through Teacher Support: Two Experiences", Education, IEEE Transactions on, On page(s): 109 - 115 Volume: 52, Issue: 1, Feb. 2009.
3. Abel, F.; Bittencourt, I.I.; Costa, E.; Henze, N.; Krause, D.; Vassileva, J. "Recommendations in Online Discussion Forums for E-Learning Systems", Learning Technologies, IEEE Transactions on, On page(s): 165 - 176 Volume: 3, Issue: 2, April-June 2010
4. Zhuhadar, L.; Romero, E.; Wyatt, R. "The Effectiveness of Personalization in Delivering E-learning Classes", Advances in Computer-Human Interactions, 2009. ACHI '09. Second International Conferences on, On page(s): 130 - 135
5. Salehi, M.; Kmalabadi, I.N. "Attribute-based recommender system for learning resource by learner preference tree", Computer and Knowledge Engineering (ICCKE), 2012 2nd International eConference on, On page(s): 133 - 138
6. Salehi, M.; Kmalabadi, I.N. "Attribute-based recommender system for learning resource by learner preference tree", Computer and Knowledge Engineering (ICCKE), 2012 2nd International eConference on, On page(s): 133 - 138
7. Khribi, M.K.; Jemni, M.; Nasraoui, O. "Automatic Recommendations for E-Learning Personalization Based on Web Usage Mining Techniques and Information Retrieval", Advanced Learning Technologies, 2008. ICALT '08. Eighth IEEE International Conference on, On page(s): 241 - 245
8. Primo, T.T.; Vicari, R.M.; Bernardi, K.S. "User profiles and Learning Objects as ontology individuals to allow reasoning and interoperability in recommender systems", Global Engineering Education Conference (EDUCON), 2012 IEEE, On page(s): 1 - 9.
9. Aciar, Silvana, and Gabriela Aciar. "Analyzing User's Comments to Peer Recommendations in Virtual Communities." Computational Intelligence in Data Mining-Volume 1. Springer India, ISBN 978-81-322-2204-0. 2015. 583-592.