



TRABAJO FIN DE MÁSTER
MÁSTER UNIVERSITARIO OFICIAL EN CIENCIA DE DATOS
E INGENIERÍA DE COMPUTADORES

Sistemas de recomendación de películas para grupos

Autor

Juan José Sierra González

Director

Juan Manuel Fernández Luna



ESCUELA TÉCNICA SUPERIOR DE INGENIERÍAS INFORMÁTICA Y DE
TELECOMUNICACIÓN

Granada, enero de 2020

Sistemas de recomendación de películas para grupos

Autor

Juan José Sierra González

Director

Juan Manuel Fernández Luna

Sistemas de recomendación de películas para grupos

Juan José Sierra González

Palabras clave: incluir palabras clave...

Resumen

Incluir resumen...

Group recommender systems for movies

Juan José Sierra González

Keywords: include keywords...

Abstract

Include abstract...

Yo, **Juan José Sierra González**, alumno del Máster Universitario Oficial en Ciencia de Datos e Ingeniería de Computadores de la **Escuela Técnica Superior de Ingenierías Informática y de Telecomunicación de la Universidad de Granada**, con DNI 76589592Y, autorizo la ubicación de la siguiente copia de mi Trabajo Fin de Máster en la biblioteca del centro para que pueda ser consultada por las personas que lo deseen.

Fdo: Juan José Sierra González

Granada a 9 de enero de 2020.

D. **Juan Manuel Fernández Luna**, Profesor del Área de Uncertainty Treatment in Artificial Intelligence del Departamento de Ciencias de la Computación e Inteligencia Artificial de la Universidad de Granada.

Informa:

Que el presente trabajo, titulado ***Sistemas de recomendación de películas a grupos***, ha sido realizado bajo su supervisión por **Juan José Sierra González**, y autoriza la defensa de dicho trabajo ante el tribunal que corresponda.

Y para que conste, expide y firma el presente informe en Granada a 9 de enero de 2020.

El director:

Juan Manuel Fernández Luna

Agradecimientos

Incluir agradecimientos...

Índice general

1. Introducción	1
1.1. Motivación	1
1.2. Objetivos	2
2. Revisión de la literatura: estado del arte	5
Bibliografía	7

Índice de figuras

Índice de tablas

Capítulo 1

Introducción

En la actualidad, la informática juega un papel fundamental en el día a día de la mayoría de personas. No solo en el ámbito profesional, donde ya lleva más años asentada, sino que cada vez es más común encontrar soluciones informáticas en problemas de la vida cotidiana. Donde antes podían requerirse varios minutos de búsqueda entre páginas de enciclopedias, ahora se resuelven consultas de todo tipo en cuestión de segundos gracias a los grandes volúmenes de información que se almacenan en Internet, y que comparten (y amplían) a diario millones de usuarios.

Pero la información no siempre se presenta de forma clara y concisa, y a menudo se acumulan enormes cantidades de datos aparentemente inservibles. Sin embargo, mediante el manejo de tan grandes cantidades de dichos datos se puede obtener información útil que de otra forma sería prácticamente imposible de detectar. La **ciencia de datos**, encargada de obtener toda esa información compleja, abundante y difusa y transformarla en información útil, está siendo utilizada cada vez con más asiduidad para todo tipo de finalidades. Este estudio se centrará en su aplicación para **recomendar** elementos a los usuarios, o lo que más comúnmente se conoce como **sistemas de recomendación**.

1.1. Motivación

Los sistemas de recomendación ya forman parte de la gran mayoría de plataformas web que ofrecen bienes o servicios. Mientras un usuario ojea un elemento de un catálogo en una página web es habitual que encuentre las secciones de “Productos similares” y “Productos recomendados”. Mientras que la primera se basa puramente en características comunes entre los elementos, la segunda es **una recomendación personalizada** para cada usuario.

Debido a que las empresas generan beneficio directamente de realizar una buena recomendación a sus clientes, optimizar el funcionamiento de estos sistemas de recomendación resulta una tarea muy importante. Si bien los sistemas clásicos funcionan bien en términos generales, a cada empresa le interesa que su sistema se comporte de forma correcta en su ámbito, y es en estos detalles donde más interesantes son las mejoras. Y es que para una multinacional, una mejora en sus recomendaciones de un mero 1 % (que puede parecer ínfima) resulta en una inyección de ingresos muy sustancial.

Este estudio analizará uno de los ámbitos más comunes, las **recomendaciones de películas**. Como principal ejemplo, la plataforma de ocio audiovisual **Netflix** es una de las más utilizadas del momento, cuenta con un sistema de recomendación de películas y series muy elaborado, y sin embargo, solo es capaz de recomendar **para un único usuario**. No solo en el ámbito de las películas, sino en casi cualquiera, generalmente las recomendaciones se hacen de forma individual, y hay que buscar un sistema especializado si se quiere recomendar a un grupo de personas. Existen multitud de implementaciones de **sistemas de recomendación a grupos** que tratan de optimizar la elección para satisfacer lo máximo posible a todos los miembros del grupo, o al menos lograr un equilibrio. Este trabajo se esforzará en estudiar las soluciones más actuales a este problema y cómo han mejorado lo que ofrecen las soluciones más clásicas, enmarcadas en el ámbito de las películas.

1.2. Objetivos

Como se ha mencionado, ese Trabajo de Fin de Máster pone el foco en los sistemas de recomendación de películas a grupos. Existen múltiples soluciones a este problema, que han ido mejorando con el paso de los años y el progreso logrado mediante investigación. En este estudio se pondrá especial atención en aquellos más recientes, que en esencia serán los que aúnen de forma más eficiente dichos progresos. Una vez estudiados algunos de estos modelos, se tratará de crear una serie de versiones propias que mejoren lo existente.

El plan de ruta del trabajo será partir de un **modelo básico**, o *baseline*, que será el que marque el valor mínimo que debe tener el modelo propio implementado. Se entiende que un modelo peor que un baseline jamás debe ser considerado como bueno, ya que se trata de algoritmos altamente probados y contra numerosas situaciones distintas, y que funcionan bien en términos generales.

Con el modelo básico seleccionado, se procederá a realizar una investigación en la literatura, descubriendo cómo se ha producido la evolución en el

campo de los sistemas de recomendación a grupos, cuáles son las propuestas más innovadoras en los últimos años y, principalmente, cuáles han dado un mejor resultado. Entre estas últimas se seleccionará una entre las más recientes y las que más interesantes y adaptables resulten de cara a futuras modificaciones.

Habiendo elegido un modelo baseline y un modelo actual, el siguiente paso será implementarlos. Se tratará de ajustar lo máximo posible a la información recabada en los *papers* de los algoritmos, para ser fiel a la implementación original de los autores. Por desgracia, ninguna de las versiones más modernas y funcionales es de código abierto y es imposible encontrar una implementación que no haya sido creada a interpretación del *paper*.

Una vez se hayan implementado los dos modelos, se crearán una serie de nuevas soluciones que pongan en práctica lo aprendido durante el estudio en la literatura y las ideas surgidas al mismo tiempo. Se tratará de adaptar el modelo actual en la medida de lo posible, pero utilizando las nuevas técnicas que se crea que puedan suponer un incremento en la eficacia del método.

Por último, con todas las implementaciones hechas, se realizará un estudio experimental donde se pondrán a prueba todos los sistemas de recomendación bajo diferentes circunstancias y parámetros. Se recogerán los datos y posteriormente se evaluará cuál es la mejor solución y si alguna de las soluciones ideadas para este trabajo han logrado mejorar lo existente.

Capítulo 2

Revisión de la literatura: estado del arte

En este capítulo se expondrá una revisión de los estudios y propuestas más innovadoras relacionadas con los **sistemas de recomendación a grupos**, y en particular, de aquellos orientados al ámbito de las películas. Se comenzará haciendo un breve inciso sobre cuándo y cómo se originaron los sistemas de recomendación, así como la particularidad de aquellos específicos de recomendación a grupos. Después se tratará su evolución hasta nuestros días, realizando un comentario analítico acerca de lo que han supuesto los sistemas de recomendación y cómo se trasladan actualmente dichos sistemas a situaciones cotidianas, con algunos claros ejemplos que todo el mundo utiliza. Entre ellos se destacarán los que abarcan la **recomendación de películas**, como tema principal del estudio. Por último, se mencionarán algunos de los sistemas de recomendación de películas a grupos más actuales, eligiendo además el que será tomado como referencia para la experimentación como representante del estado del arte de estos recomendadores.

Queda ya muy atrás el primer sistema de recomendación de la historia, que vio la luz en 1992, cuando Goldberg et al. (1992) presentaron su artículo “*Using collaborative filtering to Weave an Information tapestry*” [1]. En este trabajo se acuña por primera vez el término “**filtrado colaborativo**” (en inglés, *collaborative filtering*), refiriéndose así a la técnica que permite filtrar elementos a un usuario basándose en lo que otros usuarios han opinado de él anteriormente. Esta técnica sienta las bases de los sistemas de recomendación, y aún a día de hoy se encuentra presente en la mayoría de ellos. La otra alternativa al filtrado colaborativo es el **filtrado basado en contenido** (en inglés, *content-based filtering*). Esta técnica no tiene en cuenta las valoraciones que otros usuarios han dado a los elementos del sistema, sino que se limita a comparar los elementos por características y recomendar aquellos si-

milares. En este experimento, sin embargo, se tratarán únicamente métodos que utilizan el filtrado colaborativo.

Una vez se hubo contemplado el vasto rango de opciones que albergaban los sistemas de recomendación, a través de numerosos trabajos y estudios, se contempló la posibilidad de recomendar ya no solo a individuos, sino a grupos de personas. Se trataba de una situación a la que una sociedad podía enfrentarse en multitud de facetas, y que sin duda merecía la pena explorar. La pionera, y una de las más famosas y reconocidas propuestas sobre este tema, fue **PolyLens** [2]. En esta publicación de 2001, O'Connor et al. plantearon **un sistema de recomendación de películas para grupos**, derivado de la base de datos de valoraciones de películas **MovieLens**, en la que varios trabajos anteriores se habían fundado. Este sistema abrió la veda de un nuevo subgrupo dentro de la ya amplia categoría de sistemas de recomendación.

Como se ha comentado anteriormente, las recomendaciones a películas han levantado mucho interés desde el primer momento. El contenido de ocio audiovisual, junto a los catálogos de tiendas *online*, son los casos de mayor demanda en cuanto a sistemas de recomendación en la sociedad actual. Plataformas como **Netflix**, **Spotify** o **HBO** son ejemplos de los primeros, que utilizan estos sistemas para tratar de mantener al usuario satisfecho ofreciéndole nuevo contenido que sea similar al que hayan visto. Por su parte, **Amazon** o **AliExpress** son ejemplos de los segundos, grandes plataformas que dan cabida a todo tipo de productos y que son capaces de recomendar a sus clientes otros productos similares a los que están buscando. Gracias al sistema de recomendación pertinente, incluso consiguen ofrecer nuevos artículos que detectan que el usuario puede querer comprar, basándose en lo que anteriormente ha buscado o comprado.

Tal fue el impacto de los sistemas de recomendación que la propia plataforma Netflix lanzó en 2006 lo que se conocería como el **Netflix Prize** [3], una competición de sistemas que, basándose únicamente en el filtrado colaborativo (no conociendo datos adicionales sobre películas ni sobre usuarios, solamente las valoraciones que estos habían dado a cada película) fueran capaces de predecir acertadamente dichas valoraciones. El objetivo principal era superar al por aquel entonces actual sistema de recomendación de Netflix, conocido como **Cinematch**, ofreciendo un premio de 1 millón de dólares a aquel sistema capaz de superarlo en un 10 %. Adicionalmente, como la competición se estimaba que duraría hasta 2011, se concedería un premio anual de 50.000 dólares al método ganador al final de cada año.

Tras varios tira y afloja entre algunos equipos, e incluso agrupaciones entre equipos para forjar mejores ideas juntos, finalmente en 2009 el equipo '*BellKor's Pragmatic Chaos*' se alzó ganador y recibió el gran premio [4], superando al algoritmo anterior por 10.05 %. De esta forma tan original

Netflix cambió su sistema de recomendación promoviendo una iniciativa de investigación en este campo, y por supuesto previendo que este avance en su algoritmo les serviría para incrementar la satisfacción de sus clientes, y por ende, para aumentar su valor de mercado.

Bibliografía

- [1] D. Goldberg, D. Nichols, B.M. Oki, and D. Terry. Using collaborative filtering to Weave an Information tapestry. *Communications of the ACM*, 35:61–70, 1992.
- [2] M. O'Connor, D. Cosley, J.A. Konstan, and J. Riedl. PolyLens: A Recommender System for Groups of Users, 2001.
- [3] Netflix Prize. <https://www.netflixprize.com/>.
- [4] Netflix Prize Forum: Grand Prize awarded to team BellKor's Pragmatic Chaos. https://www.netflixprize.com/community/topic_1537.html.

