



Aplicación de comercio electrónico utilizando un algoritmo de filtrado colaborativo

Algoritmo

RECOMENDADOR
RECOMENDADOR

SISTEMA

Filtrado colaborativo

Algoritmo k-NN

RECOMENDADOR
RECOMENDADOR

SISTEMA

Yeimi Duran Xochimitl*
Selene Hernández
Rodríguez**

*Estudiante de Maestría en Ingeniería en el Instituto Tecnológico de Puebla (ITP).

**Docente en el Instituto Tecnológico de Puebla (ITP) en la División de Estudios de Posgrado e Investigación (DEPI).



SUMARIO: 1. Resumen/Abstract; 2. Introducción; 3. Antecedentes; 4. Sistema recomendador propuesto; 5. Análisis de resultados; 6. Conclusiones; 7. Fuentes de consulta.

1. RESUMEN

Con el crecimiento actual de las Tecnologías de la Información y Comunicación (TIC) y la capacidad de almacenamiento de los dispositivos computacionales, se han desarrollado aplicaciones web cada vez más robustas. Tal es el caso de los sistemas recomendadores, que permiten a los usuarios ver productos (o ítems) que probablemente serían de su interés entre una gran cantidad de información disponible en estos sitios, en los cuales generalmente el número de productos es muy grande. Para realizar estas recomendaciones a un usuario, los sistemas recomendadores basados en filtrado colaborativo utilizan la información de usuarios similares, suponiendo que un usuario similar es aquél que ha calificado los mismos ítems con calificaciones similares. Para encontrar los usuarios similares se han utilizado diversos enfoques; uno de estos enfoques ha sido utilizar los algoritmos de clasificación. En este trabajo se seleccionó el algoritmo de clasificación de k-vecinos más cercanos (k-NN) debido a su rapidez y buen funcionamiento. Para evaluar el grado de similitud entre usuarios, se seleccionó la función de correlación de Pearson. El sistema recomendador propuesto se implementa dentro de una aplicación web de venta de ropa.

PALABRAS CLAVES: Sistema recomendador, filtrado colaborativo, algoritmo k-NN, comercio electrónico.

ABSTRACT

With the current growth of Information and Communication Technology (ICT) and the storage capacity of computational devices, more robust applications have been developed. Such is the case of recommender systems, which allow users to see products (or items) that are likely to be of their interest from a large amount of information available in these sites, where the number of products is generally very large. To make these recommendations to a user, the recommender systems based on collaborative filtering uses information of similar users, assuming that a similar user is one who has qualified same items with similar qualifications. To find similar users, different approaches have been used; one of these approaches have been to use the classification algorithms. In this paper the k-nearest neighbors (k-NN) classifier was selected, due to its speed and good performance. To evaluate the degree of similarity between users, Pearson correlation function was selected. The proposed recommender system is implemented in a web application of clothes.

KEYWORDS: Recommender system, collaborative filtering, k-NN algorithm, e-Commerce applications.





2. INTRODUCCIÓN

Actualmente el desarrollo de las aplicaciones web ha tenido un gran crecimiento debido a los avances de las Tecnologías de la Información y Comunicación (TIC) y la capacidad de almacenamiento actual de los dispositivos computacionales. Con estos avances las aplicaciones web actuales nos permiten obtener beneficios tales como: estar informados en todo momento, comunicarse desde cualquier lugar en cualquier instante, comprar cualquier producto de nuestro interés en línea, entre otros. En particular, este último punto ha tenido un gran auge en Internet, ya que resulta bastante útil poder comprar por medio de Internet los productos de acuerdo a nuestro interés sin tomar en cuenta obstáculos geográficos.

Sin embargo, debido a la gran cantidad de información disponible en estos sitios web comerciales, resulta complicado para los usuarios tomar decisiones. Por ejemplo, la página web para renta de películas Netflix tiene más de 17,000 películas para seleccionar (Bennett, 2007) y la página de ventas de productos Amazon de acuerdo a un reporte financiero de Amazon.com en enero del 2010 (Ekstrand, 2010), tiene más de 41,000 títulos de libros en su tienda de Kindle. Por esta razón, hoy en día se cuenta con aplicaciones inteligentes, tales como los sistemas recomendadores, los cuales tienen como objetivo proporcionar sugerencias personalizadas de algún elemento (ítems) útil y de interés para cada uno de los usuarios analizando sus características con la información disponible de cada uno de ellos. Estos sistemas recomendadores han sido utilizados en diversas aplicaciones web, tales como: venta de artículos libros, música, programas de televisión, revistas, películas, entre otros productos o servicios.

En este trabajo se presenta el desarrollo de un sistema recomendador para una aplicación web de venta de ropa, basado en el enfoque de filtrado colaborativo, este enfoque, tiene como objetivo sugerir nuevos ítems o predecir la utilidad de cierto ítem para un usuario en particular, basándose en sus elecciones anteriores y en las elecciones de otros usuarios (usuarios similares) con un historial similar de calificaciones (Sarwar, 2001). Para encontrar los usuarios similares se han utilizado diversos algoritmos; uno de estos son los algoritmos de clasificación. En este trabajo se utiliza el algoritmo de k-vecinos más cercanos (k-NN) el cual es un método de clasificación supervisada que sirve para predecir un valor, en este caso, se predice la calificación que un usuario asignaría a un ítem. Para calcular la predicción, se realiza una comparación exhaustiva entre el usuario al que se le desea hacer la recomendación (usuario activo) y el resto de los usuarios con el objetivo de encontrar la similitud entre usuarios (algunos otros enfoques se basan en la similitud entre ítems). Para la comparación entre usuarios se utiliza una función de similitud, en este trabajo se usa la función de correlación de Pearson, con la cual aquél usuario que minimiza o maximiza esta función será el usuario más cercano. El sistema recomendador propuesto se aplica a bases de datos con calificaciones asignadas por los usuarios a los ítems consultados o comprados, estas calificaciones están en el intervalo de números enteros $[a, b]$, donde a significa que el ítem no fue del agrado del usuario, mientras que b significa que sí le gusto. Esta información es utilizada para sugerir a los usuarios ítems que aún no han consultado, basándose en la idea de que usuarios similares son aquellos que



tienden a calificar bajo o alto los mismos ítems. Para realizar pruebas con el sistema recomendador desarrollado en este trabajo, se utilizó una base de datos de ropa compuesta por 150 usuarios y 25 productos, y otras dos bases de datos obtenidas de Internet, las cuales contienen información de calificaciones asignadas por usuarios a ciertos ítems. Una de estas bases de datos consiste en calificaciones asignadas a películas por distintos usuarios (MovieLens) y la otra base de datos consiste en calificaciones asignadas a bromas (Jester). Para validar el funcionamiento de este sistema recomendador propuesto se calculó el error absoluto medio (MAE).

3. ANTECEDENTES

En esta sección se va a describir el trabajo existente que se ha llevado a cabo relacionado con los sistemas recomendadores en diferentes aplicaciones.

Un sistema recomendador (Adomavicius, 2005), se puede definir como una herramienta que ayuda a los usuarios a encontrar ítems que pudiera ser de su agrado, basado en información de su perfil o bien en información de otros usuarios. Existen varias formas en las que un usuario puede expresar su agrado hacia un ítem, por ejemplo: opiniones, valoraciones o calificaciones. Los sistemas recomendadores adoptan dos tipos de evaluación, los cuales son:

- Evaluación explícita, la cual se refiere a que el usuario asigna una puntuación a cada ítem ya sea por medio de estrellas, like/dislike, calificaciones numéricas entre algún rango mínimo y máximo, entre otros.
- Evaluación implícita, en la cual no se requiere de una intervención del usuario activo. El sistema detecta en número de veces que un usuario hace clic a un ítem.

Generalmente un sistema de recomendación compara el perfil del usuario con algunas características de referencia de los ítems, y busca predecir la calificación que el usuario le asignaría a un ítem que aún el sistema no ha considerado. Estas características pueden basarse en la relación o acercamiento del usuario con un ítem o en el ambiente social del mismo usuario. Estos sistemas han tenido gran aceptación debido a que los usuarios están acostumbrados a recibir recomendaciones de amigos y colegas. Por lo tanto, se deben tener en cuenta diferentes aspectos para diseñar sistemas de recomendación. A continuación se describen algunos de estos aspectos.

Uno de los aspectos importantes es el volumen de información, ya que éste depende de las recomendaciones, es decir, necesitan de una gran cantidad de información sobre usuarios e ítems para poder realizar recomendaciones de calidad. Otro aspecto es establecer un perfil de intereses de las recomendaciones ya que puede ocasionar problemas en casos de imparcialidad. Además, la privacidad de los usuarios debe tomarse en cuenta, ya que no a todos les gusta exhibir sus preferencias.





Los sistemas de recomendación están resultando de vital importancia para el marketing personalizado ya que quieren lograr los objetivos de reducir el tiempo de búsqueda para un ítem y conseguir una mayor efectividad en las búsquedas, por lo tanto, una mayor satisfacción en los clientes. Para lograr estos objetivos, todos los sistemas de recomendación llevan a cabo dos tareas:

Predecir. Los sistemas de recomendación predicen una serie de ítems o servicios en los que un usuario o cliente en particular podría estar interesado.

Recomendar los m-mejores ítems. Los sistemas de recomendación identifican los m ítems en los que el usuario estaría más interesado. Cabe aclarar que m se refiere a un conjunto de ítems.

Los sistemas recomendadores se pueden dividir de acuerdo a su funcionamiento ver figura 1.

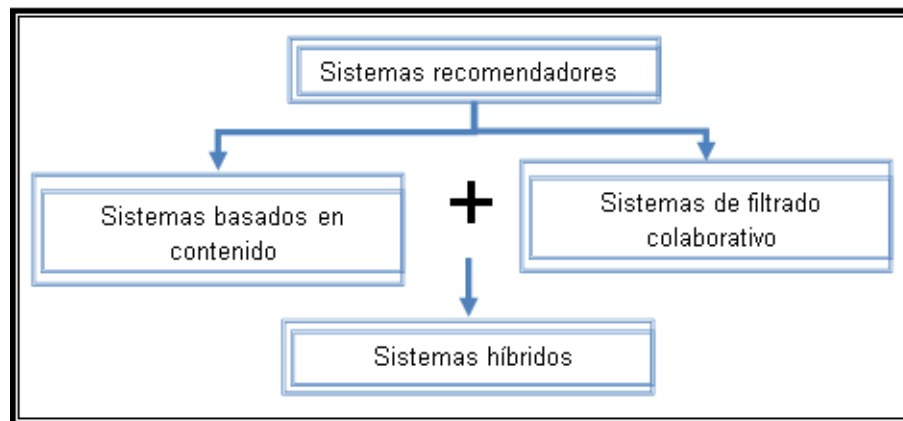


Figura 1. Clasificación de sistemas recomendadores. Fuente: Elaboración propia.

Sistemas basados en Contenido: El objetivo de estos algoritmos (Adomavicius, 2005 y Mladenic, 1999) es mostrar al usuario información de acuerdo al historial que se crea con los ítems que más ha visitado.

Sistemas de Filtrado Colaborativo (FC): El objetivo de estos algoritmos (Bell, 2007) es tomar en cuenta las preferencias de usuarios similares o ítems similares para recomendar nuevos ítems a un usuario. Estos sistemas a su vez se clasifican en dos grupos (Breese, 1998), los: basados en memoria (o basados en usuarios) y los basados en modelo (o basados en ítems). Los basados en memoria utilizan el conjunto completo de opiniones para encontrar los usuarios más similares al usuario activo, al cual se le quiere hacer la recomendación. Para ello se utiliza el algoritmo k-vecinos más cercanos (k-NN). Y los basados en modelo utilizan técnicas de minería de datos para desarrollar un modelo de calificaciones de cada usuario, el cual es empleado para predecir las preferencias ellos.

Sistemas recomendadores Híbridos: Estos sistemas (Christakou, 2005) utilizan una combinación de los dos enfoques anteriores.



Los sistemas recomendadores utilizan algún algoritmo de clasificación para encontrar los usuarios o items más similares. A continuación se describen algunos de estos algoritmos:

Redes neuronales. Se aplicaron al filtrado colaborativo por primera vez (Bilsus, 1998). Más recientemente, se propone un enfoque (Nasraoui, 2004), en el cual un perceptrón multicapa es usado para predecir URLs que se dan como recomendación a los usuarios de acuerdo a su nivel de interés.

Máquinas de soporte vectorial (SVM). En términos generales, en este enfoque cada usuario es representado por un vector compuesto por las calificaciones de los ítems.

Algoritmo k Vecinos más Cercanos (k-NN). Se introduce el algoritmo k-NN (Adomavicius, 2005) en el contexto de los sistemas recomendadores, debido a que este algoritmo tiene la ventaja de ser conceptualmente muy relacionado con la idea de los sistemas de filtrado colaborativo, es decir, encontrar los usuarios con las mismas preferencias es la misma tarea que encontrar los vecinos más cercanos a los usuarios o items. Otra ventaja de utilizar k-NN contra algún otro clasificador es que no requiere generar un modelo a partir de la información de las calificaciones, por lo que es posible incorporar fácilmente nuevas calificaciones a la matriz de calificaciones, la cual es la estructura de entrada al algoritmo k-NN, a diferencia de clasificadores que generan estructuras de árbol o redes neuronales a partir de los datos de entrada (en este caso es la matriz de calificaciones).

Existen otros algoritmos de clasificación, no obstante, los más destacados son los que se describieron anteriormente. En este trabajo se utiliza el algoritmo k-NN.

4. SISTEMA RECOMENDADOR PROPUESTO

Este artículo está enfocado a la aplicación de un sistema de filtrado colaborativo con evaluación basada en usuarios, utilizado el desarrollo de un sistema web para la gestión y comercio de una tienda virtual de ropa. En la Figura 2 se muestra el diagrama general del sistema propuesto.

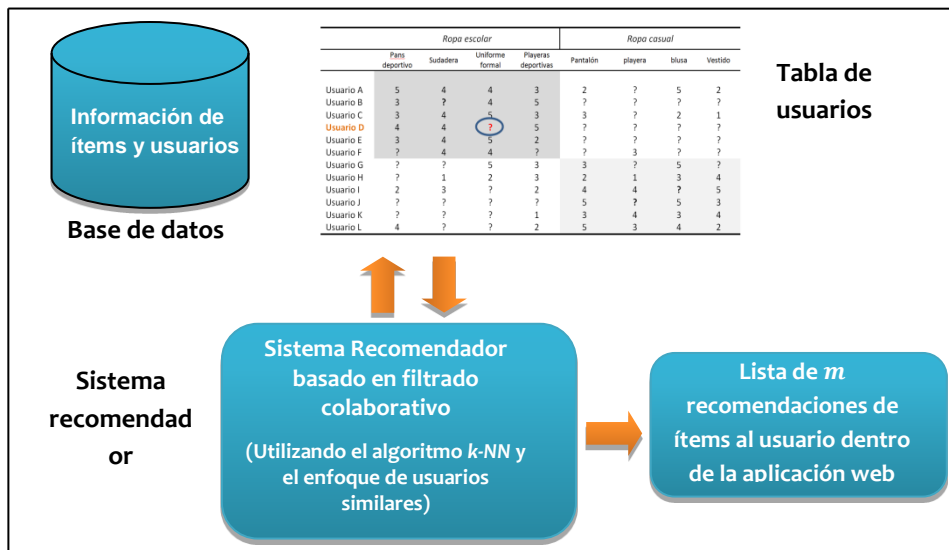


Figura 2. Diagrama general del sistema recomendador propuesto, basado en un algoritmo de filtrado colaborativo. Fuente: Elaboración propia.

A continuación, se da una breve descripción de cada uno de los módulos que se proponen para el sistema recomendador:

Base de datos. La función que tiene este módulo, es almacenar el perfil de cada uno de los usuarios a los que se les desea recomendar ítems, el perfil del usuario contiene datos como: nombre, edad, sexo y ocupación, así mismo se almacenará información de los ítems.

Tabla de usuarios. Es la tabla que se localiza dentro de la base de datos, donde las columnas corresponden a los ítems (en este caso la ropa) y las filas corresponden a los usuarios, los cuales han asignado alguna calificación a los ítems. (figura 3).

Sistema Recomendador. Este módulo contiene el algoritmo propuesto (ver figura 3), el cual tiene como objetivo predecir la calificación que el usuario activo asignaría a cada uno de los ítems que aún no ha calificado (o visto), basándose en la información de las calificaciones asignadas a estos ítems por los usuarios similares, encontrados con el algoritmo de clasificación de k-Vecinos más Cercanos (k-NN). Finalmente, la lista de m ítems con calificación aproximada más alta es mostrada al usuario activo para sugerirle ítems.

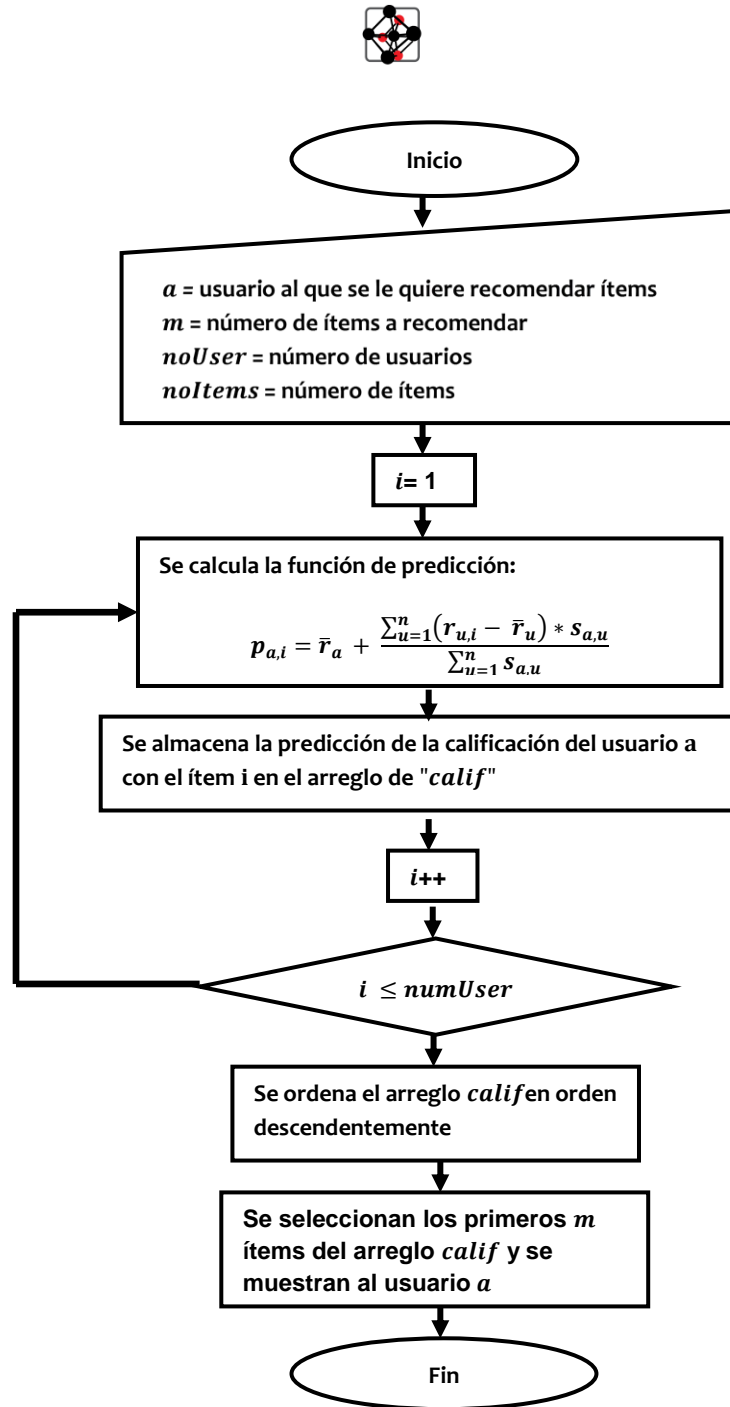


Figura 3. Algoritmo de filtrado colaborativo basado en usuarios.
Fuente: Elaboración propia.



Para medir la similaridad entre usuarios ($s_{a,u}$) se pueden utilizar varias funciones, tales como: la distancia euclidiana, la correlación de Pearson (Herlocker, 1999), Spearman (Herlocker, 1999), Coseno (Sarwar, 2000) y Entropía (Press, 1986). Como función de similaridad entre usuarios, se seleccionó y programó la función de Pearson, ya que de acuerdo al estado del arte relacionado con sistemas de filtrado colaborativo se han obtenido buenos resultados con esta función (ver ecuación1).

$$S_{a,u} = \frac{\sum_{i=1}^m (r_{a,i} - \bar{r}_a) * (r_{u,i} - \bar{r}_u)}{\sigma_a * \sigma_u} \quad (1)$$

Donde: $S_{a,u}$ es la similaridad que hay entre un usuario activo (a) con el resto de los demás usuarios (u), a es el usuario al que se le desea realizar la recomendación o también llamado usuario activo, u representa a cualquier otro usuario dentro de un conjunto de usuarios, excepto al usuario activo (a), i representa a un ítem, r es la calificación que un usuario asigna a un ítem, \bar{r}_a es la calificación promedio de un usuario activo (a), \bar{r}_u es la calificación promedio de algún usuario (u), $r_{a,i}$ es la calificación que un usuario activo (a) asigna al ítem (i), $r_{u,i}$ es la calificación que asigna cualquier usuario (u) a un ítem (i).

Finalmente, la función de similaridad (en este caso Pearson) se utiliza para calcular la predicción de la calificación que un usuario asignaría a un ítem. Esta predicción (ver ecuación 2) se calcula para todos los ítems que el usuario no ha calificado.

$$p_{a,i} = \bar{r}_a + \frac{\sum_{u=1}^n (r_{u,i} - \bar{r}_u) * s_{a,u}}{\sum_{u=1}^n s_{a,u}} \quad (2)$$

Donde: $p_{a,i}$ es la predicción de un usuario activo (a) a un ítem (i), \bar{r}_a es la calificación promedio que un usuario activo (a) asigna a los ítems que evalúa, $r_{u,i}$ es la calificación que los usuarios (u) asignan a un ítem (i), \bar{r}_u es la calificación promedio de los usuarios (u) y $s_{a,u}$ es la similaridad de un usuario activo (a) con un usuario (u), la cual se calculo anteriormente (ver ecuación 1).

Para ejemplificar el funcionamiento del algoritmo mostrado en la Figura 3, suponga que se tienen los datos mostrados en la Tabla 1, en la cual se muestra un ejemplo de usuarios que califican ropa. En esta tabla se puede observar que cierto grupo de usuarios tiene preferencias por la ropa escolar (los usuarios A-E), mientras que otro grupo de usuarios tiene preferencias por la ropa casual (los usuarios I-M). En esta tabla también se puede observar que algunos usuarios no han consultado todos los modelos de ropa



(representadas por un signo ?). Supongamos que en particular, queremos sugerir nuevos modelos de ropa al usuario D.

Tabla 1. Matriz de usuarios por ítems para almacenar calificaciones.

	Ropa escolar				Ropa casual			
	Pans deportivo	Sudadera	Uniforme formal	Playera deportiva	Pantalón mezclilla	Playera	Blusa	Vestido
Usuario A	5	4	4	3	2	3	1	2
Usuario B	3	?	4	5	?	?	3	?
Usuario C	3	4	5	3	3	4	2	1
Usuario D	4	4	?	5	2	?	?	2
Usuario E	3	4	5	2	?	2	1	3
Usuario F	?	4	4	?	3	3	4	5
Usuario G	5	?	5	3	3	?	5	?
Usuario H	4	5	2	3	2	1	3	4
Usuario I	2	3	?	2	4	4	?	5
Usuario J	?	?	?	?	5	?	5	3
Usuario K	?	?	2	1	3	4	3	4
Usuario L	2	?	?	2	5	3	4	2
Usuario M	3	2	1	3	3	5	4	5

Con el algoritmo que se mostró anteriormente en la Figura 3, se calculará la predicción de la calificación que el usuario D asignaría a la ropa que aún no ha consultado. Por lo tanto, la predicción de la calificación obtenida de acuerdo al algoritmo es: $D = 4$. Finalmente, se le mostrará como recomendación la lista de los m modelos de ropa con calificación aproximada más alta, recordando que como objetivo se tenían que localizar los vecinos más cercanos a D, que en este caso es el usuario C y el usuario E. En conclusión, la lista con los m modelos de ropa serán los modelos que sus vecinos (usuario C y E) calificaron alto, determinando que fueron modelos de agrado para los usuarios C y D, así mismo, se considera que, serán modelos que posiblemente también le agraden al usuario D.

5. ANÁLISIS DE RESULTADOS

Para realizar pruebas con el sistema propuesto, se creó una base de datos de ropa (BD ropa) compuesta por 150 usuarios y 25 productos, y se utilizaron otras dos bases de datos obtenidas de Internet (ver Tabla 2), las cuales contienen información de calificaciones asignadas por usuario a ciertos ítems (por ejemplo, la matriz mostrada en la Tabla 2). Una de estas bases de datos consiste en calificaciones asignadas a 1682 películas por 943 usuarios (MovieLens) y la otra base de datos consiste en calificaciones asignadas a 100 chistes por 24983 usuarios (Jester).

**Tabla 2.** Descripción de las bases de datos utilizadas

Base de datos	Número de usuarios	Número de usuarios	Intervalo de las calificaciones	Símbolo para indicar ítems no calificados aún
BD ropa	150	25	De 1 a 5	0
MovieLens	943	1682	De 1 a 5	0
Jester	24983	100	De -10 a 10	99

Para evaluar el funcionamiento del sistema recomendador propuesto se calcularon algunas calificaciones asignadas por algunos usuarios a ciertos ítems, de las cuales previamente se conocía su valor, este proceso se realizó para el 10% de la información disponible de cada base de datos mencionada anteriormente. Es decir, se utilizó el 90% de las calificaciones como base para que pudiera funcionar el sistema recomendador propuesto y el 10% restante se utilizó para pruebas. En este trabajo se utilizó el valor del error medio absoluto (MAE), de acuerdo a la ecuación 3, para calcular el error obtenido por el sistema recomendador propuesto.

$$MAE = \frac{\sum_{i=1}^N |p_i - r_i|}{N} \quad (3)$$

Donde: p_i es la calificación obtenida por la predicción del sistema recomendador propuesto, r_i es la calificación real conocida y N es el número de calificaciones que se predijeron, que en este caso corresponde al 10% de cada base de datos.

En la Tabla 3 se muestra el error MAE para la evaluación del sistema recomendador propuesto, con las bases de datos descritas anteriormente, utilizando la función de similitud de Correlación de Pearson.

Tabla 3. Error Medio Absoluto (MAE) obtenido de evaluar el sistema recomendador desarrollado, con el 10% de la información disponible para pruebas y el 90% como base para el funcionamiento del sistema recomendador propuesto.

Función de similitud utilizada dentro de la función de predicción	Base de datos		
	BD ropa	MovieLens	Jester
Correlación de Pearson	0.37	0.77	1.21



A partir de la Tabla 3 se puede observar que en el caso de las bases de datos BD ropa y MovieLens, las cuales están calificadas en el intervalo del 1 al 5, se obtiene un error de 0.37 (para BD ropa) y 0.77 (para MovieLens). Esto significa, por ejemplo, que en lugar de obtener una calificación real de 3, el sistema recomendador propuesto nos estaría regresando una calificación en el intervalo de 2.63 y 3.37, para la base de datos BD ropa y un intervalo de 2.23 y 3.77 para la base de datos MovieLens. Para la base de datos BD ropa el error no es muy grande ya que los números comprendidos en este intervalo se redondean a la calificación de 3. Para la base de datos MovieLens el error es poco significativo ya que al redondear si se aleja un poco de la calificación real 3. Para el caso de jester, suponiendo una calificación real de 8 el sistema recomendador propuesto regresaría una calificación aproximada en el intervalo de 6.79 a 9.21. Como se puede ver, a partir de estos tres resultados el valor aproximado se acerca mucho al valor real.

De acuerdo a los resultados obtenidos con el funcionamiento del sistema recomendador que se propone en este trabajo y las pruebas de error que se realizaron con tres diferentes bases de datos, se considera que el sistema recomendador será de mucha utilidad a los usuarios para realizar búsquedas más rápidas sin tener la necesidad de consultar todos los modelos de ropa que hay dentro de la aplicación web a la que será implementado. Este sistema recomendador, ya que con la ayuda del algoritmo k-NN y la función de Correlación de Pearson se localizarán los usuarios similares al usuario activo y de ellos se harán las posibles recomendaciones, de modo que se evitarán las búsquedas exhaustivas.

6. CONCLUSIONES

Con el crecimiento y auge actual de las aplicaciones web de Internet enfocadas al comercio electrónico, surge la idea de implementar un sistema recomendador dentro de una aplicación web para la venta de ropa. En este trabajo se hace uso del enfoque de filtrado colaborativo con las principales tareas de predecir una calificación y en contra los mejores ítems para recomendar al usuario activo.

Sin embargo, la calidad y el número de vecinos que puede tener un usuario activo está directamente relacionado con la cantidad de información que se puede recolectar de cada usuario con la aplicación web. Sin embargo, la calidad de vecinos depende del número de ítems en común con el usuario activo.

Finalmente, durante las pruebas que se realizaron con el sistema recomendador, observamos que cuando aumenta el número de vecinos, se mejora la precisión, mientras, por otro lado, si se considera un mínimo de ítems en común con el usuario activo, es posible seleccionar mejores vecinos.





7. FUENTES DE CONSULTA

- Adomavicius, G. and Tuzhilin, A. (2005). "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions". Recuperado de: <http://dl.acm.org/citation.cfm?id=1070611.1070751>. Fecha de consulta: 06/12/12
- Bell, Robert M., and Koren, Yehuda. (Octubre 2007). "Scalable collaborative filtering with jointly derived neighborhood interpolation weights". Recuperado de: <http://ieeexplore.ieee.org>. Fecha de consulta: 06/12/12.
- Bennett, James., Lanning, Stan. (2007). "The netflix prize". Recuperado de: <http://www.cs.uic.edu/~liub/KDD-cup-2007/NetflixPrize-description>. Fecha de consulta: 20/04/13.
- BilSus, D., Pazzani, Michael J. (1998). "Learning collaborative information filters". Recuperado de: <http://www.aaai.org/Papers/Workshops/1998/WS-98-08/WS98-08-005.pdf>. Fecha de consulta: 07/05/13.
- Breese, John S., Heckerman, David., and Kadie, Carl. (1998) "Empirical Analysis of Predictive Algorithms for Collaborative Filtering". Recuperado de <http://research.microsoft.com/pubs/69656/tr-98-12.pdf>. Fecha de consulta 20/04/13.
- Christakou, C., Stafylopatis, A. (enero 2007) "A hybrid movie recommender system based on neural networks". En: Proceedings 5th International Conference on Intelligent Systems Design and Applications, ISDA, pp.500-505.
- Ekstrand, D. Michael, Riedl, John T., and Konstan, Joseph A., (2010). "Collaborative filtering recommender systems". En: Human-Computer Interaction. 4(2) pp.81-173.
- Herlocker, Jonathan L., Konstan, Joseph A., Borchers Al., and Riedl, John. (1999). "An Algorithmic Framework for Performing Collaborative Filtering. Recuperado de: <http://dl.acm.org/citation.cfm?id=963772>. Fecha de consulta: 06/12/12.
- Mladenic, D. (1999). "Text-learning and Related Intelligent Agents: A Survey". Recuperado de: <http://www.computer.org/csdl/mags/ex/1999/04/x4044-abs.html>. Fecha de consulta: 20/01/13.
- Nasraoui, O. and Pavuluri, M. (Mayo 2004). "Accurate Web Recommendations Based on Profile-Specific URL-Predictor Neural Networks". En: In Proceedings of the International World Wide Web Conference.
- Press, William., Flannery, Brian., Teukolsky, Saul., and Vettering, William. (1986) "Numerical recipes: The art of Scientific Computing". New York: Cambridge University Press.
- Sarwar, Badrul. Karypis, George. Konstan, Joseph. and Riedl, John. (2001). "Item-based Collaborative Filtering Recommendation Algorithms". Recuperado de: <http://dl.acm.org/citation.cfm?id=372071>. Fecha de consulta: 06/12/12.



Sarwar, Badrul. Karypis, George. Konstan, Joseph. and Riedl, John. (2000). "Analysis of recommendation algorithms for e-commerce". Recuperado de: <http://dl.acm.org/citation.cfm?id=352887>. Fecha de consulta: 06/12/12.