

Sistema de Recomendación Híbrido

Carlos A. Bresó Sotto
Marian S. Álvarez Suri

Facultad de Matemática y Computación
Universidad de La Habana, La Habana

<https://github.com/marians002/IRS-Final-Project.git>

Abstract

En el ámbito de los sistemas de recomendación, los enfoques híbridos han demostrado ser altamente efectivos al combinar las fortalezas de múltiples técnicas de recomendación. Este trabajo presenta un sistema de recomendación híbrido que integra métodos basados en contenido y filtrado colaborativo para mejorar la precisión y relevancia de las recomendaciones. El sistema propuesto utiliza un modelo de filtrado colaborativo para capturar las preferencias de los usuarios a partir de sus interacciones pasadas, mientras que el enfoque basado en contenido analiza las características intrínsecas de los ítems para ofrecer recomendaciones personalizadas. La combinación de estos métodos se realiza mediante un esquema de ponderación adaptativa que ajusta la contribución de cada técnica según el contexto y el perfil del usuario. Los resultados experimentales muestran que el sistema híbrido supera a los enfoques individuales en términos de precisión y satisfacción del usuario, destacando su potencial para aplicaciones en diversas áreas como el comercio electrónico, la recomendación de películas y la educación en línea.

Keywords: Sistema de Recomendación Híbrido, filtrado colaborativo, filtrado basado en contenido.

Contents

1	Introducción	3
1.1	Motivación	3
1.2	Objetivos	4
1.3	Estructura del Informe	4
2	Estado del Arte	5
2.1	Sistemas de Recomendación	5
2.1.1	Filtrado Colaborativo	6
2.1.2	Filtrado Basado en Contenido	6
2.1.3	Sistemas de Recomendación Híbridos	6
2.2	Métodos de Hibridación	6
2.2.1	Ponderación de Pesos	6
2.2.2	Mezcla de Modelos	6
2.2.3	Sistema Híbrido Alternante	6
2.2.4	Combinación de Características	6
3	Metodología	7
3.1	Descripción del Sistema	7
3.2	Datos Utilizados	7
3.3	Implementación	7
3.3.1	Algoritmos Utilizados	7
3.3.2	Integración de Técnicas	8
4	Resultados y Discusión	9
4.1	Evaluación del Sistema	9
4.1.1	Métricas de Evaluación	9
4.1.2	Resultados Obtenidos	9
4.2	Análisis de Resultados	9
4.2.1	Precisión	9
4.2.2	Diversidad	9
4.2.3	Comparación con Otros Sistemas	9

5	Conclusiones y Trabajo Futuro	10
5.1	Conclusiones	10
5.2	Limitaciones	10
5.3	Trabajo Futuro	10
A	Apéndices	11
A.1	Código Fuente	11
A.2	Detalles Adicionales	11

Chapter 1

Introducción

1.1 Motivación

Los sistemas de recomendación han emergido como uno de los pilares fundamentales en la estrategia digital de las grandes compañías a nivel mundial, transformándose en herramientas indispensables para mejorar la experiencia del usuario y maximizar la eficiencia de los servicios en línea. Empresas líderes en el sector tecnológico, como Amazon, Netflix y Spotify, han adoptado estos sistemas para ofrecer a sus usuarios recomendaciones personalizadas, lo que no solo incrementa la satisfacción y fidelidad del cliente, sino que también impulsa significativamente las ventas y el compromiso. La capacidad de estos sistemas para analizar grandes volúmenes de datos en tiempo real y predecir las preferencias de los usuarios con alta precisión les permite a las compañías estar un paso adelante, asegurando una oferta más ajustada a las necesidades y gustos de su audiencia. En un mundo cada vez más saturado de información y opciones, los sistemas de recomendación se han convertido en una herramienta clave para filtrar y personalizar el contenido, demostrando ser un recurso valioso para mantener la competitividad y liderazgo de las grandes empresas en el mercado global.

El proyecto en cuestión aborda la creación de un sistema de recomendación híbrido, una solución innovadora que combina las fortalezas del filtrado colaborativo y el filtrado basado en contenido para superar las limitaciones inherentes a cada uno de estos enfoques por separado. El filtrado colaborativo aprovecha las interacciones y preferencias de los usuarios para hacer recomendaciones, basándose en la premisa de que aquellos con gustos similares en el pasado tendrán intereses parecidos en el futuro. Por otro lado, el filtrado basado en contenido se centra en las características de los ítems para hacer recomendaciones, lo que permite una personalización más

detallada, pero puede limitar la diversidad de las recomendaciones y reforzar la cámara de eco. Al integrar ambos enfoques, nuestra propuesta no solo mejora la precisión de las recomendaciones, sino que también enriquece la experiencia del usuario al ofrecer una variedad más amplia y relevante de opciones.

1.2 Objetivos

Nuestro trabajo se distingue de otros en el campo de los sistemas de recomendación por la incorporación de un enfoque innovador que mejora la diversidad de los elementos recomendados. A diferencia de los métodos tradicionales que se centran únicamente en las preferencias explícitas del usuario, hemos añadido un conjunto de recomendaciones que no se encuentran entre las preferidas por este. Dicha estrategia no solo enriquece la experiencia del usuario al introducirle nuevos y variados contenidos, sino que también mitiga el problema de la homogeneidad en las recomendaciones, fomentando así una exploración más amplia de las opciones disponibles.

El objetivo general de este trabajo es desarrollar un sistema de recomendación híbrido para películas que no solo se base en las preferencias explícitas del usuario, sino que también incorpore elementos que podrían resultar atractivos a pesar de su diferencia con el historial previo o con los gustos de personas con intereses similares. Esta estrategia busca enriquecer la diversidad de las recomendaciones cinematográficas, permitiendo descubrir nuevas películas y opciones que de otro modo podrían pasar desapercibidas.

1.3 Estructura del Informe

Chapter 2

Estado del Arte

2.1 Sistemas de Recomendación

En el ámbito de los sistemas de recomendación híbridos, el estado del arte ha evolucionado significativamente, reflejando un creciente interés en superar las limitaciones de los enfoques tradicionales mediante la integración de múltiples métodos de recomendación. Según el análisis presentado en [1], los sistemas de recomendación híbridos ofrecen una solución efectiva para superar las limitaciones de los sistemas basados únicamente en filtrado colaborativo o contenido, pues han demostrado ser especialmente eficaces en abordar problemas como el arranque frío y la escasez de datos, ofreciendo recomendaciones más precisas y personalizadas. Los estudios destacan la diversidad de técnicas utilizadas para combinar diferentes enfoques, desde métodos ponderados hasta la fusión de características, mostrando una tendencia hacia la experimentación y la innovación en el campo [2]. Existen varias formas de combinar estos métodos, entre las cuales destacan:

1. Ponderación de Pesos
2. Mezcla de Modelos
3. Sistema Híbrido Alternante
4. Combinación de Características

Estas revisiones sistemáticas subrayan la importancia de los sistemas de recomendación híbridos en el panorama actual, identificando áreas clave para la investigación futura, como la mejora de los algoritmos de fusión y la exploración de nuevas fuentes de datos para enriquecer las recomendaciones. La continua evolución de estos sistemas promete no solo mejorar la experiencia

del usuario en plataformas digitales, sino también abrir nuevas vías para el avance tecnológico en la inteligencia artificial y el análisis de datos [3].

2.1.1 Filtrado Colaborativo

2.1.2 Filtrado Basado en Contenido

2.1.3 Sistemas de Recomendación Híbridos

2.2 Métodos de Hibridación

2.2.1 Ponderación de Pesos

En este enfoque, las recomendaciones de diferentes sistemas se combinan asignando un peso a cada una de ellas. La importancia o puntuación de cada producto se calcula a partir de los resultados obtenidos por todas las técnicas de recomendación presentes en el sistema. Un ejemplo de este tipo de hibridación es el sistema P-Tango [2], que ajusta los pesos de acuerdo con las opiniones de los usuarios sobre los resultados mostrados.

2.2.2 Mezcla de Modelos

En este método, las recomendaciones provenientes de varias técnicas de recomendación se muestran al mismo tiempo. Esto permite que el usuario vea una combinación de recomendaciones basadas en diferentes enfoques, lo que puede mejorar la diversidad y la calidad de las recomendaciones.

2.2.3 Sistema Híbrido Alternante

Este enfoque utiliza criterios específicos para alternar entre diferentes técnicas de recomendación. Por ejemplo, primero se emplean técnicas basadas en contenido y, si no se obtienen buenos resultados, el sistema cambia a filtrado colaborativo. Esto permite que las recomendaciones sean sensibles a las fortalezas y debilidades de cada método [4].

2.2.4 Combinación de Características

En este método, las características de las fuentes de datos de varias técnicas de recomendación se combinan en un único algoritmo de recomendación. Esto permite aprovechar la información de diferentes fuentes para generar recomendaciones más precisas.

Chapter 3

Metodología

3.1 Descripción del Sistema

El sistema desarrollado es un recomendador híbrido que combina múltiples modelos de recomendación con pesos específicos. Este sistema se encarga de generar recomendaciones personalizadas para los usuarios basándose en sus interacciones previas con los ítems.

3.2 Datos Utilizados

Se utiliza el conjunto de datos MovieLens 100K, que contiene 100,000 calificaciones de películas realizadas por 943 usuarios sobre 1682 películas. Este conjunto de datos incluye información sobre las películas, como títulos, géneros y fechas de lanzamiento, así como las calificaciones otorgadas por los usuarios [5].

3.3 Implementación

3.3.1 Algoritmos Utilizados

Se tomaron los siguientes algoritmos de recomendación de la librería cornac [6] para implementar el sistema híbrido:

- SVD (Singular Value Decomposition): Un modelo de factorización de matrices que descompone la matriz de calificaciones en matrices de características latentes.
- ItemKNN (Item-based K-Nearest Neighbors): Un modelo basado en la similitud entre ítems, utilizando la similitud coseno.

- BPR (Bayesian Personalized Ranking): Un modelo de optimización para el ranking personalizado basado en la probabilidad bayesiana.

3.3.2 Integración de Técnicas

El sistema híbrido combina los modelos mencionados utilizando un esquema de ponderación. Los pesos asignados a cada modelo son (6, 3, 1) para SVD, BPR y ItemKNN, respectivamente. Además, se calcula la similitud entre las películas basándose en sus géneros utilizando la similitud de coseno. Esta información de similitud se utiliza para mejorar las recomendaciones, añadiendo películas no tan similares a la última película que el usuario calificó con el objetivo de lograr mayor diversidad de las recomendaciones.

Chapter 4

Resultados y Discusión

4.1 Evaluación del Sistema

4.1.1 Métricas de Evaluación

4.1.2 Resultados Obtenidos

4.2 Análisis de Resultados

4.2.1 Precisión

4.2.2 Diversidad

4.2.3 Comparación con Otros Sistemas

Chapter 5

Conclusiones y Trabajo Futuro

5.1 Conclusiones

5.2 Limitaciones

5.3 Trabajo Futuro

Appendix A

Apéndices

A.1 Código Fuente

A.2 Detalles Adicionales

Bibliography

- [1] E. Çano and M. Morisio, “Hybrid recommender systems: A systematic literature review,” *Intelligent Data Analysis*, vol. 21, no. 6, pp. 1487–1524, 2017.
- [2] L. G. P. Cordón, “Modelos de recomendación con falta de información: aplicaciones al sector turístico,” *Consejería de Turismo, Comercio y Deporte, Junta de Andalucía*, vol. 1, no. N/A, p. N/A, 2009, tesis doctoral dirigida por Luis Martínez López, Universidad de Jaén, Escuela Politécnica Superior, Departamento de Informática.
- [3] D. Roy and M. Dutta, “A systematic review and research perspective on recommender systems,” *Journal of Big Data*, vol. 9, no. 1, p. Article number: 59, 2022.
- [4] C. V. Sacin, “Propuesta de una metodología de aplicación de técnicas de descubrimiento del conocimiento para la ayuda al estudiante en entornos de enseñanza superior,” *Universidad Autónoma de Madrid, Escuela Politécnica Superior, Departamento de Ingeniería Informática*, vol. N/A, no. N/A, p. N/A, 2010, tesis doctoral dirigida por D. Álvaro Ortigosa.
- [5] “Movielens dataset,” (<https://grouplens.org/datasets/movielens>).
- [6] C. Team, “Cornac: A comparative framework for multimodal recommender systems,” <https://github.com/PreferredAI/cornac>, 2018.